			JPO; DERWENT; IBM_TDB			
S27	7	"709"/230-234.ccls. and (upconvert$3 or upshift$3 or downshift$3) with (stream$3 or (video with strem$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:21
S28	7	"709"/230-234.ccls. and (upconvert$3 or upshift$3 or downshift$3) with (stream$3 or (video with stream$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:22
S29	14	(real near2 network).as.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:32
S30	142	(sure-stream or sure adj stream or surestream)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:35
S31	26	(sure-stream or surestream)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:35
S32	2	"7054948".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:37
S33	1	"7054948".pn. and surestream	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 17:37
S34	9	"7406527".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:01
S35	855	(((multiple or plurality or many or three or two) near5 (version or	US-PGPUB; USPAT;	OR	ON	2009/02/09 18:02

		format$4 or copies or cop$4))) same (stream$3 or (video with stream$3)) same server	USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB			
S36	855	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same (stream$3 or (video with stream$3)) same server and709/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:03
S37	855	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same (stream$3 or (video with stream$3)) same server and709/231.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:03
S38	855	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same (stream$3 or (video with stream$3)) same server and709/231.ccls. and @ad<"20040430"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:03
S39	94	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same (stream$3 or (video with stream$3)) same server and (709/231.ccls. and @ad<"20040430")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:04
S40	0	S39 and (upconvert$3 or upshift$3 or downshift$3) with (stream$3 or (video with stream$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:08
S41	0	S39 and (upconvert$3 or upshift$3 or downshift$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 18:08
S42	121	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same3 (stream$3 or (video with stream$3)) same server and (709/231.ccls. and @ad<"20040430")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:25
S43	0	((dynamically near3 chang$4) with (((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4)))) same3 (stream$3 or (video with stream$3)) same server and (709/231.ccls. and	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:30

		@ad<"20040430")				
S44	0	((dynamically near3 chang$4) with (((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4)))) same3 (stream$3 or (video with stream$3)) same server	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:31
S45	2	((dynamically) with (((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4)))) same3 (stream$3 or (video with stream$3)) same server	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:33
S46	0	((dynamical) with (((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4)))) same3 (stream$3 or (video with stream$3)) same server	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:34
S47	5	((dynamic$5) with (((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4)))) same3 (stream$3 or (video with stream$3)) same server	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:34
S48	3	S47 not S45	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:36
S49	782	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same3 (stream$3 or (video with stream$3)) same server and (@ad<"20040430")	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:45
S50	146	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same3 (stream$3 or (video with stream$3)) same server and (@ad<"20040430") and ((dynamic$5 or automatical$5) near3 (upshift$3 or downshift$3 or upconvert$3 or downcovnert$3 or chan$4))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:56
S51	78	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same3 (stream$3 or (video with stream$3)) same server and (@ad<"20040430") and ((dynamic$5 or automatical$5) near3 (upshift$3 or downshift$3 or upconvert$3 or downcovnert$3 or chan$4)) and	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 19:58

		"709"/$$$.ccls.				
S52	29	(((multiple or plurality or many or three or two) near5 (version or format$4 or copies or cop$4))) same3 (stream$3 or (video with stream$3)) same server and (@ad<"20040430") and ((dynamic$5 or automatical$5) near3 (upshift$3 or downshift$3 or upconvert$3 or downcovnert$3 or chan$4)) and "709"/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 20:01
S53	7913	"709""/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 20:02
S54	1660	(multi-rate and (stream$3 or (video with stream$3)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 20:04
S55	11	(multi-rate and (stream$3 or (video with stream$3))).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/09 20:04
S56	6	("20020174434" | "20030067872" | "20040030797").PN.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:44
S57	32	("20020174434" | "20030067872" | "20040030797").PN. or ("20020073167" | "20050123058" | "20070024705" | "20080028428" | "20080056373" | "6091775" | "6091777" | "6510553" | "6731600" | "6732183" | "6968387" | "7075986" | "7349976").PN.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:45
S58	21	S57 and (stream$3 or (video with stream$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:45
S59	1	S57 and (stream$3 or (video with stream$3)) and ((plurality or many or multiple) near5 (version or format or cop$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:46

S60	4	S57 and (stream$3 or (video with stream$3)) and ((plurality or many or multiple) near5 (version or format or cop$3 or quality or (high and low)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:47
S61	670510	(plurality or many or three or two or multiple) near5 (file or cop$3 or version or format$3 or compress$5)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:51
S62	662	((plurality or many or three or two or multiple) near5 (file or cop$3 or version or format$3 or compress$5)) with ((automatic$4 or dynamic$5) near5 (switch$3 or chang$3 or upshift$3 or downshift$3 or upconvert$3 or downconvert$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 14:52
S64	72	S62 and "709"/$$$.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:02
S65	62	S62 and "709"/$$$.ccls. and @ad<"20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:05
S66	5	S62 and "709"/230-234.ccls. and @ad<"20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:06
S67	424	((plurality or many or three or two or multiple) near5 (file or cop$3 or version or format$3 or compress$5)) with ((automatic$4) near5 (switch$3 or chang$3 or upshift$3 or downshift$3 or upconvert$3 or downconvert$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:07
S68	425	((plurality or many or three or two or multiple) near5 (file or cop$3 or version or format$3 or compress$5 or (bit adj2 rate or bit-rate))) with ((automatic$4) near5 (switch$3 or chang$3 or upshift$3 or downshift$3 or upconvert$3 or downconvert$3))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:08
S69	5	S68 and "709"/230-234.ccls. and @ad<"20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO;	OR	ON	2009/02/11 15:09

			DERWENT; IBM_TDB			
S70	12589	(stream$3 or (video with stream$3)) and ((plurality or many or multiple) near8 automat$5)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:10
S71	6937	(stream$3 or (video with stream$3)) and ((plurality or many or multiple) near4 automat$5)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:10
S72	26	(stream$3 or (video with stream$3)) and ((plurality or many or multiple) near4 automat$5) and 709/231.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:11
S73	0	(stream$3 or (video with stream$3)) and ((plurality or many or multiple) near4 automat$5) and 709/231.ccls. and (upshift$3 or downshift$3 or upconvert$3 or downconvert$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:11
S74	2	(stream$3 or (video with stream$3)) and ((plurality or many or multiple) near4 automat$5) and 709/230-234.ccls. and (upshift$3 or downshift$3 or upconvert$3 or downconvert$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:12
S75	5236	709/231-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:12
S76	4485	709/231-234.ccls. and @ad<"20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:12
S77	691	S76 and ((multiple or plurality or many or three) near5 files)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:13
S78	591	S76 and ((multiple or plurality or many or three) near3 files)	US-PGPUB; USPAT; USOCR;	OR	ON	2009/02/11 15:14

			FPRS; EPO; JPO; DERWENT; IBM_TDB			
S79	233	S76 and ((multiple or plurality or many or three) near3 copies)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:14
S80	1	S76 and ((multiple or plurality or many or three) near3 copies) near5 (encod$3 or bit-rate or qualilty)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:15
S81	9	S76 and ((multiple or plurality or many or three) near3 (copies or files)) near5 (encod$3 or bit-rate or qualilty)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:25
S82	1	S76 and ((multiple or plurality or many or three) near3 (copies))with (encod$3 or bit-rate or qualilty)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:27
S83	1	S76 and ((multiple or plurality or many or three) near3 (copies)) with (encod$3 or bit-rate or qualilty)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:27
S84	233	S76 and ((multiple or plurality or many or three) near3 (copies))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:28
S85	1	S76 and ((multiple or plurality or many or three) near3 (copies)) near5 (bit adj3 rate)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:28
S86	6	@ad<"20050408" and ((multiple or plurality or many or three) near3 (copies)) near5 (bit adj3 rate)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:29
S87	251	(dynamic near5 stream$3).ti.	US-PGPUB;	OR	ON	2009/02/11

			USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB			15:32
S88	169	(dynamic near5 stream$3).ti. and @ad<"20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:34
S89	10	(dynamic near5 stream$3).ti. and @ad<"20050408" and 709/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:34
S90	2	"5832300".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:40
S91	0	(dautomatic$4 near5 stream$3).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:41
S92	326	(automatic$4 near5 stream$3).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:41
S93	241	@ad<"20050408" and S92	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:42
S94	0	(automatic$4 near5 stream$3 with (upshif$3 or downshif$3 or upconvert$3 or downconvert$3)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:43
S95	0	(automatic$4 near5 stream$3 with (upshift$3 or downshift$3 or upconvert$3 or downconvert$3)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT;	OR	ON	2009/02/11 15:43

			IBM_TDB			
S96	0	(automatic$4 with stream$3 with (upshift$3 or downshift$3 or upconvert$3 or downconvert$3)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:43
S97	3	(stream$3 with (upshift$3 or downshift$3 or upconvert$3 or downconvert$3)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:43
S98	0	(stream$3 with (video with stream$3)).ti. and (automatic$4 near4 switch$3 near4 different near4 quality)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:52
S99	0	(stream$3 with (video with stream$3)) and (automatic$4 near4 switch$3 near4 different near4 quality)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:52
S100	1081374	(dynamic$3 and switch$3 and stream$3 or video)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:57
S101	271819	(dynamic$3 and switch$3 and stream$3 or video).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:58
S102	0	(dynamic$3 and switch$3 and stream$3 and video).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:58
S103	0	(dynamic$3 and switch$3 and (stream$3 and video)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:59
S104	0	(dynamic$3 and (switch$3 or shivt$3 or upshift$3) and (stream$3 and video)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO;	OR	ON	2009/02/11 15:59

			JPO; DERWENT; IBM_TDB			
S105	5	(automatic$4 and (switch$3 or shivt$3 or upshift$3) and (stream$3 and video)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:59
S106	6	(automatic$4 and (switch$3 or shift$3 or upshift$3) and (stream$3 and video)).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 15:59
S107	108	(automatic$4 and (switch$3 or shift$3 or upshift$3) and (stream$3 and video))	DERWENT	OR	ON	2009/02/11 16:00
S108	6	(automatic$4 near5 (switch$3 or shift$3 or upshift$3) near5 (stream$3 and video))	DERWENT	OR	ON	2009/02/11 16:02
S109	1	(automatic$4 near5 (copies or files or versions) near5 (stream$3 and video))	DERWENT	OR	ON	2009/02/11 16:02
S110	11	(multiple or many or three or plurality) near5 ((files or version) near5 (quality or bit-rate or bit adj rate))	DERWENT	OR	ON	2009/02/11 16:03
S111	365	(multiple or many or three or plurality) near5 ((files or version) near5 (quality or bit-rate or bit adj rate))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:04
S112	37	S111 and 709/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:04
S113	108249	(multiple or many or three or plurality) near5 ((files or version))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:10
S114	1240	S113 and 709/230-234.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:11
S115	1038	S113 and 709/230-234.ccls. and @ad< "20050408"	US-PGPUB; USPAT;	OR	ON	2009/02/11 16:11

			USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB			
S116	11	S113 and 709/230-234.ccls. and (upshift$3 or downshift$3 or upcovert$3 or downconvert$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:12
S117	3	S113 and 709/230-234.ccls. and ((upshift$3 or downshift$3 or upcovert$3 or downconvert$3) with (quality))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:14
S118	3	709/230-234.ccls. and ((upshift$3 or downshift$3 or upcovert$3 or downconvert$3) with (quality with (video or stream$3)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:14
S119	7	((upshift$3 or downshift$3 or upcovert$3 or downconvert$3) with (quality with (video or stream$3)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:15
S120	50419	microsoft.as.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:15
S121	981	microsoft.as. and (stream$3).ti.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:16
S122	650	microsoft.as. and (stream$3).ti. and @ad< "20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:16
S123	6	microsoft.as. and (stream$3 with quality).ti. and @ad< "20050408"	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:17

S124	0	(automatica$3 near5 switch$3 near5 (different near5 qualit$5) near5 (files or version))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:19
S125	0	(automatic$5 near5 switch$3 near5 (different near5 qualit$5) near5 (files or version))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:20
S126	0	(automatic$5 near5 switch$3 near5 ((different near5 qualit$5) near5 (files or version)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:20
S127	0	((automatic$5 near5 (switch$3 or upshift$3 or downshift$3 or upconvert$3)) with ((different near5 qualit$5) near5 (files or version)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:20
S128	6	((automatic$5 near5 (switch$3 or upshift$3 or downshift$3 or upconvert$3)) with ((qualit$5) near5 (files or version)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/11 16:22
S129	2	"20050262257".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:34
S130	0	"20050262257".pn. and (ration or average)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:38
S131	1	"20050262257".pn. and (ratio or average)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:39
S132	0	"20050262257".pn. and (sub-parts or subparts)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO;	OR	ON	2009/02/13 17:39

			DERWENT; IBM_TDB			
S133	1	"20050262257".pn. and (safe$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:43
S134	2	"2002174434".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:49
S135	2	"20020174434".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:49
S136	2	"20020144276".pn.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 17:51
S137	8	("20020078209" | "20040117459" | "20050021616" | "20060031523").PN.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 18:33
S138	1	("20020078209" | "20040117459" | "20050021616" | "20060031523").PN. and clip$4	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 18:33
S139	1	"20020188745".pn. and play$3	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 18:37
S140	0	"20020188745".pn. and tcp	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:00
S141	0	"20020188745".pn. and http	US-PGPUB; USPAT; USOCR;	OR	ON	2009/02/13 19:00

			FPRS; EPO; JPO; DERWENT; IBM_TDB			
S142	1	"20020188745".pn. and web	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:00
S143	2	"20020188745".pn. and internet	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:01
S144	4765975	((plurality or mutiple or two or threee or many) near5 tcp mear5 connection)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:03
S145	4766104	((plurality or multiple or two or threee or many) near5 tcp mear5 connection)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:03
S146	2131	(((plurality or multiple or two or threee or many) near5 (tcp near5 connection)))	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:03
S147	30	(((plurality or multiple or two or threee or many) near5 (tcp near5 connection))) and 709/231.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:04
S148	12	(((plurality or multiple) near5 (tcp near5 connection))) and 709/231.ccls.	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:07
S149	0	"20020188745".pn. and cache	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:22
S150	0	"20020188745".pn. and operator	US-PGPUB;	OR	ON	2009/02/13

			USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB			19:22
S151	1	"20020188745".pn. and separate	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:25
S152	1	"20050262257".pn. and (cache)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:25
S153	1	"20050262257".pn. and (saf$3)	US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB	OR	ON	2009/02/13 19:36

EAST Search History (Interference)

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L161	0	(media player video playback multipl tcp connection subsequent).clm.	US-PGPUB; USPAT; UPAD	AND	ON	2014/07/31 01:19
L162	1	(media player video playback multiple tcp connection subsequent).clm.	US-PGPUB; USPAT; UPAD	AND	ON	2014/07/31 01:19
L163	1	(media player video playback multiple tcp connection subsequent).clm.	US-PGPUB; USPAT; UPAD	AND	ON	2014/07/31 01:19
L164	0	(media player video playback multiple tcp connection subsequent generat$5).clm.	US-PGPUB; USPAT; UPAD	AND	ON	2014/07/31 01:20

7/ 31/ 2014 1:20:27 AM
C:\ Users\ ndonabed\ Documents\ EAST\ Workspaces\ 11116783_amendment.wsp

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

NOTICE OF ALLOWANCE AND FEE(S) DUE

110240 7590 08/12/2014
Ingrassia Fisher & Lorenz, P.C. (DISH Digital)
7010 East Cochise Road
Scottsdale, AZ 85253

EXAMINER
DONABED, NINOS

ART UNIT	PAPER NUMBER
2444	

DATE MAILED: 08/12/2014

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/116,783	04/28/2005	R. Drew Major	066.0141 (P2004-04-01.1)	5535

TITLE OF INVENTION: APPARATUS, SYSTEM, AND METHOD FOR ADAPTIVE-RATE SHIFTING OF STREAMING CONTENT

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	11/12/2014

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), to: Mail **Mail Stop ISSUE FEE**
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

110240 7590 08/12/2014
Ingrassia Fisher & Lorenz, P.C. (DISH Digital)
7010 East Cochise Road
Scottsdale, AZ 85253

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
11/116,783	04/28/2005	R. Drew Major	066.0141 (P2004-04-01.1)	5535

TITLE OF INVENTION: APPARATUS, SYSTEM, AND METHOD FOR ADAPTIVE-RATE SHIFTING OF STREAMING CONTENT

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	SMALL	$480	$0	$0	$480	11/12/2014

EXAMINER	ART UNIT	CLASS-SUBCLASS
DONABED, NINOS	2444	709-231000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
- ☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
- ☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent): ☐ Individual ☐ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
- ☐ Issue Fee
- ☐ Publication Fee (No small entity discount permitted)
- ☐ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
- ☐ A check is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☐ The Director is hereby authorized to charge the required fee(s), any deficiency, or credits any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
- ☐ Applicant certifying micro entity status. See 37 CFR 1.29
  NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
- ☐ Applicant asserting small entity status. See 37 CFR 1.27
  NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
- ☐ Applicant changing to regular undiscounted fee status.
  NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

PTOL-85 Part B (10-13) Approved for use through 10/31/2013. OMB 0651-0033 U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

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.
11/116,783	04/28/2005	R. Drew Major	066.0141 (P2004-04-01.1)	5535

110240 7590 08/12/2014

Ingrassia Fisher & Lorenz, P.C. (DISH Digital)
7010 East Cochise Road
Scottsdale, AZ 85253

EXAMINER	
DONABED, NINOS	

ART UNIT	PAPER NUMBER
2444	

DATE MAILED: 08/12/2014

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. 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, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. 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.

