			USOCR; EPO; JPO; DERWENT; IBM_TDB			
S22	92	US-1234567-$.DID. OR US-5423046-$.DID. OR US-5881315-$.DID. OR US-5930768-$.DID. OR US-6108699-$.DID. OR US-6202080-$.DID. OR US-6546403-$.DID. OR US-6578068-$.DID. OR US-6751698-$.DID. OR US-6782537-$.DID. OR US-6968335-$.DID. OR US-7093004-$.DID. OR US-7136924-$.DID. OR US-7174381-$.DID. OR US-7249357-$.DID. OR US-7334232-$.DID. OR US-7437460-$.DID. OR US-7472193-$.DID. OR US-7554949-$.DID. OR US-7937455-$.DID. OR US-8082289-$.DID. OR US-8140612-$.DID. OR US-8402083-$.DID. OR US-8676877-$.DID. OR US-20020049859-$.DID. OR US-20030005266-$.DID. OR US-20030051062-$.DID. OR US-20030135621-$.DID. OR US-20030195931-$.DID. OR US-20030195938-$.DID. OR US-20040110209-$.DID. OR US-20040157203-$.DID. OR US-20040252710-$.DID. OR US-20050021751-$.DID. OR US-20050038852-$.DID. OR US-20050060237-$.DID. OR US-20050076105-$.DID. OR US-20050108394-$.DID. OR US-20050180095-$.DID. OR US-20060053216-$.DID. OR US-20060106931-$.DID. OR US-20070044099-$.DID. OR US-20070073705-$.DID. OR US-20080250347-$.DID. OR US-20080281997-$.DID. OR JP-1234567-$.DID. OR JP-0887473-$.DID. OR JP-2002117010-$.DID. OR US-6131999-$.DID.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/11/29 23:39
S23	4	(computer with cluster) same (nodes) same (single-node with kernel)	USPAT	OR	OFF	2016/06/17 19:19
S24	10	(computer with cluster) same (nodes) same (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S25	11	(computer with cluster) same (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S26	11	(computer with cluster) and (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S27	224	(computer with cluster) and (nodes) and (single with kernel)	US-PGPUB; USPAT; EPO; JPO;	OR	OFF	2016/06/17 19:20

			DERWENT; IBM_TDB			
S28	145	(computer with cluster) same (nodes) and (single with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:21
S29	6	(computer with cluster) same (nodes) and (single with kernel) and 709/223.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:22
S30	10	(computer with cluster) same (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 16:08
S31	79	(computer with cluster) and (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 18:41
S32	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 12:43
S33	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:47
S34	1505	(computer with cluster) and kernel and nodes and request	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S35	271	S34 and mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S36	178750	S34 and mathematical smae kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S37	10	S34 and mathematical same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S38	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:54
S39	29	(single with node with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO;	OR	OFF	2016/07/08 16:10

			DERWENT; IBM_TDB			
S40	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 16:11
S41	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:18
S42	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:24
S43	180800	mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S44	11	mathematica near5 kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S45	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:40
S46	10	("8402080").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:43
S47	2	("8849889").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44
S48	101	("20020049859" | "20030005266" | "20030051062" | "20030135621" | "20030195931" | "20030195938" | "20040110209" | "20040157203" | "20040252710" | "20050021751" | "20050038852" | "20050060237" | "20050076105" | "20050108394" | "20050180095" | "20060053216" | "20060106931" | "20070044099" | "20070073705" | "20080250347" | "20080281997" | "5423046" | "5881315" | "5930768" | "6108699" | "6202080" | "6546403" | "6578068" | "6751698" | "6782537" | "6968335" | "7093004" | "7136924" | "7174381" | "7249357" | "7334232" | "7437460" | "7472193" | "7554949" | "7937455" | "8082289" | "8140612" | "8402083" | "8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44

S49	17	mathematica with cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:50
S50	43	mathematica same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:51
S51	537678	stephen wolfram mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S52	50	(stephen near2 wolfram).in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S53	1	S52 and mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S54	215	mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S55	74	S54 and cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S56	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:44
S57	0	("gridMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S58	0	("gridwithMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S59	19	grid with Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S60	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16

S61	14	gridMathematica and wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:21
S62	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S63	2578	Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S64	215	Mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S65	15749	wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S66	8450	wolfram.in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S67	15	S66 and mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S68	42	S66 and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S69	2	("20070073705").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:44
S70	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:46
S71	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/11/07 13:23
S72	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:24

S73	0	(" (computerwithnodeswithclister)same(peer-to-peer)and@ad<20060101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:47
S74	0	(computer with nodes with clister) same (peer-to-peer) and @ad<"20060101"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:47
S75	3	(computer with nodes with cluster) same (peer-to-peer) and @ad<"20060101"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:48
S76	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/02 17:12
S77	12	("8601101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/02 17:12
S78	44	(computer with cluster) and kernel and (mathematical with expression)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:39
S79	0	(computer with clusters) same kernel same mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:47
S80	30	(computer with clusters) and kernel same mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:47
S81	9	(computer with clusters) same kernel same mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:48
S82	95	(computer with clusters) same mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:50
S83	2	(computer with clusters) same mathematical same expressions	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/18 22:50
S84	0	cluster near5 mathematical near5 evaluate	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:19

S85	2	cluster with mathematical with evaluate	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:20
S86	2	("20050038852").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:22
S87	20	("20050021751" | "20070073705" | "20070094532" | "6968359" | "8601101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:35
S88	70	(US-0005068-$ | US-0015460-$ | US-0026601-$ | US-0059473-$ | US-0097300-$ | US-0098447-$ | US-0124363-$ | US-0154789-$ | US-0188088-$ | US-0254984-$ | US-1234567-$ | US-1368948-$ | US-20030005068-$ | US-20040098447-$ | US-20040254984-$ | US-20050015460-$ | US-20050026601-$ | US-20050097300-$ | US-20050154789-$ | US-20050188088-$ | US-20060059473-$ | US-20070124363-$ | US-25376674-$ | US-5195580-$ | US-6006259-$ | US-6067609-$ | US-6604134-$ | US-7031944-$ | US-7177874-$ | US-7315877-$ | US-7502915-$ | US-7519652-$ | US-7533141-$ | US-7533389-$ | US-7568131-$ | US-7598390-$ | US-7644130-$ | US-7757236-$ | US-7788314-$ | US-7835022-$ | US-8117288-$ | US-9405564-$).DID.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:36
S89	4802	(computer with cluster with nodes)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:49
S90	30	(computer with cluster with nodes) same (mathematica$)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:50
S91	12	(computer with cluster with nodes) same (mathematica$) and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:50
S92	230	(computer with cluster) same (mathematica$)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:51
S93	28	(computer with cluster) same (mathematica$) and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:52
S94	230	(computer with cluster) same (mathematica$)	US-PGPUB; USPAT;	OR	OFF	2018/02/19 11:54

			EPO; JPO; DERWENT; IBM_TDB			
S95	17	(computer with cluster) same (mathematica$) same result$3	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:55
S96	8	(computer with cluster) same (mathematica$ with expression)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 11:56
S97	15	(high with performance) same cluster same (mathematics mathematica)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:07
S98	114	(high with performance) same cluster and (mathematics mathematica)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:08
S99	2	("7418470").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:12
S100	2	("20030195938").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:13
S101	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:42
S102	12	("8601101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:42
S103	2	("20030195938").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 12:42
S104	2	("20030195938").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2018/02/19 13:24

EAST Search History (Interference)

<This search history is empty>

2/19/2018 4:17:24 PM
C:\Users\hasres\Documents\EAST\Workspaces\14181112.wsp

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

CPC- SEARCHED		
Symbol	**Date**	**Examiner**

CPC COMBINATION SETS - SEARCHED		
Symbol	**Date**	**Examiner**

US CLASSIFICATION SEARCHED			
Class	**Subclass**	**Date**	**Examiner**
709	223	07/07/2016	HA

* See search history printout included with this form or the SEARCH NOTES box below to determine the scope of the search.

SEARCH NOTES		
Search Notes	**Date**	**Examiner**
Searched East (US Patents, USPGPubs, EPO, JPO, IBM, Derwent)	07/07/2016	HA
See East Search attached	07/07/2016	HA
Inventor search performed	07/07/2016	HA
Consulted Justin Rider (QAS) regarding Alice 101 rejection	07/08/2016	HA
Searched updated	03/01/2017	HA
Search updated	02/18/2018	HA

INTERFERENCE SEARCH			
US Class/ CPC Symbol	**US Subclass / CPC Group**	**Date**	**Examiner**

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	07/07/2016	03/01/2017	02/18/2018						
	1	-	-	-						
	2	✓	✓	✓						
	3	✓	✓	✓						
	4	✓	✓	✓						
	5	✓	✓	✓						
	6	✓	✓	✓						
	7	✓	✓	✓						
	8	✓	✓	✓						
	9	✓	✓	✓						
	10	✓	✓	✓						
	11	✓	✓	✓						
	12		✓	✓						
	13		✓	✓						
	14		✓	✓						
	15		✓	✓						
	16		✓	✓						
	17		✓	✓						
	18		✓	✓						
	19		✓	✓						
	20		✓	✓						
	21		✓	✓						
	22		✓	✓						
	23		✓	✓						
	24		✓	✓						
	25		✓	✓						
	26		✓	✓						
	27		✓	✓						
	28		✓	✓						
	29		✓	✓						
	30		✓	✓						
	31		✓	✓						
	32			✓						
	33			✓						
	34			✓						
	35			✓						
	36			✓						

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	07/07/2016	03/01/2017	02/18/2018						
	37			✓						
	38			✓						
	39			✓						
	40			✓						
	41			✓						

Bibliographic Data

Application No: 14181112

Foreign Priority claimed: ○ Yes ◉ No

35 USC 119 (a-d) conditions met: ☐ Yes ☑ No | ☐ Met After Allowance

Verified and Acknowledged: Examiner's Signature | Initials

Title: CLUSTER COMPUTING

FILING or 371(c) DATE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
02/14/2014 **RULE**	709	2449	ZTANN.002P1C3

APPLICANTS

Advanced Cluster Systems, Inc., Aliso Viejo, CA,

INVENTORS

Zvi Tannenbaum, Newport Beach, CA, UNITED STATES

Dean E. Dauger, Huntington Beach, CA, UNITED STATES

CONTINUING DATA

This application is a CON of 13423063 03/16/2012 PAT 8676877

13423063 is a CON of 12040519 02/29/2008 PAT 8140612

12040519 is a CIP of 11744461 05/04/2007 PAT 8082289

11744461 has PRO of 60850908 10/11/2006

11744461 has PRO of 60813738 06/13/2006

FOREIGN APPLICATIONS

IF REQUIRED, FOREIGN LICENSE GRANTED**

03/04/2014

STATE OR COUNTRY

UNITED STATES

ADDRESS

KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614
UNITED STATES

FILING FEE RECEIVED

$1,900

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 1 OF 2	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	6,006,259	11-20-1998	Adelman et al.	
	2	6,067,609	05-23-2000	Meeker et al.	
	3	6,195,680	02-27-2001	Goldszmidt et al.	
	4	6,604,134	08-05-2003	Haury	
	5	7,031,944	04-18-2006	Tanioka	
	6	7,177,874	02-13-2007	Jardin	
	7	7,315,877	01-01-2008	Bhanot et al.	
	8	7,502,915	03-10-2009	Jacob et al.	
	9	7,519,652	04-14-2009	Page et al.	
	10	7,533,141	05-12-2009	Nadgir et al.	
	11	7,533,389	05-12-2009	Verbeke et al.	
	12	7,568,131	07-28-2009	Vertes	
	13	7,644,130	01-05-2010	Magro et al.	
	14	7,698,390	04-13-2010	Harkness et al.	
	15	7,757,236	07-13-2010	Singh	
	16	7,788,314	08-31-2010	Holt	
	17	7,835,022	11-16-2010	Matsumoto	
	18	8,117,288	02-14-2012	Bhanot et al.	
	19	9,405,564	08-02-2016	Muellers et al.	
	20	2003/0005068	01-02-2003	Nickel et al.	
	21	2004/0098447	05-20-2004	Verbeke et al.	
	22	2004/0254984	12-16-2004	Dinker	
	23	2005/0097300	05-05-2005	Gildea et al.	
	24	2005/0015460	01-20-2005	Goyle et al.	
	25	2005/0154789	07-14-2005	Fellenstein et al.	
	26	2005/0188088	08-25-2005	Fellenstein et al.	
	27	2006/0026601	02-02-2006	Solt	
	28	2006/0059473	03-16-2006	Moler	
	29	2007/0124363	05-31-2007	Lurie et al.	

Examiner Signature /HERMON ASRES/	Date Considered 02/19/2018

*Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A/

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 2 OF 2	Attorney Docket No.	ZTANN.002P1C3

FOREIGN PATENT DOCUMENTS						
Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	30	EP 1368948	12-10-2003	Chan et al.		
	31	WO 2004/086270	10-07-2004	Katz et al.		

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	32	U.S. Provisional Application No. 60/799,474, filed May 10, 2006, titled "Graphical Interface for Monitoring Status of a Concurrent Computing Process".	

25376674

Examiner Signature	/HERMON ASRES/	Date Considered	04/11/2018

*Examiner: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A/

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Asres, Hermon
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS					
Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear

FOREIGN PATENT DOCUMENTS						
Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	**T[1]**

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	**T[1]**
	1	English translation of MATSUMURA, et al., "Construction of distributed computing system for large-scale matrices intended or reduction of communication blocks," The Special Interest Group Technical Reports of IPSJ, Information Processing Society of Japan, March 5, 1998, Vol. 98, No. 18, pp. 19-24.	T

27161567

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Asres, Hermon
Art Unit	:	2449
Conf. No.	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Pursuant to 37 CFR 1.56, an Information Disclosure Statement listing a reference is provided herewith. Reference number 1 on the accompanying listing was previously submitted as reference number 5 in the Information Disclosure Statement filed on September 22, 2017. Applicant subsequently obtained an English language translation of the reference and submits the translation for consideration.

If the Examiner would like additional information regarding this reference, or if anything is unclear, the Examiner is invited to contact the undersigned for assistance.

No Disclaimers

To the extent that anything in the Information Disclosure Statement or the listed references could be construed as a disclaimer of any subject matter supported by the present application, Applicant hereby rescinds and retracts such disclaimer.

Timing of Disclosure

This Information Disclosure Statement is being filed within three months of the filing date or date of national phase entry, with an RCE or before receipt of a First Office Action after an RCE, and no fee is believed to be required.

The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment, to Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: December 1, 2017

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Registered Practitioner
Customer No. 20995
(949) 760-0404

27161561

Electronic Acknowledgement Receipt

EFS ID:	31109662
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Jennifer Woo
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	01-DEC-2017
Filing Date:	14-FEB-2014
Time Stamp:	20:47:41
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		IDS_ZTANN_002P1C3.pdf	44590 6cc56685a33b083b38b47f1a35df84d59df256f1	yes	3

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Information Disclosure Statement (IDS) Form (SB08)	3	3
	Transmittal Letter	1	2

Warnings:

Information:

2	Non Patent Literature	MATSUMURA_ConstructionOf Distributed_16pages.pdf	45338 e114421e85b1c9922bcad4e54e51808ea5cf7d97	no	16

Warnings:

Information:

Total Files Size (in bytes):	89928

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Tannenbaum, Zvi
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Asres, Hermon
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	2004/0195572	10-07-2004	Kato et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	2	JP 11-328134 w/ English Abstract	11-30-1999	Hitachi Ltd.		
	3	JP 2004-061359 w/ English Abstract	02-26-2004	Ricoh Co. Ltd.		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	4	KONISHI, et al., "Performance Evaluation of Parallel Computing Tool for MATLAB on PC Grid Environment," IPSJ SIG Technical Reports, August 5, 2005, Vol. 2005, No. 84. 9 pages.	
	5	MATSUMURA, et al. "Construction of distributed computing system for large-scale matrices intended or reduction of communication blocks," August 5, 1998, Vol. 98, No. 18. 9 pages.	
	6	TATEBE, et al., "Efficient Implementation of MPI Using Remote Memory Write," Transactions of Information Processing Society of Japan, May 1999, Vol. 40, No. 5. 14 pages.	

Examiner Signature	Date Considered
***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

J-PlatPat
Japan Platform for Patent Information

Help desk (9:00-21:00)
(+81)3-6666-8801
helpdesk@j-platpat.inpit.go.jp

Japanese Top page Help list Site map JPO

National Cen
Industrial Pro
Information a

Patent & Utility Model Design Trademark Trial & Appea

Top page Patent & Utility Model Patent & Utility Model Number Search

Patent & Utility Model Number Search

Back Help Search List

You can retrieve a variety of patent and utility model gazettes by their numbers.

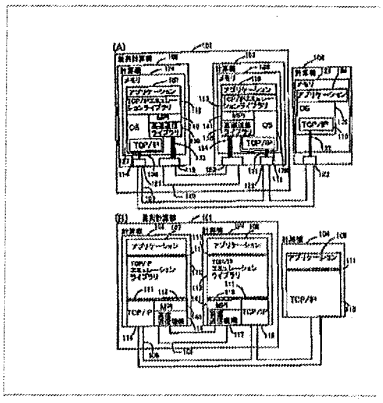
Selected Gazette

Previous Document 1/2 Next Doc

JP,11-328134,A

PAJ Detail Image

(11)Publication number	11-328134
(43)Date of publication of application	30.11.1999
(51)Int.Cl.	G06F 15/163 G06F 13/00
(21)Application number	10-132364
(22)Date of filing	14.05.1998
(71)Applicant	HITACHI LTD
(72)Inventor	IMAKI TSUNEYUKI SAGAWA NOBUTOSHI



(54)METHOD FOR TRANSMITTING AND RECEIVING DATA BETWEEN COMPUTERS

(57)Abstract
PROBLEM TO BE SOLVED: To utilize a message passing type stream communication from a TCP/IP application program.
SOLUTION: An emulation library 112 provided for each computer decides whether or not a communication party process that a communication request from an application program 107 specifies is inside or outside a parallel computer 101 and when the process is inside the computer, a fast communication library 135 is actuated to transmit and receive data by a message passing type communication system through a fast internal communication network 105. The data are transmitted and received so as to actualize a stream communication. When the communication party is an external computer 104 of the parallel computer 101, the library 112 requests a communication of a TCP/IP process routine 114 according to TCP/IP.

Previous Document 1/2 Next Doc

To return to the top of

Notice on Usage Privacy Policy Questionnaire

(19)日本国特許庁（ＪＰ）　(12) 公開特許公報（Ａ）　(11)特許出願公開番号

特開平11－328134

(43)公開日　平成11年(1999)11月30日

(51)Int.Cl.⁶	識別記号	ＦＩ	
Ｇ０６Ｆ　15/163		Ｇ０６Ｆ　15/16	３１０Ｒ
13/00	３５１	13/00	３５１Ｅ

審査請求　未請求　請求項の数12　ＯＬ　（全 23 頁）

(21)出願番号　特願平10－132364

(22)出願日　平成10年(1998)５月14日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地

(72)発明者　今木　常之
東京都国分寺市東恋ケ窪一丁目280番地
株式会社日立製作所中央研究所内

(72)発明者　佐川　暢俊
東京都国分寺市東恋ケ窪一丁目280番地
株式会社日立製作所中央研究所内

(74)代理人　弁理士　高橋　明夫　（外１名）

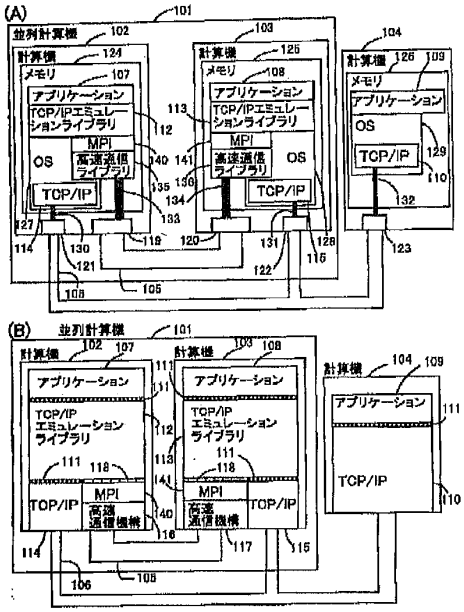
(54)【発明の名称】　計算機間データ送受信方法

(57)【要約】

【課題】ＴＣＰ／ＩＰアプリケーションプログラムからメッセージパッシング型のストリーム通信を利用可能にする。

【解決手段】各計算機に設けられたエミュレーションライブラリ１１２が、アプリケーションプログラム１０７からの通信要求が指定する通信相手プロセスが並列計算機１０１の内部であるか外部であるかを判別し、内部であれは高速通信ライブラリ１３５を起動して、高速内部通信網１０５を介してメッセージパッシング型通信方式でもってデータを送受信する。このデータの送受信は、ストリーム通信を実現するように実行される。相手先が並列計算機１０１の外部計算機１０４であるときには、ライブラリ１１２は、ＴＣＰ／ＩＰ処理ルーティン１１４にＴＣＰ／ＩＰに従い通信を要求する。

図１



1

【特許請求の範囲】

【請求項1】第1の計算機上で実行されている第1のアプリケーションプログラムが発行した複数の送信命令により指定される複数の送信データのそれぞれを、第2の計算機上で実行されている第2のアプリケーションプログラムが発行した複数の受信命令に応答して、第2の計算機上に設けられたエミュレーションライブラリの制御下でメッセージパッシング型の通信で受信し、

上記複数の送信データの連なりからなるひと繋がりのデータの内、上記複数の受信命令により指定されるサイズ部分に区分して得られる部分をそれぞれの受信命令が指定する複数のバッファに格納するように、受信されたそれぞれの送信データを上記エミュレーションライブラリの制御下で処理する計算機間データ送受信方法。

【請求項2】メッセージパッシング型の通信を実行するための第1の通信ライブラリを有する第1の計算機上で実行されている第1のアプリケーションプログラムが発行した複数の送信命令に応答して、それらの送信命令により指定される複数のバッファに保持された複数の送信データの送信をそれぞれ要求する複数の送信命令を、上記第1の計算機に設けられた第1のエミュレーションライブラリにより上記第1の通信ライブラリに対して発行し、

メッセージパッシング型の通信を実行するための第2の通信ライブラリを有する第2の計算機上で実行されている第2のアプリケーションプログラムが発行した複数の受信命令に応答して、上記複数の送信データをメッセージパッシング型の通信で受信するための複数の受信命令を、上記第2の計算機に設けられた第2のエミュレーションライブラリにより上記第2の通信ライブラリに対して発行し、

第1のエミュレーションライブラリが発行した上記複数の送信命令により指定される複数の送信データの連なりからなるひと繋がりのデータを、上記複数の受信命令により指定されるサイズ部分に区分してそれぞれの受信命令が指定する複数のバッファに格納するように、第2のエミュレーションライブラリが発行した上記複数の受信命令により受信される上記複数の送信データの第2の計算機内での格納位置およびそれぞれの送信データの受信後の第2の計算機内での移動を、上記第2のエミュレーションライブラリにより制御する計算機間データ送受信方法。

【請求項3】メッセージパッシング型の通信を実行するための第1の通信ライブラリを有する第1の計算機上で実行されている第1のアプリケーションプログラムが発行した複数回の送信命令に応答して、それらの送信命令により指定される複数のバッファに保持された複数の送信データの送信をそれぞれ要求する複数の送信命令を、上記第1の計算機に設けられた第1のエミュレーションライブラリにより上記第1の通信ライブラリに対して発行し、

2

メッセージパッシング型の通信を実行するための第2の通信ライブラリを有する第2の計算機上で実行されている第2のアプリケーションプログラムが発行した複数の受信命令に応答して、上記複数の送信データをメッセージパッシング型の通信で受信するための複数の受信命令を、上記第2の計算機に設けられた第2のエミュレーションライブラリにより上記第2の通信ライブラリに対して発行し、

第1のエミュレーションライブラリが発行した上記複数の送信命令が指定する複数の送信データを連ねて構成されるひと繋がりのデータを、受信側のアプリケーションプログラムが発行した複数回の受信命令により指定される複数のバッファに、それぞれの受信命令が指定するサイズ部分に区分して格納できるように、上記第2のエミュレーションライブラリにより（a）第2のエミュレーションライブラリが発行した上記複数の受信命令の各々が指定する、受信データを格納するバッファのアドレスを、現在処理中の第2のアプリケーションプログラムが発行した受信命令が指定するバッファに属するアドレスもしくは上記第2の通信ライブラリが利用可能なバッファに属するアドレスとなるように定め、かつ（b）、上記第2の通信ライブラリが利用可能な上記バッファのアドレスを指定する、第2のエミュレーションライブラリが発行した受信命令により受信した送信データのその後の上記第1の計算機内での移動を制御し、

第2のエミュレーションライブラリが発行した上記複数の受信命令は、それぞれ第1のエミュレーションライブラリが発行した上記複数の送信命令の一つに対応して発行され、

第2のエミュレーションライブラリが発行した上記複数の受信命令の各々が指定する受信データのサイズは、第1のエミュレーションライブラリが発行した上記複数の送信命令の内の対応する一つの送信命令が指定する送信データのサイズに等しくなるように設定される計算機間データ送受信方法。

【請求項4】（a）送信側のアプリケーションが指定するいずれかの送信データの送信を要求する、上記メッセージパッシング型の通信により定められた送信命令を送信側の計算機内で発行し、（b）上記送信命令が指定する送信データ長を、その送信データの送信前に、受信側の計算機上で動作する受信側のエミュレーションライブラリにより検知し、（c）受信側のアプリケーションが発行した受信命令によって指定された受信データ長と上記送信データ長とを、上記受信側のエミュレーションライブラリにより比較し、（d）上記比較の結果に応じて、上記受信側の計算機内のエミュレーション用のバッファに上記送信データを受信することを要求するかもしくは上記送信データを上記アプリケーションバッファに受信することを要求する、上記メッセージパッシング型

の通信により定められた受信命令を上記受信側のエミュレーションライブラリにより発行し、（ｅ）上記送信命令と受信側のエミュレーションライブラリにより発行した上記受信命令とに応答して、上記送信データを上記送信側の計算機と上記受信側の計算機により、上記送信側の計算機から上記エミュレーション用のバッファもしくは上記アプリケーションバッファ内に転送する計算機間データ送受信方法。

【請求項５】（ｆ）上記送信データが上記エミュレーションバッファに送信されたときには、そのエミュレーションバッファから上記受信データ長分だけのデータを上記受信命令が指定したアプリケーションバッファ内に、上記受信側のエミュレーションライブラリによりコピーするステップをさらに有する請求項４に記載の計算機間データ送受信方法。

【請求項６】（ｇ）上記ステップ（ｅ）により、上記送信データが上記アプリケーションバッファ内に転送された場合、上記送信側の計算機が後続の送信データに関してステップ（ａ）を実行するのを待ち、（ｈ）上記後続の送信データに関して、ステップ（ａ）が実行された場合、その後続の送信データと上記受信命令が要求した受信データ長の残りの長さのデータに関して上記ステップ（ｂ）以降のステップを実行するステップをさらに有する請求項５に記載の計算機間データ送受信方法。

【請求項７】（ｉ）上記ステップ（ａ）を実行する前に、上記エミュレーションバッファ内に受信済みでいずれのアプリケーションバッファにも転送されていないデータが残っているか否かを判別し、（ｊ）そのデータが残っているときには、そのバッファから上記アプリケーションに上記受信データ長を超えない範囲でコピーするステップをさらに有する請求項６に記載の計算機間データ送受信方法。

【請求項８】第１の通信網で相互に結合されている複数の計算機を有し、少なくともその一部の計算機が第２の通信網に相互に接続されている計算機システムにおいて、（ａ）上記一部の計算機の内の第１の計算機上で動作する第１のアプリケーションプログラムから、第２の計算機上で動作するいずれかの第２のアプリケーションプログラムとの間のデータ送受信に使用する通信路を、第１の通信網を使用するために定められた通信規約に従って動作する処理ルーチンにより決定し、（ｂ）第１のアプリケーションプログラムからの、上記決定された通信路に対する接続要求をエミュレーションライブラリにより処理し、その処理の中で、上記第２のアプリケーションプログラムとの通信に上記第２の通信網を利用可能か否かを判別し、（ｃ）上記第２の通信網が利用可能と判別されたときには、上記通信路を上記エミュレーションライブラリ内に登録し、（ｄ）その後上記第１のアプリケーションプログラムがデータ送受信命令を発行したときに、その送受信命令により指定された通信路が、上記エミュレーションライブラリ内に登録されているか否かを判別し、（ｅ）その指定された通信路が、上記エミュレーションライブラリ内に登録されているときには、上記エミュレーションライブラリにより、上記送受信命令が要求する送受信処理を上記第２の通信網を使用して実行し、（ｆ）その指定された通信路が、上記エミュレーションライブラリ内に登録されていないときには、上記エミュレーションライブラリにより、上記送受信命令が要求する送受信処理の実行を上記処理ルーティンに要求し、上記処理ルーチンにより、その送受信処理を上記第１の通信網を使用して実行するデータ送受信方法。

【請求項９】上記通信規約はＴＣＰ／ＩＰであり、上記通信路の決定は、ソケットの生成およびソケットへの名前付けの処理を含む請求項８記載のデータ送受信方法。

【請求項１０】上記第２の通信網を使用した通信はメッセージパッシング型の通信である請求項８記載のデータ送受信方法。

【請求項１１】第１の通信網で相互に結合されている複数の計算機を有し、少なくともその一部の計算機が第１の通信網よりも高速にデータを転送可能な第２の通信網に相互に接続され、上記一部の計算機の内の第１の計算機上で動作する第１のアプリケーションプログラムから、第２の計算機上で動作するいずれかの第２のアプリケーションプログラムとの間のデータ送受信に、上記第２の通信網を使用可能であるときには、これを使用し、そうでないときには上記第２の通信網を使用する計算機システムにおいて、

第１のアプリケーションプログラムに宛てて他のアプリケーションプログラムから送信されるデータの有無を検出するときに、上記第２の通信網を介して送信されるデータの検出処理と、上記第１の通信網を介して送信されるデータの検出処理とを、ノンブロックで交互に繰り返し、

両検出処理を１回繰り返す毎に両検出処理を中断し、上記第１の計算機上で動作している他のプロセスに処理を譲渡する送信データ検出方法。

【請求項１２】第１の通信網で相互に結合されている複数の計算機を有し、少なくともその一部の計算機が第１の通信網よりも高速にデータを転送可能な第２の通信網に相互に接続され、上記一部の計算機の内の第１の計算機上で動作する第１のアプリケーションプログラムから、第２の計算機上で動作するいずれかの第２のアプリケーションプログラムとの間のデータ送受信に、上記第２の通信網を使用可能であるときには、これを使用し、そうでないときには上記第２の通信網を使用する計算機システムにおいて、

第１のアプリケーションプログラムに宛てて他のアプリケーションプログラムから送信されるデータの有無を検出するときに、二つのスレッドを生成し、

上記第２の通信網を介して送信されるデータの検出処理と、上記第１の通信網を介して送信されるデータの検出処理とを、それぞれのスレッド上で独立に行なうデータ検出方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、複数の種類の通信網により接続された複数の計算機を有する計算機システムにおける、計算機間のデータ送受信方法に係り、特にメッセージパッシング型の通信を実行するのに好適な、計算機間のデータ送受信方法に関する。

【０００２】

【従来の技術】計算機間の通信規約としてＴＣＰ／ＩＰが極めて一般的に使用されている。ＴＣＰ／ＩＰを使用して他のプログラムと通信するように構成されたプログラムを以下ではＴＣＰ／ＩＰアプリケーションプログラムと呼ぶ。このＴＣＰ／ＩＰアプリケーションプログラムからＴＣＰ／ＩＰ以外の通信規約を使用できるようにした計算機システムも存在する。すなわち、計算機システムを構成する計算機が性能の異なる複数のネットワークにより構成され、そのシステムの内の計算機が他の計算機と通信するときに、それらの計算機がその通信にこれらのネットワークのいずれを使用するかに依存して、ＴＣＰ／ＩＰまたは他の通信規約を使用するシステムも提案されている。例えば、バークレー大のSteven H.Rodriguesらが開発したシステムは、ＴＣＰ／ＩＰプロトコルで使用可能な広域のネットワークと、それより狭い領域でより簡潔でオーバヘッドが低い通信規約を使用する局所的なネットワークからなり、局所的なネットワークに接続された計算機同士は、この通信規約を使用して通信し、広域のネットワークに接続された計算機同士は、ＴＣＰ／ＩＰ規約に従って通信する。たとえば、"High-Performance Local-Area Communication With Fast Socket", USENIX '97 Annual Technical Conference pp.257-274)参照。

【０００３】この計算機システムには、ＴＣＰ／ＩＰアプリケーションプログラムからもこの通信規約を利用できるようにするために、ＴＣＰ／ＩＰエミュレーションライブラリが用意されている。上記システムではこのＴＣＰ／ＩＰエミュレーションライブラリは、高速ソケット（Ｆａｓｔ　Ｓｏｃｋｅｔ）と呼ばれている。

【０００４】高速ソケットは、簡潔な通信規約としてワークステーションクラスタ向けに開発されたアクティブメッセージ（Ａｃｔｉｖｅ　Ｍｅｓｓａｇｅ）を用いる。例えば 文献 T. von Eicken, D. E. Culler, S. C. Goldstein, K. E. Schauser "Active Messages: a Mechanism for Integrated Communication and Computation", in Proceedings of the 19th International Symposium on Computer Architecture, May 1992 pp.256-266 参照。アクティブメッセージでは、データを送る側のアプリケーションプログラムが受け手側のアプリケーションプログラムに割り込み、受け手側がその割り込みを契機にデータ受信処理を行うという方法でデータをやり取りする。

【０００５】並列計算機は、複数の計算機が互いに通信しながら協調して１つの問題を解決することを目的とした計算機である。この目的を満たすため、一般に並列計算機の内部の各計算機は高速の内部高速通信網により相互に接続され、並列計算機内の少なくとも一部の計算機は、さらに、ＬＡＮ等のより広域のネットワークにより外部の計算機に接続されている。

【０００６】広域のネットワークに使用される通信規約は主としてＴＣＰ／ＩＰである。並列計算機は、内部高速通信網を使用するために、各計算機に他の内部計算機と内部高速通信網を介して通信するための高速な通信ハードウェアとそのハードウエアを使用するための高速通信ライブラリを設けている。現在、多くの並列計算機に採用されている通信規約は、メッセージパッシング型の通信である。メッセージパッシング型の通信は、送信アプリケーションプログラムが発行する送信命令と受信アプリケーションプログラムが発行する受信命令が一対一に対応付けられた時に通信が行われるという通信である。多くの場合、この通信方式は並列計算機内部における高速通信ハードウェアに適している。このメッセージパッシング型通信を実現するために使用される高速通信ライブラリには、リモートメモリ書き込みライブラリあるいはＰＵＴライブラリと呼ばれているものが主に使用されている。

【０００７】並列計算機内部の通信のオーバヘッドは、高速通信ハードウェアを使用することによりかなり減じることができる。アクティブメッセージで用いる割り込み型の通信方式では、割り込みのオーバヘッドが目立ってしまう。しかるに、メッセージパッシング型通信方式は、割り込みを前提としないので、並列計算機の内部での通信にはアクティブメッセージ型通信よりも適している。

【０００８】

【発明が解決しようとする課題】前述のように、並列計算機の内部高速通信網を使用するための通信規約は一般にはメッセージパッシング型である。しかし、並列計算機をビジネスの分野で利用する場合に使用するビジネス用のアプリケーションプログラムは多くはＴＣＰ／ＩＰ規約を使用するように構成されている。したがって、そのようなアプリケーションプログラムは、そのままではメッセージパッシング型通信方式をそのまま利用することはできない。しかも、メッセージパッシング型通信方式を利用した状態で、ストリーム通信を実現する方法が知られていない。また、割込を使用するアクティブメッセージ型通信で使用されるストリーム通信の実現方法をそのままメッセージパッシング型通信でのストリーム通

信に利用するわけには行かない。

【0009】従って本発明の目的は、メッセージパッシング型通信を実行するように構成されている計算機上で動作する複数のアプリケーションプログラム間でストリーム通信を実行可能にする計算機間データ送受信方法を提供することである。

【0010】本発明のより具体的な目的は、第1の計算機ネットワークを介して、ＴＣＰ／ＩＰのごとき、メッセージパッシング型通信と異なる通信規約を用いて通信を実行するように構成されている複数のアプリケーションプログラム間で、その計算機ネットワークよりも高速な通信網を使用して、かつ、メッセージパッシング型通信を実行可能にする計算機間データ送受信方法を提供することである。

【0011】さらに、本発明の他の目的は、第1の通信網とそれより高速の第2の通信網に接続された計算機上で動作し、他のアプリケーションプログラムとの間で、たとえばＴＣＰ／ＩＰ通信規約のような第1の通信規約に基づいて通信するように構成されたアプリケーションプログラムが、第1の通信網に接続された他の計算機上で動作する他のアプリケーションプログラムとの間で、その通信規約に基づいた通信を実行することを可能にするとともに、第2の通信網に接続された他の計算機上で動作する他のアプリケーションプログラムとの間でその第2の通信網を使用した高速の通信を実行する行うことを可能にする計算機間データ送受信方法を提供することである。

【0012】さらに、本発明の他の具体的な目的は、上記第2の通信網を使用した通信を、メッセージパッシング型の通信とすることができる計算機間データ送受信方法を提供することである。

【0013】

【課題を解決するための手段】上記目的を達成するために、本発明による計算機間データ送受信方法では、第1の計算機上で実行されている第1のアプリケーションプログラムが発行した複数の送信命令により指定される複数の送信データのそれぞれを、第2の計算機上で実行されている第2のアプリケーションプログラムが発行した複数の受信命令に応答して、メッセージパッシング型の通信で受信する。この受信を第2の計算機上に設けられたエミュレーションライブラリにより制御する。

【0014】さらに、上記複数の送信データの連なりからなるひと繋がりのデータの内、上記複数の受信命令により指定されるサイズ部分に区分して得られる部分をそれぞれの受信命令が指定する複数のバッファに格納するというストリーム通信を実現するように、受信されたそれぞれの送信データを処理する。この処理を、上記エミュレーションライブラリにより制御する。

【0015】より具体的には、本発明による計算機間データ送受信方法は、以下の処理を実行する。

【0016】（a）送信側アプリケーションが一回の送信命令によって送信しようとするデータの長さを、受信側エミュレーションライブラリにより検知し、（b）上記の送信データ長を、アプリケーションが一回の受信命令によって指定しているデータ受け取り長と受信側エミュレーションライブラリにより比較し、（c）上記の比較で、送信データ長がデータ受け取り長より長ければ、受信側エミュレーションライブラリが、一旦メモリ上に確保したバッファ領域にデータの全てを受信して、そこからアプリケーションが指定したデータ受け取り長分だけのデータをアプリケーション領域にコピーし、送信データ長がデータ受け取り長以下であれば、受信側エミュレーションライブラリが、アプリケーション領域にデータを直接受信する。

【0017】（d）アプリケーションが受信命令を発行した時に、バッファ領域にデータが残っている場合には、そこからアプリケーション領域にデータをコピーする。

【0018】さらにより具体的には、本発明によるデータ送受信方法を実現するために、ＴＣＰ／ＩＰエミュレーションライブラリを用意する。このライブラリは、ＴＣＰ／ＩＰのソケットアプリケーションプログラムインタフェースと同一のインタフェースを持ち、通信相手が並列計算機の外部であったら従来のシステムコールを、内部であったらＭＰＩ等の並列計算機用メッセージパッシング型通信方式による高速通信網を用いる、という切り分けを行う。すなわち、このＴＣＰ／ＩＰエミュレーションライブラリには以下のような特徴を持たせる。

【0019】（1）並列計算機内部で通信する場合は、メッセージパッシング型通信方式に適した通信手順を用いて、ＴＣＰ／ＩＰと同等のストリーム通信サービスを提供する。

【0020】（2）ＴＣＰ／ＩＰエミュレーションライブラリとして適当な手段を用いて通信方法を切り分ける。

【0021】（3）外部通信のデータと内部通信のデータをスピンループによって検出する。もしくは外部通信のデータと内部通信のデータを別々のスレッドで検出する。

【0022】

【発明の実施の形態】＜従来の技術とその問題点＞本発明の実施の形態を説明する前に、従来の技術とその問題点を説明する。

【0023】（1）ＴＣＰ／ＩＰ

図13にＴＣＰ／ＩＰの階層を示す。ＴＣＰ／ＩＰでは、下からリンク層301，ＩＰ層302，ＴＣＰ層303，アプリケーション層306の4階層を定めている。実際にはリンク層301の下に物理層が存在するが簡単化のためにこの層には言及しない。以下では、簡単化のためにアプリケーション層より下の層303，30

２，３０１からなる層群３０４をまとめてＴＣＰ／ＩＰ層と呼ぶ。一般にＴＣＰ／ＩＰが定める手順で通信を行なうために必要なプログラム（以下、ＴＣＰ／ＩＰ処理ルーチンと呼ぶ）はＯＳに含まれる。ＴＣＰ／ＩＰが定める手順で通信を実行するには、ＴＣＰ／ＩＰ層３０４を構成する複数の層の各々によりそれぞれ特定の処理が実行される。本明細書では、これらの層がそれぞれ実行する複数の処理あるいはそれらの処理を実行する複数のプログラムルーチンを総称してＴＣＰ／ＩＰ処理ルーチンと呼ぶ。

【００２４】アプリケーションプログラムがＴＣＰ／ＩＰ処理ルーチンを利用するためのインタフェースとしては、一般にソケットアプリケーションプログラミングインタフェース（ソケットＡＰＩ）１１１が用いられる。アプリケーションプログラムがＴＣＰ／ＩＰ処理ルーチンを利用するために、ソケットライブラリと呼ばれるシステムコール群がＯＳの一部として設けられている。

【００２５】なお、本明細書では、データ受信のためのシステムコールおよび関数あるいはデータ送信のためのシステムコールおよび関数を呼び出すことを、それぞれの機能と組み合わせて受信命令あるいは送信命令を発行すると記載することがある。ソケットアプリケーションプログラムインタフェースは、ソケットライブラリのアプリケーションプログラムインタフェースとして定められている。ソケットアプリケーションプログラムインタフェースに従って記述された、ＴＣＰ／ＩＰで通信を行なうアプリケーションプログラム（以下、ＴＣＰ／ＩＰアプリケーションプログラムと呼ぶ）の例を以下に示す。

【００２６】

（サーバ側）

```
sa=socket(AF_INET,SOCK_STREAM,0);
bind(sa, server, slen);
listen(sa, 5);
sa1=accept(sa, &client, &clen);
read(sa1, buffer1, length1);
read(sa1, buffer2, length2);
        .
        .
        .
close(sa1);
close(sa);
```

（クライアント側）

```
sb0=socket(AF_INET,SOCK_STREAM,0);
connect(sb0, server, slen);
write(sb0, buffer0, length0);
write(sb0, buffer3, length3);
        .
        .
        .
close(sb0);
```

まず、他と通信を行おうとするサーバ側のＴＣＰ／ＩＰアプリケーションプログラムおよびクライアント側のＴＣＰ／ＩＰアプリケーションプログラムは、それが実行中の計算機に設けられたＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールソケット（ｓｏｃｋｅｔ）を呼び出す。呼び出されたシステムコールソケットは、通信端の役目を果たす、ソケットと呼ばれるオブジェクトを生成し、ソケット記述子を返す。上の例ではサーバ側、クライアント側のＴＣＰ／ＩＰアプリケーションプログラムは、それぞれに返されるソケット記述子をｓａまたはｓｂ０として受け取る。

【００２７】ソケット記述子は、アプリケーションプログラム内でソケット毎に一意に決定されるソケットの識別子（ＩＤ）であり、整数値からなる。生成されたソケットに対する操作は、全てこのソケット記述子を指定して行う。以下、特に断らない限り、ソケットとソケット記述子は同義であるとして説明する。上のシステムコールの引数ＡＦ＿ＩＮＥＴは、アドレスファミリインタネットを表し、インタネットを介した通信にソケットを使用することを示す。さらに、引数ＳＯＣＫ＿ＳＴＲＥＡＭは、ストリーム通信を要求する。最後の引数はプロトコルを指定する引数である。この値が０のときには、プロトコルはその引数より前の二つの引数により決まる。今の場合には、ＴＣＰ／ＩＰプロトコルを使用することになる。ソケットが生成された後、サーバ側とクライアント側のＴＣＰ／ＩＰアプリケーションプログラムは異なる手順でソケットの接続処理を行う。

【００２８】サーバ側は、ＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールバインド（ｂｉｎｄ）を呼び出す。呼び出されたシステムコールバインドは、引数で指定されたソケットｓａに引数で指定された名前を対応付ける。名前は、ＩＰアドレスとポート番号の組合わせからなる。上の例では、ｓｌｅｎで指定される大きさを持つ構造体ｓｅｒｖｅｒに格納されたＩＰアドレスとポート番号の組からなる名前にソケット識別子ｓａが対応付けられる。サーバ側は、さらにＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールリスン（ｌｉｓｔｅｎ）を呼び出す。

【００２９】このシステムコールリスンは、引数で指定されたソケットｓａを接続要求の受領のためのソケットに設定する。このシステムコールの第２の引数は、その接続要求および他の接続要求を一時的に保持するのに用いるキューに要求するサイズを表し、今の場合には、このコールは、５個の接続要求を保持可能なキューを要求

している。

【００３０】以上の処理の結果、論理的には、サーバ側とクライアント側の間に通信路が決定されたことになる。

【００３１】その後、サーバ側は、ＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールアクセプト（ａｃｃｅｐｔ）を呼び出す。このシステムコールは、引数により指定されたソケットｓａを接続要求の待ち状態にする。このシステムコールの第２，第３の引数は、待つべき接続要求の発行元クライアント側のＩＰアドレスと長さを表す。

【００３２】一方、クライアント側は、ＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールコネクト（ｃｏｎｎｅｃｔ）を呼び出す。このシステムコールは、引数で指定された名前、今の場合にはサーバ側のソケットｓａに付けられた名前（構造体ｓｅｒｖｅｒに格納された、ＩＰアドレスｓｉｎ＿ａｄｄｒとポート番号ｓｉｎ＿ｐｏｒｔの組み合わせ）に対して引数で指定されたソケットｓｂ０を接続する。

【００３３】さらに、サーバ側の計算機では、先に呼び出されたシステムコールアクセプトが、この接続要求を受信し、システムコールアクセプトに対する前述の引数が指定するソケットｓａ１をこのクライアント側との通信のためのソケットとして新たに生成する。こうして、サーバ側のソケットｓａ１とクライアント側のソケットｓｂ０との間で通信路が確立されたことになる。

【００３４】ソケット間の接続が確立された状態でクライアント側がサーバ側にデータを送信する時には、クライアント側は、ＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールライト（ｗｒｉｔｅ）を呼び出す。このシステムコールに対する引数では、送信すべきデータを保持するのに用いるバッファのアドレスｂｕｆｆｅｒ０とそのバッファの長さｌｅｎｇｔｈ０を指定する。システムコールライドは、このバッファ内のデータをこのシステムコールが指定するソケットｓｂ０を用いて送信する。サーバ側はＴＣＰ／ＩＰ処理ルーチンに含まれたシステムコールリード（ｒｅａｄ）を呼び出す。このシステムコールは、その引数で指定されたソケットｓａ１を介して転送されたデータを受信し、引数で指定されたアドレスｂｕｆｆｅｒ１のバッファに書き込む。このようにして、２つのアプリケーションプログラムの間でデータが転送される。もし必要ならば、クライアント側は複数の後続のデータをそれぞれ送信するために複数のライトシステムコールを呼び出し、サーバ側はそれらの後続のデータを受信するために一つまたは複数のリードシステムコールを呼び出す。

【００３５】ソケット記述子は、ファイルへの入出力や標準入出力などを行う際に用いられる、ファイル記述子の一種として定義されている。このためソケットアプリケーションプログラムインタフェースでは、ファイル等の入出力を行う場合と同様のインタフェースを通して、データの送受信を行うことができる。

【００３６】サーバ側とクライアント側のプログラムの間の通信が終了したときには、サーバ側は、システムコールクローズ（ｃｌｏｓｅ）を呼び出し、引数で指定されたソケットｓａ１を閉鎖する。サーバ側は、さらに同じシステムコールクローズを再度呼び出し、システムコールクローズは今度はソケットｓａを閉鎖する。クライアント側も同様にシステムコールクローズを呼び出して、このシステムコールはソケットｓｂ０を閉鎖する。

【００３７】（２）　高速ソケットにおける通信方法の切り替え

従来のＴＣＰ／ＩＰでは、システムコールバインドによってＩＰアドレスとポート番号からなる名前（ｓｅｒｖｅｒ）をソケット（ｓａ）に対応付けていた。高速ソケットでは、この時、さらに高速通信専用のソケットを以下のように生成する。まず、アプリケーションがシステムコールバインドの呼び出し時に指定する名前のポート番号（構造体ｓｅｒｖｅｒに格納されたポート番号ｓｉｎ＿ｐｏｒｔ）にハッシュ関数を施して、新たにシャドウポート番号を導き出す。次に、システムコールソケットを呼び出して、シャドウソケットを新規に生成する。最後にバインドシステムコールを呼び出して、シャドウソケットにシャドウポート番号を対応付ける。高速通信を行なう場合は、サーバおよびクライアントがａｃｃｅｐｔおよびｃｏｎｎｅｃｔを呼び出して接続するときに、このシャドウソケットを用いることで高速通信専用の通信路を貼る。

【００３８】この方法では、システムコールバインドやシステムコールリスンが呼び出される時に、特別な処理が必要となる。

【００３９】（３）ＴＣＰ／ＩＰ処理ルーチンによるストリーム通信

ＴＣＰ／ＩＰのソケットライブラリは、ストリーム通信をサービスする。ストリーム通信とは、送信側のアプリケーションプログラムが一連のライトシステムコールにより送る複数のデータをひと繋がりのデータストリームとして処理し、そのデータストリームを受信側のアプリケーションプログラムが発行する一つあるいは複数のリードシステムコールが指定する任意の長さを有する一つあるいは複数のデータに切り分けて受け取るという通信方法である。

【００４０】図１４を用いて従来のストリーム通信の動作を説明する。ここでは、送信アプリケーションプログラム８０１が受信アプリケーションプログラム８０２にデータを送る例を示している。送信アプリケーションプログラム８０１は、まず第１のライトシステムコールを呼び出してそのプログラムのバッファ８０３内の５０キロバイト（ＫＢ）のデータを送信する（８０５）。なお、図１４では単位ＫＢは簡単化のために省略してい

る。送信側のＴＣＰ／ＩＰ処理ルーチンは、バッファ８０３内のこの送信データを一旦ＯＳ内に確保したバッファ（図示せず）にコピーし、複数のパケットに分割して、受信側ＯＳに送信する。 通常パケットのサイズは４０～１５００バイトである。受信側ＯＳは、それらのパケットをＯＳ内に確保した複数のバッファ（図示せず）に受け取り、これらのパケットをリスト状に繋いで、データストリームを再構成する。送信アプリケーションプログラム８０１は、さらに第２のライトシステムコールを発行してそのプログラムの他のバッファ８０４内の８０ＫＢのデータを送信する（８０６）。このときも送信側のＴＣＰ／ＩＰ処理ルーチンと受信側のＴＣＰ／ＩＰ処理ルーチンは同様に動作して、このデータを先に送信されたデータと組み合わせて一つのデータストリームとしてＯＳ内の前述のバッファに保持する。

【００４１】図において、８０９はＯＳ９００内に保持されたこのデータストリームを模式的に表す。先頭のデータ８０７、後続のデータ８０８はそれぞれ上記第１、第２のライトシステムコールにより送信されたそれぞれ５０ＫＢ、８０ＫＢのデータを表す。ストリーム通信では、受信アプリケーションプログラムは、これらのデータを一本の連続した１３０ＫＢのデータストリーム８０９として捉える。送信アプリケーションプログラム８０１内のバッファ８０３、８０４内のデータは、ＯＳ９００内のデータストリーム８０９内のデータ部分８０７，８０９をそれぞれ保持する。このための、受信アプリケーションプログラム８０２は、第１のリードシステムコールを発行して、ストリームデータ８０９内の５０ＫＢの先頭データ８０７の内の３０ＫＢを受信アプリケーションプログラム８０２の３０ＫＢのバッファ８１２に受け取ることを要求する（８１４）。受信側のＴＣＰ／ＩＰ処理ルーチンは、このリードシステムコールに応答して、このデータストリーム８０９から先頭の３０ＫＢのデータ８１０をバッファ８１２にコピーする。受信アプリケーションプログラム８０２がさらに第２のリードシステムコールを発行すると、このシステムコールが指定する長さに従って、ストリームデータ８０９内にある残りの１００ＫＢのデータ８１１をこのシステムコールが指定するバッファ８１３にコピーする（８１５）。

【００４２】(4) アクティブメッセージ型通信でのストリーム通信

上記高速ソケットを使用したシステムでは、アクティブメッセージ型の通信を採用しながらストリーム通信を実現している。すなわち、このシステムでは、送信側のアプリケーションプログラムから送信要求が発行されると、この送信要求は、受信側のアプリケーションプログラムから受信要求が発行されるのを待たないで実行される。送信要求は、受信側の計算機に割り込みを発生し、この割り込みにより割り込みハンドラーが起動され、送信データは、この割り込みハンドラー内のバッファに一旦受信される。受信側のアプリケーションプログラムが受信要求を発行すると、割り込みハンドラーは、既に受信されたデータの内、この受信要求が要求する大きさのデータを、受信側のアプリケーションプログラムのバッファに転送する。もし、受信側のアプリケーションプログラムが要求するサイズが、受信済みのデータのサイズより小さければ、受信要求の処理が終了する。割り込みハンドラーに保持された残りのデータは、受信側のアプリケーションプログラムから新たな受信要求が発行されたときに、そのアプリケーションプログラムのバッファに転送される。

【００４３】逆に、受信側のアプリケーションプログラムが要求するサイズが、受信済みのデータのサイズより大きければ、割り込みハンドラーは、送信側のアプリケーションプログラムがその後データを送信してきたときに、そのデータを、受信側のアプリケーションプログラムに供給する。受信側の割り込みハンドラーは、送信側のアプリケーションプログラムが送信要求を発行する前に、受信側のアプリケーションプログラムから後続の受信要求が発行されたときにも、同様にその後送信側のアプリケーションプログラムが後続のデータをその後送信してきたときに、その後続のデータを受信側のアプリケーションプログラムに供給する。もし、この後続のデータが、上記最初の受信要求が要求する不足のデータのサイズあるいは上記後続の受信要求が要求するデータのサイズより大きければ、割り込みハンドラーは、あまりのデータをさらに後続の受信要求のために保持する。

【００４４】こうして、送信側のアプリケーションプログラムが発行する複数の送信要求により送信される複数のデータを、受信側のアプリケーションプログラムが発行する一連の受信要求に応答して受信側のアプリケーションプログラムに供給する。こうして、この方法では、割り込みハンドラー内のバッファに送信データを一旦保持することにより、ストリーム通信を実現している。

【００４５】しかし、割込を使用しないメッセージパッシング型の通信では、このような方法を使用してストリーム通信を実現できない。

【００４６】以下、本発明に係る計算機間データ送受信方法を図面に示したいくつかの実施の形態を参照してさらに詳細に説明する。なお、以下においては、同じ参照番号は同じものもしくは類似のものを表わすものとする。また、発明の第２の実施の形態以降においては、発明の第１の実施の形態との相違点を主に説明するに止める。

【００４７】＜発明の実施の形態１＞

（１）装置の概要

図１は、本発明に係る計算機間送受信方法を実行するための計算機システムの一例を示す。図において、並列計算機１０１の内部の２台の計算機１０２，１０３と、１台の外部計算機１０４とがお互いに通信網で繋がってい

ると仮定する。実際には、並列計算機１０１の内部および外部の計算機の台数はそれぞれ任意である。内部の計算機１０２と１０３は内部高速通信網１０５で繋がっており、内部の計算機１０２，１０３にはそれぞれ内部高速通信網１０５専用のネットワークインタフェースハードウェア１１９，１２０が存在する。内部高速通信網１０５は、複数のパケットを互いに並列にかつ高速に転送可能なネットワーク、たとえばハイパクロスバスイッチなどにより構成される。また、内部の計算機１０２，１０３の全てと外部の計算機１０４はグローバルな通信網１０６に繋がっており、各計算機にはそれぞれ通信網１０６専用のネットワークインタフェースハードウェア１２１，１２２，１２３が存在する。

【００４８】データの送受信は、各計算機のメモリ１２４，１２５，１２６にロードされているアプリケーションプログラム１０７，１０８，１０９の間で行われる。また、各メモリにはアプリケーションプログラムの他にＯＳ１２７，１２８，１２９がロードされており、それぞれのＯＳの中にはＴＣＰ／ＩＰ処理ルーチン１１４，１１５，１１０が存在する。ＴＣＰ／ＩＰが定める手順で通信を実行するには、ＴＣＰ／ＩＰ層３０４を構成する各層によりそれぞれ処理が実行される。ＴＣＰ／ＩＰ処理ルーチンは、ＴＣＰ／ＩＰが定める手順で通信を実行するためにこれらの層がそれぞれ実行する複数の処理の総称で、これらのＴＣＰ／ＩＰ処理ルーチン１１４，１１５，１１０自体は公知のものと同じであり、既に述べたような、システムコールで呼び出し可能な複数の関数を含んでいる。ＴＣＰ／ＩＰ処理ルーチン１１４，１１５，１１０は、広域通信を目的として、グローバルな通信網１０６専用のネットワークインタフェースハードウェア１２１，１２２，１２３によって通信を行う。

【００４９】計算機１０２のメモリ１２４には、さらにＴＣＰ／ＩＰエミュレーションライブラリ１１２とメッセージパッシング型ライブラリ１４０と高速通信ライブラリ１３５とがロードされている。同様に、計算機１０３のメモリ１２５には、ＴＣＰ／ＩＰエミュレーションライブラリ１１３とメッセージパッシング型ライブラリ１４１と高速通信ライブラリ１３６とがロードされている。

【００５０】計算機１０２，１０３上の高速通信ライブラリ１３５，１３６は並列計算機１０１の内部での高速通信を目的として、内部高速通信網１０５専用のネットワークインタフェースハードウェア１１９，１２０によって通信を行うためのライブラリであり、リモートメモリ書き込みライブラリあるいはＰＵＴライブラリと呼ばれているものが多く使用される。本実施の形態でも高速通信ライブラリ１３５、１３６にはこのＰＵＴライブラリを使用する。しかし、本発明はこのライブラリに限定されるのではなく、他のライブラリたとえばＰＵＴ／ＧＥＴライブラリと呼ばれるライブラリも使用可能である。

【００５１】一般に、高速内部通信網１０５を通した通信はグローバルな通信網１０６を通して通信する場合に比べて格段に速い。そこでＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３は、内部計算機上のアプリケーションプログラム同士がＴＣＰ／ＩＰ通信を行おうとする際には、ＴＣＰ／ＩＰ処理ルーチン１１４，１１５ではなくメッセージパッシング型ライブラリ１４０，１４１と高速通信ライブラリ１３５，１３６と内部高速通信網１０５を使用してメッセージパッシング型の通信を実現するように構成され、それでもって通信の高速化を図る。

【００５２】メッセージパッシング型ライブラリ１４０、１４１は、ＴＣＰ／ＩＰエミュレーションライブラリ１１２または１１３からの要求にしたがって、高速通信ライブラリ１３５または１３６を起動するためのライブラリである。メッセージパッシング型ライブラリ１４０、１４１は、一般的にはアプリケーションプログラム（本実施の形態においてはＴＣＰ／ＩＰエミュレーションライブラリ１１２、１１３）に対してメッセージパッシング型のインタフェースを有するライブラリである。

【００５３】ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３は、さらに、このメッセージパッシング型の通信においても従来のＴＣＰ／ＩＰ処理ルーチンが提供していたのと同じくストリーム通信を実現し、それでもって通信の高速化を図る。ＴＣＰ／ＩＰ処理ルーチン１１４，１１５は、ＯＳ１２７，１２８の機能であるため、従来技術ではアプリケーションプログラムがこれらを利用する際には必ずコンテクストスイッチのオーバーヘッドが発生するが、本実施の形態では、高速通信ライブラリ１３５，１３６を使うので、ＯＳを介さないためオーバーヘッドを回避でき、それにより、より一層の通信の高速化も期待できる。

【００５４】

（２）論理構成と構成要素間のインタフェース図１（Ｂ）は、上記ハードウェア構成を論理構成として表現した図である。この図では、以降の説明に関係の無い部分は全て省略してある。また、高速通信ライブラリ１３５，１３６と内部高速通信専用ハードウェア１１９，１２０をひとまとめにして高速通信機構１１６，１１７と表現している。さらに、各構成要素間のインタフェース１１１および１１８を新たに示している。計算機１０２，１０３，１０４や並列計算機１０１を表す四角はそれぞれ、その中の論理構成要素が一台の計算機または並列計算機１０１上で実行されることを示している。

【００５５】計算機１０４上では従来通り、アプリケーションプログラム１０９がＴＣＰ／ＩＰ１１０と、ソケットアプリケーションプログラムインタフェース１１１でリンクされている。並列計算機１０１の内部の計算機１０２，１０３上で動くアプリケーションプログラム１

０７，１０８は、ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３にソケットアプリケーションプログラムインタフェース１１１でもってリンクされている。また、ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３は、ＯＳの機能である従来のＴＣＰ／ＩＰ処理ルーチン１１４，１１５にソケットアプリケーションプログラムインタフェース１１１でもってリンクされ、同時に、高速通信機構１１６，１１７にＭＰＩ仕様のインタフェース１１８でもってリンクされている。

【００５６】（３）アプリケーションプログラム１０７，１０８

本実施の形態では、並列計算機１０１内のいずれかの計算機１０２上で動作しているアプリケーションプログラム例えば１０７が、いずれかの他の計算機上で動作している他のアプリケーションプログラムと通信する場合、当該他のアプリケーションプログラムが、並列計算機１０１内の計算機１０３上で動作しているアプリケーションプログラム例えば１０８か並列計算機１０１外の計算機１０４上で動作しているアプリケーションプログラム例えば１０９かによって、高速通信機構１１６とＴＣＰ／ＩＰ処理ルーチン１１４を使い分ける。並列計算機１０１内の異なる計算機１０２、１０３上で動作する二つ＊のアプリケーションプログラム１０７、１０８が高速通信機構１１６、１１７を使用して相互に通信するためには、それらのアプリケーションプログラム間でデータを実際に送受信するときだけでなく、それぞれのアプリケーションプログラムのためのソケットを相互に接続するときにも高速通信機構１１６、１１７を利用するように特別の処理を行う必要がある。ＴＣＰ／ＩＰエミュレーションライブラリ１１２または１１３がこの特別の処理を実行するための複数の関数を含む。本実施の形態では、ＴＣＰ／ＩＰエミュレーションライブラリ１１２、１１３内に設けられた関数の名前にＥＭＵ＿という接頭辞をつけ、ＴＣＰ／ＩＰ処理ルーチン１１４、１１５または１１０内に設けられた前述の関数の名前と区別する。

【００５７】具体的には、並列計算機１０１内の計算機１０２、１０３上で動作しているアプリケーションプログラム１０７または１０８の内、サーバ側およびクライアント側として動作するアプリケーションプログラムはそれぞれ以下のプログラムを実行するように生成される。

【００５８】

```
（サーバ側）                          （クライアント側）
sa=socket(AF_INET,SOCK_STREAM,0);
                                  sb0=socket(AF_INET,SOCK_STREAM,0);
bind(sa, server, slen);
listen(sa, 5);
sa1=EMU_accept(sa, &client, &clen);  EMU_connect(sb0, server, slen);
EMU_read(sa1, buffer1, length1);     EMU_write(sb0, buffer0, length0);
EMU_read(sa1, buffer1, length1);     EMU_write(sb0, buffer0, length0);
        ・                                  ・
        ・                                  ・
        ・                                  ・
close(sa1);
close(sa);                           close(sb0);
```

サーバ側アプリケーションプログラムは従来と同様にｓｏｃｋｅｔ、ｂｉｎｄ、ｌｉｓｔｅｎに対するシステムコールを呼び出す。これらのシステムコールは対応するＴＣＰ／ＩＰ処理ルーチンにより従来と同様に処理される。クライアント側のアプリケーションプログラムも従来と同様にｓｏｃｋｅｔシステムコールを発行する。このシステムコールも対応するＴＣＰ／ＩＰ処理ルーチンにより従来と同様に処理される。こうしてサーバ側とクライアント側に対して従来と同様にソケットｓａ、ｓｂ０が生成される。

【００５９】その後サーバ側アプリケーションプログラムは、従来のシステムコールアクセプトに代えて、そのアプリケーションプログラムが動作している計算機内に設けられたＴＣＰ／ＩＰエミュレーションライブラリ内に設けられた関数エミュレーションアクセプト（ＥＭＵ＿ａｃｃｅｐｔ）を呼び出す。クライアント側のアプリケーションプログラムは、従来の関数コネクトに代えて、そのアプリケーションプログラムが動作している計算機内に設けられたＴＣＰ／ＩＰエミュレーションライブラリ内に設けられた関数エミュレーションコネクト（ＥＭＵ＿ｃｏｎｎｅｃｔ）を呼び出す。さらに、サーバ側アプリケーションプログラムは、従来のシステムコールリードに代えて、ＴＣＰ／ＩＰエミュレーションライブラリ内に設けられた関数エミュレーションリード（ＥＭＵ＿ｒｅａｄ）を呼び出し、クライアント側アプリケーションプログラムは、従来のシステムコールライトに代えて、対応するＴＣＰ／ＩＰエミュレーションライブラリ内に設けられた関数エミュレーションライト（ＥＭＵ＿ｗｒｉｔｅ）を呼び出す。以下、これらの新たな関数が行う処理を説明する。

【0060】（4）TCP／IPエミュレーションライブラリによるソケットの接続
図2において、サーバ側アプリケーションプログラム，クライアント側アプリケーションプログラムが、それぞれ上述のEMU__accept，EMU__connect関数を呼び出すと（処理501，502）、これらのシステムコールによってサーバ側のTCP／IPエミュレーションライブラリ内に設けられた関数EMU__acceptとクライアント側のTCP／IPエミュレーションライブラリ内に設けられたEMU__connect呼び出される。これらのシステムコールの引数は、サーバ側アプリケーションプログラム，クライアント側アプリケーションプログラムが、それぞれシステムコールacept、connectを呼び出し、ソケットsa、sb0の接続をTCP／IP処理ルーチン114，115に要求するときと同じである。

【0061】呼び出された関数EMU__acceptとEMU__connectは、まずacceptシステムコール，connectシステムコールをそれぞれ発行する（処理503，504）。これらのシステムコールの引数は、関数EMU__accept，EMU__connectのそれぞれに対する引数がそのまま使用される。これによりサーバ側のTCP／IP処理ルーチン内に設けられたシステムコールacceptとクライアント側のTCP／IP処理ルーチン内に設けられたシステムコールconnectが呼び出され、従来と同様にコールされたシステムコールacceptはコールされたシステムコールconnectによって発行された接続要求を受領し、ソケットsa1が生成され、ソケットsa1とソケットsb0が通信路106を介して接続された状態になる。

【0062】その後、関数EMU__accept，EMU__connectは、それぞれ相手のIPアドレスが並列計算機101の内部の計算機のアドレスであるか否かを確認する（処理506，507）。相手が並列計算機101の内部である場合は、ソケット記述子を内部テーブル410または411に登録する（処理510，511）。今の場合には、サーバ側のアプリケーションプログラム107用のソケットsa1、クライアント側のアプリケーションプログラム用のソケットsb0がそれぞれ内部テーブル410、411にそれぞれ登録される。さらに、TCP／IPエミュレーションライブラリ112、113は、並列計算機101内における各計算機の識別子などの高速通信機構116，117を用いるために必要なデータを交換する。このデータ交換のために、アプリケーションプログラム107、108はそれぞれ関数read，writeシステムコールを発行する（処理512、513、515、516）。ここでのデータ交換は、既に接続されているソケットsa1とsb0を利用し、従来と同様にTCP／IP処理ルーチン114，115と通信路106を介して行われる。その後それぞれサーバ側、クライアント側のアプリケーションプログラム107，108にリターンする（処理518，519）。このとき、サーバ側のTCP／IPエミュレーションライブラリ112はアプリケーションプログラム107に生成されたソケットsa1の識別子を戻す。

【0063】処理506，507において、相手が並列計算機101の内部の計算機でないことが判明したときは、関数EMU__connectはクライアント側のアプリケーションプログラム108に返り、関数EMU__acceptはソケットsa1をサーバ側のアプリケーションプログラム107に返し、そのプログラムに戻る（処理508，509）。こうして、サーバ側のアプリケーションプログラムに対するソケットsa1とクライアント側のアプリケーションプログラムに対するソケットsb0は、それぞれに対応するTCP／IP処理ルーチンと通信路106を介して接続される。

【0064】なお、図3は、本実施の形態によりアプリケーションプログラム107，108，109が生成したソケットがお互いに接続されている一つの状態を示している。ここでは、アプリケーションプログラム107のソケットSA1（402）とアプリケーションプログラム108のソケットSB0（403）とが接続され（407）、アプリケーションプログラム108のソケットSB1（404）とアプリケーションプログラム109のソケットSC0（405）が接続され（408）、アプリケーションプログラム109のソケットSC1（406）とアプリケーションプログラム107のソケットSA0（401）が接続されている（409）。ここで、TCP／IPエミュレーションライブラリ112，113は、内部ソケットテーブル410，411を保持している。内部ソケットテーブ410、411には、ソケットの接続先が内部計算機である場合に、そのソケット記述子を登録する。例えば、アプリケーションプログラム107のソケットSA1（402）の接続先であるソケットSB0（403）は、内部計算機103の上で動くアプリケーションプログラム108のソケットなので、SA1を内部ソケットテーブル410に登録する。同様に、ソケットSB0の接続先であるソケットSA1は、内部計算機102の上で動くアプリケーションプログラム107のソケットなので、SB0を内部ソケットテーブル411に登録する。一方、ソケットSA0（401）の接続先であるソケットSC1（406）は、外部計算機上アプリケーションプログラム109のソケットなので、SA0は内部ソケットテーブル410には登録しない。同様に、SC0に接続されているSB1も内部ソケットテーブル411には登録しない。

【0065】（5）内部通信と外部通信の切り分け方法
その後、サーバ側のアプリケーションプログラムとクラ

イアント側のアプリケーションプログラムはデータ通信を開始する。これらの二つのプログラムの内の一方および他方は、データ送信および受信のために関数ＥＭＵ＿ｗｒｉｔｅ，ＥＭＵ＿ｒｅａｄをそれぞれ呼び出す。先に示したクライアント側とサーバ側のプログラムの例では、サーバ側のアプリケーションプログラムが、関数ＥＭＵ＿ｒｅａｄを呼び出し、クライアント側のアプリケーションプログラムが関数ＥＭＵ＿ｗｒｉｔｅを呼び出している。これらで指定する引数は、ＴＣＰ／ＩＰ処理ルーチン１１４，１１５に含まれたシステムコールｗｒｉｔｅ、ｒｅａｄに対する引数と同じである。

【００６６】図４を参照するに、ＥＭＵ＿ｒｅａｄ，ＥＭＵ＿ｗｒｉｔｅが呼び出されると（処理７０１、７０２）、それぞれの関数は、それぞれの引数で指定されたソケットｓａ１、ｓｂ０のソケット記述子が対応する内部ソケットテーブル４１０、４１１にそれぞれ登録されているか否かを判定する（処理７０３，７０４）。それぞれのソケット識別子が内部ソケットテーブル４１０，４１１にそれぞれ登録されているときには、後に詳細に説明する手順で高速通信機構１１６，１１７、高速内部通信網１０５を用いてメッセージパッシング方式の通信を行う（処理７０７，７０８）。それぞれのソケット識別子が内部ソケットテーブル４１０、４１１に登録されていなかったら、そのまま従来のｒｅａｄ，ｗｒｉｔｅシステムコールをそれぞれ指定されたソケットに対して発行する（処理７０５，７０６）。これらのシステムコールは対応するＴＣＰ／ＩＰ処理ルーチン１１４，１１５により処理され、グローバル通信路１０６を用いた通信がそれらのＴＣＰ／ＩＰ処理ルーチン１１４，１１５により実行される。