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 disclosure of these records is required by the Freedom of Information Act.
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 inspection 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.

Notice of Allowability	Application No. 11/116,783	Applicant(s) MAJOR ET AL.	
	Examiner NINOS DONABED	Art Unit 2444	AIA (First Inventor to File) Status No

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *board descision dated 7/23/2014*.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on_____.
2. ☐ 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.
3. ☒ The allowed claim(s) is/are *20,21,23-26,35-38 and 40-50*. As a result of the allowed claim(s), 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 .
4. ☐ 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)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☒ Notice of References Cited (PTO-892)
2. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
5. ☐ Examiner's Amendment/Comment
6. ☒ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/NINOS DONABED/
Primary Examiner, Art Unit 2444

1. The present application is being examined under the pre-AIA first to invent provisions.

Information Disclosure Statement

2. The information disclosure statement filed 7/29/2014 fails to comply with the provisions of 37 CFR 1.97, 1.98 and MPEP § 609 because prosecution is closed. It has been placed in the application file, but the information referred to therein has not been considered as to the merits. Applicant is advised that the date of any re-submission of any item of information contained in this information disclosure statement or the submission of any missing element(s) will be the date of submission for purposes of determining compliance with the requirements based on the time of filing the statement, including all certification requirements for statements under 37 CFR 1.97(e). See MPEP § 609.05(a).

REASONS FOR ALLOWANCE

The following is an examiner's statement of reasons for allowance: Claims 20,21,23-26,35-38 and 40-50 are allowable over the prior art of record: the closest prior art of record (Birney "Intelligent Streaming") does not teach or suggest in detail "streaming by a media player operating on an end user station a video from a set of one or more servers, wherein each of a plurality of different copies of the video encoded at different bit rates is stored as multiple files on the set of servers, wherein each of the multiple files yields a different portion of the video on playback, wherein the multiple files across the different copies yield the same portions of the video on playback, each of said files having a time index such that the files whose playback is the same portion of

the video for each of the different copies have the same time index in relation to the beginning of the video, and wherein the streaming comprises: requesting by the media player a plurality of sequential ones of the files of one of the copies from the set of servers over a plurality of Transmission Control Protocol (TCP) connections based on the time indexes; automatically requesting by the media player from the set of servers over the plurality of TCP connections subsequent portions of the video by requesting for each such portion one of the files from one of the copies dependent upon successive determinations by the media player to shift the playback quality to a higher or lower quality one of the different copies, said automatically requesting including, repeatedly generating a set of one or more factors indicative of the current ability to sustain the streaming of the video using the files from different ones of the copies, wherein the set of one or more factors relate to the performance of the network; and making the successive determinations to shift the playback quality based on at least one of the set of factors to achieve continuous playback of the video using the files of the highest quality one of the copies determined sustainable at that time; and presenting the video by playing back with the media player on the end user station the requested files in order of ascending playback time" in combination with all the elements of the independent claim as argued by the Applicant.

Birney teaches multiple copies of the same stream being stored on a server at different bit rates with the same time index used to switch between the stream based on network bandwidth limitations. Whereas, stated above, Applicant's claimed invention states "streaming by a media player operating on an end user station a video from a set

of one or more servers, wherein each of a plurality of different copies of the video encoded at different bit rates is stored as multiple files on the set of servers, wherein each of the multiple files yields a different portion of the video on playback, wherein the multiple files across the different copies yield the same portions of the video on playback, each of said files having a time index such that the files whose playback is the same portion of the video for each of the different copies have the same time index in relation to the beginning of the video, and wherein the streaming comprises: requesting by the media player a plurality of sequential ones of the files of one of the copies from the set of servers over a plurality of Transmission Control Protocol (TCP) connections based on the time indexes; automatically requesting by the media player from the set of servers over the plurality of TCP connections subsequent portions of the video by requesting for each such portion one of the files from one of the copies dependent upon successive determinations by the media player to shift the playback quality to a higher or lower quality one of the different copies, said automatically requesting including, repeatedly generating a set of one or more factors indicative of the current ability to sustain the streaming of the video using the files from different ones of the copies, wherein the set of one or more factors relate to the performance of the network; and making the successive determinations to shift the playback quality based on at least one of the set of factors to achieve continuous playback of the video using the files of the highest quality one of the copies determined sustainable at that time; and presenting the video by playing back with the media player on the end user station the requested files in order of ascending playback time." So as indicated in the above statements.

Applicant's arguments have been considered persuasive, in light of the claims limitation as well as the enabling portions of the specification.

Examiner also notes the board decision and reversal on 7/23/2014.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Any inquiry concerning this communication or earlier communications from the examiner should be directed to NINOS DONABED whose telephone number is (571)272-8757. The examiner can normally be reached on Monday-Friday, 7:30 AM-5:00 PM EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Peter-Anthony Pappas can be reached on (571) 272-7646. 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.

/NINOS DONABED/
Primary Examiner, Art Unit 2444

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11116783	MAJOR ET AL.
	Examiner	Art Unit
	NINOS DONABED	2444

CPC- SEARCHED		
Symbol	Date	Examiner

CPC COMBINATION SETS - SEARCHED		
Symbol	Date	Examiner

US CLASSIFICATION SEARCHED			
Class	Subclass	Date	Examiner
709	246	05/07/2008	ND

SEARCH NOTES		
Search Notes	Date	Examiner
See EAST Search History	05/07/2008	ND
Conducted double patenting search	05/07/2008	ND
Consulted with SPE William Vaughn	11/13/2008	ND
See updated EAST search	12/13/2009	ND
Consulted with SPE William Vaughn regarding Final	02/12/2009	ND
Consulted with SPE William Vaughn regarding Advisory	03/29/2009	ND
Consulted with SPE William Vaughn regarding Advisory	5/7/2009	ND
Consulted with Primary Paul Kang	02/10/2009	ND
See updated east search	7/30/2009	ND
Consulted with SPe William Vaughn regarding search	7/30/2009	ND
Consulted SPE William Vaughn regarding Final	2/4/2010	nd
Consulted SPE William Vaughn regarding Appeal	1/27/2011	nd
Consulted with Qas Mark Radtke regarding IDS	7/30/2014	nd
Conducted 101 claim analysis	7/31/2014	nd
Conducted 112 2nd paragraph analysis	7/31/2014	nd
Conducted 112 6th paragraph claim analysis	7/31/2014	nd
Conducted double patenting search (EAST, eDAN, PALM) *Inventor **Assignee	7/31/2014	nd
East Search	7/31/2014	nd

SEARCH NOTES		
Search Notes	**Date**	**Examiner**
Plus Serch	7/31/2014	nd
NPL Seach (Google, Google Scholar)	7/31/2014	nd
709/246,231 + texted searched1	7/31/2014	nd

INTERFERENCE SEARCH			
US Class/ CPC Symbol	**US Subclass / CPC Group**	**Date**	**Examiner**
	Conducted Interference Search - See EAST Search History	7/31/2014	nd

Notice of References Cited	Application/Control No. 11/116,783	Applicant(s)/Patent Under Reexamination MAJOR ET AL.	
	Examiner NINOS DONABED	Art Unit 2444	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
*	A	US-2005/0033855	02-2005	Moradi et al.	709/231
	B	US-			
	C	US-			
	D	US-			
	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	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.

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11116783	MAJOR ET AL.
	Examiner	**Art Unit**
	NINOS DONABED	2444

CPC						
Symbol					**Type**	**Version**
H04N		21	/	845	I	2013-01-01
H04N		21	/	84	F	2013-01-01
H04N		21	/	64792	I	2013-01-01
H04N		21	/	64769	I	2013-01-01
H04N		21	/	643	I	2013-01-01
H04N		21	/	25808	I	2013-01-01
H04N		21	/	2662	I	2013-01-01
			/			
			/			
			/			
			/			
			/			
			/			
			/			
			/			

CPC Combination Sets							
Symbol					**Type**	**Set**	**Ranking**	**Version**
			/					
			/					

NONE		**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	21	
/NINOS DONABED/ Primary Examiner.Art Unit 2444	7/31/2014	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	20	7

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11116783	MAJOR ET AL.
	Examiner	**Art Unit**
	NINOS DONABED	2444

US ORIGINAL CLASSIFICATION		INTERNATIONAL CLASSIFICATION	
CLASS	SUBCLASS	CLAIMED	NON-CLAIMED
709	231	G 06 F 15/16 (2006.01.01)	

CROSS REFERENCE(S)					
CLASS	SUBCLASS (ONE SUBCLASS PER BLOCK)				
709	246				

NONE		**Total Claims Allowed:** 21
(Assistant Examiner)	(Date)	
/NINOS DONABED/ Primary Examiner.Art Unit 2444	7/31/2014	O.G. Print Claim(s): 20 / O.G. Print Figure: 7
(Primary Examiner)	(Date)	

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	11116783	MAJOR ET AL.
	Examiner	**Art Unit**
	NINOS DONABED	2444

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

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
-	1	-	17	-	33	20	49								
-	2	-	18	-	34	21	50								
-	3	-	19	7	35										
-	4	1	20	8	36										
-	5	2	21	9	37										
-	6	-	22	10	38										
-	7	3	23	-	39										
-	8	4	24	11	40										
-	9	5	25	12	41										
-	10	6	26	13	42										
-	11	-	27	14	43										
-	12	-	28	15	44										
-	13	-	29	16	45										
-	14	-	30	17	46										
-	15	-	31	18	47										
-	16	-	32	19	48										

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 21	
/NINOS DONABED/ Primary Examiner.Art Unit 2444 (Primary Examiner)	7/31/2014 (Date)	O.G. Print Claim(s) 20	O.G. Print Figure 7

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 April 2004 (29.04.2004)

PCT

(10) International Publication Number
WO 2004/036824 A1

(51) International Patent Classification[7]: H04L 9/06

(21) International Application Number:
PCT/FI2003/000752

(22) International Filing Date: 10 October 2003 (10.10.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
20021820 14 October 2002 (14.10.2002) FI

(71) Applicant *(for all designated States except US)*: **NOKIA CORPORATION** [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **CURCIO, Igor, D.D.** [IT/FI]; Hatanpään valtatie 12 C 57, FIN-33100 Tampere (FI). **LUNDAN, Miikka** [FI/FI]; Männikönkatu 4 A 7, FIN-33820 Tampere (FI).

(74) Agent: **FORSSÉN & SALOMAA OY**; Eerikinkatu 2, FIN-00100 Helsinki (FI).

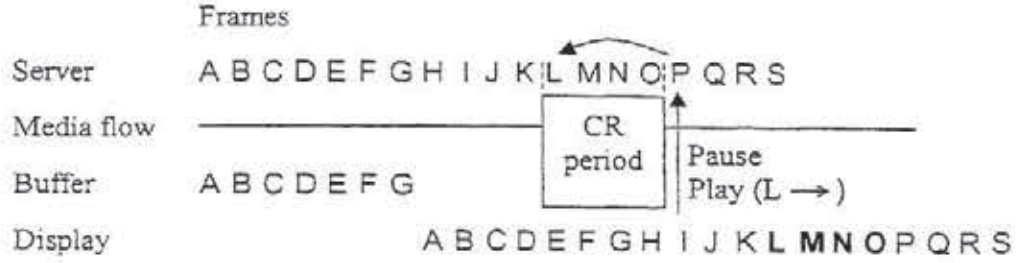
(81) Designated States *(national)*: AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, EG, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, 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, NI, NO, NZ, OM, PG, PH, PL, PT (utility model), PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

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

Published:
— *with international search 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.

(54) Title: STREAMING MEDIA



(57) Abstract: The invention relates to a method for streaming of media from a streaming server (111) to a mobile client device (101) over an air-interface, wherein, after a cell reselection, the method comprises requesting the streaming server (111) to resend streaming media which the mobile client device (101) was not able to receive due to the cell reselection.

WO 2004/036824 A1

STREAMING MEDIA

FIELD OF THE INVENTION

The invention relates to streaming of media from a streaming server to a mobile client device over an air-interface.

BACKGROUND OF THE INVENTION

A packet switched streaming service PSS is currently being standardized for mobile environment by 3GPP (3rd Generation Partnership Project). An example of a communications system capable of streaming media (streaming video and/or audio) is shown in Figure 1. The system comprises a streaming server 111 which is coupled to an IP-network (Internet Protocol) 104. The IP-network 104 may be, for example, the Internet or a service provider operator's intranet (an intranet network belonging to the operator's domain). The IP-network 104 is coupled to a core network 103 of a mobile communications network via a G_i interface. The mobile communications network also has a radio access network (RAN) 102 coupled to the core network 103. The radio access network 102 provides mobile communications devices 101 with access to the mobile communications network over an air-interface. Said access may be provided either by circuit switched means (circuit switched voice or data call) or packet switched means or both. In the following, GPRS (General Packet Radio Service) is used as an example of a packet switched means to communicate over the air-interface. Concerning GPRS, the generic term radio access network (RAN) is considered to comprise base (transceiver) stations (BS) and base station controllers (BSC).

In streaming media a sequence of 'moving images' (i.e. video) or sound (i.e. audio) or a sequence of 'moving images' together with sound (i.e. multimedia) is sent in a compressed form from the streaming server 111 to a mobile communications device 101 (hereinafter referred to as a client device 101). In contrast to the

technique in which an entire media file has to be arrived at the client before it can be played, the streaming technique enables sending of media (video and/or audio) from the streaming server 111 to the client device 101 in a continuous manner and playing of the media as it arrives at the client.

When media is streamed in cellular mobile communications systems such as the system presented in Figure 1, new problems specific to the mobile environment arise. These problems are mostly due to the different restrictions of mobile systems. One such problem arises in a cell reselection (i.e. handover) situation presented in Figure 2.

In Figure 2, a first base station BS1 of the radio access network 102 serves the area of a first cell 201 whereas a second base station BS2 of the radio access network 102 serves the area of a second cell 202. Before cell reselection, the client device 101 is served by the first base station BS1, i.e. the client device 101 has an active radio link with the first base station BS1. This situation is shown on the left side of the Figure 2. After the cell reselection, the client device 101 is served by the second base station BS2. The radio link with the first base station BS1, shown by a dotted line, is dropped. This situation is shown on the right side of the Figure 2.

The cell reselection (CR) can be divided in three periods in time:

a) Pre-CR period
b) CR period
c) Post-CR period.

During the pre-CR period, received signal quality in the first cell weakens and the client device 101 starts cell reselection signaling. The client device 101 is able to receive streaming media via the first base station BS1 during that period. During the CR period, the actual cell reselection is performed. The client device 101 can

not receive streaming media during that period. During the post-CR period, the client device 101 is able to receive streaming media via the second base station BS2.

The cell reselection may cause a long break in service. For example, a cell reselection (CR period) in a case where GPRS is used as radio bearer may take up to 30-40 seconds. This will affect an ongoing streaming session. For example, part of the packets carrying streaming media may be lost and/or freezing of the played media may occur at the client device 101. If the streaming media comprises a video stream, the freezing means that a still image will appear and stay for a period of time on a client device display before the streaming media is again received and can be played again. If the streaming media comprises an audio stream, the freezing means that no sound is played back (i.e. silence) until the CR period is over.

Accordingly, there is a need to find appropriate means for reducing the effects of cell reselection on an ongoing streaming session.

SUMMARY OF THE INVENTION

According to a first aspect of the invention, there is provided a method for streaming of media from a streaming server to a mobile client device over an air-interface, wherein the method comprises:
requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

A request requesting the streaming server to send the streaming media may be sent from the mobile client device to the streaming server before or after the mentioned cell reselection.

The mentioned request is not limited to requesting to send only the media which

the mobile client device is (or was) not able to receive due to the cell reselection (during a cell reselection period). In an embodiment, the request may be construed such that what is actually requested is: resending of a portion of a media content not received as well as sending any remaining portion of the media content.

The term media is considered to mean either video or audio or another media, such as still image, or any combination thereof, i.e multimedia.

Preferably, a resending request is sent in response to the cell reselection. The re-sending request is preferably an application layer request sent between a mobile client device application and a streaming server application.

In a preferred embodiment, the streaming server is requested to send at an increased speed for a determined period of time in order to increase at the client device a degree of fullness of a temporary store, such as a buffer, in which the streaming media is temporarily stored before playback. Preferably, this is done by by requesting the streaming server to switch from sending a higher bit rate media stream to sending a lower bit rate media stream at an increased speed. The higher and lower bit rate streams are preferably provided by a multi-rate codec.

According to a second aspect of the invention, there is provided a mobile client device for receiving streaming media from a streaming server over an air-interface, the mobile client device comprising:
means for requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

Preferably, the air-interface couples said mobile client device to a mobile communications network. The mobile client device preferably comprises a cellular mobile phone.

According to a third aspect of the invention, there is provided a streaming server

for sending streaming media to a mobile client device over an air-interface, the streaming server comprising:
means for receiving a request requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection; and
means for acting upon the received request.

According to a fourth aspect of the invention, there is provided a system comprising a streaming server and a mobile client device, for streaming of media from the streaming server to the mobile client device over an air-interface, the system comprising, at the mobile client device:
means for requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection, the system further comprising, at the streaming server:
means for receiving the request; and
means for acting upon the received request.

According to a fifth aspect of the invention, there is provided a computer program executable in a mobile client device, the computer program comprising:
program code for causing the mobile client device to request the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection.