【００６７】なお、図５は、図３で示したソケットの接続状態において、各アプリケーションプログラム１０７，１０８，１０９同士が通信を行う場合のデータの流れを示して、上記切り分け方法を説明するための全体構成図である。アプリケーションプログラム１０７とアプリケーションプログラム１０８が通信する場合は、ソケットｓａ１（４０２）とソケットｓｂ０（４０３）を通信端として用いる。これらのソケットは、ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３が保持する内部ソケットテーブル４１０，４１１に登録されている。そこでＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３は、データ通信処理にＴＣＰ／ＩＰ処理ルーチン１１４，１１５ではなく高速通信機構１１６，１１７を利用する（６０１）。一方、アプリケーションプログラム１０８とアプリケーションプログラム１０９が通信する場合は、ソケットＳＢ１（４０４）とソケットＳＣ０（４０５）を通信端として用いる。アプリケーションプログラム１０８とリンクされているＴＣＰ／ＩＰエミュレーションライブラリ１１３が保持する内部ソケットテーブル４１１には、ＳＢ１は登録されていない。そこでＴＣＰ／ＩＰエミュレーションライブラリ１１３は、データ通信処理に、ＴＣＰ／ＩＰ処理ルーチン１１５をそのまま利用する（６０２）。これによって、外部のＴＣＰ／ＩＰ処理ルーチン１１０とのデータ交換が可能となる。

【００６８】本方式により、バインドシステムコール、リスンシステムコールは従来のＴＣＰ／ＩＰ処理ルーチンから変更せずにそのまま用いて、内部通信と外部通信の切り分けを実現できる。

【００６９】（６）メッセージパッシング型ライブラリ１４０，１４１

並列計算機の各計算機に使用される内部高速通信網１０５を使用するための通信ハードウェア１１９、１２０および高速通信ライブラリ１３５，１３６はベンダ特有である場合が多いので、その利用方法もマシンによって様々である。よって、各マシンの高速通信ハードウェアと高速通信ライブラリを利用したアプリケーションプログラムを作ろうとする場合、マシンに特化した汎用性の低いプログラムにならざるを得なかった。これに対して、並列計算機内部の通信ハードウェアを利用して通信するためのライブラリを用意し、このライブラリのアプリケーションプログラムインタフェースを標準として規定することで、アプリケーションプログラムの汎用性を高めようという活動が世界中で活発である。

【００７０】このメッセージ型の通信を使用するための汎用のインタフェースとして現在広く使用されているインタフェースは、ＭＰＩと呼ばれるメッセージパッシングインタフェース（ＭＰＩ－Ｍｅｓｓａｇｅ　Ｐａｓｓｉｎｇ　Ｉｎｔｅｒｆａｃｅ）である。例えば、文献："MPI: Message Passing Interface Standard version1.1", MPI Forum, University of Tennessee, 1995 参照。このインタフェースは、メッセージパッシング型の通信を実現するための上に述べた高速通信ライブラリが存在することを前提としているものであり、このインタフェースを使用しても、メッセージパッシング型の通信は、基本的には上記高速通信ライブラリにより実現されることには変わらない。

【００７１】多くの並列計算機ベンダが、ＭＰＩに準拠したアプリケーションプログラムが並列計算機内部の高速通信ハードウェアを使用できるようにするための高速通信ライブラリを提供している。

【００７２】本明細書では、メッセージ型の通信を使用するためのインタフェースをメッセージパッシング型インタフェースと呼び、そのインタフェースを有するライブラリをメッセージパッシング型ライブラリと呼ぶ。特に、ＭＰＩ仕様のインタフェースをＭＰＩあるいはＭＰＩインタフェースあるいはメッセージパッシングインタフェースと呼び、そのインタフェースを有するライブラリをＭＰＩライブラリあるいはメッセージパッシングインタフェースライブラリと呼ぶことがある。

【００７３】本実施の形態では、メッセージパッシング型ライブラリ１４０、１４１として多くの並列計算機で利用可能である標準のＭＰＩ仕様により定められたインタフェースでもってコマンドあるいはデータを交換するライブラリを使用する。それでもって、ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３の汎用性を高める。しかし、他の仕様のインタフェースを使用してもよい。

＊【００７４】ＭＰＩに準拠した従来のアプリケーションプログラムの記述例を以下に示す。本実施の形態では、ＴＣＰ／ＩＰエミュレーションライブラリ１１２、１１３は、メッセージパッシング型ライブラリ１４０、１４１を起動する部分に関しては、以下に示すプログラム部分を有する。このプログラム部分のより詳細は、後にふれる。

【００７５】

```
MPI_Init(&argc, &argv);
MPI_Comm_rank(MPI_COMM_WORLD, &rank);
if ( rank == 0 )          /* receiver   */
  MPI_Recv(buffer, length, type, 1, tag, comm, status);
else if ( rank == 1)      /* sender   */
  MPI_Send(buffer, length, type, 0, tag, comm);
MPI_Finalize();
```

ＭＰＩでは、通信しようとする全プロセスを一斉に立ち上げる。この時、各プロセスにはランクと呼ばれるプロセス識別子が決定される。上の例では、２つのプロセス、すなわち送信側プロセスｓｅｎｄｅｒと受信側プロセスｒｅｃｅｉｖｅｒとを立ち上げる場合を示している。この時、それぞれのプロセスにはランク０と１が付けられる。各プロセスは、まずＭＰＩ初期化関数ＭＰＩ＿Ｉｎｉｔによってｍｐｉの初期化を行い、ＭＰＩ通信ランク関数ＭＰＩ＿Ｃｏｍｍ＿ｒａｎｋによって自分のランクを取得する。その後、各ランクごとの処理を行う。上の例では、ランク１のプロセスがＭＰＩ送信関数ＭＰＩ＿Ｓｅｎｄによって送るデータを、ランク０のプロセスがＭＰＩ受信関数ＭＰＩ＿Ｒｅｃｖで受け取る場合を示している。

【００７６】（７）高速通信機構を用いたメッセージパッシング型の通信

高速通信機構１１６，１１７を用いるデータ受信処理７０７，データ送信処理７０８の詳細を説明する前に、これらの処理の実行に必要な、高速通信機構１１６，１１７を用いたメッセージパッシング型の通信の概要を説明する。

【００７７】一般にメッセージパッシング型の通信を実行するには、送信側の計算機がデータを送信するための送信関数を実行し、受信側の計算機がそのデータを受信するための受信関数を実行する。後続のデータをそれらの計算機の間で転送するために、それらの関数が繰り返し実行される。本実施の形態では、メッセージパッシング型の通信の具体的な説明として、ＭＰＩ仕様によるメッセージパッシング通信を説明する。

【００７８】ＭＰＩ仕様によるメッセージパッシング通信では、送信側の計算機は送信関数ＭＰＩ＿Ｓｅｎｄを実行し、受信側の計算機は受信関数ＭＰＩ＿Ｒｅｃｖを実行する。今の場合、送信処理７０８が起動されると、送信側のＴＣＰ／ＩＰエミュレーションライブラリ１１３は、先に記載したＭＰＩ仕様のアプリケーションプログラムの例にあるように、送信関数ＭＰＩ＿Ｓｅｎｄを呼び出す。なお、送信関数ＭＰＩ＿Ｓｅｎｄを呼び出す前に初期処理としてＭＰＩ＿Ｉｎｉｔ、ＭＰＩ＿Ｃｏｍｍ＿ｒａｎｋ等の関数の呼び出しを実行する必要があるが、これらの処理は、先に記載した関数ＥＭＵ＿ａｃｃｅｐｔ，ＥＭＵ＿ｃｏｎｎｅｃｔ内の内部テーブル登録処理（５１０，５１１）と同時に行なう。一方、受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２は、内部高速通信網１０５を用いるために受信関数ＭＰＩ＿Ｒｅｃｖを呼び出す。

【００７９】送信側のＴＣＰ／ＩＰエミュレーションライブラリ１１３が上記送信処理７０８内で呼び出す関数ＭＰＩ＿Ｓｅｎｄの引数で指定するバッファアドレスｂｕｆｆｅｒおよびバッファ長ｌｅｎｇｔｈは、送信側のアプリケーションプログラム１０８が発行した関数ＥＭＵ＿ｗｒｉｔｅの対応する引数に等しくされる。一方、受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２が受信処理７０７内で呼び出す関数ＭＰＩ＿Ｒｅｃｖの引数で指定するバッファアドレスｂｕｆｆｅｒおよびバッファ長ｌｅｎｇｔｈは、後に説明するように受信側のアプリケーションプログラム１０７が発行した、関数ＥＭＵ＿ｒｅａｄの引数が指定するバッファアドレスおよびバッファ長もしくは受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１３内に設けられる特定のバッファのアドレスおよびサイズに等しくされる。

【００８０】送信側のメッセージパッシング型ライブラリ１４１内の関数ＭＰＩ＿Ｓｅｎｄは、起動されると、対応する高速通信ライブラリ１３６に対し、関数ＭＰＩ＿Ｓｅｎｄに対する引数が指定するアドレスのバッファからその引数が指定する長さのデータを内部高速通信網１０５を介して受信側の計算機１２４に転送することを要求する。受信側のメッセージパッシング型ライブラリ１４０内の関数ＭＰＩ＿Ｒｅｃｖは、起動されると、対応する高速通信ライブラリ１３５に対し、その高速通信ライブラリ１３５に、この転送されたデータを内部高速

通信網１０５を介して受信し、関数ＭＰＩ＿Ｒｅｃｖに対する引数が指定するアドレスのバッファにその引数が指定する長さのデータを書き込むことを要求する。

【００８１】起動された送信側の高速転送ライブラリ１３６と受信側の高速転送ライブラリ１３５は、要求されたデータの送信と受信をそれ自体公知の方法により内部高速通信網１０５を介して行う。このデータ転送は具体的には以下のように行われる。一般にはリモートメモリ書き込みコマンドあるいはＰＵＴコマンドといわれるコマンドが使用される。以下では、このコマンドをＰＵＴコマンドと呼ぶ。内部高速通信網１０５を介したデータ転送を行う通信方法は３つのＰＵＴコマンドにより行われる。

【００８２】まず、送信側のＴＣＰ／ＩＰエミュレーションライブラリ１１３から関数ＭＰＩ＿Ｓｅｎｄが呼び出されると、送信側のメッセージパッシング型ライブラリ１４１は、この引数が指定したバッファ長（これはすなわち送信すべきデータの長さである）を含むデータの属性情報を含むヘッダの送信を送信側の高速転送ライブラリ１３６に要求する第１のＰＵＴコマンドを発行する。この高速転送ライブラリ１３６は、受信側の計算機１０２にこのヘッダを送信する。計算機１０２内の内部高速通信専用ハードウェア１１９が、このデータを受信側のメモリ１２４内の所定の位置に直接書き込む。

【００８３】一方、受信側ＴＣＰ／ＩＰエミュレーションライブラリ１１２から受信命令ＭＰＩ＿Ｒｅｃｖが呼び出されると、受信側のメッセージパッシング型ライブラリ１４０は、まず送信側の計算機１０３からヘッダがすでに送られてきているか調べる。このために後述するデータ検査命令ＭＰＩ＿ｐｒｏｂｅが使用される。ヘッダが送られてきていないときには、その受信命令の実行は終了する。ヘッダがすでに送られてきているときには、そのヘッダ内のバッファ長から送信データが受信可能か否かを判定する。すなわち、その送信データの長さが、上記受信命令ＭＰＩ＿Ｒｅｃｖが指定したバッファ長以下であるか否かが判定される。ＭＰＩ仕様のメッセージパッシング通信に限らず、一般にメッセージパッシング型の通信では、送信命令が指定した送信データ長が、受信命令が指定した受信バッファサイズより大きいときには、受信側の計算機はその送信データを受信しないようになっている。もし受信側の計算機が受信側のバッファのサイズを超える受信データを受信したときには、そのバッファ以外のメモリ領域が受信データにより破壊する恐れがあるためである。受信側のメッセージパッシング型ライブラリ１４０は、上記判定の結果、送信データが受信可能であると判断したときには、受信命令ＭＰＩ＿Ｒｅｃｖに対する引数が指定した受信バッファのアドレスの送信を受信側の高速通信ライブラリ１３５に要求する第２のＰＵＴコマンドを発行する。この高速通信ライブラリ１３５はこのコマンドに対する応答としてバッファアドレスを送信側の計算機１０３に送信する。１０２内の内部高速通信専用ハードウェア１１９は、このアドレスを送信側のメモリ１２５内の所定の位置に書き込む。なお、上記判定の結果、送信データが受信可能でないときには、受信側のメッセージパッシング型ライブラリはバッファドレスを返さないで、エラーを送信側の計算機に通知する。

【００８４】最後に、送信側のメッセージパッシング型ライブラリは、送信データと受信されたバッファアドレスの送信を送信側の高速転送ライブラリ１３６に要求する第３のＰＵＴコマンドを発行する。高速転送ライブラリ１３６は、このデータとバッファアドレスを受信側の計算機１０２に送信する。受信側の計算機１０２内の内部高速通信専用ハードウェア１１９は、このバッファアドレスとデータを受信し、そのアドレスを有するバッファに受信したデータを直接書き込む。こうして、データの転送が終了する。

【００８５】なお、内部高速通信網１０５を介したデータ通信方法は以上の方法に限らず他の方法も使用可能である。たとえば並列計算機によっては、高速通信ライブラリはＰＵＴコマンドの他にリモートメモリ書き込みコマンドあるいはＧＥＴコマンドも実行可能である。このような並列計算機の場合には、送信側のメッセージパッシング型ライブラリ１４１が上記第３のＰＵＴコマンドを発行するのに代えて、受信側のメッセージパッシング型ライブラリ１４０がＧＥＴコマンドを発行し、受信側の高速通信ライブラリ１３５が送信側のメモリ１２５から送信データを読み出す処理を実行する。

【００８６】（８）メッセージパッシング型の通信におけるストリーム通信

図４に示されたデータ受信処理７０７，送信処理７０８の詳細を図７、８に示した具体例を適宜参照して説明する。図４において、すでに述べたように送信側のアプリケーションプログラム１０８が、関数ＥＭＵ＿ｗｒｉｔｅを呼び出した結果（７０２）、送信処理７０８が起動されると、メッセージパッシング型ライブラリ１４１に含まれた関数ＭＰＩ＿Ｓｅｎｄが呼び出される（７８１）。図７に示した例では、関数ＥＭＵ＿ｗｒｉｔｅ呼び出し時（７０２）の引数ではバッファア８０３のアドレスＳｂｕｆｆと、５０キロバイト（ＫＢ）のバッファサイズを指定すると仮定する。関数ＭＰＩ＿Ｓｅｎｄ呼び出し時（７８１）の引数が指定するバッファアドレスとバッファサイズはこれらの値に等しくされる。以下では使用する関数の他の引数の説明は簡単化のために省略する。また、指定するバッファサイズの単位はＫＢであると仮定し、関数の引数を図示するときには、この単位ＫＢは簡単化のために図示しない。メッセージパッシング型ライブラリ１４１はこの関数呼び出し７８１に応答して、データの転送のために、既に述べたように高速通信機構１１７にヘッダの送信を指示し、その後データの

送信を指示する。

【００８７】一方、図４において、すでに述べたように受信側のアプリケーションプログラム１０７が、関数ＥＭＵ＿ｒｅａｄを呼び出した結果（７０１）、受信処理７０７が起動されると、この受信処理７０７では図６に示す処理がなされる。なお、図７に示した例では、関数ＥＭＵ＿ｒｅａｄの呼び出し７０１の引数が指定するバッファアドレスはＲｂｕｆｆであり、バッファサイズは３０ＫＢであり、送信関数ＥＭＵ＿ｗｒｉｔｅの呼び出し７０２が指定したバッファサイズより小さいサイズを指定していると仮定する。

【００８８】この受信処理７０７では、まず、受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２内に設けられるデータ受信用の特定のバッファ（後に示すバッファ９０１）に受信済みでまだ受信側のアプリケーションプログラムに転送されていないデータが残っているかを判定する命令を発行する（処理２０２）。この命令は図７には示されていない。今仮定しているように最初に関数ＥＭＵ＿ｒｅａｄが実行されたときにはこの判定の結果は否定的となる。その後、送信データ検知命令ＭＰＩ＿ｐｒｏｂｅ７７１が発行され、受信側のアプリケーションプログラムに宛てて送信されようとするデータがあるかを判別する（処理２０６）。この命令は、具体的には、このアプリケーションプログラム１０７に宛てて送信されるべきデータに関する、既に説明したヘッダが高速通信機構１１６により受信済みであるか否かを判定する命令である。もしこのヘッダが送信側のアプリケーションプログラム１０８からまだ送信されていない場合には、受信処理７０８は、処理を終了し受信側アプリケーションプログラム１０７に戻る（処理２０７）。受信側アプリケーションプログラム１０７は受信が失敗したときの処理を実行する。たとえば、受信が成功するまで関数７０１を繰り返し呼び出す。

【００８９】ヘッダが送信側のアプリケーションプログラム１０８からすでに送信済みであると仮定すると、そのヘッダが指定する送信データが受信側のアプリケーションプログラムのバッファ８１２に入りきるか否かが判定される（処理２０８）。図７の例では送信データのサイズは、５０ＫＢであり、受信側のバッファ８１２のサイズは３０ＫＢであり、この判定の結果は否定的となる。このような場合に、受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２が、この送信データを受信側のアプリケーションプログラム１０７のバッファ８１２に受信するＭＰＩ＿Ｒｅｃｖ命令を発行すると、受信側のメッセージパッシング型ライブラリ１４０は、通常はＭＰＩ仕様によりこの命令をエラーとして処理するかもしくは送信データの内、受信側のアプリケーションプログラムのバッファ８１２に入りきらない分を捨ててしまう。

【００９０】これを防ぐために、本実施の形態では受信側アプリケーションプログラム１０７とリンクしている受信側ＴＣＰ／ＩＰエミュレーションライブラリ１１２内に、特別なバッファ９０１を用意し、受信側ＴＣＰ／ＩＰエミュレーションライブラリ１１２はここに送信データをこのバッファ９０１に一旦受信することを要求する。すなわち、このバッファ９０１のアドレスＥｂｕｆｆと全送信データのサイズ５０ＫＢとを指定する、関数ＭＰＩ＿Ｒｅｃｖを呼び出す（７７２）。

【００９１】この関数呼び出し７７２に応答して、受信側の高速通信機構１１６と送信側の高速通信機構１１７は、すでに述べたようにしてデータを送受信し、受信側の高速通信機構１１６は、このデータを上記バッファ９０１に書き込む。受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２はその後、この受信データの内、関数ＥＭＵ＿Ｒｅａｄ呼び出し時の引数で指定されたサイズ３０ＫＢのデータを受信側のアプリケーションプログラム１０７のバッファ８１２にコピーするメモリコピー命令（ＭＥＭＣＰＹ（Ｒｂｕｆｆ、Ｅｂｕｆｆ、３０ＫＢ））７７３を発行する（処理２０９）。その後処理はアプリケーションプログラムに戻る（処理２１１）。こうして、ＴＣＰ／ＩＰエミュレーションライブラリ１１２内のバッファ９０１に受信されたデータの一部９０６が残っている状態で、アプリケーションプログラム間の送受信が完了する。

【００９２】その後さらに送信側アプリケーションプログラム１０８がアドレスＳｂｕｆｆ’のバッファ８０４のデータ８０ＫＢを送信するために関数ＥＭＵ＿ｗｒｉｔｅ（Ｓｂｕｆｆ’，８０ＫＢ）（７９０）（図７（Ｂ））を呼び出すと、同様にして送信処理７０８が実行され、この処理の中で関数ＭＰＩ＿Ｓｅｎｄ（Ｓｂｕｆｆ’，８０）（７９１）が呼び出される。一方、受信側アプリケーションプログラム１０７もアドレスＲｂｕｆｆ’のバッファ８１３に１００ＫＢのデータを受信するために関数ＥＭＵ＿ｒｅａｄ（Ｒｂｕｆｆ’，１００ＫＢ）（７７５）を呼び出すと、これに対しても同様に受信７０７が実行される。

【００９３】この受信処理７０７は、最初の判定処理２０２における判定では、受信側のＴＣＰ／ＩＰエミュレーションライブラリ１１２内の受信用のバッファ９０１内に未転送のデータ９０６が残っていると判断される。その結果、メモリコピー命令ｍｅｍｃｐｙ（Ｒｂｕｆｆ’，Ｅｂｕｆｆ＋３０ＫＢ，２０ＫＢ）（７７６）を発行して、このデータ９０６を受信側のアプリケーションプログラム１０７内のバッファ８１３の先頭領域９０７にコピーする（処理２０３）。このメモリコピーするデータの長さは、バッファ９０１に保持されているデータの長さの内、関数ＥＭＵ＿ｒｅａｄに対する上記第２の関数呼び出しが指定するバッファ長を超えない範囲に設定される。今の場合にはバッファ９０１に保持されているデータの長さが２０ＫＢであり、受信を要求された

データの長さが１００ＫＢより小さいので、このデータ２０ＫＢが全てコピーされる。

【００９４】その後受信側アプリケーションプログラム１０７が要求する８０ＫＢの残りのデータをさらに受信するために、処理２０６が実行される。この処理２０６ではすでに説明したように、送信データがあるか否かをか調べる命令ＭＰＩ＿Ｐｒｏｂｅ（）が発行される。具体的には、送信データに対するヘッダが受信済みであるか否かが判定される。今の場合に、送信側の関数呼び出し７９１がすでに実行済みであると仮定すると、８０ＫＢの送信データがあることが判明する。その場合には、この送信データのサイズが受信側アプリケーションプログラム１０２が指定するバッファ８１３に入りきるか否かが判定される（処理２０８）。今の場合には受信側のアプリケーションプログラムのバッファ８１３の残りの領域９０８のサイズは８０ＫＢなので、送信データはバッファ８１３のこの領域９０８に入りきる。

【００９５】この結果、関数ＭＰＩ＿Ｒｅｃｖ（Ｒｂｕｆｆ’＋２０ＫＢ，８０ＫＢ）（７７８）が呼び出される。この関数呼び出しは、バッファ８１３の残りの領域９０８のアドレスと送信データのサイズ８０ＫＢを指定する。こうして、この送信データがバッファ８１３の領域９０８に直接受信される（処理２１０）。その後、受信を要求されたデータが全て受信されたか否かが判定される（処理２１２）。今の場合は判定の結果が肯定的であるので、受信処理７０７は終了し、処理はアプリケーションプログラムに戻る（処理２１１）。

【００９６】以上の手順により、複数の関数ＥＭＵ＿ｗｒｉｔｅの呼び出し（７０２，７９０）が指定する送信データをひと繋がりのデータストリームとして複数の関数の呼び出しＥＭＵ＿ｒｅａｄ（７０１、７７５）により受信する、ストリーム通信を実現することができる。

【００９７】送信側のアプリケーションプログラムが指定する送信データのサイズが受信側のアプリケーションプログラムが指定するバッファのサイズよりも小さいときでも、以下のようにしてストリーム通信が簡単に実現される。たとえば、送信側のアプリケーションプログラムが５０ＫＢのデータの送信を繰り返し要求し、受信側のアプリケーションプログラムが１００ＫＢのデータの受信を要求する場合のストリーム通信を図８を参照して説明する。

【００９８】送信側のアプリケーションプログラムが呼び出す関数ＥＭＵ＿ｗｒｉｔｅ（７０２）に対する送信処理７０８（図４）の中で、関数ＭＰＩ＿ｓｅｎｄ（７８１）が呼び出される。この関数呼び出しでは、送信側アプリケーションプログラムのバッファ８０３のアドレスＳｂｕｆｆとサイズ５０ＫＢを指定する。

【００９９】受信側のアプリケーションプログラムが呼び出す関数ＥＭＵ＿ｒｅａｄ（７０１）に対する受信処理７０７（図４）も、すでに述べたように図６に従い処理される。今の仮定では処理２０２での判定は失敗する。処理２０６において、データ検査命令７７１が実行されたときに、送信データがあると判定されたと仮定する。今の場合には受信側のバッファ８１２のサイズは、送信側のバッファのサイズより大きいので、処理２０８での判定の結果は肯定的となる。その結果、処理２１０が実行される。この処理では、送信データを受信側アプリケーションプログラムが指定したバッファ８１２に直接受信するための関数ＭＰＩ＿ｒｅｃｖ（７７４）が呼び出される。この関数呼び出しは、受信側アプリケーションプログラムのバッファ８１２の先頭アドレスＲｂｕｆｆと送信側のバッファ８０３のサイズ５０ＫＢを指定する。こうして、送信側のバッファ８０３内の全データが、受信側のバッファ８１２内の先頭の５０ＫＢの領域９０７に書き込まれる。次に処理２１２が実行される。今の場合、受信を要求されたデータのサイズは１００ＫＢであるのに対して、すでに受信されたデータのサイズは５０ＫＢである。したがって、要求されたデータの一部がまだ受信されていない。したがって、判定２１２の結果は否定的となり、残りのデータを受信するために処理２０６が再度実行される。

【０１００】もし、送信側のアプリケーションプログラムが次に関数ＥＭＵ＿ｗｒｉｔｅ７９０を呼び出せば、それに対する送信処理７０８（図４）の中で、関数ＭＰＩ＿ｓｅｎｄ（７９１）が同様に呼び出される。この関数呼び出しでも、送信側アプリケーションプログラムの次のバッファ８０４のアドレスＳｂｕｆｆ’とサイズ５０ＫＢを指定する。

【０１０１】上記処理２０６を繰り返したときに、すでに送信信側のアプリケーションプログラムが上記次の関数ＥＭＵ＿Ｓｅｎｄ７９１を呼び出していたならば、処理２０６での判定結果は肯定的となり、判定処理２０８に移る。今の場合には、受信側のバッファ８１２の残りの領域９１０のサイズは送信されようとするデータのサイズに等しいので、この判定の結果は肯定的となる。その結果、処理２１０が実行され、送信データを受信側のバッファ８１２の残りの領域９１０に直接書き込むための第２の受信関数７７９が呼び出される。この関数呼び出しでは、受信側アプリケーションプログラムのバッファ８１２の残りの領域９１０のアドレスＲｂｕｆｆ＋５０ＫＢと送信データのサイズ５０ＫＢとを指定する。こうして、処理２１２において、要求された全てのデータの受信が完了したと判断されるので、受信処理７０７は完了する。なお、上記処理２０６が繰り返し実行された時点で送信データが存在しないときには、受信処理７０７は終了し、処理は受信側アプリケーションプログラムに戻る。また、上記処理２０８が繰り返された時点で、処理２０８での判定結果が否定的であるときには、処理２０９が実行される。この処理の内容は、すでに説明したものと同じである。以上のごとく、送信側のアプリケ

ーションプログラムが発行した送信命令ＥＭＵ＿ｗｒｉｔｅが指定するバッファのサイズと受信側のアプリケーションプログラムが発行した受信命令ＥＭＵ＿ｒｅａｄが指定するバッファのサイズが異なっていても、また、送信側のアプリケーションプログラムが発行する送信命令ＥＭＵ＿ｗｒｉｔｅの数と受信側のアプリケーションプログラムが発行する受信命令ＥＭＵ＿ｒｅａｄの数が異なっていてもバッファストリーム通信が実現されることが分かる。

【０１０２】以上の説明から分かるように、本実施の形態によれば、並列計算機内部の様に、メッセージパッシング型通信方式の高速通信機構が提供されている計算機で動くアプリケーション同士が、ＴＣＰ／ＩＰを用いてデータ通信を行なう際に、高速通信機構の特徴を活かした高速通信が可能となる。また、それ以外の計算機上で動くアプリケーションとは、従来通りのＴＣＰ／ＩＰによる通信を保証する。利用者は、既存のＴＣＰ／ＩＰアプリケーションを一切変更する必要がない。

【０１０３】<発明の実施の形態１の変形例>本発明は、実施の形態１の内容に限定されるのではなく、以下に例示する変形例および他の変形例を含めいろいろの実施形態により実施できる。

【０１０４】（１）広域ネットワークの通信規約としてＴＣＰ／ＩＰを使用したが、これに代えて他の通信規約を用いることもできる。そのときには、ＴＣＰ／ＩＰ処理ルーチン、エミュレーションライブラリを変更する必要があるのは言うまでもない。

【０１０５】（２）実施の形態１では、メッセージパッシング型ライブラリを使用したが、これを使用しないことも可能である。このときには、エミュレーションライブラリは、直接高速通信ライブラリを呼び出すことになる。

【０１０６】（３）さらに、この通信ライブラリをなくすことも可能である。たとえば、これに代えて、専用の回路を使用することもできる。

【０１０７】（４）実施の形態１では、内部計算機は全て広域通信網に接続されると想定した。しかし、一部の内部計算機が広域通信網に接続されている場合にも同様に本発明を適用できる。

【０１０８】（５）実施の形態１では、広域通信網と内部高速通信網の両方を利用することを前提とした。しかし、本発明によるストリーム通信それ自体は、メッセージパッシング型の通信を実行可能な計算機間に適用できるものであり、したがって、このストリーム通信を実行するには、ＴＣＰ／ＩＰ通信を使用しないでメッセージパッシング型の通信のみを使用するアプリケーションプログラム間の通信にも適用できる。この場合には複数種類の通信網を使用しなくてもよい。その際には、送信側のエミュレーションライブラリは実質的には使用しない変形例も可能である。

【０１０９】<発明の実施の形態２>ＴＣＰ／ＩＰ処理ルーチンには従来からアプリケーションプログラムが使用可能な関数としてｓｅｌｅｃｔ関数が設けられている。そもそもソケット記述子はファイル記述子の一種として定義されている。このファイル記述子が指定するオブジェクトから、データを取得することが可能であるかどうかを調べるためのシステムコールとして、ｓｅｌｅｃｔ関数が用意されている。例えば、あるソケットから受け取り可能なデータが送信側から送られてきている（あるいは送られようとしている）かどうかも、ｓｅｌｅｃｔ関数によって調べることができる。具体的にはあるソケットを割り当てられているアプリケーションプログラムが、送信システムコールｗｒｉｔｅを発行したか否かが判定できる。ｓｅｌｅｃｔ関数では、見張ろうとするファイル記述子をビット列で指定する。このビット列の各ビットはそれぞれ個別のファイル記述子に対応しており、ビットを１にすることでファイル記述子を指定する。複数のビットを１にすれば、一回のｓｅｌｅｃｔ関数で複数のファイル記述子を同時に見張ることができる。ｓｅｌｅｃｔ関数は、見張っているファイル記述子の何れかがデータ受け取り可能な状態になるまでブロックする。

【０１１０】発明の実施の形態１で用いるＴＣＰ／ＩＰエミュレーションライブラリでは、並列計算機内部の通信時には従来のソケットライブラリを用いないため、ｓｅｌｅｃｔシステムコールでは、内部通信用のソケットからデータ受信可能であるかどうかを調べることができない。そこで、本実施の形態では、実施の形態１のごとくＴＣＰ／ＩＰエミュレーションライブラリを使用する計算機システムにおいても、アプリケーションプログラムが従来と同様にセレクト関数を利用可能にする。

【０１１１】本実施の形態では、内部ソケットテーブル４１０、４１１等に登録されているソケット記述子に対しては、内部通信専用のｓｅｌｅｃｔに相当する処理を行い、それ以外のファイル記述子に対しては、従来のｓｅｌｅｃｔシステムコールをそのまま用いる、という切り分けを行うセレクト関数ＥＭＵ＿ｓｅｌｅｃｔをＴＣＰ／ＩＰエミュレーションライブラリ１１３内に設ける。

【０１１２】ただし、内部通信用のｓｅｌｅｃｔとｓｅｌｅｃｔシステムコールは同時に並行して実行しなくてはならない。２つのｓｅｌｅｃｔを逐次に実行するのでは、例えば内部通信用のｓｅｌｅｃｔがデータを待っている間は、外部との通信用のソケットや標準入出力などがデータを受け取り可能な状態になった場合でも、それを検知することができないからである。

【０１１３】ｓｅｌｅｃｔの同時実行を疑似的に実現する手段として、内部用ｓｅｌｅｃｔとｓｅｌｅｃｔシステムコールをノンブロッキングで続けて発行することをスピンループで繰り返す、という方法が考えられる。し

かしこの方法を用いると、データが受け取り可能になるまで計算機を占有してしまい、同一計算機上で走っている他のプロセスに処理が渡らなくなってしまう。

【０１１４】これに対して本実施の形態では、１ループ毎に処理を他のプロセスに譲渡する命令を挿入するという方法を採る。この方法により、スピンループによる計算機の占有を避けることができる。

【０１１５】図９は、図３で示した接続状態において、アプリケーションプログラム１０８および１０９が送信命令１００２，１００３を発行して、アプリケーションプログラム１０７にデータを送信し、アプリケーションプログラム１０７側でそれらのデータの到着を、ｓｅｌｅｃｔ命令１００１によって見張っている様子を表す。ｓｅｌｅｃｔ命令１００１で指定しているビット列はソケットＳＡ０およびｓａ１に対応しているとする（１００４，１００５）。このうちｓａ１は内部ソケットテーブルに登録されているので、内部用ｓｅｌｅｃｔで見張る（１００８）。一方、ＳＡ０は内部ソケットテーブルに登録されていないので、ｓｅｌｅｃｔシステムコールで見張る（１００９）。ＴＣＰ／ＩＰエミュレーションライブラリで発行するｓｅｌｅｃｔシステムコールではｓａ１を見張る必要が無いので、アプリケーションプログラム１０７が発行したｓｅｌｅｃｔ命令１００１で指定されていたビット列に対し、ｓａ１に対応するビットを０にしたビット列を指定する。処理１００８と１００９はノンブロッキングに発行し、交互に繰り返す（１０１０，１０１１）。ただし、繰り返しの途中で処理を一旦、他のプロセスに譲渡する。

【０１１６】図９に示したｓｅｌｅｃｔ命令、内部用のｓｅｌｅｃｔ関数をアプリケーションプログラムが使用可能にするためには、ＴＣＰ／ＩＰエミュレーションライブラリ１１２，１１３等にはエミュレーションセレクト関数ＥＭＵ＿ｓｅｌｅｃｔが設けられ、アプリケーションプログラムは、これを呼び出して使用する。アプリケーションプログラムがＥＭＵ＿ｓｅｌｅｃｔ関数を呼び出す際には従来と同じく、それぞれ一つのソケットに対応するビットからなるビット列ａｐ＿ｂｉｔｓを指定する。ＴＣＰ／ＩＰエミュレーションライブラリは、この関数呼び出しに応答して図１０にともない処理を実行する。

【０１１７】まず、そのビット列ａｐ＿ｂｉｔｓに対し（処理１１０１）、内部ソケットテーブルに登録されているソケット記述子に対応したビットを０にするためのマスクをかける（処理１１０２）。ｉｎ＿ｍａｓｋは、内部ソケットテーブルに登録されている全てのソケット記述子に対応するビットが０、それ以外のビットが１であるようなビット列である。よって、ａｐ＿ｂｉｔｓにｉｎ＿ｍａｓｋをかけて作成したビット列ｅｘ＿ｂｉｔｓは、ａｐ＿ｂｉｔｓで指定されたファイル記述子のうち、内部通信用のソケット記述子を除いたビット列となる。その後、ｅｘ＿ｂｉｔｓを引数にしたｓｅｌｅｃｔシステムコールと、内部用ｓｅｌｅｃｔ処理をノンブロッキングで一回ずつ実行し（処理１１０３，１１０４）、もし、この処理で調べたファイル記述子の何れかがデータ受け取り可能状態であった場合にはリターンする（処理１１０６）。そうでない場合は、一旦他のプロセスに処理を譲渡し（処理１１０７）、再びｓｅｌｅｃｔ処理を繰り返す（処理１１０８）。

【０１１８】こうして、本実施の形態によれば、実施の形態１のように内部高速通信機構を併用する計算機システムにおいても、アプリケーションプログラムがｓｅｌｅｃｔ関数を利用可能になる。

【０１１９】<発明の実施の形態３>上記スピンループによるｓｅｌｅｃｔ関数の実現方法では、スピンループの途中に他のプロセスに処理を譲渡する処理１１０７を挿入することで計算機の占有を回避するが、同じ計算機上で処理されるプロセスの優先度が低いと、そのプロセスには処理が渡らない可能性がある。データの到着をスリープして待つブロッキングウェイトを用いればこれを避けることができるが、ブロッキングウェイトの内部通信用ｓｅｌｅｃｔ処理とブロッキングウェイトのｓｅｌｅｃｔシステムコールを、１プロセス・１スレッド上で同時に実行することはできない。

【０１２０】これに対して本実施の形態では、まず２つのスレッドを生成し、一方のスレッド上では内部用ｓｅｌｅｃｔを、もう一方ではｓｅｌｅｃｔシステムコールを実行するという方法を採る。この方法によれば、内部用ｓｅｌｅｃｔとｓｅｌｅｃｔシステムコールがそれぞれのスレッド上で独立に動作することができるため、同時にブロックしてデータを見張ることができる。

【０１２１】図１１は、図９で示したのと同じ通信状態を表している。ただし、内部用ｓｅｌｅｃｔとｓｅｌｅｃｔシステムコールの同時実行の実現方法が異なる。図１１では、これら２つのｓｅｌｅｃｔ処理は、別々のスレッドの上で実行する（１２０３，１２０４）。処理１００９と同様に、スレッド１２０３上のｓｅｌｅｃｔシステムコールでは、ｓａ１に対応するビットを０にしたビット列を指定する。

【０１２２】図１２を参照するに、本実施の形態において、アプリケーションプログラムがＥＭＵ＿ｓｅｌｅｃｔ関数を発行する際に指定したビット列ａｐ＿ｂｉｔｓに対し、内部ソケットテーブルに登録されているソケット記述子に対応したビットを０にするためのマスクをかけてｅｘ＿ｂｉｔｓを作成する処理までは図１０の処理（１１０１，１１０２）と同じである（処理１３０１，１３０２）。その後、まずノンブロッキングでｓｅｌｅｃｔシステムコール（処理１３０３）と内部用ｓｅｌｅｃｔ処理（処理１３０４）を１回ずつ行う。この処理で調べたファイル記述子の何れかがデータ受け取り可能状態であった場合にはリターンする。そうでない場合に

は、スレッドを２つ生成し（処理１３０７）、それぞれの上でｓｅｌｅｃｔシステムコール（処理１３０８）と内部用ｓｅｌｅｃｔ（処理１３０９）を実行する。これらの処理は、それぞれデータ受け取り状態になるまでブロックする。両処理のうち先にブロックが解けた方は、もう一方のスレッドをキャンセルして（処理１３１０，１３１１）リターンする（処理１３１４，１３１５）。このキャンセル処理では、スレッドを強制的に終了させるのではなく、そのスレッドがもう不要であるという印を付ける。スレッドはブロックが解けた時にこの印が付けられているかどうかを調べ（処理１３１６，１３１７）、もし付けられていれば、キャンセルされていたことになるのでそのまま消滅する（処理１３１８，１３１９）。

【０１２３】上記ｓｅｌｅｃｔ処理において、スレッド分割の前に処理１３０３および処理１３０４のノンブロッキングｓｅｌｅｃｔ処理を一回ずつ行うのは、次の理由による。もし、処理１３０１が発行される以前に内部用ソケットとそれ以外のファイル記述子が共にデータ受け取り可能な状態になっている場合、ｓｅｌｅｃｔ関数はその両方を検知できなければならない。しかし、いきなりスレッドを分割して、内部ソケットとそれ以外のファイル記述子を別々に見張り始めると、両スレッドのうち若干早く検知した方がもう一方のスレッドをキャンセルしてしまうため、片方のスレッドの状態しか検知することができない。これに対して処理１３０３および処理１３０４を実行することで、処理１３０１が発行される以前の内部ソケットとそれ以外の記述子の状態を両方とも確実に調べることができる。

【０１２４】本方式では、データの到着をスリープして待つため、実施の形態２のスピンループで待つ方法に比べると、データ検出のタイミングが遅れるが、優先度の低いプロセスに対しても、処理の妨げとなることを回避できる。

【０１２５】

【発明の効果】本発明によれば、ストリーム通信をメッセージパッシング型の通信でもって実現できる。

【０１２６】本発明の他の態様によれば、第１の通信網とそれより高速の第２の通信網に接続された計算機上で動作するアプリケーションプログラムが、第１の通信網に接続された他の計算機上で動作する他のアプリケーションプログラムとの間で第１の通信規約に基づいて通信することができ、さらに、第２の通信網に接続された他の計算機上で動作する他のアプリケーションプログラムとの間でその第２の通信網を使用した高速の通信を行うことができる。

【０１２７】さらに具体的には、上記第１の通信規約はＴＣＰ／ＩＰ通信規約を使用できる。また、上記第２の通信網を使用した通信を、メッセージパッシング型の通信とすることができる。

【図面の簡単な説明】

【図１】本発明の実施例の全体構成図。

【図２】ソケット接続のフローチャート。

【図３】ソケット接続を説明すための図。

【図４】内部・外部通信切り分け方法のフローチャート。

【図５】内部・外部通信切り分け方法を説明するための図。

【図６】ストリーム通信のフローチャート。

【図７】送受信動作に使用される命令列の一例を示す図。

【図８】送受信動作に使用される命令列の他の例を示す図。

【図９】ｓｅｌｅｃｔ機能をスピンループで実現する方法を説明するための図。

【図１０】ｓｅｌｅｃｔ機能をスピンループで実現する方法のフローチャート。

【図１１】ｓｅｌｅｃｔ機能をスレッドを用いて実現する方法を説明するための図。

【図１２】ｓｅｌｅｃｔ機能をスレッドを用いて実現する方法のフローチャート。

【図１３】従来のＴＣＰ／ＩＰ通信規約の階層図。

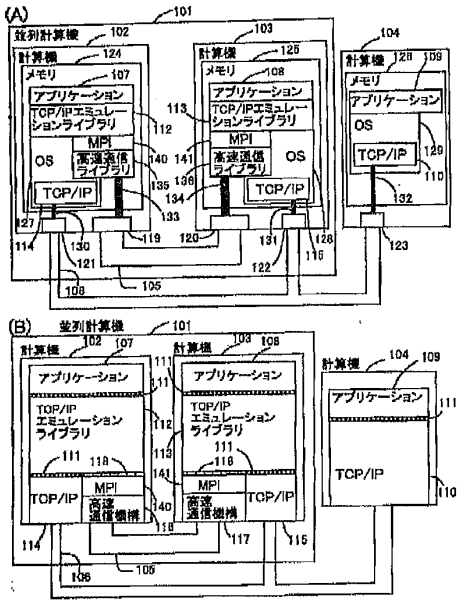
【図１４】従来のストリーム通信を説明するための図。

【符号の説明】

１０５．．．並列計算機内部通信網，１０６．．．グローバル通信網，１１１．．．ソケットアプリケーションプログラムインタフェース，１１８．．．ＭＰＩ仕様のインタフェース，１１９，１２０，１２１，１２２，１２３．．．ネットワークインタフェースハードウェア，１４０、１４１．．．メッセージパッシング型ライブラリ，４０１，４０２，４０３，４０４，４０５，４０６．．．ソケット，４０７，４０８，４０９．．．ソケットのコネクション，４１０，４１１．．．内部ソケットテーブル，６０１．．．並列計算機内部通信時におけるデータ経路，６０２．．．外部計算機との通信時におけるデータ経路，８０３，８０４，８１２，８１３．．．アプリケーションプログラムのバッファ，８０９．．．ストリーム，９０１．．．ＴＣＰ／ＩＰエミュレーションライブラリ内のバッファ。

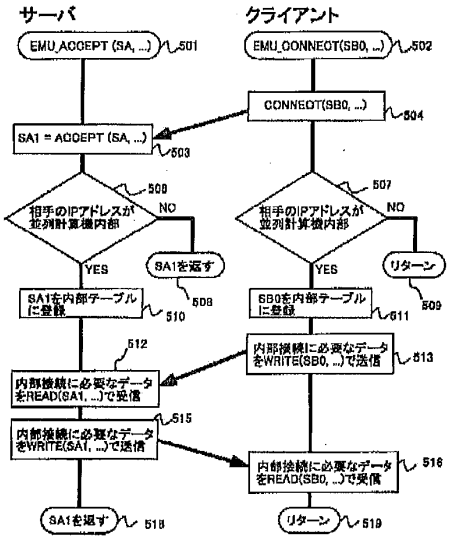
【図１】

図 1



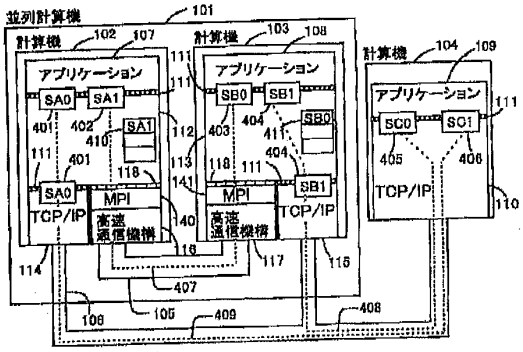
【図２】

図 2



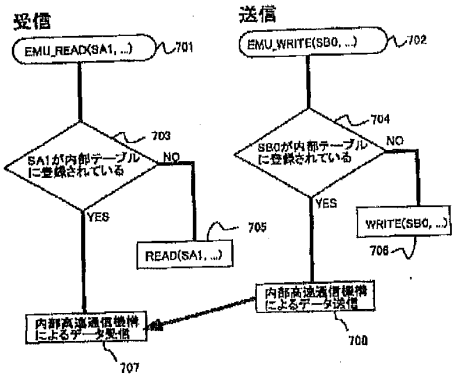
【図３】

図 3



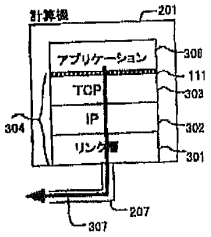
【図４】

図 4



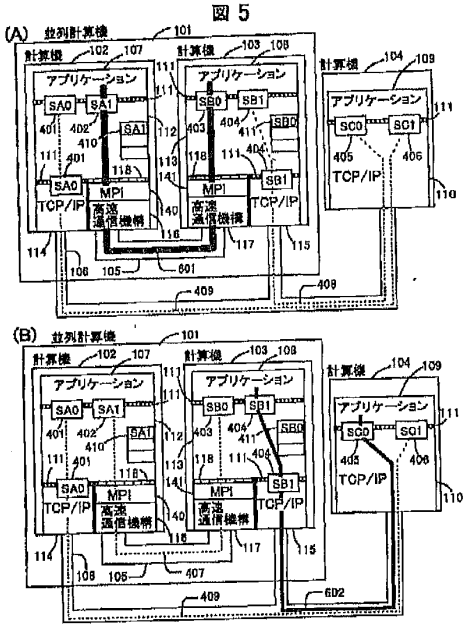
【図１３】

図 13



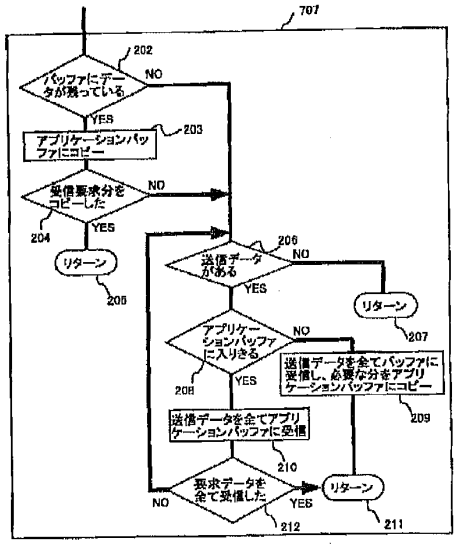
【図５】

図 5



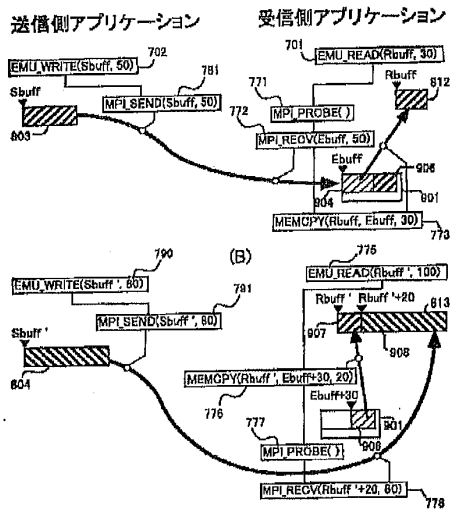
【図６】

図 6



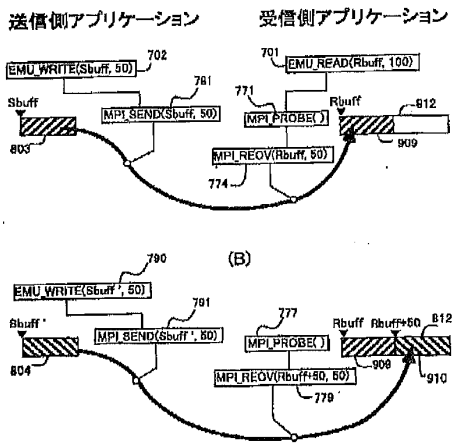
【図７】

図 7



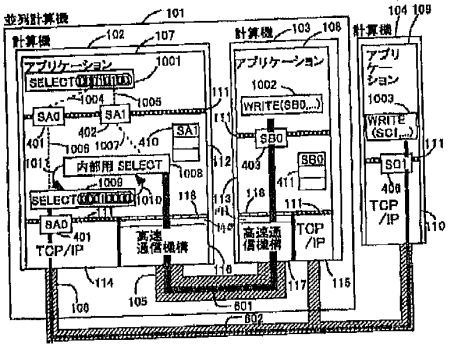
【図８】

図 8



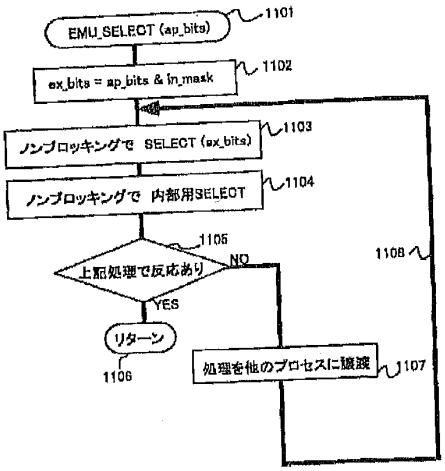
【図９】

図 9



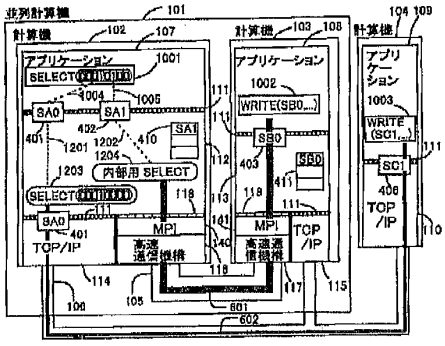
【図１０】

図 10



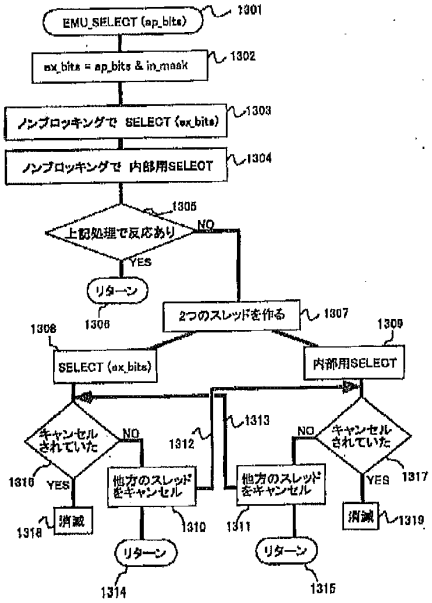
【図１１】

図 11



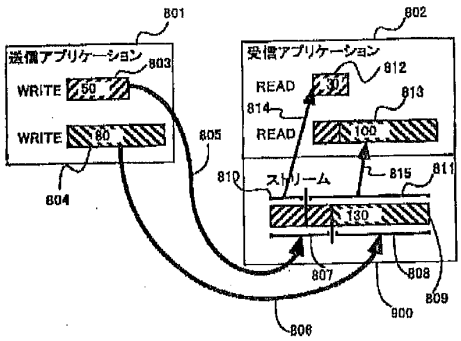
【図１２】

図 12



【図１４】

図 14


801
送信アプリケーション
803
WRITE
50
WRITE
80
804
805
806
802
受信アプリケーション
812
READ
30
813
814
READ
100
810
ストリーム
815
811
130
807
808
809
900

J-PlatPat
Japan Platform for Patent Information

Help desk (9:00-21:00)
(+81)3-6666-8801
helpdesk@j-platpat.inpit.go.jp

Japanese Top page Help list Site map JPO

National Cen
Industrial Pro
Information &

Patent & Utility Model Design Trademark Trial & Appea

Top page Patent & Utility Model Patent & Utility Model Number Search

Patent & Utility Model Number Search

Back Help Search List

You can retrieve a variety of patent and utility model gazettes by their numbers.

Selected Gazette

Previous Document 2/2 Next Doc

JP,2004-061359,A

PAJ Detail Image

(11)Publication number	2004-061359
(43)Date of publication of application	26.02.2004
(51)Int.Cl.	G01B 21/30
	G03G 5/00
	G03G 5/10
	G03G 21/00
(21)Application number	2002-221511
(22)Date of filing	30.07.2002
(71)Applicant	RICOH CO LTD
(72)Inventor	YAMAZAKI JUNICHI

(54)METHOD AND SYSTEM FOR EVALUATING SURFACE ROUGHNESS OF COMPONENT FOR IMAGE FORMING DEVICE, AND METHOD AND SYSTEM FOR CUTTING

(57)Abstract

PROBLEM TO BE SOLVED: To provide a surface roughness evaluating method for sensitively and accurately comprehending a local change or variation of a surface to be measured, in surface roughness measurement of a component for an image forming device such as a base substance for an electrophotography photoreceptor.

SOLUTION: According to this surface roughness evaluating method of a component for an image forming device, a cross- section curve defined by JIS B0601 is found on the surface condition of the component to perform multiple resolution analysis on positional data rows in a surface roughness direction at equally spaced positions on the cross-section curve, and the state of the surface roughness is evaluated at least based on the result.

Previous Document 2/2 Next Doc

To return to the top of

Notice on Usage Privacy Policy Questionnaire

(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2004-61359
(P2004-61359A)
(43)公開日　平成16年2月26日(2004.2.26)

(51)Int.Cl.⁷	F I			テーマコード（参考）
G01B 21/30	G01B 21/30	102		2F069
G03G 5/00	G03G 5/00	101		2H068
G03G 5/10	G03G 5/10		B	2H134
G03G 21/00	G03G 21/00			

審査請求 未請求　請求項の数 28　ＯＬ　（全 31 頁）

(21)出願番号　特願2002-221511(P2002-221511)
(22)出願日　平成14年7月30日(2002.7.30)

（特許庁注：以下のものは登録商標）
フロッピー

(71)出願人　000006747
株式会社リコー
東京都大田区中馬込１丁目３番６号
(74)代理人　100074505
弁理士　池浦　敏明
(72)発明者　山崎　純一
東京都大田区中馬込１丁目３番６号　株式会社リコー内

Ｆターム(参考)　2F069 AA60 BB40 JJ08 NN00 NN02 NN07 NN08
2H068 AA54 AA59 EA07 FA11
2H134 QA01 QA02

(54)【発明の名称】画像形成装置用部品の表面粗さ評価方法及び評価システム並びに切削加工方法及び切削加工システム

(57)【要約】
【課題】電子写真感光体用基体等の画像形成装置用部品の表面粗さの測定において、測定対象面の局所的な変化や変異を感度良くかつ正確に把握可能にすること表面粗さ評価方法を提供する。
【解決手段】画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列の多重解像度解析を行い、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とする画像形成装置用部品の表面粗さ評価方法。
【選択図】　なし

【特許請求の範囲】
【請求項１】
画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列の多重解像度解析を行い、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とする画像形成装置用部品の表面粗さ評価方法。
【請求項２】
多重解像度解析の方法がウェーブレット変換であることを特徴とする請求項１に記載の画像形成装置用部品の表面粗さ評価方法。
【請求項３】
多重解像度解析の方法がバンドパスフィルター処理又は短時間高速フーリエ変換であることを特徴とする請求項１に記載の画像形成装置用部品の表面粗さ評価方法。
【請求項４】
画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列のウィグナー分布を求め、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とする画像形成装置用部品の表面粗さ評価方法。
【請求項５】
画像形成装置用部品が電子写真感光体用基体であることを特徴とする請求項１～４のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。
【請求項６】
画像形成装置用部品が電子写真感光体用基体の上に塗工層を形成した電子写真感光体であることを特徴とする請求項１～４のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。
【請求項７】
画像形成装置用部品が電子写真装置用帯電ローラ、現像ローラ、定着ローラ、転写ベルト又は搬送ベルトであることを特徴とする請求項１～４のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。
【請求項８】
請求項１～７のいずれかに記載の表面粗さ評価方法と、ＪＩＳ　Ｂ０６０１に定める十点平均粗さ（Ｒｚ）、算術平均粗さ（Ｒａ）及び最大高さ（Ｒｍａｘ）のうちの少なくとも一つとで表面粗さの評価を行うことを特徴とする画像形成装置用部品の表面粗さ評価方法。
【請求項９】
請求項１～８のいずれかに記載の表面粗さ評価方法の評価結果を、あらかじめ定めておいた基準と比較することによる表面粗さの判定を行うことを特徴とする画像形成装置用部品の表面粗さ評価方法。
【請求項１０】
画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号をウェーブレット変換する機能のあるハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。
【請求項１１】
画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号をバンドパスフィルター処理又は短時間高速フーリエ変換処理する機能を有するハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。
【請求項１２】

画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号のウィグナー分布を求める機能を有するハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。

【請求項１３】

請求項１～４及び請求項９のうちのいずれかに記載の表面粗さ評価方法によって評価したことを特徴とする電子写真感光体用基体。

【請求項１４】

請求項１～４及び請求項９のうちのいずれかに記載の表面粗さ評価方法によって評価したことを特徴とする電子写真感光体。

【請求項１５】

請求項１４の電子写真感光体を搭載したことを特徴とする電子写真装置。

【請求項１６】

切削工具又は切削工具の取り付け治具に振動センサーを取り付け、切削加工時に、該振動センサーにより切削工具の振動を計測し、この計測信号から切削加工面に対応する二次元配列データーを作成し、この二次元配列データーを評価するか、又はこの二次元配列データーの信号解析を行って評価することにより、切削加工状態を把握しながら切削加工を行うことを特徴とする画像形成装置用部品の切削加工方法。

【請求項１７】

作成した二次元配列データーの解析方法が二次元フーリエ変換であることを特徴とする請求項１６に記載の画像形成装置用部品の切削加工方法。

【請求項１８】

作成した二次元配列データーの解析方法が二次元多重解像度解析であることを特徴とする請求項１６に記載の画像形成装置用部品の切削加工方法。

【請求項１９】

二次元多重解像度解析の方法が離散二次元ウェーブレット変換であることを特徴とする請求項１８に記載の画像形成装置用部品の切削加工方法。

【請求項２０】

振動を測定して作成した二次元データーを離散二次元ウェーブレット変換し、さらにこの離散二次元ウェーブレット変換した結果の一部を除くか、あるいは演算処理し、次いで二次元ウェーブレット逆変換を行い、その結果を解析評価することを特徴とする請求項１９に記載の画像形成装置用部品の切削加工方法。

【請求項２１】

二次元多重解像度解析の方法がバンドパスフィルターを使用した方法であることを特徴とする請求項１８に記載の画像形成装置用部品の切削加工方法。

【請求項２２】

二次元多重解像度解析の方法が短時間高速フーリエ変換であることを特徴とする請求項１８に記載の切削加工方法。

【請求項２３】

被加工物がアルミニウム又はアルミニウム合金であることを特徴とする請求項１６～２２のいずれかに記載の画像形成装置用部品の切削加工方法。

【請求項２４】

被加工物が電子写真感光体用基体であることを特徴とする請求項１６～２３のいずれかに記載の画像形成装置用部品の切削加工方法。

【請求項２５】

旋盤と、該旋盤で使用する切削工具の振動を検出する振動センサーと、該振動センサーからの信号を記録する記憶装置又は該振動センサーからの信号から作成した二次元データーを記録する記憶装置と、さらに離散二次元ウェーブレット変換を行うハードウェア又はソフトウェアを備えたことを特徴とする画像形成装置用部品の切削加工装置。

【請求項２６】
請求項１６～２４のいずれかに記載の切削加工方法で作成した電子写真感光体用基体。
【請求項２７】
請求項２６に記載の電子写真感光体用基体を用いた電子写真感光体。
【請求項２８】
請求項２７に記載の電子写真感光体を搭載した電子写真装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】
本発明は、画像形成装置用部品の表面粗さ評価方法及び評価システムに関し、特に電子写真感光体用基体の表面粗さの評価に適用して好適な画像形成装置用部品の表面粗さ評価方法及び評価システムに関するものである。
また、本発明は、画像形成装置用部品の切削加工方法及び切削加システムに関し、詳しくは、特に電子写真感光体基体の表面加工に適用して好適な切削加工方法及び切削加工システムに関するものである。
【０００２】
【従来の技術】
電子写真装置において、書き込み光源にレーザー等の可干渉光源を使用した場合、画像に干渉縞が発生する問題がある。この原因としては、可干渉光が感光層へ入射した際、入射光とその反射光の位相の干渉状態によって像形成に寄与する光量が変化することが挙げられる。この対策としては、基体表面を粗面化する方法が広く行われている。例えば、特開平５－２２４４３７（電子写真感光体用基体及びその表面処理方法）では、水蒸気を凍結させて得られる微細凍結粒子を気体表面に噴射して、基体表面の汚染物質を除去すると共に、基体表面を粗面化しており、この表面処理法により処理された基体を用いて形成される電子写真感光体は、レーザー光による干渉縞の形成を防止することができるとしている。
【０００３】
このように、基体表面の粗さは、画像品質にとって重要であるが、従来はＪＩＳ　Ｂ０６０１等に定める表面粗さで測定し、判断することが多かった。広く使われている測定方法としては、算術平均粗さ（Ｒａ）、最大高さ（Ｒｍａｘ）、十点平均粗さ（Ｒｚ）等がある。
【０００４】
例えば、この表面粗さ測定法で基体表面を評価した例としては、特開平５－７２７８５（電子写真感光体用基体及びその製造方法）において、しごき加工又は冷間引抜き加工によって作られたアルミニウム円筒基体を、該基体の表面が基体の軸方向のスジを全円周にわたって有し、該表面における円周方向のＲｍａｘが０．２～３μｍである電子写真感光体用アルミニウム円筒基体及びその製造方法を提案している。
また、特開平８－７６３９５（電子写真装置における感光体用アルミニウム基盤）では、光源として発光ダイオードが用いられる電子写真装置における感光体用アルミニウム基盤（基体）であって、表面粗さＲｚが０．８μｍ以下に規定されてなる基盤を提案している。
また、特開平６－１３８６８５（電子写真感光体）では、導電性支持体（基体）上に、感光層及び表面保護層を少なくとも有する電子写真感光体において、前記導電性支持体の表面粗さＲｚを０．０１μｍ以上０．５μｍ以下とし、かつ表面保護層を表面粗さＲｚを０．２μｍ以上１．２μｍ以下とすることを提案している。
【０００５】
ところが、上記のいずれの表面粗さ測定法でも、画像干渉縞対策として必要な表面粗さを定義しきれない問題があった。
【０００６】
また、従来表面粗さ表現法として使用している表面粗さＲａ、Ｒｚ等の定義では、測定長

内に飛び外れた山と谷があった場合、計算上表面粗さが大きく出る問題がある。このような問題に対し、各種の方法で表面粗さを規定する方法が検討されている。次にそれを紹介する。

【０００７】

特開平７－１０４４９７では、表面形状を表面粗さ測定装置で測定して得られる断面曲線上で、平均線を中心とした仕切り幅を規定し、この仕切り幅を超える相隣る山と谷の一対からなるピークの単位長あたりの数により表面形状を評価し、このような方法で仕切り幅を２０μｍとし単位長を１ｃｍとしたときのピークの数を１００以下とした基体を提案している。しかし、このような方法では表面粗さの評価法として充分ではなく、上記の評価法では十分な値であっても、実際に感光体を作成すると、画像異常が発生する場合があった。

【０００８】

そこで、本発明者らは表面粗さをフーリエ変換して評価する方法を特開２００１－２６５０１４、特開２００１－２８９６３０で提案した。

しかし、フーリエ変換では断面曲線中に頻度多く出現する変化をその周波数分布として捉えることはできるが、頻度が少ない変化を調べるにはさらに有効なものが望まれていた。すなわち、断面曲線中に数個の鋭い山や谷があっても、フーリエ変換ではスペクトルに出現しにくい問題がある。また、フーリエ変換した結果からは、どこでその変化が生じたのかが判らない問題があった。電子写真感光体用基体にこれらの数個の鋭い山や谷があると、画像欠陥となる問題があるが、従来の表面粗さによる方法、そしてフーリエ変換による方法では検出できなかった。

【０００９】

このような問題があるため、従来は、表面粗さ測定時に、表面粗さ計の記録チャートを保存しておき、記録チャートに記録された切削波形から判定していたが、記録チャートの傾向を読み取らねばならず、熟練を要する問題があった。

【００１０】

以上述べてきたように、従来の表面粗さ（Ｒａ、Ｒｍａｘ、Ｒｚ）評価法では測定対象面の局所的な変化や変異を感度良くかつ正確に把握できない問題がある。

また、先に述べたようにフーリエ変換では信号中に頻度多く出現する変化をその周波数分布として捉えることはできるが、頻度が少ない変化を調べるには有効ではない問題があった。また、フーリエ変換した結果からは、どこでその変化が生じたのかが判らない問題があった。

【００１１】

一方、電子写真複写機、デジタルコピア、レーザープリンター等の電子写真装置において広く使用されている、回転ドラム状の電子写真感光体用基体（以下適宜「基体」と略称する）上に感光層を設けてなる電子写真感光体を構成する基体の材料としては、低コスト、軽量、加工容易性等の利点から、アルミニウム系材料が好ましく用いられている。このアルミニウム系材料からなる回転ドラム状の基体は、一般に、管状素材の表面を切削加工して仕上げられる。

【００１２】

ここで、円筒体表面の切削加工には、基体を回転させ、バイト（刃物）を基体の軸方向に移動させて切削を行う方法と、基体を固定しておき、その周囲をバイト（刃物）が回転して切削を行う方法がある。

前者の方法は例えば、コロナ社刊精密工学講座１１　切削工学において普通旋盤として紹介されており、電子写真感光体用基体の加工方法としては、特許公報第３２１５８２９号、特許公報第２７９５３５７号、特開平７－７７８１４、特開平８－２７６３０１等で使用されている。

また、後者の方法は特開平６－３２８３０１や特開平６－３２８３０等で紹介されている。

【００１３】

一般に、旋盤は被加工物に回転を与えるための主軸台と、これに相対して被切削材の他端を支えるための心押台があり、さらに、バイトを取り付けて送りを与えるための往復台（刃物台）が載っている。こうした旋盤でバイトを使用し、被加工物を切削するときのバイト各部の角度、切削速度、送りなどの作業条件は、切りくず生成機構、切削抵抗、切削温度、バイト寿命、切削仕上面粗さ、ビビリ振動などに影響を及ぼす。

【００１４】

切削加工で画像形成装置に使用する部品、特に電子写真感光体用基体を製造する際には、切削ビビリや切削スジの無い、良好な切削面が要求される。

ここで、切削中に発生するビビリ振動には、強制ビビリ振動と自励ビビリ振動とがあることが知られている。強制ビビリ振動の原因としては、回転体の不平衡、旋盤の構造による振動などが考えられる。自励ビビリ振動の原因としては、切削抵抗の変動とそれによる旋盤、バイト、被切削材の振動特性によるものと考えられている。

【００１５】

以上のように、切削加工においては、ビビリ、スジ等の問題が発生するので、特開平８－３３６７０６、特開平９－２９５０３、特開平９－８０９１４、特開平９－１９２９５９、特開平９－２３４６３９、特開平１０－５８２１２、特開平１０－７１５０１、特開平１１－１８８５６６等多くの方法が検討されている。しかし、いずれも完璧なものとは言えなかった。

【００１６】

このような切削加工でのトラブルで発生した物の評価法としては、特開平１０－２６７７４９「切削加工における異常診断方法」では、切削加工における異常の原因を切削工具の振動から正確に特定できる異常診断方法として、切削工具の回転方向と半径方向の加速度振動から、有次元及び無次元特徴パラメータＸ、実効値Ｘｒｍｓ、ピーク値Ｘｐ、波高率Ｃ、歪み度β１、尖り度β２、交差頻度Ｎｏ、極値頻度Ｎｍ、定常度α、動揺度ε、一次及び二次平均周波数ｆ１、ｆ２とを、切削加工の推奨条件と実際の加工状態で求め、これらの値を１０回程サンプリングしてその平均値と標準偏差とを求め、さらにそれから両分布の差異を表す識別指数ＤＩを計算し、同識別指数の２以上の異常値の特徴パラメータの組み合せで原因を特定する。しかし、この評価法は処理が複雑であり、これだけでは異常診断できない問題がある。

【００１７】

電子写真感光体の基体に要求される品質は極めて高く、以上のようなビビリ対策やスジ対策を行っても、尚ビビリやスジが発生することがあった。そこで、切削後に切削加工面の検査を行う方法がいくつか提案されている。

【００１８】

特開平８－１０５７３２「研削面検査装置」では、円筒部材の外周面等を研削した際に研削面に生じるビビリの有無を簡単且つ確実に判断する装置として、被研削物をワーク受台で支持し、上方に配置した投光部から研削面に向けて光を投射するとともに、反射光を投影面に投影させ、研削面にビビリが生じていれば投影面にストライプ状の影が映し出されるので、その影の有無にてビビリの有無を判断している。しかし、この方法では切削加工後の検査は行えるものの、切削中に切削の状態を把握することはできない。

【００１９】

以上述べたように、従来の切削加工では突発的なビビリやスジの発生を防止できず、加工中の検出も困難な問題があった。

【００２０】

【発明が解決しようとする課題】

そこで、本発明は、電子写真感光体用基体等の画像形成装置用部品の表面粗さの測定において、測定対象面の局所的な変化や変異を感度良くかつ正確に把握可能にすること表面粗さ評価方法及び評価システムを提供することをその課題とする。

また、本発明は、電子写真感光体用基体等の画像形成装置用部品、特に、アルミニウムあるいはアルミニウム合金製管を切削加工して電子写真感光体用基体を作成する際に、切削

加工の状態を精度良く把握し、ビビリやスジ等の切削異常の発生を精度良く検知することを可能にし、良好な品質の切削面を作ることができる切削加工方法及び切削加工システムを提供することを別の課題とする。

【００２１】

【課題を解決するための手段】

本願第１の発明によれば、上記課題は下記の技術的手段により解決される。

（１）画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列の多重解像度解析を行い、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とする画像形成装置用部品の表面粗さ評価方法。

本方法によれば、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２２】

（２）多重解像度解析の方法がウェーブレット変換であることを特徴とする前記（１）に記載の画像形成装置用部品の表面粗さ評価方法。

本方法では、多重解像度解析の方法がウェーブレット変換であるので、高速かつ正確に多重解像度解析を行うことができ、これによって、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中ある局部的な変化や変異を検出できる。

【００２３】

（３）多重解像度解析の方法がバンドパスフィルター処理又は短時間高速フーリエ変換であることを特徴とする前記（１）に記載の画像形成装置用部品の表面粗さ評価方法。

本方法では、多重解像度解析の方法がバンドパスフィルターによるものであるので、高速かつ正確に多重解像度解析を行うことができ、これによって、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２４】

（４）画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列のウィグナー分布を求め、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とする画像形成装置用部品の表面粗さ評価方法。

本方法では、ウィグナー分布を求めてその結果に基づいて表面粗さの状態を評価するので、上記と同様、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２５】

（５）画像形成装置用部品が電子写真感光体用基体であることを特徴とする前記（１）～（４）のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。本方法では、電子写真感光体用基体に対して、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２６】

（６）画像形成装置用部品が電子写真感光体用基体の上に塗工層を形成した電子写真感光体であることを特徴とする前記（１）～（４）のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。

本方法では、電子写真感光体用基体の上に塗工層を形成した電子写真感光体に対して、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２７】

（７）画像形成装置用部品が電子写真装置用帯電ローラ、現像ローラ、定着ローラ、転写ベルト又は搬送ベルトであることを特徴とする前記（１）～（４）のいずれかに記載の画像形成装置用部品の表面粗さ評価方法。

本方法では、電子写真装置用帯電ローラ、現像ローラ、定着ローラ、転写ベルト又は搬送

ベルトに対して、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等による方法では検出できなかったような断面曲線中にある局部的な変化や変異を検出できる。

【００２８】

（８）前記（１）～（７）のいずれかに記載の表面粗さ評価方法と、ＪＩＳ　Ｂ０６０１に定める十点平均粗さ（Ｒｚ）、算術平均粗さ（Ｒａ）及び最大高さ（Ｒｍａｘ）のうちの少なくとも一つとで表面粗さの評価を行うことを特徴とする画像形成装置用部品の表面粗さ評価方法。

本方法では、前記と同様、表面粗さＲｚ、Ｒａ、Ｒｍａｘ等にのみよる方法では検出できなかったような断面曲線中に有る局部的な変化や変異を検出できる。

【００２９】

（９）前記（１）～（８）のいずれかに記載の表面粗さ評価方法の評価結果を、あらかじめ定めておいた基準と比較することによる表面粗さの判定を行うことを特徴とする画像形成装置用部品の表面粗さ評価方法。

本方法では、評価情報を別に設けた基準と比較して判定しているので、正確な評価が可能になる。

【００３０】

（１０）画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号をウェーブレット変換する機能のあるハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。

本システムでは、上記評価方法を高速かつ正確に実行することができる。

【００３１】

（１１）画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号をバンドパスフィルター処理又は短時間高速フーリエ変換処理する機能を有するハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。

本システムでは、上記システムと同様、上記評価方法を高速かつ正確に実行することができる。

【００３２】

（１２）画像形成装置用部品の表面粗さを測定するための表面粗さ測定用センサーと、該センサーからの信号を増幅する電気回路と、該電気回路から出た電気信号のウィグナー分布を求める機能を有するハードウェア又はソフトウェアと、表面粗さ測定用センサーと被測定対象を相対的に移動させる機構とからなることを特徴とする画像形成装置用部品の表面粗さ評価システム。

本システムでは、上記システムと同様、上記評価方法を高速かつ正確に実行することができる。

【００３３】

（１３）前記（１）～（４）及び（９）のうちのいずれかに記載の表面粗さ評価方法によって評価したことを特徴とする電子写真感光体用基体。

本電子写真感光体用基体は、前記表面粗さ評価方法によって評価しているので、その表面中の微細な変異や変化を検査でき、従って、電子写真感光体に用いたときに良好な画像を形成できる。

【００３４】

（１４）前記（１）～（４）及び（９）のうちのいずれかに記載の表面粗さ評価方法によって評価したことを特徴とする電子写真感光体。

本電子写真感光体は、前記（１３）の電子写真感光体用基体を用いているので、良好な画像を形成できる。

【００３５】

（１５）前記（１４）の電子写真感光体を搭載したことを特徴とする電子写真装置。
本電子写真装置は、前記（１４）の電子写真感光体を用いているので、良好な画像を形成できる。
【００３６】
（１６）切削工具又は切削工具の取り付け治具に振動センサーを取り付け、切削加工時に、該振動センサーにより切削工具の振動を計測し、この計測信号から切削加工面に対応する二次元配列データーを作成し、この二次元配列データーを評価するか、又はこの二次元配列データーの信号解析を行って評価することにより、切削加工状態を把握しながら切削加工を行うことを特徴とする画像形成装置用部品の切削加工方法。
本方法によれば、切削加工の一次元振動では存在が判らなかった変異や異常の検出が可能になる。従って、切削加工面の状態を精度良く把握しながら被加工物の切削加工をすることができるので、高品質の加工物が得られる。
【００３７】
（１７）作成した二次元配列データーの解析方法が二次元フーリエ変換であることを特徴とする前記（１６）に記載の画像形成装置用部品の切削加工方法。
本方法では、二次元配列データーの解析方法が二次元フーリエ変換であるので、精度の高い評価が可能になる。
【００３８】
（１８）作成した二次元配列データーの解析方法が二次元多重解像度解析であることを特徴とする前記（１６）に記載の画像形成装置用部品の切削加工方法。本方法では、作成した二次元配列データーの解析方法が二次元多重解像度解析であるので、評価対象二次元データーを周波数範囲ごとに分析することが可能になり、精度の高い評価が可能になる。
【００３９】
（１９）二次元多重解像度解析の方法が離散二次元ウェーブレット変換であることを特徴とする前記（１８）に記載の画像形成装置用部品の切削加工方法。
本方法では、二次元配列データーの解析方法が離散二次元ウェーブレット変換であるので、高速で精度の高い評価が可能になる。
【００４０】
（２０）振動を測定して作成した二次元データーを離散二次元ウェーブレット変換し、さらにこの離散二次元ウェーブレット変換した結果の一部を除くか、あるいは演算処理し、そして二次元ウェーブレット逆変換を行い、その結果を解析評価することを特徴とする前記（１９）に記載の画像形成装置用部品の切削加工方法。
本方法では、逆ウェーブレット変換によって余分や余剰な情報が除かれ、従って、高速で精度の高い評価が可能になる。
【００４１】
（２１）二次元多重解像度解析の方法がバンドパスフィルターを使用した方法であることを特徴とする前記（１８）に記載の画像形成装置用部品の切削加工方法。
本方法では、作成した二次元配列データーの二次元多重解像度解析の方法がバンドパスフィルターを使用した方法であるので、高速で精度の高い評価が可能になる。
【００４２】
（２２）二次元多重解像度解析の方法が短時間高速フーリエ変換であることを特徴とする前記（１８）に記載の画像形成装置用部品の切削加工方法。
本方法では、作成した二次元配列データーの二次元多重解像度解析の方法が短時間高速フーリエ変換であるので、高速で精度の高い評価が可能になる。
【００４３】
（２３）被加工物がアルミニウム又はアルミニウム合金であることを特徴とする前記（１６）～（２２）のいずれかに記載の切削加工方法。
本方法では、被切削材料がアルミニウムあるいはアルミニウム合金であっても、前記（１６）～（２２）のいずれかに示す切削加工方法を採用しているので、精度の高い切削加工評価が可能になる。

【００４４】
（２４）被加工物が電子写真感光体用基体であることを特徴とする前記（１６）～（２３）のいずれかに記載の切削加工方法。
本方法では、被加工物が高い品質を要求される電子写真感光体用基体であっても、前記（１６）～（２３）のいずれかに示す切削加工方法を使用しているので、精度の高い切削加工評価が可能になる。
【００４５】
（２５）旋盤と、該旋盤で使用する切削工具の振動を検出する振動センサーと、該振動センサーからの信号を記録する記憶装置又は該振動センサーからの信号から作成した二次元データーを記録する記憶装置と、さらに離散二次元ウェーブレット変換を行うハードウェア又はソフトウェアを備えたことを特徴とする切削加工装置。
本装置によれば、前記（１６）～（２４）に従う切削加工が可能になる。
【００４６】
（２６）前記（１６）～（２４）のいずれかに記載の切削加工方法で作成した電子写真感光体用基体。
本電子写真感光体用基体は、前記（１６）～（２４）のいずれかに記載の切削加工方法で作成しているので、切削不良や切削異常のない高品質な基体となる。
【００４７】
（２７）前記（２６）に記載の電子写真感光体用基体を用いた電子写真感光体。
本電子写真感光体は、前記（２６）の電子写真感光体用基体を用いているので、高品質な感光体となる。
（２８）前記（２７）に記載の電子写真感光体を搭載した電子写真装置。
本電子写真装置は、前記（２７）の電子写真感光体を搭載しているので、高品質な電子写真装置となる。
【００４８】
【発明の実施の形態】
以下本発明について詳細に説明する。
先ず、本願第１の発明について説明する。本願第１の発明の表面粗さ評価技術においては、画像形成装置用部品の表面の状態についてＪＩＳ　Ｂ０６０１に定める断面曲線を求め、その断面曲線上の等間隔位置における表面粗さ方向の位置データー列の多重解像度解析を行い、少なくともその結果に基づいて表面粗さの状態を評価することを特徴とするものである。
本願第１の発明においては、具体的には以下のような方法により画像形成装置用部品の評価を行う。始めに評価対象表面の状態を表面粗さ計によって測定し、ＪＩＳ　Ｂ０６０１に示す断面曲線を得る。この断面曲線は一次元のデーター配列、すなわち一次元信号と捉えることがでる。これは表面粗さ計からアナログあるいはデジタルの電気信号として得ることができる。
【００４９】
このようにして測定した信号から測定対象なる周波数成分のみを必要によりフィルタリングする。そして信号がアナログ信号である場合はＡ／Ｄ変換を行いデジタル信号に変換する。Ａ／Ｄ変換を行う場合は、少なくとも８ビット以上、好ましくは１０ビット以上の分解能で行う。Ａ／Ｄ変換速度はシステムの仕様により適時選定すれば良い。このようにしてデジタル化した数百点ないし数万点のデーターは一旦メモリに蓄えられる。
【００５０】
次に、このデーターを各種方法で多重解像度解析を行ったり、あるいはウィグナー分布を求め、評価を行う。多重解像度解析は周波数領域でスペクトルを分析しながら、同時に変動の時間的推移を解析する方法であり、例えば「ウェーブレット解析」芦野隆一、山本鎮男著、１９９７年６月、共立出版刊に説明されている。本発明において多重解像度解析は、各種の方法で行うことができ、請求項にあるように、ウェーブレット変換、バンドパスフィルターによる処理、短時間高速フーリエ変換による処理が使用可能である。

【００５１】
図１は本願第１の発明を適用した、画像形成装置用部品の表面粗さ評価装置の一構成例を模式的に示す構成図である。図中、１は電子写真感光体用基体あるいはその表面に下引き層を形成した物等の測定対象であり、２は表面粗さを測定するプローブを取り付けた治具、３は上記治具２を測定対象に沿って移動させる機構、４は表面粗さ計、５は信号解析を行うパーソナルコンピューターである。この図において、パーソナルコンピューター５によって上記の多重解像度解析やウィグナー分布計算が行われる。
【００５２】
この図は一例として示したものであり、構成は他の構成によってもかまわない。例えば、多重解像度解析やウィグナー分布計算はパーソナルコンピューターではなく、専用の数値計算プロセッサーで行っても良い。また、この処理を表面粗さ計自体で行っても良い。結果の表示は各種の方法が使用可能であり、ＣＲＴや液晶画面に表示しても良く、あるいは印字出力を行ったりしても良い。また、他の装置に電気信号として送信しても良く、フロッピーディスクやＭＯディスクに保存しても良い。
【００５３】
次に、本願第１の発明における信号処理方法を簡単に説明する。
始めに、短時間フーリエ変換（Ｓｈｏｒｔ　Ｔｉｍｅ　Ｆｏｕｒｉｅｒ　Ｔｒａｎｓｆｏｒｍ）は非定常信号の周波数成分の時間変化を捉えるために考え出された方法であり、短時間ごとに信号を切り出し、フーリエ変換したものである。フーリエ変換では精度を良くするためにはサンプル数を少なくする（取り出す時間長を短くする）必要があるが、この場合、周波数に対する精度（周波数分解能）は低下する。短時間フーリエ変換では、切り出し時間窓とフーリエ変換の長さを別個に設定して必要な周波数分解能を保ったまま時間分解能を向上させている。短時間フーリエ変換の結果の絶対二乗値の分布をスペクトログラム（Ｓｐｅｃｔｒｏｇｒａｍ）という。
ここで、短時間フーリエ変換は、その処理アルゴリズムを高速化した短時間高速フーリエ変換でも良い。
【００５４】
次にウェーブレット変換による方法を説明する。
関数ｆ（ｔ）のウェーブレット変換は（数２）で表される。
【数１】

$$W(b,a)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty}f(t)\overline{\psi\left(\frac{t-b}{a}\right)}\mathrm{d}t$$

ここで、
Ｗ（ｂ，ａ）　ウェーブレット変換
ψ（ｔ）　マザーウェーブレット
ａ　スケールパラメータ
ｂ　トランスレートパラメータ
である。
（数１）は連続関数のウェーブレット変換、すなわち連続ウェーブレット変換である。本発明では、サンプリングを一定間隔に行うので、離散系であり、離散ウェーブレット変換を行う。離散ウェーブレット変換において、ウェーブレット係数$c_{j,k}$は（数２）で表される。
【数２】

$$c_{i,k} = 2^{-j/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(2^{-j}t - k)}\mathrm{d}t$$

である。また、スケーリング係数$d_{j,k}$は（数３）で表される。

【数３】

$$d_{j,k} = 2^{-j/2} \int_{-\infty}^{\infty} f(t)\overline{\phi(2^{-j}t - k)}\mathrm{d}t$$

（数３）においてφ（ｔ）はスケーリング関数である。

また、（数２）、（数３）においてｊはレベルであり、元信号に対する解像度である階層の番号を示す。ウェーブレット係数$c_{j,k}$は信号の周波と時間分布を示す。また、スケーリング関数$d_{j,k}$は元信号のｊ次の解像度の離散化表現である。

【００５５】

離散ウェーブレット変換では、データーを（数４）によって計算する。

【数４】

$$c_{j+1,k} = \frac{1}{\sqrt{2}} \sum_{1} q_{1-2k} d_{j,1}$$

$$d_{j+1,k} = \frac{1}{\sqrt{2}} \sum_{1} p_{1-2k} d_{j,1}$$

（数４）において、係数群ｐ及びｑはウェーブレット変換のための変換基底であり、それぞれローパスフィルター、ハイパスフィルターの機能がある。従って、（ｊ＋１）次のスケーリング係数$d_{j+1,k}$は、ｊ次のスケーリング係数$d_{j,k}$より１つ下の解像度表現となり、解析可能な周波数及び時間的な解像度がｊ次の１／２になる。

【００５６】

一方、（ｊ＋１）次のウェーブレット係数$c_{j+1,k}$はｊ次のスケーリング係数$d_{j,k}$をハイパスフィルターに通すことにより得られ、スケーリング係数$d_{j+1,k}$と$d_{j,k}$の間の周波数成分を表すことになる。

【００５７】

図２はウェーブレット変換の処理フローを示した図であり、元信号１０（Ｓｏｕｒｃｅ Ｓｉｇｎａｌ）をハイパスフィルター（ＨＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って、信号１１に示されるＨ成分（Ｈ　ｐａｒｔ）を得る。

また、元信号１０（Ｓｏｕｒｃｅ　Ｓｉｇｎａｌ）をローパスフィルター（ＬＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って信号１２を得る。

このようにして得た信号１２をハイパスフィルター（ＨＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って、信号１３に示されるＬＨ成分（ＬＨ　ｐａｒｔ）を得る。

また、信号１２をローパスフィルター（ＬＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って信号１４を得る。

このようにして得た信号１４をハイパスフィルター（ＨＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って、信号１５に示されるＬＬＨ成分（ＬＬＨ　ｐａｒｔ）を得る。
また、信号１４をローパスフィルター（ＬＰＦ）に通し、さらにデーターをひとつ置きに間引く処理（ＳＳ）を行って、信号１６に示されるＬＬＬ成分（ＬＬＬ　ｐａｒｔ）を得る。
図３はこのようにして多重解像度解析処理を行った結果を示した図であり、元信号はＬＬＬ成分、ＬＬＨ成分、ＬＨ成分、Ｈ成分の４成分に分解される。

【００５８】
ここで、ウェーブレット変換は直交ウェーブレット変換と非直交ウェーブレット変換に分類することが可能であり、このいずれを用いても良い。

【００５９】
直交ウェーブレット変換では、ウェーブレット関数は実数形のみが用いられることが多い。このウェーブレット関数としては、ドビッシー（Ｄａｕｂｅｃｉｅｓ）関数、ハール（Ｈａｒｒ）関数、メーヤー（Ｍｅｙｅｒ）関数、シムレット（Ｓｙｍｌｅｔ）関数、そしてコイフレット（Ｃｏｉｆｌｅｔ）関数等が使用可能である。ここでＤａｕｂｅｃｉｅｓはドベシィと表記することがある。これらの直交ウェーブレット変換では、演算した絶対値にローパスフィルターなどにより包絡線（エンベロープ）処理を行えば強度に相当する情報が得られる。

【００６０】
非直交ウェーブレット関数には、複素数形ウェーブレットと実数形ウェーブレットを用いるものがある。複素数形ウェーブレット関数としてはガウス形ウェーブレット関数がある。この複素数形ウェーブレット関数を用いた場合、ウェーブレット変換結果に対して絶対値を演算することにより強度が得られる。実数形ウェーブレット関数としては、メキシカンハット関数、フレンチハット関数等があるが、これを使用して得たウェーブレット変換結果に対して絶対値を演算しても強度は得られない。しかし、演算した絶対値にローパスフィルター等で包絡線（エンベロープ）処理を行うことにより強度に相当する値を得ることが可能である。

【００６１】
ウェーブレット変換結果の二乗絶対値はスカログラム（Ｓｃａｌｏｇｒａｍ）と呼ばれ、スカログラムで示すこともできる。短時間フーリエ変換から求められるスペクトログラムは定周波数バンド分析であるが、ウェーブレット変換によるスペクトログラムは定対数バンド幅分析である。

【００６２】
次にバンドパスフィルターによる処理であるが、これは数種のバンドパスフィルターをあらかじめ用意しておき、信号を順次これらのフィルターに通過させて、信号の多重解像度解析を行う方法である。バンドパスフィルターによる方法では、信号はデジタル信号である必要は無く、アナログ信号であってもかまわない。アナログ信号のバンドパスフィルター処理では、処理の高速化が行える利点がある。

【００６３】
最後に、ウィグナー分布について説明する。時間信号ｆ（ｔ）のウィグナー分布は（数５）で表される。

【数５】

$$W(t,f)=\int_{-\infty}^{\infty} f\left(t+\frac{\tau}{2}\right)\cdot f^{*}\left(t-\frac{\tau}{2}\right)\cdot \exp(-j2\tau)d\tau$$

ここで、＊は複素共役を示す。ウィグナー（Ｗｉｇｎｅｒ）分布はウィグナーービレ（Ｗｉｇｎｅｒ－Ｖｉｌｌｅ）分布と呼ぶ場合もある。本発明では、信号は位置の関数である

が、時間の関数と読み替えてウィグナー分布の計算を行うことが可能である。

【００６４】

これらのバンドパスフィルター処理、ウェーブレット変換、短時間フーリエ変換、ウィグナー分布計算は各種の方法で行うことが可能である。例えばソフトウェアで行う場合、ｍａｔｈｅｍａｔｉｃａではウェーブレット変換を行うパッケージで可能であるＷａｖｅｌｅｔ　Ｅｘｐｌｏｒｅ等を併用して計算できる。また、ＭＡＴＬＡＢではウェーブレット変換を行うパッケージ（Ｗａｖｅｌｅｔ　Ｔｏｏｌ　Ｂｏｘ）等を併用して計算できる。また、Ｃ言語等でプログラミングしても計算が可能であり、専用の数値演算プロセッサーによっても計算が可能である。

【００６５】

本発明において、バンドパスフィルター処理結果、ウェーブレット変換結果、短時間フーリエ変換結果、ウィグナー分布計算結果は別に測定を行った正常なもの、あるいは、正常ではないものの結果と照合比較することによって、判定を行うことができる。

【００６６】

本発明者の測定では、表面粗さ計は東京精密社製サーフコム５７０Ａを使用し、パーソナルコンピューターはＩＢＭ社製パーソナルコンピューターを使用し、サーフコム５７０ＡとＩＢＭ製パーソナルコンピューターの間はＲＳ－２３２－Ｃケーブルで接続した。サーフコムからパーソナルコンピューターに送られた表面粗さデーターの処理とその多重解像度解析、ウィグナー分布の計算等は、Ｃ言語で作成したソフトウェアで行った。

【００６７】

以下、本発明による表面粗さ評価方法で評価を行った電子写真感光体用基体を用いた電子写真感光体及び該電子写真感光体を搭載した電子写真装置の構成例について説明する。なお、電子写真装置は、狭義の意味での、後述する電子写真プロセスカートリッジを包含する。

【００６８】

図４は、本発明による電子写真感光体の層構成を模式的に表わす断面図であり、導電性の電子写真感光体用基体（以下単に基体とも称する）２１上に、下引き層２２を介して、電荷発生材料を主成分とする電荷発生層２３と、電荷輸送材料を主成分とする電荷輸送層４とが積層された構成をとっている。更に、図５は電荷輸送層２４の上に保護層２５を設けた構成になっている。

【００６９】

導電性基体（支持体）２１としては、体積抵抗10^{10}Ωｃｍ以下の導電性を示すもの、例えば、アルミニウム、ニッケル、クロム、ニクロム、銅、金、銀、白金などの金属、酸化スズ、酸化インジウムなどの金属酸化物を、蒸着又はスパッタリングにより、フィルム状もしくは円筒状のプラスチック、紙に被覆したもの、あるいは、アルミニウム、アルミニウム合金、ニッケル、ステンレス等の板及びそれらを押し出し、引き抜き等の工法で素管化後、切削、超仕上げ、研磨等で表面処理した管等を使用することができる。また、特開昭５２－３６０１６に開示されたエンドレスニッケルベルト、エンドレスステンレスベルトも導電性基体２１として用いることができる。

【００７０】

この他、上記基体上に導電性粉体を適当な結着樹脂に分散して塗工したものも、本発明の導電性基体２１として用いることができる。この導電性粉体としては、カーボンブラック、アセチレンブラック、また、アルミニウム、ニッケル、鉄、ニクロム、銅、亜鉛、銀等の金属粉或いは導電性酸化スズ、ＩＴＯ等の金属酸化物粉体等が挙げられる。また、同時に用いられる結着樹脂には、ポリスチレン、スチレン－アクリロニトリル共重合体、スチレン－ブタジエン共重合体、スチレン－無水マレイン酸共重合体、ポリエステル、ポリ塩化ビニル、塩化ビニル－酢酸ビニル共重合体、ポリ酢酸ビニル、ポリ塩化ビニリデン、ポリアリレート樹脂、フェノキシ樹脂、ポリカーボネート、酢酸セルロース樹脂、エチルセルロース樹脂、ポリビニルブチラール、ポリビニルホルマール、ポリビニルトルエン、ポリ－Ｎ－ビニルカルバゾール、アクリル樹脂、シリコーン樹脂、エポキシ樹脂、メラミン

樹脂、ウレタン樹脂、フェノール樹脂、アルキッド樹脂等の熱可塑性、熱硬化性樹脂又は光硬化性樹脂が挙げられる。このような導電性層は、これらの導電性粉体と結着樹脂を適当な溶剤、例えば例えばＴＨＦ（テトラヒドロフラン）、ＭＤＣ（ジクロロメタン）、ＭＥＫ（メチルエチルケトン）、トルエン等に分散して塗布することにより設けることができる。更に、適当な円筒基体上にポリ塩化ビニル、ポリプロピレン、ポリエステル、ポリスチレン、ポリ塩化ビニリデン、ポリエチレン、塩化ゴム、テフロン（商標名）などの素材に前記導電性粉体を含有させた熱収縮チューブによって導電性層を設けてなるものも、本発明の導電性基体２１として良好に用いることができる。

【００７１】

導電性基体２１の加工方法としては、各種の切削加工、研削加工、研磨加工が可能であり、それらの加工法の組み合わせも有効である。

【００７２】

次に感光層について説明する。感光層は単層でも積層でもよいが、説明の都合上、まず電荷発生層２３と電荷輸送層２４から構成される場合から述べる。

【００７３】

電荷発生層２３は、電荷発生材料を主成分とする層である。電荷発生材料には、顔料、染料などの有機材料が用いられ、その代表例として、モノアゾ顔料、ジスアゾ顔料、トリスアゾ顔料、ペリレン系顔料、ペリノン系顔料、キナクリドン系顔料、キノン系縮合多環化合物、スクアリック酸系染料、フタロシアニン系顔料、ナフタロシアニン系顔料、アズレニウム塩系染料等が挙げられ用いられる。電荷発生材料は、単独であるいは２種以上混合して用いられる。

【００７４】

電荷発生層２３に用いられる結着樹脂としては、ポリアミド、ポリウレタン、エポキシ樹脂、ポリケトン、ポリカーボネート、シリコーン樹脂、アクリル樹脂、ポリビニルブチラール、ポリビニルホルマール、ポリビニルケトン、ポリスチレン、ポリスルホン、ポリ－Ｎ－ビニルカルバゾール、ポリアクリルアミド、ポリビニルベンザール、ポリエステル、フェノキシ樹脂、塩化ビニル－酢酸ビニル共重合体、ポリ酢酸ビニル、ポリフェニレンオキシド、ポリアミド、ポリビニルピリジン、セルロース系樹脂、カゼイン、ポリビニルアルコール、ポリビニルピロリドン等が挙げられる。結着樹脂の量は、電荷発生物質１００重量部に対し２０～２００重量部、好ましくは５０～１５０重量部が適当である。

【００７５】

ここで用いられる溶剤としては、例えばイソプロパノール、アセトン、メチルエチルケトン、シクロヘキサノン、テトラヒドロフラン、ジオキサン、エチルセルソルブ、酢酸エチル、酢酸メチル、ジクロロメタン、ジクロロエタン、モノクロロベンゼン、シクロヘキサン、トルエン、キシレン、リグロイン等が挙げられる。塗布液の塗工法としては、浸漬塗工法、スプレーコート、ビートコート、ノズルコート、スピナーコート、リングコート等の方法を用いることができる。電荷発生層２３の膜厚は０．０１～５μｍ程度が適当であり、好ましくは０．１～２μｍである。

【００７６】

電荷輸送層２４は、電荷輸送物質及び結着樹脂を適当な溶剤に溶解ないし分散し、これを電荷発生層２３上に塗布、乾燥することにより形成できる。また、必要により可塑剤、レベリング剤、酸化防止剤等を添加することもできる。レベリング剤としては、ジメチルシリコーンオイル、メチルフェニルシリコーンオイルなどのシリコーンオイル類や、側鎖にパーフルオロアルキル基を有するポリマーあるいはオリゴマーが使用され、その使用量は結着樹脂に対して０～１重量％程度が適当である。

【００７７】

電荷輸送物質には、正孔輸送物質と電子輸送物質とがある。電子輸送物質としては、例えば、クロルアニル、ブロムアニル、テトラシアノエチレン、テトラシアノキノジメタン、２，４，７－トリニトロ－９－フルオレノン、２，４，５，７－テトラニトロ－９－フルオレノン、２，４，５，７－テトラニトロキサントン、２，４，８－トリニトロチオキサ

ントン、２，６，８－トリニトロ－４Ｈ－インデノ［１，２－ｂ］チオフェン－４－オン、１，３，７－トリニトロジベンゾチオフェン－５，５－ジオキサイド、ベンゾキノン誘導体等の電子受容性物質が挙げられる。

【００７８】

正孔輸送物質としては、ポリ－Ｎ－カルバゾール及びその誘導体、ポリ－γ－カルバゾリルエチルグルタメート及びその誘導体、ピレン－ホルムアルデヒド縮合物及びその誘導体、ポリビニルピレン、ポリビニルフェナントレン、ポリシラン、オキサゾール誘導体、オキサジアゾール誘導体、イミダゾール誘導体、モノアリールアミン誘導体、ジアリールアミン誘導体、トリアリールアミン誘導体、スチルベン誘導体、α－フェニルスチルベン誘導体、ベンジジン誘導体、ジアリールメタン誘導体、トリアリールメタン誘導体、９－スチリルアントラセン誘導体、ピラゾリン誘導体、ジビニルベンゼン誘導体、ヒドラゾン誘導体、インデン誘導体、ブタジエン誘導体、ピレン誘導体等、ビススチルベン誘導体、エナミン誘導体等その他公知の材料が挙げられる。これらの電荷輸送物質は、単独で又は２種以上混合して用いられる。

【００７９】

結着樹脂としては、ポリスチレン、スチレン－アクリロニトリル共重合体、スチレン－ブタジエン共重合体、スチレン－無水マレイン酸共重合体、ポリエステル、ポリ塩化ビニル、塩化ビニル－酢酸ビニル共重合体、ポリ酢酸ビニル、ポリ塩化ビニリデン、ポリアリレート、フェノキシ樹脂、ポリカーボネート、酢酸セルロース樹脂、エチルセルロース樹脂、ポリビニルブチラール、ポリビニルホルマール、ポリビニルトルエン、ポリ－Ｎ－ビニルカルバゾール、アクリル樹脂、シリコーン樹脂、エポキシ樹脂、メラミン樹脂、ウレタン樹脂、フェノール樹脂、アルキッド樹脂等の熱可塑性又は熱硬化性樹脂が挙げられる。

【００８０】

電荷輸送物質の量は、結着樹脂１００重量部に対し、２０～３００重量部、好ましくは４０～１５０重量部が適当である。また、電荷輸送層の膜厚は、５～５０μｍ程度とすることが好ましい。

【００８１】

ここで用いられる溶剤としては、テトラヒドロフラン、ジオキサン、トルエン、ジクロロメタン、モノクロロベンゼン、ジクロロエタン、シクロヘキサノン、メチルエチルケトン、アセトンなどが挙げられる。

【００８２】

また、電荷輸送層には電荷輸送物質としての機能と、バインダー樹脂の機能をもった高分子電荷輸送物質も良好に使用される。これら高分子電荷輸送物質から構成される電荷輸送層は、耐摩耗性に優れたものである。高分子電荷輸送物質としては、公知の材料が使用できるが、トリアリールアミン構造を主鎖及び／又は側鎖に含むポリカーボネートが良好に用いられる。例えば、特開２０００－１０３９８４の（１）～（１０）式で表わされる高分子電荷輸送物質が良好に用いられる。

【００８３】

また、本発明において、電荷輸送層２４に可塑剤やレベリング剤を添加してもよい。可塑剤としては、ジブチルフタレート、ジオクチルフタレートなど一般の樹脂の可塑剤として使用されているものがそのまま使用でき、その使用量としては結着樹脂に対して０～３０重量％程度が適当である。レベリング剤としては、ジメチルシリコーンオイル、メチルフェニルシリコーンオイルなどのシリコーンオイル類や、側鎖にパーフルオロアルキル基を有するポリマーあるいはオリゴマーが使用され、その使用量は結着樹脂に対して０～１重量％程度が適当である。

【００８４】

本発明における電子写真感光体には、図４、図５に示すように、導電性基体２１と感光層（２３、２４）との間に下引き層２２を設けることができる。下引き層２２は一般には樹脂を主成分とするが、これらの樹脂はその上に感光層を溶剤で塗布することを考えると、一般の有機溶剤に対して耐溶剤性の高い樹脂であることが望ましい。このような樹脂とし

ては、ポリビニルアルコール、カゼイン、ポリアクリル酸ナトリウム等の水溶性樹脂、共重合ナイロン、メトキシメチル化ナイロン等のアルコール可溶性樹脂、ポリウレタン、メラミン樹脂、フェノール樹脂、アルキッド－メラミン樹脂、エポキシ樹脂等、三次元網目構造を形成する硬化型樹脂などが挙げられる。

【００８５】

また、下引き層２２にはモアレ防止、残留電位の低減等のために、酸化チタン、シリカ、アルミナ、酸化ジルコニウム、酸化スズ、酸化インジウム等で例示できる金属酸化物の微粉末を加えてもよい。これらの下引き層２２は、前述の感光層の場合と同様、適当な溶媒、塗工法を用いて形成することができる。

【００８６】

さらに、本発明における下引き層２２として、シランカップリング剤、チタンカップリング剤、クロムカップリング剤等を使用することもできる。この他、下引き層２２にはAl_2O_3を陽極酸化にて設けたものや、ポリパラキシリレン（パリレン）等の有機物や、SiO_2、SnO_2、TiO_2、ＩＴＯ、CeO_2等の無機物を真空薄膜作製法にて設けたものも良好に使用できる。この他にも公知のものを用いることができる。下引き層２２の膜厚は０～５μｍが適当である。

【００８７】

図５に示すように、本発明の電子写真感光体には、感光層保護の目的で、保護層２５を感光層の上に設けることもある。保護層２５に使用する材料としては、ＡＢＳ樹脂、ＡＣＳ樹脂、オレフィン－ビニルモノマー共重合体、塩素化ポリエーテル、アリル樹脂、フェノール樹脂、ポリアセタール、ポリアミド、ポリアミドイミド、ポリアクリレート、ポリアリルスルホン、ポリブチレン、ポリブチレンテレフタレート、ポリカーボネート、ポリエーテルスルホン、ポリエチレン、ポリエチレンテレフタレート、ポリイミド、アクリル樹脂、ポリメチルペンテン、ポリプロピレン、ポリフェニレンオキシド、ポリスルホン、ポリスチレン、ＡＳ樹脂、ブタジエン－スチレン共重合体、ポリウレタン、ポリ塩化ビニル、ポリ塩化ビニリデン、エポキシ樹脂等の樹脂が挙げられる。保護層２５にはその他、耐摩耗性を向上する目的で、ポリテトラフルオロエチレンのような弗素樹脂、シリコーン樹脂及びこれら樹脂に酸化チタン、酸化スズ、チタン酸カリウム等の無機材料を分散したもの等を添加することができる。

【００８８】

保護層２５の形成方法としては、通常の塗布法が採用される。なお、保護層２５の厚さは、０．１～７μｍ程度が適当である。また、以上の他に真空薄膜作製法にて形成したａ－Ｃ、ａ－ＳｉＣなど公知の材料も保護層２５として用いることができる。

【００８９】

本発明においては、感光層（２３、２４）と保護層２５との間に中間層を設けることも可能である。中間層には、一般にバインダー樹脂を主成分として用いる。これら樹脂としては、ポリアミド、アルコール可溶性ナイロン、水酸化ポリビニルブチラール、ポリビニルブチラール、ポリビニルアルコールなどが挙げられる。中間層の形成法としては、前述のごとく通常の塗布法が採用される。なお、中間層の厚さは０．０５～２μｍ程度が適当である。

【００９０】

次に、上記電子写真感光体を搭載した本発明による電子写真装置について説明する。

図６は、該電子写真装置の一例を模式的に説明する概略図であり、下記するような変形例も本発明の範疇に属するものである。

図６に示す電子写真装置は、ドラム状の電子写真感光体３１のまわりに、帯電機構３２、露光光源３３、現像機構３４、転写機構３５、クリーニング機構３７が配置されている。転写機構３５において、転写材３８にはトナーが転写され、これは定着機構３６で定着される。

【００９１】

上記の電子写真装置を使用した電子写真方法においては、電子写真感光体３１は、反時計

方向に回転して、帯電機構３２で正又は負に帯電され、露光光源３３からの露光によって、静電潜像を電子写真感光体３１上に形成する。

【００９２】

帯電機構３２には、コロトロン、スコロトロン、固体帯電器（ソリッドステートチャージャ）、帯電ローラなどをはじめとする公知の帯電手段を用いることができる。転写機構３５には、一般の帯電器が使用できるが、転写チャージャと分離チャージャを併用したものが効果的である。

【００９３】

また露光光源３３、及び図示されていないが、除電光源等で使用する光源としては、蛍光灯、タングステンランプ、ハロゲンランプ、水銀灯、ナトリウム灯、発光ダイオード（ＬＥＤ）、半導体レーザー（ＬＤ）、エレクトロルミネッセンス（ＥＬ）等の発光物を使用することができる。そして、所望の波長域の光のみを照射するために、シャープカットフィルター、バンドパスフィルター、近赤外カットフィルター、ダイクロイックフィルター、干渉フィルター、色温度変換フィルターなどの各種フィルターを用いることもできる。かかる光源等は、図６に図示した構成の他に、光照射を併用した転写工程、除電工程、クリーニング工程、或いは前露光等の工程を設けることにより、感光体に光が照射される際にも用いることができる。

【００９４】

感光体に正又は負の帯電を施して画像露光を行った場合、感光体上には正又は負の静電潜像が形成される。これを負又は正に帯電した極性のトナー（検電微粒子）で現像すればポジ画像が得られるし、逆に正又は負に帯電した極性のトナーで現像すればネガ画像が得られる。かかる現像には、公知の方法を適用することができ、また除電手段にも公知の方法が用いられる。

【００９５】

この例においては、導電性基体はドラム状のものとして示されているが、シート状、エンドレスベルト状のものを使用することができる。クリーニング前チャージャとしては、コロトロン、スコロトロン、固体帯電器（ソリッドステートチャージャ）、帯電ローラなどをはじめとする公知の帯電手段を用いることができる。

また転写チャージャ及び分離チャージャには、通常上記の帯電手段を使用することができる。

クリーニング機構には、ファーブラシ、マグファーブラシなどをはじめとする公知のブラシやポリウレタン製ブレードを使用することができる。

【００９６】

以上に示すような電子写真装置に代表される本発明の画像形成手段は、複写装置、ファクシミリ、プリンタなどの装置内に固定して組み込まれていてもよいが、プロセスカートリッジの形でそれら装置内に組み込まれてもよい。プロセスカートリッジとは、感光体を内蔵し、他に帯電手段、露光手段、現像手段、転写手段、クリーニング手段、除電手段などを含んだ一つの装置（部品）である。プロセスカートリッジの形状等は多く挙げられるが、一般的な例として図７に示すものが挙げられる。

【００９７】

図７に本発明による電子写真用プロセスカートリッジを示す。

図７において、３１は電子写真感光体、３２は帯電手段、３３は画像露光光源、３４は現像手段、３５は転写手段、３８は紙等の転写材、３６は定着機構、３７はクリーニング機構、３９はプロセスカートリッジの容器を示す。

【００９８】

この図は一構造例を示したものであり、各手段は図に示した以外の形態でも良い。

例えば、帯電手段３２はコロトロン、スコロトロン、帯電ロール等の公知の帯電手段が使用可能である。

画像露光、及び図示されていない前露光光の光源には、蛍光燈、タングステンランプ、ハロゲンランプ、水銀灯、ナトリウム灯、発光ダイオード（ＬＥＤ）、半導体レーザー（Ｌ

Ｄ）、エレクトロルミネッセンス（ＥＬ）などの発光手段を使用することができる。
また、所定の波長域の光のみを照射するために、シャープカットフィルター、バンドパスフィルター、近赤外カットフィルター、ダイクロックフィルター、干渉フィルター、色温度変換フィルターなどの各種フィルターが使用可能である。クリーニング機構３７は、クリーニングブレードだけで行われることもあり、クリーニングブラシ、もしくはブレードと併用されることもある。
図７に示すプロセスカートリッジにおいて、クリーニング手段等がプロセスカートリッジに含まれなくても良い。また、図では内蔵していない発光手段や転写手段をプロセスカートリッジに内蔵していても良い。

【００９９】
次に本願第２の発明について説明する。
本願第２の発明は、画像形成装置用部品の切削加工において、切削工具又は切削工具の取り付け治具に振動センサーを取り付け、切削加工時に、該振動センサーにより切削工具の振動を計測し、この計測信号から切削加工面に対応する二次元配列データーを作成し、この二次元配列データーを評価するか、又はこの二次元配列データーの信号解析を行って評価することにより、切削加工状態を把握しながら切削加工を行うことを特徴とするものである。

【０１００】
本願第２の発明では、切削時に切削工具（例えばバイト）の振動データーを計測し、計測して得た一次元データーを被加工物の表面に対応する二次元データーとして配列する。この二次元データーには、切削加工時の被加工物の１回転に１回しか発生しないような発生頻度の少ない変異であっても、特定の箇所には変異の発生が集中するので、検出が容易となる。

【０１０１】
図８は本発明を実施するのに好適な旋盤とバイト振動データーの処理装置の構成図である。図８において、４１は被加工物、４２はバイト取り付け部である。バイト取り付け部４２には振動センサー（図示せず）が取り付けられており、その信号は４３のケーブルで４４のバイトの信号を処理する信号処理装置に伝えられる。そして、４５は被加工物４１を回転させるモーター、４６はバイトを載せた台を左右に動作させるモーターである。

【０１０２】
図９はバイトの振動信号データーから本発明の請求項１６、請求項１７に示す二次元データーを作成する過程の図であり、Ｉはバイトの振動信号のデーターである。Ｉにおいて、ｎの部分はｎ回転目、ｎ＋１の部分はｎ＋１回転目、ｎ＋２の部分はｎ＋２回転目、ｎ＋３の部分はｎ＋３回転目であり、このデーターが連続している。
図９のＩＩはＩのデーターを１回転ごとに並べた図であり、このように二次元データーが作成される。
図９のＩＩＩはＩＩを二次元配列にしたデーターであり、１行目はｎ回転目、２行目はｎ＋１回転目、３行目はｎ＋２回転目、４行目はｎ＋３回転目であり、以降このデーターが連続する。

【０１０３】
次に、本願第２の発明における信号処理方法を説明する。
始めに二次元フーリエ変換による方法から説明する。画像のｘ軸方向の空間周波数をω_x、軸方向の空間周波数をω_yとする。この場合、二次元フーリエ変換は（数６）で求まる。

【数６】

$$G(\omega_x,\omega_y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)e^{-i2\pi(\omega_x x+\omega_y y)}dxdy$$

ただし、実際の画像データーの場合は離散二次元フーリエ変換となる。

【０１０４】

画像データーマトリックスを二次元フーリエ変換する場合、マトリックスの行と列の点数は同じであることが好ましい。また、マトリックスの行と列の点数は２の階乗であることが好ましく、５１２以上が良い。データー点数が２の階乗でなくともフーリエ変換は可能であるが、フーリエ変換の計算速度が低下するので好ましくない。データー点数が２の階乗でない場合は、２の階乗になるように０の列あるいは行を入れて２の階乗になるようにすることが好ましい。

【０１０５】

次に本願第２の発明で用いるウェーブレット変換による方法を説明するが、一次元ウェーブレット変換については、第１の発明で説明したものと同じであるので重複説明を避けるため、その説明は省略する。

【０１０６】

一次元ウェーブレット変換は第１の発明で説明したと同様に行うが、二次元ウェーブレット変換は二次元データーに対して縦方向と横方向（列方向と行方向）に行うことで求めることができる。

【０１０７】

図１０は離散二次元ウェーブレット変換を行った場合のフローであり、元データーS_0は低周波成分L_1、縦方向高周波成分V_1、横方向高周波成分H_1、斜め方向高周波成分T_1に分けられる。このようにして得た低周波成分L_1を同様にしてウェーブレット変換することにより、低周波成分L_2、縦方向高周波成分V_2、横方向高周波成分H_2、斜め方向高周波成分T_2に分けられる。同様にして低周波成分をウェーブレット変換してゆき、低周波成分L_n、縦方向高周波成分V_n、横方向高周波成分H_n、斜め方向高周波成分T_nを得ることができる。

【０１０８】

ここで、離散二次元ウェーブレット変換を行う場合、変換する画像のサイズは作成した二次元画像マトリックスを一度に行う必要はなく、分割して処理しても良い。

【０１０９】

本願第２の発明において、被評価面を測定して、その二次元画像データーマトリックスを作成し、ウェーブレット変換するが、ウェーブレット変換は二次元画像データーマトリックスのデーターがすべて揃ってから行う必要は無く、二次元画像データーマトリックスの作成を行いつつ、マトリックスが用意できた部分についてのウェーブレット変換を行っても良い。

【０１１０】

同様に、フーリエ変換も、二次元画像データーマトリックスのデーターがすべて揃ってから行う必要は無く、二次元画像データーマトリックスの作成を行いつつ、マトリックスが用意できた部分についての二次元フーリエ変換を行っても良い。

【０１１１】

本願第２の発明において、二次元画像データーマトリックスの作成、ウェーブレット変換、フーリエ変換は各種の手段で実施可能であり、例えば、マイクロプロセッサー、ＤＳＰ（デジタルシグナルプロセッサ）、ＡＳＩＣ、あるいは専用ＩＣやＬＳＩが使用可能である。例えばソフトウェアで行う場合、ｍａｔｈｅｍａｔｉｃａではウェーブレット変換を行うパッケージで可能であるＷａｖｅｌｅｔ　Ｅｘｐｌｏｒｅを用いることによって計算できる。また、ＭＡＴＬＡＢではウェーブレット変換を行うパッケージ（Ｗａｖｅｌｅｔ　Ｔｏｏｌ　Ｂｏｘ）を用いて計算できる。また、Ｃ言語等でプログラミングしても計算が可能であり、専用の数値演算プロセッサーによっても計算が可能である。本発明者の検討ではＣ＋＋言語で作成したプログラムで計算した。

【０１１２】

本願第２の発明における画像形成装置用部品として典型的なものは電子写真感光体用基体であるが、このような基体としてはアルミニウム又はアルミニウム合金が好ましく使用さ

れる。具体的にはＡ１１００、Ａ３０３０、Ａ６０６３合金が特に好ましく使用される。このような基体は切削加工を施した後、陽極酸化や化成処理を施してもよい。

【０１１３】

本願第２の発明による電子写真感光体及び電子写真装置としては、好ましく適用される基体を除き、第１の発明で説明したものと同様であるので、その説明は省略する。

本願請求項１～９の発明において、演算処理の結果はコンピューターによる比較処理でも良く、あるいは演算処理結果を印字出力するかあるいは画面出力し、その出力を人間が見て判断しても良い。

【０１１４】

【実施例】

以下、本発明を実施例及び比較例により詳細に説明する。

【０１１５】

＜第１の発明＞

先ず、実施例及び比較例で測定対象とする電子写真感光体用基体の作成法と、その基体から電子写真感光体を作成する方法を説明し、次に作成した電子写真感光体の画像評価結果を説明する。そして、本発明による測定を実施例と比較例で説明し、この測定結果と画像評価結果を照合して本発明の効果を検証する。以下順に説明してゆく。

【０１１６】

（測定試料の用意）

本発明に従う測定を行う試料として、切削加工で３条件の加工を行い、電子写真感光体用基体３種を作成した。また、センタレス研削加工で２条件の加工を行い、電子写真感光体用基体２種を作成した。以下その条件を示す。

【０１１７】

測定試料Ａ：アルミニウム合金ＪＩＳ規格Ａ６０６３材をポートホール押出し法により外径φ３０．２ｍｍ、内径φ２８．５ｍｍのパイプ状に連続押出しし、それを長さ２５４ｍｍにカットして円筒シリンダーとした。この管を昌運社製旋盤でＲバイトを使用して外径３０．０ｍｍに切削加工した。このようにして作製した電子写真感光体用基体を測定試料Ａとする。この基体の表面粗さを測定したところ、１０本測定し、平均でＲｚは１．６２であった。

【０１１８】

測定試料Ｂ：アルミニウム合金ＪＩＳ規格Ａ６０６３材をポートホール押出し法により外径φ３０．２ｍｍ、内径φ２８．５ｍｍのパイプ状に連続押出しし、それを長さ２５４ｍｍにカットして円筒シリンダーとした。この管を昌運社製旋盤で測定試料Ａ切削時とは異なるＲフラットバイトを使用して外径３０．０ｍｍに切削加工した。このようにして作製した電子写真感光体用基体を測定試料Ｂとする。この基体の表面粗さを測定したところ、１０本測定し、平均でＲｚは１．３１であった。

【０１１９】

測定試料Ｃ：バイトが摩滅したＲバイトを使用した以外は測定試料Ａと同様な方法で切削加工した。このようにして作製した電子写真感光体用基体を測定試料Ｃとする。この基体の表面粗さを測定したところ、１０本測定し、平均でＲｚは１．５５であった。

【０１２０】

測定試料Ｄ：アルミニウム合金ＪＩＳ規格Ａ６０６３材をポートホール押出し法により外径φ３０．２ｍｍ、内径φ２８．５ｍｍのパイプ状に連続押出しし、それを長さ２５４ｍｍにカットして円筒シリンダーとした。この管をミクロン精機製センタレス研削盤で粗センタレス研削を行った後、仕上げンタレス研削を行って外径３０．０ｍｍに加工した。センタレス研削条件は以下に示す。このようにして作製した電子写真感光体用基体を測定試料Ｄとする。この基体の表面粗さを測定したところ、１０本測定し、平均でＲｚは２．７８であった。

【０１２１】

（粗センタレス加工）

研削輪の粒度　　８００メッシュ
研削輪の砥粒材質　炭化珪素質（ＳｉＣ）
研削輪の回転数　　１２５０ｒｐｍ
粗研削送り速度　　０．００７６ｍｍ／秒
（仕上げセンタレス加工）
研削輪の粒度　　１０００メッシュ
研削輪の砥粒材質　炭化珪素質（ＳｉＣ）
研削輪の回転数　　１２５０ｒｐｍ
仕上げ研削送り速度　０．００２２ｍｍ／秒

【０１２２】

測定試料Ｅ：アルミニウム合金ＪＩＳ規格Ａ６０６３材をポートホール押出し法により外径φ３０．２ｍｍ、内径φ２８．５ｍｍのパイプ状に連続押出しし、それを長さ２５４ｍｍにカットして円筒シリンダーとした。この管を測定試料Ｃと同様にしてセンタレス研削を行って外形３０．０ｍｍに加工した。ここで、仕上げ研削に使用した研削輪は目詰まりを起こしていた。このようにして作製した電子写真感光体用基体を測定試料Ｅとする。この基体の表面粗さを測定したところ、１０本測定し、平均でＲｚは１．７１であった。

【０１２３】

（電子写真感光体の作成）

次に測定試料Ａ～Ｅから電子写真感光体を以下の手順で作成した。

（１）洗浄

測定試料Ａ～Ｅの基体を、ジェット水流を用いた水洗浄装置にて洗浄し、表面に付着しているオイル分を除去した。その際、界面活性剤として、常盤化学（株）のケミコールＣ（商品名）及び超音波発振機を併用し、ジェット洗浄後に純水にて再洗浄して界面活性剤を完全に除去してから乾燥を行った。

（２）下引き層の形成

次いで、この基体面に下記の組成からなる樹脂塗料を浸漬法で塗布し、次いで１５０℃で１５分間加熱し、熱硬化させて、基体面に厚さ５μｍの下引き層を形成させた。

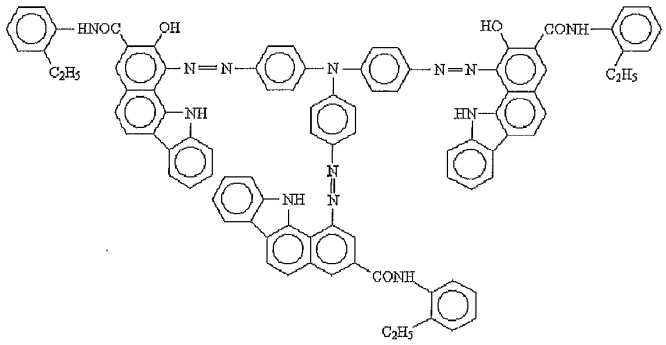
酸化チタン	２０	重量部
アルキッド樹脂	１０	重量部
メラミン樹脂	１０	重量部
メチルエチルケトン	６０	重量部

（３）電荷発生層の形成

次いで、この下引き層上に、電荷発生層を積層形成するために下記の組成からなる樹脂塗料（塗工樹脂液）を調製し、上記基体に同じく浸漬法でこの樹脂塗料を塗布し、１００℃で１０分間乾燥し、下引き層上に電荷発生層を積層形成させた。

ブチラール樹脂（ＵＣＣ社製ＸＹＨＬ）	１重量部
ジスアゾ顔料［下記（１）式］	９重量部
シクロヘキサノン	３０重量部
テトラヒドロフラン（ＴＨＦ）	３０重量部

【化１】



・・・(1)

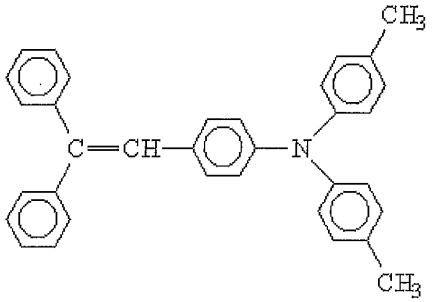
（４）電荷輸送層の形成

さらに、この電荷発生層上に電荷輸送層を積層形成するために下記の組成からなる樹脂塗料（塗工樹脂液）を調製し、上記基体に同じく浸漬法でこの樹脂塗料を塗布し、塗布後１２０℃で１５分間乾燥し、電荷発生層上に電荷輸送層を積層形成させた。

ポリカーボネート樹脂	１０重量部
電荷移動剤［下記（２）式］	１０重量部
ジクロロメタン	８０重量部

なお、ポリカーボネート樹脂は帝人社製のパンライトＫ－１３００を使用した。

【化２】



・・・(2)

【０１２４】

このようにして測定試料Ａの基体から電子写真感光体Ａを作成した。また、測定試料Ｂの基体から電子写真感光体Ｂを作成した。同様にして電子写真感光体Ｃ，Ｄ、Ｅを作成した。

【０１２５】

（画像評価）

作成した電子写真感光体Ａ～Ｅを図６に示す電子写真装置に取り付け、画像評価を行った。その結果を表１に示す。表１から判るように切削加工法で作成したＡ、Ｂ、ＣではＡが優れており、本発明の目的とする基体評価でＡとＢ、Ｃの判別ができれば良いことになる。また、センタレス加工法で作成したＤ、ＥではＤが優れており、同上に本発明の目的とする基体評価でＤとＥの区別が出来れば良いことになる。

【０１２６】

【表１】

電子写真感光体	画像評価結果
電子写真感光体Ａ	良好
電子写真感光体Ｂ	白ベタ画像に直径０．３ｍｍ程度の黒ポチ有り。
電子写真感光体Ｃ	一部にハーフトーンムラ有り。
電子写真感光体Ｄ	良好
電子写真感光体Ｅ	一部にハーフトーンムラ有り。

【０１２７】
画像評価で上のような結果が出たが、本発明の効果を検証するために測定試料Ａ～Ｅの基体の評価を行った。この実施例と比較例を以下に説明する。
【０１２８】
［実施例１］
測定試料Ａ～Ｅを図１に示した装置で表面粗さを測定し、得た断面曲線データーのウェーブレット変換を行った。
実施例１の測定例として、測定試料Ａの断面曲線を図１１に、そのウェーブレット変換を行った結果を図１２に示す。また、測定試料Ｄの断面曲線を図１３に、そのウェーブレット変換を行った結果を図１４に示す。
【０１２９】
［実施例２］
測定試料Ａ～Ｅを図１に示した装置で表面粗さを測定し、得た断面曲線データーをウェーブレット変換し、その結果のスカログラム（ｓｃａｒｏｇｒａｍ）を求めた。
実施例２の測定例として、測定試料Ａの元データーとそのスカログラムを図１５に示す。また、試料Ｄの元データーとそのスカログラムを図１６に示す。
【０１３０】
［実施例３］
測定試料Ａ～Ｅを図１に示した装置で表面粗さを測定し、得た断面曲線データーを短時間フーリエ変換した。
【０１３１】
［実施例４］
測定試料Ａ～Ｅを図１に示した装置で表面粗さを測定し、得た断面曲線データーのウィグナー分布を求めた。
実施例４の測定例として、測定試料Ａのウィグナー分布を図１７に、更にそれを三次元表示した図を図１８に示す。また、測定試料Ｄのウィグナー分布を図１９に、更にそれを三次元表示した図を図２０に示す。
【０１３２】
［比較例１］
測定試料Ａ～ＥのＪＩＳ　Ｂ０６０１に示す方法で行い、断面曲線を得た。
【０１３３】
［比較例２］
比較例１で求めた断面曲線のフーリエ変換を行った。
【０１３４】
実施例１～４、比較例１、２の評価結果を表２に示す。ここで、表２において、実施例１～４はコンピューターで測定試料Ａ、Ｂ、Ｃの測定結果を演算処理した結果を更にコンピューターで比較した結果である。
【０１３５】
【表２】

	評価方法	測定試料A、B、Cの差の検出	測定試料DとEの差の検出	総合評価
実施例1	ウェーブレット	A、B、Cの差を明確に検出できた。	DとEの差を明確に検出できた。	良好に検出
実施例2	スカログラム	A、B、Cの差を明確に検出できた。	DとEの差を明確に検出できた。	良好に検出
実施例3	短時間FFT	A、B、Cの差を明確に検出できた。	DとEの差を明確に検出できた。	良好
実施例4	ウィグナー分布	A、B、Cの差を検出できた。	DとEの差を検出できた。	良好に検出
比較例1	断面波形	A、B、Cの差を検出できたが、熟練が必要である。	DとEの差を検出できない。	良好できず
比較例2	フーリエ変換	A、B、Cの差を検出できない。	DとEの差を検出できない。	良好できず

【０１３６】
以上の結果から、本願第１の発明による方法によって、切削加工法で作成した測定試料Ａ、Ｂ、Ｃ間の差を検出できた。また、センタレス研削法で作成した測定試料Ｄ、Ｅの差を検出できたので、本発明の効果を検証できた。
＜第２の発明＞
【０１３７】
［実施例５］
Ａ３００３アルミニウム合金を押し出し加工で作成した外径３０．５ｍｍ、内径２８．５ｍｍ、長さ３４０ｍｍのアルミニウム合金を図８に示す旋盤に取り付けた。
旋盤に取り付けたアルミニウム合金管を５０００ｒｐｍで回転させつつ、バイトには５ｍｍの焼結ダイヤモンド製バイトを使用し、灯油を噴霧しながら、バイト送り速度５０ｍｍ／ｓｅｃで切削加工を行い、このときの振動を測定した。このとき、振動センサーのサンプリング周波数は８０ＫＨｚで行った。振動を検出した信号をＡ／Ｄ変換し、ＩＢＭ社製パーソナルコンピューターに取り込んで、切削表面に対応した二次元データーを作成した。このデーターは数値のマトリックスなので、視認しやすいように画像データーに変換した結果の例を図２１で示す。図２１で判るように、縞状パターンが確認できた。
ここで、バイトを先端が摩滅したバイトに交換し、切削ビビリが発生気味にして切削を行い、切削振動データーから同様にして画像を作成したところ、ビビリに対応した大きく、濃い縞が確認できた。また、切削加工時の巣と思われる信号の変異も確認できた。その様子を図２２に拡大図として示す。
【０１３８】
［実施例６］
実施例５と同様にして切削加工を行い、切削振動を取り込んで二次元データーを作成した。この二次元データーを二次元フーリエ変換した結果を図２３に示す。また、これを三次元表示した結果を図２４に示す。
バイトを研磨不良のバイトに交換し、切削スジが発生するようにして切削を行い、このときの振動を取り込んで二次元データーを作成し、この二次元データーを二次元フーリエ変換した。その結果にはスジによると思われる低い周波数成分が確認できた。
【０１３９】
［実施例７］
実施例５と同様にして切削加工を行い、切削振動を取り込んで二次元データーを作成した。次に、この二次元データーの二次元離散ウェーブレット変換を行った。このウェーブレット変換ではドビッシー関数を使用した。この結果を図２５に示す。

次に、バイトを先端が摩滅したバイトに交換し、切削ビビリが発生気味にして切削を行い、切削振動データーから同様にして二次元データーを作成し、更にこれを同様にウェーブレット変換したところ、ウェーブレット変換結果にはビビリに対応した縞が確認できた。このウェーブレット変換ではコフィレット関数を使用した。

【０１４０】

［実施例８］

実施例５と同様にして切削加工を行い、切削振動を取り込んで二次元データーを作成した。次に、この二次元データーの二次元離散ウェーブレット変換を行った。このウェーブレット変換ではドベシィ関数を使用した。

次に、バイトを研磨不良のバイトに交換し、切削スジが発生するようにして切削を行い、このときの振動を取り込んで二次元データーを作成し、更にこれを同様にウェーブレット変換したところ、ウェーブレット変換結果には切削スジに対応したと思われる縞が確認できた。

【０１４１】

［比較例３］

実施例５と同様にして切削加工を行い、切削振動を取り込んで一次元データーを作成した。この一次元データーをフーリエ変換した。

次に、バイトを研磨不良のバイトに交換し、切削スジが発生するようにして切削を行い、このときの振動を取り込んで一次元データーを作成し、この一次元データーをフーリエ変換した。この二つのフーリエ変換結果を比較したが、差は認められなかった。

【０１４２】

［比較例４］

実施例５と同様にして切削加工を行い、切削振動を取り込んで一次元データーを作成した。この一次元データーの一次元離散ウェーブレット変換を行った。次に、バイトを研磨不良のバイトに交換し、切削スジが発生するようにして切削を行い、このときの振動を取り込んで一次元データーを作成し、この一次元データーの一次元離散ウェーブレット変換を行った。この二つのウェーブレット変換結果を比較したが、差は認められなかった。

【０１４３】

【発明の効果】

本願第１の発明によれば、従来の表面粗さ評価方法すなわちＲｚ、Ｒａ、Ｒｍａｘでは把握できなかった測定対象面の局所的な変化や微細な変異も把握可能となり、画像異常が発生することが無く、均質で高画質な画像を形成できる画像形成装置を提供することができる。

本願第２の発明によれば、画像形成装置用部品、特に、アルミニウムあるいはアルミニウム合金製管の切削加工において、切削加工の状態を精度良く把握し、切削不良や切削異常の発生を検知しながら切削加工を行うことができるので、良好な品質の切削面を有する画像形成装置用部品、ひいては均質で高画質な画像を形成できる画像形成装置を提供することができる。

【図面の簡単な説明】

【図１】本発明を実施するのに好適な表面粗さ評価システムの構成図

【図２】ウェーブレット変換の処理フローの図

【図３】ウェーブレット変換の概念の図

【図４】電子写真感光体の層構成の図

【図５】電子写真感光体の別の層構成の図

【図６】電子写真装置の構成図

【図７】電子写真用装置用プロセスカートリッジの構成図

【図８】本発明を実施するのに適な切削加工システムの構成図

【図９】切削加工時に測定した振動データーから二次元データーを作成するフローの図

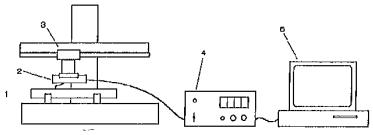
【図１０】二次元ウェーブレット変換の処理フロー図

【図１１】切削によって作成した基体の断面曲線の図

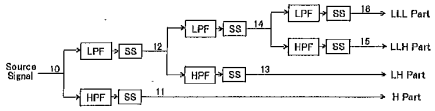
【図１２】図１１の断面曲線のデーターをウェーブレット変換によって多重解像度解析した結果の図
【図１３】図１１の断面曲線のデーターをウェーブレット変換して求めたスカログラムの図
【図１４】図１１の断面曲線のデーターのウィグナー分布を求めた図
【図１５】図１１の断面曲線のデーターのウィグナー分布の三次元表示の図
【図１６】切削によって作成した基体の断面曲線の図
【図１７】図１６の断面曲線のデーターをウェーブレット変換によって多重解像度解析した結果の図
【図１８】図１６の断面曲線のデーターをウェーブレット変換して求めたスカログラムの図
【図１９】図１６の断面曲線のデーターのウィグナー分布を求めた図
【図２０】図１６の断面曲線のデーターのウィグナー分布の三次元表示の図
【図２１】実施例５で作成した二次元データーを二次元濃度グラフで示した図
【図２２】実施例５で作成した二次元データーを二次元濃度グラフで示した図の部分拡大図
【図２３】図２１のデーターを二次元フーリエ変換した結果の図
【図２４】図２１のデーターを二次元フーリエ変換した結果を三次元表示した図
【図２５】図２１のデーターを離散二次元ウェーブレット変換した結果の図
【符号の説明】
１　測定対象
２　表面粗さを測定するプローブを取り付けた治具
３　上記治具を測定対象に沿って移動させる機構
４　表面粗さ計
５　信号解析を行うパーソナルコンピューター
１０　元信号
１１　ローパスフィルターを通過した信号
１２　ハイパスフィルターを通過した信号
１３　ローパスフィルターを通過した信号
１４　ハイパスフィルターを通過した信号
１５　ローパスフィルターを通過した信号
１６　ハイパスフィルターを通過した信号
２１　電子写真感光体用基体
２２　下引き層
２３　電荷発生層
２４　電荷輸送層
２５　保護層
３１　電子写真感光体
３２　帯電機構
３３　露光光源
３４　現像機構
３５　転写機構
３６　定着機構
３７　クリーニング機構
３８　転写材
３９　プロセスカートリッジのケース
４１　被加工物
４２　バイト取り付け台
４３　振動センサーからの信号を伝えるケーブル
４４　信号処理装置

４５　主軸駆動用モーター
４７　バイト取り付け台送り駆動用モーター

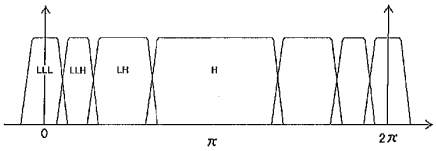
【図１】



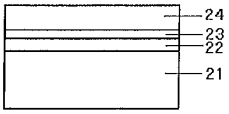
【図２】



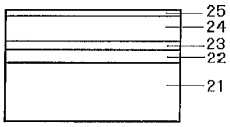
【図３】



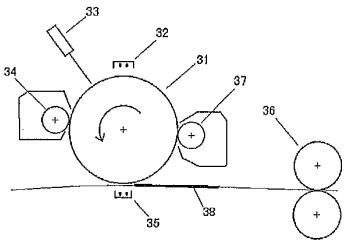
【図４】



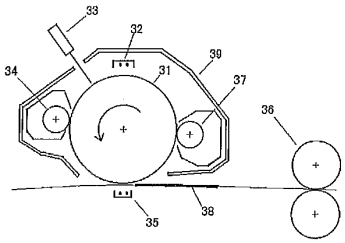
【図５】



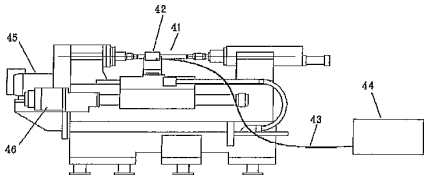
【図６】



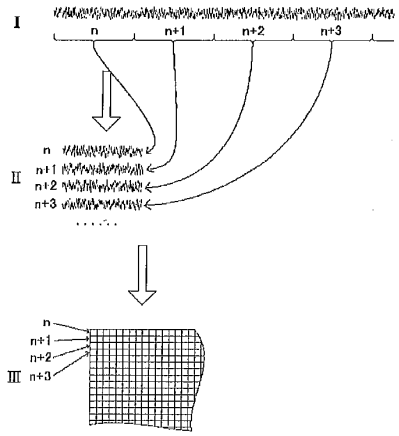
【図７】



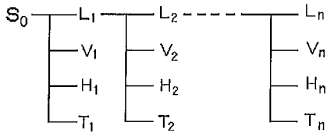
【図８】



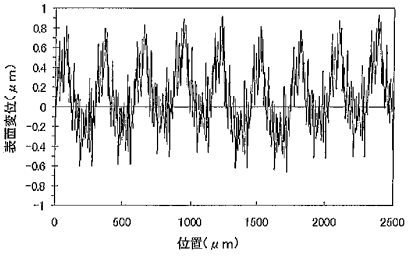
【図９】



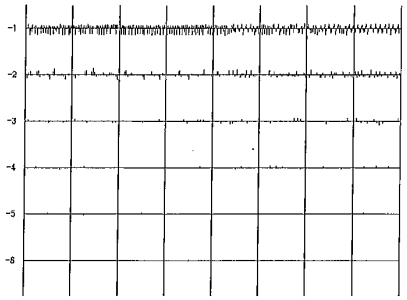
【図１０】



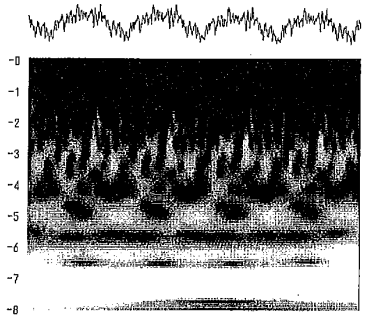
【図１１】



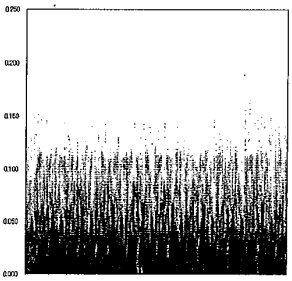
【図１２】



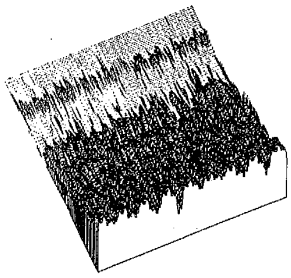
【図１３】



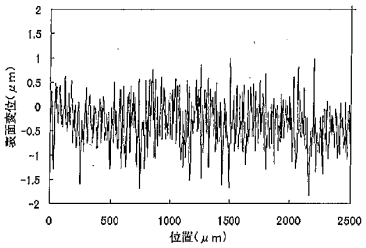
【図１４】



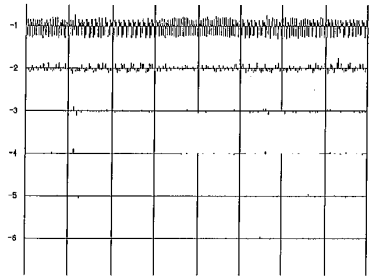
【図１５】



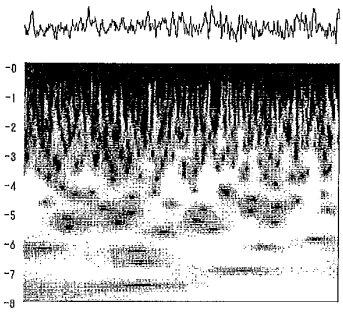
【図１６】



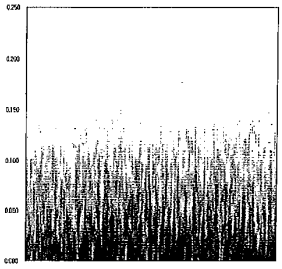
【図１７】



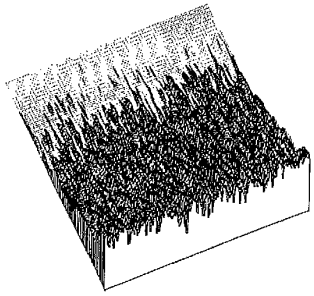
【図１８】



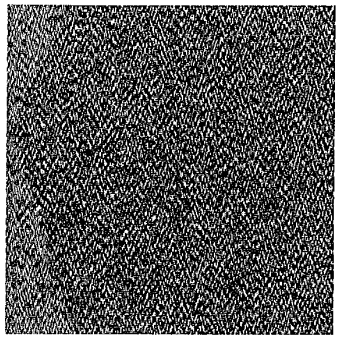
【図１９】



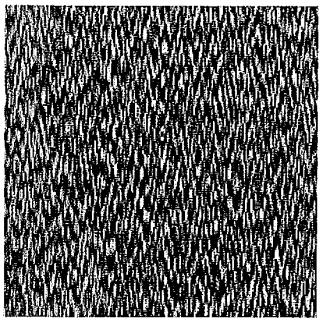
【図２０】



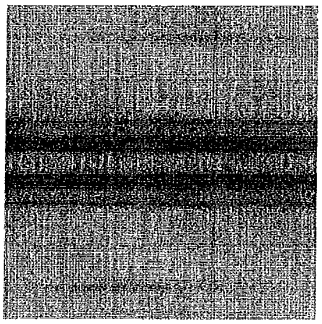
【図２１】



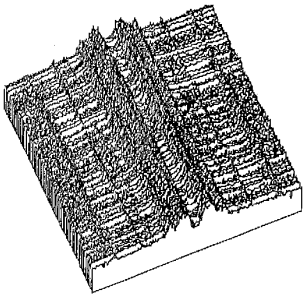
【図２２】



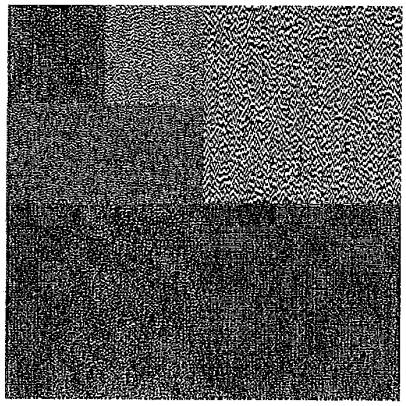
【図２３】



【図２４】



【図２５】



Electronic Acknowledgement Receipt

EFS ID:	30448191
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Jennifer Woo
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	22-SEP-2017
Filing Date:	14-FEB-2014
Time Stamp:	15:38:11
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		IDS_ZTANN_002P1C3.pdf	145458 c4f20c14b6ffcaa108a5d21a7a49a399642756d4	yes	3

	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Transmittal Letter		1	2	
	Information Disclosure Statement (IDS) Form (SB08)		3	3	
Warnings:					
Information:					
2	Non Patent Literature	KONISHI_Performance_Eval_of_Parallel_Comp_Aug_5_2005.pdf	2458551 e5fd3c8319c5521abdd51e4566e982b4b0c6d52d	no	9
Warnings:					
Information:					
3	Non Patent Literature	MATSUMURA_Construction_of_a_distributed_Aug_5_1998.pdf	2539138 503c6f92c4622807d02eca6c1f87d57846c302e3	no	9
Warnings:					
Information:					
4	Non Patent Literature	TATEBE_Efficient_Implementation_of_MPI_May_1999.pdf	5953948 9abc536e865709460e66ea38f8a51fbd5d3cc639	no	14
Warnings:					
Information:					
5	Foreign Reference	JP11328134_w_Abs.pdf	2115122 6058cd4e9b62c41985ea20a7156d71bbd0ac5919	no	24
Warnings:					
Information:					
6	Foreign Reference	JP2004061359_w_Abs.pdf	2118271 bd3389d9451afb86485c649dc1d50f0987120956	no	32
Warnings:					
Information:					
		Total Files Size (in bytes):	15330488		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE STATEMENT

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Asres, Hermon
Art Unit	:	2449
Conf. No.	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Pursuant to 37 CFR 1.56, an Information Disclosure Statement listing references is provided herewith. Copies of any listed foreign and non-patent literature references are being submitted. Any foreign references may also include English abstract(s) and/or machine translation(s), but no representation is made as to their accuracy.

If the Examiner would like additional information regarding these references or if anything is unclear, the Examiner is invited to contact the undersigned for assistance.