According to a sixth aspect of the invention, there is provided a computer program executable in a streaming server, the computer program comprising:
program code for causing the streaming server to receive a request requesting the streaming server to send streaming media which the mobile client device is not able to receive due to a cell reselection; and
program code for acting upon the received request.

Dependent claims contain preferable embodiments of the invention. The subject

matter contained in dependent claims relating to a particular aspect of the invention is also applicable to other aspects of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention will now be described by way of example with reference to the accompanying drawings in which:

Figure 1 shows a communications system capable of streaming media;

Figure 2 shows a cell reselection situation in the communications system of Figure 1;

Figure 3 shows a client device buffer in a situation in which the client device buffer is shorter than the time that cell reselection takes;

Figure 4 shows a client device buffer in a situation in which the client device buffer is longer than the time that cell reselection takes;

Figure 5 shows another illustration of the situation which is behind Figure 4;

Figure 6 shows preferred action in a preferred embodiment of the invention;

Figure 7 illustrates a preferred embodiment of the invention;

Figure 8 illustrates a preferred embodiment of the invention;

Figure 9 shows a client device according to a preferred embodiment of the invention;

Figure 10 shows a streaming server according to a preferred embodiment of the

invention; and

Figure 11 illustrates three ways of sending messages according to preferred embodiments of the invention.

DETAILED DESCRIPTION

The system shown in Figure 1 can also be used in the preferred embodiments of the invention (the same applies to Figure 2). Accordingly, the system comprises a streaming server 111 which is coupled to an IP-network 104. The IP-network 104 may be, for example, the Internet or a service provider operator's intranet. The IP-network 104 is coupled to a core network 103 of a mobile communications network via a G_i interface. The mobile communications network also has a radio access network (RAN) 102 coupled to the core network 103. The radio access network 102 provides mobile communications devices 101 with access to the mobile communications network over an air-interface. Said access may be provided either by circuit switched means (circuit switched voice or data call) or packet switched means or both. In the following, GPRS (General Packet Radio Service) is used as an example of a packet switched means to communicate over the air-interface.

In order to better understand the preferred embodiments of the invention, let us consider the following example.

In this example, a mobile communications device 101 (again referred to as a client device 101) has established a streaming session with the streaming server 111 when being served by the first base station BS1. In this example and in the following description, mostly a video streaming session is considered. However, a corresponding consideration applies to an audio streaming session as well as to a multimedia (e.g. video with audio) streaming session.

RTSP (Real Time Streaming Protocol) protocol is used in the streaming session

setup. Once the session has been established, the streaming itself may be performed (i.e. media flow may be sent) according to RTP (Real time Transport Protocol) or another protocol. However, if it is desired to make a change in the established session, this will again be done by using RTSP.

At the beginning of the established streaming session, the received streaming media (media flow), transmitted in packets (or frames), is buffered (i.e. temporarily stored) in a buffer at the client device (hereinafter referred to as a client device buffer). When the client device buffer is full, playing (playback) of the streaming media is started at one end of the buffer while the buffer is continuously filled at the other end with received streaming media. Therefore, the buffer should remain nearly full during normal operation of the system.

The client device now undergoes a cell reselection from the first cell 201 to the second cell 202 (Fig. 2). It is to be noted that, although the radio access network (RAN) 102 is illustrated in Figure 2 by a single cloud only, it may comprise different radio access networks. It may comprise a GPRS (GPRS-only) RAN, a 3G (third generation) RAN or a combination thereof. The first cell 201 (served by the base station BS1) and the second cell 202 (served by the base station BS2) may therefore be, for example, cells of the GPRS RAN or 3G RAN. Accordingly, the cell reselection (handover) can occur, for example, between base stations belonging to GPRS RAN, between base stations belonging to 3G RAN or between a base station belonging to GPRS RAN and a base station belonging to 3G RAN.

The client device 101 can not receive streaming media during the CR period. Therefore, when the CR period begins the client device buffer starts emptying because the buffer content is further played while the continuous filling is stopped. Depending on the size of the client device buffer, two different cases can be identified:

i) The client device buffer (in time) is shorter than the time that cell reselection

takes, i.e. the time to empty a full client device buffer is shorter than the CR period,

ii) The client device buffer (in time) is longer than the time that cell reselection takes, i.e. the time to empty a full client device buffer is longer than the CR period.

In the first case (Figure 3), the buffer will become totally empty during the CR period. Since the buffer becomes empty, the client device 101 has no media to play anymore. Therefore, the client device 101 starts rebuffering. Rebuffering can be started during the post-CR period. In order to get the rebuffering started the client device 101 can use an RTSP PAUSE/PLAY method. In this method, the client device 101 first sends an RTSP PAUSE message to the streaming server 111. The PAUSE message will cause the server 111 to pause sending of the media flow. Subsequently, the client device 101 sends an RTSP PLAY message which requests the streaming server 111 to start resending the streaming media from the last received frame onwards. Rebuffering starts when the client again starts receiving the streaming media. When the client device buffer is full, playing of the streaming media is restarted from the last received frame onwards. In the meantime, the last received frame is shown as a still image on a client device display. However, in spite of the drawback of showing a still image, no packets are effectively lost since the media flow which is not received at the client device during the CR period is resent from the server 111 to the client device 101 after the CR period.

In the second case (Figure 4), the whole content of the buffer has not been played during the CR period. Therefore, the buffer only becomes partially empty during the CR period, i.e. there is still streaming media (DATA) in the buffer when the CR period ends. During the post-CR period (after the CR period) playing of the streaming media stored in the buffer before the CR period (i.e. DATA*) is further continued at one end of the buffer while the other end is filled with the streaming

media received after the CR period (i.e. DATA**). If nothing is done, DATA* will eventually end and playing is continued with the first frame of data DATA**. No still image is shown. However, the first frame of DATA** is not the frame that should be played right after the last frame of DATA*. Since the client device 101 could not receive streaming media during the CR period it has not even received the frame that should be showed after the last frame of DATA*. There exists a temporal gap in the buffered media. As a result, the client device 101 will skip showing the sequence of streaming media (video) which could not be received due to the cell reselection Concerning an audio stream, the same effect occurs, i.e. the client device 101 will skip playing sound which could not be received.

The second case is further illustrated with the aid of Figure 5. It is to be noted that Figure 5 only shows a very simplified case. The time is considered to flow from left to right in Figure 5. The streaming media sequence that the server sends comprises video frames A-S. The frames are received at the client device and buffered before playing. The playing of the media starting from frame A starts when the buffer is full (i.e. when the frames A-G have been stored in the buffer), the frames (moving video picture) are subsequently shown on a client device display. The buffering of the frames following frame G is not shown in Figure 5 for clarity reasons.

The media flow from the server to the client device is not possible during the CR period. Therefore, the client device does not receive the frames L-O. Otherwise, the media flow is perfectly received at the client device. Now that the the CR period is shorter than the buffer size, the showing of the frame (frame K) last received before the CR period will be followed by the showing of the frame (frame P) first received after the CR period. Video frames L-O are not shown at all. Concerning an audio stream, the same effect occurs, i.e. the user hears a pause and a temporal gap in the sound.

There exists two type of frames: intra frames and inter frames. The intra frames

contain all necessary information of an image whereas inter-frames only contain changes or predicted changes compared to a previous image. The previous image may be an intra frame or an inter frame. Therefore, if the frame P is an intra frame the shown picture only warps ahead in time from frame K to frame P. In this case the frames L-O are not simply shown to the user. But if the frame P is an inter frame severe distortion in the shown (moving) picture is likely to occur before the next intra frame is received and played.

A preferred embodiment of the invention as shown in Figure 6 concentrates on the above mentioned problem relating to the second case. In this embodiment, the buffer is longer than the duration of the CR period. Therefore, the buffer does not become totally empty during the CR period but only partially empty. It is to be noted that Figure 6 only shows a very simplified case. The time is considered to flow from left to right in Figure 6. The streaming media sequence that the server sends comprises frames A-S. The frames are received at the client device and buffered before playing. The playing of the media starting from frame A starts when the buffer is full, the frames (moving video picture) are subsequently shown on a client device display. The buffering of the frames following frame G is not shown in Figure 6 for clarity reasons. The media flow from the server to the client device is not possible during the CR period. Therefore, the client device does not receive the frames L-O. Otherwise, the media flow is perfectly received at the client device.

When the CR period is over the client device knows exactly what is the last frame that it has received before the CR period (i.e. during the pre-CR period). In this embodiment it is frame K. The client device then requests, right after the CR period has been ended, the streaming server to start resending the streaming media from the last received frame onwards. Upon receiving the request the streaming server starts resending the media flow. The first frame to be sent is in this embodiment frame L. Any frames sent and received after the CR period but before the resending has started (possibly some frames starting from frame P) are

ignored by the client device. When the streaming media (of frames A-K) stored in the buffer before the CR period ends, then the playing is continued with the first frame (frame L) of the resent streaming media received after the CR period. Therefore, no discontinuity in the shown video picture should appear to the user. In Figure 6, the resent video frames L-O which would have otherwise failed to have been received at the client device are shown in bold font. Correspondingly, concerning an audio stream, the streaming server is requested to start resending the audio stream at a point at which reception has been stopped in the client device. No discontinuity in playback should appear.

In practice, the resending request can be performed by the RTSP PAUSE/PLAY method. In this method, the client device first sends, after the CR period has been ended, an RTSP PAUSE message to the streaming server. The PAUSE message will cause the server to pause sending of the media flow. However, the playing of the streaming media received at the client device is not paused unless the buffer gets empty (which should not be the case). Subsequently, the client device sends an RTSP PLAY message to the streaming server 111. The PLAY message contains information on the starting point of the resending. At the end of the CR period the client device knows the time of the last received frame. Based on that, the client device determines the starting point before sending the PLAY message. The PLAY message causes the server to start resending.

An example of the PAUSE message is as follows:

```
PAUSE rtsp://example.com/foo RTSP/1.0
CSeq: 6
Session: 354832
```

The PAUSE message informs the server that a change is coming. An example of the subsequently sent PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-

The message field 'Range' tells the starting point of resending. In this example, the starting point is 28 seconds from the start of the streaming media sequence. In relation to the embodiment shown in Figure 6, this point would be exactly the time of frame L in the streaming media sequence.

In the embodiment just described, the frames (packets) that were lost during the CR period will be resent from the server to the client devide. Also, since the buffer size is longer than the duration of the CR period, no interruption in the shown video picture should appear and the user experience is maximized.

However, as the case is such that during normal streaming the degree of fullness of the buffer does not change, since the buffer is emptied (played) at the same rate as it is filled, this may cause an additional problem. In the embodiment just described, after the cell reselection, the client device buffer stays emptier than before, for the reasons as described in the foregoing. Therefore if, for example, a new cell reselection is to be performed in a near future, the emptier buffer may cause the same drawback as discussed in connection with the first case above (i.e. the case in which the client device buffer was shorter than the time that cell reselection takes).

A preferred embodiment of the invention concentrates on the above identified problem. In this embodiment, to ensure smooth behaviour of the shown video picture (correspondingly: played sound), the client device buffer is filled after the CR period, for a period of time, at a higher rate than it is emptied (played). This period may be called a filling period. When the filling period is over the buffer is full again and normal streaming in which the buffer is filled at the same rate as

emptied is resumed.

It is to be noted that normally raising the degree of fullness of the buffer would require the playing to be paused. In the present embodiment, playing is not paused but the degree of fullness of the buffer can still be raised due to the clever buffer management in which the buffer is filled, during playing, at a rate higher than the playing rate.

To raise the degree of fullness of the buffer without pausing the playback, the client device, in this embodiment, requests the server to switch to sending a lower bit rate streaming media sequence, but to use in the actual sending the same transmission bit rate (hereinafter referred to as an original transmission bit rate) as before. In order to reach the original transmission bit rate the client device requests the server to speed up the transmission of the lower bit rate sequence by a speeding factor. Speeding up the transmission bit rate causes more data to be written in the buffer than read out from the buffer. Therefore, the degree of fullness of the buffer increases as desired.

In other words, the server is asked to switch from sending the original sequence encoded at a first bit rate to sending a new corresponding sequence encoded at a second bit rate lower than the first bit rate and to increase the transmission bit rate of the new sequence in order to reach the original transmission bit rate (bandwidth). It is to be noted that the bit rate at which a media stream is encoded (and will be decoded) is a different concept than the transmission bit rate. The bit rate at which the media stream has been encoded has an effect on the picture quality. If the media stream has been encoded at a higher bit rate, this means that more bits have been used in encoding compared to encoding at a lower bit rate. This typically results in a better picture quality. The transmission bit rate, on the other hand, is the bit rate at which the media stream is actually sent and it depends on the available bandwidth.

The request to switch to sending a lower bit rate sequence and to speed up the transmission bit rate can be delivered by using the RTSP PAUSE/PLAY method. When the client device desires to fill the client device buffer during the post-CR period, it first sends a PAUSE message, corresponding to the one presented already in the preceding description, to the server to indicate that a change is coming. Subsequently it generates a PLAY message and sends it to the server.

An example of such a PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-40.00
Bandwidth: 20000
Speed: 1.5

This message contains two optional message fields, namely 'Bandwidth' and 'Speed' to be understood at the client device and streaming server. These fields are already specified as optional fields by the IETF (Internet Engineering Task Force) in the standard RFC 2326 (Real Time Streaming Protocol).

The message field 'Bandwidth' tells the server to change into sending a lower bit rate sequence (here: a sequence whose bit rate is 20 kbps) and the message field 'Speed' tells the server to speed up the sending by a speeding factor (here: 1.5). The message field 'Range' tells the starting point and stopping point in the streaming media sequence in time units (calculated from the beginning of the streaming media sequence(s)).

The PLAY message shown would suit for an example in which the server initially sends a 30 kbps bit rate sequence at an original transmission bit rate of 30 kbps and in which the server, in order to fill the buffer, is desired to switch into sending

a sequence whose bit rate is 20 kbps and to speed up the sending at a speeding factor 1.5 to reach the original transmission bit rate of 30 kbps during a 12 seconds filling period (i.e. between the time instants of 28 and 40 seconds of the media sequence).

If the client knows the possible bit rate options (they are typically communicated from the server to the client during session setup, for example, with the aid of a 200 OK response to an RTSP DESCRIBE message) it can calculate the speeding factor needed to reach the original bitrate using the following formula:

$$speeding\ factor = \frac{original\ transmission\ bit\ rate}{new\ transmission\ bit\ rate}$$

When the buffer is full, the client device sends another RTSP PAUSE and PLAY message pair to the server. An example of the PLAY message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 67
Session: 354832
Range: npt=40.00-
Bandwidth: 30000
Speed: 1.0

This exemplary message requests the server to start sending the (original) 30 kbps bit rate media sequence (Bandwidth: 30000) at the original transmission bit rate 30 kbps (Speed: 1.0) starting from the time instant 40 seconds of the media sequence (Range: npt=40.00-).

The client device can calculate the length of the buffer filling period using the following formulas:

$$filling\ period = \frac{LowSeqTime}{speeding\ factor}$$

wherein

$$LowSeqTime = (BufferSize - BufferData) \cdot \frac{original\ transmission\ bit\ rate}{original\ transmission\ bit\ rate - new\ transmission\ bit\ rate}$$

In these formulas, *LowSeqTime* indicates the duration of the lower bit rate sequence playing time in the client device, *BufferSize* indicates the size of the buffer in seconds and *BufferData* indicates the data left in the buffer in seconds.

The filling of the buffer is illustrated in Figure 7. It is to be noted that Figure 7 only shows a very simplified case. All delays, for example, have been ignored. The time is considered to flow from left to right in Figure 7. It can be seen that the buffer starts emptying when the CR period starts. Upon receiving the first PLAY message the server switches to sending a media sequence having a lower bit rate (here: 20 kbps) and speeds up the sending by a speeding factor (here again: 1.5). Since the speeding factor is 1.5, the buffer filling takes twice the time that the emptying period lasted. For example, if the emptying period lasted for 25 seconds, the filling period should be 50 seconds with this speeding factor. When the buffer is full the second PLAY message is sent and the server switches to sending the original media sequence having a higher bit rate (here: 30 kbps) and continues sending at the original transmission bit rate.

In the foregoing, it has been described that, for example, a video stream typically comprises both intra and inter frames, wherein the intra frames are 'independent frames' containing all necessary information of an image whereas inter-frames only contain changes or predicted changes compared to a previous image. In the following preferred embodiment of the invention, the timing of the switch from the lower bit rate sequence to the original bit rate sequence is more closely de-

scribed with this respect.

In this embodiment the purpose is to time, whenever possible, the switch from the lower bit rate sequence to the original bit rate sequence to take place at the point of an intra frame of the original bit rate sequence (or, more generally, at the point of an intra frame of the sequence to which the sending is switched). In this way, prediction errors and/or jumps in media (frame) flow can be avoided. Figure 8 illustrates this embodiment. It is to be noted that Figure 8 only shows a simplified case. The time is considered to flow from left to right in Figure 8. In this embodiment, the first switch from the original bit rate sequence to the lower bit rate sequence is effected at point 28 seconds with the aid of a first PAUSE/PLAY message pair.