No Disclaimers

To the extent that anything in the Information Disclosure Statement or the listed references could be construed as a disclaimer of any subject matter supported by the present application, Applicant hereby rescinds and retracts such disclaimer.

Timing of Disclosure

This Information Disclosure Statement is being filed within 3 months after the filing of a Request for Continued Examination (RCE), and no fee is believed to be required.

The Commissioner is authorized to charge any additional fees which may be required or to credit any overpayment to Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: September 22, 2017

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Registered Practitioner
Customer No. 20995
(949) 760-0404

26739290

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3	6874

20995 7590 08/08/2017
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

EXAMINER
ASRES, HERMON

ART UNIT	PAPER NUMBER
2449	

NOTIFICATION DATE	DELIVERY MODE
08/08/2017	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jayna.cartee@knobbe.com
efiling@knobbe.com

PTOL-90A (Rev. 04/07)

Applicant-Initiated Interview Summary	Application No.	Applicant(s)	
	14/181,112	TANNENBAUM ET AL.	
	Examiner	Art Unit	
	HERMON ASRES	2449	

All participants (applicant, applicant's representative, PTO personnel):

(1) *HERMON ASRES*.

(2) _____.

(3) *Lance Smemoe (REG NO 66,152)*.

(4) _____.

Date of Interview: *25 July 2017*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☒ No.
If Yes, brief description: _____.

Issues Discussed ☐101 ☐112 ☐102 ☒103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *2*.

Identification of prior art discussed: *Block and Singh*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Applicant doscussed proposed amendments to Independent claim 2 including adding "program code for a single node kernel that when executed causes the hardware processor to interpret user instructions to evaluate mathematical expressions and to produce results of evaluation and a mechanism for the nosed to communcuate results of evaluation with each other using a peer-to-peer architecture" Examiner indicated to applicant that the Block and Singh references will be reviewed and a search will also be updated. No agreement was reached.

Applicant recordation instructions: The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☒ Attachment

/HERMON ASRES/ Examiner, Art Unit 2449	/VIVEK SRIVASTAVA/ Supervisory Patent Examiner, Art Unit 2449

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

M E M O R A N D U M

TO: Examiner Hermon Asres
direct: 571-272-4257
email: hermon.asres@uspto.gov

FROM: Lance Smemoe, Reg. #66,152

RE: Agenda for telephone interview on ~~November 7, 2016~~July 25, 2017 at 2:00 p.m.

Eastern time

Application No.: 14/181,112

Attorney Docket No.: ZTANN.002P1C3

INFORMAL DOCUMENT FOR DISCUSSION ONLY – DO NOT ENTER

I. Discuss Applicant's disclosure and claimed inventions

II. Discuss references cited in Office Action

A. Block et al. (US 2005/0021751)
B. Singh (US 8,601,101)

III. Discuss proposed claim amendments

2. (Currently amended) A computer cluster comprising:

a plurality of nodes comprising a hardware processor, wherein one or more of the nodes is configured to receive ~~receives~~ a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel that, when executed, causes the hardware processor to interpret user instructions, to evaluate mathematical expressions, and to produce results of evaluation; and

a mechanism for the nodes to communicate results of evaluation with each other using a peer-to-peer architecture;

wherein at least one of the nodes ~~returns a result~~ is configured to return at least one result of evaluation to a user interface or a script.

5. (Previously presented) The computer cluster of claim 2, wherein each of the nodes comprises one or more cluster node modules.

Knobbe Martens

INFORMAL DOCUMENT FOR DISCUSSION ONLY – DO NOT ENTER

32. (New) The computer cluster of claim 5, wherein each of the plurality of nodes implements asynchronous calls that enable the single-node kernel to perform computation tasks while the cluster node modules are simultaneously communicating with one another.

33. (New) The computer cluster of claim 5, wherein intercommunication among the plurality of single-node kernels during thread execution is enabled by the plurality of cluster node modules, and wherein the computer cluster is configured to permit exchange of information between nodes during the course of a parallel computation.

34. (New) The computer cluster of claim 5, wherein each of the single-node kernels are configured to call a send command involving creating a packet containing:

an expression to be sent as payload; and

where the expression should be sent;

wherein, upon reception of the send command by a local cluster node module associated with the single-node kernel, the local cluster node module is configured to decode the packet and forward a payload to the cluster node module specified in the packet;

wherein each of the single-node kernels is configured to call a receipt command involving creating a packet specifying which message to test for completion and to then wait for a reply expression to evaluate;

wherein, upon reception of the receipt command by the local cluster node module associated with the single-node kernel, the local cluster node module is configured to decode the packet and use a message specifier to search for any matching expressions listed as completed in a received message queue;

wherein, upon determining that such a completed expression is found, the local cluster node module is configured to send the completed expression to a local single-node kernel associated with the cluster node module in response to the receipt command, and

wherein each of the single-node kernels is configured to receive the completed expression and to update variables used by the single-node kernel during evaluation of expressions.

INFORMAL DOCUMENT FOR DISCUSSION ONLY – DO NOT ENTER

35. (New) The computer cluster of claim 2, wherein the plurality of nodes are configured to permit exchange of information between nodes during the course of parallel computation.

36. (New) The computer cluster of claim 2, wherein each of the plurality of nodes comprises instructions executable by the hardware processor and configured to implement asynchronous behavior, wherein the instructions comprise:

a first instruction to asynchronously send a payload to another node;

a second instruction to asynchronously receive a payload from another node; and

a third instruction to search for a payload matching a message specifier.

26363986

Formatted: Left

Formatted: Font: 8 pt

PTO/SB/30EFS (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

REQUEST FOR CONTINUED EXAMINATION(RCE)TRANSMITTAL
(Submitted Only via EFS-Web)

Application Number	14181112	Filing Date	2014-02-14	Docket Number (if applicable)	ZTANN.002P1C3	Art Unit	2449
First Named Inventor	Zvi Tannenbaum			Examiner Name	Hermon Asres		

This is a Request for Continued Examination (RCE) under 37 CFR 1.114 of the above-identified application.
Request for Continued Examination (RCE) practice under 37 CFR 1.114 does not apply to any utility or plant application filed prior to June 8, 1995, or to any design application. The Instruction Sheet for this form is located at WWW.USPTO.GOV

SUBMISSION REQUIRED UNDER 37 CFR 1.114

Note: If the RCE is proper, any previously filed unentered amendments and amendments enclosed with the RCE will be entered in the order in which they were filed unless applicant instructs otherwise. If applicant does not wish to have any previously filed unentered amendment(s) entered, applicant must request non-entry of such amendment(s).

☐ Previously submitted. If a final Office action is outstanding, any amendments filed after the final Office action may be considered as a submission even if this box is not checked.

- ☐ Consider the arguments in the Appeal Brief or Reply Brief previously filed on _____
- ☐ Other _____

☒ Enclosed

- ☒ Amendment/Reply
- ☒ Information Disclosure Statement (IDS)
- ☐ Affidavit(s)/ Declaration(s)
- ☐ Other _____

MISCELLANEOUS

☐ Suspension of action on the above-identified application is requested under 37 CFR 1.103(c) for a period of months _____
(Period of suspension shall not exceed 3 months; Fee under 37 CFR 1.17(i) required)

☐ Other _____

FEES

☒ **The RCE fee under 37 CFR 1.17(e) is required by 37 CFR 1.114 when the RCE is filed.**
The Director is hereby authorized to charge any underpayment of fees, or credit any overpayments, to Deposit Account No 111410

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

☒ Patent Practitioner Signature

Applicant Signature

PTO/SB/30EFS (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Signature of Registered U.S. Patent Practitioner			
Signature	/Lance D. Smemoe/	Date (YYYY-MM-DD)	2016-08-08
Name	Lance D. Smemoe	Registration Number	66152

This collection of information is required by 37 CFR 1.114. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

REMARKS

In the Office Action dated March 9, 2017, Claims 2-31 were rejected, as discussed below. In this Amendment, claims 2, 28, and 31 are amended, and new claims 32-41 are added. No new matter is added. After the amendments, claims 2-41 are pending for consideration. Claims 2, 28, 31, and 37 are independent claims.

Telephone Interview

Applicant thanks the Examiner for the courteous and helpful telephone interview held on July 25, 2017 (summarized above).

Response to Double Patenting Rejection

Claim 2 is rejected on the ground of non-statutory (obviousness-type) double patenting over claims 1 and 8 of U.S. Patent No. 8,676,877. Applicant respectfully requests that the double patenting rejection be held in abeyance until the other rejections are overcome.

Response to Rejections under 35 U.S.C. § 103

The Office Action rejects Claims 2, and 4-31 under 35 U.S.C. § 103, arguing that they are unpatentable in view of U.S. Publication No. 2005/0021751 to Block and U.S. Patent No. 8,601,101 to Singh. Applicant respectfully traverses the rejections.

Applicant has amended the claims in an effort to expedite allowance of the application. No new matter is added. The claims as amended recite that the program code for the single-node kernel causes a hardware processor to evaluate mathematical expressions and to produce results of mathematical expression evaluation. Amended claim 1 recites a mechanism for the nodes to communicate results of mathematical expression evaluation with each other using a peer-to-peer architecture. Amended claim 28 recites a mechanism for the nodes to communicate results of mathematical expression evaluation with each other using asynchronous calls. Amended claim 31 recites a mechanism for the nodes to communicate results of mathematical expression evaluation with each other. The amended claims recite that at least one of the nodes is configured to return at least one result of evaluation to a user interface or a script.

The amended claims are patentable over the cited references because the cited references in combination do not disclose the computer cluster with the unique combination of features

recited in each claim. For example, the cited references do not disclose at least the combination of features identified in the preceding paragraph. Neither Block nor Singh discloses a computer cluster node having a mechanism for communicating results of mathematical expression evaluation with other nodes of the computer cluster or a computer cluster node that returns at least one result of mathematical expression evaluation to a user interface or a script. For example, Singh does not disclose that the nodes return results of mathematical expression evaluation to the cluster configuration service 1510 described in column 14, lines 61-66 of the reference.

For at least these reasons, Applicant respectfully submits that Claims 2-31 are allowable over the cited references and respectfully requests that the rejections of these claims be withdrawn.

New claims 32-41

New claims 32-41 are added. No new matter is added. Dependent claims 32-36 are patentable because each depends from a patentable claim and because each recites additional combinations of features that are patentable. New independent claim 37 is patentable because the cited references do not disclose a computer cluster node with the combination of functionality recited in claim 37. Dependent claims 38-41 are patentable because of their dependence on claim 37 and because each recites additional combinations of features that are patentable.

No Disclaimers or Disavowals

Applicant respectfully submits that the claims are in condition for allowance. Furthermore, any remarks in support of patentability of one claim should not be imputed to any other claim, even if similar terminology is used. Any remarks referring to only a portion of a claim should not be understood to base patentability on that portion or that the limitation discussed is essential or critical; rather, patentability must rest on each claim taken as a whole. Applicant respectfully traverses each of the Examiner's rejections and each of the Examiner's assertions regarding what the prior art shows or teaches, even if not expressly discussed herein. Although the present communication may include alterations to the application or claims, or characterizations of claim scope or referenced art, applicant is not conceding in this application that previously pending claims are not patentable over the cited references. Rather, any

alterations or characterizations are being made to facilitate expeditious prosecution of this application. Applicant reserves the right to pursue at a later date any previously pending or other broader or narrower claims that capture any subject matter supported by the present disclosure, including subject matter found to be specifically disclaimed herein or by any prior prosecution. Accordingly, reviewers of this or any parent, child or related prosecution history shall not reasonably infer that applicant has made any disclaimers or disavowals of any subject matter supported by the present application.

CONCLUSION

For the foregoing reasons, it is respectfully submitted that the rejections set forth in the outstanding Office Action are inapplicable to the present claims. Accordingly, issuance of a Notice of Allowance is requested.

The undersigned has made a good faith effort to respond to all of the rejections in the case and to place the claims in condition for immediate allowance. Nevertheless, if any undeveloped issues remain or if any issues require clarification, the Examiner is respectfully requested to call the undersigned at 949-721-5280.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: August 8, 2017

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

25486182

SUMMARY OF INTERVIEW

Attendees, Date and Type of Interview

A telephonic interview was conducted on July 25, 2017 and attended by Examiner Hermon Asres and Applicant's representative, Lance D. Smemoe (reg. no. 66,152).

Exhibits and/or Demonstrations

None

Identification of Claims Discussed

The pending claims, including claim 1, were discussed. New claim 34 as set forth in the claim amendments presented herein was also discussed.

Identification of Cited/Disclosed Art Discussed

The art of record, including U.S. Patent No. 8,601,101 (Singh) and U.S. Patent Application Publication No. 2005/0021751 (Block et al.), was discussed.

Proposed Amendments

Substantially as set forth herein.

Principal Arguments and Other Matters

The claims as amended recite subject matter that is novel and non-obvious in view of the art of record. Specific arguments are set forth in the following remarks.

Results of Interview

The examiner indicated that the arguments and amendments appeared to overcome the rejections. The examiner will reconsider the allowability of the claims after the Applicant files a formal response.

AMENDMENTS TO THE CLAIMS

The following listing of the claims replaces all previous listings of the claims.

1. (Canceled)

2. (**Currently amended**) A computer cluster comprising:

a plurality of nodes comprising a hardware processor, wherein one or more of the nodes are configured to receive ~~receives~~ a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel that, when executed, causes the hardware processor to interpret user instructions, to evaluate mathematical expressions, and to produce results of mathematical expression evaluation; and

a mechanism for the nodes to communicate results of mathematical expression evaluation with each other using a peer-to-peer architecture;

wherein at least one of the nodes ~~returns a result~~ is configured to return at least one result of mathematical expression evaluation to a user interface or a script.

3. (Previously presented) The computer cluster of claim 2, wherein the single-node kernel comprises a Mathematica kernel.

4. (Previously presented) The computer cluster of claim 2, wherein the mechanism comprises a message passing interface.

5. (Previously presented) The computer cluster of claim 2, wherein each of the nodes comprises one or more cluster node modules.

6. (Previously presented) The computer cluster of claim 2, wherein each single-node kernel is stored in a non-transitory computer-readable medium and configured to accept and execute a request.

7. (Previously presented) The computer cluster of claim 6, wherein each of the nodes comprises one or more cluster node modules, wherein each of the cluster node modules comprises instructions stored in a non-transitory computer-readable medium, and wherein the instructions, when executed by the hardware processor, cause the cluster node module to communicate with the single-node kernel and with one or more other cluster node modules.

8. (Previously presented) The computer cluster of claim 7, wherein the plurality of cluster node modules act as a cluster.

9. (Previously presented) The computer cluster of claim 7, wherein the plurality of cluster node modules communicate with one another to act as a cluster.

10. (Previously presented) The computer cluster of claim 9, wherein the computer cluster includes the user interface.

11. (Previously presented) The computer cluster of claim 10, wherein each cluster node module accepts instructions from the user interface and interprets one or more of the instructions.

12. (Previously presented) The computer cluster of claim 2, wherein the cluster initialization process comprises establishing communication among two or more of the nodes.

13. (Previously presented) The computer cluster of claim 2, wherein the cluster initialization process comprises configuring access by one or more of the nodes to a computer-readable medium comprising program code for the single-node kernel.

14. (Previously presented) The computer cluster of claim 13, wherein the cluster initialization process comprises establishing message-passing support among the nodes in the cluster.

15. (Previously presented) The computer cluster of claim 13, wherein the cluster initialization process comprises launching cluster node modules by the cluster.

16. (Previously presented) The computer cluster of claim 15, wherein the cluster initialization process comprises assigning a processor identification number to each of the cluster node modules.

17. (Previously presented) The computer cluster of claim 2, wherein the cluster initialization process comprises:

launching cluster node modules by the cluster; and

after launching the cluster node modules, configuring access by one or more of the nodes to a non-transitory computer-readable medium comprising program code for the single-node kernel.

18. (Previously presented) The computer cluster of claim 2, wherein the cluster initialization process comprises:

launching cluster node modules by the cluster;

after launching the cluster node modules, establishing communication among two or more of the nodes; and

after establishing communication among two or more of the nodes, assigning a processor identification number to each of the cluster node modules.

19. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to communicate at least some of the user instructions.

20. (Previously presented) The computer cluster of claim 19, wherein one or more of the nodes are configured to communicate at least some of the user instructions to one or more single-node kernels.

21. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to accept user instructions via one or more of the nodes.

22. (Previously presented) The computer cluster of claim 21, wherein one or more of the nodes are configured to communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other.

23. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to transmit at least some of the user instructions that originate from a user interface.

24. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to parallelize at least some of the user instructions before communicating at least some of the user instructions to one or more single-node kernels.

25. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to accept user instructions before communicating at least some of the user instructions to one or more single-node kernels.

26. (Previously presented) The computer cluster of claim 2, wherein one or more of the nodes are configured to:

accept user instructions;

after accepting user instructions, communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other; and

after communicating at least some of the user instructions using the mechanism, communicate at least some of the user instructions to one or more single-node kernels.

27. (Previously presented) The computer cluster of claim 2, wherein the plurality of nodes are configured to communicate with one another to interpret and translate commands for execution by a plurality of single-node kernels.

28. (**Currently amended**) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes are configured to receive:

a command to start a cluster initialization process for the computer cluster, wherein the cluster initialization process comprises establishing communication among two or more of the nodes; and

an instruction from a user interface or a script;

and

a mechanism for the nodes to communicate results of mathematical expression evaluation with each other using asynchronous calls;

wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel configured to interpret user instructions, to evaluate mathematical expressions, and to produce results of mathematical expression evaluation; and

wherein at least one of the nodes ~~returns a result~~ is configured to return at least one result of mathematical expression evaluation to the user interface or the script.

29. (Previously presented) The computer cluster of claim 28, wherein the asynchronous calls comprise a first command to create a first packet containing:

an expression to be sent as payload; and

a target node where the expression should be sent;

wherein the first command is configured to be called from within a single-node kernel;

wherein the single-node kernel is configured to send the first packet to a local cluster node module; and

wherein the local cluster node module is configured to forward the expression to the target node.

30. (Previously presented) The computer cluster of claim 29, wherein the asynchronous calls comprise a second command to create a second packet containing:

a location where the expression is expected to be received; and

a sending node from which the expression is expected to be received;

wherein the second command is configured to be called from within a single-node kernel;

wherein the single-node kernel is configured to send the second packet to a local cluster node module; and

wherein the local cluster node module is configured to store the second packet contents in a message receiving queue.

31. (**Currently amended**) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes are configured to receive:

a command to start a cluster initialization process for the computer cluster, wherein the cluster initialization process comprises establishing communication among two or more of the nodes; and

an instruction from a user interface or a script;

and

a mechanism for the nodes to communicate results of mathematical expression evaluation with each other;

wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel configured to interpret user instructions, to evaluate mathematical expressions, to produce results of mathematical expression evaluation, and to interpret the user instructions into commands that are executable by a special purpose microprocessor; and

wherein at least one of the nodes ~~returns a result~~ is configured to return at least one result of evaluation to the user interface or the script.

32. (**New**) The computer cluster of claim 5, wherein each of the plurality of nodes implements asynchronous calls that enable the single-node kernel to perform computation tasks while the cluster node modules are simultaneously communicating with one another.

33. (**New**) The computer cluster of claim 5, wherein intercommunication among the plurality of single-node kernels during thread execution is enabled by the plurality of cluster node modules, and wherein the computer cluster is configured to permit exchange of information between nodes during the course of a parallel computation.

34. (**New**) The computer cluster of claim 5, wherein each of the single-node kernels are configured to call a send command involving creating a packet containing:

an expression to be sent as payload; and

where the expression should be sent;

wherein, upon reception of the send command by a local cluster node module associated with the single-node kernel, the local cluster node module is configured to decode the packet and forward a payload to the cluster node module specified in the packet;

wherein each of the single-node kernels is configured to call a receipt command involving creating a packet specifying which message to test for completion and to then wait for a reply expression to evaluate;

wherein, upon reception of the receipt command by the local cluster node module associated with the single-node kernel, the local cluster node module is configured to decode the packet and use a message specifier to search for any matching expressions listed as completed in a received message queue;

wherein, upon determining that such a completed expression is found, the local cluster node module is configured to send the completed expression to a local single-node kernel associated with the cluster node module in response to the receipt command, and

wherein each of the single-node kernels is configured to receive the completed expression and to update variables used by the single-node kernel during evaluation of expressions.

35. **(New)** The computer cluster of claim 2, wherein the plurality of nodes are configured to permit exchange of information between nodes during the course of parallel computation.

36. **(New)** The computer cluster of claim 2, wherein each of the plurality of nodes comprises instructions executable by the hardware processor and configured to implement asynchronous behavior, wherein the instructions comprise:

a first instruction to asynchronously send a payload to another node;

a second instruction to asynchronously receive a payload from another node; and

a third instruction to search for a payload matching a message specifier.

37. **(New)** A computer cluster node for evaluating expressions in parallel with other computer cluster nodes, the computer cluster node comprising:

a hardware processor configured to access one or more non-transitory memory devices comprising program code for a single-node kernel that, when executed, causes

the hardware processor to interpret user instructions, to evaluate mathematical expressions, and to produce results of mathematical expression evaluation;

a user connection interface configured to receive a command to start a cluster initialization process for a computer cluster; and

a mechanism to communicate results of evaluation with other computer cluster nodes using a peer-to-peer architecture;

wherein the user connection interface is configured to return at least one result of mathematical expression evaluation to a user interface or a script.

38. (**New**) The computer cluster node of claim 37, wherein the one or more non-transitory memory devices comprise program code for performing a parallel fast Fourier transform, wherein the program code causes the hardware processor to evaluate a command to perform a Fourier transform on an array comprising a first data portion that is stored on the computer cluster node and a second data portion that is not stored on the computer cluster node.

39. (**New**) The computer cluster node of claim 37, wherein the computer cluster node is configured to permit exchange of information with other computer cluster nodes during the course of parallel computation.

40. (**New**) The computer cluster node of claim 37, wherein the computer cluster node comprises instructions executable by the hardware processor and configured to implement asynchronous behavior, wherein the instructions comprise:

a first instruction to asynchronously send a payload to another node;

a second instruction to asynchronously receive a payload from another node; and

a third instruction to search for a payload matching a message specifier.

41. (**New**) The computer cluster node of claim 37, wherein the hardware processor comprises a special purpose microprocessor.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Hermon Asres
Art Unit	:	2449
Conf. No.	:	6874

RESPONSE TO OFFICE ACTION DATED MARCH 9, 2017

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office Action dated March 9, 2017, please reconsider the rejections in view of the following remarks.

Amendments to the Claims are reflected in the listing of claims which begins on page 2.

Summary of Interview begins on page 9.

Remarks/Arguments begin on page 10.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 October 2004 (07.10.2004)

PCT

(10) International Publication Number
WO 2004/086270 A2

(51) International Patent Classification[7]: G06F 17/60

(21) International Application Number: PCT/IL2004/000255

(22) International Filing Date: 18 March 2004 (18.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
155112 27 March 2003 (27.03.2003) IL

(71) Applicant *(for all designated States except US)*: **ECI TELECOM LTD.** [IL/IL]; 30 Hasivim Street, 49517 PETACH TIKVA (IL).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KATZ, Igor** [IL/IL]; 128/4 Hey Beiyar Street, Rosh Ha'ayin 48256 (IL). **GOZMAN, Gennady** [IL/IL]; 9b Arie Ben Eliezer Street, 72285 RAMLA (IL). **YEKUTIELI, Oren** [IL/IL]; 18 Nahal Hayarkon, Modi'in 71700 (IL).

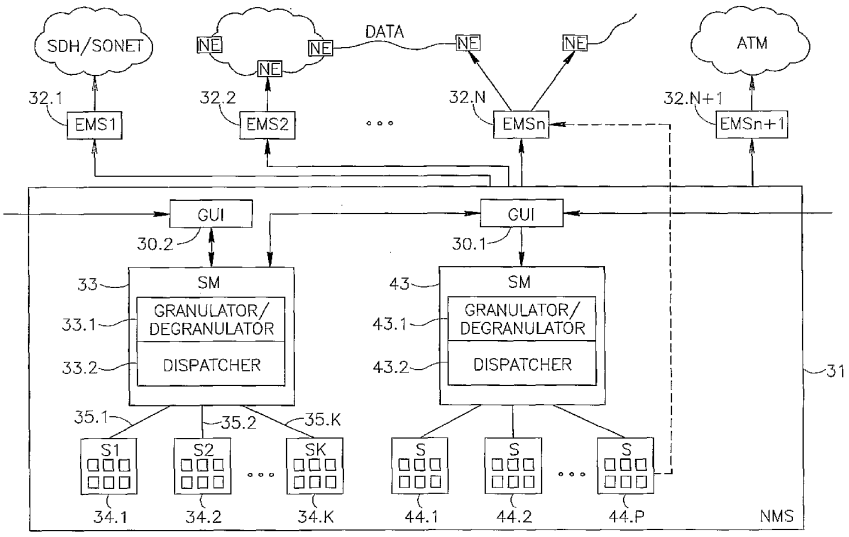
(74) Agents: **INGEL, Gil** et al.; Eci Telecom Ltd., Patent And Trademark Dept., 30 Hasivim Street, Petach Tikva 49517 (IL).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: TECHNIQUE FOR NETWORK MANAGEMENT



(57) Abstract: A network management technology for managing a client requested communication task having sub-task granularity, comprising: establishing K communication links with K communication servers, and maintaining all the K communication links active; dismantling the communication task into L sub-tasks having the same character; performing the L sub-tasks to obtain a series of results; processing the series of results to obtain a final result for the communication task for further communicating said final result to the client.

WO 2004/086270 A2

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Technique for network management

Field of the invention

The present invention relates to the field of client/server interaction in computing systems, where, for example, one computing device ("the client") requests another computing device ("the server") to perform part of the client's work. More particularly, the present invention relates to the computing systems serving telecommunication networks .

Background of the invention

A so-called "distributed" computing allows implementing the known principle of client/server interaction.

The distributed computing allows one process (a "client") running on a machine to delegate some of its work to another process (a "server") running on another machine that might be, for example, better suited to perform that work. The client and server might also be two processes running on the same physical machine.

The benefits of client/server computing have been even further enhanced by the use of a well-known computer programming technology called Common Object Request Broker Architecture, (or CORBA for short), which allows an object oriented programming (OOP) client and a server to be located on different (heterogeneous) "platforms". CORBA provides the object-oriented bridge between the client and the server. A platform is a combination of the specific hardware/software/operating system/communication protocol, which a machine uses to do its work. CORBA allows the client application program and server application

program to operate on their own platforms without worrying how the client application's work requests will be communicated and accepted by the server application. Likewise, the server application does not have to worry about how the CORBA system will receive, translate and send the server application's processing results back to the requesting client application.

Details of how CORBA techniques have been integrated with heterogeneous client/server systems are explained in U.S. Pat. No.5,440,744 and European Published Patent Application No. EP 0 677,943 A2. These latter two publications are hereby incorporated by reference. However, an example of the basic architecture will be given below for contextual understanding of the described environment. In this example, the client and server are on different machines, but, the same software architecture applies like they are on the same machine.

As shown in FIG. 1, the client computer has an application program 40 running on its operating system. The application program 40 will periodically require work to be performed on the server computer 20 and/or data to be returned from the server 20 for subsequent use by the application program 40. It is irrelevant whether the requests for communications services to be carried out by the server are instigated by user interaction with the first application program 40, or whether the application program 40 operates independently of user interaction and makes the requests automatically during the running of the program.

When the client computer 10 wishes to make a request for the server computer 20's services, the first application program 40 informs the first logic means 50 of the service required. The first logic means 50 then handles the task of finding and establishing the necessary communications with the second computer 20. The connection with the

server computer 20 is established via the connection means 80, and a request is subsequently sent of to the second logic means 90.

The second logic means 90 then passes the request on to the second application program 100 (service application program) running on the server computer 20 so that the service application 100 can perform the specific task required by that request, such as running a data retrieval procedure. Once this task has been completed, the service application may need to send results back to the first computer 10. Oftentimes, duplication of server resources (in the form of plural servers) is provided to enable "the server" to handle a large number of simultaneous client requests and to provide for fault tolerance (if one server is defective another can "step in" and take over). The clients think they are communicating with a single "server", but in fact, a plurality (or group) of servers is actually handling the requests (these servers could be located on separate machines, or they could be separate processes running on the same machine). In this context, client requests destined for "the server" are directed to a central gateway or a network manager (common to all of the server machines making up "the server") which performs workload management and load balancing to distribute client requests to the various servers in the group according to any well-known workload balancing policy (or combination of a plurality thereof), such as the round robin policy, where each server is assigned a request in turn. The current workload of each server in the group is also a common factor used in scheduling a new request (i.e., a new request is assigned to a server with a low current workload, wherein the current workload of a server is estimated based on a number of requests sent by the client to this specific server). Specific details of the balancing policy/policies to be used are

selected by the person configuring "the server" system architecture, typically by a systems administrator.

US patent 6,003,083 describes a group of object servers configured for workload management (shown in Fig. 2). A suitable method of performing workload management amongst a group of server objects within an object-oriented client/server computing environment, includes steps of: 1) receiving a client request; 2) accessing a pointer previously exported by the group of servers, the pointer pointing to a workload management unit; 3) performing workload management using the workload management unit on the client request in order to select which server of the group should be used to satisfy the request, and 4) forwarding the client request to the selected server.

In Fig. 2, the server group 23 exports an interoperable object reference (IOR) to each client 21a-21c so that the clients will know how to contact the server group in order to make a request thereof. When a client 21b makes a request to the server group 23, the client first contacts the server process containing an IP router 22. IP router 22 then uses workload management policies to make a decision as to which object server in the group should be selected to serve this particular client request. This workload management decision takes place at the TCP/IP level. That is, the IP router 22 looks at the IP addresses (IP1 to IPn) for each object server 1-n in the server group 23, makes a workload management decision to select a particular object server in the server group and then simply forwards on the request from IPR to the IP address of the selected object server. The router 22 performs workload balancing amongst the IP addresses of the various object servers in the server group 23.

Alternatively, the workload management can be performed by pointing the requesting client to a dynamic domain name server (DNS).

It should be emphasized that the manager unit such as IP router or dynamic domain server forwards each particular request to a specific server.

To the best of the Applicant's knowledge, in all presently known solutions using distributed computing, a client's request has always been treated as an integral entity.

Object of the invention

It is the object of the present invention to provide a way of allowing a greater number of clients to access the resources of a workload managed server group in a more efficient (robust) way.

Summary of the invention

The idea proposed by the Inventors can be formulated in a number of aspects.

According to the first aspect of the invention, there is proposed a technology for a distributed client/server computing system comprising a number of clients and a group (cluster) of K servers managed by a servers manager, wherein

a client's request (a task) having sub-task granularity is decomposed at the management level of the servers group (preferably at the servers manager) to obtain a number L of sub-tasks having identical characters;

communication links are established and maintained between the servers manager and said K servers of the group for a time period not less than the period of handling the initial task,

the sub-tasks are distributed by the servers manager between the servers for execution, via the mentioned constantly maintained communication links,

replies of the servers upon execution of the sub-tasks are collected by said servers manager using the communication links, and composed so as to obtain a result of execution of the initial task.

The distinctive features of the inventive concept are in that each of the servers takes care of one or more sub-tasks of the initial task received by the servers manager, and in that the communication links are established in advance (preferably, even before the first task is obtained) and maintained constantly active, at least while performing the sub-tasks up to completing the task.

Preferably, the above-described client-server computing system is a Network Management System (NMS).

It should be noted that a number of various servers (servers managers, with or without a lower level group or cluster of servers, respectively suitable for different tasks) might change in the NMS in a dynamic manner. Say, an additional server to a cluster can be created in the NMS on demand.

According to the preferred implementation of the above technology, the method is performed by using CORBA tools.

It should be mentioned that CORBA, as a rule, keeps open a communication link between a client and a server, when a task is forwarded to the server for execution. However, if load balancing operations are to be performed, i.e., when tasks are to be distributed between different servers according to the known methods such as those mentioned in the background, a particular connection (communication) link is established "ad hoc" with one of the potential servers, so that to close it after the task is completed.

In the proposed method, more than one communication links are opened and maintained open simultaneously between the

servers' manager and the servers, for example to perform a number of subtasks extracted from the initial task.

Optionally and preferably, the character of the sub-tasks may be identical to character of the initial task.

Since the servers of the group in the proposed concept perform relatively short sub-task(s) and may substitute one another if one of them fails, the system has a high degree of fault tolerance and almost does not reduce its performance in such a critical case. Indeed, if one server fails and another server is made to perform the sub-task of the faulty server, the total lost of processing time is quite low since the sub-tasks are relatively short. Contrary to that, in a regular client-server system with a number of servers processing non-granulated tasks, both the fault tolerance and the performance speed strongly depend on duration of the tasks. Therefore, in case one of the servers fail in the regular client-server system, the longer the task the slower the performance.

Preferably, the computing system is oriented to the purposes of network management. The clients, the servers manager and the servers may be positioned at platforms situated at different locations in a telecommunication network. However, according to the preferred embodiment, the clients, the servers manager and the group of servers are situated at one and the same location and form together a network management system (NMS) responsible for the whole telecommunication network.

It should be noted that the number of sub-tasks formed from the initial task may be and is preferably greater than the number of servers in the group (said servers being capable of executing tasks and sub-tasks of a particular character).

The sub-tasks may be distributed to the servers by portions, wherein each portion comprises one or more sub-tasks to be executed in parallel at a suitable server.

It is proposed that the step of distribution of the sub-tasks between the servers for execution be performed on a first-to reply basis, i.e. the server, which faster reports that one or more sub-tasks (previously forwarded to it) has been executed, receives one or more additional sub-tasks from the servers manager. This feature is considered one of the novel features of the proposed invention since, contrary to the known approach where a new request is assigned to a server with a low current workload (i.e. with a low number of requests sent to the server up to the moment), the Inventors propose assigning one or more sub-tasks from a queue to a server which may have a relatively high current workload, but working faster.

This approach enables judging about efficiency of the server in a very simple way, namely - based on receiving its response that a part of the previously obtained work is executed.

In view of the above, and according to a second aspect of the invention, there is also proposed a technology for managing a plurality of communication tasks, which technology includes:

receiving a plurality of communication tasks at a servers manager; distributing said plurality of communication tasks from said servers manager to a plurality of K servers ;

receiving, at the servers manager, a new task in addition to said plurality of communication tasks; and

dispatching said new task to one of said plurality of K **servers** on a first-to-reply basis.

For both of the mentioned aspects, the technology preferably comprises a step of pre-determining a number C of tasks or sub-tasks which can be sent to a particular server by a particular client. Any server is capable of performing sub-tasks (tasks) in parallel. A kind of monitoring that a particular server performs no more than the agreed number C sub-tasks (tasks) for a specific client is performed at the servers manager, by dispatching the sub-tasks (tasks) and further, if non-dispatched sub-tasks (tasks) still remain, by the above-mentioned first-to reply load balancing. Actually, after the number C has been reached, the client (via its SM) is allowed to add no more tasks or sub-tasks to its specific server, then the number of replies on the executed tasks/subtasks which have been received from said server.

Before distributing the plurality of communication tasks to the plurality of K servers, the technology preferably comprises opening K communication links between each of the respective K servers and the servers manager and maintaining them open up to the last of said tasks is executed.

Preferably, at least a number of said plurality of communication tasks have sub-task granularity, where said number is an integer greater than one, and the technology also comprises the steps of:

dismantling each of said number of communication tasks having sub-task granularity into a plurality of sub-tasks having the same character to provide a total of X separate sub-tasks, where X is an integer greater than one;

providing said X separate sub-tasks to said at least some of the plurality of K servers;

performing said X separate sub-tasks to obtain a series of results;

collecting and processing said series of results to obtain said number of final results to the communication tasks having sub-task granularity.

The above technology is also preferably performable by using CORBA tools.

According to any one of the above aspects, the step of collecting replies (results) from the servers is followed by identifying and arranging the replies to compose the result of execution of the initial task, and further forwarding the result to the client, from which the request (task) was received. The above steps are preferably performed at the servers manager, and the servers manager generally comprises a granulating block responsible for disassembling the task into the sub-tasks (if possible), and a dispatcher block or proxy which is responsible for distributing the sub-tasks between the servers; in the opposite direction, the dispatcher block (proxy) takes care of collecting the replies from the servers, and the granulating block is capable of assembling the result of execution of the initial task from the replies received from the servers.

The tasks which are preferably considered in the frame of the present invention are tasks having sub-task granularity.

The tasks to be carried out by the proposed technology preferably, though not exclusively, are so-called communication tasks being typical for network management.

For example, a task of determining a communication trail can be presented as a task having the sub-task granularity since, to determine a particular trail, a number of sub-trails are to be pointed out and computed, and the process of computing any of the sub-trails has the

same character and features as the process of computing another of the sub-trails.

However, tasks which do not have the sub-task-granularity but have the character suitable for the group of servers can still be handled by the proposed system; such tasks can be distributed (per task) between the servers by the servers manager; the loading balance may be performed by using the proposed first to reply basis. In other words, the proposed architecture is suitable for carrying out both the non-granulated tasks and the granulated tasks in case that non-granulated task has the same character as the sub-tasks. If neither server from the described group of servers is capable of performing a specific task, or this specific task cannot be granulated into sub-tasks suitable for the group of servers (which may or may not be a case), the servers manager should recognize such a fact and forward this specific task to another server capable of handling such specific tasks. That another server may control its own group of servers for handling the mentioned specific tasks.

Actually, there may be a kind of sub-task-granulated tasks, where such a task can be divided into a number of sub-tasks of the same character and a portion of the task having a different character.

To perform such a task, the portion having a different character can be sent to another server (like it was described for a task having a different character) communicating with the servers manager. This another server may (or may not) have its own group of lower level servers.

In the frame of the present application, the client(s) from which the servers manager receives one or more tasks is (are) preferably represented by one or more parts of the Network Management System

(NMS). For example, GUI (Graphical User Interface) being part of the NMS may represent a client by visualizing a list of required trails.

Further, it should be understood that each server from the group (cluster) of servers in the network management system NMS may itself perform functions of the servers manager with respect to servers of yet a lower level. Such lower level servers can be, for example, Element Manager Systems (EMS) in the network, capable of performing requests which are forwarded to them by one or more servers of the NMS. (EMS is usually responsible for a number of network elements). The requests may be performed as are, or upon being separated into sub-requests according to the principles described above. Replies of EMS will be collected at suitable servers. For example, if the initial task received by the servers manager of NMS from one of its clients is to determine a trail, servers of the NMS will process sub-trails of the trail as sub-tasks of the initial task. Each of the sub-tasks may be presented as a request of the server to its lowest level server (in our example, EMS) so that a number of EMS will be capable of processing sections of the sub-trails (as sub-requests or sub-sub-tasks). The sub-requests, upon being executed, will be collected and then arranged into replies to sub-tasks at the servers of the NMS.

It should be appreciated, that the Network Management System can be built according to a multi-level model, where the described principles of executing tasks having sub-task granularity are used at least at one of the management levels.

The concepts of the described technology will further be formulated both in terms of a method, and in terms of a system.

Brief description of the drawings

The invention will further be described with the aid of the following non-limiting drawings, in which:

Fig. 1 (prior art) illustrates one known architecture for performing the distributed computing.

Fig. 2 (prior art) shows another known architecture that enables performing the distributed computing with workload balancing.

Fig. 3 illustrates one embodiment of architecture (a system) that enables network management according to the invention, using the method of distributed computing.

Fig. 4 schematically illustrates another embodiment of a system according to the present invention, with a multi-level management principle.

Detailed description of the preferred embodiments

Fig. 3 illustrates an embodiment of a network architecture capable of implementing the inventive method. Network Management System NMS (marked 31) is responsible for a combined network comprising SDH/SONET, data and ATM network sections. Element Manager Systems ("N" EMS.k blocks marked 32.1 – 32.N) are controlled by the NMS 31. On the other hand, each EMS controls its related network element(s) NE and/or a suitable network section. Operation of the Network Management System 31 should be understood as a process of handling requests (tasks) received from clients which, in this example, are represented by GUI (Graphical User Interfaces) 30.1 and 30.2 forming part of the NMS 31. GUI is in communication with a human operator on one hand, and with a network via a number of servers on the other hand.

In this particular example we discuss telecommunication tasks having sub-task granularity. The sub-tasks are such fragments of the task, which have the same character and therefore need similar means for their execution. According to the invention, the sub-tasks can be separated from the task and handled in parallel, using a number of similar processing means in the NMS. For example, if any of the clients (GUI) applies to the NMS with a task of determining a trail, it can be divided into a number of sub-tasks of determining sub-trails. If both clients apply with such tasks, both of the tasks will be divided into similar sub-tasks. The clients are connected to a processor called Server Manager (SM marked 33) of the NMS 31, where a granulator/degranulator 33.1 is responsible for de-granulating the task(s) into a number of sub-tasks, and a dispatcher 33.2 is responsible for temporary storing and distributing the sub-tasks between a group of K lower level processors (servers) S1 – SK marked 34.1 – 34.K. Small squares indicated in the servers 34 symbolize their hardware/software means capable of carrying out various sub-tasks. Sets of sub-tasks in the servers 34 are similar, and correspond to the set of tasks, which the Servers Manager 33 is operative to resolve. Actually, NMS 31 may be adapted for resolving other tasks, by means of other processing means, for example by means of an additional servers manager (31) and its group of servers (33). Communication links, which can be established between the server manager SM 33 and the servers S1-S.K, are indicated as 35.1 – 35.K.

When a task is received at the Servers Manager SM 33, the SM is perceived by the client as its direct server. Actually, the servers manager SM 33, acts as a proxy responsible for distributing the sub-tasks between servers 341, 34.2, ...34.K for execution. A number of communication

links 35 (so-called network interface) already exists between the SM and a group of relevant servers, and the distribution is performed upon selecting the servers by the SM when the task is received.

The sub-tasks can be distributed by portions of one or more sub-tasks per one server. Results of the execution are received by the SM 33 via the links 35 which remain active. The sub-tasks which remain not yet distributed, can be sent to one or more servers by the SM 33 according to a so-called first-to-reply principle of workload balance proposed by the Inventors, i.e. the faster a particular server works, the more additional sub-tasks it receives for execution. The system usually sets a limitation on a number (C) of sub-tasks which can be sent in parallel by a particular client (GUI in this example) to a specific server. It should be noted that usually, the similar limitation can be set for another client so that each particular server may serve different clients while the load balancing can be performed either per client, or per a common pool of sub-tasks. For example, the SM 33, while handling a firstly received task or a couple of tasks, may start carrying out one or more further tasks having the character suitable for taking care by the servers 34.1 – 34.K. Such an additional task may, for example, arrive from another client (say, GUI 30.2). It is important, that while some of the servers terminate handling sub-tasks of one task, some of the servers may already start handling sub-tasks of another task. Actually, a number of tasks may be received and de-granulated into a pool of sub-tasks, and these sub-tasks may be dispatched by the SM 33 between the servers of the group. SM 33 will then be responsible for tracking the relationship: which sub-task relates to which task (and, of course, to which client).

In this example, the SM 33 is responsible for collecting replies from the servers (block 33.2) and for combining the replies into an integrated

response to the task (block 33.1) that is then sent to the client (GUI) from which the task was received. Based on the resulting reply to the task, NMS 31 may initiate control of the suitable network sections. Ways of implementation of the task solution obtained in the NMS are not discussed in detail in the present application. Schematically, direction of implementing the obtained solution is illustrated by arrows outgoing from the NMS 31 towards Element Management Systems EMS 32.1 – 32.N+1. For example, in the most usual network task "create/activate path", the path can be "created" by a pathfinder utilizing SM 33 and servers 34, and then "activated" via another servers manager (SM 43) associated to it lower level servers (44.1 ...44.P) respectively connected to a plurality of EMS 32 (one dotted connection is shown). The server managers 33 and 43 may be interconnected directly, that option is illustrated in Fig. 4.

Fig. 4 shows yet another example of NMS architecture. The servers manager SM 33 is connected to one group of servers 34 (34.1...34.K). Let the servers manager SM 33 comprises a task discriminator and a task portion discriminator (not shown). In this configuration, the servers manager SM 33 is also connected to an additional servers manager SM 37 of the same level, and the SM 37 has a further group of servers 38.1, 38.2 of the lower level. Such an additional server 37 with its group of servers may be intended, for example, for handling tasks or task portions having the character different than that of the tasks/subtasks which can be handled by servers 33 and 34. The SM 37 may receive its work both independently (from a particular client), or from the SM 33. The tasks assigned to the SM 37 may be granulated or not; may be dispatched to the servers 38.1-38.2 or not. The distribution (dispatching) thereof to the

servers 38.1, 38.2 can also be performed on the first to reply basis. Tasks that by any reason arrived to the SM37 (or portions of such tasks), but which are unsuitable for treatment by the SM37, can be forwarded to another server of the same level. This function can be implemented in the SM 33 by cooperation of the task/task portion discriminator and a dispatcher/proxy (not shown). A bi-directional arrow between the blocks 33 and 37 illustrates the mentioned option, though yet another SM of the same level may exist for intercepting tasks differing both from those handled by SM 37, and those handled by SM 33.

For example, GUI presents a task of determining a trail through two different types of network. This task can be de-granulated at SM33 so that SM33 with its group of servers be responsible for carrying out the portion of task relating to one type of the network, while an additional one or more servers would be intended for performing a remaining portion of the task related to the other network type.

In another example, GUI may send to SM 37 a task "create/activate a path". The SM 37 finds the path consisting of a number of path sections defined by servers 38.1 and 38.2, and sends description of the defined path to the SM 33 for activation; the activation is performed by utilizing servers_34.1 – 34.K and a number of suitable EMS of the lower level which create real connections in and between the required network elements. The proposed features of granulating the task, dispatching sub-tasks between the servers via open links and maintaining load balance are applicable in this case as well.

All the communication links 35 (interface) established in the NMS 31are maintained active during the time of execution of all the tasks, sub-tasks and sub-sub-tasks, and returning the results to the SM 33.

To this end, Fig. 4 illustrates a multi-level management system, comprising the architecture where the servers S1 – SK represent middle level managers which are provided, in turn, with a plurality of lower level management blocks. Such lower level management blocks may be, for example, Element Management Systems: EMSA, EMSA1,....EMSZ, EMSZ.1 For example, these EMSA – EMSZ blocks are management systems of network elements, every EMS is responsible for network elements inside a particular network section if it is included in the sub-trail which is processed by the suitable server S1 – SM. Any of the servers S1 – SM is capable of managing the sub-tasks in such a manner that those sub-tasks which have sub-sub-task granularity can be distributed between the EMSA – EMSZ servers by portions of the sub-tasks, i.e. by sub-sub-tasks. The middle level servers may utilize all the features described above with respect to the higher level of the network management, such as the feature of workload balancing and the feature of maintaining the communication links active up to resuming all the tasks received at the upper level.

It should be noted that, for performing the above-mentioned process, the lower level managers EMSA...EMSZ controlled by servers S1 – S.K should belong to one and the same type of network which suits to the type handled by the SM 33. In other words, if a sub-task is to calculate a sub-trail in the ATM network section, the EMS managers executing portions of the sub-task may be only the EMS managers of the ATM network sections. It is also important to note that EMS blocks are not mutually replaceable if one of them fails, since each of them is responsible for its specific section of the network. However, one EMS block may serve different servers 34 if more than one particular servers

need processing with respect to the specific network section managed by a specific EMS (connections shown by dotted lines).

The architecture illustrated in Fig. 4 is multi-level also due to the fact that NMS 31 can be served by two sub-NMS (SNMS) 41 and 42. In our example, sub-NMS 41 and 42 are responsible for different sub-networks of the network governed by NMS 31. In turn, structure of the SNMS 41 and 42 may also be multi-level.

Though the present invention has been described with the aid of examples of particular embodiments, it should be appreciated that other modifications of the described configurations can be proposed and claimed based on the disclosure, and are to be considered part of the present invention.

CLAIMS:

1. A network management method for managing a client requested communication task having sub-task granularity, the method comprising steps of:

establishing K communication links with K communication servers, where K is an integer greater than one, and maintaining all said K communication links active

dismantling said communication task into L sub-tasks having the same character, where L is an integer greater than one;

providing said L sub-tasks to at least some of the K communication servers;

performing said L sub-tasks to obtain a series of results;

processing said series of results to obtain a final result for the communication task for further communicating said final result to the client.

2. The method according to claim 1, further comprising the step of maintaining load balancing among said communication servers while performing said L sub-tasks.

3. The method according to Claim 2, wherein the load balancing is performed on the first-to-reply basis.

4. A network management method for managing a plurality of communication tasks, the method comprising the steps of:

receiving a plurality of communication tasks at a servers manager; distributing said plurality of communication tasks from said servers manager to a plurality of K servers ;

receiving, at the servers manager, a new task in addition to said plurality of communication tasks; and

dispatching said new task to one of said plurality of K servers on a first-to-reply basis.

5. The method according to Claim 4, comprising preliminarily establishing communication links with at least some of said plurality of servers, and maintaining all said communication links active up to completing the last of said tasks.

6. The method according to claim 4 or 5, wherein two or more of said plurality of communication tasks have sub-task granularity, and the method also comprises the steps of:

dismantling each of said number of communication tasks having sub-task granularity into a plurality of sub-tasks having the same character to provide a total of X separate sub-tasks, where X is an integer greater than one;

providing said X separate sub-tasks to said at least some of said plurality of K servers;

performing said X separate sub-tasks to obtain a series of results;

collecting and processing said series of results to obtain said number of final results to the communication tasks having sub-task granularity.

7. The method according to any one of Claims 1 to 6, further comprising steps of

- selecting a number C of tasks or sub-tasks allowed for dispatching from a particular client to a particular server, and

-dispatching not more additional tasks or sub-tasks from said client to said particular server, than a number of replies on the executed tasks/subtasks received by said client from said server.

8. The method according to any one of Claims 2 to 7, implementing CORBA tools.

9. A network management system for managing a client requested communication task having sub-task granularity, the system comprising:

a network interface operative to establish K communication links with K communication servers and to maintain all said K communication links active, where K is an integer greater than one; and

a servers manager being a processor operatively associated with the network interface and operative to perform the following:

dismantle said communication task into L sub-tasks having identical character, where L is an integer greater than one,

provide said L sub-tasks to at least some of the K communication servers for performance of the L sub-tasks by the at least some of the K communication servers,

receive, via the network interface, a series of results resulting from performance of the L sub-tasks by the at least some of the K communication servers,

process said series of results to obtain a final result to the communication task; and

communicate said final result to the client.

10. The system according to claim 9, wherein said processor is also operative to maintain load balancing among said K communication servers while performing said L sub-tasks.

11. The system according to claim 10 wherein said load balancing is performed by said processor on the first-to-reply basis.

12. A network management system for managing a plurality of communication tasks, the apparatus comprising:

a plurality of servers; and

a processor being a servers manager associated with said plurality of servers and operative to perform the following:

receive said plurality of communication tasks;

distribute said plurality of communication tasks among said plurality of servers;

receive a new task in addition to said plurality of communication tasks; and

dispatch said new task to one of said plurality of servers on a first-to-reply basis.

13. The system according to Claim 12, further comprising a network interface operatively controlled by said processor and operative to establish communication links with at least some of said plurality of servers, and to maintain all said communication links active up to completing performance of said tasks.

14. The system according to claim 12 or 13, wherein at least some of said plurality of communication tasks have sub-task granularity, and wherein:

the servers manager being operative to dismantle each task having sub-task granularity into a plurality of sub-tasks to provide a total of X separate sub-tasks having the same character , where X is an integer greater than one; and

said servers manager is also operative to dispatch said X separate sub-tasks to said at least some of the plurality of servers for performance thereby, and to collect and process a series of results received from said at least some of the plurality of servers and resulting from performance of said X separate sub-tasks, thereby to obtain final results to the tasks having sub-task granularity, and

said servers manager is also operative to output said final results to clients.

15. The network management system according to any one of Claims 9 to 14, being further capable of

- selecting a number C of tasks or sub-tasks allowed for dispatching from a particular client to a particular server, and

-dispatching not more additional tasks or sub-tasks from said client to said particular server than a number of replies on the executed tasks/subtasks received by said client from said server.

16. The system according to any one of Claims 9 to 15, implementing CORBA tools.

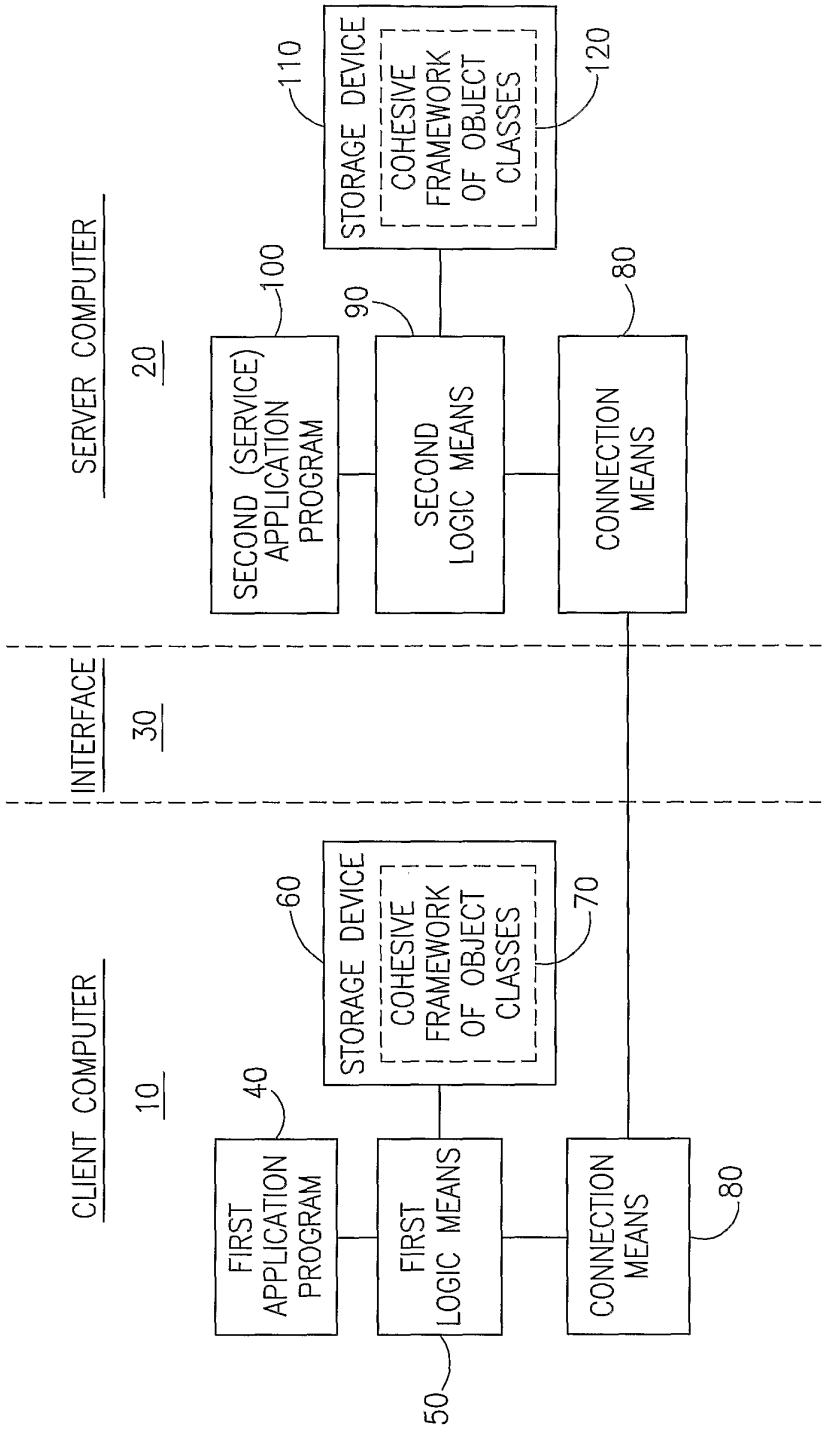


FIG.1
PRIOR ART

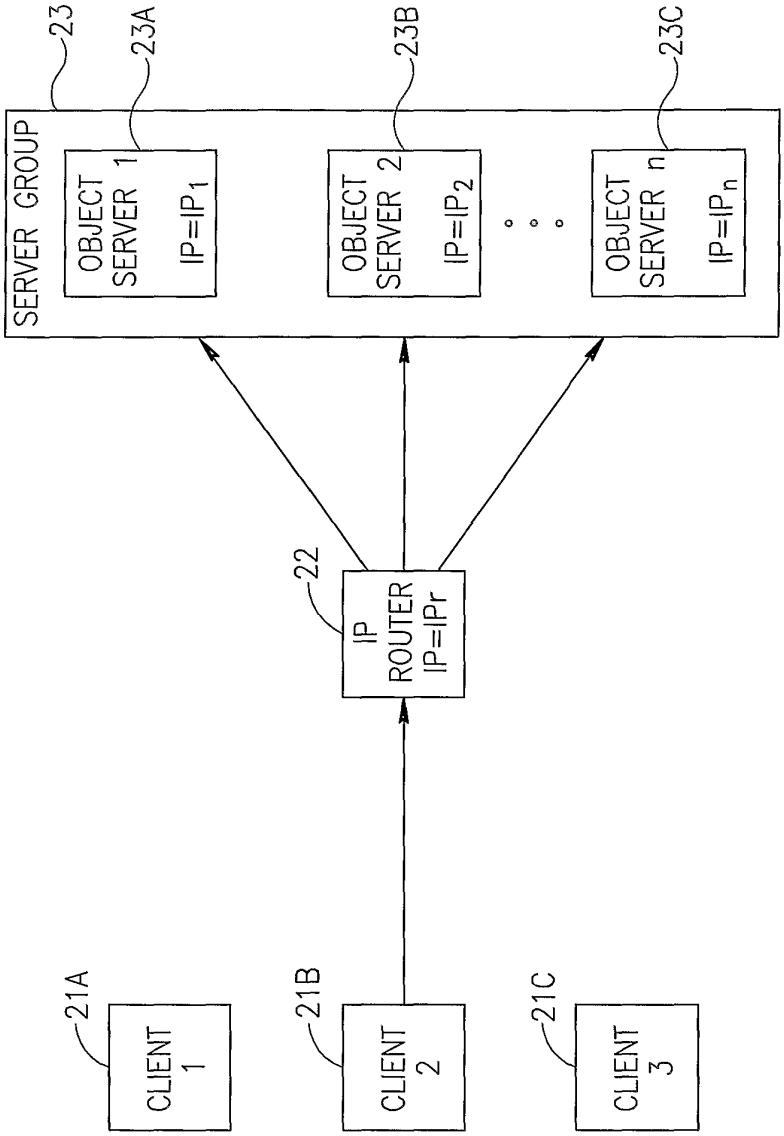


FIG.2
PRIOR ART

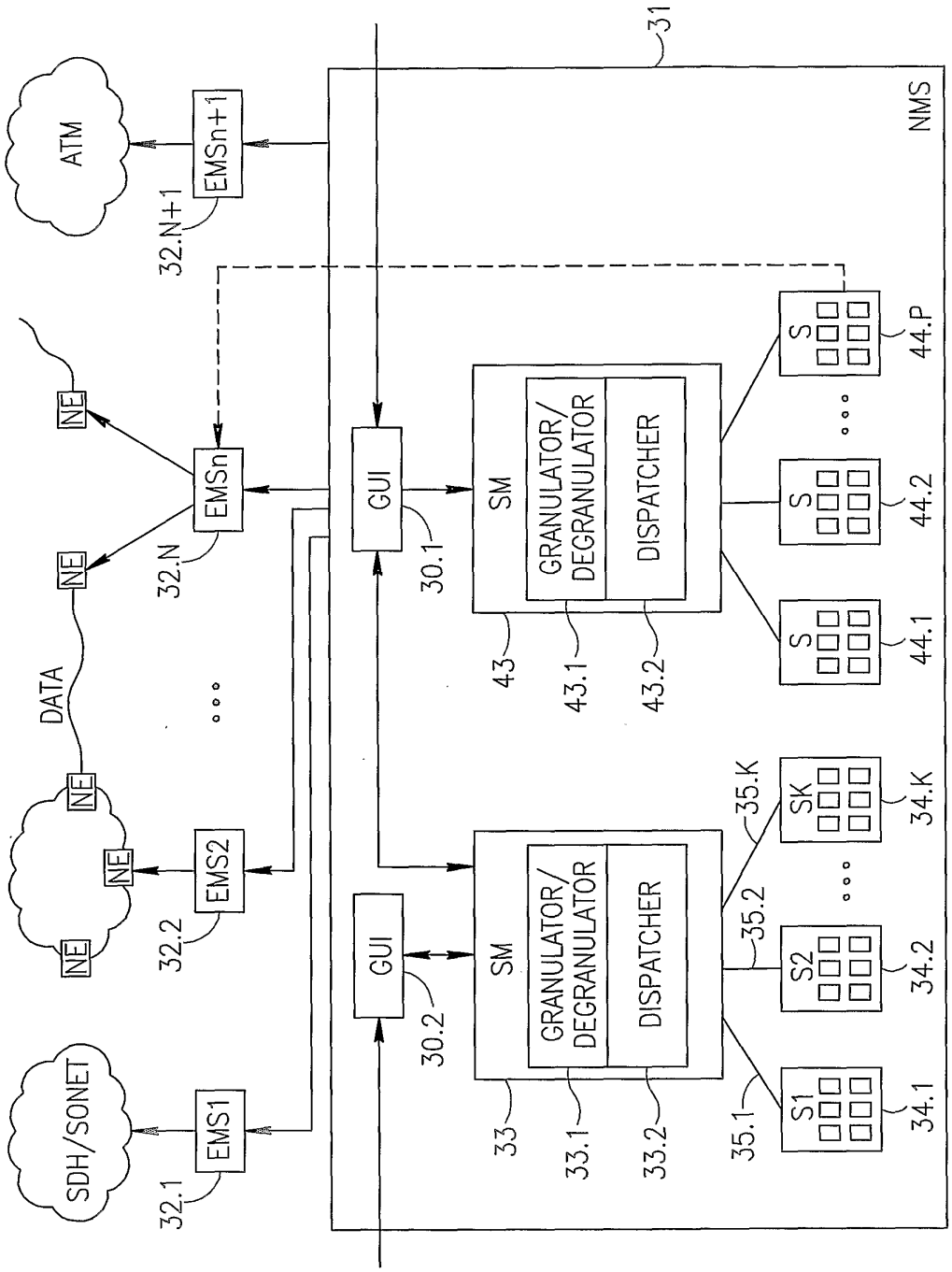

ATM
EMSn+1
32.N+1
31
NMS
NE
EMSn
32.N
GUI
30.1
SM
GRANULATOR/
DEGRANULATOR
DISPATCHER
43
43.1
43.2
S
44.1
44.2
44.P
DATA
NE
EMS2
32.2
GUI
30.2
SM
GRANULATOR/
DEGRANULATOR
DISPATCHER
33
33.1
33.2
35.1
35.2
35.K
S1
S2
SK
34.1
34.2
34.K
SDH/SONET
EMS1
32.1

FIG.3

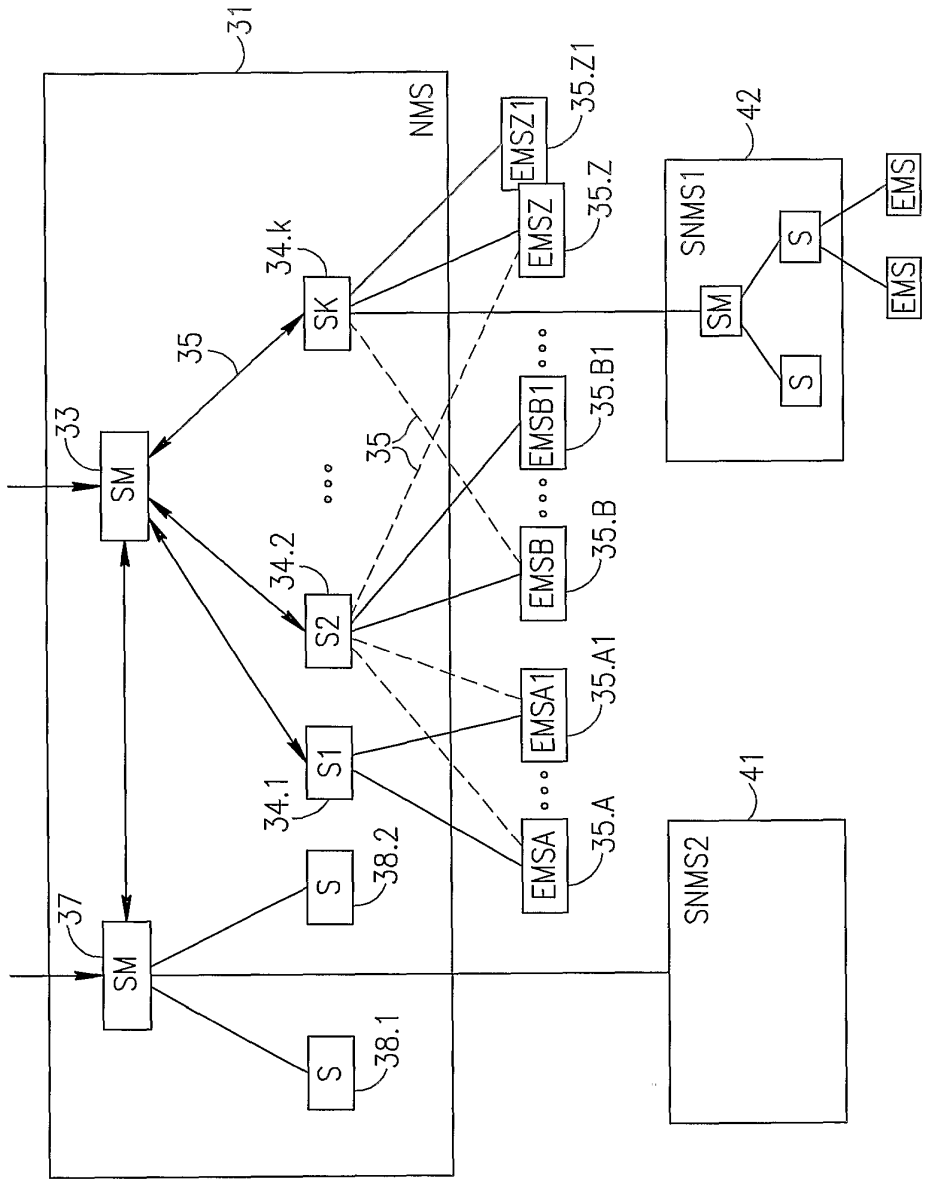

31
NMS
33
SM
37
SM
35
34.k
SK
34.2
S2
34.1
S1
S
38.2
S
38.1
35
EMSZ1
35.Z1
EMSZ
35.Z
EMSB1
35.B1
EMSB
35.B
EMSA1
35.A1
EMSA
35.A
42
SNMS1
SM
S
S
EMS
EMS
41
SNMS2

FIG.4

Espacenet

Bibliographic data: EP1368948 (A2) — 2003-12-10

METHOD AND APPARATUS FOR LARGE PAYLOAD DISTRIBUTION IN A NETWORK

Inventor(s): SIM SIEW YONG [US]; CHAN DESMOND CHO-HUNG [US] ± (SIM, SIEW, YONG, ; CHAN, DESMOND, CHO-HUNG)

Applicant(s): PRISMEDIA NETWORKS INC [US] ± (PRISMEDIA NETWORKS, INC)

Classification: - international: ***G06F17/30; H04L12/803; H04L29/06;*** (IPC1-7): G06F17/30; H04L29/06
- cooperative: **G06F17/30194; G06F17/30233; H04L47/125; H04L67/1095**

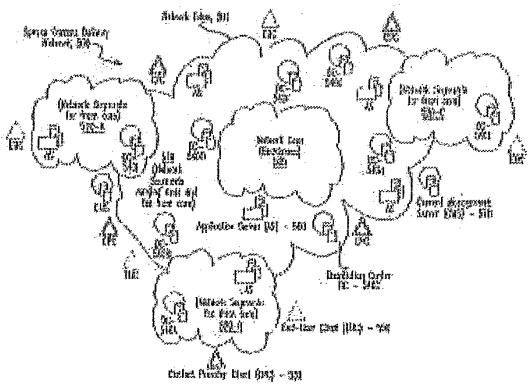
Application number: EP20010989004 20011026 Global Dossier

Priority number(s): WO2001US32638 20011026 ; US20000266286P 20001026 ; US20010681644 20010515 ; US20010681665 20010517 ; US20010681668 20010518 ; US20010681669 20010518 ; US20010681671 20010518 ; US20010681672 20010518 ; US20010681673 20010518

Also published as: WO0235799 (A2) WO0235799 (A3) CN1481635 (A) CN100525288 (C) AU2444802 (A)

Abstract not available for EP1368948 (A2)
Abstract of corresponding document: WO0235799 (A2)

Large payload files are selectively partitioned in blocks and theblocks distributed to a plurality of distribution stations at theedge of the network qualified to have the data. Each qualified station decides how much and what portion of the content to save locally, based on information such as network location and environment, usage, popularity, and other distribution criteria defined by the content provider. Different pieces of a large payload file may be



available from different nodes, however, when a user requests access to the large payload file, for example, through an application server, a virtual file control system creates an illusion that the entire file is present at the connected node. However, since only selective portions of the large payload file may actually be resident at that node's storage at the time of request, a cluster of distribution servers at the distribution station may download the non-resident portions of the file as the applicationserver is servicing the user. The download may be in parallel andusually from the least congested nodes. New nodes added to the network learn from other nodes in the network what content they should have and downlaod the required content, in a desired amount, onto their local storage devices from the nearest and least congested nodes without interrupting network operation. Each node manages its local storage and decides what content to prune based on information such as usage patterns.

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Veröffentlichungsnummer:
(11) Publication number: EP 1 368 948 A0
(11) Numéro de publication:

Internationale Anmeldung veröffentlicht durch die Weltorganisation für geistiges Eigentum unter der Nummer:

WO 02/035799 (art. 158 des EPÜ).

International application published by the World Intellectual Property Organisation under number:

WO 02/035799 (art. 158 of the EPC).

Demande internationale publiée par l'Organisation Mondiale de la Propriété sous le numéro:

WO 02/035799 (art. 158 de la CBE).

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 May 2002 (02.05.2002)

PCT

(10) International Publication Number
WO 02/035799 A3

(51) International Patent Classification[7]: H04L 29/06, G06F 17/30

(21) International Application Number: PCT/US01/32638

(22) International Filing Date: 26 October 2001 (26.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/266,286	26 October 2000 (26.10.2000)	US
09/681,644	15 May 2001 (15.05.2001)	US
09/681,665	17 May 2001 (17.05.2001)	US
09/681,668	18 May 2001 (18.05.2001)	US
09/681,669	18 May 2001 (18.05.2001)	US
09/681,671	18 May 2001 (18.05.2001)	US
09/681,672	18 May 2001 (18.05.2001)	US
09/681,673	18 May 2001 (18.05.2001)	US

(71) Applicant: PRISMEDIA NETWORKS, INC. [US/US]; 3080 North First Street, Second Floor, San Jose, CA 95134 (US).

(72) Inventors: SIM, Siew, Yong; 10435 Sterling Boulevard, Cupertino, CA 95014 (US). CHAN, Desmond, Cho-Hung; 55 Devonshire Avenue, Mountain View, CA 94043 (US).

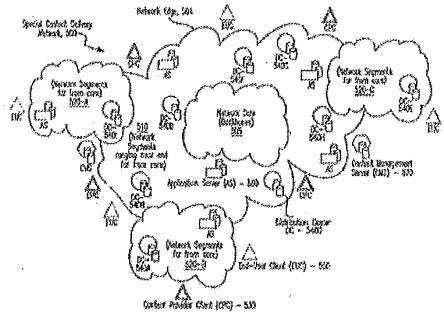
(74) Agents: RAY, Michael, B. et al.; Sterne, Kessler, Goldstein & Fox P.L.L.C., Suite 600, 1100 New York Avenue, Washington, DC 20005-3934 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR LARGE PAYLOAD DISTRIBUTION IN A NETWORK



(57) Abstract: Large payload files are selectively partitioned in blocks and theblocks distributed to a plurality of distribution stations at theedge of the network qualified to have the data. Each qualified station decides how much and what portion of the content to save locally, based on information such as network location and environment, usage, popularity, and other distribution criteria defined by the content provider. Different pieces of a large payload file may be available from different nodes, however, when a user requests access to the large payload file, for example, through an application server, a virtual file control system creates an illusion that the entire file is present at the connected node. However, since only selective portions of the large payload file may actually be resident at that node's storage at the time of request, a cluster of distribution servers at the distribution station may download the non-resident portions of the file as the applicationserver is servicing the user. The download may be in parallel andusually from the least congested nodes. New nodes added to the network learn from other nodes in the network what content they should have and downlaod the required content, in a desired amount, onto their local storage devices from the nearest and least congested nodes without interrupting network operation. Each node manages its local storage and decides what content to prune based on information such as usage patterns.