As illustrated in Figure 8, the position of intra frames in the original bit rate sequence may differ from the ones of the lower bit rate sequence. However, based on:

- the time required to totally fill the buffer, i.e. '*filling period*',
- the time distance of two adjacent intra frames in the original sequence, i.e. an intra frame rate, '$IFrameTime_{Original}$', and
- the time at which the first switch from the original bit rate sequence to the lower bit rate sequence is effected, i.e. '*SwitchTime*',

it can be calculated a suitable point in time for swiching back to the original bit rate sequence. The parameter '$IFrameTime_{Original}$' can be calculated in the client device from the original bit rate sequence before switching to the lower bit rate sequence. The calculation of the suitable point in time for swiching back to the original bit rate sequence can be performed using the following formula:

$$SeqChangeTime = IFrameTime_{Original} * \left\lfloor \frac{SwitchTime + filling\ period}{IFrameTime_{Original}} \right\rfloor$$

In this formula, '*SeqChangeTime*' indicates the point in time for swiching back to the original bit rate sequence and the brackets indicate a *floor*-function (or *trunc*-function) which truncates the fractional part of the value computed in the brackets. For example, if the time required to totally fill the buffer, i.e. '*filling period*' is 16 seconds, the time at which the first switch is effected, i.e. '*SwitchTime*' is 28 seconds and the time difference between two adjacent intra frames in the original sequence, i.e. '$IFrameTime_{Original}$' is 5 seconds, then the point in time for swiching back to the original bit rate sequence is 40 seconds (*SeqChangeTime*=5*floor((28+16)/5) = 40 s). Accordingly, the second PAUSE/PLAY message pair is sent at point 40 seconds, as indicated in Figure 8. In this embodiment, the starting point and stopping point to be placed in the message field 'Range' of the PLAY message are 28 seconds and 40 seconds, respectively.

It is to be noted that, due to the intra frame rate, it might not in all cases be possible to totally fill the buffer. For example, in the example just described, it would have taken 4 seconds more to totally fill the buffer. However, the formula gives the closest suitable intra frame to the buffer filling period.

In another embodiment, it may well be that, at the point of the last received frame before the CR period, there is not an intra frame in the lower bit rate media sequence. In this embodiment, the starting point of the resending request is adjusted by a necessary amount of frames (or time) backwards such that there will be an intra frame in the lower bit rate sequence at the starting point. In this case, a set of last received frames before the CR period, which belong to a time period subsequent to the starting point, are ignored at the client device in order to guarantee a continuous playback.

In another embodiment of the invention, the switches between two different bit rate sequences are performed without the timing method(s) just described. In this embodiment, the first switching point (from higher to lower bit rate sequence) is directly determined by the last received frame and the second switching point (from lower to higher bit rate sequence) is directly determined by the the time required to totally fill the buffer (*'filling period'*). The switching points can therefore end up being at the point of an intra or inter frame. Should a switching point end up being at the point of an inter frame (for example a P-frame), a small prediction error may occur in the played media, however, in this embodiment, the buffer can be totally filled.

Another way to perform switching is to use so called "switch frames". These are frames which contain "difference" information between corresponding frames in different bit rate sequences. In this embodiment, a bridge and switch between the two sequences is performed with the aid of these frames using the difference information.

Figure 9 shows a client device 101 according to a preferred embodiment of the invention. The client device may be a mobile station of a cellular radio telephone network. The client device 101 comprises a processing unit MCU, a radio frequency part RF, and a user interface UI. The radio frequency part RF and the user interface UI are coupled to the processing unit MCU. The user interface UI typically comprises a display, one or more speaker and a keyboard (not shown) with the aid of which a user can use the device 101.

The processing unit MCU comprises a processor (not shown), a memory 210 and computer software. The software has been stored in the memory 210. The aforementioned client device buffer 240 is also comprised in the memory 210. The processor controls, in accordance with the software, the operation of the client device 101, such as receiving of the streaming media sent from the server 111 and sending of requests to the server 111 via the radio frequency part RF, reading and

writing of received streaming media (video and/or audio) in the buffer 240, and presenting of the received streaming video on the display and audio on the one or more speaker of the user interface UI. A suitable size of the buffer (in time) can be, for example, 1.5 or 2 times the maximum (or average) CR period time.

The software comprises a streaming client software application 220 (hereinafter referred to as client software 220), a protocol stack 230 for implementing the necessary protocol layers such as an RTP layer, an RTSP layer, an SDP (Session Description Protocol) layer, a TCP layer (Transmission Control Protocol), an IP layer and, below the IP-layer, lower protocol layers. In addition, the software comprises as a part of the client software 220 a media player for playing the received media.

The processor generates, based on the client software 220 the aforementioned PAUSE and PLAY messages and sends them to the server 111 via the radio frequency part RF. The processor also performs, based on the client software 220, the necessary calculations relating to the speeding factor, relating to the buffer filling period and relating to the suitable point in time for swiching from a lower bit rate sequence back to the original bit rate sequence.

The generation and sending of the resending request (PLAY message) and other appropriate action is triggered by a cell reselection event occurred in the mobile client device 101. The event can be detected by the client software 220 by reception of an asynchronous message from an API (Application Programming Interface) provided by the lower layers of the protocol stack 230. Alternatively or in addition, the event can be detected by monitoring the buffer level, i.e. the degree of fullness of the client device buffer 240. In this case, if the buffer 240 does not receive data for a certain amount X of time (depending on the implementation, the parameter X can be defined as a constant value, and it is a threshold for the client software 220 to understand that the cell reselection event has occurred), and if the client device 101 afterwards starts to receive data after a certain variable amount

Y of time (wherein Y > X, and Y is the real duration of the CR period), then the client can trigger the action described in this description.

Figure 10 shows a streaming server 111 according to a preferred embodiment of the invention. The streaming server 111 comprises a processing unit CPU, a first memory 310, an IP network interface 350, and a second memory 360. The first memory 310, the IP network interface 350, and the second memory 360 are coupled to the processing unit CPU.

The processing unit CPU controls, in accordance with computer software stored in the first memory 310, the operation of the streaming server 111, such as processing of requests received from the client device 101 and the sending of pre-recorded media streams, stored in the second memory (disk) 360, to the client device 101 via the IP network interface 350.

The software comprises a streaming server software application 320 (hereinafter referred to as server software 320), a protocol stack 330 for implementing the necessary protocol layers such as an RTP layer, an RTSP layer, an SDP layer, a TCP layer, an IP layer and lower protocol layers.

The PAUSE and PLAY messages sent from the client device 101 are received via the IP network interface 350. A processor (not shown) of the processing unit CPU processes the messages according to the server software 320 and the protocol stack 330 and takes appropriate action.

The present invention provides means for reducing the effects of cell reselection on an ongoing streaming session. It is to be noted that, according to preferable embodiments of the invention, a request for resending (e.g. a PLAY message) is sent on the application layer, i.e. between the client software application 220 and the server software application 320. Preferably RTSP over TCP (Transmission Control Protocol) or RTSP over another reliable protocol is used in transferring

application layer requests from the client device to the streaming server. The reception of the messages at the streaming server can therefore basically be guaranteed.

Although it has been described that the cell reselection would be performed between two base stations, it should be noted that a cell reselection can be performed also between two sectors of one and the same base station. Also, it should be noted that depending on the implemention, it may well be that separate messages (PAUSE messages) pausing the sending of streaming media are not needed. In an alternative embodiment, both to stop sending streaming media and to start resending streaming media is caused by a single appropriate message.

Further, in relation to the buffer filling embodiment, it has been described in the foregoing that when the server switches to sending a lower bit rate sequence, the original transmission bit rate would be maintained. However, in an alternative embodiment of the invention, a transmission bit rate higher than the original transmission bit rate is used during the filling period in order to fill the buffer more quickly. In this embodiment, it is assumed that a larger bandwidth can be requested by the mobile client device and that a larger bandwidth is actually available in the radio access network.

Further, in relation to the buffer filling embodiment, the bandwidth information of the sequence can be sent to the streaming server 111 also by other means than the RTSP field 'Bandwidth' (for example, by requesting a specific sequence encoded at a bitrate known to the client device 101). In this case, the field 'Bandwidth' is not used, the fields 'Speed' and 'Range' are recalculated according to the actual known sequence bit rate.

It is possible that the network (air-interface) bandwidth changes during the buffer filling period. If the client device 101 supports bandwidth adaptation that involves bit stream switching, the client device 101 should pause the buffer filling (with an

RTSP PAUSE message) before sending a bandwidth adaptation message to the streaming server 111. After the bandwidth adaptation operations are over, the client device can start the buffer filling again and recalculate the 'Bandwidth', 'Speed' and 'Range' values according to the bit rate and timing information of the new media stream.

Also, in relation to the buffer filling embodiment, it should be noted that, in an alternative embodiment, the second PAUSE/PLAY messages are not sent at all but the switch back to sending the original media sequence at the original transmission bit rate will be performed automatically by the server based on the stopping point information contained in the first PLAY message.

The resending request mentioned in the various embodiments of the invention may, in certain cases, actually be a sending request. One such case is considered as an alternative embodiment of the invention. In this embodiment, the client device (101) knowing beforehand that a cell reselection is going to occur in a very near future sends a PAUSE message to the streaming server (111) before the beginning of the cell reselection period (i.e. during the pre-CR period). Sending of the PAUSE message is triggered by a cell reselection initiation event notified to the client software 220 by means of the lower layer API. The PAUSE message causes the streaming server 111 to stop sending the streaming media. When the CR-period is over, the client device 101 then sends a PLAY message causing the streaming server to start sending at the point at which sending was stopped before the CR period. Playing of the streaming media is not stopped at the client device 101 in between, and if the client device buffer 240 has been selected to be longer in time than the time that cell reselection takes, the buffer 240 does not become totally empty during cell reselection and the user does not experience any jumps or interruption in playback. The PLAY message may contain a request to send at an increased speed in order to increase the degree of fullness of the buffer 240.

In yet another embodiment, the mobile client device 101 sends, before the

CR period, to the streaming server 111 an appropriate message requesting the streaming server 111 to stop sending the streaming media and to start sending again at a suitable point in time after the CR period. Sending of the message is triggered, during the pre-CR period, by a cell reselection initiation event notified to the client software 220 by means of the lower layer API. The mobile client device 101, knowing approximately the time that the cell reselection is going to take, estimates a suitable point in time at which it is again able to receive data. It inserts this information to the message so as to prevent the streaming server 111 from starting to send again too early.

An example of the above mentioned message is as follows:

PLAY rtsp://example.com/foo RTSP/1.0
CSeq: 7
Session: 354832
Range: npt=28.00-40.00; **time=19970123T153600Z**
Bandwidth: 20000
Speed: 1.5

This is an RTSP PLAY message having a header field 'Time'. The value of that field schedules the start of the "future" sending of the streaming media.

Figure 11 illustrates (especially relating to the buffer filling embodiments) the following three ways in which messages may be sent:

Case 1: The first PAUSE and PLAY messages are sent after the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

Case 2: The first PAUSE message is sent before the CR period. The associated PLAY message is sent after the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

Case 3: The first PAUSE message is sent before the CR period. The associated PLAY message is sent after the PAUSE message but before the CR period. The second PAUSE and PLAY messages are sent when the buffer 240 is full.

The sending of the second PAUSE message is not necessary, if the first PLAY message contains a closed 'Range' field.

Particular implementations and embodiments of the invention have been described. It is clear to a person skilled in the art that the invention is not restricted to details of the embodiments presented above (e.g. message names and message field names), but that it can be implemented in other embodiments using equivalent means without deviating from the characteristics of the invention. The scope of the invention is only restricted by the attached patent claims.

CLAIMS:

1. A method for streaming of media from a streaming server (111) to a mobile client device (101) over an air-interface, wherein the method comprises:

requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

2. A method according to claim 1, wherein the streaming server is provided with a starting point at which to start sending.

3. A method according to claim 1, wherein streaming server (111) sends the streaming media which the mobile client device (101) is not able to receive due to said cell reselection as well as a remaining portion of streaming media in response to the request.

4. A method according to claim 1, wherein the cell reselection comprises a cell reselection period during which the mobile client device (101) is not able to receive streaming media, the method comprising:

sending from the mobile client device (101) to the streaming server (111), after the cell reselection period, a resending request which requests the streaming server (111) to resend streaming media which the mobile client device (101) was not able to receive during the cell reselection period.

5. A method according to claim 4, wherein the resending request is generated according to RTSP protocol (Real Time Streaming Protocol).

6. A method according to claim 4, wherein the resending request is implemented by an RTSP PAUSE/PLAY message pair.

7. A method according to claim 1, wherein the streaming media is temporarily stored in a temporary store (240), such as a buffer, at the client device (101)

before playing.

8. A method according to claim 7, wherein the temporary store (240) has a size longer in time than a cell reselection period.

9. A method according to claim 7, wherein the streaming server is requested to send streaming media at a rate higher than the playing rate of that media so as to increase a degree of fullness of the temporary store (240).

10. A method according to claim 9, wherein a bandwidth or desired transmission bit rate with speeding factor is communicated to the streaming server (111) in a request.

11. A method according to claim 9, wherein the streaming media is stored at the mobile client device (101) at a rate higher than the playing rate.

12. A method according to claim 9, wherein the streaming server (111) is subsequently requested to resume an original configuration.

13. A method according to claim 7, wherein a degree of fullness of the temporary store (240) decreases during the cell reselection, and the streaming server is requested to send the not received streaming media although the temporary store (240) has not become totally empty, and said requesting is performed without pausing playback at the mobile client device (101).

14. A method according to claim 1, wherein the streaming server has a set of media streams available for transmission in which the same media content has been encoded at different bit rates.

15. A method according to claim 14, wherein information on the available set of media streams is beforehand communicated to the mobile client device (101)

in a streaming session setup.

16. A method according to claim 15, wherein the streaming server (111) is requested to switch from sending a higher bit rate media stream to sending a lower bit rate media stream at an increased speed.

17. A method according to any preceding claim, wherein the streaming media comprise one of the following: a video stream, an audio stream, another stream of single media, a multimedia stream.

18. A method according to any preceding claim, wherein the streaming server (111) sends streaming media to the mobile client device (101) via a mobile communications network.

19. A method according to any preceding claim, wherein the mobile communications network comprises a mobile packet radio network, such as a GPRS (General Packet Radio Service) network.

20. A method according to any preceding claim, wherein said cell reselection is performed between two base stations (BS1, BS2) which are selected from a group comprising: base stations belonging to a GPRS system, base stations belonging to a third generation mobile communications system.

21. A mobile client device (101) for receiving streaming media from a streaming server (111) over an air-interface, the mobile client device (101) comprising:

means (220, 230, MCU) for requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

22. A streaming server (111) for sending streaming media to a mobile client device (101) over an air-interface, the streaming server (111) comprising:

means (350) for receiving a request requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection; and

means (320, 330, CPU) for acting upon the received request.

23. A system comprising a streaming server (111) and a mobile client device (101), for streaming of media from the streaming server (111) to the mobile client device (101) over an air-interface, the system comprising, at the mobile client device (101):

means (220, 230, MCU) for requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection, the system further comprising, at the streaming server (111):

means (350) for receiving the request; and

means (320, 330, CPU) for acting upon the received request.

24. A computer program executable in a mobile client device (101), the computer program comprising:

program code for causing the mobile client device (101) to request the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection.

25. A computer program executable in a streaming server (111), the computer program comprising:

program code for causing the streaming server (111) to receive a request requesting the streaming server (111) to send streaming media which the mobile client device (101) is not able to receive due to a cell reselection; and

program code for acting upon the received request.