WO 02/035799 A3

Published:
— *with international search report*

(88) Date of publication of the international search report:
25 September 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/32638

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04L29/06 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 H04L G06F H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	GRUBER S ET AL: "Design Considerations for an RTSP-Based Prefix-Caching Proxy for Multimedia Streams" AT&T LABS RESEARCH, 7 September 1999 (1999-09-07), XP002194289 page 2, paragraphs 2,3 page 8, paragraph 3; figure 1 ---	1,9,46, 50,51
A	BABAK S NOGHAMI ET AL: "A Novel Approach to Reduce Latency on the Internet: Component-Based Download" DEPT. OF ELECTRICAL AND COMPUTER ENGINEERING, UNIVERSITY OF MANITOBA, June 2000 (2000-06), XP002204898 page 1, last paragraph page 2, paragraph 2 -page 3, paragraph 1 ---	1,9,46, 50,51

-/---

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

'A' document defining the general state of the art which is not considered to be of particular relevance
'E' earlier document but published on or after the international filing date
'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
'O' document referring to an oral disclosure, use, exhibition or other means
'P' document published prior to the international filing date but later than the priority date claimed

'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
'&' document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 October 2002	30. 01. 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Olmos, J

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 038 601 A (VAN DER RIJN DANIEL J G ET AL) 14 March 2000 (2000-03-14) column 4, line 13 -column 6, line 55 column 9, line 43 -column 9, line 52; figure 3	1,9,46, 50,51
A	US 5 751 968 A (COHEN ALON) 12 May 1998 (1998-05-12) column 2, line 6 -column 2, line 67 column 4, line 41 -column 4, line 51 column 5, line 5 -column 6, line 49; figures 1-7	1,9,46, 50,51
A	EP 0 967 559 A (LUCENT TECHNOLOGIES INC) 29 December 1999 (1999-12-29) column 2, line 28 -column 2, line 53 column 3, line 31 -column 4, line 5 column 4, line 55 -column 5, line 2 column 6, line 23 -column 7, line 4 column 8, line 22 -column 8, line 30; figures 1,2	1,9,46, 50,51
A	ALLEN MIU ET AL: "Performance Analysis of a Dynamic Parallel Downloading Scheme from Mirror Sites Throughout the Internet" 6892 TERM PAPER, XX, XX, December 1999 (1999-12), pages 1-13, XP002204897 abstract page 2, left-hand column, paragraph 2 page 3, left-hand column, paragraph 3 page 5, right-hand column, last paragraph page 11, right-hand column, last paragraph	1,9,46, 50,51
A	RODRIGUEZ P ET AL: "Parallel-access for mirror sites in the Internet" INFOCOM 2000. NINETEENTH ANNUAL JOINT CONFERENCE OF THE IEEE COMPUTER AND COMMUNICATIONS SOCIETIES. PROCEEDINGS. IEEE TEL AVIV, ISRAEL 26-30 MARCH 2000, PISCATAWAY, NJ, USA,IEEE, US, 26 March 2000 (2000-03-26), pages 864-873, XP010376176 ISBN: 0-7803-5880-5 page 866, left-hand column, paragraph 5 page 867, right-hand column, last paragraph -page 868, left-hand column, paragraph 5	1,9,46, 50,51
A	US 6 105 029 A (SHIEH JOHNNY MENG-HAN ET AL) 15 August 2000 (2000-08-15) column 5, line 16 -column 5, line 62; figures 2,3	1,9,46, 50,51

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/32638

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	US 6 081 840 A (ZHAO YAN) 27 June 2000 (2000-06-27) column 4, line 42 -column 4, line 59 column 9, line 55 -column 9, line 65; figures 1,7,8 ------	1,9,46, 50,51

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/32638

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-13,46-54

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (1)) (July 1998)

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-13,46-54

   Three methods and two apparatus of creating a virtual file, using said associations, for presentation to a client requesting said file content, said virtual file providing an illusion to said client that said file content is contiguous in said network node.

2. Claims: 14-27,34-45

   Three methods and two apparatus of sending a search request to determine which nodes in a network have content for a requesting node, said network having a plurality of nodes; receiving a response to said search request from each of one or more responding nodes having said content.

3. Claims: 28-33

   A method and an apparatus of distributing one or more tasks to one or more available servers in a network node, when one server becomes designated as an unavailable server.

4. Claims: 55-70

   Two methods and an apparatus of providing distribution criteria for a content; wherein each qualified node conforms to said distribution criteria.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/32638

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
US 6038601	A	14-03-2000	AU	8578898	A	10-02-1999
			EP	0996893	A1	03-05-2000
			WO	9904345	A1	28-01-1999
US 5751968	A	12-05-1998	IL	115263	A	11-04-1999
			AU	712229	B2	04-11-1999
			AU	6888796	A	17-04-1997
			CA	2230259	A1	03-04-1997
			EP	0850451	A2	01-07-1998
			WO	9712447	A2	03-04-1997
			JP	11512893	T	02-11-1999
			NZ	316616	A	28-02-2000
EP 0967559	A	29-12-1999	US	2001052015	A1	13-12-2001
			EP	0967559	A1	29-12-1999
			JP	2000078208	A	14-03-2000
US 6105029	A	15-08-2000	NONE			
US 6081840	A	27-06-2000	NONE			

Form PCT/ISA/210 (patent family annex) (July 1992)

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 1 OF 2	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	6,006,259	11-20-1998	Adelman et al.	
	2	6,067,609	05-23-2000	Meeker et al.	
	3	6,195,680	02-27-2001	Goldszmidt et al.	
	4	6,604,134	08-05-2003	Haury	
	5	7,031,944	04-18-2006	Tanioka	
	6	7,177,874	02-13-2007	Jardin	
	7	7,315,877	01-01-2008	Bhanot et al.	
	8	7,502,915	03-10-2009	Jacob et al.	
	9	7,519,652	04-14-2009	Page et al.	
	10	7,533,141	05-12-2009	Nadgir et al.	
	11	7,533,389	05-12-2009	Verbeke et al.	
	12	7,568,131	07-28-2009	Vertes	
	13	7,644,130	01-05-2010	Magro et al.	
	14	7,698,390	04-13-2010	Harkness et al.	
	15	7,757,236	07-13-2010	Singh	
	16	7,788,314	08-31-2010	Holt	
	17	7,835,022	11-16-2010	Matsumoto	
	18	8,117,288	02-14-2012	Bhanot et al.	
	19	9,405,564	08-02-2016	Muellers et al.	
	20	2003/0005068	01-02-2003	Nickel et al.	
	21	2004/0098447	05-20-2004	Verbeke et al.	
	22	2004/0254984	12-16-2004	Dinker	
	23	2005/0097300	05-05-2005	Gildea et al.	
	24	2005/0015460	01-20-2005	Goyle et al.	
	25	2005/0154789	07-14-2005	Fellenstein et al.	
	26	2005/0188088	08-25-2005	Fellenstein et al.	
	27	2006/0026601	02-02-2006	Solt	
	28	2006/0059473	03-16-2006	Moler	
	29	2007/0124363	05-31-2007	Lurie et al.	

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 2 OF 2	Attorney Docket No.	ZTANN.002P1C3

FOREIGN PATENT DOCUMENTS						
Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	30	EP 1368948	12-10-2003	Chan et al.		
	31	WO 2004/086270	10-07-2004	Katz et al.		

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	32	U.S. Provisional Application No. 60/799,474, filed May 10, 2006, titled "Graphical Interface for Monitoring Status of a Concurrent Computing Process".	

25376674

Examiner Signature	Date Considered
***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Hermon Asres
Art Unit	:	2449
Conf. No.	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Pursuant to 37 CFR 1.56, an Information Disclosure Statement listing references is provided herewith. Copies of any listed foreign and non-patent literature references are being submitted.

No Disclaimers

To the extent that anything in the Information Disclosure Statement or the listed references could be construed as a disclaimer of any subject matter supported by the present application, Applicant hereby rescinds and retracts such disclaimer.

Timing of Disclosure

This Information Disclosure Statement is being filed with an RCE and no fee is believed to be required.

The Commissioner is hereby authorized to charge any additional fees which may be required or to credit any overpayment to Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: August 8, 2017

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

25376708

Electronic Patent Application Fee Transmittal

Application Number:	14181112
Filing Date:	14-Feb-2014
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Filer:	Lance Dean Smemoe/Debora Swies
Attorney Docket Number:	ZTANN.002P1C3

Filed as Small Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
CLAIMS IN EXCESS OF 20	2202	10	40	400
INDEPENDENT CLAIMS IN EXCESS OF 3	2201	1	210	210
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 2 months with $0 paid	2252	1	300	300
Miscellaneous:				
RCE- 1st Request	2801	1	600	600
		Total in USD ($)		**1510**

Electronic Acknowledgement Receipt

EFS ID:	30020902
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Chelsea Burdeno
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	08-AUG-2017
Filing Date:	14-FEB-2014
Time Stamp:	18:25:34
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$1510
RAM confirmation Number	080917INTEFSW18261000
Deposit Account	111410
Authorized User	Chelsea Burdeno

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.16 (National application filing, search, and examination fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Request for Continued Examination (RCE)	ZTANN_002P1C3_RCE.pdf	697596 4e90686d38b809bddc9e690d00f32bd6870da7b3	no	3
Warnings:					
Information:					
2		ZTANN_002P1C3_OAResponse.pdf	153864 aa1fe871706d2b6f91d86298f6c9bbe52df831d7	yes	12

Multipart Description/PDF files in .zip description

Document Description	Start	End
Applicant Arguments/Remarks Made in an Amendment	10	12
Applicant summary of interview with examiner	9	9
Claims	2	8
Amendment Submitted/Entered with Filing of CPA/RCE	1	1

Warnings:					
Information:					
3	Foreign Reference	WO2004086270.pdf	1325089 7016822d73c8a3cfbe97ff33074272e98553d384	no	30
Warnings:					
Information:					
4	Non Patent Literature	US_60_799474.pdf	1448836 35c6f8e4cd8a504d4534a411eb29d9c3cde03d84	no	45
Warnings:					
Information:					

5	Foreign Reference	EP_1368948.pdf	688020 3710a958834bf3e54570c85a79c1dc7590c94e41	no	11
Warnings:					
Information:					
6		ZTANN_002P1C3__IDS.pdf	262882 10a7a92338864501a415f6c59537ae5c8ca540cd	yes	4
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Information Disclosure Statement (IDS) Form (SB08)		3	4	
	Transmittal Letter		1	2	
Warnings:					
Information:					
7	Fee Worksheet (SB06)	fee-info.pdf	35239 0838db6950dfc23a7b163119f5c044fd005b362a	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	4611526		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3	6874

20995 7590 03/09/2017
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

EXAMINER
ASRES, HERMON

ART UNIT	PAPER NUMBER
2449	

NOTIFICATION DATE	DELIVERY MODE
03/09/2017	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jayna.cartee@knobbe.com
efiling@knobbe.com

PTOL-90A (Rev. 04/07)

Office Action Summary	Application No. 14/181,112	Applicant(s) TANNENBAUM ET AL.	
	Examiner HERMON ASRES	Art Unit 2449	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *11/21/2016*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims*

5)☒ Claim(s) *2-31* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *2-31* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date _____.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

DETAILED ACTION

The following is a Final office action in response to the Amendments filed on November 21, 2016.

Claims 2, 3, 6, 7, and 10 have been amended.

Claims 12-31 have been added.

Claims 2-31 are pending.

Response to Arguments

Double Patenting Rejection

The double patenting rejection will not be withdrawn because applicant in remarks indicated that the double patenting rejection be held in abeyance until the other rejections are overcome.

35 U.S.C. 112 Rejections

Examiner withdraws the 112 rejection of claims 2, 6, and 7 in view of applicant's amendment/remarks filed on 11/21/2016.

35 U.S.C. 101 Rejections

Examiner withdraws the 101 rejection of claims 2-11 (previously rejected for being directed to a judicial exception without significantly more) in view of applicant's amendment/remarks filed on 11/21/2016.

35 U.S.C. 103 Rejections

Applicant's argument filed 11/21/2016 has been fully considered but they are deemed not persuasive in view of new grounds of rejection.

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the claims at issue are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*,

686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the reference application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO internet Web site contains terminal disclaimer forms which may be used. Please visit http://www.uspto.gov/forms/. The filing date of the application will determine what form should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to http://www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

Claim 2 is rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1 and 8 of U.S. Patent No. 8,676,877. Although the conflicting claims are not identical, they are not patentably distinct from each other because:

Instant Application		US Patent 8,676,877	
Claim	**Limitation**	**Claim**	**Limitation**
2, 28, and 31	**A computer cluster comprising:**	**1 and 8**	**A system for performing an instruction received from a front end by executing commands on one or more special purpose microprocessors, the system comprising:**
	a plurality of nodes, comprising a hardware processor wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and		**a plurality of nodes, wherein each node is configured to access a computer-readable memory system comprising program code for a single-node kernel module, ……..**
	wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel that when executed causes the hardware processor to interpret user instructions; and a mechanism for the nodes to communicate with each other using a peer-to-peer architecture; wherein at least one of the nodes returns a result to a user interface or a script		**wherein each cluster node module is stored in a computer-readable memory system and configured to communicate with a single-node kernel and with one or more other cluster node modules, to accept instructions, and to interpret at least some of the instructions such that the plurality of cluster node modules communicate with one another in order to act as a cluster in executing commands using one or more hardware processors**

Claim Rejections - 35 USC § 103

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 2, and 4-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Block et al. (U.S. PGPub 2005/0021751), in view of Singh (US Patent 8,601,101).

As per claim 2, Block teaches a computer cluster (Block, ….cluster data port services within a cluster infrastructure. See paragraph [0025]) comprising:

a plurality of nodes, comprising a hardware processor wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]) that when executed causes the hardware processor to interpret user instructions; and a mechanism for the nodes to communicate with each other. (Block, establishment of multiple network connections between a source node, a target node and one or more

backup nodes in such a manner that a cluster data port is effectively utilized as single data port. See paragraph [0025]).

Block doesn't explicitly teach that the nodes communicate with each other using a peer-to-peer architecture and also that one of the nodes returns a result to a user interface or a script. In analogous art Singh teaches nodes that communicate with each other using a peer-to-peer architecture. (Singh, …….each node in a cluster including a peer-to-peer communication channel compiling and maintaining its own cluster membership are disclosed……. each cluster node may be coupled to every other node and commonly-accessible storage through a network. …..Based on the cluster membership information, the joining node may request a peer-to-peer connection with each potential cluster member. See Column 2 line 26-38). Singh also teaches one of the nodes returns a result to a user interface or a script (Singh, The cluster configuration service 1510A may include a user interface to allow a system administrator to perform various node management functions through administrative console 1500. When the administrator makes changes to the node 240A configuration, the cluster configuration service 1510A may record all of the modifications that are made. See Column 14 line 61-66).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to take the teaching of Singh and apply them on the teaching of Block because when each node on the cluster communication channel is in peer-to-peer connection, each node can detect the connection failure based on the socket communication with the failed node and can update its cluster membership data without

the need for a central entity to pass an updated membership list among nodes. (Singh, See column 13 line 23-28).

As per claim 4, Block-Singh teaches wherein the mechanism comprises a message passing interface. (Block, a single data port instance may also support target to source message sends. See paragraph [0050]).

As per claim 5, Block-Singh teaches wherein each of the nodes comprises one or more cluster node modules. (Block, cluster data port services within a cluster infrastructure to provide reliable and efficient communications between nodes. See paragraph [0010]).

As per claim 6, Block-Singh teaches wherein each single-node kernel is stored in a non-transitory computer-readable medium and configured to accept and execute a request. (Block, The routines executed to implement ……operating system or a specific application, component, program, object, module or sequence of instructions, will also be referred to herein as "computer program code," or simply "program code." See paragraph [0039], [0049]).

As per claim 7, Block-Singh teaches wherein each of the nodes comprises one or more cluster node modules, wherein each of the cluster node modules comprises instructions stored in a non-transitory computer- readable medium and wherein the

instructions when executed by the hardware processor cause the cluster node module to communicate with the single-node kernel and with one or more other cluster node modules. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044], [0039]).

As per claim 8, Block-Singh teaches wherein the plurality of cluster node modules act as a cluster. (Block, utilize cluster data port services within a cluster infrastructure to provide reliable and efficient communications between nodes in a clustered computer system. See paragraph [0010]).

As per claim 9, Block-Singh teaches wherein the plurality of cluster node modules communicate with one another to act as a cluster. (Block, Nodes 12, 14 and 16 are typically coupled together via a clustering network 18. See paragraph [0028]).

As per claim 10, Block-Singh teaches wherein the computer cluster includes the user interface. (Block, Cluster manager application 60 that provides the user interface whereby a user such as a systems administrator can manage clustering operations in the system. See paragraph [0037]).

As per claim 11, Block-Singh teaches wherein each cluster node module accepts instructions from the user interface and interprets one or more of the instructions. (Block, Cluster manager application 60 that provides the user interface whereby a user such as a systems administrator can manage clustering operations in the system. See paragraph [0037], [0039]).

As per claim 12, Block-Singh teaches wherein the cluster initialization process comprises establishing communication among two or more of the nodes. (Bloack, communications between nodes in a clustered computer system. See paragraph [0010]).

As per claim 13, Block-Singh teaches wherein the cluster initialization process comprises configuring access by one or more of the nodes to a computer-readable medium comprising program code for the single-node kernel. (Block, The cluster data port services described hereinafter, on the other hand, provide an abstracted transport service that encapsulates and manages the establishment of multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or kernel component) like a single data port or data pipe. See paragraph [0044]).

As per claim 14, Block-Singh teaches wherein the cluster initialization process comprises establishing message-passing support among the nodes in the cluster.

(Block, data port services may also be configured to provide synchronous and asynchronous caller send models, as well as support message encryption. As such, the herein-described services may be used to provide a general messaging service that allows a variety of operating system or kernel components to make use of the services in a clustered environment. See paragraph [0045]).

As per claim 15, Block-Singh teaches wherein the cluster initialization process comprises launching cluster node modules by the cluster. (Block, see paragraph [0011], [0012]).

As per claim 16, Block-Singh teaches wherein the cluster initialization process comprises assigning a processor identification number to each of the cluster node modules. (Block, Data port internal object 72 also has a data port address table 78, which stores remote IP address, node ID pairs relating to the available IP addresses through which a particular node (identified by the node ID) is coupled to the clustering network. See paragraph [0059]).

As per claim 17, Block-Singh teaches wherein the cluster initialization process comprises: launching cluster node modules by the cluster; and after launching the cluster node modules, configuring access by one or more of the nodes to a non-transitory computer-readable medium comprising program code for the single-node kernel. (Block, data port services may also be configured to provide synchronous and

asynchronous caller send models, as well as support message encryption. As such, the herein-described services may be used to provide a general messaging service that allows a variety of operating system or kernel components to make use of the services in a clustered environment. See paragraph [0045]).

As per claim 18, Block-Singh teaches wherein the cluster initialization process comprises: launching cluster node modules by the cluster; after launching the cluster node modules, establishing communication among two or more of the nodes; (Block, a cluster data port consistent with the invention may be used to facilitate the transfer of large volumes of data between a source node and specified target nodes in a clustering environment. See paragraph [0012]) and after establishing communication among two or more of the nodes, assigning a processor identification number to each of the cluster node modules. (Block, Data port internal object 72 also has a data port address table 78, which stores remote IP address, node ID pairs relating to the available IP addresses through which a particular node (identified by the node ID) is coupled to the clustering network. See paragraph [0059]).

As per claim 19, Block-Singh teaches wherein one or more of the nodes are configured to communicate at least some of the user instructions. (Block, establishment of multiple network connections between a source node, a target node and one or more backup nodes in such a manner that a cluster data port is effectively utilized as single data port. See paragraph [0025], [0044]).

As per claim 20, Block-Singh teaches wherein one or more of the nodes are configured to communicate at least some of the user instructions to one or more single-node kernel. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]).

As per claim 21, Block-Singh teaches wherein one or more of the nodes are configured to accept user instructions via one or more of the nodes. (Block, see paragraph [0025], [0044]).

As per claim 22, Block-Singh teaches wherein one or more of the nodes are configured to communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other. (Block, see paragraph [0025], [0044]).

As per claim 23, Block-Singh teaches wherein one or more of the nodes are configured to transmit at least some of the user instructions that originate from a user interface. (Singh, The cluster configuration service 1510A may include a user interface to allow a system administrator to perform various node management functions through administrative console 1500. When the administrator makes changes to the node 240A configuration, the cluster configuration service 1510A may record all of the modifications that are made. See Column 14 line 61-66).

As per claim 24, Block-Singh teaches wherein one or more of the nodes are configured to parallelize at least some of the user instructions before communicating at least some of the user instructions to one or more single-node kernels. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]).

As per claim 25, Block-Singh teaches wherein one or more of the nodes are configured to accept user instructions before communicating at least some of the user instructions to one or more single-node kernels. (Singh, The cluster configuration service 1510A may include a user interface to allow a system administrator to perform various node management functions through administrative console 1500. When the administrator makes changes to the node 240A configuration, the cluster configuration service 1510A may record all of the modifications that are made. See Column 14 line 61-66).

As per claim 26, Block-Singh teaches wherein one or more of the nodes are configured to: accept user instructions; after accepting user instructions, communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other; (Block, see paragraph [0025], [0044]). and after communicating at least some of the user instructions using the mechanism, communicate at least some of the user instructions to one or more single-node kernels.

(Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]). (Also see Singh Column 14 line 61-66).

As per claim 27, Block-Singh teaches wherein the plurality of nodes are configured to communicate with one another to interpret and translate commands for execution by a plurality of single-node kernels. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]). (Also see Singh Column 14 line 61-66).

As per claim 28, Block teaches a computer cluster comprising: a plurality of nodes, (Block, ….cluster data port services within a cluster infrastructure. See paragraph [0025]) wherein one or more of the nodes are configured to receive:

a command to start a cluster initialization process for the computer cluster, wherein the cluster initialization process comprises establishing communication among two or more of the nodes; (Block, establishment of multiple network connections between a source node, a target node and one or more backup nodes in such a manner that a cluster data port is effectively utilized as single data port. See paragraph [0025]).

and a mechanism for the nodes to communicate with each other using asynchronous calls; wherein each of the nodes is configured to access a non-transitory

computer- readable medium comprising program code for a single-node kernel configured to interpret user instructions; and wherein at least one of the nodes returns a result to the user interface or the script. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]).

Block doesn't explicitly teach receiving an instruction from a user interface or a script. In analogous art Singh teaches receiving instruction from a user interface in a computer cluster. (Singh, The cluster configuration service 1510A may include a user interface to allow a system administrator to perform various node management functions through administrative console 1500. When the administrator makes changes to the node 240A configuration, the cluster configuration service 1510A may record all of the modifications that are made. See Column 14 line 61-66).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to take the teaching of Singh and apply them on the teaching of Block because it would allow a system administrator to perform various node management functions through administrative console. (Singh, … allow a system administrator to perform various node management functions through administrative console 1500. When the administrator makes changes to the node 240A configuration, the cluster configuration service 1510A may record all of the modifications that are made. See column 14 line 61-66).

As per claim 29, Block-Singh teaches wherein the asynchronous calls comprise a first command to create a first packet containing: an expression to be sent as payload; and a target node where the expression should be sent; (Block, provide synchronous and asynchronous caller send models, as well as support message encryption. As such, the herein-described services may be used to provide a general messaging service that allows a variety of operating system or kernel components to make use of the services in a clustered environment. See paragraph [0045]) wherein the first command is configured to be called from within a single-node kernel; wherein the single-node kernel is configured to send the first packet to a local cluster node module; and wherein the local cluster node module is configured to forward the expression to the target node. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]). (Also see Singh Column 14 line 61-66).

As per claim 30,

[Rejection rational for claim 29 is applicable].

As per claim 31,

[Rejection rational for claim 28 is applicable].

Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Block et al. (US PGPub 2005/0021751) in view of Singh (US Patent 8,601,101) and further in view of Gray (USPGPub 2007/0073705).

As per claim 3, Block-Singh teaches the computer cluster of claim 2 but doesn't teach a Mathematica kernel. In analogous art Gray teaches a Mathematica kernel. (Gray, The front end may include an interactive document referred to as a notebook similar to those often used with MATHEMATICA.RTM. software systems. A notebook may include input to be sent to the kernel 104 and output received from the kernel, as well as text, graphics, palettes, etc. see paragraph [0031]).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to take the teaching of Gray and apply them on the teaching of Block because the kernel and the front end may be implemented on a same computing system or on different computing systems that are communicatively coupled to one another. (Gray, see paragraph [0030]).

Conclusion

THIS ACTION IS MADE FINAL. Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to HERMON ASRES whose telephone number is (571)272-4257. The examiner can normally be reached on Monday - Friday 7:30 am to 5:00 pm est.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, VIVEK SRIVASTAVA can be reached on 571-272-7304. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/HERMON ASRES/
Examiner, Art Unit 2449

/VIVEK SRIVASTAVA/
Supervisory Patent Examiner, Art Unit 2449

Notice of References Cited	Application/Control No. 14/181,112	Applicant(s)/Patent Under Reexamination TANNENBAUM ET AL.	
	Examiner HERMON ASRES	Art Unit 2449	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-2005/0021751 A1	01-2005	Block, Timothy Roy	G06F9/54	709/225
*	B	US-2007/0073705 A1	03-2007	Gray; Theodore W.	G06F9/4443	1/1
*	C	US-8,601,101 B1	12-2013	Singh; Servesh Pratap	H04L69/40	370/254
*	D	US-6,968,359 B1	11-2005	Miller; Robert	G06F9/5061	709/205
*	E	US-2007/0094532 A1	04-2007	SENGUPTA; Sabyasachi	G06F11/3632	714/5.1
	F	US-				
	G	US-				
	H	US-				
	I	US-				
	J	US-				
	K	US-				
	L	US-				
	M	US-				

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	7,716,323	05-11-2010	Gole et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	2	HAYNOS, MATT, "Perspectives on grid: Grid computing--next-generation distributed computing", January 27, 2004, IBM Developer Works, http://www-106.ibm.com/developerworks/library/gr-heritage	
	3	"Pingpong MPI Benchmark - SEM vs 'grid'," January 2009, Dauger Research, Inc., http://daugerresearch.com/pooch/mathmatica/	
	4	Wolfram Research, "Mathematica Parallel Computing Toolkit", January 2005, pp. 1-95.	

24039252

Examiner Signature /HERMON ASRES/	Date Considered 03/01/2017

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A/

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

CPC- SEARCHED

Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED

Symbol	Date	Examiner

US CLASSIFICATION SEARCHED

Class	Subclass	Date	Examiner
709	223	07/07/2016	HA

SEARCH NOTES

Search Notes	Date	Examiner
Searched East (US Patents, USPGPubs, EPO, JPO, IBM, Derwent)	07/07/2016	HA
See East Search attached	07/07/2016	HA
Inventor search performed	07/07/2016	HA
Consulted Justin Rider (QAS) regarding Alice 101 rejection	07/08/2016	HA
Searched updated	03/01/2017	HA

INTERFERENCE SEARCH

US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S1	6	((Zvi) near2 (Tannenbaum)).INV.	USPAT	OR	OFF	2015/07/24 20:25
S2	7	((Dean) near2 (Dauger)).INV.	USPAT	OR	OFF	2015/07/24 20:25
S3	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S4	24	S3 and interpret	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S5	28950	709/223.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S6	5291	H04L41/0213.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S7	1692	H04L29/08072.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S8	27776	H04L29/06.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S9	6793	H04L41/22.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S10	8156	H04L41/12.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S11	0	S3 and S6	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28

S12	0	S3 and S7	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S13	15	S3 and S8	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S14	0	S3 and S9	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S15	2	S3 and S10	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S16	99	("20020049859" | "20030005266" | "20030051062" | "20030135621" | "20030195931" | "20030195938" | "20040110209" | "20040157203" | "20040252710" | "20050021751" | "20050038852" | "20050060237" | "20050076105" | "20050108394" | "20050180095" | "20060053216" | "20060106931" | "20070044099" | "20070073705" | "20080250347" | "20080281997" | "5423046" | "5881315" | "5930768" | "6108699" | "6202080" | "6546403" | "6578068" | "6751698" | "6782537" | "6968335" | "7093004" | "7136924" | "7174381" | "7249357" | "7334232" | "7437460" | "7472193" | "7554949" | "7937455" | "8082289" | "8140612" | "8402083" | "8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 12:54
S17	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:19
S18	10	(computer with cluster) same nodes same kernel and mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:20
S19	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:23
S20	10	("8402080").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/11/29 23:38
S21	2	("8849889").PN.	US-PGPUB; USPAT;	OR	OFF	2015/11/29 23:38

			USOCR; EPO; JPO; DERWENT; IBM_TDB			
S22	92	US-1234567-$.DID. OR US-5423046-$.DID. OR US-5881315-$.DID. OR US-5930768-$.DID. OR US-6108699-$.DID. OR US-6202080-$.DID. OR US-6546403-$.DID. OR US-6578068-$.DID. OR US-6751698-$.DID. OR US-6782537-$.DID. OR US-6968335-$.DID. OR US-7093004-$.DID. OR US-7136924-$.DID. OR US-7174381-$.DID. OR US-7249357-$.DID. OR US-7334232-$.DID. OR US-7437460-$.DID. OR US-7472193-$.DID. OR US-7554949-$.DID. OR US-7937455-$.DID. OR US-8082289-$.DID. OR US-8140612-$.DID. OR US-8402083-$.DID. OR US-8676877-$.DID. OR US-20020049859-$.DID. OR US-20030005266-$.DID. OR US-20030051062-$.DID. OR US-20030135621-$.DID. OR US-20030195931-$.DID. OR US-20030195938-$.DID. OR US-20040110209-$.DID. OR US-20040157203-$.DID. OR US-20040252710-$.DID. OR US-20050021751-$.DID. OR US-20050038852-$.DID. OR US-20050060237-$.DID. OR US-20050076105-$.DID. OR US-20050108394-$.DID. OR US-20050180095-$.DID. OR US-20060053216-$.DID. OR US-20060106931-$.DID. OR US-20070044099-$.DID. OR US-20070073705-$.DID. OR US-20080250347-$.DID. OR US-20080281997-$.DID. OR JP-1234567-$.DID. OR JP-0887473-$.DID. OR JP-2002117010-$.DID. OR US-6131999-$.DID.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/11/29 23:39
S23	4	(computer with cluster) same (nodes) same (single-node with kernel)	USPAT	OR	OFF	2016/06/17 19:19
S24	10	(computer with cluster) same (nodes) same (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S25	11	(computer with cluster) same (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S26	11	(computer with cluster) and (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S27	224	(computer with cluster) and (nodes) and (single with kernel)	US-PGPUB; USPAT; EPO; JPO;	OR	OFF	2016/06/17 19:20

			DERWENT; IBM_TDB			
S28	145	(computer with cluster) same (nodes) and (single with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:21
S29	6	(computer with cluster) same (nodes) and (single with kernel) and 709/223.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:22
S30	10	(computer with cluster) same (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 16:08
S31	79	(computer with cluster) and (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 18:41
S32	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 12:43
S33	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:47
S34	1505	(computer with cluster) and kernel and nodes and request	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S35	271	S34 and mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S36	178750	S34 and mathematical smae kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S37	10	S34 and mathematical same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S38	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:54
S39	29	(single with node with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO;	OR	OFF	2016/07/08 16:10

			DERWENT; IBM_TDB			
S40	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 16:11
S41	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:18
S42	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:24
S43	180800	mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S44	11	mathematica near5 kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S45	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:40
S46	10	("8402080").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:43
S47	2	("8849889").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44
S48	101	("20020049859" | "20030005266" | "20030051062" | "20030135621" | "20030195931" | "20030195938" | "20040110209" | "20040157203" | "20040252710" | "20050021751" | "20050038852" | "20050060237" | "20050076105" | "20050108394" | "20050180095" | "20060053216" | "20060106931" | "20070044099" | "20070073705" | "20080250347" | "20080281997" | "5423046" | "5881315" | "5930768" | "6108699" | "6202080" | "6546403" | "6578068" | "6751698" | "6782537" | "6968335" | "7093004" | "7136924" | "7174381" | "7249357" | "7334232" | "7437460" | "7472193" | "7554949" | "7937455" | "8082289" | "8140612" | "8402083" | "8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44

S49	17	mathematica with cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:50
S50	43	mathematica same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:51
S51	537678	stephen wolfram mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S52	50	(stephen near2 wolfram).in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S53	1	S52 and mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S54	215	mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S55	74	S54 and cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S56	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:44
S57	0	("gridMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S58	0	("gridwithMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S59	19	grid with Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S60	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16

S61	14	gridMathematica and wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:21
S62	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S63	2578	Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S64	215	Mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S65	15749	wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S66	8450	wolfram.in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S67	15	S66 and mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S68	42	S66 and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S69	2	("20070073705").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:44
S70	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:46
S71	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/11/07 13:23
S72	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:24

S73	0	(" (computerwithnodeswithclister)same(peer-to-peer)and@ad<20060101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:47
S74	0	(computer with nodes with clister) same (peer-to-peer) and @ad<"20060101"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:47
S75	3	(computer with nodes with cluster) same (peer-to-peer) and @ad<"20060101"	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/01 19:48
S76	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/02 17:12
S77	12	("8601101").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2017/03/02 17:12

EAST Search History (Interference)

<This search history is empty>

3/2/2017 9:23:40 PM
C:\Users\hasres\Documents\EAST\Workspaces\14181112.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	07/07/2016	03/01/2017							
	1	-	-							
	2	✓	✓							
	3	✓	✓							
	4	✓	✓							
	5	✓	✓							
	6	✓	✓							
	7	✓	✓							
	8	✓	✓							
	9	✓	✓							
	10	✓	✓							
	11	✓	✓							
	12		✓							
	13		✓							
	14		✓							
	15		✓							
	16		✓							
	17		✓							
	18		✓							
	19		✓							
	20		✓							
	21		✓							
	22		✓							
	23		✓							
	24		✓							
	25		✓							
	26		✓							
	27		✓							
	28		✓							
	29		✓							
	30		✓							
	31		✓							

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Asres, Hermon
Art Unit	:	2449
Conf. No.	:	6874

RESPONSE TO OFFICE ACTION DATED JULY 20, 2016

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

In response to the Office Action dated July 20, 2016, please reconsider the rejections in view of the following remarks.

Amendments to the Claims are reflected in the listing of claims which begins on page 2.

Summary of Interview begins on page 7.

Remarks/Arguments begin on page 8.

AMENDMENTS TO THE CLAIMS

1. (Canceled)

2. (Currently amended) A computer cluster comprising:

a plurality of nodes comprising a hardware processor, wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel that, when executed, causes the hardware processor ~~module configured~~ to interpret user instructions; and

a mechanism for the nodes to communicate with each other using a peer-to-peer architecture;

wherein at least one of the nodes returns a result to a user interface or a script.

3. (Currently amended) The computer cluster of claim 2, wherein the single-node kernel ~~module~~ comprises a Mathematica kernel.

4. (Previously presented) The computer cluster of claim 2, wherein the mechanism comprises a message passing interface.

5. (Previously presented) The computer cluster of claim 2, wherein each of the nodes comprises one or more cluster node modules.

6. (Currently amended) The computer cluster of claim 2, wherein each single-node kernel ~~module~~ is stored in a non-transitory computer-readable medium and configured to accept and execute a request.

7. (Currently amended) The computer cluster of claim 6, wherein each of the nodes comprises one or more cluster node modules,[[and]] wherein each of the cluster node modules [[is]] comprises instructions stored in a non-transitory computer-readable medium, and ~~configured to~~ wherein the instructions, when executed by the hardware processor, cause the cluster node module to communicate with the single-node kernel and with one or more other cluster node modules.

8. (Previously presented) The computer cluster of claim 7, wherein the plurality of cluster node modules act as a cluster.

9. (Previously presented) The computer cluster of claim 7, wherein the plurality of cluster node modules communicate with one another to act as a cluster.

10. (Currently amended) The computer cluster of claim 9, wherein the computer cluster includes[[a]] the user interface.

11. (Previously presented) The computer cluster of claim 10, wherein each cluster node module accepts instructions from the user interface and interprets one or more of the instructions.

12. (New) The computer cluster of claim 2, wherein the cluster initialization process comprises establishing communication among two or more of the nodes.

13. (New) The computer cluster of claim 2, wherein the cluster initialization process comprises configuring access by one or more of the nodes to a computer-readable medium comprising program code for the single-node kernel.

14. (New) The computer cluster of claim 13, wherein the cluster initialization process comprises establishing message-passing support among the nodes in the cluster.

15. (New) The computer cluster of claim 13, wherein the cluster initialization process comprises launching cluster node modules by the cluster.

16. (New) The computer cluster of claim 15, wherein the cluster initialization process comprises assigning a processor identification number to each of the cluster node modules.

17. (New) The computer cluster of claim 2, wherein the cluster initialization process comprises:

launching cluster node modules by the cluster; and

after launching the cluster node modules, configuring access by one or more of the nodes to a non-transitory computer-readable medium comprising program code for the single-node kernel.

18. (New) The computer cluster of claim 2, wherein the cluster initialization process comprises:

launching cluster node modules by the cluster;

after launching the cluster node modules, establishing communication among two or more of the nodes; and

after establishing communication among two or more of the nodes, assigning a processor identification number to each of the cluster node modules.

19. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to communicate at least some of the user instructions.

20. (New) The computer cluster of claim 19, wherein one or more of the nodes are configured to communicate at least some of the user instructions to one or more single-node kernels.

21. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to accept user instructions via one or more of the nodes.

22. (New) The computer cluster of claim 21, wherein one or more of the nodes are configured to communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other.

23. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to transmit at least some of the user instructions that originate from a user interface.

24. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to parallelize at least some of the user instructions before communicating at least some of the user instructions to one or more single-node kernels.

25. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to accept user instructions before communicating at least some of the user instructions to one or more single-node kernels.

26. (New) The computer cluster of claim 2, wherein one or more of the nodes are configured to:

accept user instructions;

after accepting user instructions, communicate at least some of the user instructions using the mechanism for the nodes to communicate with each other; and

after communicating at least some of the user instructions using the mechanism, communicate at least some of the user instructions to one or more single-node kernels.

27. (New) The computer cluster of claim 2, wherein the plurality of nodes are configured to communicate with one another to interpret and translate commands for execution by a plurality of single-node kernels.

28. (New) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes are configured to receive:

a command to start a cluster initialization process for the computer cluster, wherein the cluster initialization process comprises establishing communication among two or more of the nodes; and

an instruction from a user interface or a script;

and

a mechanism for the nodes to communicate with each other using asynchronous calls;

wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel configured to interpret user instructions; and

wherein at least one of the nodes returns a result to the user interface or the script.

29. (New) The computer cluster of claim 28, wherein the asynchronous calls comprise a first command to create a first packet containing:

an expression to be sent as payload; and

a target node where the expression should be sent;

wherein the first command is configured to be called from within a single-node kernel;

wherein the single-node kernel is configured to send the first packet to a local cluster node module; and

wherein the local cluster node module is configured to forward the expression to the target node.

30. (New) The computer cluster of claim 29, wherein the asynchronous calls comprise a second command to create a second packet containing:

a location where the expression is expected to be received; and

a sending node from which the expression is expected to be received;

wherein the second command is configured to be called from within a single-node kernel;

wherein the single-node kernel is configured to send the second packet to a local cluster node module; and

wherein the local cluster node module is configured to store the second packet contents in a message receiving queue.

31. (New) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes are configured to receive:

a command to start a cluster initialization process for the computer cluster, wherein the cluster initialization process comprises establishing communication among two or more of the nodes; and

an instruction from a user interface or a script;

and

a mechanism for the nodes to communicate with each other;

wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel configured to interpret user instructions and interpret the user instructions into commands that are executable by a special purpose microprocessor; and

wherein at least one of the nodes returns a result to the user interface or the script.

SUMMARY OF INTERVIEW

Attendees, Date and Type of Interview

A telephonic interview was conducted on November 7, 2016 and attended by Examiner Hermon Asres and Applicant's representative, Lance D. Smemoe (reg. no. 66,152).

Exhibits and/or Demonstrations

None

Identification of Claims Discussed

Claims 2-11

Identification of Cited/Disclosed Art Discussed

None

Proposed Amendments

Substantially as set forth herein

Principal Arguments and Other Matters

Applicant's representative presented proposed amendments and argued that the rejections under 35 USC § 112 and 35 USC § 101 are inapplicable to the claims as amended. Specific arguments are set forth in the following remarks.

Results of Interview

Applicant agreed to submit a formal response with claim amendments. The Examiner will reconsider the rejections and confirm the allowability of the amended claims.

REMARKS

In the Office Action dated July 20, 2016, Claims 2-11 were rejected, as discussed below. In this Amendment, claims 2, 3, 6, 7, and 10 are amended, and new claims 12-31 are added. No new matter is added. After the amendments, claims 2-31 are pending for consideration. Claims 2, 28, and 31 are independent claims.

Telephone Interview

Applicant thanks the Examiner for the courteous and helpful telephone interview held on November 7, 2016 (summarized above).

Response to Double Patenting Rejection

Claim 2 is rejected on the ground of non-statutory (obviousness-type) double patenting over claims 1 and 8 of U.S. Patent No. 8,676,877. Applicant respectfully requests that the double patenting rejection be held in abeyance until the other rejections are overcome.

Response to Rejections under 35 U.S.C. § 112

Claims 2, 6, and 7 are rejected under 35 U.S.C. § 112 as allegedly being indefinite for failing to particularly point out and distinctly claim the subject matter. Applicant respectfully disagrees with the rejections because the specification discloses structure, material, or acts that correspond to the claimed subject matter. Nevertheless, to expedite allowance of the application, Applicant has amended claims 2, 6, and 7 to remove the language identified by the examiner as invoking 35 U.S.C. § 112, sixth paragraph. Accordingly, Applicant respectfully requests that the rejections be withdrawn.

Response to Rejections under 35 U.S.C. § 101

Claims 2-11 are rejected under 35 U.S.C. § 101 as allegedly being directed to non-statutory subject matter. Applicant respectfully traverses the rejection. The Supreme Court of the United States has stated that a claim that recites "an improvement to the functioning of the computer itself" is directed to significantly more than an abstract idea. *Alice Corp. v. CLS Bank*, 134 S. Ct. at 2359. Claim 2 as amended recites improvements that improve operation of the claimed computer cluster. For example, claim 2 recites that "the nodes . . . communicate with

each other using a peer-to-peer architecture." The peer-to-peer architecture improves operation of the computer because it makes intermediate results of computation available to other nodes, thus permitting some computations to be solved using multi-node processing that are not possible with other types of distributed computing. *See, e.g.,* paragraphs [0005] and [0062] of the application as filed. As another example, the peer-to-peer architecture can increase performance of the computer compared to other architectures. *See, e.g.,* "Pingpong MPI Benchmark – SEM vs 'grid'," cited as reference no. 3 in the Information Disclosure Statement filed herewith. This information demonstrates that claim 2 recites an improvement to the functioning of the computer itself and, thus, meets the requirements of subject-matter patent-eligibility under 35 U.S.C. § 101. Applicant respectfully requests withdrawal of the rejections.

Response to Rejections under 35 U.S.C. § 102 and 35 U.S.C. § 103

The Office Action rejects Claims 2, and 4-11 under 35 U.S.C. § 102, arguing that they are anticipated by U.S. Patent Publication No. 2005/0021751 to Block et al. (hereafter "Block"). Applicant respectfully traverses the rejections.

The Office Action rejects Claim 3 under 35 U.S.C. § 103(a), arguing that it is unpatentable in view of Block and U.S. Patent Publication No. 2007/0073705 to Gray. Applicant respectfully traverses the rejection.

Applicant respectfully submits that Block does not disclose every feature recited in claim 2 or the claims that depend from claim 2. For example, Block does not disclose at least a computer cluster comprising "a mechanism for the nodes to communicate with each other using a peer-to-peer architecture," in combination with the other features recited in claim 2. As another example, Block does not disclose at least that "at least one of the nodes returns a result to a user interface or a script," in combination with the other features recited in claim 2. Indeed, Block does not disclose that the disclosed target nodes and backup nodes share data with one another using a peer-to-peer architecture.

Each of dependent Claims 3-11 is patentable because each depends directly or indirectly from claim 2 and because of the additional features recited in each claim.

For at least these reasons, Applicant respectfully submits that Claims 2-11 are allowable over the cited references and respectfully requests that the rejections of these claims be withdrawn.

New claims 12-31

New claims 12-31 are added in this amendment. No new matter is added. Each of claims 12-27 depends directly or indirectly from claim 2 and is patentable for that reason and for the additional features recited in each claim. Each of claims 28-31 is patentable because the cited references do not anticipate or render obvious the unique combination of features recited in each of claims 28-31.

No Disclaimers or Disavowals

Applicant respectfully submits that the claims are in condition for allowance. Furthermore, any remarks in support of patentability of one claim should not be imputed to any other claim, even if similar terminology is used. Any remarks referring to only a portion of a claim should not be understood to base patentability on that portion or that the limitation discussed is essential or critical; rather, patentability must rest on each claim taken as a whole. Applicant respectfully traverses each of the Examiner's rejections and each of the Examiner's assertions regarding what the prior art shows or teaches, even if not expressly discussed herein. Although the present communication may include alterations to the application or claims, or characterizations of claim scope or referenced art, applicant is not conceding in this application that previously pending claims are not patentable over the cited references. Rather, any alterations or characterizations are being made to facilitate expeditious prosecution of this application. Applicant reserves the right to pursue at a later date any previously pending or other broader or narrower claims that capture any subject matter supported by the present disclosure, including subject matter found to be specifically disclaimed herein or by any prior prosecution. Accordingly, reviewers of this or any parent, child or related prosecution history shall not reasonably infer that applicant has made any disclaimers or disavowals of any subject matter supported by the present application

CONCLUSION

For the foregoing reasons, it is respectfully submitted that the rejections set forth in the outstanding Office Action are inapplicable to the present claims. Accordingly, issuance of a Notice of Allowance is requested.

The undersigned has made a good faith effort to respond to all of the rejections in the case and to place the claims in condition for immediate allowance. Nevertheless, if any undeveloped issues remain or if any issues require clarification, the Examiner is respectfully requested to call the undersigned below.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: November 21, 2016

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

24428740

INFORMATION DISCLOSURE STATEMENT

First Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Hermon Asres
Art Unit	:	2449
Conf. No.	:	6874

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Pursuant to 37 CFR 1.56, an Information Disclosure Statement listing references is provided herewith. Copies of any listed foreign and non-patent literature references are being submitted.

No Disclaimers

To the extent that anything in the Information Disclosure Statement or the listed references could be construed as a disclaimer of any subject matter supported by the present application, Applicant hereby rescinds and retracts such disclaimer.

Timing of Disclosure

This Information Disclosure Statement is being filed after receipt of a First Office Action, but before the mailing date of a Final Action and before the mailing date of a Notice of Allowance. This Statement is accompanied by the fees set forth in 37 CFR 1.17(p). The Commissioner is hereby authorized to charge any additional fees which may be required or to credit any overpayment to Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: November 21, 2016

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

24039383

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	February 14, 2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Hermon Asres
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	7,716,323	05-11-2010	Gole et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	2	HAYNOS, MATT, "Perspectives on grid: Grid computing--next-generation distributed computing", January 27, 2004, IBM Developer Works, http://www-106.ibm.com/developerworks/library/gr-heritage	
	3	"Pingpong MPI Benchmark - SEM vs 'grid'," January 2009, Dauger Research, Inc., http://daugerresearch.com/pooch/mathmatica/	
	4	Wolfram Research, "Mathematica Parallel Computing Toolkit", January 2005, pp. 1-95.	

24039252

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

Electronic Patent Application Fee Transmittal

Application Number:	14181112
Filing Date:	14-Feb-2014
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Filer:	Lance Dean Smemoe/Debora Swies
Attorney Docket Number:	ZTANN.002P1C3

Filed as Small Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
CLAIMS IN EXCESS OF 20	2202	10	40	400
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Extension-of-Time:				
Extension - 1 month with $0 paid	2251	1	100	100
Miscellaneous:				
SUBMISSION- INFORMATION DISCLOSURE STMT	2806	1	90	90
		Total in USD ($)		**590**

Electronic Acknowledgement Receipt

EFS ID:	27579539
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Christina Graul
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	21-NOV-2016
Filing Date:	14-FEB-2014
Time Stamp:	20:06:29
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	CARD
Payment was successfully received in RAM	$590
RAM confirmation Number	112216INTEFSW20081700
Deposit Account	111410
Authorized User	Christina Graul

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.16 (National application filing, search, and examination fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		ZTANN_002P1C3_Response.pdf	143658 65e3f9edcebc4fcff095592ca1a83c08efb7c1a9	yes	11

Multipart Description/PDF files in .zip description

Document Description	Start	End
Amendment/Req. Reconsideration-After Non-Final Reject	1	1
Claims	2	6
Applicant summary of interview with examiner	7	7
Applicant Arguments/Remarks Made in an Amendment	8	11

Warnings:

Information:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
2		ZTANN_002P1C3_IDS.pdf	212224 634f80b1c4201d81bc09686e16f96505eb23f72e	yes	3

Multipart Description/PDF files in .zip description

Document Description	Start	End
Transmittal Letter	1	2
Information Disclosure Statement (IDS) Form (SB08)	3	3

Warnings:

Information:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
3	Non Patent Literature	Wolfram_MathmaticsParallelComputingToolkit_2005.pdf	2731164 797b5edda1f45e6d82c146e910e78a76751f542a	no	95

Warnings:

Information:

4	Non Patent Literature	Haynos__PerspectivesOn-Grid.pdf	7077722 2cc5e8ea1e822fc63d7f69381e3e6408c8d5c3b7	no	7
Warnings:					
Information:					
5	Non Patent Literature	Pingpong_MPIBenchmark.PDF	1403241 040f73de0b6e8a5fba3797bba8ebdb942fef8306	no	1
Warnings:					
Information:					
6	Fee Worksheet (SB06)	fee-info.pdf	32049 b0cd6e3895ea6d301c0c9e8872be27bea06f3545	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	11600058		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD
Substitute for Form PTO-875

Application or Docket Number	Filing Date	
14/181,112	02/14/2014	☐ To be Mailed

ENTITY: ☐ LARGE ☒ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT

11/21/2016	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	* 30	Minus	** 20	= 10	X $40 =	400
Independent (37 CFR 1.16(h))	* 3	Minus	***3	= 0	X $210 =	0
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					TOTAL ADD'L FEE	**400**

AMENDMENT

	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
☐ Application Size Fee (37 CFR 1.16(s))						
☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
					TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

LIE
ANTHONY TYSON

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3	6874

20995 7590 11/16/2016
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

EXAMINER
ASRES, HERMON

ART UNIT	PAPER NUMBER
2449	

NOTIFICATION DATE	DELIVERY MODE
11/16/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jayna.cartee@knobbe.com
efiling@knobbe.com

PTOL-90A (Rev. 04/07)

Applicant-Initiated Interview Summary	Application No.	Applicant(s)	
	14/181,112	TANNENBAUM ET AL.	
	Examiner	Art Unit	
	HERMON ASRES	2449	

All participants (applicant, applicant's representative, PTO personnel):

(1) *HERMON ASRES*.

(2) _____.

(3) *Lance Smemoe (REG NO 66,152)*.

(4) _____.

Date of Interview: *07 November 2016*.

Type: ☒ Telephonic ☐ Video Conference
☐ Personal [copy given to: ☐ applicant ☐ applicant's representative]

Exhibit shown or demonstration conducted: ☐ Yes ☒ No.
If Yes, brief description: _____.

Issues Discussed ☒101 ☒112 ☒102 ☐103 ☐Others
(For each of the checked box(es) above, please describe below the issue and detailed description of the discussion)

Claim(s) discussed: *2*.

Identification of prior art discussed: *Block*.

Substance of Interview
(For each issue discussed, provide a detailed description and indicate if agreement was reached. Some topics may include: identification or clarification of a reference or a portion thereof, claim interpretation, proposed amendments, arguments of any applied references etc...)

Applciant discussed proposed amendment to independent claim 1 including adding "a mechanism for the nodes to communicate with each other using a peer-to-peer architecture". Applicant also added a non-transitory language to the independent claim to address the 101 rejection. Applicant further indicated that the "module configured to" will be rewritten to overcome the 112 rejection. Examiner reminded applicant about the pending double patenting rejection and applicant indicated that if the proposed amendment when filed doesn't overcome the double patenting rejection a terminal disclaimer will be filed. No agreement was reached.

Applicant recordation instructions: The formal written reply to the last Office action must include the substance of the interview. (See MPEP section 713.04). If a reply to the last Office action has already been filed, applicant is given a non-extendable period of the longer of one month or thirty days from this interview date, or the mailing date of this interview summary form, whichever is later, to file a statement of the substance of the interview

Examiner recordation instructions: Examiners must summarize the substance of any interview of record. A complete and proper recordation of the substance of an interview should include the items listed in MPEP 713.04 for complete and proper recordation including the identification of the general thrust of each argument or issue discussed, a general indication of any other pertinent matters discussed regarding patentability and the general results or outcome of the interview, to include an indication as to whether or not agreement was reached on the issues raised.

☒ Attachment

/HERMON ASRES/ Examiner, Art Unit 2449	/VIVEK SRIVASTAVA/ Supervisory Patent Examiner, Art Unit 2449

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)

In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case. It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

M E M O R A N D U M

TO: Examiner Hermon Asres
direct: 571-272-4257
email: hermon.asres@uspto.gov

FROM: Lance Smemoe, Reg. #66,152

RE: Agenda for telephone interview on November 7, 2016 at 2:00 p.m. Eastern time

Application No.: 14/181,112

Attorney Docket No.: ZTANN.002P1C3

INFORMAL DOCUMENT FOR DISCUSSION ONLY – DO NOT ENTER

A. Rejections under 35 USC § 112 of claims 2, 6, and 7

Proposed claim amendments

2. (Currently amended) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a non-transitory computer-readable medium comprising program code for a single-node kernel ~~module~~ configured to interpret user instructions; and

a mechanism for the nodes to communicate with each other using a peer-to-peer architecture;

wherein at least one of the nodes returns a result to a user interface module or a script.

6. (Currently amended) The computer cluster of claim 2, wherein each single-node kernel ~~module~~ is stored in a non-transitory computer-readable medium and configured to accept and execute a request.

7. (Currently amended) The computer cluster of claim 6, wherein each of the nodes comprises one or more cluster node modules, and each of the cluster node modules is stored in a non-transitory computer-readable medium and configured to communicate with the single-node kernel and with one or more other cluster node modules.

INFORMAL DOCUMENT FOR DISCUSSION ONLY – DO NOT ENTER

Non-limiting example embodiments from the specification

[0093] A kernel module 206 typically includes program code for interpreting high-level code, commands, and/or instructions supplied by a user or a script into low-level code, such as, for example, machine language or assembly language. In one embodiment, each cluster node module 204a-e is connected to all other cluster node modules, while each kernel module 206a-e is allocated and connected only to one cluster node module 204. In one embodiment, there is one cluster node module-kernel module pair per processor. For example, in an embodiment of a computer cluster 100 including single-processor computer systems, each cluster node module-kernel module pair could reside on a single-processor computer. If a computer contains multiple processors or processing cores, it may contain multiple cluster node module-kernel module pairs, but the pairs can still communicate over the cluster node module's network connections.

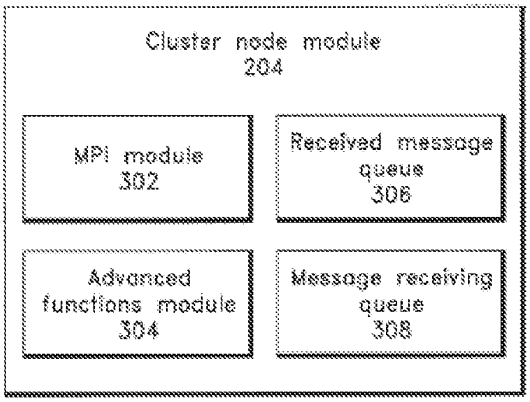


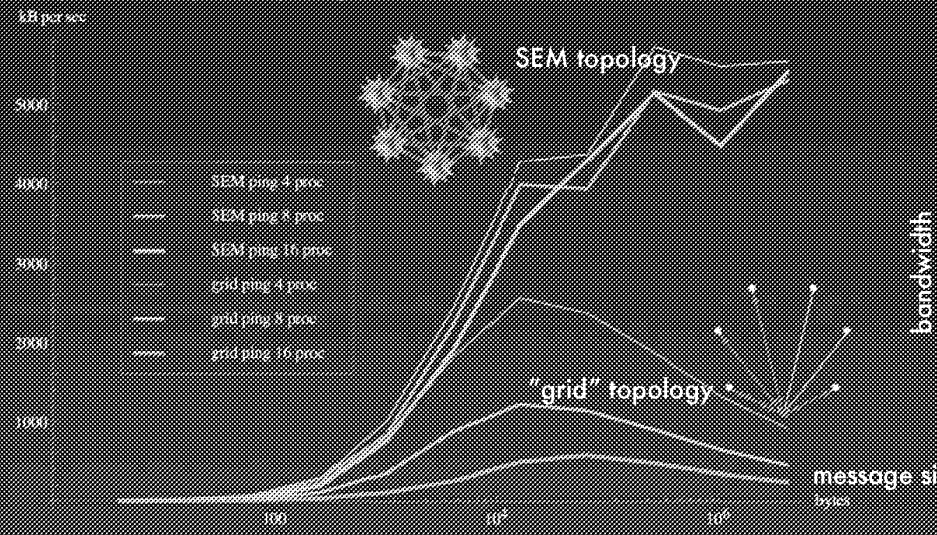
FIG. 3

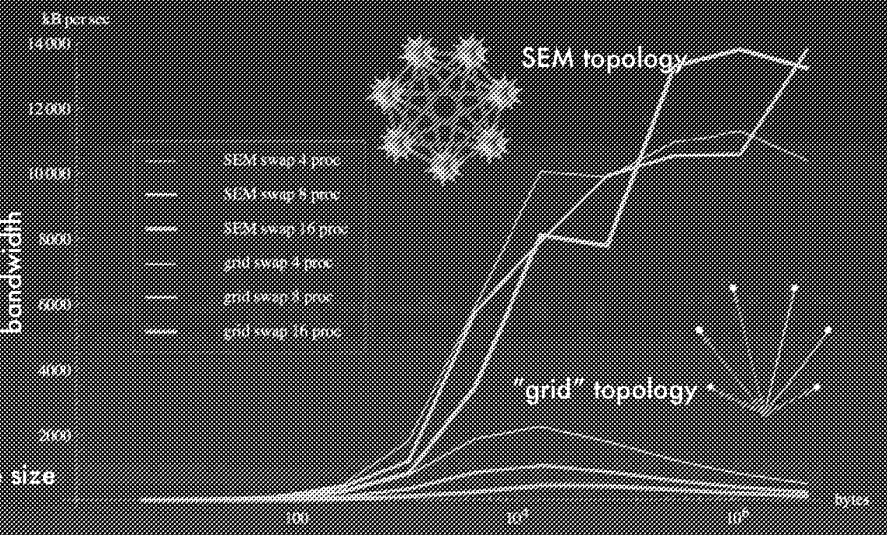
B. Rejections under 35 USC § 101 of Claims 2-11

Pingpong MPI Benchmark - SEM vs. "grid"

The Supercomputing Engine for Mathematica (SEM), like modern supercomputers, supports an "all-to-all" communications network between *Mathematica* kernels. Practitioners in high-performance computing know the communications network ultimately limits the addressable problem size. SEM provides supercomputing infrastructure for Wolfram's *Mathematica*.

We tested the performance of SEM versus gridMathematica using the pingpong and swap communications benchmarks on a cluster with 4, 8, and 16 processors. While SEM's performance remained steady for increased cluster size, gridMathematica's performance decreased due to its "master" kernel becoming a bottleneck for all communications.





A high-performance computing benchmark called "pingpong", developed by Viktor K. Decyk of the UCLA Plasma Physics Group, stresses the network of a parallel computer in two stages after dividing the processing elements into the even and odd processors. The pingpong stage has each even one send a message to an odd one (0 to 1, 2 to 3, ...) and back again. The message is varied in size from a few bytes to many megabytes while each operation is timed and averaged. The second stage performs a "swap", meaning each pair of processes sends and receives messages from each other simultaneously, if possible. Again the message size is varied and the results are timed and averaged.

© 2002 Dauger Research, Inc.

Patent Pending

http://dauger.com/pooch/mathematica/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3	6874

20995 7590 07/20/2016
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

EXAMINER
ASRES, HERMON

ART UNIT	PAPER NUMBER
2449	

NOTIFICATION DATE	DELIVERY MODE
07/20/2016	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

jayna.cartee@knobbe.com
efiling@knobbe.com

PTOL-90A (Rev. 04/07)

Office Action Summary	Application No. 14/181,112	Applicant(s) TANNENBAUM ET AL.	
	Examiner HERMON ASRES	Art Unit 2449	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *02/14/2014*.

☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims*

5)☒ Claim(s) *2-11* is/are pending in the application.

5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6)☐ Claim(s) _____ is/are allowed.

7)☒ Claim(s) *2-11* is/are rejected.

8)☐ Claim(s) _____ is/are objected to.

9)☐ Claim(s) _____ are subject to restriction and/or election requirement.

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to PPHfeedback@uspto.gov.

Application Papers

10)☐ The specification is objected to by the Examiner.

11)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

Certified copies:

a)☐ All b)☐ Some** c)☐ None of the:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☒ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b) Paper No(s)/Mail Date *09/10/2014, 11/18/2015*.

3) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

4) ☐ Other: _____.

DETAILED ACTION

Claims 2-11 have been examined and are rejected.

Information Disclosure Statement

The information disclosure statement (IDS) submitted on 11/18/2015, and 09/10/2014 is in compliance with the provisions of 37 CFR 1.97. Accordingly, the information disclosure statement is being considered by the examiner.

Double Patenting

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory double patenting rejection is appropriate where the claims at issue are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the reference application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement. A terminal disclaimer must be signed in compliance with 37 CFR 1.321(b).

The USPTO internet Web site contains terminal disclaimer forms which may be used. Please visit http://www.uspto.gov/forms/. The filing date of the application will determine what form should be used. A web-based eTerminal Disclaimer may be filled out completely online using web-screens. An eTerminal Disclaimer that meets all requirements is auto-processed and approved immediately upon submission. For more information about eTerminal Disclaimers, refer to http://www.uspto.gov/patents/process/file/efs/guidance/eTD-info-I.jsp.

Claim 2 is rejected on the ground of nonstatutory double patenting as being unpatentable over claims 1 and 8 of U.S. Patent No. 8,676,877. Although the conflicting claims are not identical, they are not patentably distinct from each other because:

	Instant Application	**US Patent 8,676,877**	
Claim	**Limitation**	**Claim**	**Limitation**
2	**A computer cluster comprising:**	**1 and 8**	**A system for performing an instruction received from a front end by executing commands on**

			one or more special purpose microprocessors, the system comprising:
	a plurality of nodes, wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a computer-readable medium comprising program code for a single-node kernel module configured to interpret user instructions; and a mechanism for the nodes to communicate with each other.		**a plurality of nodes, wherein each node is configured to access a computer-readable memory system comprising program code for a single-node kernel module, ……..wherein each cluster node module is stored in a computer-readable memory system and configured to communicate with a single-node kernel and with one or more other cluster node modules, to accept instructions, and to interpret at least some of the instructions such that the plurality of cluster node modules communicate with one another in order to act as a cluster in executing commands using one or more hardware processors**

Claim Rejections - 35 USC § 112

The following is a quotation of 35 U.S.C. 112 (pre-AIA), second paragraph:
The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

The following is a quotation of pre-AIA 35 U.S.C. 112, sixth paragraph:
An element in a claim for a combination may be expressed as a means or step for performing a specified function without the recital of structure, material, or acts in support thereof, and such claim shall be construed to cover the corresponding structure, material, or acts described in the specification and equivalents thereof.

Claim 2, 6 and 7 is rejected under 35 U.S.C. 112 second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which the inventor(s) or applicant regards as the invention.

As per claim 2, 6 and 7 "***module configured to***" is a limitations that invokes 35 U.S.C. 112, sixth paragraph. The written disclosure fails to disclose the corresponding structure, material, or acts for the claimed function. Applicant's specification is devoid of sufficient disclosure of structure for these terms as required by 112 second paragraph.

Rationale for invoking §112 6¶

Examiners will apply § 112, ¶ 6 to a claim limitation that meets the following conditions:

(1) The claim limitation uses the phrase "means for" or "step for" or a non-structural term that does not have a structural modifier;

(2) The phrase "means for" or "step for" or the non-structural term recited in the claim is modified by functional language; and

(3) The phrase "means for" or "step for" or the non-structural term recited in the claim is not modified by sufficient structure, material, or acts for achieving the specified function. This modifies the 3-prong analysis in MPEP § 2181, which will be revised in due course. See Supplemental Examination, 76 FR at 7167.

"When the claim limitation does not use the phrase "means for" or "step for," examiners should determine whether the claim limitation uses a nonstructural term (a term that is simply a substitute for the term "means for"). Examiners will apply § 112, ¶6 to a claim limitation that uses a nonstructural term associated with functional language, unless the nonstructural term is (1) preceded by a structural modifier, defined in the specification as a particular structure or known by one skilled in the art, that denotes the type of structural device (e.g., "filters"), or (2) modified by sufficient

structure or material for achieving the claimed function. The following is a list of non-structural terms that may invoke § 112, ¶6: "mechanism for," "module for," "device for," "unit for," "component for," "element for," "member for," "apparatus for," "machine for," or "system for." This list is not exhaustive, and other non-structural terms may invoke § 112, ¶6." See id.

Applicant is required to:

(a) Amend the claim so that the claim limitation will no longer be a means (or step, or non-structure terms) plus function limitation under 35 U.S.C. 112, sixth paragraph; or
(b) Amend the written description of the specification such that it expressly recites what structure, material, or acts perform the claimed function without introducing any new matter (35 U.S.C. 132(a)).

If applicant is of the opinion that the written description of the specification already implicitly or inherently discloses the corresponding structure, material, or acts so that one of ordinary skill in the art would recognize what structure, material, or acts perform the claimed function, applicant is required to clarify the record by either:
(a) Amending the written description of the specification such that it expressly recites the corresponding structure, material, or acts for performing the claimed function and clearly links or associates the structure, material, or acts to the claimed function, without introducing any new matter (35 U.S.C. 132(a)); or

(b) Stating on the record what the corresponding structure, material, or acts, which are implicitly or inherently set forth in the written description of the specification, perform the claimed function. For more information, see 37 CFR 1.75(d) and MPEP §§ 608.01(o) and 2181.

Claim Rejections - 35 USC § 101

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 2-11 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter. The claim(s) do not fall within at least one of the four categories of patent eligible subject matter because the claimed invention is directed to a judicial exception (i.e., a law of nature, a natural phenomenon, or an abstract idea) without significantly more. Claim(s) 2-11 are directed to computer cluster for "***a plurality of nodes wherein one or more of the nodes receives a command to start a cluster initialization process and a single-node kernel to interpret user instructions"*** which under the Mayo test can be viewed as merely insignificant extrasolution activity to the judicial exception, e.g., mere data gathering in conjunction with a law of nature or abstract idea" (see Mayo, 132 S. Ct. 1297-1301). The claims do not recite additional elements that are sufficient to amount to significantly more than the judicial exception as the claims do not recite an improvement to another technology or technical field, an improvement to the functioning of a computer itself, or meaningful

limitations beyond generally linking the use of an abstract idea to a particular technological environment. The "***computer cluster***", "***nodes***", and "**single-node kernel**" are recited as performing generic computer functions routinely used in computer applications. Generic computer components recited as performing generic computer functions that are well-understood, routine, and conventional activities amount to no more than implementing the abstract idea with a computerized system.

Claims not specifically mentioned are rejected by virtue of dependency and because they do not obviate the above-recited deficiencies.

Limitation that may be enough to qualify as "significantly more"

1. Improvements to another technology /technical fields;34
2. Improvements to the functioning of the computer itself;35
3. Applying the judicial exception with, or by user of a particular machine; 36
4. Effecting a transformation or reduction of a particular article to a different state or thing; 37
5. Adding a specific limitation other than what is well-understood, routine and conventional in the field, or adding unconventional steps that confine the claim to a particular useful application.38
6. Other meaningful limitations beyond generally linking the use of an abstract idea to a particular technological environment.39

34 (Id., slip op. at 15: e.g., a mathematical formula applied in a specific rubber molding process (citing Diamond v. Diehr, 450 U.S. 175,177-178 (1981)). In this case, the application was for a "[process] for molding raw, uncured synthetic rubber into cured precision products." The process of curing synthetic rubber depends on a number of factors including time, temperature and thickness of the mold. Using the Arrhenius equation it is possible to calculate when to open the press and to remove the cured, molded rubber. The problem was that there was, at the time the invention was made, no disclosed way to obtain an accurate measure of the temperature without opening the press. The invention solved this problem by using embedded thermocouples to constantly check the temperature, and then feeding the measured values into a computer. The computer then used the Arrhenius equation to calculate when sufficient energy had been absorbed so that the molding machine should open the press.)

35 (Id., at 2359 Alice (using a computer to obtain data, adjust account balances, and issue automated instructions - is “[p]urely ‘conventional. Mayo, 566 U. S., at _. They do not, for example, purport to improve the functioning of the computer itself or effect an improvement in any other technology or technical field, see Alice Corp. v. CIS Banks Inti. 110 USPG2d 1976 (U.S. 2014) “[o]n their face, they are drawn to the concept of intermediated settlement, i.e., the use of a third party to mitigate settlement risk. Like the risk hedging in Bilski, the concept of intermediated settlement is “ ‘a fundamental economic practice long prevalent in our system of commerce,” ibid., and the use of a third-party intermediary (or “clearing house”) is a building block of the

modern economy. Thus, intermediated settlement, like hedging, is an “abstract idea” beyond §101 ’s scope, pp. 7-10.))

36 (Bilski, 130 S. Ct. at 3227 ("The Court's precedents establish that the machine-or-transformation test is a useful and important clue, and investigative tool, for determining whether some claimed inventions are processes under §101.")

37 (Diehr, 450 U.S. at 184 (“That respondents’ claims [to a specific rubber molding process] involve the transformation of an article, in this case raw, uncured synthetic rubber, into a different state or thing cannot be disputed. The respondent's claims describe in detail a step-by-step method for accomplishing such, beginning with the loading of a mold with raw, uncured rubber and ending with the eventual opening of the press at the conclusion of the cure.”). See also Benson, 409 U.S. at 70 (“Transformation and reduction of an article ‘to a different state or thing’ is the clue to the patentability of a process claim that does not include particular machines. So it is that a patent in the process of ‘manufacturing fat acids and glycerine from fatty bodies by the action of water at a high temperature and pressure’ was sustained in Tilghman, 102 U.S. at 721”).

38 (Mayo, 132 S. Ct. at 1299, 1302 (claim ineligible because the recited “instructions add nothing specific to the laws of nature other than what is well-understood, routine, conventional activity, previously engaged in by those in the field,” which was “[u]nlike, say, a typical patent on a new drug or a new way of using an existing drug”)).

39 (Alice Corp., 134 S. Ct. at 2360 (As to its system claims, petitioner emphasizes that those claims recite “specific hardware” configured to perform “specific computerized functions.” Brief for Petitioner 53. But what petitioner characterizes as specific hardware - a “data processing system” with a “communications controller” and “data storage unit,” for example, see App. 954,958, 1257 - is purely functional and generic. Nearly every computer will include a “communications controller” and “data storage unit” capable of performing the basic calculation, storage, and transmission functions required by the method claims. As a result, none of the hardware recited "offers a meaningful limitation beyond generally linking 'the use of the [method] to a particular technological environment,' that is, implementation via computers" (citing Bilski v, Kappos 561 U.S. at 610, 611));

Limitation that were not found to be enough to qualify as "significantly more"

1. Adding the words "encoding" (or an equivalent) with an abstract idea, or mere instructions to implement an abstract idea on a computer.40

2. Simply appending well-understood, routine and conventional activities previously known to the industry, specified at a high level of generality, to the judicial exception, e.g., a claim to an abstract idea requiring no more than a generic computer to perform generic computer functions that are well understood, routine and conventional activities previously known to the industry;41

3. Adding insignificant extra-solution activity to the judicial exception, e.g., mere data gathering in conjunction with a law of nature or abstract idea; 42or

4. Generally linking the use of the judicial exception to a particular technological environment or field of use.43

40 (Id., at 2358 (simply implementing a mathematical principle on a physical machine, namely a computer[i]s not a patentable application of that principle." citing Mayo, 132 S. Ct. at 1301))

41 (Id., at 2359 (using a computer to obtain data, adjust account balances, and issue automated instructions); Mayo, 132 S. Ct. at 1300 (telling a doctor to measure metabolite levels in the blood using any known process).

42 (Mayo, 132 S. Ct. at 1297-98 (measuring metabolites of a drug administered to a patient); Flook, 437 U.S. at 589-90 (1978) (adjusting an alarm limit variable to a figure computed according to a mathematical formula)

43 (Mayo, 132 S. Ct. at 1300-01 (citing Bilski, 130 S.Ct. 3223-24 (limiting hedging to use in commodities and energy markets); Flook, 437 U.S. at 589-90.

*a copy of the federal register can be found at

http://www.gpo.gov/fdsys/pkg/FR-2014-12-16/pdf/2014-29414.pdf

Claims 2-11 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

The claims recite "***a computer-readable medium***". The specification is silent regarding the meaning of "***a computer-readable medium***". This can include signal per se. Thus, applying the broadest reasonable interpretation in light of the specification and taking into account the meaning of the words in their ordinary usage they would be understood by one of ordinary skill in the art (MPEP §2111), the claim as a whole covers both transitory and not transitory media. A transitory media does not fall into any of the 4 categories of invention (Process, Machine, Manufacture, or composition of matter). Therefore claim 1 is rejected under 35 U.S.C. 101 for being directed towards non-statutory subject matter. The applicant is respectfully suggested to amend the claims to overcome the 35 USC § 101 rejection.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of pre-AIA 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 2, and 4-11 are rejected under 35 U.S.C. 102(b) as being anticipated by Block et al. (US PGPub 2005/0021751).

As per claim 2, Block teaches a computer cluster (Block, ….cluster data port services within a cluster infrastructure. See paragraph [0025]) comprising:

a plurality of nodes, wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a computer-readable medium comprising program code for a single-node kernel module (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044]) configured to interpret user instructions; and a mechanism for the nodes to communicate with each other. (Block, establishment of multiple network connections between a source node, a target node and one or more backup nodes in such a manner that a cluster data port is effectively utilized as single data port. See paragraph [0025]).

As per claim 4, Block teaches wherein the mechanism comprises a message passing interface. (Block, a single data port instance may also support target to source message sends. See paragraph [0050]).

As per claim 5, Block teaches wherein each of the nodes comprises one or more cluster node modules. (Block, cluster data port services within a cluster infrastructure to provide reliable and efficient communications between nodes. See paragraph [0010]).

As per claim 6, Block teaches wherein each single-node kernel module is stored in a computer-readable medium and configured to accept and execute a request.

(Block, The routines executed to implement ……operating system or a specific application, component, program, object, module or sequence of instructions, will also be referred to herein as "computer program code," or simply "program code." See paragraph [0039], [0049]).

As per claim 7, Block teaches wherein each of the nodes comprises one or more cluster node modules, and each of the cluster node modules is stored in a computer-readable medium and configured to communicate with the single-node kernel and with one or more other cluster node modules. (Block, …..multiple logical or TCP connections to a designated target node with designated backup target nodes, which appears to the user (i.e., an operating system or ***kernel component) like a single data port or data pipe***. See paragraph [0044], [0039]).

As per claim 8, Block teaches wherein the plurality of cluster node modules act as a cluster. (Block, utilize cluster data port services within a cluster infrastructure to provide reliable and efficient communications between nodes in a clustered computer system. See paragraph [0010]).

As per claim 9, Block teaches wherein the plurality of cluster node modules communicate with one another to act as a cluster. (Block, Nodes 12, 14 and 16 are typically coupled together via a clustering network 18. See paragraph [0028]).

As per claim 10, Block teaches wherein the computer cluster includes a user interface. (Block, Cluster manager application 60 that provides the user interface whereby a user such as a systems administrator can manage clustering operations in the system. See paragraph [0037]).

As per claim 11, Block teaches wherein each cluster node module accepts instructions from the user interface and interprets one or more of the instructions. (Block, Cluster manager application 60 that provides the user interface whereby a user such as a systems administrator can manage clustering operations in the system. See paragraph [0037], [0039]).

Claim Rejections - 35 USC § 103

The following is a quotation of pre-AIA 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Block et al. (US PGPub 2005/0021751) in view of Gray (USPGPub 2007/0073705).

As per claim 3, Block teaches the computer cluster of claim 2 but doesn't teach a Mathematica kernel. In analogous art Gray teaches a Mathematica kernel. (Gray, The front end may include an interactive document referred to as a notebook similar to those

often used with MATHEMATICA.RTM. software systems. A notebook may include input to be sent to the kernel 104 and output received from the kernel, as well as text, graphics, palettes, etc. see paragraph [0031]).

Therefore it would have been obvious to one of ordinary skill in the art at the time the invention was made to take the teaching of Gray and apply them on the teaching of Block because the kernel and the front end may be implemented on a same computing system or on different computing systems that are communicatively coupled to one another. (Gray, see paragraph [0030]).

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. This includes:

U.S. PGPub (2007/0094532), which describes Kernel Debugging in a cluster computing system.

U.S. Patent (6,968,359), which describes merge protocol for cluster computer system.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to HERMON ASRES whose telephone number is (571)272-4257. The examiner can normally be reached on Monday - Friday 7:30 am to 5:00 pm est.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, VIVEK SRIVASTAVA can be reached on 571-272-7304. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/HERMON ASRES/
Examiner, Art Unit 2449

/VIVEK SRIVASTAVA/

Supervisory Patent Examiner, Art Unit 2449

Notice of References Cited	Application/Control No. 14/181,112	Applicant(s)/Patent Under Reexamination TANNENBAUM ET AL.	
	Examiner HERMON ASRES	Art Unit 2449	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	CPC Classification	US Classification
*	A	US-2005/0021751 A1	01-2005	Block, Timothy Roy	G06F9/54	709/225
*	B	US-2007/0073705 A1	03-2007	Gray; Theodore W.	G06F9/4443	1/1
*	C	US-2007/0094532 A1	04-2007	SENGUPTA; Sabyasachi	G06F11/3632	714/5.1
*	D	US-6,968,359 B1	11-2005	Miller; Robert	G06F9/5061	709/205
	E	US-				
	F	US-				
	G	US-				
	H	US-				
	I	US-				
	J	US-				
	K	US-				
	L	US-				
	M	US-				

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	CPC Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
	U	
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	February 14, 2014
	First Named Inventor	Tannenbaum, Zvi
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Asres, Hermon
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	8,402,080	03-19-2013	Tannenbaum et al.	
	2	8,849,889	09-30-2014	Tannenbaum et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	3	JP 2005-063033 w/ English Abstract	10-03-2005	Toshiba Corp		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]

22047750

Examiner Signature /Hermon Asres/	Date Considered 07/07/2016

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Place a check mark in this area when an English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 1 OF 3	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	5,423,046	06-06-1995	Nunnelley, et al.	
	2	5,881,315	03-09-1999	Cohen	
	3	5,930,768	07-1999	Hooban	
	4	6,108,699	08-22-2000	Moiin	
	5	6,202,080	03-13-2001	Lu et al.	
	6	6,546,403	04-08-2003	Carlson, Jr. et al.	
	7	6,578,068	06-10-2003	Bowman-Amuah	
	8	6,751,698	06-15-2004	Deneroff et al.	
	9	6,782,537	08-24-2004	Blackmore et al.	
	10	6,968,335	11-22-2005	Bayliss et al.	
	11	7,093,004	08-15-2006	Bernardin et al.	
	12	7,136,924	11-14-2006	Dauger	
	13	7,174,381	02-06-2007	Gulko et al.	
	14	7,249,357	07-24-2007	Landman et al.	
	15	7,334,232	02-19-2008	Jacobs et al.	
	16	7,437,460	10-14-2008	Chidambaran et al.	
	17	7,472,193	12-30-2008	Dauger	
	18	7,554,949	06-30-2009	Chen	
	19	7,937,455	05-03-2011	Saha et al.	
	20	8,082,289 (Attorney Docket No. ZTANN.002A)	12-20-2011	Tannenbaum et al.	U.S. Application No. 11/744,461 and its entire prosecution history**
	21	8,140,612 (Attorney Docket No. ZTANN.002CP1)	03-20-2012	Tannenbaum et al.	U.S. Application No. 12/040,519 and its entire prosecution history**
	22	8,402,083 (Attorney Docket No. ZTANN.002CP1C1)	03-19-2013	Tannenbaum et al.	U.S. Application No. 12/465,519 and its entire prosecution history**
	23	8,676,877 (Attorney Docket No. ZTANN.002P1C2)	03-18-2014	Tannenbaum et al.	U.S. Application No. 13/423,063 and its entire prosecution history**
	24	2002/0049859	04-25-2002	Bruckert et al.	

Examiner Signature	/Hermon Asres/	Date Considered	07/07/2016

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A./

[1] Place a check mark in this area when an English language translation is attached.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 2 OF 3	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	25	2003/0005266	01-02-2003	Dauger	
	26	2003/0051062	03-13-2003	Circenis	
	27	2003/0135621	07-17-2003	Romagnoli	
	28	2003/0195931	10-16-2003	Dauger	
	29	2003/0195938	10-16-2003	Howard et al.	
	30	2004/0110209	06-10-2004	Yokota, et al.	
	31	2004/0157203	08-12-2004	Dunk	
	32	2004/0252710	03-17-2005	Barsness et al.	
	33	2005/0021751	01-27-2005	Block et al.	
	34	2005/0038852	02-17-2005	Howard	
	35	2005/0060237	03-17-2005	Barsness et al.	
	36	2005/0076105	04-07-2005	Keohane et al.	
	37	2005/0108394	05-19-2005	Braun et al.	
	38	2005/0180095	08-18-2005	Ellis	
	39	2006/0053216	03-09-2006	Deokar et al.	
	40	2006/0106931	05-18-2006	Richoux	
	41	2007/0044099	02-22-2007	Rajput	
	42	2007/0073705	03-29-2007	Gray	
	43	2008/0250347	10-09-2008	Gray et al.	
	44	2008/0281997	11-13-2008	Archer et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	45	JP 08-87473 A	02-04-1996	Toshiba Corp		
	46	JP 2002117010 w/ English Abstract	04-19-2002	International Business Mach Corp		

Examiner Signature	/Hermon Asres/	Date Considered	07/07/2016

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] Place a check mark in this area when an English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A./

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 3 OF 3	Attorney Docket No.	ZTANN.002P1C3

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	47	"gridMathematica 1.1: Grid Computing Gets a Speed Boost from Mathematica 5", The Mathematica Journal, Vol. 9, No. 2, 2004.	
	48	"Wolfram gridMathematica", Wolfram Research, Inc., 2007.	
	49	CARNS et al., "An Evaluation of Message Passing Implementations on Beowulf Workstations", Aerospace Conference, March 6-13, 1999, IEEE 1999, Vol. 5, pp. 41-54.	
	50	DAUGER et al., "Plug-and-Play Cluster Computing using Mac OS X", IEEE International Conference, December 1-4, 2003, Paper appears in Cluster Computing January 8, 2004, pp. 430-435.	
	51	DAUGER et al., Plug-and Play Cluster Computing: High-Performance Computing For the Mainstream", IEEE Computing in Science and Engineering, March/April 2005, pp 27-33.	
	52	HAMSCHER et al., "Evaluation of Job-Scheduling Strategies for grid Computing", LNCS: Lecture Notes in Computer Science, 2000, pp. 191-202.	
	53	International Search Report dated September 11, 2008, International Application No. PCT/US07/70585 (Attorney Docket No. ZTANN.002VPC).	
	54	JAHANZEB et al., "Libra: a computational economy-based job scheduling system for clusters", Software Practice and Experience, February 24, 2004, Vol. 34, pp. 573-590.	
	55	JAIN et al., "Data Clustering: A Review", ACM Computing Surveys, September 1999, Vol. 31, No. 3.	
	56	KEPNER, Jeremy et al., "Parallel Matlab: The Next Generation", August 20, 2004, MIT Lincoln Laboratory, Lexington, MA.	
	57	KIM, Hahn et al., "Introduction to Parallel Programming and pMatlab v2.0", 2011, MIT Lincoln Laboratory, Lexington, MA.	
	58	MAEDER, R., "Mathematica Parallel Computing Toolkit: Unleash the Power of Parallel Computing", Wolfram Research, January 2005.	
	59	TEPENEU and IDA, "MathGridLink – A bridge between Mathematica and the Grid", Nippon Sofutowea Kagakkai Taikai Ronbunshu, 2003, Vol. 20, pp. 74-77.	
	60	Wolfram, Stephen: THE MATHEMATICA BOOK 5TH Edition; Wolfram Research, Inc. 2003	

** Signature indicates consideration of publication and file history. The Examiner has access to these materials through the PTO computer systems. If additional copies are desired, please notify Applicants through their attorneys

18642854

Examiner Signature	/Hermon Asres/	Date Considered	07/07/2016

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /H.A./

[1] Place a check mark in this area when an English language translation is attached.

EAST Search History

EAST Search History (Prior Art)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
S1	6	((Zvi) near2 (Tannenbaum)).INV.	USPAT	OR	OFF	2015/07/24 20:25
S2	7	((Dean) near2 (Dauger)).INV.	USPAT	OR	OFF	2015/07/24 20:25
S3	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S4	24	S3 and interpret	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S5	28950	709/223.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:26
S6	5291	H04L41/0213.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S7	1692	H04L29/08072.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S8	27776	H04L29/06.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:27
S9	6793	H04L41/22.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S10	8156	H04L41/12.cpc.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S11	0	S3 and S6	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28

S12	0	S3 and S7	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S13	15	S3 and S8	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S14	0	S3 and S9	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S15	2	S3 and S10	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/24 20:28
S16	99	("20020049859" | "20030005266" | "20030051062" | "20030135621" | "20030195931" | "20030195938" | "20040110209" | "20040157203" | "20040252710" | "20050021751" | "20050038852" | "20050060237" | "20050076105" | "20050108394" | "20050180095" | "20060053216" | "20060106931" | "20070044099" | "20070073705" | "20080250347" | "20080281997" | "5423046" | "5881315" | "5930768" | "6108699" | "6202080" | "6546403" | "6578068" | "6751698" | "6782537" | "6968335" | "7093004" | "7136924" | "7174381" | "7249357" | "7334232" | "7437460" | "7472193" | "7554949" | "7937455" | "8082289" | "8140612" | "8402083" | "8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 12:54
S17	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:19
S18	10	(computer with cluster) same nodes same kernel and mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:20
S19	177	(computer with cluster) same nodes same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/07/25 13:23
S20	10	("8402080").PN.	US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/11/29 23:38
S21	2	("8849889").PN.	US-PGPUB;	OR	OFF	2015/11/29

			USPAT; USOCR; EPO; JPO; DERWENT; IBM_TDB			23:38
S22	92	US-1234567-$.DID. OR US-5423046-$.DID. OR US-5881315-$.DID. OR US-5930768-$.DID. OR US-6108699-$.DID. OR US-6202080-$.DID. OR US-6546403-$.DID. OR US-6578068-$.DID. OR US-6751698-$.DID. OR US-6782537-$.DID. OR US-6968335-$.DID. OR US-7093004-$.DID. OR US-7136924-$.DID. OR US-7174381-$.DID. OR US-7249357-$.DID. OR US-7334232-$.DID. OR US-7437460-$.DID. OR US-7472193-$.DID. OR US-7554949-$.DID. OR US-7937455-$.DID. OR US-8082289-$.DID. OR US-8140612-$.DID. OR US-8402083-$.DID. OR US-8676877-$.DID. OR US-20020049859-$.DID. OR US-20030005266-$.DID. OR US-20030051062-$.DID. OR US-20030135621-$.DID. OR US-20030195931-$.DID. OR US-20030195938-$.DID. OR US-20040110209-$.DID. OR US-20040157203-$.DID. OR US-20040252710-$.DID. OR US-20050021751-$.DID. OR US-20050038852-$.DID. OR US-20050060237-$.DID. OR US-20050076105-$.DID. OR US-20050108394-$.DID. OR US-20050180095-$.DID. OR US-20060053216-$.DID. OR US-20060106931-$.DID. OR US-20070044099-$.DID. OR US-20070073705-$.DID. OR US-20080250347-$.DID. OR US-20080281997-$.DID. OR JP-1234567-$.DID. OR JP-0887473-$.DID. OR JP-2002117010-$.DID. OR US-6131999-$.DID.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2015/11/29 23:39
S23	4	(computer with cluster) same (nodes) same (single-node with kernel)	USPAT	OR	OFF	2016/06/17 19:19
S24	10	(computer with cluster) same (nodes) same (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S25	11	(computer with cluster) same (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S26	11	(computer with cluster) and (nodes) and (single-node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:20
S27	224	(computer with cluster) and (nodes) and (single with kernel)	US-PGPUB; USPAT;	OR	OFF	2016/06/17 19:20

			EPO; JPO; DERWENT; IBM_TDB			
S28	145	(computer with cluster) same (nodes) and (single with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:21
S29	6	(computer with cluster) same (nodes) and (single with kernel) and 709/223.ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/06/17 19:22
S30	10	(computer with cluster) same (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 16:08
S31	79	(computer with cluster) and (single with node with kernel)	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/07 18:41
S32	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 12:43
S33	2	("8676877").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:47
S34	1505	(computer with cluster) and kernel and nodes and request	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S35	271	S34 and mathematical	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S36	178750	S34 and mathematical smae kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S37	10	S34 and mathematical same kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:53
S38	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 15:54
S39	29	(single with node with kernel) same cluster	US-PGPUB; USPAT;	OR	OFF	2016/07/08 16:10

			EPO; JPO; DERWENT; IBM_TDB			
S40	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 16:11
S41	22	(mathematical with kernel) same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:18
S42	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 17:24
S43	180800	mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S44	11	mathematica near5 kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:39
S45	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:40
S46	10	("8402080").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:43
S47	2	("8849889").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44
S48	101	("20020049859" | "20030005266" | "20030051062" | "20030135621" | "20030195931" | "20030195938" | "20040110209" | "20040157203" | "20040252710" | "20050021751" | "20050038852" | "20050060237" | "20050076105" | "20050108394" | "20050180095" | "20060053216" | "20060106931" | "20070044099" | "20070073705" | "20080250347" | "20080281997" | "5423046" | "5881315" | "5930768" | "6108699" | "6202080" | "6546403" | "6578068" | "6751698" | "6782537" | "6968335" | "7093004" | "7136924" | "7174381" | "7249357" | "7334232" | "7437460" | "7472193" | "7554949" | "7937455" | "8082289" | "8140612" | "8402083" |	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:44

		"8676877").PN.				
S49	17	mathematica with cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:50
S50	43	mathematica same cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 18:51
S51	537678	stephen wolfram mathematica kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S52	50	(stephen near2 wolfram).in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:42
S53	1	S52 and mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S54	215	mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S55	74	S54 and cluster	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:43
S56	13	mathematica with kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/08 19:44
S57	0	("gridMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S58	0	("gridwithMathematica").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S59	19	grid with Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:16
S60	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT;	OR	OFF	2016/07/09 10:16

			IBM_TDB			
S61	14	gridMathematica and wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:21
S62	20	gridMathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S63	2578	Mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S64	215	Mathematica and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:22
S65	15749	wolfram	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S66	8450	wolfram.in.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:29
S67	15	S66 and mathematica	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S68	42	S66 and kernel	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:30
S69	2	("20070073705").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:44
S70	2	("20050021751").PN.	US-PGPUB; USPAT; EPO; JPO; DERWENT; IBM_TDB	OR	OFF	2016/07/09 10:46

EAST Search History (Interference)

<This search history is empty>

7/9/2016 12:33:14 PM
C:\Users\hasres\Documents\EAST\Workspaces\14181112.wsp

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	07/07/2016								
	1	-								
	2	✓								
	3	✓								
	4	✓								
	5	✓								
	6	✓								
	7	✓								
	8	✓								
	9	✓								
	10	✓								
	11	✓								

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	14181112	TANNENBAUM ET AL.
	Examiner	**Art Unit**
	HERMON ASRES	2449

CPC- SEARCHED

Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED

Symbol	Date	Examiner

US CLASSIFICATION SEARCHED

Class	Subclass	Date	Examiner
709	223	07/07/2016	HA

SEARCH NOTES

Search Notes	Date	Examiner
Searched East (US Patents, USPGPubs, EPO, JPO, IBM, Derwent)	07/07/2016	HA
See East Search attached	07/07/2016	HA
Inventor search performed	07/07/2016	HA
Consulted Justin Rider (QAS) regarding Alice 101 rejection	07/08/2016	HA

INTERFERENCE SEARCH

US Class/ CPC Symbol	US Subclass / CPC Group	Date	Examiner

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	February 14, 2014
	First Named Inventor	Tannenbaum, Zvi
	Art Unit	2449
(Multiple sheets used when necessary)	Examiner	Asres, Hermon
SHEET 1 OF 1	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	8,402,080	03-19-2013	Tannenbaum et al.	
	2	8,849,889	09-30-2014	Tannenbaum et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	T[1]
	3	JP 2005-063033 w/ English Abstract	10-03-2005	Toshiba Corp		

NON PATENT LITERATURE DOCUMENTS

Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]

22047750

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

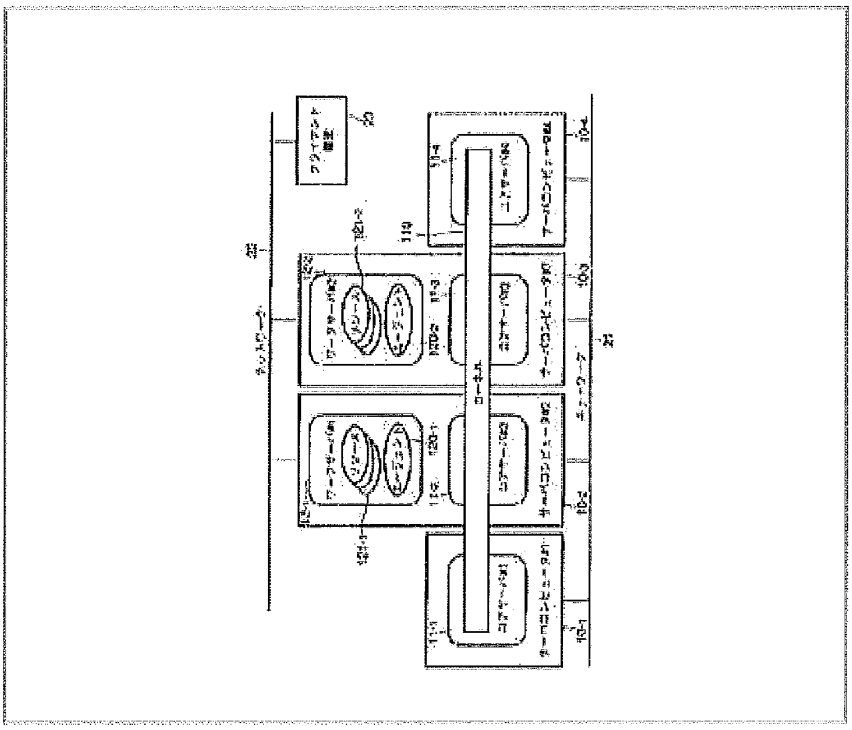
T[1] - Place a check mark in this area when an English language Translation is attached.

(11)Publication number	2005-063033
(43)Date of publication of application	10.03.2005
(51)Int.Cl.	G06F 15/173
	G06F 15/177
(21)Application number	2003-290264
(22)Date of filing	08.08.2003
(71)Applicant	TOSHIBA CORP
(72)Inventor	ENDO KOTARO

(54)CLUSTER SYSTEM AND SERVER PROGRAM

(57)Abstract

PROBLEM TO BE SOLVED: To realize cluster structure having a large number of nodes without sharply increasing communication between nodes, and to flexibly change the number of nodes. SOLUTION: On a plurality of computers which are at least part of server computers 10-1 to 10-4 constituting a cluster system, for example, computers 10-2 and 10-3, work servers 12-1 and 12-2 which perform service control for providing a service requested from a client terminal 23 operate independently. Also, on a plurality of computers which are at least part of server computers 10-1 to 10-4, for example, all the computers 10-1 to 10-4, core servers 11-1 to 11-4 operate. The core servers 11-1 to 11-4 synchronously control the work servers while communicating with each other. Thus, it is possible to execute consistent cluster control as a single kernel 110 as a whole.



(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2005-63033
(P2005-63033A)
(43)公開日　平成17年3月10日(2005.3.10)

(51)Int.Cl.⁷	F I		テーマコード（参考）
G06F 15/173	G06F 15/173		5B045
G06F 15/177	G06F 15/177	670C	

審査請求 有　請求項の数 9　ＯＬ　（全 19 頁）

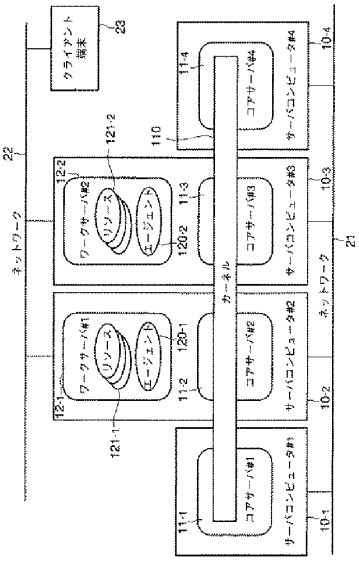
(21)出願番号　特願2003-290264(P2003-290264)
(22)出願日　平成15年8月8日(2003.8.8)

(71)出願人　000003078
株式会社東芝
東京都港区芝浦一丁目１番１号
(74)代理人　100058479
弁理士　鈴江　武彦
(74)代理人　100091351
弁理士　河野　哲
(74)代理人　100088683
弁理士　中村　誠
(74)代理人　100108855
弁理士　蔵田　昌俊
(74)代理人　100084618
弁理士　村松　貞男
(74)代理人　100092196
弁理士　橋本　良郎

最終頁に続く

(54)【発明の名称】クラスタシステム及びサーバプログラム

(57)【要約】
【課題】ノード数の多いクラスタ構成がノード間の通信の大幅な増加を招くことなく実現でき、且つノード数を柔軟に変更できるようにする。
【解決手段】クラスタシステムを構成するサーバコンピュータ１０-1～１０-4のうちの少なくとも一部の複数のコンピュータ、例えばコンピュータ１０-2，１０-3上では、それぞれクライアント端末２３から要求されたサービスを提供するためのサービス制御を行うワークサーバ１２-1，１２-2が独立して動作する。また、サーバコンピュータ１０-1～１０-4のうちの少なくとも一部の複数のコンピュータ、例えば全てのコンピュータ１０-1～１０-4上では、コアサーバ１１-1～１１-4が動作する。コアサーバ１１-1～１１-4は、互いに通信を行いながら同期してワークサーバを制御することにより、全体として１つのカーネル１１０として一貫性のあるクラスタ制御を行う。
【選択図】　図1



【特許請求の範囲】
【請求項1】
複数のコンピュータから構成されるクラスタシステムにおいて、
前記クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上でそれぞれ独立動作して、クライアント端末から要求されたサービスを提供するためのサービス制御を行うワークサーバと、
前記クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上でそれぞれ動作するコアサーバであって、他のコアサーバと互いに通信を行いながら同期して前記ワークサーバを制御することにより、他のコアサーバと一体となって１つのカーネルとして一貫性のあるクラスタ制御を行うコアサーバと
を具備することを特徴とするクラスタシステム。
【請求項2】
前記各コンピュータ上では、前記コアサーバ及びワークサーバのいずれか一方のみが動作することを特徴とする請求項１記載のクラスタシステム。
【請求項3】
前記ワークサーバは、
当該ワークサーバと接続されるコアサーバを切り替える切替手段と、
当該ワークサーバと接続されているコアサーバに対して当該ワークサーバから送信されたデータを格納するための第１のバッファと、
前記切替手段により当該ワークサーバと接続されるコアサーバが切り替えられた場合、前記第１のバッファに格納されているデータのうち、前記切替手段による切り替え前に当該ワークサーバと接続されていたコアサーバが確かに受信したと確認できないデータを、前記切替手段により当該ワークサーバと新たに接続されたコアサーバに再送する第１の再送手段とを備え、
前記コアサーバは、
前記ワークサーバ毎に設けられ、対応するワークサーバに対して当該コアサーバまたは他のコアサーバから送信されたデータを格納するための第２のバッファと、
当該コアサーバが新たにワークサーバと当該ワークサーバ内の前記切替手段により接続された場合、当該ワークサーバに対応する前記第２のバッファに格納されているデータのうち、前記切替手段による切り替え前に当該ワークサーバと接続されていたコアサーバから当該ワークサーバに送信されたものの、当該ワークサーバが確かに受信したと確認できないデータを当該ワークサーバに再送する第２の再送手段とを備えている
ことを特徴とする請求項１記載のクラスタシステム。
【請求項4】
前記ワークサーバは、
当該ワークサーバと接続されているコアサーバから送信されたデータを受信した場合に、受信応答を返す第１の送信手段と、
前記第１の再送手段による再送の対象となるデータのうち、最も古いデータが格納されている前記第１のバッファ内の位置を指し示す第１の再送ポインタ情報を保持する第１の再送ポインタ手段と、
当該ワークサーバと接続されているコアサーバに対して送信されたデータに対する受信応答が返された場合に、前記第１の再送ポインタ情報を更新する第１の再送ポインタ更新手段とを更に備え、
前記コアサーバは、
当該コアサーバと接続されているワークサーバから送信されたデータを受信した場合に、受信応答を返す第２の送信手段と、
前記第２の再送手段による再送の対象となるデータのうち、最も古いデータが格納されている前記第２のバッファ内の位置を指し示す第２の再送ポインタ情報を保持する第２の再送ポインタ手段と、

当該コアサーバと接続されているワークサーバに対して送信されたデータに対する受信応答が返された場合に、前記第２の再送ポインタ情報を更新する第２の再送ポインタ更新手段とを更に備えている

ことを特徴とする請求項３記載のクラスタシステム。

【請求項5】

前記ワークサーバは、当該ワークサーバにおけるサービスの状態を各サービス毎に保持する第１の状態保持手段を更に備え、

前記コアサーバは、

前記クラスタシステム内の全てのワークサーバにおけるサービスの状態を各サービス毎に保持する第２の状態保持手段と、

当該コアサーバが他のコアサーバと一体となって１つのカーネルとして再起動される際に、当該コアサーバと接続されている前記ワークサーバの第１の状態保持手段に保持されている情報を収集して前記第２の状態保持手段の内容を復元する手段とを更に備えている

ことを特徴とする請求項１記載のクラスタシステム。

【請求項6】

複数のコンピュータから構成されるクラスタシステムであって、当該クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上では、クライアント端末から要求されたサービスを提供するためのサービス制御を行うワークサーバがそれぞれ独立に動作し、前記クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上ではそれぞれクラスタ制御を行うコアサーバが他のコアサーバと互いに通信を行いながら同期して動作して前記ワークサーバを制御することにより、他のコアサーバと一体となって１つのカーネルとして一貫性のあるクラスタ制御を行うクラスタシステムにおいて、前記ワークサーバを実現するためのサーバプログラムであって、

前記ワークサーバが動作するコンピュータに、

当該ワークサーバと接続されているコアサーバに対して当該ワークサーバからデータを送信した場合に、当該送信データをバッファに格納するステップと、

当該ワークサーバと接続されるコアサーバを切り替えるステップと、

当該ワークサーバと接続されるコアサーバが切り替えられた場合、前記バッファに格納されているデータのうち、切り替え前に当該ワークサーバと接続されていたコアサーバが確かに受信したと確認できないデータを、新たに接続されたコアサーバに再送するステップと

を実行させるためのサーバプログラム。

【請求項7】

前記ワークサーバが動作するコンピュータに、

当該ワークサーバと接続されているコアサーバから送信されたデータを受信した場合に、受信応答を返すステップと、

当該ワークサーバと接続されているコアサーバに対して送信されたデータに対する受信応答が返された場合に、再送の対象となるデータのうち、最も古いデータが格納されている前記バッファ内の位置を指し示す再送ポインタを更新するステップと

を更に実行させるための請求項６記載のサーバプログラム。

【請求項8】

複数のコンピュータから構成されるクラスタシステムであって、当該クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上では、クライアント端末から要求されたサービスを提供するためのサービス制御を行うワークサーバがそれぞれ独立に動作し、前記クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上ではそれぞれクラスタ制御を行うコアサーバが他のコアサーバと互いに通信を行いながら同期して動作して前記ワークサーバを制御することにより、他のコアサーバと一体となって１つのカーネルとして一貫性のあるクラスタ制御を行うクラスタシステムにおいて、前記コアサーバを実現するためのサーバプログラ

ムであって、
　前記コアサーバが動作するコンピュータに、
　当該コアサーバと接続されているワークサーバに対して当該コアサーバからデータを送信した場合に、当該送信データを当該ワークサーバに対応して設けられたバッファに格納するステップと、
　当該コアサーバが新たにワークサーバと接続された場合、当該ワークサーバに対応する前記バッファに格納されているデータのうち、それ以前に当該ワークサーバと接続されていたコアサーバから当該ワークサーバに送信されたものの、当該ワークサーバが確かに受信したと確認できないデータを当該ワークサーバに再送するステップと
　を実行させるためのサーバプログラム。
【請求項9】
　前記コアサーバが動作するコンピュータに、
　当該コアサーバと接続されているワークサーバから送信されたデータを受信した場合に、受信応答を返すステップと、
　当該コアサーバと接続されているワークサーバに対して送信されたデータに対する受信応答が返された場合に、再送の対象となるデータのうち、最も古いデータが格納されている前記バッファ内の位置を指し示す再送ポインタを更新するステップと
　を更に実行させるための請求項８記載のサーバプログラム。

【発明の詳細な説明】
【技術分野】
【0001】
　本発明は、複数のコンピュータから構成されるクラスタシステムに係り、特にクラスタ制御を実行するコアサーバと、当該コアサーバと独立して設けられ、複数のコアサーバが同期して動作することで実現される仮想マシンとしてのカーネルからの制御によって動作するワークサーバとを実現するクラスタシステム及びサーバプログラムに関する。
【背景技術】
【0002】
　近年、コンピュータ上でアプリケーションプログラムを実行することにより、ユーザ（クライアント端末）にサービスを提供するシステムが運用されている。この種のシステムでは、継続したサービスの提供が必須になっている。これに伴い、サービス（業務）を実行しているコンピュータ（サーバコンピュータ）にも高い可用性（サーバ稼働率、業務稼働率）が求められる。そこで、複数のコンピュータをクラスタ構成として、一部のコンピュータで障害が発生しても、別のコンピュータでサービスを引き継いで、システム全体が停止するのを防止する、いわゆるクラスタシステムが開発されている（例えば、非特許文献１参照）。
【0003】
　クラスタシステムを構成するには、各コンピュータ上にクラスタマネージャが必要となる。クラスタマネージャは、クラスタの制御と、アプリケーションを起動・停止する制御（サービス制御）とを実行する。クラスタシステムでは、複数のコンピュータ（ノード）で分散してクラスタマネージャによるクラスタ制御が行われる。ここでは、クラスタ全体としての視点でクラスタ制御が行われること、つまり各コンピュータで分散して行われるクラスタ制御が、全体として一貫性のある制御となっていることが必要となる。そこでクラスタシステムでは、各コンピュータでのクラスタ制御が、互いに通信を行いながら同期して（連携して）行われる。つまり、クラスタ制御が各コンピュータで多重化して行われる。これにより高可用性（高業務稼働率）が実現される。
【0004】
　ところがクラスタシステムにおいてスプリットブレインと呼ばれる状況が発生すると、多重化された処理に一貫性がなくなり、システム全体に致命的な影響を及ぼす。このスプリットブレインは、コンピュータの障害検出が誤って行われたときに発生する。例えばクラスタシステム内の２つのコンピュータ群の間で互いに通信できない状態となった場合、

それぞれのコンピュータ群は、互いに障害検出し、独立して動作を継続する。これにより、アプリケーションやデータが破壊される虞がある。そこで、スプリットブレインの発生を防止するために、ｎ台のコンピュータで処理を多重化し、ｆ台（ｆは３ｆ＜ｎを満足する最大の整数）のコンピュータが故障しても、残りのｎ－ｆ台のコンピュータが多重化処理を行うことにより、多数決により正しい処理結果を決定できるようにした技術が提案されている（例えば、特許文献１参照）。ここで、ｆ＜ｎ／３であることから、ｎ－ｆ＞ｎ－ｎ／３、即ちｎ－ｆ＞２ｎ／３である。このため、特許文献１に記載された技術は、少なくともｎ台のコンピュータのうちのｎ－ｆ台が正常であること、つまりｎ台のコンピュータのうちの２／３を超えるコンピュータが正常であることが要求されることから、２／３定足数アルゴリズムと呼ばれる。

【特許文献1】特開２００３－６７２１５（段落００５６乃至００６４、図４）

【非特許文献1】金子哲夫、森良哉、「クラスタソフトウェア」、東芝レビュー、Vol.54 No.12(1999)、p.18-21

【発明の開示】

【発明が解決しようとする課題】

【0005】

上記したように従来のクラスタシステムでは、各コンピュータ上で動作するクラスタマネージャが互いに通信を行ってクラスタ制御を行うことで、システム全体として一貫性のある制御を実現している。また各コンピュータ上で動作するクラスタマネージャにより、ユーザにサービスを提供するためのアプリケーションを起動・停止する制御（サービス制御）が行われる。

【0006】

このように従来のクラスタシステムでは、クラスタマネージャは、クラスタを構成する全てのノード（コンピュータ）に存在する。このため、各ノード上のクラスタマネージャが一貫性のある制御を行おうとすると、全てのノード間で当該制御のための通信が発生する。その結果、ノード数が多くなると、一貫性ある制御のためのノード間の通信が増加して、通信に起因するオーバヘッド（ネットワーク負荷）が大きくなる。よって従来は、少ないノード数（せいぜい１６ノード程度）のクラスタシステムしか実用的ではなかった。また従来のクラスタシステムでは、各ノード上のクラスタマネージャが、クラスタ制御と、アプリケーションを起動・停止する制御（サービス制御）とを実行している。このため従来のクラスタシステムでは、クラスタ制御の一貫性を保ちながら動的にノード数を増やすための仕組みが複雑であり、ノードの動的な追加に対応しきれていないという問題もある。

【0007】

本発明は上記事情を考慮してなされたものでその目的は、ノード数の多いクラスタ構成がノード間の通信の大幅な増加を招くことなく実現できるクラスタシステム及びサーバプログラムを提供することにある。

本発明の他の目的は、ノード数を柔軟に変更可能なクラスタシステム及びサーバプログラムを提供することにある。

【課題を解決するための手段】

【0008】

本発明の１つの観点によれば、複数のコンピュータから構成されるクラスタシステムが提供される。このクラスタシステムは、クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上でそれぞれ独立動作して、クライアント端末から要求されたサービスを提供するためのサービス制御を行うワークサーバと、上記クラスタシステムを構成する複数のコンピュータのうちの少なくとも一部の複数のコンピュータ上でそれぞれ動作するコアサーバであって、他のコアサーバと互いに通信を行いながら同期して上記ワークサーバを制御することにより、他のコアサーバと一体となって１つのカーネルとして一貫性のあるクラスタ制御を行うコアサーバとを備えることを特徴とする。

【0009】
このような構成においては、サービス制御とクラスタ制御とがそれぞれ独立のワークサーバとコアサーバによって行われることから、ワークサーバが配置されるサーバコンピュータとコアサーバが配置されるサーバコンピュータ（ワークサーバが動作するサーバコンピュータと同じコンピュータであっても構わない）とを柔軟に決定できる。このため、ノード数の多いクラスタ構成がノード間の通信の大幅な増加を招くことなく実現できると共にノード数を柔軟に変更することができる。
【0010】
ここで、上記ワークサーバには、当該ワークサーバと接続されるコアサーバを切り替える切替手段と、当該ワークサーバと接続されているコアサーバに対して当該ワークサーバから送信されたデータを格納するための第１のバッファと、上記切替手段により当該ワークサーバと接続されるコアサーバが切り替えられた場合、上記第１のバッファに格納されているデータのうち、上記切替手段による切り替え前に当該ワークサーバと接続されていたコアサーバが確かに受信したと確認できないデータを、上記切替手段により当該ワークサーバと新たに接続されたコアサーバに再送する第１の再送手段とを持たせるとよい。一方、上記コアサーバには、上記ワークサーバ毎に設けられ、対応するワークサーバに対して当該コアサーバまたは他のコアサーバから送信されたデータを格納するための第２のバッファと、当該コアサーバが新たにワークサーバと当該ワークサーバ内の上記切替手段により接続された場合、当該ワークサーバに対応する上記第２のバッファに格納されているデータのうち、上記切替手段による切り替え前に当該ワークサーバと接続されていたコアサーバから当該ワークサーバに送信されたものの、当該ワークサーバが確かに受信したと確認できないデータを当該ワークサーバに再送する第２の再送手段とを持たせるとよい。
【0011】
このような構成においては、ワークサーバと接続されているコアサーバの障害が発生しても、他のコアサーバとの接続にスムーズに切り替えることができ、しかもデータを失うことが防止できるため、コアサーバの障害に対して耐障害性を持つことができる。
【0012】
また、上記ワークサーバには、当該ワークサーバにおけるサービスの状態を各サービス毎に保持する第１の状態保持手段を持たせ、上記コアサーバには、上記クラスタシステム内の全てのワークサーバにおけるサービスの状態を各サービス毎に保持する第２の状態保持手段と、当該コアサーバが他のコアサーバと一体となって１つのカーネルとして再起動される際に、当該コアサーバと接続されているワークサーバの第１の状態保持手段に保持されている情報を収集して上記第２の状態保持手段の内容を復元する手段を持たせるならば、カーネルの再起動後も、クラスタの状態を維持して制御を続けることができる。
【発明の効果】
【0013】
本発明によれば、サービス制御を行うワークサーバとクラスタ制御を行うコアサーバとが互いに分離独立していることから、クラスタシステム上でワークサーバとコアサーバを柔軟に配置できる。このため本発明によれば、、ノード数の多いクラスタ構成がノード間の通信の大幅な増加を招くことなく実現でき、しかもノード数を柔軟に変更できる。
【発明を実施するための最良の形態】
【0014】
以下、本発明の実施の形態につき図面を参照して説明する。
［第１の実施形態］
図１は本発明の第１の実施形態に係るクラスタシステムの構成を示すブロック図である。図１のクラスタシステムは、４台のサーバコンピュータ１０-1（＃１），１０-2（＃２），１０-3（＃３），１０-4（＃４）から構成される。サーバコンピュータ１０-1，１０-2，１０-3，１０-4は２つのネットワーク２１，２２により相互接続されている。
【0015】
サーバコンピュータ１０-1，１０-2，１０-3，１０-4上では、クラスタ制御を行うコア

サーバ１１-1（＃１），１１-2（＃２），１１-3（＃３），１１-4（＃４）が動作する。コアサーバ１１-1～１１-4は、サーバコンピュータ１０-1～１０-4が、対応するサーバプログラム（クラスタ制御のためのコアサーバプログラム）を実行することにより実現される。

【0016】

また、サーバコンピュータ１０-1～１０-4のうちのサーバコンピュータ１０-2，１０-3上では、サービス制御（つまりアプリケーションを起動・停止する制御）を行うワークサーバ１２-1（＃１），１２-2（＃２）が動作する。ワークサーバ１２-1，１２-2は、サーバコンピュータ１０-2，１２-3が、対応するサーバプログラム（サービス制御のためのワークサーバプログラム）を実行することにより実現される。

【0017】

ネットワーク２２には、ワークサーバ１２-1，１２-2に対してサービスの実行を要求するクライアント端末２３が接続されている。ワークサーバ１２-1，１２-2とクライアント端末２３との間の通信は、ネットワーク２２を介して行われる。なお、図１では、作図の都合上、ネットワーク２２に１つのクライアント端末２３が接続されている例が示されている。しかし、ネットワーク２２には、複数のクライアント端末が接続されるのが一般的である。

【0018】

サーバコンピュータ１０-1～１０-4上の各コアサーバ１１-1～１１-4は、従来のクラスタシステムにおける各コンピュータ上で動作するクラスタマネージャのクラスタ制御機能と同様のクラスタ制御機能を有し、他のコアサーバと一体となって（多重化して動作して）、クラスタシステムの制御を実行する。一体となったコアサーバ１１-1～１１-4は、カーネル１１０と呼ぶ１つのバーチャルマシン（仮想的な実行環境）を形成する。カーネル１１０は、個々のサーバコンピュータ１０-1～１０-4上で動作するコアサーバ１１-1～１１-4が連携して形成される。このためカーネル１１０は、サーバコンピュータ１０-1～１０-4にまたがって存在していると考えることができる。つまり、個々のコアサーバ１１-1～１１-4がカーネルなのではなく、コアサーバ１１-1～１１-4が一体となってカーネル１１０が構築され、よってクラスタシステムにはひとつだけカーネルが存在する。

【0019】

コアサーバ１１-1～１１-4が一体となって同期（連携）して動作する（つまり多重化して動作する）のに必要な、コアサーバ１１-1～１１-4間の通信には、ネットワーク２１が用いられる。ここで、ネットワーク２１には、上記多重化動作の高速化のために、ワークサーバ１２-1，１２-2とクライアント端末２３との間の通信に用いられるネットワーク２２よりも高速のネットワークが用いられる。なお、高速化が要求されない場合には、コアサーバ１１-1～１１-4間の通信にもネットワーク２２を用いても構わない。

【0020】

サーバコンピュータ１０-2，１０-3上のワークサーバ１２-1，１２-2は、カーネル１１０からの制御に従って動作し、サービス（アプリケーション）の起動・停止を制御する。つまり、ワークサーバ１２-1，１２-2ではサービスが実行される。このサービスについて、ウェブサービスを例に説明する。まずウェブサービスを実行するには、当該サービスに対するクライアント端末２３からのリクエストに応答してコンテンツを配信する役割を持つウェブサーバが必要である。またクライアント端末からウェブサーバにリクエストを届けるためにはアドレス、例えばＩＰ（Internet Protocol）アドレスが必要である。提供されるコンテンツを保存するためのファイルシステムも必要である。つまりウェブサービスの提供には、ウェブサーバ、ＩＰアドレス、ファイルシステムなどの、物理的、もしくは論理的な実体が必要となる。これらの実体は、「リソース」と呼ばれる。つまり、ワークサーバ１２-1，１２-2では、それぞれ複数のリソース１２１-1，１２１-2を組み合わせて実行することにより、結果としてサービスが実行される。ワークサーバ１２-1，１２-2において、リソース１２１-1，１２１-2の起動・停止を実際に行う手段をエージェント１２０-2と呼ぶ。エージェント１２０-1，１２０-2は、カーネル１１０から制御の指示を受

け取り、それに従ってリソース１２１-1，１２１-2を制御し、結果をカーネル１１０に返す。

【0021】

このように、図１のクラスタシステムには２台のワークサーバ１２-1，１２-2が存在し、それぞれが独立して稼働する。このため、カーネル１１０が、ワークサーバ１２-1，１２-2をいかに制御し、同期をとるかという点が重要である。そこで本実施形態では、クラスタシステムの制御がカーネル１１０によって一元的に行われる構成を適用している。カーネル１１０は、コアサーバ１１-1～１１-4で多重化実行されるバーチャルマシンである。このため、一部のコアサーバに障害が発生した場合でも、クラスタシステムの制御を継続することが可能である。

【0022】

コアサーバ１１-1～１１-4における多重化実行のアルゴリズムには、前記特許文献１に記載されているスプリットブレインを起こさない２／３定足数アルゴリズムが適用される。本実施形態のクラスタシステムは、ｎ＝４，ｆ＝１であり、２／３定足数アルゴリズムが適用可能な最も少ない数のコンピュータで構成される。つまり、本実施形態のクラスタシステムは２／３定足数アルゴリズムが適用可能な最小クラスタ構成をとる。これにより、ｆ（＝１）台のコアサーバの停止障害だけでなく、ビザンティン故障にも耐障害性を実現できる。このように、カーネル１１０は、超高信頼なバーチャルマシンとなっており、信頼性の高いクラスタシステムの制御を実現する。

【0023】

さて本実施形態では、クラスタシステム内の全てのサーバコンピュータ、つまり４台のサーバコンピュータ１０-1～１０-4上でコアサーバ１１-1～１１-4が稼働させられて、カーネル１１０が実現される。これにより上記したように、２／３定足数アルゴリズムを適用するのに必要な最少のコンピュータ数（ｎ＝４）でクラスタシステムが構築される。また、この４台のサーバコンピュータ１０-1～１０-4のうちの２台のサーバコンピュータ１０-2，１０-3上でワークサーバ１２-1，１２-2が稼働させられて、クラスタシステムのサービスが実行される。このため、サーバコンピュータ１０-2，１０-3には、それぞれそのサービスに必要なリソース１２１-1，１２１-2と、それに見合ったハードウェア／ソフトウェアが必要になるものの、サーバコンピュータ１０-1，１０-4ではコアサーバ１１-1，１１-2しか動作しないため、当該サーバコンピュータ１０-1，１０-4は比較的小さなコンピュータで十分である。

【0024】

以上のことから、図１のクラスタシステムでは、サービス制御に関しては、２つのワークサーバ１２-1，１２-2間でのフェイルオーバによる可用性の向上が期待できる。一方、クラスタ制御に関しては、４つのコアサーバ１１-1～１１-4による２／３定足数アルゴリズムによる可用性の向上が期待できる。

【0025】

図２は、図１のクラスタシステム内のコアサーバ及びワークサーバの構成を、コアサーバ１１-1，１１-4及びワークサーバ１２-1について示すブロック図である。図２に示すように、コアサーバ１１-1，１１-4は、同期処理部１１１-1，１１１-4と、状態管理テーブル１１２-1，１１２-4と、送信部１１３-1，１１３-4と、受信部１１４-1，１１４-4と、バッファ１１５-1，１１５-4と、再送ポインタ（再送ポインタ保持部）１１６-1，１１６-4と、再送部１１７-1，１１７-4とを備えている。なお、図２では省略されているが、コアサーバ１１-2，１１-3も、上記コアサーバ１１-1，１１-4と同様の構成を備えている。また、図２では省略されているが、これらの構成はワークサーバ毎に用意される。また、少なくともバッファがワークサーバ毎に用意される構成であっても構わない。

【0026】

同期処理部１１１-i（ｉ＝１～４）は、コアサーバ１１-1～１１-4間の一貫性を保つために、ネットワーク２１を介して通信する。状態管理テーブル１１２-iは、図１のクラスタシステム内の全てのワークサーバ、即ちワークサーバ１２-1（＃１），１２-2（＃２）

における各サービスの状態を管理するのに用いられる。

【0027】

送信部１１３-iは、図1のクラスタシステム内のワークサーバ１２-j（ｊ＝１または２）へのデータ送信、及び当該ワークサーバ１２-jから送信されたデータを受信したことを当該ワークサーバ１２-jへ通知するための受信応答を行う。受信応答は、受信データに付される当該データを識別するために当該データに割り当てられた識別情報、例えばシーケンス番号を通知することで行われる。送信部１１３-iは、データ送信時または受信応答時には、送信データまたは受信データのシーケンス番号を同期処理部１１１-iにより他のコアサーバに通知させる。

【0028】

受信部１１４-iは、ワークサーバ１２-jから送信されたデータの受信、及び当該ワークサーバ１２-jへのデータ送信に対する当該ワークサーバ１２-jからの受信応答の受信を行う。受信部１１４-iは、ワークサーバ１２-jからのデータの受信時または受信応答の受信時に、受信データまたは受信応答で通知されたシーケンス番号を同期処理部１１１-iにより他のコアサーバに通知させる。

【0029】

バッファ１１５-iは、送信部１１３-iからワークサーバ１２-jに送信されたデータをシーケンス番号と対応付けて格納するのに用いられる。バッファ１１５-iは、ワークサーバ１２-j毎に用意される。バッファ１１５-iは、再送ポインタ（ＲＰ）１１６-iと、図示せぬ書き込みポインタ（ＷＰ）とによって管理される。再送ポインタ１１６-iは、バッファ１１５-iに格納されている、ワークサーバ１２-jが確かに受信したと確認できないデータのうち、最も古いデータの位置を指す。書き込みポインタは、次にバッファ１１５-iに書き込むべき送信データの格納先を指す。再送部１１７-iは、ワークサーバ１２-jと接続されるコアサーバが他のコアサーバからコアサーバ１１-iに切り替えられた場合に、それまでワークサーバ１２-jと接続されていたコアサーバが送信したデータのうち、当該ワークサーバ１２-jで確かに受信したと確認できないデータを、再送ポインタ１１６-iに従って取り出して当該ワークサーバ１２-jに再送する。

【0030】

一方、ワークサーバ１２-1（１２-j）内のエージェント１２０-1（１２０-j）は、切替部１２１と、状態テーブル１２２と、送信部１２３と、受信部１２４と、バッファ１２５と、再送ポインタ（再送ポインタ保持部）１２６と、再送部１２７とを備えている。なお、図２では省略されているが、ワークサーバ１２-2内のエージェント１２０-2も、上記ワークサーバ１２-1内のエージェント１２０-1と同様の構成を備えている。

【0031】

切替部１２１は、ワークサーバ１２-1（内のエージェント１２０-1）をコアサーバ１１-1～１１-4のいずれかと接続する。切替部１２１は、動作しているコアサーバを見つけ、当該コアサーバに接続する。接続しているコアサーバが停止または無応答になった場合、切替部１２１は別のコアサーバに接続し直す。状態テーブル１２２は、ワークサーバ１２-1における各サービスの状態を管理するのに用いられる。

【0032】

送信部１２３は、ワークサーバ１２-1（内のエージェント１２０-1）と接続されているコアサーバ１１-iへのデータ送信、及び当該コアサーバ１１-iから送信されたデータを受信したことを当該コアサーバ１１-iへ通知するための受信応答を行う。受信部１２４は、コアサーバ１１-iから送信されたデータの受信、及び当該コアサーバ１１-iへのデータ送信に対する当該コアサーバ１１-iからの受信応答の受信を行う。

【0033】

バッファ１２５は、送信部１２３からコアサーバ１１-iに送信されたデータをシーケンス番号と対応付けて格納するのに用いられる。バッファ１２５は、再送ポインタ（ＲＰ）１２６と、書き込みポインタ（ＷＰ）１２８（図３参照）とによって管理される。再送ポインタ１２６は、バッファ１２５に格納されている、コアサーバ１１-iが確かに受信した

と確認できないデータのうち、最も古いデータの位置を指す。書き込みポインタは、次にバッファ１２５に書き込むべき送信データの格納先を指す。再送部１２７は、ワークサーバ１２-1（内のエージェント１２０-1）と接続されるコアサーバが切替部１２１によって切り替えられた場合に、それまで当該ワークサーバ１２-1（内のエージェント１２０-1）と接続されていたコアサーバに送信したデータのうち、当該コアサーバで確かに受信したと確認できないデータを、再送ポインタ１２６に従って取り出して、新たにワークサーバ１２-1（内のエージェント１２０-1）と接続されるワークサーバに再送する。

【0034】

図3は、バッファ１２５と再送ポインタ（ＲＰ）１２６及び書き込みポインタ（ＷＰ）１２８との関係を示す。バッファ１２５は、当該バッファ１２５の領域がサイクリックに用いられる循環バッファである。バッファ１２５の領域のうち、再送ポインタ１２６で指定される位置から書き込みポインタ１２８で指定される位置の直前の位置までの領域（図３においてハッチングが施されている部分）のデータが、再送部１２７による再送の対象となる。バッファ１１５-iと再送ポインタ（ＲＰ）１１６-i及び図示せぬ書き込みポインタ（ＷＰ）との関係も同様である。

【0035】

図4は、コアサーバ１１-iが有する状態管理テーブル１１２-iのデータ構造例を示す。図4に示す状態管理テーブル１１２-iの例では、当該状態管理テーブル１１２-iの各エントリには、クラスタシステム内の各ワークサーバ１２-1（＃１），１２-2（＃２）における各サービスの状態が保持される。

【0036】

図5は、ワークサーバ１２-1（内のエージェント１２０-1）が有する状態テーブル１２２のデータ構造例を示す。図5に示す状態テーブル１２２の例では、当該状態テーブル１２２の各エントリには、当該状態テーブル１２２を有するワークサーバ１２-1（＃１）における各サービスの状態が保持される。

【0037】

本実施形態では、従来のクラスタマネージャとは異なって、クラスタ制御を行うコアサーバ１１-1～１１-4とサービス制御を行うワークサーバ１２-1，１２-2とが分離独立して設けられる。このため、コアサーバ１１-1～１１-4とワークサーバ１２-1，１２-2内のエージェント１２０-1，１２０-2とは、制御するものと制御されるものとの関係として、互いに通信する。また、コアサーバ１１-1～１１-4間では、一貫性を保つための通信が行われる。ところが、例えばワークサーバ１２-1がコアサーバ１１-1と接続された状態で、当該コアサーバ１１-1に障害が発生したために、当該ワークサーバ１２-1が他のコアサーバに接続し直された場合には、コアサーバ１１-1～１１-4とワークサーバ１２-1，１２-2とが分離独立していることに起因してデータが失われる虞がある。しかし本実施形態では、以下に述べるワークサーバ１２-1及びコアサーバ１１-1におけるデータ送受信処理により、データが失われるのを防止している。

【0038】

そこで、本実施形態の動作について、ワークサーバ１２-1及びコアサーバ１１-1におけるデータ送受信処理を例に説明する。今、図2に示すように、ワークサーバ１２-1（内のエージェント１２０-1）がコアサーバ１１-1～１１-4のうちのコアサーバ１１-1と接続されているものとする。この状態で、ワークサーバ１２-1内のエージェント１２０-1はコアサーバ１１-1とネットワーク２１を介して通信を行う。なお、ワークサーバ１２-1がコアサーバ１１-2と接続されている場合、両サーバ１２-1，１１-2は同一のサーバコンピュータ１０-2上で動作していることから、両サーバ１２-1，１１-2間の通信はネットワーク２１を介すことなく、当該サーバコンピュータ１０-2内で行われる。勿論、上記両サーバ１２-1，１１-2が同一のサーバコンピュータ１０-2上で動作していても、両サーバ１２-1，１１-2間の通信がネットワーク２１を介して行われる構成としてもよい。

【0039】

まず、ワークサーバ１２-1（内のエージェント１２０-1）に設けられた送信部１２３の

動作について、図６のフローチャートを参照して説明する。送信部１２３は、コアサーバ１１-1に対して送信すべきデータが無い場合（ステップＳ１）、受信応答が必要か否かを判定する（ステップＳ２）。ここでは、コアサーバ１１-1から送信されたデータを受信部１２４が受信してから一定期間を経過しても受信応答を返していない場合、受信応答が必要であると判定される。もし、受信応答が必要な場合、送信部１２３は、受信応答を目的のコアサーバ、即ちワークサーバ１２-1と接続されているコアサーバ１１-1に返す（ステップＳ３）。ここでの受信応答により、ワークサーバ１２-1内の受信部１２４によって受信されながら受信応答が未だ返されていないデータのうち、最も最近に受信されたデータ（つまり最も最近に受信が確認されたデータ）に付されていたシーケンス番号（以下、受信確認シーケンス番号と称する）が通知される。

【0040】

一方、コアサーバ１１-1に対して送信すべきデータが有る場合にも（ステップＳ１）、送信部１２３は受信応答が必要か否かを判定する（ステップＳ４）。ここでの受信応答が必要か否かの判定は、先のステップＳ２とは異なり、まだ受信応答が返されていないデータが有るか否かによって行われる。送信すべきデータが有り、且つ受信応答が必要な場合、送信部１２３は、受信応答付きの送信データをコアサーバ１１-1に送信する（ステップＳ５）。受信応答付きの送信データとは、受信確認シーケンス番号（最も最近に受信が確認されたデータに付されていたシーケンス番号）を含むヘッダ部を持つ送信データをいう。この送信データのヘッダ部には、当該送信データに割り当てられたシーケンス番号（送信データシーケンス番号）も含まれている。この受信応答付きの送信データの送信により、通信回数を減らすことができる。送信部１２３は、送信したデータを当該データに割り当てられたシーケンス番号に対応付けて、書き込みポインタ１２８で指定されるバッファ１２５内の位置に格納する（ステップＳ６）。ここで、シーケンス番号はコアサーバ１１-1へのデータ送信の都度１インクリメントされる。また、書き込みポインタ１２８は、バッファ１２５に送信データを格納する都度、次に送信データを格納すべき位置を指すように更新される。

【0041】

次に、コアサーバ１１-1に対して送信すべきデータは有るが（ステップＳ１）、受信応答の必要がない場合（ステップＳ４）、送信部１２３は、送信データシーケンス番号が付された送信データをコアサーバ１１-1に送信する（ステップＳ７）。そして送信部１２３は、送信したデータを送信データシーケンス番号に対応付けて、書き込みポインタ１２８で指定されるバッファ１２５内の位置に格納する（ステップＳ６）。

【0042】

次に、ワークサーバ１２-1（内のエージェント１２０-1）に設けられた受信部１２４の動作について、図７のフローチャートを参照して説明する。受信部１２４は、ワークサーバ１２-1と接続されているコアサーバ１１-1から何らかの情報を受信すると（ステップＳ１１）、その情報に受信応答が含まれているか否かを判定する（ステップＳ１２）。もし、受信情報に受信応答が含まれている場合、即ち受信応答それ自体または受信応答付き送信データを受信した場合、受信部１２４は、再送ポインタ１２６を受信応答により通知された受信確認シーケンス番号に応じて更新する（ステップＳ１３）。即ち受信部１２４は、受信確認シーケンス番号に一致するシーケンス番号に対応付けてバッファ１２５に格納されている送信データの次の位置を指すように、再送ポインタ１２６を更新する（進める）。このことは、コアサーバ１１-1に送信したデータのうち、当該コアサーバ１１-1で確かに受信したと確認できたデータを、バッファ１２５から削除したことと等価である。もし、更新後の再送ポインタ１２６で指定されるバッファ１２５内の位置（第１の位置と称する）と書き込みポインタ１２８で指定されるバッファ１２５内の位置（第２の位置と称する）とが一致していない場合、バッファ１２５内の第１の位置から第２の位置の直前の位置までに格納されているデータが、現時点において、コアサーバ１１-1で確かに受信したと確認できないデータとなる。

【0043】

次に、ワークサーバ１２-1（内のエージェント１２０-1）と接続されたコアサーバ１１-1に設けられた送信部１１３-1の動作について、図８のフローチャートを参照して説明する。送信部１１３-1は、ワークサーバ１２-1（内のエージェント１２０-1）に対して送信すべきデータが無い場合（ステップＳ２１）、受信応答が必要か否かを判定する（ステップＳ２２）。ここでは、ワークサーバ１２-1から送信されたデータを受信部１１４-1が受信してから一定期間を経過しても受信応答を返していない場合、受信応答が必要であると判定される。もし、受信応答が必要な場合、送信部１１３-1は、受信確認シーケンス番号を同期処理部１１１-1によりシステム内の他の全てのコアサーバ、即ちコアサーバ１１-2～１１-4に通知する（ステップＳ２３）。そして送信部１１３-1は、ワークサーバ１２-1内の送信部１２３によるステップＳ３と同様にして、受信応答を目的のワークサーバ、即ちコアサーバ１１-1と接続されているワークサーバ１２-1に返す（ステップＳ２４）。ここでの受信応答により、コアサーバ１１-1内の受信部１１４-1によって受信されながら受信応答が未だ返されていないデータのうち、最も最近に受信が確認されたデータに付されていたシーケンス番号（つまり受信確認シーケンス番号）が通知される。

【0044】

一方、ワークサーバ１２-1に対して送信すべきデータが有る場合にも（ステップＳ２１）、送信部１１３-1は受信応答が必要か否かを判定する（ステップＳ２５）。ここでの受信応答が必要か否かの判定は、先のステップＳ２２とは異なり、まだ受信応答が返されていないデータが有るか否かによって行われる。送信すべきデータが有り、且つ受信応答が必要な場合、送信部１１３-1は、送信データと受信確認シーケンス番号とを同期処理部１１１-1によりシステム内の他の全てのコアサーバ１１-2～１１-4に通知する（ステップＳ２６）。次に送信部１１３-1は、ワークサーバ１２-1内の送信部１２３によるステップＳ５と同様にして、受信応答付きの送信データを目的のワークサーバ１２-1に返す（ステップＳ２７）。そして送信部１１３-1は、送信部１２３によるステップＳ６と同様にして、送信したデータを当該データに割り当てられたシーケンス番号に対応付けて、書き込みポインタ（図示せず）で指定されるバッファ１１５-1内の位置に格納する（ステップＳ２８）。ここで、シーケンス番号はワークサーバ１２-1へのデータ送信の都度１インクリメントされる。また、書き込みポインタは、バッファ１１５-1に送信データを格納する都度、次に送信データを格納すべき位置を指すように更新される。一方、コアサーバ１１-1以外のコアサーバ、即ちコアサーバ１１-2～１１-4においても、コアサーバ１１-1内の送信部１１３-1から同期処理部１１１-1を通して通知された送信データと受信確認シーケンス番号とから、コアサーバ１１-1におけるのと同様の動作が行われる。これにより、例えば、コアサーバ１１-4内のバッファ１１５-4の内容、バッファ１１５-4用の再送ポインタ１１６-4及びバッファ１１５-4用の書き込みポインタが、コアサーバ１１-1内のバッファ１１５-1の内容、バッファ１１５-1用の再送ポインタ１１６-1及びバッファ１１５-1用の書き込みポインタに一致化される。

【0045】

次に、ワークサーバ１２-1に対して送信すべきデータは有るが（ステップＳ２１）、受信応答の必要がない場合（ステップＳ２５）、送信部１１３-1は送信データシーケンス番号が付された送信データを同期処理部１１１-1によりシステム内の他の全てのコアサーバ１１-2～１１-4に通知する（ステップＳ２９）。次に送信部１１３-1は、この送信データシーケンス番号が付された送信データをコアサーバ１１-1に送信する（ステップＳ３０）。そして送信部１１３-1は、送信したデータを送信データシーケンス番号に対応付けて、書き込みポインタで指定されるバッファ１１５-1内の位置に格納する（ステップＳ２８）。一方、コアサーバ１１-2～１１-4においても、コアサーバ１１-1内の送信部１１３-1から同期処理部１１１-1を通して通知された送信データと送信データシーケンス番号とから、コアサーバ１１-1におけるのと同様の動作が行われる。これにより、例えば、コアサーバ１１-4内のバッファ１１５-4の内容、及びバッファ１１５-4用の書き込みポインタが、コアサーバ１１-1内のバッファ１１５-1の内容、及びバッファ１１５-1用の書き込みポインタに一致化される。

【0046】
次に、コアサーバ１１-1に設けられた受信部１１４-1の動作について、図９のフローチャートを参照して説明する。受信部１１４-1は、コアサーバ１１-1と接続されているワークサーバ１２-1から何らかの情報を受信すると（ステップＳ３１）、その受信情報が受信応答のみ（つまり受信応答それ自体）であるか否かを判定する（ステップＳ３２）。もし、受信応答それ自体を受信した場合、受信部１１４-1は、受信応答により通知された受信確認シーケンス番号を同期処理部１１１-1により他の全てのコアサーバ１１-2～１１-4に通知する（ステップＳ３３）。そして受信部１１４-1は、ワークサーバ１２-1内の受信部１２４によるステップＳ１３と同様にして、再送ポインタ１１６-1を受信応答により通知された受信確認シーケンス番号に応じて更新する（ステップＳ３４）。一方、コアサーバ１１-2～１１-4においても、コアサーバ１１-1内の受信部１１４-1から同期処理部１１１-1を通して通知された受信確認シーケンス番号から、コアサーバ１１-1におけるのと同様の動作が行われる。これにより、例えば、コアサーバ１１-4内の再送ポインタ１１６-4が、コアサーバ１１-1内の再送ポインタ１１６-1に一致化される。
【0047】
一方、受信情報が受信応答それ自体ではない場合、受信部１１４-1は、受信情報が受信応答付き受信データであるか否かを判定する（ステップＳ３５）。もし、受信情報が受信応答付き受信データである場合、受信部１１４-1は、受信データと受信確認シーケンス番号とを同期処理部１１１-1によりシステム内の他の全てのコアサーバ１１-2～１１-4に通知する（ステップＳ３６）。そして受信部１１４-1は、受信情報が受信応答のみの場合と同様に、再送ポインタ１１６-1を受信応答により通知された受信確認シーケンス番号に応じて更新する（ステップＳ３４）。一方、コアサーバ１１-2～１１-4においても、コアサーバ１１-1内の受信部１１４-1から同期処理部１１１-1を通して通知された受信確認シーケンス番号から、コアサーバ１１-1におけるのと同様の動作が行われる。これにより、例えば、コアサーバ１１-4内の再送ポインタ１１６-4が、コアサーバ１１-1内の再送ポインタ１１６-1に一致化される。
【0048】
以上により、コアサーバ１１-1～１１-4内のバッファ１１５-1～１１５-4の内容、バッファ１１５-1～１１５-4用の再送ポインタ１１６-1～１１６-4及びバッファ１１５-1～１１５-4用の書き込みポインタの一貫性が保たれる。
【0049】
このような状態で、ワークサーバ１２-1（内のエージェント１２０-1）と接続されているコアサーバ１１-1の障害が発生したものとする。この場合、コアサーバ１１-1が停止したり、無応答となることから、ワークサーバ１２-1（内のエージェント１２０-1）に設けられた切替部１２１は、ワークサーバ１２-1と接続するコアサーバをコアサーバ１１-1から他の正常に動作する例えばコアサーバ１１-4に切り替える。
【0050】
すると、ワークサーバ１２-1（内のエージェント１２０-1）に設けられた再送部１２７は、再送ポインタ１２６で指定されるバッファ１２５内の第１の位置から書き込みポインタ１２８で指定されるバッファ１２５内の第２の位置の直前の位置までに格納されているデータ（図３においてハッチングされた領域のデータ）を、前にワークサーバ１２-1と接続されていたコアサーバ１１-1で確かに受信したと確認できないデータであるとして、切替部１２１による切替でワークサーバ１２-1に新たに接続されたコアサーバ１１-4に送信（再送）する。このデータがコアサーバ１１-4で受信されると、そのデータがクラスタシステム内の他のコアサーバに同期処理部１１１-4により通知され、受信データに関し、コアサーバ１１-1～１１-4間で一貫性が保たれる。
【0051】
一方、ワークサーバ１２-1と新たに接続されたコアサーバ１１-4では、再送部１１７-1が、上記ワークサーバ１２-1内の再送部１２７と同様の送信動作（再送動作）を行う。即ち再送部１１７-1は、再送ポインタ１１６-4で指定されるバッファ１１５-4内の第１の位

置から書き込みポインタで指定されるバッファ１１５-4内の第２の位置の直前の位置までに格納されているデータ、つまりワークサーバ１２-1が前にコアサーバ１１-1と接続されていた状態で確かに受信したと確認できないデータを、当該ワークサーバ１２-1に送信（再送）する。

【0052】

これらの、ワークサーバ１２-1内の切替部１２１及び再送部１２７と、コアサーバ１１-4内の再送部１１７-4の動作によって、コアサーバ１１-1～１１-4のうちのいずれが障害で停止した場合（ここではコアサーバ１１-1が停止）でも、スムーズに切り替えが行われ、且つデータのロスが発生しないで済む。つまり、これらの機能により、コアサーバとワークサーバとを分離したことに起因して発生する可能性のある問題を解消できる。

【0053】

また本実施形態では、各ワークサーバ１２-1，１２-2（のエージェント１２０-1，１２０-2）には、当該ワークサーバ１２-1，１２-2におけるサービスの状態を保持する状態テーブル（ワークサーバ１２-1では状態テーブル１２２）が設けられている。このため、カーネル１１０の再起動後に、当該カーネル１１０が各ワークサーバ１２-1，１２-2の状態テーブルの内容（つまりワークサーバ１２-1，１２-2におけるサービスの状態）を収集することで、クラスタシステム内の全てのワークサーバ１２-1，１２-2におけるサービスの状態を復元する。具体的には、カーネル１１０を構成するコアサーバ１１-1～１１-4のうち、ワークサーバ１２-1，１２-2と接続されているコアサーバ１１-2，１１-3がワークサーバ１２-1，１２-2におけるサービスの状態を収集する。収集したサービスの状態は、各コアサーバ１１-1～１１-4の同期処理部１１１-1～１１１-4により各コアサーバ１１-1～１１-4に通知され、各コアサーバ１１-1～１１-4内の状態管理テーブル１１２-1～１１２-4に反映される。これにより、カーネル１１０はクラスタの制御を再開できる。つまり本実施形態においては、カーネル１１０を一時的に停止し、その後に再起動することが可能になる。これにより、コアサーバのバージョンアップ、コアサーバの移動などが可能となり、クラスタ構成における柔軟性が増す。