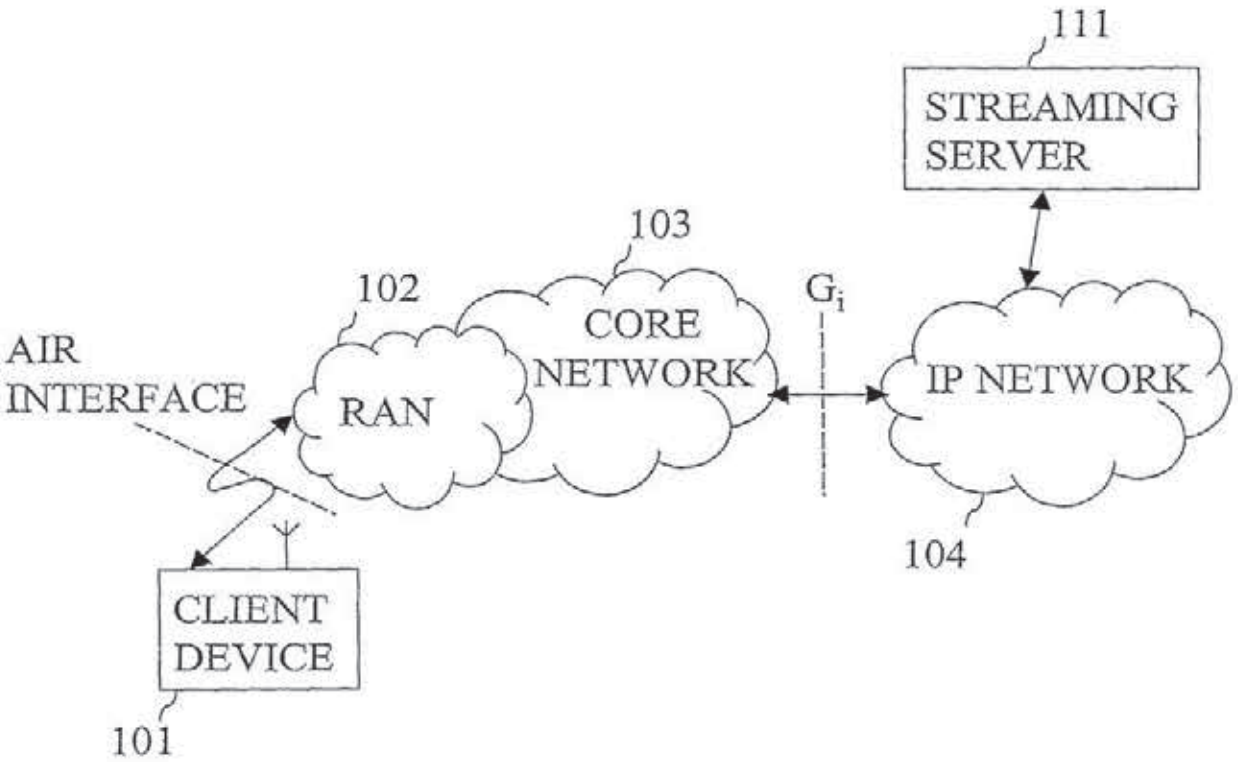

111
STREAMING SERVER
103
102
CORE NETWORK
G_i
AIR INTERFACE
RAN
IP NETWORK
104
CLIENT DEVICE
101

FIG. 1

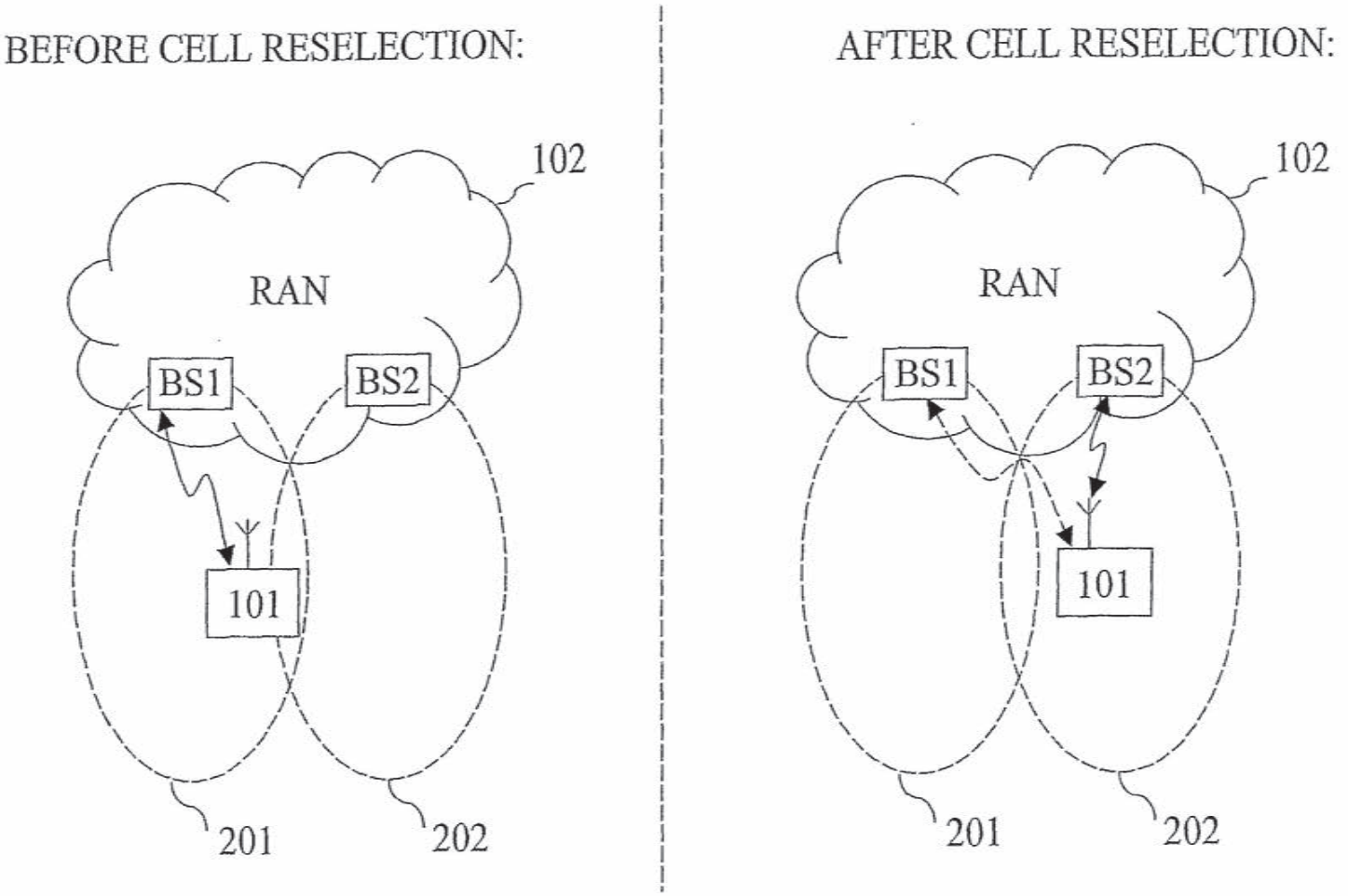


FIG. 2

Ex. 1018-1564

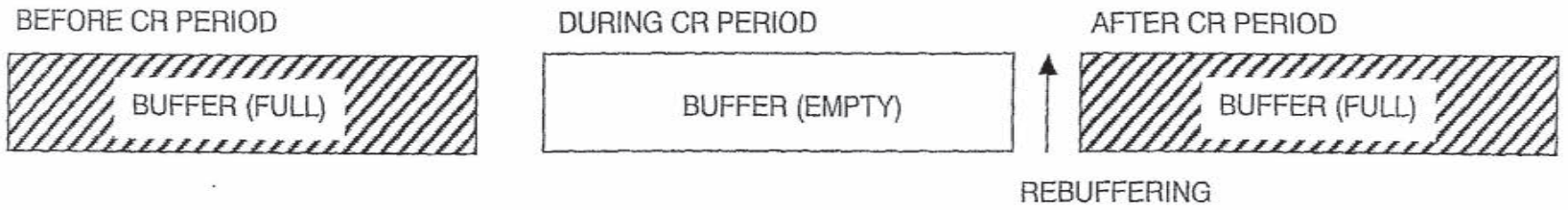

BEFORE CR PERIOD
DURING CR PERIOD
AFTER CR PERIOD
BUFFER (FULL)
BUFFER (EMPTY)
BUFFER (FULL)
REBUFFERING

FIG. 3

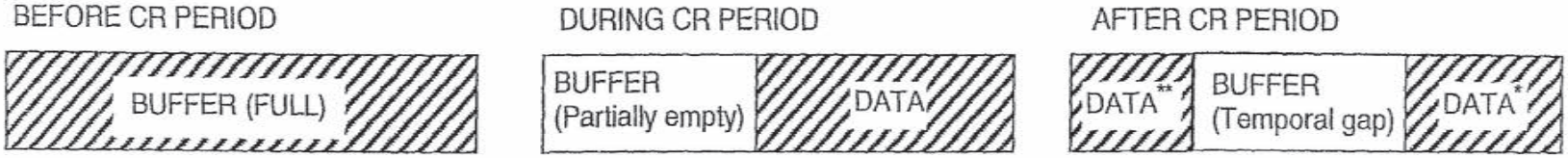

BEFORE CR PERIOD
DURING CR PERIOD
AFTER CR PERIOD
BUFFER (FULL)
BUFFER (Partially empty)
DATA
DATA**
BUFFER (Temporal gap)
DATA*

FIG. 4

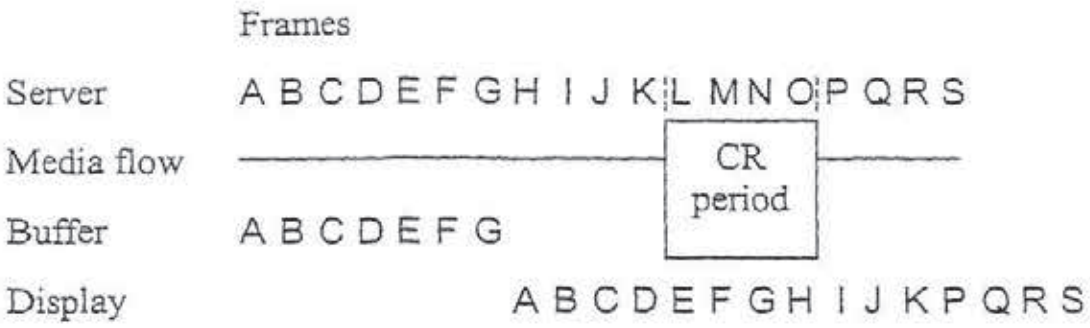

Frames
Server
A B C D E F G H I J K L M N O P Q R S
Media flow
CR period
Buffer
A B C D E F G
Display
A B C D E F G H I J K P Q R S

FIG. 5

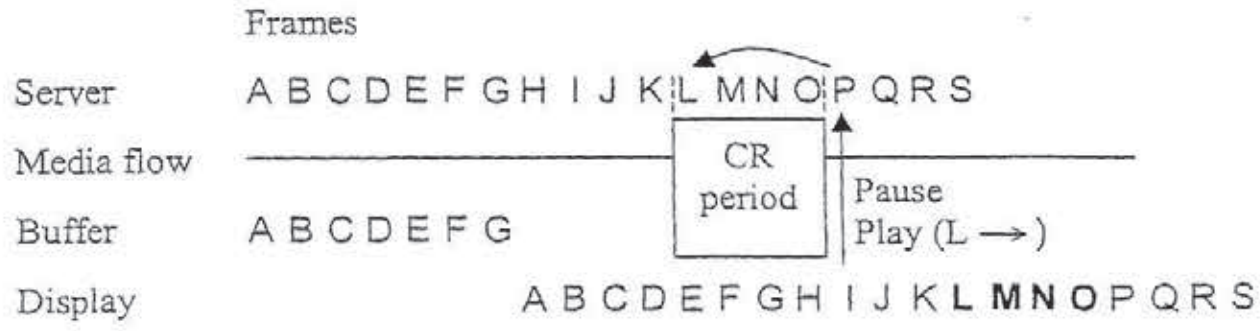

Frames
Server
A B C D E F G H I J K L M N O P Q R S
Media flow
CR period
Pause
Play (L →)
Buffer
A B C D E F G
Display
A B C D E F G H I J K L M N O P Q R S

FIG. 6

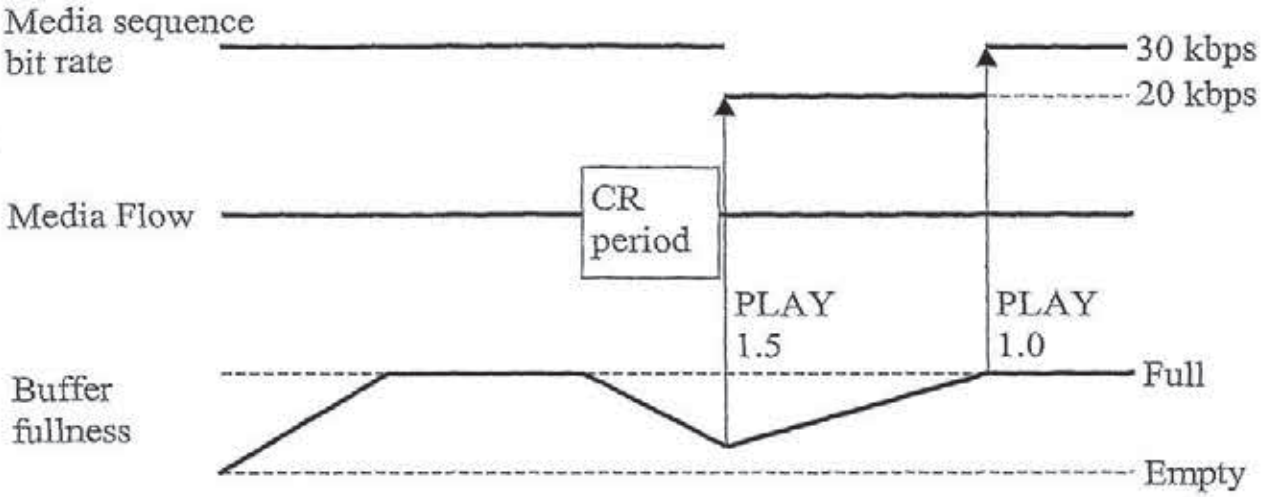

Media sequence bit rate
30 kbps
20 kbps
Media Flow
CR period
PLAY 1.5
PLAY 1.0
Buffer fullness
Full
Empty

FIG. 7

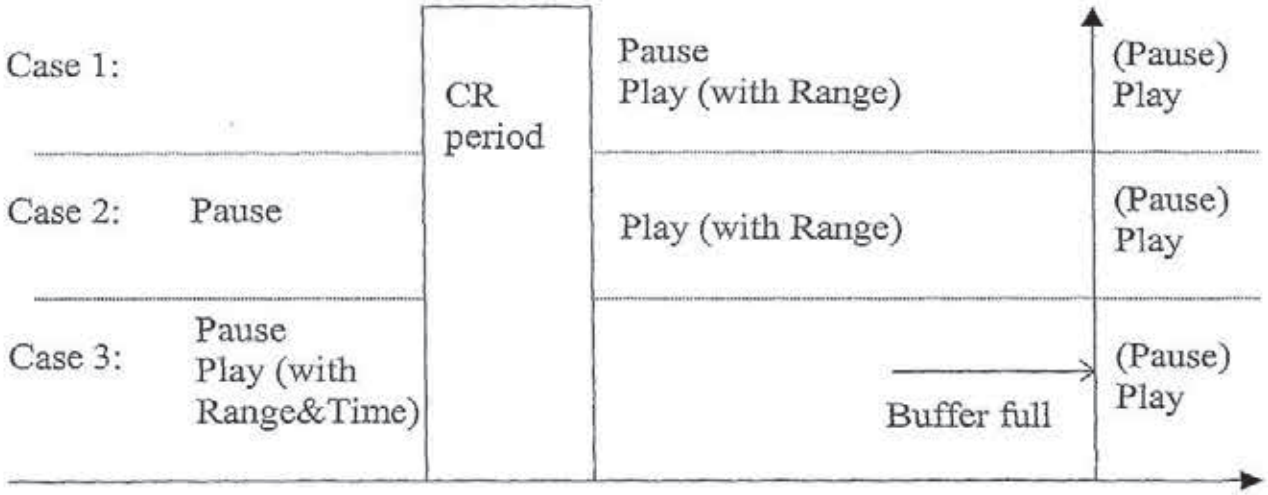

Case 1:
CR period
Pause
Play (with Range)
(Pause) Play
Case 2:
Pause
Play (with Range)
(Pause) Play
Case 3:
Pause
Play (with Range&Time)
Buffer full
(Pause) Play
time

FIG. 11

FIG. 8

Ex. 1018-1568

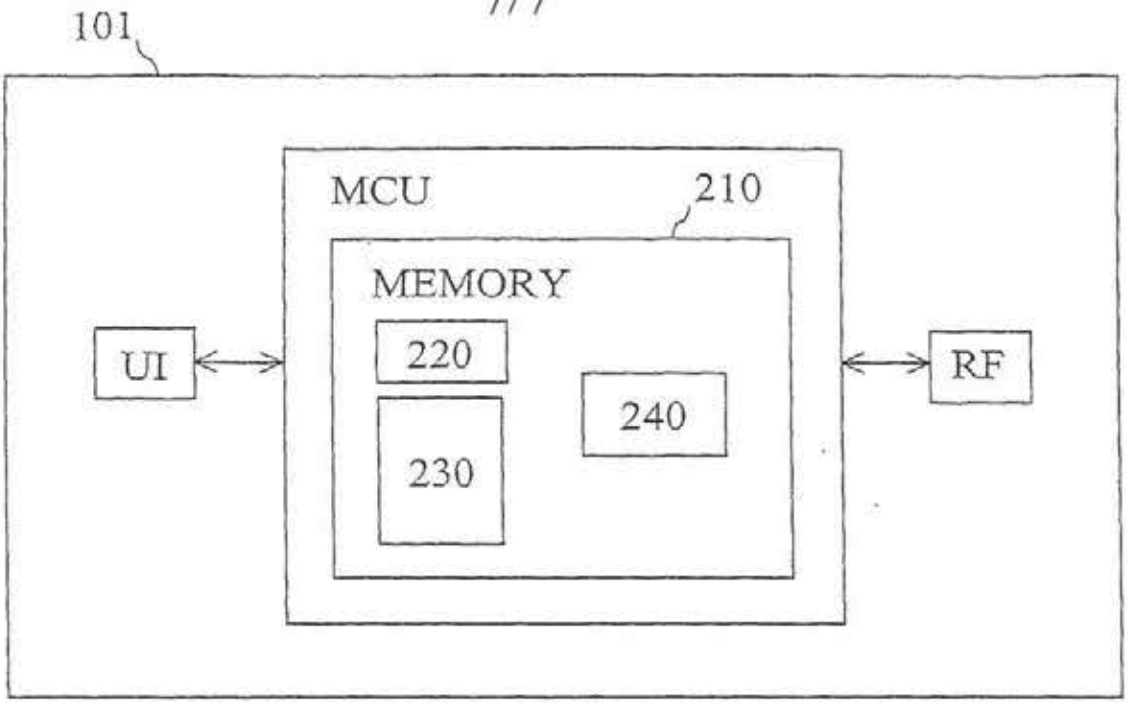

101
MCU
210
MEMORY
220
230
240
UI
RF

FIG. 9

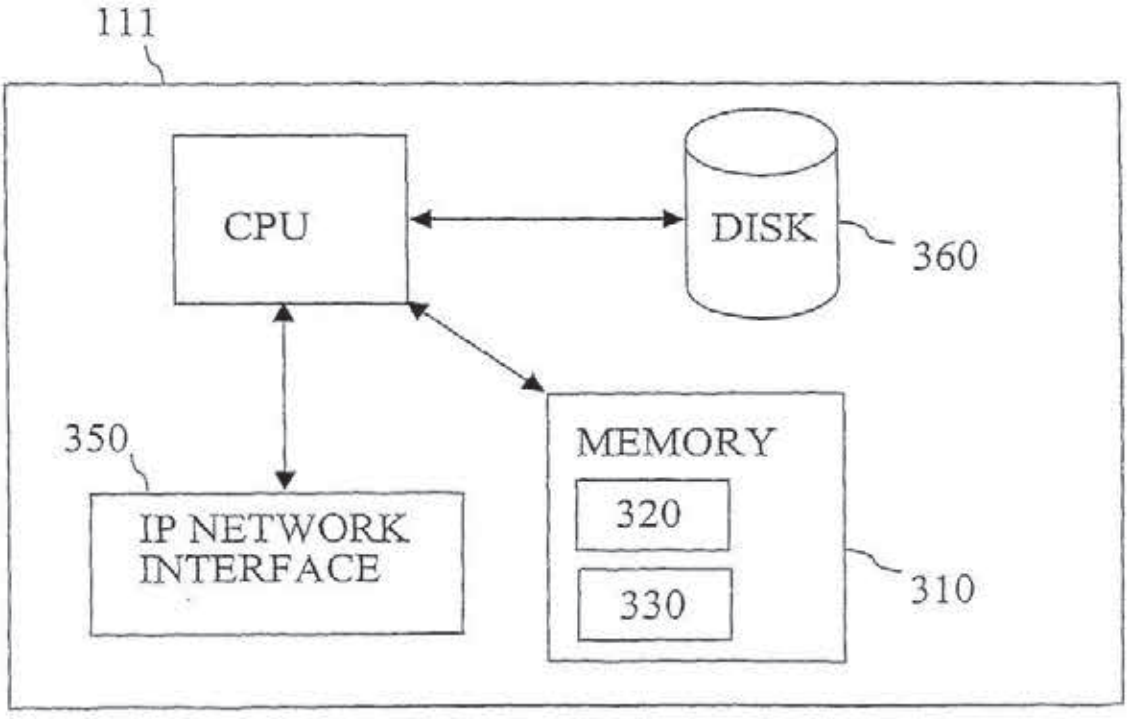

111
CPU
DISK
360
350
IP NETWORK
INTERFACE
MEMORY
320
330
310

FIG. 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00752

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: H04L 9/06
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)

IPC7: H04L, H04Q, H04B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, 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.
P,A	US 2003009576 A1 (APOSTOLOPOULOS, J.G. ET AL), 9 January 2003 (09.01.03), paragraphs [0009] - [0011]; figures 8-11	1-25
A	3GPP TS 26.233 V5.0.0: 3rd Generation Partnership project; Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS); General description (Release 5), March 2002, pages 1-12, see section 4, "Usage scenarios" and section 5, "General service architecture".	1-25

[X] Further documents are listed in the continuation of Box C. [X] See patent family 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 application or patent 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
8 December 2003	16 -12- 2003
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Fredrik Blomqvist /OGU Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00752

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 0130090 A2 (TELEFONAKTIEBOLAGET LM ERICSSON (PUBL)), 26 April 2001 (26.04.01), page 2, line 29 - page 4, line 19 -- --------	1-25

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

06/09/03

International application No.
PCT/FI 03/00752

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2003009576 A1	09/01/03	WO 03005761 A	16/01/03
WO 0130090 A2	26/04/01	AU 1065601 A	30/04/01
		EP 1224829 A	24/07/02

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 395 014 A1

(12) EUROPEAN PATENT APPLICATION

(43) Date of publication:
03.03.2004 Bulletin 2004/10

(51) Int Cl.7: **H04L 29/06**, H04L 29/08

(21) Application number: **02019052.6**

(22) Date of filing: **27.08.2002**

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LI LU MC NL PT SE SK TR
Designated Extension States:
AL LT LV MK RO SI

(71) Applicant: **MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.**
Kadoma-city, Osaka 571-8501 (JP)

(72) Inventors:
- **Rey, José Luis**
  64287 Darmstadt (DE)
- **Burmeister, Carsten**
  22081 Hamburg (DE)
- **Hakenberg, Rolf**
  64295 Darmstadt (DE)

(74) Representative: **Grünecker, Kinkeldey,**
Stockmair & Schwanhäusser Anwaltssozietät
Maximilianstrasse 58
80538 München (DE)

(54) A method of transmitting data streams with data segments of variable length

(57) A method of transmitting data streams between a sending host and a receiving host using a transport protocol comprising the steps of making available data streams at the sending host in the form of mono-rate or multi-rate content streams having the same content but different encoding rates, establishing a connection between the sending host and the receiving host, transmitting data streams with a selected encoded rate to the receiving host, and separating the data streams into data segments of variable length at the sending host as required during transmission.

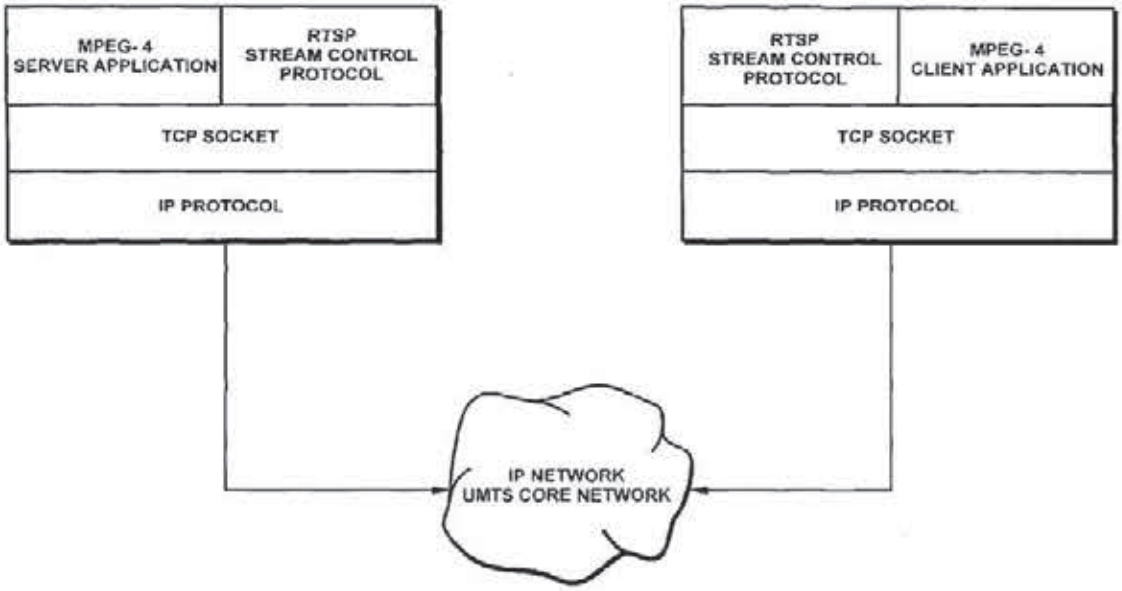


Fig. 1

Printed by Jouve, 75001 PARIS (FR)

EP 1 395 014 A1

Description

[0001] The present invention relates to a method of transmitting data streams between a sending host and a receiving host using a transport protocol such as TCP.

[0002] In network communications between a sending host and a receiving host, the transmission of data packets is mostly done on a best-effort basis. In the communication link, the available bandwidth and the current link conditions vary and hence, the end hosts must be provided with mechanisms to probe and adjust themselves to the current link conditions.

[0003] With the development of 3rd generation networks, so-called streaming applications are gaining importance. Streaming applications are understood as a continuous stream of data, e.g. video or audio data which is typically sensitive to time delays. In order that the user will be able to receive multimedia data streams on a portable receiver, for example a mobile phone, several requirements on the transport mechanism will be imposed in order that the delivered data meets with the resulting time constraints. These constraints are difficult to meet in a best-effort environment, which is subject to bandwidth variations and delays.

[0004] To cope with the problem of bandwidth adaption, it is common praxis to make several data streams with different encoding rates but same content available at the sending host. According to the available bandwidth, it is then switched between the data streams in order to adjust the rate to the available bandwidth. In other words, multi-rate content streams at the sending host are transmitted and switching operations between streams are performed depending on the available bandwidth. The current stream is forwarded to the receiving host in one piece and consequently, for every switching operation, the ongoing connection must be torn down and a new one must be set up. As a result, this operation is inefficient and decreases the data throughput.

[0005] Alternatively, it is possible to transmit only one stream of data to the receiving host and to discard data packets based on priority and time constraints. However, this leads to a loss of data and is therefore not an optimal solution.

[0006] The object subject to the present invention is to minimize the latency of the switching operation and to ensure fast recovery from an incorrect switching decision.

[0007] This object is solved according to the present invention with a method as set forth in claim 1. Preferred embodiments of the method are subject to various dependent claims.

[0008] The method of the present invention supposes the existence of multi-rate content data streams at the sending host. With the method of the present invention, it is possible to quickly switch between data streams that are variably segmented at the application layer as required during transmission. Instead of using data streams of fixed segment length, the data segments given to the socket are made smaller before, during and after the switching. For each content several video streams encoded with different rates are stored at the video server. Under normal sending conditions, i.e. the transmission has reached some equilibrium where the bandwidth is enough and the losses are not dominating, the video content is passed to the sending TCP socket in segments of the same size (irregardless of the stream encoding rate, the segment size is a function of the link, not of the encoding rate). Only when the server is notified (or knows by other means) that a stream switch might be needed, then it reduces the size of these segments.

[0009] Further, with the method of the present invention, it is neither necessary to modify the originally encoded data at the sending host or to modify the transport protocol nor to introduce additional control messages, making the method easy to implement.

[0010] Preferably, the switching is performed between data streams of different rate and the variable segmentation of the data streams is performed around the switching instant. Hence, a quick stream switching from the old to the new data stream can be performed. Further, preferably, the segmentation during, before or after the switching instance is made finer, such that the switching operation is more resilient to an incorrect decision and enables a quicker, further switching operation.

[0011] Advantageously, the step of switching between streams involves a feedback message (e.g. an RTSP command) i.e. communication between hosts when a switching operation is needed and when to make the necessary preparations therefor, for example, to start with finer segmentation.

[0012] It is furthermore of advantage if the feedback message includes at least one of the parameters, which comprise the length of data segment, an identifier of the data to be sent or a stream control command which makes the operation smoother and less susceptible to data losses.

[0013] Preferably, the length of the data segments is gradually increased after a switching instant to reduce interface communication, i.e. between application and TCP socket.

[0014] The present invention will be more fully understood from the following detailed description with reference to the accompanying drawings which show:

Figure 1: the basic architecture of a sending host and a receiving host connected by a communications network in which the method subject to the present invention can be performed,

Figure 2: a block diagram of a server architecture as the sending host in which the method of the present invention is performed, and

Ex. 1018-1574

Figure 3: a block diagram of a client architecture.

[0015] Figure 1 illustrates the basic architecture of a communications network for connecting a video server as a sending host to a video client acting as a receiving host. The communications network is, for example, a UMTS core network, an external network, such as the internet or an internal intranet over which the server and the client communicate by means of a transport protocol. As transport protocol TCP is selected as an example, but it is clear to a skilled person that any other transport protocol for transmitting data can be employed.

[0016] In a more general way, it will now be explained how TCP handles the exchange of data between the server and the client.

[0017] First of all, the video server needs to transfer its data to the protocol and tell it to send same to the client. To progressively read and send data, the protocol uses a kind of buffer with added functionalities, a so-called TCP socket. Generally, sockets are unequivocally identified by the IP addresses of the correspondents, the protocols and ports which are used. Once a connection between two sockets is established, the communication can be controlled by use of interface calls, which, for example may comprise "create a socket", "bind a socket to a transport address", "accept" or "close a connection", "send", "receive" or "write" data between the sender and the receiver.

[0018] Wide-spread socket implementations are UNIX BSD 4.2 and Windows Winsock 2.0, where equivalent interface calls are implemented.

[0019] Each interface call has parameters, e.g. for the "send" call, parameters of relevance are the socket descriptor, the identifier of the data to be sent and the length of the data stream. Once a data stream is passed in one call, no single bit of this data stream can be modified. The only way to abort transfer of this data stream is to break down the connection.

[0020] Nevertheless, in order to close the connection, the application must explicitly issue a "close" call. Interface calls can be queued as, for example, in FTP. When asked for a file, the FTP server usually issues two interface calls for this transmission, namely a "send" call indicating the socket, the identifier of the data stream and the length thereof and a "close" call identifying the socket. The two interface calls are put into the queue and when the "send" call is completed, the connection is irreversibly broken down because the "close" call was next in the queue.

[0021] Alternatively, for example, in case of TELNET, the connection is established and the socket is set to wait for incoming data, i.e. commands typed on the keyboard.

[0022] Of course, the application can be configured such that the server only issues a send call and waits for the client to close. However, this is uncommon.

[0023] Additionally, it is assumed that the TCP protocol provides a means for asynchronously signaling certain information about events in the socket. Often in the specification, the information will be an error message. In other cases, there will be information relating to the completion of processing a SEND or RECEIVE or another interface call.

[0024] Referring again to Figure 1, the video server provides, e.g. an MPEG 4 compressed data stream, i. e. a given video content, which is made available to the TCP socket in the form of mono-rate or multi-rate streams. Multi-rate streams are defined for the purpose of this and the following description, as one or several data streams with the same video content at different encoding rates. The MPEG data streams are forwarded to the TCP socket. Subsequently, by means of an IP protocol, data is transmitted to the video client. A stream control protocol, e.g. RTSP, should be used to perform such operating as PLAY, PAUSE or STOP the transmission.

[0025] With reference to Figure 2, which shows the basic architecture of a video server 200, at the MPEG server application 210, video content is made available in a buffer 211 at different coding rates. A switch 212, which receives stream control RTSP commands from RTSP block 220, such as "Play", "Pause", "Tear down" or "Change rate", selects one of the data streams which is forwarded to the TCP socket 230 in form of an MPEG file.

[0026] It is clear to a skilled person that other stream control protocols than RTSP could be used at the server and the client. The stream control protocol typically instructs to perform play, forwarding, rewinding, pause or stop functions. Additionally, if multi-rate content is available, switch to a higher or lower encoding rate can be instructed.

[0027] Once a server has received a request for a file containing a chosen content encoded at a desired rate, the server application forwards each of the segments, in which the file was previously segmented, one after another to the TCP socket by means of consecutive interface calls, e.g. consecutive SEND calls. Unlike in current FTP applications, the connection is not torn down after each data transmission process is completed. The data connection remains open, so that there is no need to build up a new connection every time a segment is transmitted. Following this, the TCP socket forms with these segments packet data units (PDUs) for transmission to the client, as it finds most appropriate

[0028] Upon establishing a connection, the TCP socket selects a segment size for the data streams which the server finds most appropriate for the link under consideration of the channel bandwidth and starts transmitting.

[0029] Initially, the size of the segments is made small before, during and after a switching operation in order to avoid a large waiting time before a further switch operation can be performed in case the previous switch decision was incorrect. Subsequently, the size of the data segments can be increased gradually throughout the connection time.

Ex. 1018-1575

[0030] Eventually, when there is a change in the available link bandwidth, the video client shall indicate a possible switch operation by means of, for example, an RTSP message and the server will start doing finer segmentation. Hence, the server switches between data streams of different encoding rates. This is done based on initial information and subsequently controlled by feedback from the client. Alternatively, this could be determined by the server, which would, however, impose processing requirements on the server. Also a combination of control mechanisms at the server and the client is possible. How the server or the client exactly makes the decision to perform a switching operation or indicates a possible anticipated switch is out of the scope of the present invention.

[0031] The basic architecture of a TCP streaming client 300 is illustrated in Figure 3. The client 330 typically consists of a TCP socket 310, an application layer buffer 320 and a buffer manager 330 for the application layer buffer.

[0032] The client 300 processes the data segments of variable length as they come. In response thereto, it sends ACKs and performs flow control in order to prevent an overflow of the TCP socket 310 while extracting the maximum bit rate possible from the link. To do this, the TCP socket indicates to the server, with each ACK, how much data it can receive until the socket is full using a window field in the header of the TCP protocol. If needed, the TCP client signals a zero-window telling the server to enter a "Persist" mode wherein no data is sent until the client allows to do so while the connection is maintained. The Persist mode and the signaling of a window field are standard functionalities of the TCP protocol. When the TCP circuit has capacity for more data, it signals this to the server. Streaming control commands are signalled to the server by means of RTSP commands.

[0033] The application layer buffer 320, also called "play-out buffer" receives data from the socket and transmits same to the MPEG client application 340. The manager 330 for the application layer buffer 320 receives a frame request from the MPEG client application and issues a socket request upon considering the state of the application layer buffer. The buffer manager 320 therefore insures that the bit rate with which the TCP socket 310 can deliver the new data to the application layer buffer 320 matches the bit rate at which the application layer buffer is emptied to the MPEG client application 340. The exact control mechanism of the buffer manager is out of the scope of the present invention.