【0054】

上記実施形態では、クラスタシステムを構成するサーバコンピュータ１０-1～１０-4のうちのサーバコンピュータ１０-1，１０-4では、コアサーバ及びワークサーバのうちのコアサーバのみが稼働し、残りのサーバコンピュータ１０-2，１０-3では、コアサーバ及びワークサーバが共に稼働することを前提としている。しかし、クラスタシステムを構成する全てのサーバコンピュータでコアサーバ及びワークサーバの両方が稼働する構成であっても、一部のコンピュータ群ではコアサーバのみが稼働し、残りのコンピュータ群ではワークサーバのみが稼働する構成であっても構わない。つまり、クラスタシステムを構成するサーバコンピュータは、コアサーバのみが稼働するサーバコンピュータ、ワークサーバのみが稼働するサーバコンピュータ、或いはコアサーバとワークサーバの両方が稼働するサーバコンピュータの何れかとなる。コアサーバとワークサーバをどのように配置するかは、クラスタ構成設計の基本である。クラスタシステムに必要な可用性や拡張性、それぞれのコンピュータの用途、性能、信頼性などを考慮し、適切な配置を決定する必要がある。以下、上記第１の実施形態とは異なるクラスタ構成について説明する。

【0055】

［第２の実施形態］

図１０は本発明の第２の実施形態に係るクラスタシステムの構成を示すブロック図である。図１０のクラスタシステムは、Ｎ台のサーバコンピュータ３０-1（＃１）～３０-N（＃Ｎ）から構成される。この例では、Ｎ台のサーバコンピュータ３０-1（＃１）～３０-N（＃Ｎ）の全てで、それぞれコアサーバ３１-1（＃１）～３１-N（＃Ｎ）とワークサーバ３２-1（＃１）～３２-N（＃Ｎ）とが稼働する。

【0056】

このように図１０のクラスタシステムは、全てのサーバコンピュータ＃１～＃Ｎでコアサーバ＃１～＃Ｎによるクラスタ制御に加えて、ワークサーバ＃１～＃Ｎによるクラスタ

システムのサービスが実行される、最密クラスタ構成をとる。このため、サーバコンピュータ＃１～＃Ｎには、そのサービスに必要なリソースと、それに見合ったハードウェア／ソフトウェアがに必要になるものの、サービスに関しては、Ｎ個のワークサーバ＃１～＃Ｎ間（つまりシステムを構成するＮ台のサーバコンピュータ＃１～＃Ｎ）でのフェイルオーバによる可用性の向上が期待できる。また、カーネルに関しては、Ｎ個のコアサーバ＃１～＃Ｎ（つまりシステムを構成するＮ台のサーバコンピュータ＃１～＃Ｎ）による２／３定足数アルゴリズムによる可用性の向上が期待できる。この図１０のクラスタシステムの構成（最密クラスタ構成）は最大の可用性を得ることができる。但し、コンピュータ数Ｎが大きい構成（例えばＮ＞１６）は、カーネルの可用性が必要以上に高く設定されることになるため好ましくない。なお、コアサーバ＃１～＃Ｎ及びワークサーバ＃１～＃Ｎの構成と動作は、前記第１の実施形態におけるコアサーバ及びワークサーバと同様である。

【0057】

［第３の実施形態］（完全分離クラスタ構成）

図１１は本発明の第３の実施形態に係るクラスタシステムの構成を示すブロック図である。図１１のクラスタシステムは、Ｎ＋Ｋ台のサーバコンピュータ４０-1（＃１）～４０-N（＃Ｎ），４０-(N+1)（＃Ｎ＋１）～４０-(N+K)（＃Ｎ＋Ｋ）から構成される。この例では、Ｎ台のサーバコンピュータ４０-1（＃１）～４０-N（＃Ｎ）で、それぞれワークサーバ４２-1（＃１）～４２-N（＃Ｎ）が稼働し、残りのＫ台のサーバコンピュータ４０-(N+1)（＃Ｎ＋１）～４０-(N+K)（＃Ｎ＋Ｋ）でコアサーバ４１-1（＃１）～４１-K（＃Ｋ）が稼働する。

【0058】

このように図１１のクラスタシステムは、ワークサーバ専用のＮ台のサーバコンピュータ＃１～＃Ｎとコアサーバ専用のＫ台のサーバコンピュータ＃Ｎ＋１～＃Ｎ＋Ｋとによる完全分離クラスタ構成をとる。このため、サーバコンピュータ＃１～＃Ｎ＋Ｋには、それぞれの役割に見合ったハードウェア／ソフトウェアが必要となる。この図１１のクラスタシステムでは、サービスに関しては、Ｎ個のワークサーバ間でのフェイルオーバによる可用性の向上が期待できる。また、カーネルに関しては、Ｋ個のコアサーバ＃１～＃Ｋによる２／３定足数アルゴリズムによる可用性の向上が期待できる。

【0059】

また、図１１のクラスタシステムは１つのサーバコンピュータ上でコアサーバ及びワークサーバの一方だけが稼働する完全分離クラスタ構成をとることから、ワークサーバ＃１～＃Ｎで動いているリソースの障害が、コアサーバ＃１～＃Ｋに影響しにくく、より信頼性の高いクラスタ構成を実現できる。また、ワークサーバ数Ｎが多い（Ｎ＞１６）システムでも、コアサーバ数Ｋを適切な値（４≦Ｋ≦１６）に調整することにより、最適なクラスタシステムを構築できる。また完全分離クラスタ構成は、最初は少ない個数のワークサーバでクラスタシステムを構成し、順次ワークサーバを追加していく拡張性の高いクラスタシステムを構築したい場合にも適している。

【0060】

なお、本発明は、上記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態に開示されている複数の構成要素の適宜な組み合せにより種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。更に、異なる実施形態に亘る構成要素を適宜組み合せてもよい。

【図面の簡単な説明】

【0061】

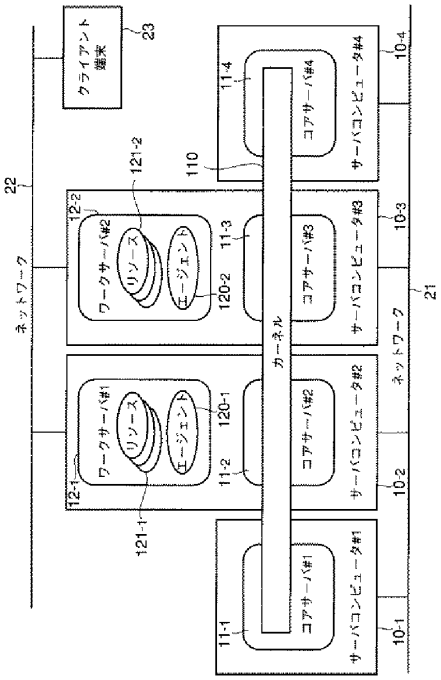
【図1】本発明の第１の実施形態に係るクラスタシステムの構成を示すブロック図。

【図2】図１のクラスタシステム内のコアサーバ及びワークサーバの構成を、コアサーバ１１-1，１１-4及びワークサーバ１２-1について示すブロック図。

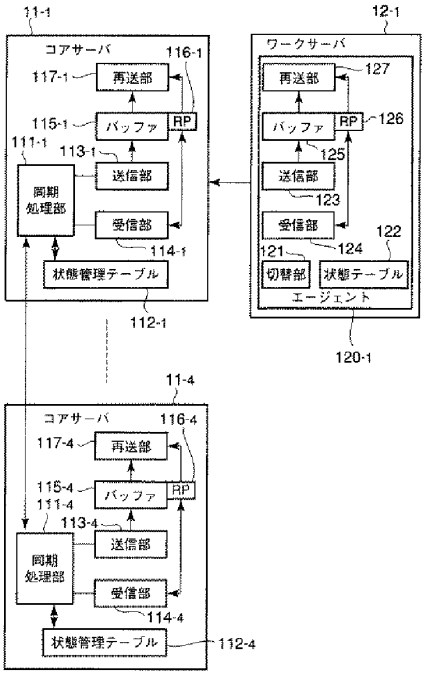
【図3】バッファ１２５と再送ポインタ（ＲＰ）１２６及び書き込みポインタ（ＷＰ）１２８との関係を示す図。

【図4】コアサーバ１１-iが有する状態管理テーブル１１２-iのデータ構造例を示す図。
【図5】ワークサーバ１２-1が有する状態テーブル１２２のデータ構造例を示す図。
【図6】ワークサーバ１２-1に設けられた送信部１２３の動作手順を示すフローチャート
。
【図7】ワークサーバ１２-1に設けられた受信部１２４の動作手順を示すフローチャート
。
【図8】コアサーバ１１-1に設けられた送信部１１３-1の動作手順を示すフローチャート
。
【図9】コアサーバ１１-1に設けられた受信部１１４-1の動作手順を示すフローチャート
。
【図10】本発明の第２の実施形態に係るクラスタシステムの構成を示すブロック図。
【図11】本発明の第３の実施形態に係るクラスタシステムの構成を示すブロック図。
【符号の説明】
【0062】
　１０-1～１０-4，３０-1～３０-N，４０-1～４０-(N+K)…サーバコンピュータ、１１-1～１１-4，３１-1～３１-N，４１-1～４１-K…コアサーバ、１２-1～１２-4，３２-1～３２-N，４２-1～４２-N…ワークサーバ、２１，２２ネットワーク、２３…クライアント端末、１１１-1～１１１-4…同期処理部、１１２-1～１１２-4…状態管理テーブル（第２の状態保持手段）、１１３-1～１１３-4…送信部、１１４-1～１１４-4…受信部、１１５-1～１１５-4…バッファ（第２のバッファ）、１１６-1～１１６-4…再送ポインタ（ＲＰ、第２の再送ポインタ保持手段）、１１７-1～１１７-4…再送部（第２の再送手段）、１２１…切替部、１２２…状態テーブル（第１の状態保持手段）、１２３…送信部、１２４…受信部、１２５…バッファ（第２のバッファ）、１２６…再送ポインタ（ＲＰ、第１の再送ポインタ保持手段）、１２７…再送部（第１の再送手段）、１２８…書き込みポインタ（ＷＰ）。

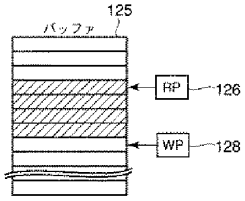
【図1】



【図2】



【図3】



【図4】

112-i(i=1～4)

状態管理テーブル

サービス#	サービス状態	
	ワークサーバ#1	ワークサーバ#2
サービス#1	状態A	状態B
サービス#2	状態B	状態C
⋮	⋮	⋮

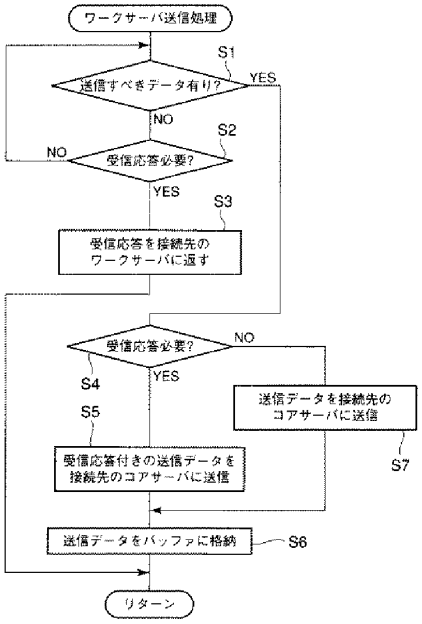
【図5】

122

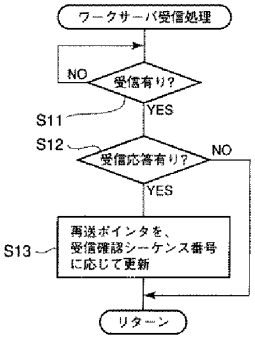
状態テーブル

サービス#	サービス状態
サービス#1	状態A
サービス#2	状態B
⋮	⋮

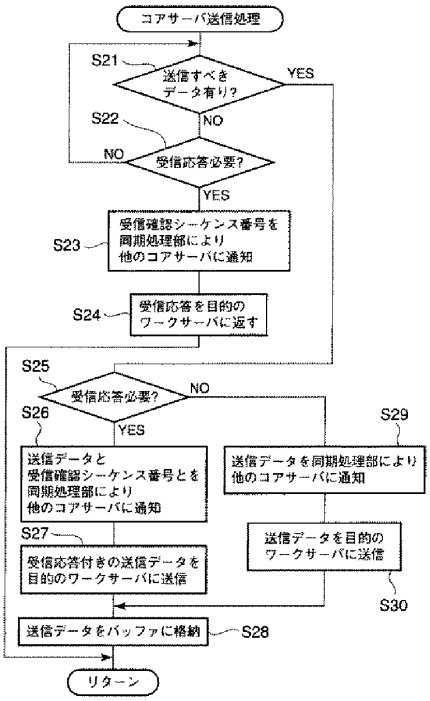
【図6】



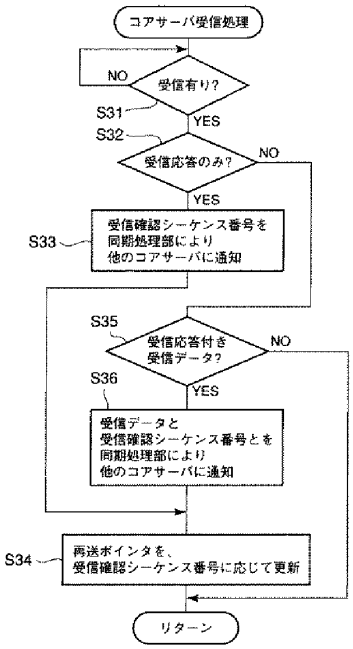
【図7】



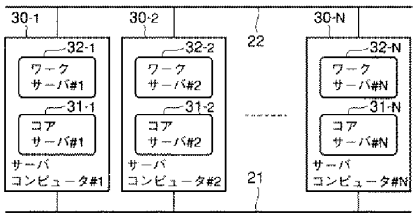
【図8】



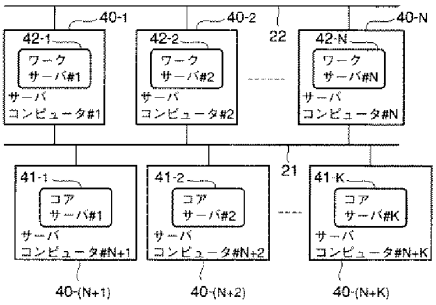
【図9】


コアサーバ受信処理
受信有り?
NO
YES
S31
S32
受信応答のみ?
NO
YES
S33
受信確認シーケンス番号を
同期処理部により
他のコアサーバに通知
S35
受信応答付き
受信データ?
NO
YES
S36
受信データと
受信確認シーケンス番号とを
同期処理部により
他のコアサーバに通知
S34
再送ポインタを、
受信確認シーケンス番号に応じて更新
リターン

【図10】


30-1
30-2
30-N
22
32-1
32-2
32-N
ワーク
サーバ#1
ワーク
サーバ#2
ワーク
サーバ#N
31-1
31-2
31-N
コア
サーバ#1
コア
サーバ#2
コア
サーバ#N
サーバ
コンピュータ#1
サーバ
コンピュータ#2
サーバ
コンピュータ#N
21

【図11】


40-1
40-2
40-N
22
42-1
42-2
42-N
ワーク
サーバ#1
ワーク
サーバ#2
ワーク
サーバ#N
サーバ
コンピュータ#1
サーバ
コンピュータ#2
サーバ
コンピュータ#N
21
41-1
41-2
41-K
コア
サーバ#1
コア
サーバ#2
コア
サーバ#K
サーバ
コンピュータ#N+1
サーバ
コンピュータ#N+2
サーバ
コンピュータ#N+K
40-(N+1)
40-(N+2)
40-(N+K)

(72)発明者　遠藤　浩太郎
東京都府中市東芝町１番地　株式会社東芝府中事業所内

Ｆターム(参考) 5B045 GG01　GG11　JJ02　JJ08　JJ26　JJ37　JJ45

Electronic Acknowledgement Receipt

EFS ID:	24120587
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Debora Swies
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	18-NOV-2015
Filing Date:	14-FEB-2014
Time Stamp:	14:37:26
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		ZTANN_002P1C3_IDS.pdf	137896 9953cb3bcd99a1f72220e8c7519edde569868e33	yes	3

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	2
	Information Disclosure Statement (IDS) Form (SB08)	3	3

Warnings:					
Information:					
2	Foreign Reference	JP2005_63033_wAbs.pdf	1743275 5201a1a40fd37445bfdf66027275ac2c82acf053	no	20
Warnings:					
Information:					
		Total Files Size (in bytes):	1881171		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

INFORMATION DISCLOSURE STATEMENT

Inventor	:	Tannenbaum, Zvi
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Asres, Hermon
Art Unit	:	2449
Conf. No.	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Pursuant to 37 CFR 1.56, an Information Disclosure Statement listing references is provided herewith. Copies of any listed foreign and non-patent literature references are being submitted. The foreign reference includes an English abstract, but no representation is made as to its accuracy.

If the Examiner would like additional information regarding these references or if anything is unclear, the Examiner is invited to contact the undersigned for assistance.

No Disclaimers

To the extent that anything in the Information Disclosure Statement or the listed references could be construed as a disclaimer of any subject matter supported by the present application, Applicant hereby rescinds and retracts such disclaimer.

Timing of Disclosure

This Information Disclosure Statement is being filed before the receipt of a First Office Action on the merits, and presumably no fee is required. If a First Office Action on the merits was mailed before the mailing date of this Statement, the Commissioner is authorized to charge the fee set forth in 37 CFR 1.17(p) to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: November 18, 2015

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

22047727

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3

CONFIRMATION NO. 6874

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

PUBLICATION NOTICE

OC000000072462299

Title:CLUSTER COMPUTING

Publication No.US-2014-0372586-A1
Publication Date:12/18/2014

NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Office of Public Records. The Office of Public Records can be reached by telephone at (703) 308-9726 or (800) 972-6382, by facsimile at (703) 305-8759, by mail addressed to the United States Patent and Trademark Office, Office of Public Records, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently http://pair.uspto.gov/. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

Doc Code: TRAN.LET

Document Description: Transmittal Letter

PTO/SB/21 (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	3

Application Number	14/181,112
Filing Date	February 14, 2014
First Named Inventor	Tannenbaum, Zvi
Art Unit	2449 - Confirmation No. 6874
Examiner Name	Asres, Hermon
Attorney Docket Number	ZTANN.002P1C3

ENCLOSURES *(Check all that apply)*

- ☐ Fee Transmittal Form
  - ☐ Fee Attached
- ☐ Amendment/Reply
  - ☐ After Final
  - ☐ Affidavits/declaration(s)
- ☐ Extension of Time Request
- ☐ Express Abandonment Request
- ☐ Information Disclosure Statement
- ☐ Certified Copy of Priority Document(s)
- ☐ Reply to Missing Parts/ Incomplete Application
  - ☐ Reply to Missing Parts under 37 CFR 1.52 or 1.53
- ☐ Drawing(s)
- ☐ Licensing-related Papers
- ☐ Petition
- ☐ Petition to Convert to a Provisional Application
- ☐ Power of Attorney, Revocation Change of Correspondence Address
- ☐ Terminal Disclaimer
- ☐ Request for Refund
- ☐ CD, Number of CD(s) _____
  - ☐ Landscape Table on CD
- ☐ After Allowance Communication to TC
- ☐ Appeal Communication to Board of Appeals and Interferences
- ☐ Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- ☐ Proprietary Information
- ☐ Status Letter
- ☑ Other Enclosure(s) (please Identify below):
  Inventor Declarations

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Knobbe, Martens, Olson & Bear, LLP		
Signature	/Lance D. Smemoe/		
Printed name	Lance D. Smemoe		
Date	December 4, 2014	Reg. No.	66,152

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature			
Typed or printed name		Date	

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/AIA/01 (06-12)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention CLUSTER COMPUTING

As the below named inventor, I hereby declare that:

This declaration is directed to:

☐ The attached application, or

☒ United States application or PCT international application number 14/181,112 filed on February 14, 2014.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Zvi Tannenbaum Date (Optional): 3/17/14

Signature:

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 1 minute to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PTO/AIA/01 (06-12)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

Title of Invention	CLUSTER COMPUTING

As the below named inventor, I hereby declare that:

This declaration is directed to:

☐ The attached application, or

☒ United States application or PCT international application number 14/181,112 filed on February 14, 2014.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Dean E. Dauger Date (Optional): 3/17/14

Signature: Dean E. Dauger

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

This collection of information is required by 35 U.S.C. 115 and 37 CFR 1.63. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 1 minute to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Electronic Acknowledgement Receipt

EFS ID:	20871208
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Gustavo Lopez
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	04-DEC-2014
Filing Date:	14-FEB-2014
Time Stamp:	18:15:08
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		Declarations_ztann_002p1c3.pdf	3082228 b0382db3d099ee16ca05a633056923a422ca1c0f	yes	3

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	1
	Oath or Declaration filed	2	3
Warnings:			
Information:			
	Total Files Size (in bytes):	3082228	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3

CONFIRMATION NO. 6874

POA ACCEPTANCE LETTER

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

OC000000070728457

Date Mailed: 09/16/2014

NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 09/10/2014.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

/dtdinh/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 14/181,112

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	70		N/A	
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	300		N/A	
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	360		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	10 minus 20 =	*	x 40 =	0.00	OR		
INDEPENDENT CLAIMS (37 CFR 1.16(h))	1 minus 3 =	*	x 210 =	0.00			
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			0.00			
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				0.00			
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	730		TOTAL	

APPLICATION AS AMENDED - PART II

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
AMENDMENT A	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
AMENDMENT A	Application Size Fee (37 CFR 1.16(s))									
AMENDMENT A	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)	OR	RATE($)	ADDITIONAL FEE($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
AMENDMENT B	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
AMENDMENT B	Application Size Fee (37 CFR 1.16(s))									
AMENDMENT B	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3

CONFIRMATION NO. 6874

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

NOTICE

OC000000070648085

Date Mailed: 09/11/2014

INFORMATIONAL NOTICE TO APPLICANT

Applicant is notified that the above-identified application contains the deficiencies noted below. No period for reply is set forth in this notice for correction of these deficiencies. However, if a deficiency relates to the inventor's oath or declaration, the applicant must file an oath or declaration in compliance with 37 CFR 1.63, or a substitute statement in compliance with 37 CFR 1.64, executed by or with respect to each actual inventor no later than the expiration of the time period set in the "Notice of Allowability" to avoid abandonment. See 37 CFR 1.53(f).

The item(s) indicated below are also required and should be submitted with any reply to this notice to avoid further processing delays.

- A properly executed inventor's oath or declaration has not been received for the following inventor(s):
  Zvi Tannenbaum
  Dean E. Dauger

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
14/181,112	02/14/2014	2447	800	ZTANN.002P1C3	10	1

CONFIRMATION NO. 6874

UPDATED FILING RECEIPT

OC000000070648084

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

Date Mailed: 09/11/2014

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Zvi Tannenbaum, Newport Beach, CA;
Dean E. Dauger, Huntington Beach, CA;

Applicant(s)

Advanced Cluster Systems, Inc., Aliso Viejo, CA

Assignment For Published Patent Application

Advanced Cluster Systems, Inc., Aliso Viejo, CA

Power of Attorney: None

Domestic Priority data as claimed by applicant

This application is a CON of 13/423,063 03/16/2012 PAT 8676877
which is a CON of 12/040,519 02/29/2008 PAT 8140612
which is a CIP of 11/744,461 05/04/2007 PAT 8082289
which claims benefit of 60/813,738 06/13/2006
and claims benefit of 60/850,908 10/11/2006

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

Permission to Access - A proper **Authorization to Permit Access to Application by Participating Offices** (PTO/SB/39 or its equivalent) has been received by the USPTO.

If Required, Foreign Filing License Granted: 03/04/2014

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/181,112**

Projected Publication Date: 12/18/2014

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****
Title

CLUSTER COMPUTING

Preliminary Class

709

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

Doc Code: PA..
Document Description: Power of Attorney

PTO/AIA/82B (07-13)
Approved for use through 11/30/2014. OMB 0651-0051
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

POWER OF ATTORNEY BY APPLICANT

I hereby revoke all previous powers of attorney given in the application identified in either the attached transmittal letter or the boxes below.

Application Number	Filing Date

(Note: The boxes above may be left blank if information is provided on form PTO/AIA/82A.)

[X] I hereby appoint the Patent Practitioner(s) associated with the following Customer Number as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above: 20995

OR

[ ] I hereby appoint Practitioner(s) named in the attached list (form PTO/AIA/82C) as my/our attorney(s) or agent(s), and to transact all business in the United States Patent and Trademark Office connected therewith for the patent application referenced in the attached transmittal letter (form PTO/AIA/82A) or identified above. (Note: Complete form PTO/AIA/82C.)

Please recognize or change the correspondence address for the application identified in the attached transmittal letter or the boxes above to:

[X] The address associated with the above-mentioned Customer Number

OR

[ ] The address associated with Customer Number:

OR

Firm or Individual Name					
Address					
City		State		Zip	
Country					
Telephone		Email			

I am the Applicant (if the Applicant is a juristic entity, list the Applicant name in the box):

Advanced Cluster Systems, Inc.

[ ] Inventor or Joint Inventor (title not required below)

[ ] Legal Representative of a Deceased or Legally Incapacitated Inventor (title not required below)

[X] Assignee or Person to Whom the Inventor is Under an Obligation to Assign (provide signer's title if applicant is a juristic entity)

[ ] Person Who Otherwise Shows Sufficient Proprietary Interest (e.g., a petition under 37 CFR 1.46(b)(2) was granted in the application or is concurrently being filed with this document) (provide signer's title if applicant is a juristic entity)

SIGNATURE of Applicant for Patent

The undersigned (whose title is supplied below) is authorized to act on behalf of the applicant (e.g., where the applicant is a juristic entity).

Signature		Date (Optional)	
Name	Zvi Tannenbaum		
Title	Chief Executive Officer		

NOTE: Signature - This form must be signed by the applicant in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. If more than one applicant, use multiple forms.

[X] Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

INFORMATION DISCLOSURE STATEMENT

Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Not yet assigned
Art Unit	:	2447
Conf. No.	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

References and Listing

Submitted herewith in the above-identified application is an Information Disclosure Statement listing references for consideration. Listed references are of record in U.S. Patent Application No. 13/423,063, filed March 16, 2012, which is the parent of this Continuation application, and is relied upon for an earlier filing date under 35 USC 120. Copies of the references are not submitted pursuant to 37 CFR 1.98(d).

Citation numbers 20-23 identify commonly owned issued patents.[1]

Timing of Disclosure

This Information Disclosure Statement is being filed before the receipt of a First Office Action on the merits, and presumably no fee is required. If a First Office Action on the merits was mailed before the mailing date of this Statement, the Commissioner is authorized to charge the fee set forth in 37 CFR 1.17(p) to Deposit Account No. 11-1410.

[1] Please consider whether the disclosure, claims, and prosecution history of these items are relevant to the patentability of the present application. The Examiner has access to these materials through the PTO computer systems. If additional copies are desired, please notify Applicants through their attorneys. Signature on the attached Supplemental IDS Form PTO/SB-08 indicates consideration of the above, including file histories.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: September 10, 2014

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

18847890

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 1 OF 3	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	1	5,423,046	06-06-1995	Nunnelley, et al.	
	2	5,881,315	03-09-1999	Cohen	
	3	5,930,768	07-1999	Hooban	
	4	6,108,699	08-22-2000	Moiin	
	5	6,202,080	03-13-2001	Lu et al.	
	6	6,546,403	04-08-2003	Carlson, Jr. et al.	
	7	6,578,068	06-10-2003	Bowman-Amuah	
	8	6,751,698	06-15-2004	Deneroff et al.	
	9	6,782,537	08-24-2004	Blackmore et al.	
	10	6,968,335	11-22-2005	Bayliss et al.	
	11	7,093,004	08-15-2006	Bernardin et al.	
	12	7,136,924	11-14-2006	Dauger	
	13	7,174,381	02-06-2007	Gulko et al.	
	14	7,249,357	07-24-2007	Landman et al.	
	15	7,334,232	02-19-2008	Jacobs et al.	
	16	7,437,460	10-14-2008	Chidambaran et al.	
	17	7,472,193	12-30-2008	Dauger	
	18	7,554,949	06-30-2009	Chen	
	19	7,937,455	05-03-2011	Saha et al.	
	20	8,082,289 (Attorney Docket No. ZTANN.002A)	12-20-2011	Tannenbaum et al.	U.S. Application No. 11/744,461 and its entire prosecution history**
	21	8,140,612 (Attorney Docket No. ZTANN.002CP1)	03-20-2012	Tannenbaum et al.	U.S. Application No. 12/040,519 and its entire prosecution history**
	22	8,402,083 (Attorney Docket No. ZTANN.002CP1C1)	03-19-2013	Tannenbaum et al.	U.S. Application No. 12/465,519 and its entire prosecution history**
	23	8,676,877 (Attorney Docket No. ZTANN.002P1C2)	03-18-2014	Tannenbaum et al.	U.S. Application No. 13/423,063 and its entire prosecution history**
	24	2002/0049859	04-25-2002	Bruckert et al.	

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

PTO/SB/08 Equivalent

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 2 OF 3	Attorney Docket No.	ZTANN.002P1C3

U.S. PATENT DOCUMENTS

Examiner Initials	Cite No.	Document Number *Number - Kind Code (if known)* Example: 1,234,567 B1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear
	25	2003/0005266	01-02-2003	Dauger	
	26	2003/0051062	03-13-2003	Circenis	
	27	2003/0135621	07-17-2003	Romagnoli	
	28	2003/0195931	10-16-2003	Dauger	
	29	2003/0195938	10-16-2003	Howard et al.	
	30	2004/0110209	06-10-2004	Yokota, et al.	
	31	2004/0157203	08-12-2004	Dunk	
	32	2004/0252710	03-17-2005	Barsness et al.	
	33	2005/0021751	01-27-2005	Block et al.	
	34	2005/0038852	02-17-2005	Howard	
	35	2005/0060237	03-17-2005	Barsness et al.	
	36	2005/0076105	04-07-2005	Keohane et al.	
	37	2005/0108394	05-19-2005	Braun et al.	
	38	2005/0180095	08-18-2005	Ellis	
	39	2006/0053216	03-09-2006	Deokar et al.	
	40	2006/0106931	05-18-2006	Richoux	
	41	2007/0044099	02-22-2007	Rajput	
	42	2007/0073705	03-29-2007	Gray	
	43	2008/0250347	10-09-2008	Gray et al.	
	44	2008/0281997	11-13-2008	Archer et al.	

FOREIGN PATENT DOCUMENTS

Examiner Initials	Cite No.	Foreign Patent Document *Country Code-Number-Kind Code* Example: JP 1234567 A1	Publication Date MM-DD-YYYY	Name of Patentee or Applicant	Pages, Columns, Lines Where Relevant Passages or Relevant Figures Appear	**T[1]**
	45	JP 08-87473 A	02-04-1996	Toshiba Corp		
	46	JP 2002117010 w/ English Abstract	04-19-2002	International Business Mach Corp		

Examiner Signature	Date Considered

***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

T[1] - Place a check mark in this area when an English language Translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT	Application No.	14/181,112
	Filing Date	02-14-2014
	First Named Inventor	Zvi Tannenbaum
	Art Unit	2447
(Multiple sheets used when necessary)	Examiner	Not yet assigned
SHEET 3 OF 3	Attorney Docket No.	ZTANN.002P1C3

NON PATENT LITERATURE DOCUMENTS			
Examiner Initials	Cite No.	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc.), date, page(s), volume-issue number(s), publisher, city and/or country where published.	T[1]
	47	"gridMathematica 1.1: Grid Computing Gets a Speed Boost from Mathematica 5", The Mathematica Journal, Vol. 9, No. 2, 2004.	
	48	"Wolfram gridMathematica", Wolfram Research, Inc., 2007.	
	49	CARNS et al., "An Evaluation of Message Passing Implementations on Beowulf Workstations", Aerospace Conference, March 6-13, 1999, IEEE 1999, Vol. 5, pp. 41-54.	
	50	DAUGER et al., "Plug-and-Play Cluster Computing using Mac OS X", IEEE International Conference, December 1-4, 2003, Paper appears in Cluster Computing January 8, 2004, pp. 430-435.	
	51	DAUGER et al., Plug-and Play Cluster Computing: High-Performance Computing For the Mainstream", IEEE Computing in Science and Engineering, March/April 2005, pp 27-33.	
	52	HAMSCHER et al., "Evaluation of Job-Scheduling Strategies for grid Computing", LNCS: Lecture Notes in Computer Science, 2000, pp. 191-202.	
	53	International Search Report dated September 11, 2008, International Application No. PCT/US07/70585 (Attorney Docket No. ZTANN.002VPC).	
	54	JAHANZEB et al., "Libra: a computational economy-based job scheduling system for clusters", Software Practice and Experience, February 24, 2004, Vol. 34, pp. 573-590.	
	55	JAIN et al., "Data Clustering: A Review", ACM Computing Surveys, September 1999, Vol. 31, No. 3.	
	56	KEPNER, Jeremy et al., "Parallel Matlab: The Next Generation", August 20, 2004, MIT Lincoln Laboratory, Lexington, MA.	
	57	KIM, Hahn et al., "Introduction to Parallel Programming and pMatlab v2.0", 2011, MIT Lincoln Laboratory, Lexington, MA.	
	58	MAEDER, R., "Mathematica Parallel Computing Toolkit: Unleash the Power of Parallel Computing", Wolfram Research, January 2005.	
	59	TEPENEU and IDA, "MathGridLink – A bridge between Mathematica and the Grid", Nippon Sofutowea Kagakkai Taikai Ronbunshu, 2003, Vol. 20, pp. 74-77.	
	60	Wolfram, Stephen: THE MATHEMATICA BOOK 5TH Edition; Wolfram Research, Inc. 2003	

** Signature indicates consideration of publication and file history. The Examiner has access to these materials through the PTO computer systems. If additional copies are desired, please notify Applicants through their attorneys

18642854

Examiner Signature	Date Considered
***Examiner:** Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through citation if not in conformance and not considered. Include copy of this form with next communication to applicant.	

T[1] - Place a check mark in this area when an English language Translation is attached.

Please Direct All Correspondence to Customer Number 20995

RESCISSION OF ANY PRIOR DISCLAIMERS AND REQUEST TO REVISIT ART

Inventor	:	Zvi Tannenbaum
App. No	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Not yet assigned
Art Unit	:	2447
Conf #	:	6874

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

The claims of the present application are different and possibly broader in scope than the claims pursued in the parent application(s). To the extent any prior amendments or characterizations of the scope of any claim or referenced art could be construed as a disclaimer of any subject matter supported by the present disclosure, Applicant hereby rescinds and retracts such disclaimer. Accordingly, the references previously considered in the parent application(s) may need to be re-visited.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Knobbe, Martens, Olson & Bear, LLP

Respectfully submitted,

Dated: September 10, 2014

/Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

18848199

Electronic Acknowledgement Receipt

EFS ID:	20098469
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Christina Graul
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	10-SEP-2014
Filing Date:	14-FEB-2014
Time Stamp:	14:29:32
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		ZTANN_002P1C3_POA.pdf	1234235 1cf2dc6e5e8724b22ad22abab8083e4fb21b11bb	yes	2

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	1
	Power of Attorney	2	2

Warnings:

Information:

2		ZTANN_002P1C3_IDS.pdf	349269 9d858e89175812e484abc68ebbd2af5dbb6d8835	yes	5

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Transmittal Letter	1	2
	Information Disclosure Statement (IDS) Form (SB08)	3	5

Warnings:

Information:

3	Miscellaneous Incoming Letter	ZTANN_002P1C3_HAKIM.pdf	16435 7b341654385cddf1926d329758aedbe520ff393b	no	1

Warnings:

Information:

Total Files Size (in bytes):	1599939

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Doc Code: PA..
Document Description: Power of Attorney

PTO/AIA/82A (07-13)
Approved for use through 11/30/2014. OMB 0651-0051
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FOR POWER OF ATTORNEY TO ONE OR MORE REGISTERED PRACTITIONERS

NOTE: This form is to be submitted with the Power of Attorney by Applicant form (PTO/AIA/82B) to identify the application to which the Power of Attorney is directed, in accordance with 37 CFR 1.5, unless the application number and filing date are identified in the Power of Attorney by Applicant form. If neither form PTO/AIA/82A nor form PTO/AIA82B identifies the application to which the Power of Attorney is directed, the Power of Attorney will not be recognized in the application.

Application Number	14/181,112
Filing Date	February 14, 2014
First Named Inventor	Zvi Tannenbaum
Title	CLUSTER COMPUTING
Art Unit	2447 / Confirmation no. 6874
Examiner Name	Not yet assigned
Attorney Docket Number	ZTANN.002P1C3

SIGNATURE of Applicant or Patent Practitioner

Signature	/Lance D. Smemoe/	Date (Optional)	
Name	Lance D. Smemoe	Registration Number	66,152
Title (if Applicant is a juristic entity)			
Applicant Name (if Applicant is a juristic entity)			

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. If more than one applicant, use multiple forms.

☑ *Total of 1 forms are submitted.

This collection of information is required by 37 CFR 1.131, 1.32, and 1.33. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 3 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Doc Code: TRAN.LET

Document Description: Transmittal Letter

PTO/SB/21 (07-09)
Approved for use through 07/31/2012. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

TRANSMITTAL FORM

(to be used for all correspondence after initial filing)

Total Number of Pages in This Submission	4

Application Number	14/181,112
Filing Date	February 14, 2014
First Named Inventor	Zvi Tannenbaum
Art Unit	2447
Examiner Name	Not yet assigned
Attorney Docket Number	ZTANN.002P1C3

ENCLOSURES *(Check all that apply)*

- ☐ Fee Transmittal Form
  - ☐ Fee Attached
- ☑ Amendment/Reply
  - ☐ After Final
  - ☐ Affidavits/declaration(s)
- ☐ Extension of Time Request
- ☐ Express Abandonment Request
- ☐ Information Disclosure Statement
- ☐ Certified Copy of Priority Document(s)
- ☑ Reply to Missing Parts/ Incomplete Application
  - ☑ Reply to Missing Parts under 37 CFR 1.52 or 1.53
- ☐ Drawing(s)
- ☐ Licensing-related Papers
- ☐ Petition
- ☐ Petition to Convert to a Provisional Application
- ☐ Power of Attorney, Revocation Change of Correspondence Address
- ☐ Terminal Disclaimer
- ☐ Request for Refund
- ☐ CD, Number of CD(s) _____
  - ☐ Landscape Table on CD
- ☐ After Allowance Communication to TC
- ☐ Appeal Communication to Board of Appeals and Interferences
- ☐ Appeal Communication to TC **(Appeal Notice, Brief, Reply Brief)**
- ☐ Proprietary Information
- ☐ Status Letter
- ☐ Other Enclosure(s) (please Identify below):

Remarks

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT

Firm Name	Knobbe, Martens, Olson & Bear, LLP		
Signature	/Lance D. Smemoe/		
Printed name	Lance D. Smemoe		
Date	09-08-2014	Reg. No.	66,152

CERTIFICATE OF TRANSMISSION/MAILING

I hereby certify that this correspondence is being facsimile transmitted to the USPTO or deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date shown below:

Signature			
Typed or printed name		Date	

This collection of information is required by 37 CFR 1.5. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and1.14. This collection is estimated to 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor	:	Zvi Tannenbaum
App. No.	:	14/181,112
Filed	:	February 14, 2014
For	:	CLUSTER COMPUTING
Examiner	:	Not yet assigned
Art Unit	:	2447
Conf. No.	:	6874

PRELIMINARY AMENDMENT

Mail Stop Amendment
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Preliminary to examination on the merits, please amend the above-identified application as follows:

Amendments to the Claims are reflected in the listing of claims which begins on page 2.

Remarks begin on page 3.

AMENDMENTS TO THE CLAIMS

1. (Canceled)

2. (New) A computer cluster comprising:

a plurality of nodes, wherein one or more of the nodes receives a command to start a cluster initialization process for the computer cluster, and wherein each of the nodes is configured to access a computer-readable medium comprising program code for a single-node kernel module configured to interpret user instructions; and

a mechanism for the nodes to communicate with each other.

3. (New) The computer cluster of claim 2, wherein the single-node kernel module comprises a Mathematica kernel.

4. (New) The computer cluster of claim 2, wherein the mechanism comprises a message passing interface.

5. (New) The computer cluster of claim 2, wherein each of the nodes comprises one or more cluster node modules.

6. (New) The computer cluster of claim 2, wherein each single-node kernel module is stored in a computer-readable medium and configured to accept and execute a request.

7. (New) The computer cluster of claim 6, wherein each of the nodes comprises one or more cluster node modules, and each of the cluster node modules is stored in a computer-readable medium and configured to communicate with the single-node kernel and with one or more other cluster node modules.

8. (New) The computer cluster of claim 7, wherein the plurality of cluster node modules act as a cluster.

9. (New) The computer cluster of claim 7, wherein the plurality of cluster node modules communicate with one another to act as a cluster.

10. (New) The computer cluster of claim 9, wherein the computer cluster includes a user interface.

11. (New) The computer cluster of claim 10, wherein each cluster node module accepts instructions from the user interface and interprets one or more of the instructions.

REMARKS

Amendments

Prior to examination on the merits, please amend the above-identified application as set forth above. Applicant submits that no new matter is added by these amendments.

No Disclaimers or Disavowals

Although the present communication may include alterations to the application or claims, or characterizations of claim scope or referenced art, Applicant is not conceding in this application that previously pending claims are not patentable over the cited references. Rather, any alterations or characterizations are being made to facilitate expeditious prosecution of this application. Applicant reserves the right to pursue at a later date any previously pending or other broader or narrower claims that capture any subject matter supported by the present disclosure, including subject matter found to be specifically disclaimed herein or by any prior prosecution. Accordingly, reviewers of this or any parent, child or related prosecution history shall not reasonably infer that Applicant has made any disclaimers or disavowals of any subject matter supported by the present application.

Please charge any additional fees, including any fees for additional extension of time, or credit overpayment to Deposit Account No. 11-1410.

Respectfully submitted,

KNOBBE, MARTENS, OLSON & BEAR, LLP

Dated: September 8, 2014

By: /Lance D. Smemoe/

Lance D. Smemoe
Registration No. 66,152
Attorney of Record
Customer No. 20995
(949) 760-0404

18394794

Electronic Patent Application Fee Transmittal

Application Number:	14181112
Filing Date:	14-Feb-2014
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Filer:	Lance Dean Smemoe/Debora Swies
Attorney Docket Number:	ZTANN.002P1C3

Filed as Small Entity

Utility under 35 USC 111(a) Filing Fees

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Utility filing Fee (Electronic filing)	4011	1	70	70
Utility Search Fee	2111	1	300	300
Utility Examination Fee	2311	1	360	360
Pages:				
Claims:				
Miscellaneous-Filing:				
Late Filing Fee for Oath or Declaration	2051	1	70	70
Petition:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Extension - 4 months with $0 paid	2254	1	1100	1100
Miscellaneous:				
		Total in USD ($)		**1900**

Electronic Acknowledgement Receipt

EFS ID:	20076296
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/Christine Showalter
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	08-SEP-2014
Filing Date:	14-FEB-2014
Time Stamp:	17:05:06
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	Credit Card
Payment was successfully received in RAM	$1900
RAM confirmation Number	3946
Deposit Account	111410
Authorized User	KNOBBE MARTENS OLSON AND BEAR

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

Charge any Additional Fees required under 37 C.F.R. Section 1.16 (National application filing, search, and examination fees)

Charge any Additional Fees required under 37 C.F.R. Section 1.17 (Patent application and reexamination processing fees)

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		ZTANN_002P1C3_RespMP.pdf	205080 87cf396480e666dc0e2a32473469b0e61d0762fb	yes	4
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Applicant Response to Pre-Exam Formalities Notice		1	1	
	Preliminary Amendment		2	2	
	Claims		3	3	
	Applicant Arguments/Remarks Made in an Amendment		4	4	
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	38672 8c8eb0825748a232f04b63a2afe7fc1d1c2c81f3	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	243752		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 14/181,112	Filing Date 02/14/2014	☐ To be Mailed

ENTITY: ☐ LARGE ☒ SMALL ☐ MICRO

APPLICATION AS FILED – PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*	X $ =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*	X $ =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	

APPLICATION AS AMENDED – PART II

AMENDMENT	09/08/2014	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 10	Minus	** 20	= 0	X $0 =	0
	Independent (37 CFR 1.16(h))	* 1	Minus	***3	= 0	X $0 =	0
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	**0**

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=	X $ =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=	X $ =	
	☐ Application Size Fee (37 CFR 1.16(s))						
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))						
						TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

LIE
/YUVANNA CHAPHIV/

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**
If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 14/181,112

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY FEE($)
BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A	N/A	70		N/A	
SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A	N/A	300		N/A	
EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A	N/A	360		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	1 minus 20 =	*	x 40 =	0.00	OR		
INDEPENDENT CLAIMS (37 CFR 1.16(h))	1 minus 3 =	*	x 210 =	0.00			
APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).			0.00			
MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))				0.00			
* If the difference in column 1 is less than zero, enter "0" in column 2.			TOTAL	730		TOTAL	

APPLICATION AS AMENDED - PART II

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	SMALL ENTITY RATE($)	SMALL ENTITY ADDITIONAL FEE($)	OR	OTHER THAN SMALL ENTITY RATE($)	OTHER THAN SMALL ENTITY ADDITIONAL FEE($)
AMENDMENT A	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
AMENDMENT A	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
AMENDMENT A	Application Size Fee (37 CFR 1.16(s))									
AMENDMENT A	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA	RATE($)	ADDITIONAL FEE($)	OR	RATE($)	ADDITIONAL FEE($)
AMENDMENT B	Total (37 CFR 1.16(i))	*	Minus	**	=	x =		OR	x =	
AMENDMENT B	Independent (37 CFR 1.16(h))	*	Minus	***	=	x =		OR	x =	
AMENDMENT B	Application Size Fee (37 CFR 1.16(s))									
AMENDMENT B	FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							OR		
						TOTAL ADD'L FEE		OR	TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING or 371(c) DATE	GRP ART UNIT	FIL FEE REC'D	ATTY.DOCKET.NO	TOT CLAIMS	IND CLAIMS
14/181,112	02/14/2014	2447	0.00	ZTANN.002P1C3	1	1

CONFIRMATION NO. 6874

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

FILING RECEIPT

OC000000066930359

Date Mailed: 03/06/2014

Receipt is acknowledged of this non-provisional patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF APPLICANT, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection. Please verify the accuracy of the data presented on this receipt. **If an error is noted on this Filing Receipt, please submit a written request for a Filing Receipt Correction. Please provide a copy of this Filing Receipt with the changes noted thereon. If you received a "Notice to File Missing Parts" for this application, please submit any corrections to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections**

Inventor(s)

Zvi Tannenbaum, Newport Beach, CA;
Dean E. Dauger, Huntington Beach, CA;

Applicant(s)

Advanced Cluster Systems, Inc., Aliso Viejo, CA

Assignment For Published Patent Application

Advanced Cluster Systems, Inc., Aliso Viejo, CA

Power of Attorney: None

Domestic Priority data as claimed by applicant

This application is a CON of 13/423,063 03/16/2012 PAT 8676877
which is a CON of 12/040,519 02/29/2008 PAT 8140612
which is a CIP of 11/744,461 05/04/2007 PAT 8082289
which claims benefit of 60/813,738 06/13/2006
and claims benefit of 60/850,908 10/11/2006

Foreign Applications for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.

Permission to Access - A proper **Authorization to Permit Access to Application by Participating Offices** (PTO/SB/39 or its equivalent) has been received by the USPTO.

If Required, Foreign Filing License Granted: 03/04/2014

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 14/181,112**

Projected Publication Date: To Be Determined - pending completion of Missing Parts

Non-Publication Request: No

Early Publication Request: No
**** SMALL ENTITY ****
Title

CLUSTER COMPUTING

Preliminary Class

709

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications: No

PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

LICENSE FOR FOREIGN FILING UNDER

Title 35, United States Code, Section 184

Title 37, Code of Federal Regulations, 5.11 & 5.15

GRANTED

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

NOT GRANTED

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

SelectUSA

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The U.S. offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to promote and facilitate business investment. SelectUSA provides information assistance to the international investor community; serves as an ombudsman for existing and potential investors; advocates on behalf of U.S. cities, states, and regions competing for global investment; and counsels U.S. economic development organizations on investment attraction best practices. To learn more about why the United States is the best country in the world to develop technology, manufacture products, deliver services, and grow your business, visit http://www.SelectUSA.gov or call +1-202-482-6800.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NUMBER	FILING OR 371(C) DATE	FIRST NAMED APPLICANT	ATTY. DOCKET NO./TITLE
14/181,112	02/14/2014	Zvi Tannenbaum	ZTANN.002P1C3

CONFIRMATION NO. 6874

FORMALITIES LETTER

20995
KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

OC000000066930360

Date Mailed: 03/06/2014

NOTICE TO FILE MISSING PARTS OF NONPROVISIONAL APPLICATION

FILED UNDER 37 CFR 1.53(b)

Filing Date Granted

Items Required To Avoid Abandonment:

An application number and filing date have been accorded to this application. The item(s) indicated below, however, are missing. Applicant is given **TWO MONTHS** from the date of this Notice within which to file all required items below to avoid abandonment. Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR 1.136(a).

- The statutory basic filing fee is missing.
- The application search fee must be submitted.
- The application examination fee must be submitted.
- Surcharge as set forth in 37 CFR 1.16(f) must be submitted.
  The surcharge is due for any one of:
  - late submission of the basic filing fee, search fee, or examination fee,
  - late submission of inventor's oath or declaration,
  - filing an application that does not contain at least one claim on filing, or
  - submission of an application filed by reference to a previously filed application.

SUMMARY OF FEES DUE:

The fee(s) required within **TWO MONTHS** from the date of this Notice to avoid abandonment is/are itemized below. Small entity discount is in effect. If applicant is qualified for micro entity status, an acceptable Certification of Micro Entity Status must be submitted to establish micro entity status. (See 37 CFR 1.29 and forms PTO/SB/15A and 15B.)

- $ **70** basic filing fee.
- $ **70** surcharge.
- $ **300** search fee.
- $ **360** examination fee.
- $(**0.00**) previous unapplied payment amount.
- $ **800** TOTAL FEE BALANCE DUE.

Items Required To Avoid Processing Delays:

Applicant is notified that the above-identified application contains the deficiencies noted below. No period for reply is set forth in this notice for correction of these deficiencies. However, if a deficiency relates to the inventor's

oath or declaration, the applicant must file an oath or declaration in compliance with 37 CFR 1.63, or a substitute statement in compliance with 37 CFR 1.64, executed by or with respect to each actual inventor no later than the expiration of the time period set in the "Notice of Allowability" to avoid abandonment. See 37 CFR 1.53(f).

- A properly executed inventor's oath or declaration has not been received for the following inventor(s):
  Zvi Tannenbaum
  Dean E. Dauger

Replies must be received in the USPTO within the set time period or must include a proper Certificate of Mailing or Transmission under 37 CFR 1.8 with a mailing or transmission date within the set time period. For more information and a suggested format, see Form PTO/SB/92 and MPEP 512.

Replies should be mailed to:

Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

Registered users of EFS-Web may alternatively submit their reply to this notice via EFS-Web, including a copy of this Notice and selecting the document description "Applicant response to Pre-Exam Formalities Notice". https://sportal.uspto.gov/authenticate/AuthenticateUserLocalEPF.html

For more information about EFS-Web please call the USPTO Electronic Business Center at **1-866-217-9197** or visit our website at http://www.uspto.gov/ebc.

If you are not using EFS-Web to submit your reply, you must include a copy of this notice.

/mkanno/

Office of Data Management, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

PTO/AIA/14 (11-13)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

Secrecy Order 37 CFR 5.2

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

Inventor Information:

Inventor 1 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Zvi		Tannenbaum	

Residence Information (Select One) ◉ US Residency ◯ Non US Residency ◯ Active US Military Service

City	Newport Beach	State/Province	CA	Country of Residence i	US

Mailing Address of Inventor:

Address 1	86 Sidney Bay Drive		
Address 2			
City	Newport Beach	State/Province	CA
Postal Code	92657	Country i	US

Inventor 2 Remove

Legal Name

Prefix	Given Name	Middle Name	Family Name	Suffix
	Dean	E.	Dauger	

Residence Information (Select One) ◉ US Residency ◯ Non US Residency ◯ Active US Military Service

City	Huntington Beach	State/Province	CA	Country of Residence i	US

Mailing Address of Inventor:

Address 1	3582 Venture Drive		
Address 2			
City	Huntington Beach	State/Province	CA
Postal Code	92649	Country i	US

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

Correspondence Information:

Enter either Customer Number or complete the Correspondence Information section below.
For further information see 37 CFR 1.33(a).

PTO/AIA/14 (11-13)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

☐ **An Address is being provided for the correspondence Information of this application.**			
Customer Number	20995		
Email Address	efiling@knobbe.com	Add Email	Remove Email

Application Information:

Title of the Invention	CLUSTER COMPUTING		
Attorney Docket Number	ZTANN.002P1C3	**Small Entity Status Claimed**	☒
Application Type	Nonprovisional		
Subject Matter	Utility		
Total Number of Drawing Sheets (if any)	5	**Suggested Figure for Publication (if any)**	

Publication Information:

☐	Request Early Publication (Fee required at time of Request 37 CFR 1.219)
☐	**Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

Please Select One:	◉ Customer Number	○ US Patent Practitioner	○ Limited Recognition (37 CFR 11.9)
Customer Number	20995		

Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, or 365(c) or indicate National Stage entry from a PCT application. Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the application number blank.

Prior Application Status	Pending		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
	Continuation of	13/423063	2012-03-16
Prior Application Status	Patented		Remove

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)	Patent Number	Issue Date (YYYY-MM-DD)
13/423063	Continuation of	12/040519	2008-02-29	8140612	2012-03-20
Prior Application Status	Patented			Remove	
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)	Patent Number	Issue Date (YYYY-MM-DD)
12/040519	Continuation in part of	11/744461	2007-05-04	8082289	2011-12-20
Prior Application Status	Expired			Remove	

Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
11/744461	Claims benefit of provisional	60/813738	2006-06-13
Prior Application Status	Expired		Remove
Application Number	Continuity Type	Prior Application Number	Filing Date (YYYY-MM-DD)
11/744461	Claims benefit of provisional	60/850908	2006-10-11

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. Add

Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55(d). When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX) [i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(h)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Remove

Application Number	Country [i]	Filing Date (YYYY-MM-DD)	Access Code [i] (if applicable)

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. Add

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

Authorization to Permit Access:

☒ Authorization to Permit Access to the Instant Application by the Participating Offices

If checked, the undersigned hereby grants the USPTO authority to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the World Intellectual Property Office (WIPO), and any other intellectual property offices in which a foreign application claiming priority to the instant patent application is filed access to the instant patent application. See 37 CFR 1.14(c) and (h). This box should not be checked if the applicant does not wish the EPO, JPO, KIPO, WIPO, or other intellectual property office in which a foreign application claiming priority to the instant patent application is filed to have access to the instant patent application.

In accordance with 37 CFR 1.14(h)(3), access will be provided to a copy of the instant patent application with respect to: 1) the instant patent application-as-filed; 2) any foreign application to which the instant patent application claims priority under 35 U.S.C. 119(a)-(d) if a copy of the foreign application that satisfies the certified copy requirement of 37 CFR 1.55 has been filed in the instant patent application; and 3) any U.S. application-as-filed from which benefit is sought in the instant patent application.

In accordance with 37 CFR 1.14(c), access may be provided to information concerning the date of filing this Authorization.

Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

PTO/AIA/14 (11-13)
Approved for use through 01/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

Applicant 1 Remove

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

- (●) Assignee
- ( ) Legal Representative under 35 U.S.C. 117
- ( ) Joint Inventor
- ( ) Person to whom the inventor is obligated to assign.
- ( ) Person who shows sufficient proprietary interest

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor :

If the Applicant is an Organization check here. [X]

Organization Name	Advanced Cluster Systems, Inc.		
Mailing Address Information:			
Address 1	65 Enterprise		
Address 2			
City	Aliso Viejo	**State/Province**	CA
Country i	US	Postal Code	92656
Phone Number		Fax Number	
Email Address			

Additional Applicant Data may be generated within this form by selecting the Add button. Add

Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not subsitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

Assignee 1

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication . An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

Remove

If the Assignee or Non-Applicant Assignee is an Organization check here. [ ]

Application Data Sheet 37 CFR 1.76	Attorney Docket Number	ZTANN.002P1C3
	Application Number	
Title of Invention	CLUSTER COMPUTING	

Prefix	**Given Name**	Middle Name	**Family Name**	Suffix

Mailing Address Information For Assignee including Non-Applicant Assignee:

Address 1			
Address 2			
City		**State/Province**	
Country i		Postal Code	
Phone Number		Fax Number	
Email Address			

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button. Add

Signature:

Remove

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications

Signature	/Lance D. Smemoe/			Date (YYYY-MM-DD)	2014-02-14
First Name	Lance D.	Last Name	Smemoe	Registration Number	66152

Additional Signature may be generated within this form by selecting the Add button. Add

This collection of information is required by 37 CFR 1.76. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 23 minutes to complete, including gathering, preparing, and submitting the completed application data sheet form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

CLUSTER COMPUTING

INCORPORATION BY REFERENCE TO ANY PRIORITY APPLICATIONS

[0001] Any and all applications for which a foreign or domestic priority claim is identified in the Application Data Sheet as filed with the present application are incorporated by reference under 37 CFR 1.57 and made a part of this specification.

BACKGROUND

Field

[0002] The present disclosure relates to the field of cluster computing generally and to systems and methods for adding cluster computing functionality to a computer program, in particular.

Description of Related Art

[0003] Computer clusters include a group of two or more computers, microprocessors, and/or processor cores ("nodes") that intercommunicate so that the nodes can accomplish a task as though they were a single computer. Many computer application programs are not currently designed to benefit from advantages that computer clusters can offer, even though they may be running on a group of nodes that could act as a cluster. Some computer programs can run on only a single node because, for example, they are coded to perform tasks serially or because they are designed to recognize or send instructions to only a single node.

[0004] Some application programs include an interpreter that executes instructions provided to the program by a user, a script, or another source. Such an interpreter is sometimes called a "kernel" because, for example, the interpreter can manage at least some hardware resources of a computer system and/or can manage communications between those resources and software (for example, the provided instructions, which can include a high-level programming language). Some software programs include a kernel that is designed to communicate with a single node. An example of a software package that includes a kernel that is designed to communicate with a single node is Mathematica® from Wolfram Research, Inc. ("Mathematica"). Mathematics software packages from other vendors and other types of software can also include such a kernel.

[0005] A product known as gridMathematica, also from Wolfram Research, Inc., gives Mathematica the capability to perform a form of grid computing known as "distributed computing." Grid computers include a plurality of nodes that generally do not communicate with one another as peers. Distributed computing can be optimized for workloads that consist of many independent jobs or packets of work, which do not need to share data between the jobs during the computational process. Grid computers include at least one node known as a master node that manages a plurality of slave nodes or computational nodes. In gridMathematica, each of a plurality of kernels runs on a single node. One kernel is designated the master kernel, which handles all input, output, and scheduling of the other kernels (the computational kernels or slave kernels). Computational kernels receive commands and data only from the node running the master kernel. Each computational kernel performs its work independently of the other computational kernels and intermediate results of one job do not affect other jobs in progress on other nodes.

SUMMARY

[0006] Embodiments described herein have several features, no single one of which is solely responsible for their desirable attributes. Without limiting the scope of the invention as expressed by the claims, some of the advantageous features will now be discussed briefly.

[0007] Some embodiments described herein provide techniques for conveniently adding cluster computing functionality to a computer application. In one embodiment, a user of a software package may be able to achieve higher performance and/or higher availability from the software package by enabling the software to benefit from a plurality of nodes in a cluster. One embodiment allows a user to create applications, using a high-level language such as Mathematica, that are able to run on a computer cluster having supercomputer-like performance. One embodiment provides access to such high-performance computing through a Mathematica Front End, a command line interface, one or more high-level commands, or a programming language such as C or FORTRAN.

[0008] One embodiment adapts a software module designed to run on a single node, such as, for example, the Mathematica kernel, to support cluster computing, even when the software module is not designed to provide such support. One embodiment provides

parallelization for an application program, even if no access to the program's source code is available. One embodiment adds and supports Message Passing Interface ("MPI") calls directly from within a user interface, such as, for example, the Mathematica programming environment. In one embodiment, MPI calls are added to or made available from an interactive programming environment, such as the Mathematica Front End.

[0009] One embodiment provides a computer cluster including a first processor, a second processor, and a third processor. The cluster includes at least one computer-readable medium in communication at least one of the first processor, the second processor, or the third processor. A first kernel resides in the at least one computer-readable medium and is configured to translate commands into code for execution on the first processor. A first cluster node module resides in the at least one computer-readable medium. The first cluster node module is configured to send commands to the first kernel and receives commands from a user interface. A second kernel resides in the at least one computer-readable medium. The second kernel is configured to translate commands into code for execution on the second processor. A second cluster node module resides in the at least one computer-readable medium. The second cluster node module is configured to send commands to the second kernel and communicates with the first cluster node module. A third kernel resides in the at least one computer-readable medium. The third kernel is configured to translate commands into code for execution on the third processor. A third cluster node module resides in the at least one computer-readable medium. The third cluster node module is configured to send commands to the third kernel and configured to communicate with the first cluster node module and the second cluster node module. The first cluster node module comprises a data structure in which messages originating from the second and third cluster node modules are stored.

[0010] Another embodiment provides a computer cluster that includes a plurality of nodes and a software package including a user interface and a single-node kernel for interpreting program code instructions. A cluster node module is configured to communicate with the single-node kernel and other cluster node modules. The cluster node module accepts instructions from the user interface and interprets at least some of the instructions such that several cluster node modules in communication with one another act as a cluster. The cluster

node module appears as a single-node kernel to the user interface. In one embodiment, the single-node kernel includes a Mathematical kernel. In some embodiments, the user interface can include at least one of a Mathematica front end or a command line. In some embodiments, the cluster node module includes a toolkit including library calls that implement at least a portion of MPI calls. In some embodiments, the cluster node module includes a toolkit including high-level cluster computing commands. In one embodiment, the cluster system can include a plurality of Macintosh® computers ("Macs"), Windows®-based personal computers ("PCs"), and/or Unix/Linux-based workstations.

[0011] A further embodiment provides a computer cluster including a plurality of nodes. Each node is configured to access a computer-readable medium comprising program code for a user interface and program code for a single-node kernel module configured to interpret user instructions. The cluster includes a plurality of cluster node modules. Each cluster node module is configured to communicate with a single-node kernel and with one or more other cluster node modules, to accept instructions from the user interface, and to interpret at least some of the user instructions such that the plurality of cluster node modules communicate with one another in order to act as a cluster. A communications network connects the nodes. One of the plurality of cluster node modules returns a result to the user interface.

[0012] Another embodiment provides a method of evaluating a command on a computer cluster. A command from at least one of a user interface or a script is communicated to one or more cluster node modules within the computer cluster. Each of the one or more cluster node modules communicates a message based on the command to a respective kernel module associated with the cluster node module. Each of the one or more cluster node modules receives a result from the respective kernel module associated with the cluster node module. At least one of the one or more cluster node modules responds to messages from other cluster node modules.

[0013] Another embodiment provides a computing system for executing Mathematica code on multiple nodes. The computing system includes a first node module in communication with a first Mathematica kernel executing on a first node, a second node module in communication with a second Mathematica kernel executing on a second node,

and a third node module in communication with a third Mathematica kernel executing on a third node. The first node module, the second node module, and the third node module are configured to communicate with one another using a peer-to-peer architecture. In some embodiments, each of the first node module, the second node module, and third node module includes a data structure for maintaining messages originating from other node modules and a data structure for maintaining data specifying a location to which an message is expected to be received and an identifier for a node from which the message is expected to be sent.

BRIEF DESCRIPTION OF THE DRAWINGS

[0014] A general architecture that implements the various features are described with reference to the drawings. The drawings and the associated descriptions are provided to illustrate embodiments and not to limit the scope of the disclosure. Throughout the drawings, reference numbers are re-used to indicate correspondence between referenced elements.

[0015] FIG. 1 is a block diagram of one embodiment of a computer cluster.

[0016] FIG. 2 is a block diagram showing relationships between software modules running on one embodiment of a computer cluster.

[0017] FIG. 3 is a block diagram of one embodiment of a cluster node module.

[0018] FIG. 4 is a flow chart showing one embodiment of a cluster initialization process.

[0019] FIG. 5 is a flow chart showing one embodiment of the operation of a cluster node module.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[0020] For purposes of illustration, some embodiments are described herein in the context of cluster computing with Mathematica software. The present disclosure is not limited to a single software program; the systems and methods can be used with other application software such as, for example, Maple®, MATLAB®, MathCAD®, Apple Shake®, Apple® Compressor, IDL®, other applications employing an interpreter or a kernel, Microsoft Excel®, Adobe After Effects®, Adobe Premiere®, Adobe Photoshop®, Apple Final Cut Pro®, and Apple iMovie®. Some figures and/or descriptions, however, relate to embodiments of computer clusters running Mathematica. The system can include a variety of uses, including but not limited to students, educators, scientists, engineers, mathematicians,

researchers, and technicians. It is also recognized that in other embodiments, the systems and methods can be implemented as a single module and/or implemented in conjunction with a variety of other modules. Moreover, the specific implementations described herein are set forth in order to illustrate, and not to limit, the disclosure.

I. Overview

[0021] The cluster computing system described herein generally includes one or more computer systems connected to one another via a communications network or networks. The communications network can include one or more of a local area network ("LAN"), a wide area network ("WAN"), an intranet, the Internet, etc. In one embodiment, a computer system comprises one or more processors such as, for example, a microprocessor that can include one or more processing cores ("nodes"). The term "node" refers to a processing unit or subunit that is capable of single-threaded execution of code. The processors can be connected to one or more memory devices such as, for example, random access memory ("RAM"), and/or one or more optional storage devices such as, for example, a hard disk. Communications among the processors and such other devices may occur, for example, via one or more local buses of a computer system or via a LAN, a WAN, a storage area network ("SAN"), and/or any other communications network capable of carrying signals among computer system components. In one embodiment, one or more software modules such as kernels, run on nodes within the interconnected computer systems. In one embodiment, the kernels are designed to run on only a single node. In one embodiment, cluster node modules communicate with the kernels and with each other in order to implement cluster computing functionality.

[0022] FIG. 1 is a block diagram of one embodiment of a computer cluster 100 wherein computer systems 110, 120, 130 communicate with one another via a communications network 102. Network 102 includes one or more of a LAN, a WAN, a wireless network, an intranet, or the Internet. In one embodiment of the computer cluster, computer system 110 includes processors 112a, 112b, memory 114, and optional storage 116. Other computer systems 120, 130 can include similar devices, which generally communicate with one another within a computer system over a local communications architecture such as a local bus (not shown). A computer system can include one or more processors, and each

processor can contain one or more processor cores that are capable of single-threaded execution. Processor cores are generally independent microprocessors, but more than one can be included in a single chip package. Software code designed for single-threaded execution can generally run on one processor core at a time. For example, single-threaded software code typically does not benefit from multiple processor cores in a computer system.

[0023] FIG. 2 is a block diagram showing relationships among software modules running on one embodiment of a computer cluster 100. In the embodiment shown in FIG. 2, the kernel modules 206a-e are designed for single-threaded execution. For example, if each of the processors 112a, 112b, 122a, 122b, 132 shown in FIG. 1 includes only one processor core, two kernel modules (for example, kernel modules 206a, 206b) loaded into the memory 114 of computer system 110 could exploit at least some of the processing bandwidth of the two processors 112a, 112b. Similarly, two kernel modules 206c, 206d loaded into the memory 124 of computer system 120 could exploit at least some of the processing bandwidth of the two processors 122a, 122b. Likewise, the bandwidth of processor 132 of computer system 130 could be utilized by a single instance of a cluster node module 204e loaded into the computer system's memory 134.

[0024] In the embodiment shown in FIG. 2, each of the kernel modules 206a-e is in communication with a single cluster node module 204a-e, respectively. For example, the kernel module 206a is in communication with the cluster node module 204a, the kernel module 206b is in communication with the cluster node module 206b, and so forth. In one embodiment, one instance of a cluster node module 204a-e is loaded into a computer system's memory 114, 124, 134 for every instance of a kernel module 206a-e running on the system. As shown in FIG. 2, each of the cluster node modules 204a-e is in communication with each of the other cluster node modules 204a-e. For example, one cluster node module 204a is in communication with all of the other cluster node modules 204b-e. A cluster node module 204a may communicate with another cluster node module 204b via a local bus (not shown) when, for example, both cluster node modules 204a-b execute on processors 112a, 112b within the same computer system 110. A cluster node module 204a may also communicate with another cluster node module 204c over a communications network 102

when, for example, the cluster node modules 204a, c execute on processors 112a, 122a within different computer systems 110, 120.

[0025] As shown in FIG. 2, an optional user interface module 202 such as, for example, a Mathematica front end and/or a command line interface, can connect to a cluster node module 204a. The user interface module can run on the same computer system 110 and/or the same microprocessor 112a on which the cluster node module 204a runs. The cluster node modules 204a-e provide MPI calls and/or advanced cluster functions that implement cluster computing capability for the single-threaded kernel modules. The cluster node modules 204a-e are configured to look and behave like a kernel module 206a from the perspective of the user interface module 202. Similarly, the cluster node modules 204a-e are configured to look and behave like a user interface module 202 from the perspective of a kernel module 206a. The first cluster node module 204a is in communication with one or more other cluster node modules 204b, 204c, and so forth, each of which provides a set of MPI calls and/or advanced cluster commands. In one embodiment, MPI may be used to send messages between nodes in a computer cluster.

[0026] Communications can occur between any two or more cluster node modules (for example, between a cluster node module 204a and another cluster node module 204c) and not just between "adjacent" kernels. Each of the cluster node modules 204a-e is in communication with respective kernel modules 206a-e. Thus, the cluster node module 204a communicates with the kernel module 206a. MPI calls and advanced cluster commands are used to parallelize program code received from an optional user interface module 208 and distribute tasks among the kernel modules 206a-e. The cluster node modules 204a-e provide communications among kernel modules 206a-e while the tasks are executing. Results of evaluations performed by kernel modules 206a-e are communicated back to the first cluster node module 204a via the cluster node modules 204a-e, which communicates them to the user interface module 208.

[0027] Intercommunication among kernel modules 206a-e during thread execution, which is made possible by cluster node modules 204a-e, provides advantages for addressing various types of mathematic and scientific problems, for example. Intercommunication provided by cluster computing permits exchange of information between

nodes during the course of a parallel computation. Embodiments of the present disclosure provide such intercommunication for software programs such as Mathematica, while grid computing solutions can implement communication between only one master node and many slave nodes. Grid computing does not provide for communication between slave nodes during thread execution.

[0028] For purposes of providing an overview of some embodiments, certain aspects, advantages, benefits, and novel features of the invention are described herein. It is to be understood that not necessarily all such advantages or benefits can be achieved in accordance with any particular embodiment of the invention. Thus, for example, those skilled in the art will recognize that the invention can be embodied or carried out in a manner that achieves one advantage or group of advantages as taught herein without necessarily achieving other advantages or benefits as can be taught or suggested herein.

II. Computer Cluster 100

[0029] As shown in FIG. 1, one embodiment of a cluster system 100 includes computer systems 110, 120, 130 in communication with one another via a communications network 102. A first computer system 110 can include one or more processors 112a-b, a memory device 114, and an optional storage device 116. Similarly, a second computer system 120 can include one or more processors 122a-b, a memory device 124, and an optional storage device 126. Likewise, a third computer system 130 can include one or more processors 132, a memory device 134, and an optional storage device 136. Each of the computer systems 110, 120, 130 includes a network interface (not shown) for connecting to a communications network 102, which can include one or more of a LAN, a WAN, an intranet, a wireless network, and/or the Internet.

A. Computer System 110

[0030] In one embodiment, a first computer system 110 communicates with other computer systems 120, 130 via a network 102 as part of a computer cluster 100. In one embodiment, the computer system 110 is a personal computer, a workstation, a server, or a blade including one or more processors 112a-b, a memory device 114, an optional storage device 116, as well as a network interface module (not shown) for communications with the network 102.

1. Processors 112a-b

[0031] In one embodiment, the computer system 110 includes one or more processors 112a-b. The processors 112a-b can be one or more general purpose single-core or multi-core microprocessors such as, for example, a Pentium® processor, a Pentium® II processor, a Pentium® Pro processor, a Pentium® III processor, Pentium® 4 processor, a Core Duo® processor, a Core 2 Duo® processor, a Xeon® processor, an Itanium® processor, a Pentium® M processor, an x86 processor, an Athlon® processor, an 8051 processor, a MIPS® processor, a PowerPC® processor, an ALPHA® processor, etc. In addition, one or more of the processors 112a-b can be a special purpose microprocessor such as a digital signal processor. The total number of processing cores (for example, processing units capable of single-threaded execution) within all processors 112a-b in the computer system 110 corresponds to the number of nodes available in the computer system 110. For example, if the processors 112a-b were each Core 2 Duo® processors having two processing cores, computer system 110 would have four nodes in all. Each node can run one or more instances of a program module, such as a single-threaded kernel module.