[0034] Finally, from the MPEG client application, the data is transferred to a display 350 for visualization for the user.

Claims

1. A method of transmitting data streams between a sending host and a receiving host using a transport protocol comprising the steps of:

   making available data streams at the sending host in form of mono-rate or multi-rate content streams having same content but different encoding rates,

   establishing a connection between the sending host and the receiving host for the duration of transmission,

   transmitting data streams of the same content with a selected encoding rate to the receiving host, and

   separating the data streams in data segments of variable length at the sending host as required during the established connection.

2. The method according to claim 1, further comprising the steps of switching between data streams of different rate and performing the variable segmentation of the data streams around the switching instant.

3. The method according to claim 1 or 2, wherein the data streams are finer segmented during, before or after the switching instant.

4. The method according to claim 2, wherein the step of switching between streams involves a feedback message.

5. The method according to claim 4, wherein the feedback message includes signalling at least one of the parameters which comprise data segment length, an identifier of the data to be sent or a stream control command.

6. The method according to one of claims 1-5, wherein the receiving host communicates to the sending host when a stream switching is needed.

7. The method according to one of claims 1-6, wherein the sending host determines the timing of the switching instant dependent on the available bandwidth of the communication link.

8. The method according to one of claims 1-7, wherein the length of the data segments is gradually increased after a switching instant.

9. The method according to one of claims 1-8, wherein the received data streams at the receiving host are temporarily stored in an internal buffer.

10. The method according to one of claims 1-9, wherein the receiving host issues stream control commands

Ex. 1018-1576

to the sending host depending on the state of the internal buffer.

Ex. 1018-1577

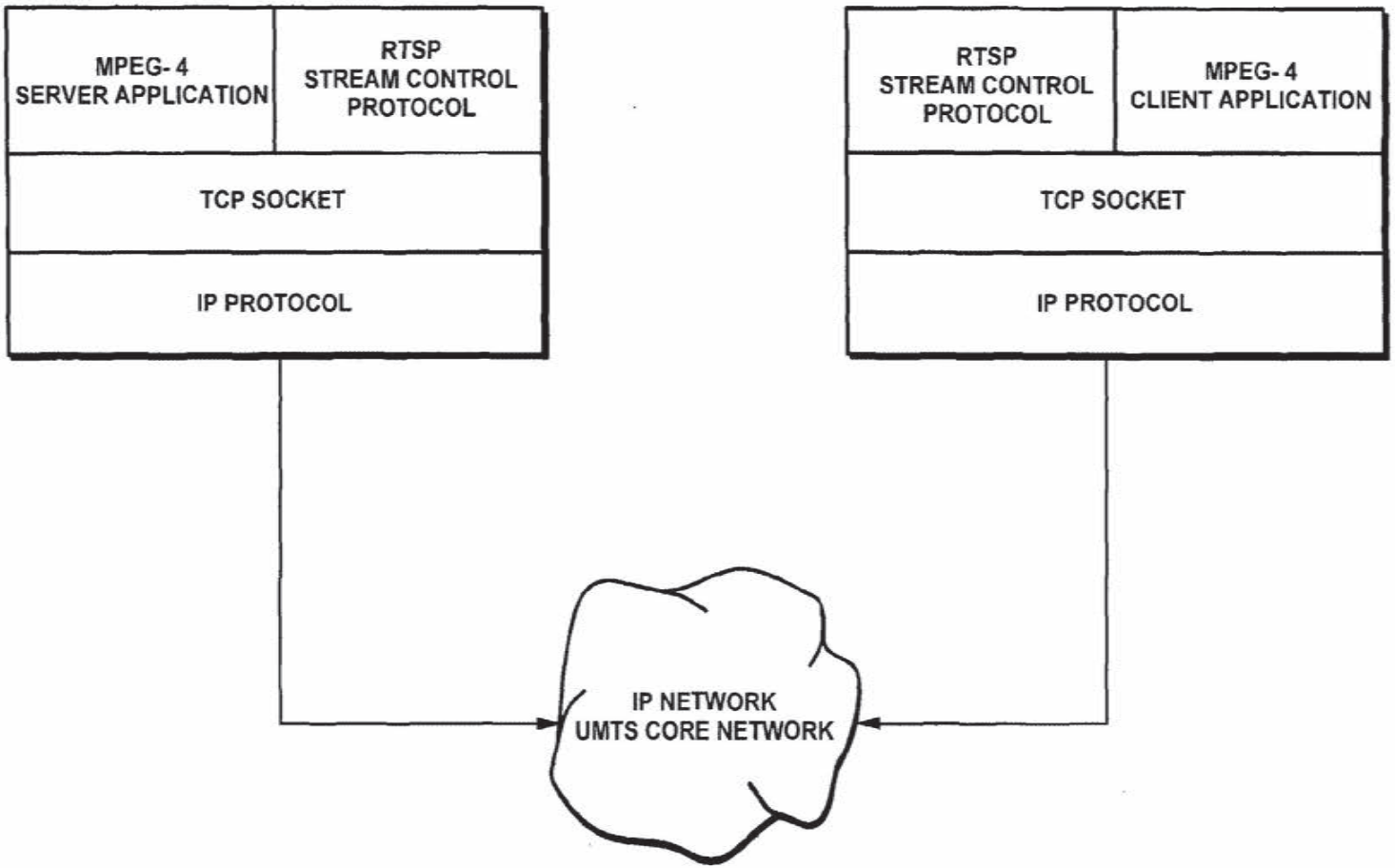


Fig. 1

Ex. 1018-1578

Ex. 1018-1579

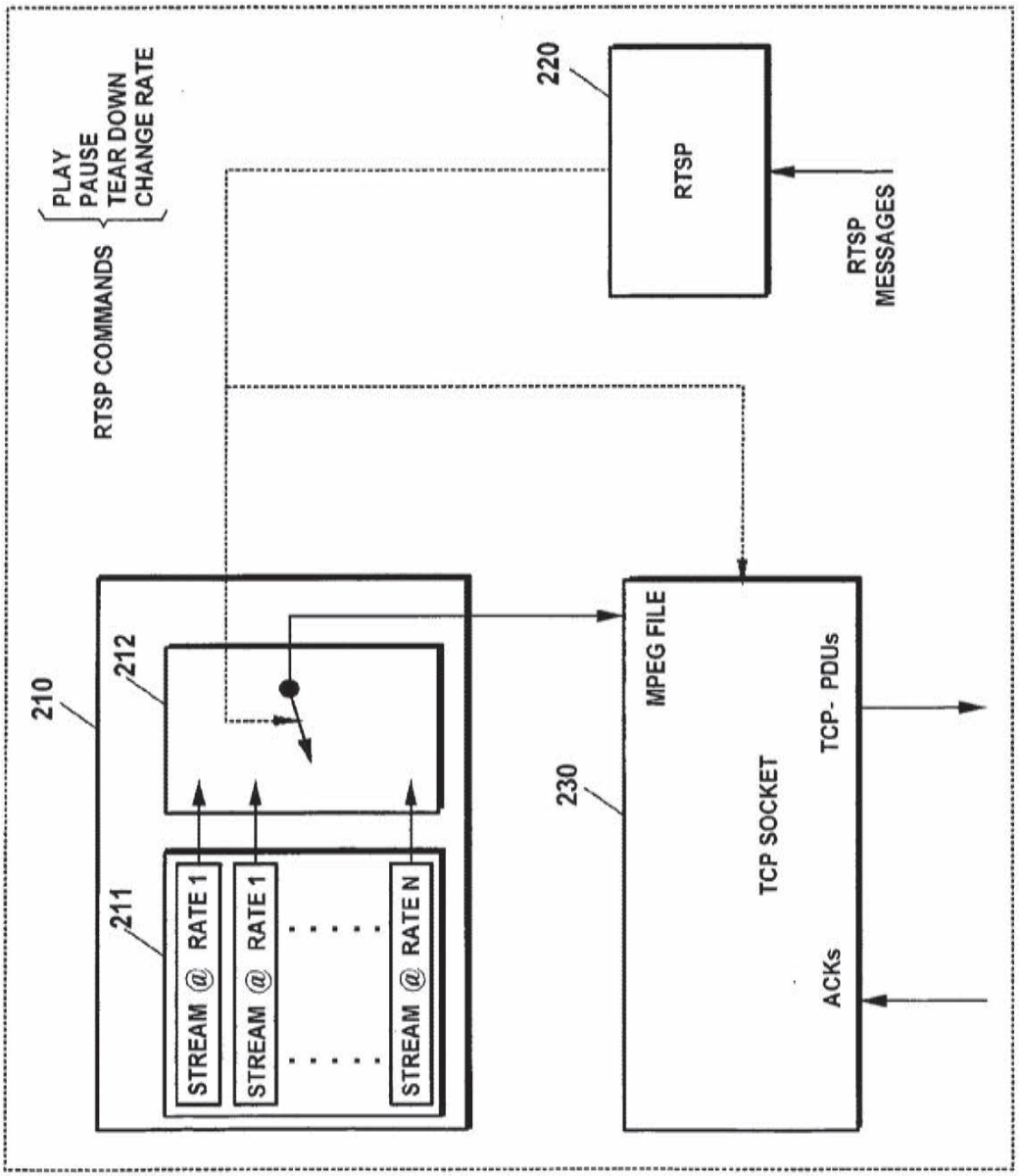


Fig. 2

Ex. 1018-1580

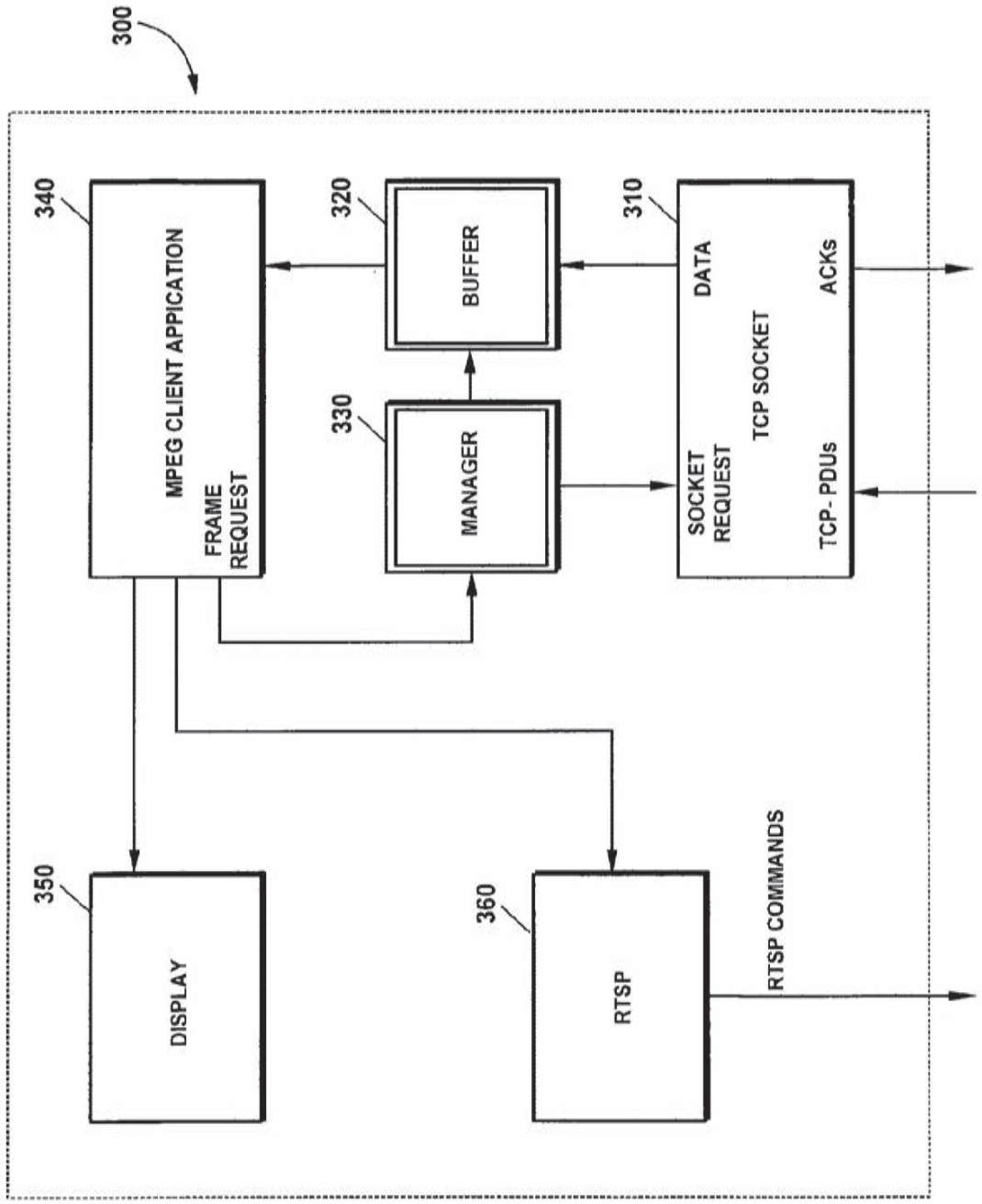


Fig. 3

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 01 9052

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
Y	CONKLIN G J ET AL: "VIDEO CODING FOR STREAMING MEDIA DELIVERY ON THE INTERNET" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE INC. NEW YORK, US, vol. 11, no. 3, March 2001 (2001-03), pages 269-281, XP001093477 ISSN: 1051-8215 * page 271, left-hand column, line 35 - line 44; figures 1,2 * * page 270, left-hand column, line 11 - line 20 * ---	1-10	H04L29/06 H04L29/08
Y	EP 0 786 885 A (NOKIA TELECOMMUNICATIONS OY) 30 July 1997 (1997-07-30) * abstract * * page 2, column 2, line 14 - line 34 * ---	1-10	
X	WO 02 054284 A (LEISTAD GEIRR I ;DANKWORTH HARALD (NO); FAST SEARCH & TRANSFER ASA) 11 July 2002 (2002-07-11) * page 2, line 31 - page 3, line 7 * * page 4, line 6 - page 5, line 6; figure 1 * * page 9, line 18 - page 10, line 15 * * page 11, line 3 - line 28; figure 3 * --- -/--	1	TECHNICAL FIELDS SEARCHED (Int.Cl.7) H04L

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	7 January 2003	Michael, T

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Ex. 1018-1581

European Patent Office

EUROPEAN SEARCH REPORT

Application Number
EP 02 01 9052

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	HUI J Y ET AL: "CLIENT-SERVER SYNCHRONIZATION AND BUFFERING FOR VARIABLE RATE MULTIMEDIA RETRIEVALS" IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE INC. NEW YORK, US, vol. 14, no. 1, 1996, pages 226-237, XP000548824 ISSN: 0733-8716 * page 228, left-hand column, line 15 - line 33 * * page 228, right-hand column, line 14 - line 32 * ---	10	
A	JONGWON KIM ET AL: "TCP-friendly Internet video streaming employing variable frame rate encoding and interpolation" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, vol. 10, no. 7, October 2000 (2000-10), pages 1164-1177, XP002226167 ISSN: 1051-8215 * page 1164, right-hand column, line 25 - line 40; figure 1 * ---	7	TECHNICAL FIELDS SEARCHED (Int.Cl.7)
A	WO 02 052859 A (DURRANT SIMON ;LAKE JULIAN (GB); MULROY PAT (GB); PA CONSULTING SE) 4 July 2002 (2002-07-04) * page 3, line 12 - line 30 * -----	7	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
MUNICH	7 January 2003	Michael, T

CATEGORY OF CITED DOCUMENTS

X : particularly relevant if taken alone
Y : particularly relevant if combined with another document of the same category
A : technological background
O : non-written disclosure
P : intermediate document

T : theory or principle underlying the invention
E : earlier patent document, but published on, or after the filing date
D : document cited in the application
L : document cited for other reasons

& : member of the same patent family, corresponding document

EPO FORM 1503 03.82 (P04C01)

Ex. 1018-1582

ANNEX TO THE EUROPEAN SEARCH REPORT ON EUROPEAN PATENT APPLICATION NO. EP 02 01 9052

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

07-01-2003

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0786885 A	30-07-1997	FI 960382 A	27-07-1997
		EP 0786885 A2	30-07-1997
		US 5914960 A	22-06-1999
WO 02054284 A	11-07-2002	NO 20010056 A	05-07-2002
		WO 02054284 A1	11-07-2002
WO 02052859 A	04-07-2002	WO 02052859 A1	04-07-2002

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
30 January 2003 (30.01.2003)

PCT

(10) International Publication Number
WO 03/009581 A1

(51) International Patent Classification[7]: H04N 5/00, 7/24

(21) International Application Number: PCT/GB02/03307

(22) International Filing Date: 19 July 2002 (19.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
01306206.2 19 July 2001 (19.07.2001) EP

(71) Applicant *(for all designated States except US)*: **BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY** [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **JACOBS, Richard, James** [GB/GB]; 15 Through Duncans, Woodbridge, Suffolk IP12 4EA (GB). **WALKER, Matthew, David** [GB/GB]; 9 Thirlmere Court, Felixstowe, Suffolk IP11 9SN (GB). **MIRAS, Dimitrios** [GR/GR]; 12 Proti Parodos Avras, 24100 Kalamata (GR).