2. Network Interface Module

[0032] The computer system 110 can also include a network interface module (not shown) that facilitates communication between the computer system 110 and other computer systems 120, 130 via the communications network 102.

[0033] The network interface module can use a variety of network protocols. In one embodiment, the network interface module includes TCP/IP. However, it is to be appreciated that other types of network communication protocols such as, for example, Point-to-Point Protocol ("PPP"), Server Message Block ("SMB"), Serial Line Internet Protocol ("SLIP"), tunneling PPP, AppleTalk, etc., may also be used.

3. Memory 114 and Storage 116

[0034] The computer system 110 can include memory 114. Memory 114 can include, for example, processor cache memory (such as processor core-specific or cache memory shared by multiple processor cores), dynamic random-access memory ("DRAM"), static random-access memory ("SRAM"), or any other type of memory device capable of storing computer data, instructions, or program code. The computer system 110 can also

include optional storage 116. Storage 116 can include, for example, one or more hard disk drives, floppy disks, flash memory, magnetic storage media, CD-ROMs, DVDs, optical storage media, or any other type of storage device capable of storing computer data, instructions, and program code.

4. Computer System 110 Information

[0035] The computer system 110 may be used in connection with various operating systems such as: Microsoft® Windows® 3.X, Windows 95®, Windows 98®, Windows NT®, Windows 2000®, Windows XP®, Windows CE®, Palm Pilot OS, OS/2, Apple® MacOS®, MacOS X®, MacOS X Server®, Disk Operating System (DOS), UNIX, Linux®, VxWorks, or IBM® OS/2®, Sun OS, Solaris OS, IRIX OS operating systems, etc.

[0036] In one embodiment, the computer system 110 is a personal computer, a laptop computer, a Blackberry® device, a portable computing device, a server, a computer workstation, a local area network of individual computers, an interactive kiosk, a personal digital assistant, an interactive wireless communications device, a handheld computer, an embedded computing device, or the like.

[0037] As can be appreciated by one of ordinary skill in the art, the computer system 110 may include various sub-routines, procedures, definitional statements, and macros. Each of the foregoing modules are typically separately compiled and linked into a single executable program. However, it is to be appreciated by one of ordinary skill in the art that the processes that are performed by selected ones of the modules may be arbitrarily redistributed to one of the other modules, combined together in a single module, made available in a shareable dynamic link library, or partitioned in any other logical way.

B. Computer System 120

[0038] In one embodiment, a second computer system 120 communicates with other computer systems 110, 130 via a network 102 as part of a computer cluster 100. In one embodiment, the computer system 120 is a personal computer, a workstation, a server, or a blade including one or more processors 122a-b, a memory device 124, an optional storage device 126, as well as a network interface module (not shown) for communications with the network 102.

1. Processors 112a-b

[0039] In one embodiment, the computer system 120 includes one or more processors 122a-b. The processors 122a-b can be one or more general purpose single-core or multi-core microprocessors such as a Pentium® processor, a Pentium® II processor, a Pentium® Pro processor, a Pentium® III processor, Pentium® 4 processor, a Core Duo® processor, a Core 2 Duo® processor, a Xeon® processor, an Itanium® processor, a Pentium® M processor, an x86 processor, an Athlon® processor, an 8051 processor, a MIPS® processor, a PowerPC® processor, an ALPHA® processor, etc. In addition, the processors 122a-b can be any special purpose microprocessors such as a digital signal processor. The total number of processing cores (for example, processing units capable of single-threaded execution) within all processors 122a-b in the computer system 120 corresponds to the number of nodes available in the computer system 120. For example, if the processors 122a-b were each Core 2 Duo® processors having two processing cores, computer system 120 would have four nodes in all. Each node can run one or more instances of a program module, such as a single-threaded kernel module.

2. Network Interface Module

[0040] The computer system 120 can also include a network interface module (not shown) that facilitates communication between the computer system 120 and other computer systems 110, 130 via the communications network 102.

[0041] The network interface module can use a variety of network protocols. In one embodiment, the network interface module includes TCP/IP. However, it is to be appreciated that other types of network communication protocols such as, for example, Point-to-Point Protocol ("PPP"), Server Message Block ("SMB"), Serial Line Internet Protocol ("SLIP"), tunneling PPP, AppleTalk, etc., may also be used.

3. Memory 124 and Storage 126

[0042] The computer system 120 can include memory 124. Memory 124 can include, for example, processor cache memory (such as processor core-specific or cache memory shared by multiple processor cores), dynamic random-access memory ("DRAM"), static random-access memory ("SRAM"), or any other type of memory device capable of storing computer data, instructions, or program code. The computer system 120 can also

include optional storage 126. Storage 126 can include, for example, one or more hard disk drives, floppy disks, flash memory, magnetic storage media, CD-ROMs, DVDs, optical storage media, or any other type of storage device capable of storing computer data, instructions, and program code.

4. Computer System 120 Information

[0043] The computer system 120 may be used in connection with various operating systems such as: Microsoft® Windows® 3.X, Windows 95®, Windows 98®, Windows NT®, Windows 2000®, Windows XP®, Windows CE®, Palm Pilot OS, OS/2, Apple® MacOS®, MacOS X®, MacOS X Server®, Disk Operating System (DOS), UNIX, Linux®, VxWorks, or IBM® OS/2®, Sun OS, Solaris OS, IRIX OS operating systems, etc.

[0044] In one embodiment, the computer system 120 is a personal computer, a laptop computer, a Blackberry® device, a portable computing device, a server, a computer workstation, a local area network of individual computers, an interactive kiosk, a personal digital assistant, an interactive wireless communications device, a handheld computer, an embedded computing device, or the like.

[0045] As can be appreciated by one of ordinary skill in the art, the computer system 120 may include various sub-routines, procedures, definitional statements, and macros. Each of the foregoing modules are typically separately compiled and linked into a single executable program. However, it is to be appreciated by one of ordinary skill in the art that the processes that are performed by selected ones of the modules may be arbitrarily redistributed to one of the other modules, combined together in a single module, made available in a shareable dynamic link library, or partitioned in any other logical way.

C. Computer System 130

[0046] In one embodiment, a third computer system 130 communicates with other computer systems 110, 120 via a network 102 as part of a computer cluster 100. In one embodiment, the computer system 130 is a personal computer, a workstation, a server, or a blade including one or more processors 132, a memory device 134, an optional storage device 136, as well as a network interface module (not shown) for communications with the network 102.

1. Processors 112a-b

[0047] In one embodiment, the computer system 130 includes a processor 132. The processor 132 can be a general purpose single-core or multi-core microprocessors such as a Pentium® processor, a Pentium® II processor, a Pentium® Pro processor, a Pentium® III processor, Pentium® 4 processor, a Core Duo® processor, a Core 2 Duo® processor, a Xeon® processor, an Itanium® processor, a Pentium® M processor, an x86 processor, an Athlon® processor, an 8051 processor, a MIPS® processor, a PowerPC® processor, or an ALPHA® processor. In addition, the processor 132 can be any special purpose microprocessor such as a digital signal processor. The total number of processing cores (for example, processing units capable of single-threaded execution) within processor 132 in the computer system 130 corresponds to the number of nodes available in the computer system 130. For example, if the processor 132 was a Core 2 Duo® processor having two processing cores, the computer system 130 would have two nodes. Each node can run one or more instances of a program module, such as a single-threaded kernel module.

2. Network Interface Module

[0048] The computer system 130 can also include a network interface module (not shown) that facilitates communication between the computer system 130 and other computer systems 110, 120 via the communications network 102.

[0049] The network interface module can use a variety of network protocols. In one embodiment, the network interface module includes TCP/IP. However, it is to be appreciated that other types of network communication protocols such as, for example, Point-to-Point Protocol ("PPP"), Server Message Block ("SMB"), Serial Line Internet Protocol ("SLIP"), tunneling PPP, AppleTalk, etc., may also be used.

3. Memory 134 and Storage 136

[0050] The computer system 130 can include memory 134. Memory 134 can include, for example, processor cache memory (such as processor core-specific or cache memory shared by multiple processor cores), dynamic random-access memory ("DRAM"), static random-access memory ("SRAM"), or any other type of memory device capable of storing computer data, instructions, or program code. The computer system 130 can also include optional storage 136. Storage 136 can include, for example, one or more hard disk

drives, floppy disks, flash memory, magnetic storage media, CD-ROMs, DVDs, optical storage media, or any other type of storage device capable of storing computer data, instructions, and program code.

4. Computer System 130 Information

[0051] The computer system 130 may be used in connection with various operating systems such as: Microsoft® Windows® 3.X, Windows 95®, Windows 98®, Windows NT®, Windows 2000®, Windows XP®, Windows CE®, Palm Pilot OS, OS/2, Apple® MacOS®, MacOS X®, MacOS X Server®, Disk Operating System (DOS), UNIX, Linux®, VxWorks, or IBM® OS/2®, Sun OS, Solaris OS, IRIX OS operating systems, etc.

[0052] In one embodiment, the computer system 130 is a personal computer, a laptop computer, a Blackberry® device, a portable computing device, a server, a computer workstation, a local area network of individual computers, an interactive kiosk, a personal digital assistant, an interactive wireless communications device, a handheld computer, an embedded computing device, or the like.

[0053] As can be appreciated by one of ordinary skill in the art, the computer system 130 may include various sub-routines, procedures, definitional statements, and macros. Each of the foregoing modules are typically separately compiled and linked into a single executable program. However, it is to be appreciated by one of ordinary skill in the art that the processes that are performed by selected ones of the modules may be arbitrarily redistributed to one of the other modules, combined together in a single module, made available in a shareable dynamic link library, or partitioned in any other logical way.

E. Communications Network 102

[0054] In one embodiment, computer systems 110, 120, 130 are in communication with one another via a communications network 102.

[0055] The communications network 102 may include one or more of any type of electronically connected group of computers including, for instance, the following networks: a virtual private network, a public Internet, a private Internet, a secure Internet, a private network, a public network, a value-added network, a wired network, a wireless network, an intranet, etc. In addition, the connectivity to the network can be, for example, a modem, Ethernet (IEEE 802.3), Gigabit Ethernet, 10-Gigabit Ethernet, Token Ring (IEEE 802.5),

Fiber Distributed Datalink Interface (FDDI), Frame Relay, InfiniBand, Myrinet, Asynchronous Transfer Mode (ATM), or another interface. The communications network 102 may connect to the computer systems 110, 120, 130, for example, by use of a modem or by use of a network interface card that resides in each of the systems.

[0056] In addition, the same or different communications networks 102 may be used to facilitate communication between the first computer system 110 and the second computer system 120, between the first computer system 110 and the third computer system 130, and between the second computer system 120 and the third computer system 130.

III. Software Modules

[0057] As shown in FIG.s 1 and 2, one embodiment of a cluster system 100 includes a user interface module 202 that is able to access a plurality of kernel modules 206a-e by communicating with a first cluster node module 204a. User interface module can be stored in a memory 114, 124, 134 while running, for example, and/or can be stored in a storage device 116, 126, 136. The first cluster node module 204a is in communication with each of the other cluster node modules 204b-e. The kernel modules 206a-e can reside in the memory of one or more computer systems on which they run. For example, the memory 114 of the first computer system 110 can store instances of kernel modules 206a-b, the memory 124 of the second computer system 120 can store instances of kernel modules 206c-d, and the memory 134 of the third computer system 130 can store an instance of kernel module 206e. The kernel modules 206a-e, which include single-threaded program code, are each associated with one of the processors 112a, 112b, 122a, 122b, 132. A cluster configuration module stored on one or more of the computer systems 110, 120, 130 or on a remote computer system, for example, can establish communication with the cluster node modules 204a-e. In one embodiment, communication between the cluster configuration module 208 and the cluster node modules 204a-e initializes the cluster node modules 204a-e to provide cluster computing support for the computer cluster 100.

A. Cluster Node Module 204

[0058] In one embodiment, the cluster node modules 204a-e provide a way for many kernel modules 206a-e such as, for example, Mathematica kernels, running on a computer cluster 100 to communicate with one another. A cluster node module 204 can

include at least a portion of an application programming interface ("API") known as the Message-Passing Interface ("MPI"), which is used in several supercomputer and cluster installations. A network of connections (for example, the arrows shown in FIG. 2) between the cluster node modules 204a-e can be implemented using a communications network 102, such as, for example, TCP/IP over Ethernet, but the connections could also occur over any other type of network or local computer bus.

[0059] A cluster node module 204 can use an application-specific toolkit or interface such as, for example, Mathematica's MathLink, Add-Ons, or packets, to interact with an application. Normally used to connect a Mathematica kernel to a user interface known as the Mathematica Front End or other Mathematica kernels, MathLink is a bidirectional protocol to sends "packets" containing messages, commands, or data between any of these entities. MathLink does not allow direct cluster computing-like simultaneous communication between Mathematica kernels during execution of a command or thread. MathLink is also not designed to perform multiple simultaneous network connections. In some embodiments, a cluster node module 204 can use an application-specific toolkit such as, for example, MathLink, for connections between entities on the same computer.

[0060] When speaking about procedures or actions on a cluster or other parallel computer, not all actions happen in sequential order, nor are they required to. For example, a parallel code, as opposed to a single-processor code of the classic "Turing machine" model, has multiple copies of the parallel code running across the cluster, typically one for each processor (or "processing element" or "core"). Such parallel code is written in such a way that different instances of the same code can communicate, collaborate, and coordinate work with each other. Multiple instances of these codes can run at the same time in parallel.

[0061] If the count of the code instances is an integer N, each instance of code execution can be labeled 0 through N-1. For example, a computer cluster can include N connected computers, each containing a processor. The first has cluster node module 0 connected with kernel module 0 running on processor 0. The next is cluster node module 1 and kernel module 1, on processor 1, and so forth for each of the N connected computers. Some steps of their procedure are collaborative, and some steps are independent. Even

though these entities are not necessarily in lock-step, they do follow a pattern of initialization, main loop behavior (for example, cluster node module operation), and shut down.

[0062] In contrast, a parallel computing toolkit (PCT) that is provided as part of the gridMathematica software package does not provide a means for instances of the same code running on different nodes to communicate, collaborate, or coordinate work among the instances. The PCT provides commands that connect Mathematica kernels in a master-slave relationship rather than a peer-to-peer relationship as enabled by some embodiments disclosed herein. A computer cluster having peer-to-peer node architecture performs computations that can be more efficient, easier to design, and/or more reliable than similar computations performed on grid computers having master-slave node architecture. Moreover, the nature of some computations may not allow a programmer to harness multi-node processing power on systems that employ master-slave node architecture.

[0063] FIG. 3 shows one embodiment of a cluster node module 204 implementing MPI calls and advanced MPI functions. In the embodiment shown in FIG. 3, cluster node module 204 includes MPI module 302, advanced functions module 304, received message queue 306, and message receiving queue 308.

1. <u>MPI module 302</u>

[0064] In one embodiment, the cluster node module 204 includes an MPI module 302. The MPI module 302 can include program code for one or more of at least five kinds of MPI instructions or calls. Selected constants, instructions, and/or calls that can be implemented by the MPI module 302 are as follows:

MPI Constants

[0065] Node identifiers are used to send messages to nodes or receive messages from them. In MPI, this is accomplished by assigning each node a unique integer ($IdProc) starting with 0. This data, with a knowledge of the total count ($NProc), makes it possible to programmatically divide any measurable entity.

Constant	Description
$IdProc	The identification number of the current processor
$NProc	The number of processors in the current cluster
$mpiCommWorld	The communicator world of the entire cluster (see MPI Communicator routines, below)
mpiCommWorld	The default communicator world for the high-level routines.

Table A

Basic MPI Calls

[0066] In one embodiment, the MPI module 302 can include basic MPI calls such as, for example, relatively low-level routines that map MPI calls that are commonly used in other languages (such as C and Fortran), so that such calls can be available directly from the Mathematica user interface 204. In some embodiments, basic MPI calls include calls that send data, equations, formulas, and/or other expressions.

[0067] Simply sending expressions from one node to another is possible with these most basic MPI calls. One node can call to send an expression while the other calls a corresponding routine to receive the sent expression. Because it is possible that the receiver has not yet called mpiRecv even if the message has left the sending node, completion of mpiSend is not a confirmation that it has been received.

Call	Description
mpiSend[expr, target, comm, tag]	Sends an expression expr to a node with the ID target in the communicator would comm, waiting until that expression has left this kernel
mpiRecv [expr, target, comm, tag]	Receives an expression into expr from a node with the ID target in the communicator world comm, waiting until the expression has arrived
mpiSendRecv[sendexpr, dest, recvexpr, source, comm]	Simultaneously sends the expression sendexpr to the node with the ID target and receives an expression into recvexpr from the node with the ID source in the communicator world comm, waiting until both operations have returned.

Table B

Asynchronous MPI Calls

[0068] Asynchronous calls make it possible for the kernel to do work while communications are proceeding simultaneously. It is also possible that another node may not be able to send or receive data yet, allowing one kernel to continue working while waiting.

Call	Description
mpiISend[expr, target, comm, tag, req]	Sends an expression expr to a processor with the ID target in the communicator world comm, returning immediately. It can be balanced with calls to mpiTest[req] until mpiTest[req] returns True.
mpiIRecv[expr, target, comm, tag, req]	Receives an expression expr from a proessor with the ID target in the communicator world comm, returning immediately. It can be balanced with calls to mpiTest[req] until mpiTest[req] returns True. The expr is not safe to access until mpiTest[req] returns True.
mpiTest[req]	Completes asynchronous behavior of mpiISend and mpiIRecv
mpWait[req]	Calls mpiTest until it returns True.
mpiWaitall[reglist]	Calls mpiWait all on every element of reqlist
mpiWaitany[reqlist]	Calls mpiTest on each element of realist until one of them returns True

Table C

[0069] The mpiISend[] command can be called from within a kernel module 206 (for example, a Mathematica kernel). It creates a packet containing the Mathematica expression to be sent as payload and where the expression should be sent. The packet itself is destined only for its local cluster node module. Once received by its local cluster node module, this packet is decoded and its payload is forwarded on to the cluster node module specified in the packet.

[0070] The mpiIRecv[] command can also be called from within a kernel module 206. It creates a packet specifying where it expects to receive an expression and from which processor this expression is expected. Once received by its local cluster node module, this packet is decoded and its contents are stored in a message receiving queue (MRQ) 308 (FIG. 3).

[0071] The mpiTest[] command can be called from within a kernel module 206. It creates a packet specifying which message to test for completion, then waits for a reply expression to evaluate. Once received by the kernel module's associated cluster node module

204, this packet is decoded and its message specifier is used to search for any matching expressions listed as completed in its received message queue (RMQ) 306. If such completed expressions are found, it is sent to its local kernel module as part of the reply in mpiTest[]. The kernel module receives this reply expression and evaluates it, which updates the kernel module's variables as needed.

[0072] Other MPI calls are built on the fundamental calls mpiISend, mpiIRecv, and mpiTest. For example, mpiBcast, a broadcast, creates instructions to send information from the broadcast processor to all the others, while the other processors perform a Recv. Similarly, high-level calls of the toolkit can be built on top of the collection of MPI calls.

Collective MPI Calls

[0073] In one embodiment, the MPI module 302 can include program code for implementing collective MPI calls (for example, calls that provide basic multi-node data movement across nodes). Collective MPI calls can include broadcasts, gathers, transpose, and other vector and matrix operations, for example. Collective calls can also provide commonly used mechanisms to send expressions between groups of nodes.

Call	Description
mpiBcast[expr, root, comm]	Performs a broadcast of expr from the root processor to all the others in the communicator world comm. An expression is expected to be supplied by the root processor, while all the others expect expr to be overwritten by the incoming expression.
mpiGather[sendexpr, recvexpr, root, comm]	All processors (including root) in the communicator comm send their expression in sendexpr to the root processor, which produces a list of these expressions, in the order according to comm, in recvexpr. On the processors that are not root, recvexpr is ignored.
mpiAllgather[sendexpr, recvexpr, comm]	All processors in the communicator comm send their expression in sendexpr, which are organized into a list of these expressions, in the order according to comm, in recvexpr on all processors in comm.
mpiScatter[sendexpr, recvexpr, root, comm]	Processor root partitions the list in sendexpr into equal parts (if possible) and places each piece in recvexpr on all the processors (including root) in the communicator world comm, according the order and size of comm.
mpiAlltoall[sendexpr, recvexpr, comm]	Each processor sends equal parts of the list in sendexpr to all other processors in the communicator world comm, which each collects from all other processors are organizes into the order according to comm.

Table D

In one embodiment, the MPI module 302 includes program code for implementing parallel sums and other reduction operations on data stored across many nodes. MPI module 302 can also include program code for implementing simple parallel input/output calls (for example, calls that allow cluster system 200 to load and store objects that are located on a plurality of nodes).

Call	Description
mpiReduce[sendexpr, recvexpr, operation, root, comm]	Performs a collective reduction operation between expressions on all processors in the communicator world comm for every element in the list in sendexpr returning the resulting list in recvexpr on the processor with the ID root.
mpiAllreduce[sendexpr, recvexpr, operation, comm]	Performs a collective reduction operation between expressions on all processors in the communicator world comm for every element in the list in sendexpr returning the resulting list in recvexpr on every processor.
mpiReduceScatter [sendexpr, recvexpr, operation, comm]	Performs a collective reduction operation between expressions on all processors in the communicator world comm for every element in the list in sendexpr, partitioning the resulting list into pieces for each processor's recvexpr.

Table E

[0074] These additional collective calls perform operations that reduce the data in parallel. The operation argument can be one of the constants below.

Constant	Description
mpiSum	Specifies that all the elements on different processors be added together in a reduction call
mpiMax	Specifies that the maximum of all the elements on different processors be chosen in a reduction call
mpiMin	Specifies that the minimum of all the elements on different processors be chosen in a reduction call

Table F

MPI Communicator Calls

[0075] In one embodiment, the MPI module 302 includes program code for implementing communicator world calls (for example, calls that would allow subsets of nodes to operate as if they were a sub-cluster). Communicators organize groups of nodes into

user-defined subsets. The communicator values returned by mpiCommSplit[] can be used in other MPI calls instead of mpiCommWorld.

Call	Description
mpiCommSize[comm]	Returns the number of processors within the communicator comm
mpiCommRank[comm]	Returns the rank of this processor in the communicator comm
mpiCommDup[comm]	Returns a duplicate communicator of the communicator comm
mpiCommSplit[comm, color, key]	Creates a new communicator into several disjoint subsets each identified by color. The sort order within each subset is first by key, second according to the ordering in the previous communicator. Processors not meant to participate in any new communicator indicates this by passing the constant mpiUndefined. The corresponding communicator is returned to each calling processor.
mpiCommMap[comm] mpiCommMap[comm, target]	Returns the mapping of the communicator comm to the processor indexed according to $mpiCommWorld. Adding a second argument returns just the ID of the processor with the ID target in the communicator comm.
mpiCommFree[comm]	Frees the communicator comm

Table G

Other MPI Support Calls

[0076] Other calls that provide common functions include:

Call	Description
mpiWtime[]	Provides wall-dock time since some fixed time in the past. There is no guarantee that this time will read the same on all processors.
mpWtick[]	Returns the time resolution of mpiWtime[]
MaxByElement[in]	For every nth element of each list of the list in, chooses the maximum according to Max[], and returns the result as one list. Used in the mpiMax reduction operation.
MinByElement[in]	For every nth element of each list of the list in, chooses the minimum according to Min[], and returns the result as one list. Used in the mpiMin reduction operation.

Table H

2. Advanced Functions Module 304

[0077] In one embodiment, the cluster node module 204 includes an advanced functions module 304. The advanced functions module 304 can include program code that provides a toolkit of functions inconvenient or impractical to do with MPI instructions and calls implemented by the MPI module 302. The advanced functions module 304 can rely at least partially on calls and instructions implemented by the MPI module 302 in the implementation of advanced functions. In one embodiment, the advanced functions module 304 includes a custom set of directives or functions. In an alternative embodiment, the advanced functions module 304 intercepts normal Mathematica language and converts it to one or more functions optimized for cluster execution. Such an embodiment can be easier for users familiar with Mathematica functions to use but can also complicate a program debugging process. Some functions implemented by the advanced functions module 304 can simplify operations difficult or complex to set up using parallel computing. Several examples of such functions that can be implemented by the advanced functions module 304 are shown below.

[0078] Built on the MPI calls, the calls that are described below provide commonly used communication patterns or parallel versions of Mathematica features. Unless

otherwise specified, these are executed in the communicator mpiCommWorld, whose default is $mpiCommWorld, but can be changed to a valid communicator at run time.

Common Divide-and-Conquer Parallel Evaluation

[0079] In one embodiment, the advanced functions module 304 includes functions providing for basic parallelization such as, for example, routines that would perform the same operations on many data elements or inputs, stored on many nodes. These functions can be compared to parallelized for-loops and the like. The following calls address simple parallelization of common tasks. In the call descriptions, "expr" refers to an expression, and "loopspec" refers to a set of rules that determine how the expression is evaluated. In some embodiments, the advanced functions module 304 supports at least three forms of loopspec, including {var, count}, where the call iterates the variable *var* from 1 to the integer *count*; {var, start, stop}, where the call iterates the variable *var* every integer from *start* to *stop*; and {var, start, stop, increment}, where the call iterates the variable *var* from *start* adding *increment* for each iteration until *var* exceeds *stop*, allowing *var* to be a non-integer.

Call	Description
ParallelDo[expr, loopspec]	Like Do[] except that it evaluates expr across the cluster, rather than on just one processor. The rules for how expr is evaluated is specified in loopspec, like in Do[].
ParallelFunctionToList[f, count] ParallelFunctionToList[f, count, root]	Evaluates the function f[i] from 1 to count, but across the cluster, and returns these results in a list. The third argument has it gather this list into the processor whose ID is root.
ParallelTable[expr, loopspec] ParallelTable[expr, loopspec, root]	Like Table[] except that it evaluates expr across the cluster, rather than on just one processor, returning the locally evalulated portion. The third argument has it gather this table in to the processor whose ID is root.
ParallelFunction[f, inputs, root]	Like f[inputs] except that it evaluates f on a subset of inputs scattered across the cluster from processor root and gathered back to root.
ParallelNintegrate[expr, loopspec] ParallelNintegrate[expr, loopspec,digits]	Like Nintegrate[] except that it evaluates a numerical integration of expr over domains partitioned into the number of processors in the cluster, then returns the sum. The third argument has each numerical integration execute with at least that many digits of precision.

Table I

Guard-Cell Management

[0080] In one embodiment, the advanced functions module 304 includes functions providing for guard-cell operations such as, for example, routines that perform nearest-neighbor communications to maintain edges of local arrays in any number of dimensions (optimized for 1-, 2-, and/or 3-D). Typically the space of a problem is divided into partitions. Often, however, neighboring edges of each partition can interact, so a "guard cell" is inserted on both edges as a substitute for the neighboring data. Thus the space a processor sees is two

elements wider than the actual space for which the processor is responsible. EdgeCell helps maintain these guard cells.

Call	Description
EdgeCell[list]	Copies the second element of list to the last element of the left processor and the second-to-last element of list to the first element of the right processor while simultaneously receiving the same from its neighbors.

Table J

Matrix and Vector Manipulation

[0081] The advanced functions module 304 can also include functions providing for linear algebra operations such as, for example, parallelized versions of basic linear algebra on structures partitioned on many nodes. Such linear algebra operations can reorganize data as needed to perform matrix and vector multiplication or other operations such as determinants, trace, and the like. Matrices are partitioned and stored in processors across the cluster. These calls manipulate these matrices in common ways.

Call	Description
ParallelTranspose[matrix]	Like Transpose[] except that it transposes matrix that is in fact represented across the cluster, rather than on just one processor. It returns the portion of the transposed matrix meant for that processor.
ParallelProduct[matrix, vector]	Evaluates the product of matrix and vector, as it would on one processor, except that matrix is represented across the cluster.
ParallelDimensions[matrix]	Like Dimensions[] except that matrix is represented across the cluster, rather than on just one processor. It returns a list of each dimension.
ParallelTr[matrix]	Like Tr[] except that the matrix is represented across the cluster, rather than on just one processor. It returns the trace of this matrix.
ParallelIdentity[rank]	Like Identity[], it generates a new identity matrix, except that the matrix is represented across the cluster, rather than on just one processor. It returns the portion of the new matrix for this processor.
ParallelOuter[f, vector1, vector2]	Like Outer[f, vectorl, vector2] except that the answer becomes a matrix represented across the cluster, rather than on just one processor. It returns the portion of the new matrix for this processor.
ParallelInverse[matrix]	Like Inverse[] except that the matrix is represented across the cluster, rather than on just one processor. It returns the inverse of the matrix.

Table K

Element Management

[0082] In one embodiment, the advanced functions module 304 includes element management operations. For example, a large bin of elements or particles cut up in space across the nodes may need to migrate from node to node based on rules or criteria (such as

their spatial coordinate). Such operations would migrate the data from one node to another. Besides the divide-and-conquer approach, a list of elements can also be partitioned in arbitrary ways. This is useful if elements need to be organized or sorted onto multiple processors. For example, particles of a system may drift out of the space of one processor into another, so their data would need to be redistributed periodically.

Call	Description
ElementManage[list, switch]	Selects which elements of list will be sent to which processors according to the function switch[] is evaluated on each element of list. If switch is a function, switch[] should return the ID of the processor that element should be sent. If switch is an integer, the call assumes that each elements is itself a list, whose first element is a number ranging from 0 to the passed argument. This call returns a list of the elements, from any processor, that is switch selected for this processor.
ElementManage[list]	Each element of list can be a list of two elements, the first being the ID of the processor where the element should be sent, while the second is arbitrary data to send. This call returns those list elements, from any and all processors, whose first element is this processors ID in a list. This call is used internally by the two-argument version of ElementManage[].

Table L

Fourier Transform

[0083] In one embodiment, the advanced functions module 304 includes program code for implementing large-scale parallel fast Fourier transforms ("FFTs"). For example, such functions can perform FFTs in one, two, and/or three dimensions on large amounts of data that are not stored on one node and that are instead stored on many nodes. Fourier transforms of very large arrays can be difficult to manage, not the least of which is the

memory requirements. Parallelizing the Fourier transform makes it possible to make use of all the memory available on the entire cluster, making it possible to manipulate problem sizes that no one processor could possibly do alone.

Call	Description
ParallelFourier[list]	Like Fourier[] except that list is a two- or three-dimensional list represented across the cluster, like for matrices, above. It returns the portion of the Fourier-transformed array meant for that processor.

Table M

Parallel Disk I/O

[0084] In one embodiment, the advanced functions module 304 includes parallel disk input and output calls. For example, data may need to be read in and out of the cluster in such a way that the data is distributed across the cluster evenly. The calls in the following table enable the saving data from one or more processors to storage and the retrieval data from storage.

Call	Description
ParallelPut[expr, filename] ParallelPut[expr, filename root] ParallelPut[expr, filename, root, comm]	Puts expr into the file with the name filename in order on processor 0. The third argument specifies that the file be written on the processor whose ID is root. The fourth uses the communicator world comm.
ParallelGet[filename] ParallelGet[filename, root] ParallelGet[filename, root,comm]	Reads and returns data from the file with the name filename on processor 0 partitioned into each processor on the cluster. The second argument specifies that the file is to be read on the processor whose ID is root. The third uses the communicator world comm.

Call	Description
ParallelBinaryPut[expr, type,filename] ParallelBinaryPut[expr, filename, root] ParallelBinaryPut[expr, filename, root, comm]	Puts expr into the file with the binary format type with the name filename in order on processor 0. The fourth argument specifies that the file be written on the processor whose ID is root. The fifth uses the communicator world comm.
ParallelBinaryGet[type, filename] ParallelBinaryGet[type, filename, root] ParallelBinaryGet[type, filename, root, comm]	Reads and returns data in the binary format type from the file with the name filename on processor 0 partitioned into each processor on the cluster. The third argument specifies that the file is to be read on the processor whose ID is root. The fourth uses the communicator world comm.
ParallelGetPerProcessor [expr, filename] ParallelGetPerProcessor [filename, root] ParallelGetPerProcessor [filename, root, comm]	Puts expr into the file with the name filename in order on processor 0, one line per processor. The third argument specifies that the file be written on the processor whose ID is root. The fourth uses the communicator world comm.
ParallelGetPerProcessor [filename] ParallelGetPerProcessor [filename, root] ParallelGetPerProcessr [filename, root, comm]	Reads and returns data from the file with the name filename on processor 0, one line for each processor. The second argument specifies that the file is to be read on the processor whose ID is root. The third uses the communicator world comm.

Table N

Automatic Load Balancing

[0085] Some function calls can take an inconsistent amount of processing time to complete. For example, in Mathematica, the call f[20] could in general take much longer to

evaluate than f[19]. Moreover, if one or more processors within the cluster are of different speeds (for example, if some operate at a core frequency of 2.6 GHz while other operate at less than one 1 GHz), one processor may finish a task sooner than another processor.

[0086] In some embodiments, the advanced functions module 304 includes a call that can improve the operation of the computer cluster 100 in such situations. In some embodiments, the root processor assigns a small subset of the possible calls for a function to each processor on the cluster 100. Whichever processor returns its results first is assigned a second small subset of the possible calls. The root processor will continue to assign small subsets of the possible calls as results are received until an evaluation is complete. The order in which the processors finish can vary every time an expression is evaluated, but the root processor will continue assigning additional work to processors as they become available.

[0087] In one illustrative example, there are 4 processors and f[1] to f[100] to evaluate. One could implement this by assigning f[1], f[2], f[3], f [4] to each of processors 0 (the root can assign to oneself) through 3. If the f[2] result came back first, then processor 1 would be assigned f[5]. If the f[4] result is returned next, f[6] would be assigned to processor 3. The assignments continue until all results are calculated. The results are organized for output back to the user.

[0088] In alternative embodiments, the subsets of possible calls can be assigned in any order, rather than sequentially, or in batches (for example, f[1], f[5], f[9] assigned to processor 1, etc.). Also, the subsets could be organized by delegation. For example, one processor node may not necessarily be in direct control of the other processors. Instead, a large subset could be assigned to a processor, which would in turn assign subsets of its work to other processors. The result would create a hierarchy of assignments like a vast army.

Call	Description
LoadBalanceFunctionToList[f, count] LoadBalanceFunctionToList[f, count, root]	Evaluates the function f[i] from 1 to count, but across the cluster using load-balancing techniques, and returns these results in a list. The third argument has it gather this list into the processor whose ID is root.

Table O

3. <u>Received Message Queue 306</u>

[0089] In one embodiment, the cluster node module 204 includes a received message queue 306. The received message queue 306 includes a data structure for storing messages received from other cluster node modules. Related data pertaining to the messages received, such as whether an expression has been completed, may also be stored in the received message queue 306. The received message queue 306 may include a queue and/or another type of data structure such as, for example, a stack, a linked list, an array, a tree, etc.

4. Message Receiving Queue 308

[0090] In one embodiment, the cluster node module 204 includes a message receiving queue 308. The message receiving queue 308 includes a data structure for storing information about the location to which an expression is expected to be sent and the processor from which the expression is expected. The message receiving queue 308 may include a queue and/or another type of data structure such as, for example, a stack, a linked list, an array, a tree, etc.

B. Cluster Configuration Module 208

[0091] Cluster configuration module 208 includes program code for initializing a plurality of cluster node modules to add cluster computing support to computer systems 110, 120, 130. United States Patent No. 7,136,924, issued to Dauger (the "'924 patent"), the entirety of which is hereby incorporated by reference and made a part of this specification, discloses a method and system for parallel operation and control of computer clusters. One method generally includes obtaining one or more personal computers having an operating system with discoverable network services. In some embodiments, the method includes obtaining one or more processors or processor cores on which a kernel module can run. As described in the '924 patent, a cluster node control and interface (CNCI) group of software applications is copied to each node. When the CNCI applications are running on a node, the cluster configuration module 208 can permit a cluster node module 204, in combination with a kernel module 206, to use the node's processing resources to perform a parallel computation task as part of a computer cluster. The cluster configuration module 208 allows extensive automation of the cluster creation process in connection with the present disclosure.

C. User Interface Module 202

[0092] In some embodiments, computer cluster 100 includes a user interface module 202, such as, for example a Mathematica Front End or a command line interface, that includes program code for a kernel module 206 to provide graphical output, accept graphical input, and provide other methods of user communication that a graphical user interface or a command-line interface provides. To support a user interface module 202, the behavior of a cluster node module 204a is altered in some embodiments. Rather than sending output to and accepting input from the user directly, the user interface module 202 activates the cluster node module 204a to which it is connected and specifies parameters to form a connection, such as a MathLink connection, between the cluster node module 204a and the user interface module 202. The user interface module's activation of the cluster node module 204a can initiate the execution of instructions to activate the remaining cluster node modules 204b-e on the cluster and to complete the sequence to start all kernel modules 206a-e on the cluster. Packets from the user interface module 202, normally intended for a kernel module 206a, are accepted by the cluster node module 204a as a user command. Output from the kernel module 206a associated with the cluster node module 204a can be forwarded back to the user interface module 202 for display to a user. Any of the cluster node modules 204a-e can be configured to communicate with a user interface module 202.

D. Kernel Module 206

[0093] A kernel module 206 typically includes program code for interpreting high-level code, commands, and/or instructions supplied by a user or a script into low-level code, such as, for example, machine language or assembly language. In one embodiment, each cluster node module 204a-e is connected to all other cluster node modules, while each kernel module 206a-e is allocated and connected only to one cluster node module 204. In one embodiment, there is one cluster node module-kernel module pair per processor. For example, in an embodiment of a computer cluster 100 including single-processor computer systems, each cluster node module-kernel module pair could reside on a single-processor computer. If a computer contains multiple processors or processing cores, it may contain multiple cluster node module-kernel module pairs, but the pairs can still communicate over the cluster node module's network connections.

IV. Cluster Computing Methods

[0094] In one embodiment, the computer cluster 100 includes a cluster initialization process, a method of cluster node module operation, and a cluster shut down process.

A. Cluster Initialization Process

[0095] In one embodiment, a cluster configuration module 202 initializes one or more cluster node modules 204 in order to provide cluster computing support to one or more kernel modules 206, as shown in FIG. 4.

[0096] At 402, cluster node modules are launched on the computer cluster 100. In one embodiment, the cluster node module 204a running on a first processor 112a (for example, where the user is located) accesses the other processors 112b, 122a-b, 132 on the computer cluster 100 via the cluster configuration module 208 to launch cluster node modules 204b-e onto the entire cluster. In an alternative embodiment, the cluster configuration module 208 searches for processors 112a-b, 122a-b, 132 connected to one another via communications network 102 and launches cluster node modules 204a-e on each of the processors 112a-b, 122a-b, 132.

[0097] The cluster node modules 204a-e establish communication with one another at 404. In one embodiment, each of the cluster node modules 204a-e establish direct connections using the MPI_Init command with other cluster node modules 204a-e launched on the computer cluster 100 by the cluster configuration module 208.

[0098] At 406, each cluster node module 204 attempts to connect to a kernel module 206. In one embodiment, each instance of the cluster node modules 204a-e locates, launches, and connects with a local kernel module via MathLink connections and/or similar connection tools, for example, built into the kernel module 206.

[0099] At 408, the cluster node modules 204 that are unconnected to a kernel module 206 are shut down. In one embodiment, each cluster node module 204 determines whether the local kernel module cannot be found or connected to. In one embodiment, each cluster node module 204 reports the failure to connect to a kernel module 206 to the other cluster node modules on computer cluster 100 and quits.

[0100] Processor identification numbers are assigned to the remaining cluster node modules 204 at 410. In one embodiment, each remaining cluster node module 204 calculates the total number of active processors (N) and determines identification numbers describing the remaining subset of active cluster node modules 204a-e and kernel modules 206a-e. This new set of cluster node module-kernel module pairs may be numbered 0 through N-1, for example.

[0101] Message passing support is initialized on the kernel modules 206a-e at 412. In one embodiment, each cluster node module 204 supplies initialization code (for example, Mathematica initialization code) to the local kernel module 206 to support message passing.

[0102] Finally, at 414, the cluster node modules 204a-e enter a loop to accept user entry. In one embodiment, a main loop (for example, a cluster operation loop) begins execution after the cluster node module 204a on the first processor 112a returns to user control while each of the other cluster node modules 204 waits for messages from all other cluster node modules 204a-e connected to the network 102.

[0103] The initialization process creates a structure enabling a way for the kernel modules 206a-e to send messages to one another. In some embodiments, any kernel module can send data to and receive data from any other kernel module within the cluster when initialization is complete. The cluster node module creates an illusion that a kernel module is communicating directly with the other kernel modules. The initialization process can create a relationship among kernel modules on a computer cluster 100 such as the one shown by way of example in FIG. 2.

B. <u>Cluster Node Module Operation</u>

[0104] In one embodiment, a cluster node module 204 implements cluster computing support for a kernel module 206 during a main loop, as shown in FIG. 5.

[0105] At 502, cluster node modules 204 wait for user commands or messages from other cluster node modules. In one embodiment, the cluster node module 204a connected to the user interface module 202 waits for a user command, while the other cluster node modules 204b-e continue checking for messages.

[0106] Once a command or message is received, the method proceeds to 504. At 504, the cluster node module 204a determines whether the message received is a quit command. If a quit command is received, the cluster node module 204a exits the loop and proceeds to a cluster node module shut down process at 505. If the message received is not a quit command, the process continues to 506.

[0107] At 506, received commands are communicated to all cluster node modules 204a-e on the computer cluster 100. In one embodiment, when a user enters a command in the user interface module 202, the cluster node module 204a connected to the user interface module 202 submits the user command to all other cluster node modules 204b-e in the computer cluster 100. The user commands can be simple (for example, "1+1"), but can also be entire subroutines and sequences of code (such as, for example, Mathematica code), including calls to MPI from within the user interface module 202 (for example, the Mathematica Front End) to perform message passing between kernel modules 206a-e (for example, Mathematica kernels). These include the fundamental MPI calls, which are implemented using specially identified messages between a cluster node module 204 and its local kernel module 206.

[0108] The message (or user command) is communicated to the kernel modules 206a-e at 508. In one embodiment, the cluster node module 204a connected to the user interface module 202 submits the user command to the kernel module 206a to which it is connected. Each of the other cluster node modules 204b-e, after receiving the message, submits the command to the respective kernel module 206b-e to which it is connected.

[0109] At 510, a cluster node module 204 receives a result from a kernel module 206. In one embodiment, once the kernel module 206 completes its evaluation, it returns the kernel module's output to the cluster node module 204 to which it is connected. Depending on the nature of the result from the kernel module, the cluster node module 204 can report the result to a local computer system or pass the result as a message to another cluster node module 204. For example, the cluster node module 204a running on the first processor 112a reports the output on its local computer system 110. For example, on the first processor 112a, cluster node module 204a only directly reports the output of kernel module 206a.

[0110] Messages from other cluster node modules 204 are responded to at 512. In one embodiment, each cluster node module (for example, the cluster node module 204a) checks for and responds to messages from other cluster node modules 204b-e and from the kernel module 206a repeatedly until those are exhausted. In one embodiment, output messages from the kernel module 206 are forwarded to output on the local computer system. Messages from other cluster node modules 204 are forwarded to a received message queue 306 ("RMQ"). Data from each entry in the message receiving queue 308 ("MRQ") is matched with entries in the RMQ 306 (see, for example, description of the mpiIRecv[] call, above). If found, data from the MRQ 308 are combined into those in the RMQ 306 and marked as "completed" (see, for example, description of the mpiTest[] call, above). This process provides the peer-to-peer behavior of the cluster node modules 204a-e. Via this mechanism, code running within multiple, simultaneously running kernel modules (for example, Mathematica kernels) can interact on a pair-wise or collective basis, performing calculations, processing, or other work on a scale larger and/or faster than one kernel could have done alone. In this manner, user-entered instructions and data specifying what work will be done via user commands can be executed more quickly and/or reliably. Once responding to messages has completed, the process returns to 502.

[0111] In some embodiments, a computer system includes software, such as an operating system, that divides memory and/or other system resources into a user space, a kernel space, an application space (for example, a portion of the user space allocated to an application program), and/or an operating system space (for example, a portion of the user space allocated to an operating system). In some embodiments, some or all of the cluster node modules 204a-e are implemented in the application space of a computer system. In further embodiments, at least some of the cluster node modules 204a-e are implemented in the operating system space of a computer system. For example, some cluster node modules in a computer cluster may operate in the application space while others operate in the operating system space.

[0112] In some embodiments, some or all of the functionality of the cluster node modules 204a-e is incorporated into or integrated with the operating system. The operating system can add cluster computing functionality to application programs, for example, by

implementing at least some of the methods, modules, data structures, commands, functions, and processes discussed herein. Other suitable variations of the techniques described herein can be employed, as would be recognized by one skilled in the art.

[0113] In some embodiments, the operating system or components of the operating system can identify and launch the front end 202 and the kernels 206. The operating system or its components can connect the front end 202 and kernels 206 to one another in the same manner as a cluster node module 204 would or by a variation of one of the techniques described previously. The operating system can also be responsible for maintaining the communications network 102 that connects the modules to one another. In some embodiments, the operating system implements at least some MPI-style calls, such as, for example, collective MPI-style calls. In some embodiments, the operating system includes an application programming interface (API) library of cluster subroutine calls that is exposed to application programs. Applications programs can use the API library to assist with launching and operating the computer cluster.

C. Cluster Shut Down Process

[0114] In one embodiment, a computer cluster 100 includes a procedure to shut down the system. If the operation process (or main loop) on the cluster node module 204a connected to the user interface module 202 detects a "Quit" or "Exit" command or otherwise receives a message from the user indicating a shut down, the sequence to shut down the cluster node modules 204a-e and the kernel modules 206a-e is activated. In one embodiment, the cluster node module 204a connected to the user interface module 202 sends a quit message to all other cluster node modules 204b-e. Each cluster node module 204 forwards the quit command to its local kernel module 206. Once its Mathematica kernel has quit, each cluster node module 204 proceeds to tear down its communication network with other cluster node modules (for example, see description of the MPI_Finalize command, above). At the conclusion of the process, each cluster node module 204 exits execution.

V. Example Operation

[0115] For purposes of illustration, sample scenarios are discussed in which the computer cluster system is used in operation. In these sample scenarios, examples of

Mathematica code are given, and descriptions of how the code would be executed by a cluster system are provided.

Basic MPI

[0116] Fundamental data available to each node includes the node's identification number and total processor count.

```
In[1]:= {$IdProc, $NProc}
Out[1]:= {0, 2}
```

[0117] The first element should be unique for each processor, while the second is generally the same for all. Processor 0 can see what other values are using a collective (see below) communications call such as mpiGather[].

```
In[2]:= mpiGather[{$IdProc, $NProc},list,0]; list
Out[2]:= {{0, 2}, {1, 2}}
```

Peer-to-Peer MPI

[0118] The mpiSend and mpiRecv commands make possible basic message passing, but one needs to define which processor to target. The following defines a new variable, targetProc, so that each pair of processors will point to each other.

```
In[3]:= targetProc=If[1==Mod[$IdProc, 2],$IdProc-1,$IdProc+1]
Out[3]:= 1
```

[0119] In this example, the even processors target its "right" processor, while the odd ones point its "left." For example, if the processors were lined up in a row and numbered in order, every even-numbered processor would pair with the processor following it in the line, and every odd-numbered processor would pair with the processor preceding it. Then a message can be sent:

```
In[4]:= If [ 1==Mod[ $IdProc , 2],mpiSend[N[Pi,22],targetProc,
      mpiCommWorld,d], mpiRecv[a,targetProc,mpiCommWorld,d]]
```

[0120] The If [] statement causes the processors to evaluate different code: the odd processor sends 22 digits of Pi, while the even receives that message. Note that these MPI calls return nothing. The received message is in the variable *a*:

```
In[5]:= a
Out[5]:= 3.1415926535897932384626
In[6]:= Clear[a]
```

[0121] The variable *a* on the odd processors would have no definition. Moreover, if $NProc is 8, processor 3 sent Pi to processor 2, processor 5 sent Pi to processor 4, and so on. These messages were not sent through processor 0, but they communicated on their own.

[0122] The mpiISend and mpiIRecv commands have a letter "I" to indicate asynchronous behavior, making it possible to do other work while messages are being sent and received, or if the other processor is busy. So, the above example could be done asynchronously:

```
In[7]:= If[1==Mod[$IdProc, 2],mpiISend[N[Pi,22],targetProc,
      mpiCommWorld,d,e],
      mpiIRecv[a,targetProc,mpiCommWorld,d,e]]
```

[0123] The variable *e* has important data identifying the message, and mpiTest[*e*] can return True before the expressions are to be accessed. At this point, many other evaluations can be performed. Then, one can check using mpiTest when the data is needed:

```
In[29]:= mpiTest[e]
Out[29]:= True
In[30]:= a
Out[30]:= 3.1415926535897932384626
In[31]:= Clear[a,e]
```

[0124] The mpiWait[*e*] command could have also have been used, which does not return until mpiTest[*e*] returns True. The power of using these peer-to-peer calls is that it becomes possible to construct any message-passing pattern for any problem.

Collective MPI

[0125] In some cases, such explicit control is not required and a commonly used communication pattern is sufficient. Suppose processor 0 has an expression in *b* that all processors are meant to have. A broadcast MPI call would do:

```
In[8]:= mpiBcast[b, 0, mpiCommWorld]
```

[0126] The second argument specifies which processor is the "root" of this broadcast; all others have their *b* overwritten. To collect values from all processors, use mpiGatherD:

```
In[9]:= mpiGather[b, c, 0, mpiCommWorld]
```

[0127] The variable *c* of processor 0 is written with a list of all the *b* of all the processors in mpiCommWorld. The temporal opposite is mpiScatter:

```
In[10]:= Clear[b]; a = {2, 4, 5, 6}; mpiScatter[a, b, 0,
      mpiCommWorld]; b
Out[10]:= {2, 4}
```

[0128] The mpiScatter command cuts up the variable *a* into even pieces (when possible) and scatters them to the processors. This is the result if $NProc = 2, but if $NProc = 4, *b* would only have {2}.

[0129] MPI provides reduction operations to perform simple computations mixed with messaging. Consider the following:

```
In[11]:= a = {{2 + $IdProc, 45[],3,{1 + $IdProc,$NProc[][];
      mpiReduce [a,d,mpiSum, 0,mpiCommWorld ]
In[12]:= d
Out[12]:= {{5, 90}, 6, {3, 4}}
```

[0130] The mpiSum constant indicates that variable *a* of every processor will be summed. In this case, $NProc is 2, so those elements that were not identical result in odd sums, while those that were the same are even.

[0131] Most of these calls have default values if not all are specified. For example each of the following calls will have the equivalent effect as the above mpiGather[] call:

```
mpiGather[b, c, 0]
mpiGather[b, c]
c = mpiGather[b]
```

High-Level Calls

[0132] High-level calls can include convenient parallel versions of commonly used application program calls (for example, Mathematica calls). For example, ParallelTable[] is like Table[], except that the evaluations are automatically performed in a distributed manner:

```
In[13]:= ParallelTable[i,{i,100},0]
Out[13]:= {1,2,3,4,5, ··· ,99,100}
```

[0133] The third argument specifies that the answers are collated back to processor 0. This is a useful, simple way to parallelize many calls to a complex function. One could define a complicated function and evaluate it over a large range of inputs:

```
In[14]:= g[x_] := Gamma[2 + 0.5*(x-1)];
      ParallelTable[g[i],{i,100},0]
```

```
Out[14]:= {1, 1.32934, 2., 3.32335, 6., 11.6317, 24., 52.3428,
      120., 287.885, 720}
```

[0134] ParallelFunctionToList[] also provides a simplified way to perform this form of parallelism.

<u>Operations with Non-Trivial Communication</u>

Matrix Operations

[0135] In some embodiments, one or more functions can help solve matrix calculations in parallel:

```
In[15]:= a = Table[i+ 3* $IdProc + 2 j, {i, 2}, {j,4}]
Out[15]:= {{3, 5, 7, 9}, {4, 6, 8, 10}}
In[16]:= t = ParallelTranspose[a]
Out[16]:= {{3, 4, 6, 7}, {5, 6, 8, 9}}
```

Fourier Transforms

[0136] A Fourier transform of a large array can be solved faster in parallel, or made possible on a cluster because it can all be held in memory. A two-dimensional Fourier transform of the above example follows:

```
In[17]:= f = ParallelFourier[a]
Out[17]:= {{32. + 0. I, -4. - 4. I, -4., -4. + 4. I}, {-3. -
      3. I, 0. + 0. I, 0., 0. + 0. I}}
```

Edge Cell Management

[0137] Many problems require interactions between partitions, but only on the edge elements. Maintaining these edges can be performed using EdgeCell[].

```
In[18]:= a = {2, 4, 5, 6, 7}+8*$IdProc
Out[18]:= {2, 4, 5, 6, 7}
In[19]:= EdgeCell[a]; a
Out[19]:= {14, 4, 5, 6, 12}
```

Element Management

[0138] In particle-based problems, items can drift through space, sometimes outside the partition of a particular processor. This can be solved with ElementManage[l:

```
In[20]:= list={{0,4},{1,3},{1,4},{0,5}}; fcn[x_]:=x[[1]]
In[21]:= ElementManage[list, fcn]
Out[21]:= {{0, 4}, {0, 5}, {0, 4}, {0, 5}}
In[21]:= ElementManage[list, 2]
```

```
Out[21]:= {{0, 4}, {0, 5}, {0, 4}, {0, 5}}
```

[0139] The second argument of ElementManage describes how to test elements of a list. The *fcn* identifier returns which processor is the "home" of that element. Passing an integer assumes that each element is itself a list, whose first element is a number ranging from 0 to the passed argument.

[0140] While the examples above involve Mathematica software and specific embodiments of MPI calls and cluster commands, it is recognized that these embodiments are used only to illustrate features of various embodiments of the systems and methods.

VI. Additional Embodiments

[0141] Although cluster computing techniques, modules, calls, and functions are disclosed with reference to certain embodiments, the disclosure is not intended to be limited thereby. Rather, a skilled artisan will recognize from the disclosure herein a wide number of alternatives for the exact selection of cluster calls, functions, and management systems. For example, single-node kernels can be managed using a variety of management tools and/or can be managed manually by a user, as described herein. As another example, a cluster node module can contain additional calls and procedures, including calls and procedures unrelated to cluster computing, that are not disclosed herein.

[0142] Other embodiments will be apparent to those of ordinary skill in the art from the disclosure herein. Moreover, the described embodiments have been presented by way of example only, and are not intended to limit the scope of the disclosure. Indeed, the novel methods and systems described herein can be embodied in a variety of other forms without departing from the spirit thereof. Accordingly, other combinations, omissions, substitutions and modifications will be apparent to the skilled artisan in view of the disclosure herein. Thus, the present disclosure is not intended to be limited by the disclosed embodiments, but is to be defined by reference to the appended claims. The accompanying claims and their equivalents are intended to cover forms or modifications as would fall within the scope and spirit of the inventions.

WHAT IS CLAIMED IS:

1. A system for performing an instruction received from a front end by executing commands on one or more special purpose microprocessors, the system comprising:

a plurality of nodes, wherein each node is configured to access a computer-readable memory system comprising program code for a single-node kernel module, and wherein each single-node kernel module is configured to interpret instructions received by the single-node kernel module into commands that are executable by a special purpose microprocessor;

a plurality of cluster node modules, wherein each cluster node module is stored in a computer-readable memory system and configured to communicate with a single-node kernel and with one or more other cluster node modules, to accept instructions, and to interpret at least some of the instructions such that the plurality of cluster node modules communicate with one another in order to act as a cluster in executing commands using one or more hardware processors; and

a communications system configured to connect the plurality of nodes;

wherein the plurality of cluster node modules cooperate to interpret and translate, as needed, the instruction for execution by a plurality of single-node kernel modules, and wherein at least one of the plurality of cluster node modules returns a result to the front end.

CLUSTER COMPUTING

ABSTRACT

In some embodiments, a computer cluster system comprises a plurality of nodes and a software package comprising a user interface and a kernel for interpreting program code instructions. In certain embodiments, a cluster node module is configured to communicate with the kernel and other cluster node modules. The cluster node module can accept instructions from the user interface and can interpret at least some of the instructions such that several cluster node modules in communication with one another and with a kernel can act as a computer cluster.

17269439

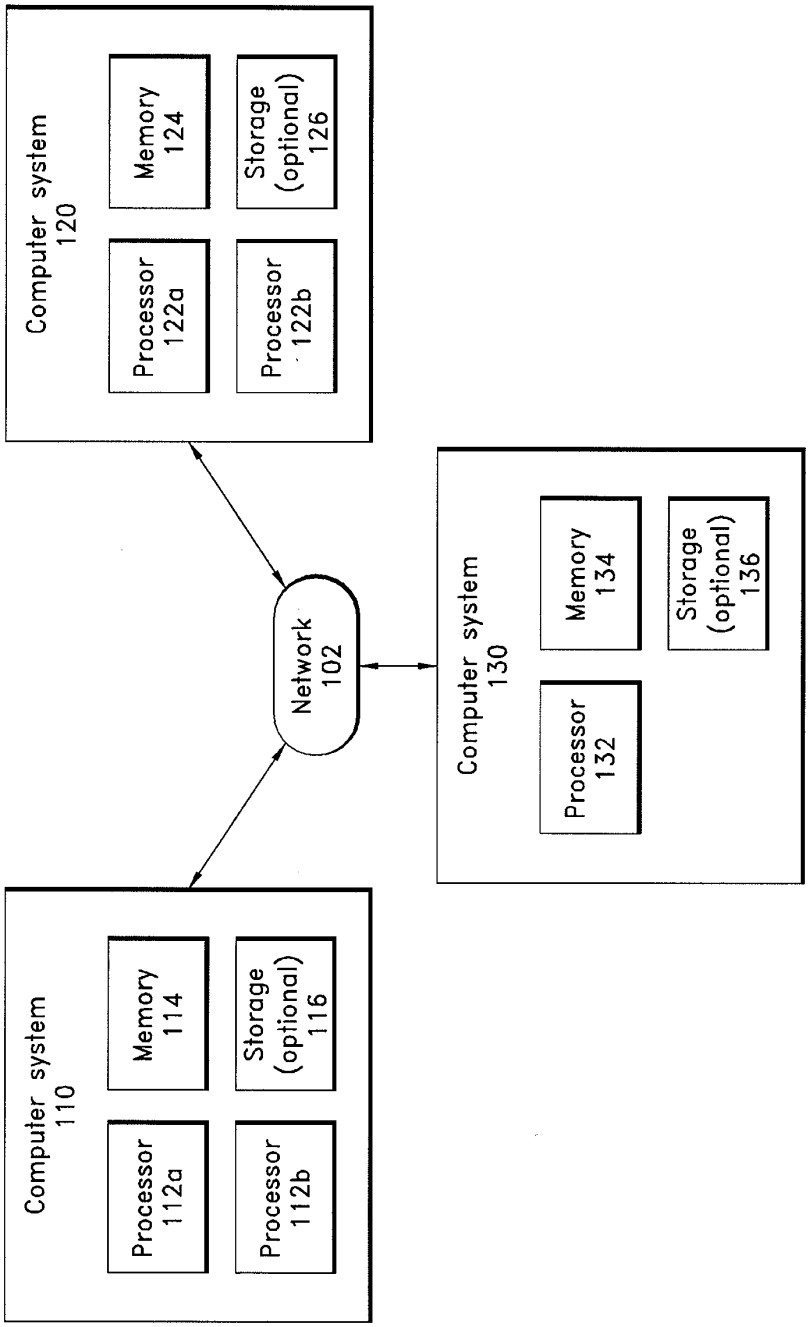

Computer system
120
Processor
122a
Memory
124
Processor
122b
Storage
(optional)
126
Network
102
Computer system
110
Processor
112a
Memory
114
Processor
112b
Storage
(optional)
116
Computer system
130
Processor
132
Memory
134
Storage
(optional)
136

FIG. 1

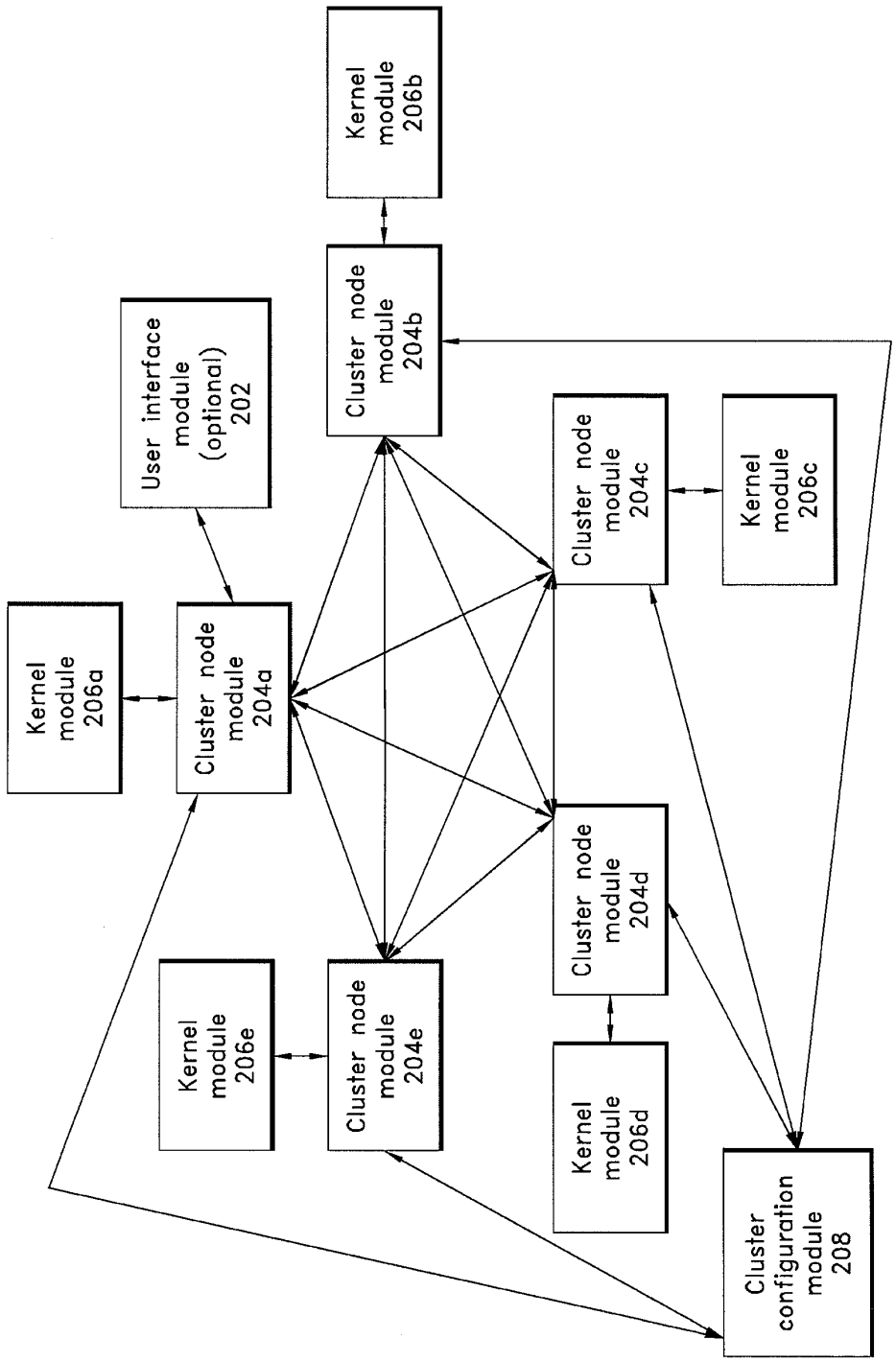

Kernel module 206a
Kernel module 206b
Kernel module 206c
Kernel module 206d
Kernel module 206e
User interface module (optional) 202
Cluster node module 204a
Cluster node module 204b
Cluster node module 204c
Cluster node module 204d
Cluster node module 204e
Cluster configuration module 208

FIG. 2

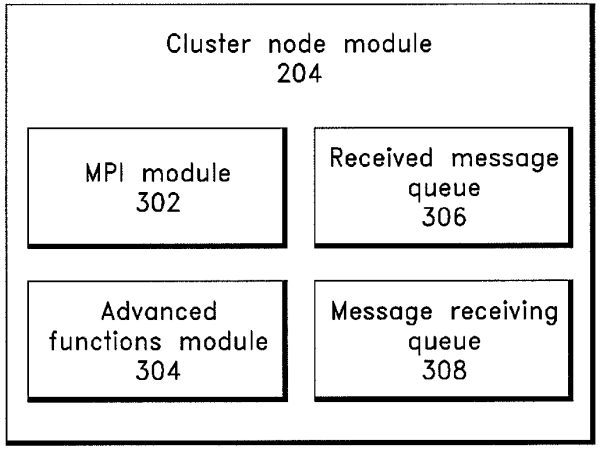

Cluster node module
204
MPI module
302
Received message
queue
306
Advanced
functions module
304
Message receiving
queue
308

FIG. 3

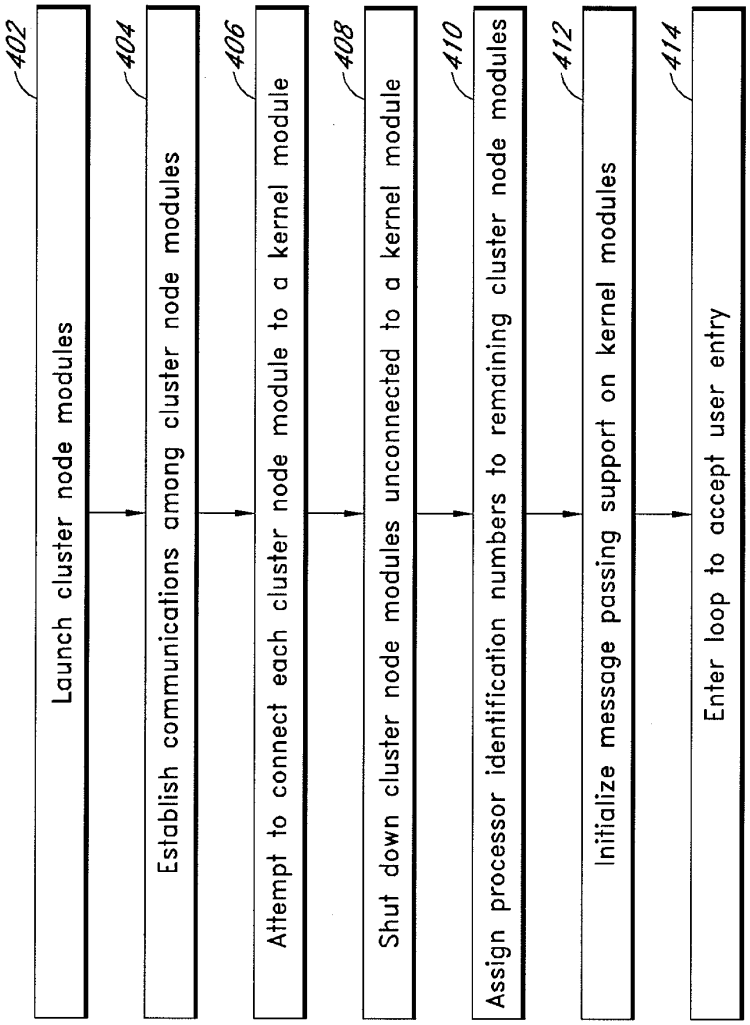

402
Launch cluster node modules
404
Establish communications among cluster node modules
406
Attempt to connect each cluster node module to a kernel module
408
Shut down cluster node modules unconnected to a kernel module
410
Assign processor identification numbers to remaining cluster node modules
412
Initialize message passing support on kernel modules
414
Enter loop to accept user entry

FIG. 4

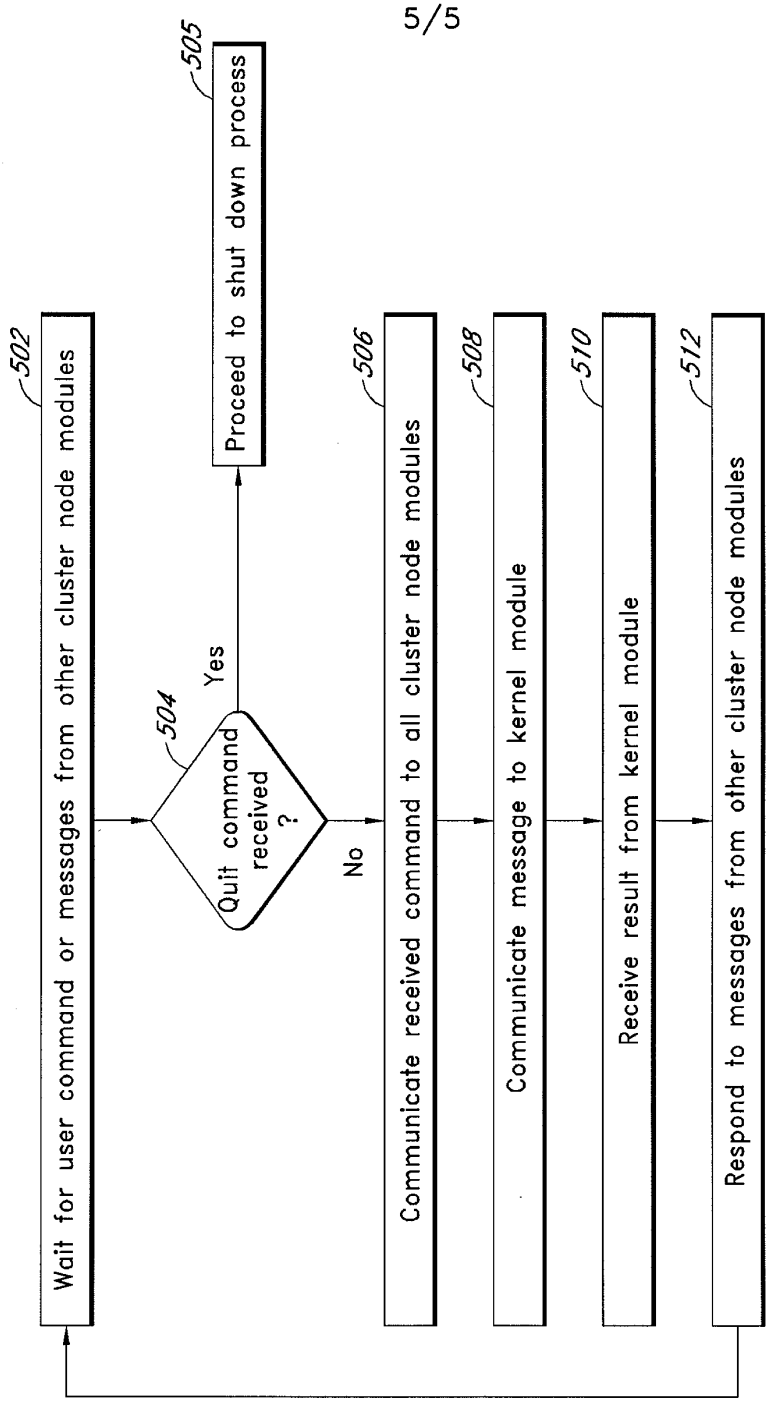

502
Wait for user command or messages from other cluster node modules
504
Quit command received ?
Yes
No
505
Proceed to shut down process
506
Communicate received command to all cluster node modules
508
Communicate message to kernel module
510
Receive result from kernel module
512
Respond to messages from other cluster node modules

FIG. 5

Electronic Acknowledgement Receipt

EFS ID:	18207034
Application Number:	14181112
International Application Number:	
Confirmation Number:	6874
Title of Invention:	CLUSTER COMPUTING
First Named Inventor/Applicant Name:	Zvi Tannenbaum
Customer Number:	20995
Filer:	Lance Dean Smemoe/ThuyQuyen Nguyen
Filer Authorized By:	Lance Dean Smemoe
Attorney Docket Number:	ZTANN.002P1C3
Receipt Date:	14-FEB-2014
Filing Date:	
Time Stamp:	16:50:23
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Application Data Sheet	ZTANN_002P1C3_ADS.pdf	1279734 79ac76f4b1c2360f9a279b29a4ccc4a811f6f36f	no	7

Warnings:

Information:

2		ZTANN_002P1C3_Spec.pdf	204657 44b9702e6ecc47b6b4ec45b0d2189277a7cb56a7	yes	48
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Specification		1	46	
	Claims		47	47	
	Abstract		48	48	
Warnings:					
Information:					
3	Drawings-only black and white line drawings	ZTANN_002P1C3_Drawings.pdf	71433 d74eb080c418e1d3f94fc9502de60f946ae706eb	no	5
Warnings:					
Information:					
		Total Files Size (in bytes):	1555824		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

<u>New Applications Under 35 U.S.C. 111</u>
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

<u>National Stage of an International Application under 35 U.S.C. 371</u>
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

<u>New International Application Filed with the USPTO as a Receiving Office</u>
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.