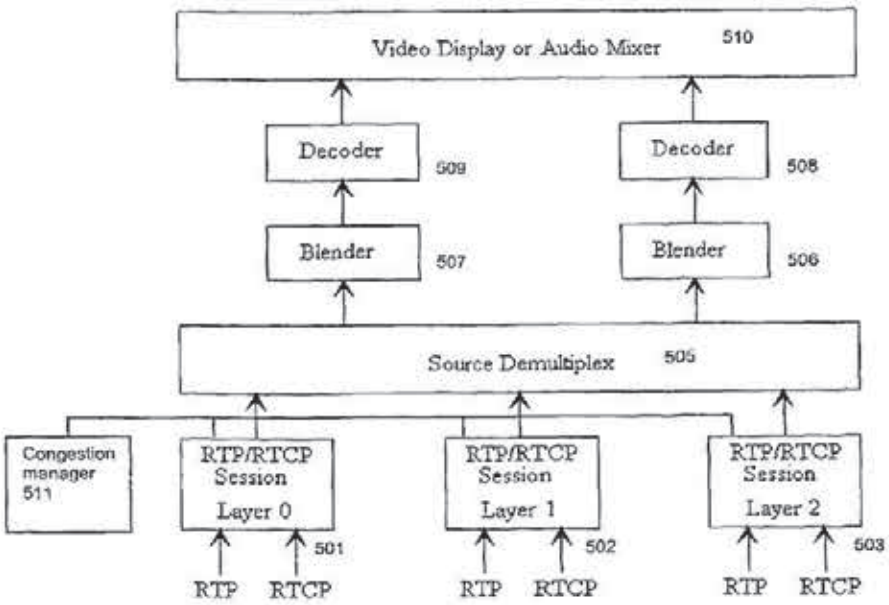
(74) Agent: **WILLIAMSON, Simeon, Paul**; BT Group Legal, Intellectual Property Department, Holborn Centre 8th Floor, 120 Holborn, London EC1N 2TE (GB).

(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, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, 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: VIDEO STREAM SWITCHING



(57) Abstract: A technique is described for providing a client on a packet based network with a stream of encoded video data. The system is able to maximise the bit-rate of the video stream by adapting to fluctuations in network capacity. The technique is characterised in that adaptation of the bit-rate of the transmitted encoded video data is timed to occur upon a scene change in the video sequence. In this way the interruption to the viewer when the perceived quality of the video sequence increases or decreases is minimised as it is 'hidden' in the scene change. The technique is described as applied to hierarchically encoded video data but equally may be applied to other encoding techniques which adapt to network conditions.

WO 03/009581 A1

Published:
— *with international search 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.

Video Stream Switching

The invention is in the field of video streaming over packet networks and in particular concerns the adaptive transmission of data in response to network congestion.

In recent years, the Internet has experienced a proliferation in transmission of real-time multimedia, mainly in the form of streamed audio-visual content, either delivered live or from pre-recorded sources. Furthermore, traditional forms of multimedia such as streaming and conferencing are being followed by applications with richer content such as Internet multi-channel TV and complex immersive environments. This increase in traffic will put strains on the network and it is therefore desirable that application programs are designed to respond to congestion if stability of the network is to be maintained. It is desirable that network conditions are monitored and output bit-rates adjusted to the available bandwidth.

There is a more immediate advantage to the user in having an adaptive output bit-rate which is that the highest possible transmission bit-rate that the network will allow is used and therefore the user will always receive the best possible image quality. There are a number of known approaches to adaptive quality video streaming, one of which is hierarchical coding. In this technique the original video data is encoded into a number of discreet streams called layers, where the first layer consists of basic data of a relatively poor quality and where successive layers represent more detailed information so that layers can be added to increase the image quality or layers can be taken away, depending on the available bandwidth.

As the bit-rate available to a session is subject to significant variations in the number of layers that are transmitted, quality fluctuations occur in the decoded image as layers are added or dropped. When layers are added or dropped frequently the fluctuations in quality may become disturbing to a viewer.

In a first embodiment of the present invention there is provided a method of operating a multimedia server, said method comprising:

providing a stream of video data representing a video sequence to an output of the multimedia server, wherein the output of the media server is connected to a packet based network,

measuring a property of the video data in order to determine the occurrence of a scene change in the video sequence,

detecting the available bandwidth on the network,

varying the bit-rate of the stream of video data,

wherein the method is characterised in that

variation in the bit-rate of the video data is controlled to occur in response to variations in the capacity of the network (203) and preferentially with a scene change in the video sequence.

The term preferentially used herein is meant to indicate that in certain circumstances (discussed in greater detail below) it may not be practical to wait for a change of scene to occur in the video sequence before varying the bit-rate of the video stream, and in such circumstances the bit-rate will be changed at points in the video stream which do not correspond to a change of scene.

The term scene change (or change of scene) is intended to refer to a sudden change in a video sequence within the space of a one or a very few frames such as typically occurs at a change of scene, whether or not there has been an actual change of scene.

Embodiments of the present invention will now be described, by way of example only, with reference to the following figures, where;

Figure 1 is a schematic diagram of the content-based inter-stream session bandwidth sharing architecture.

Figure 2 shows the order of packets in a layered coding system.

Figure 3 is a diagram of a media server.

Figure 4 is a diagram of the network interface of the media server shown in Figure 3.

Figure 5 is a diagram of a client.

Figure 6 is a diagram of a client according to a first embodiment of the present invention.

The client/ server arrangement for a known hierarchical streaming technology is shown in Figure 1. A media server 202 is provided which has access to compressed audiovisual data 201, which may be fed 'live' from an outside broadcast or may be pre-compressed and stored in a database. The data source 201 may be on the same premises as the media server 202 and linked via an intranet. The media server 202 runs on a suitable server computer and which has access to the Internet 203.

A video viewer, hitherto referred to as the client 204, running on a PC suitably configured to have access to the Internet 203, may connect to the media server 202 via the Internet 203 and thus the client 204 is able to access content. A suitable PC terminal is used.

Layered video compression is achieved with the 1998 version of H.263 but equally may be any other codec, such as MPEG4. Each layer in the hierarchy is coded in such a way as to allow the quality of individual pictures to be enhanced and their resolution to be increased, and additional pictures to be included to increase the overall picture rate, as explained with reference to Figure 2. Figure 2 illustrates a typical dependency between pictures in an H.263 scalable layered coder, with the boxes representing the frames for each layer and arrows showing the dependency between frames. The lowest row shows original, un-coded frames. The next row shows the lowest layer (Layer 0) of the hierarchy which is coded at half the frame rate of Layer 1. Frames in Layer 0 are predicted from the previously encoded frame, as in conventional video compression. Frames in Layer 1 may be predicted from the previously encoded frame in Layer 1 and, if present, the temporally simultaneous Layer 0 encoded frame. Frames in Layer 2 may be predicted from the previously encoded frame in Layer 2 and, if present, the temporally simultaneous Layer 1 or Layer 0 encoded frame. The H.263 specification allows for 15 layers; in the present embodiment server and client software is not limited in the number of layers that can be used, but in this case a database has been generated with video streams containing four layers.

Figure 3 shows the architecture of media server 201. In the present embodiment the audiovisual source is a database 301 of compressed video data and the media server 201 is responsible for reading the compressed data from a database, packetising and distributing it. The data is distributed according to the

Real-time Transport Protocol (RTP) which provides end-to-end network transport functions suitable for applications transmitting real-time data such as audio, video or simulation data. RTP does not address resource reservation and does not guarantee quality-of-service for real-time services. The data transport is augmented by a control protocol (RTCP) to allow monitoring of the data delivery and to provide minimal control and identification functionality. RTP and RTCP are designed to be independent of the underlying transport and network layers. The media server 201 has four major components: database reader 302, video RTP packetiser 303, audio RTP packetiser 305 and a network interface 304. Video and audio data are stored as files of compressed bits in the database 301. The database reader 302 retrieves and synchronises the compressed audio and video from the database 301.

The audio and video data is then sent to a RTP packetiser 303,305. Audio and video information is transmitted over the IP network using the User Datagram Protocol (UDP) and Real-time Transport Protocol (RTP) packetisation. UDP provides a checksum to detect transmission errors, but does not guarantee data delivery: packets may be lost, duplicated or re-ordered. RTP provides end-to-end delivery services, such as payload type identification, sequence numbering, time-stamping and delivery monitoring. RTP packetisers 303,305 attach the RTP Header and, in the case of video packets, the H.263 Payload Header which provides some protection from packet loss. The Payload Header contains information specific to the video stream, such as motion vector predictors, which is obtained by decoding the compressed bit stream.

The rate at which data is read from the database 301 is controlled from the network interface 304, which is illustrated in greater detail in Figure 4. With reference to Figure 4 the arrangement of the network interface 304 for the video data will now be described. A packet numbering module 401 assigns a number to each packet of data. These packet numbers are used by the client to assemble packets into the sequence required for decoding. The number of bits produced by compressing a picture is in general not constant, even when the transmission rate is constant. The first picture of a new scene will usually produce a larger number of bits than a picture where there is little movement or detail. The encoding process will use a control strategy together with an output data buffer to smooth these variations in preparation for transmission at a

constant bit rate. In the case of layered coding, each layer has its own control strategy and buffer as it has its own constant bit rate constraint to meet. The delay across each of these buffers will not be constant and will not be the same for each layer. Delay variation between the layers is therefore introduced at the source of the data. Transmission across a network may cause this delay variation to increase. A packet numbering scheme is therefore needed to ensure that a client can arrange the packets it receives into the order required for decoding. This scheme is required to handle the cases of the client not receiving all of the layers and of packet loss in the layers that it does receive.

Each layer is transmitted as an independent RTP Session on a separate IP address by a session handler 403, 405, 407, 409. The rate at which data is transmitted is controlled by the Transfer Rate Control module 402, which counts Layer 0 bytes to ensure that the correct number are transmitted in a given period of time. The transmission rate of the other layers is smoothed and locked to the rate of Layer 0 using First-In First-Out (FIFO) buffer elements 404, 406, 408.

The client 204 of the known hierarchical streaming technology will now be described with reference to Figure 5. Each RTP/RTCP Session associated with each layer of encoded data has a session handler 501, 502, 503, 504 at the client which is responsible for receiving RTP packets from the network.

The packets are then passed to a source demultiplex module 505. This receives packets from all layers and all sources, demultiplexes them by source, and routes them to a blender module 506, 507 for that source.

The blender module 506, 507 receives packets from all layers from one source in the order that they were received from the network. This may not be the order required for decoding because of packet inter-arrival jitter or packet loss. The blender module 506, 507 uses the packet numbers in the RTP headers to arrange the packets from each layer in order and then combines the packets from all layers together. The output of the blender module 506, 507 is a single stream of packets which are in the correct order for decoding. This is the same order as they come out of the packet numbering module 401 of the media server 202.

Packets are then sent to the decoder 508, 509 where the packets are decoded into 20ms blocks of audio samples or to video pictures. In the case of pictures, these are rendered to a window on the display 508.

Also provided is a congestion manager 511 to which the session handlers 501-504 report packet loss. If packets are being consistently lost, indicating network congestion, the congestion manager 511 will instruct the session handler responsible for the highest layer of compressed data to terminate the RTP/RTCP Session. Periodically the congestion manager 511 will instruct an experimental joining of a layer via the appropriate session handler to test whether there is available bandwidth in the network. If this experiment is successful, i.e. that substantially all of the packets of each layer are getting reaching the client, then the new layer will be adopted. In this way the maximum available bandwidth is employed.

It possible for the congestion manager 509 to instruct layers to be dropped and restored rapidly as the network bandwidth fluctuates. This could be annoying for a viewer and so a way of 'hiding' the shift is employed. The layered encoding method described above is adapted so that if a layer is to be dropped or added, the changeover preferentially occurs during a scene-change in the video data. A video scene is typically perceived as the number of consecutive frames within a video sequence that do not show significant changes in the video content. Within a video scene or shot, the camera action may be fixed or may exhibit a number of relatively uniform changes like panning, zooming, tracking etc. Scene changes may be recognised as abrupt transitions of the camera action or gradual transitions. In order to identify the scenes within a video sequence the assumption is that the levels of motion energy as well as those of luminance and colour do not change much between successive frames within a single scene. Techniques of scene boundary identification include pixel differencing, motion vector and block matching techniques. A very sudden change in the content of temporally adjacent frames will thus appear as a change of scene; clearly it is conceivable that such a change might not actually relate to a scene change at all, but, for example, may instead be the consequence of a large foreground object coming rapidly into view. However, such an abrupt change will nonetheless be a good place to add or remove a

session layer and thus the term scene change is meant to cover such large changes in frame content from one frame to another.

Figure 6 illustrates a system which detects congestion in the network and also imposes the extra condition of adding or dropping a layer upon a scene change. The server 601 comprises the components of the media server 202 of Figure 1 and the unit 603 comprises the components of the client 204 also shown in Figure 1. Video data from the audiovisual source is also passed to a content analyser 604 which segments the video sequence into scenes. The segmented video is then passed to an objective quality assessment module (OQAM) 605 which assesses the contribution to the perceived quality of a scene that each component layer provides.

These results are then passed to an inter layer adaptation module (ILAM) 607. The function of the ILAM 607 is to continuously calculate the number of layers that maximise the perceived quality for the session. The ILAM 607 also receives input from a sender congestion manager 608 which reports on the bandwidth available to the session on the network 203. The sender congestion manager 608 receives feedback from a client congestion manager 609 on the number of packets that have been received. If this matches the number of packets that were sent then the bandwidth is known to be the current transmission rate. If packets are being lost then the bandwidth available to the session is less than the transmission rate, in which case the sender congestion manager 608 informs the ILAM 607 that a layer should be dropped.

In order to select which layer should be dropped the ILAM 607 couples the bandwidth required by a layer with its contribution to the quality of the complete image, as calculated by the OQAM 605. The ILAM 607 performs an exhaustive search on all of the bandwidth/quality values. When the ILAM 607 has selected which layer is to be dropped from a particular scene, the timing of the drop is preferentially set to coincide with the transmission of that particular scene. In this way the drop in quality occurs with the scene change and is thus much less noticeable to a viewer than if the quality change had occurred during a scene. In the case that no packets have been lost for a preset period of time the sender congestion manager 608 will request the ILAM 607 to add in a layer to test

whether all of the available bandwidth is being employed. If no packets are lost during this experiment then the newly added layer is maintained. This process of experimentation with adding layers is continued until a significant proportion of packets are lost, in which instance the system can be confident that all of the available bandwidth is being employed. Again, the timing of the introduction of a layer is set to occur preferentially as the scene changes.

Under a given bit-rate allocation the levels of perceived quality do not change considerably within a scene, but scene cuts cause considerable changes in perceived quality, especially when the content features (spatial and motion energy) change a lot between subsequent scenes. As a consequence there will also be a significant difference in the corresponding quality scores for those successive scenes which may justify a rescheduling of the number of layers in the stream.

The invention is not limited in use to hierarchical encoding schemes. The invention is suitable for use in any encoding technique where the adaptation of transmission bit-rate to accommodate fluctuations in network capacity occurs. For instance the invention may be applied to a transcoding system where encoded data streams are transcoded from a high bit-rate to a low bit-rate or from a low bit-rate to a high bit-rate. The present invention would be suitable to reduce the impact for the viewer as the output bit-rate shifts in response to network conditions by timing the transition to occur upon a scene change in the encoded video sequence. Another example of an adaptive video streaming technique to which the invention may be applied is where multiple independent video streams of different bit-rates are transmitted. In this case the client chooses which stream to accept based on session bandwidth. The client may switch from one stream to another as bandwidth fluctuates; the present invention ensures that the switch is timed to coincide with a scene change in the encoded video stream.

Naturally, there may be times when a change of scene is such a long way away that it is advantageous to switch from one bit-rate to another other than during a change of scene. For example, consider the case where multiple independent video streams at different bit-rates are available for transmission by a media server to a client as, for example, described in co-pending European patent

application No. 00310594.7 the contents of which are hereby incorporated herein by way of reference. In such a case, the server may be capable of transmitting a first stream at a bit-rate of 500kbit/s and a second, higher quality stream at a bit-rate of 1500kbit/s. The client may initially request that the server transmit the first stream at a transmission rate of 1000kbit/s. If the network is not congested and all of the packets transmitted are successfully received by the client, the receive buffer at the client will start to fill with data at a rate of 500kbit/s, since the client will only be removing data from the buffer at a rate of 500kbit/s. After say 10 seconds of filling up the buffer at this rate, the client will have a buffer of ten seconds worth of data at which point it may decide it can attempt to receive the higher bit-rate second stream of video data from the server and thus sends an appropriate request to the server to this effect. If, however, the server is aware that the client has a receive buffer of 5 Mbytes size, it knows that it may continue sending data from the first stream at the rate of 1000kbit/s for at least another 150 seconds before the receive buffer overflows, causing problems for the client. Therefore, the server may attempt to wait for a specified period to see if a scene change occurs during this interval. If so, a switch to the second higher bit-rate signal is not made until the change of scene occurs. Of course, if there is no change of scene within the determined period, the server switches to the higher rate anyway. In this example, a period of only ten seconds is deemed appropriate as the waiting time.

Note that instead of simply waiting to see if a change of scene occurs in the specified period and switching at the end of the period if no such change of scene is detected, an alternative method would be to pre-analyse the video to be sent (clearly this only applies to pre-stored video data and not live video) and to note when changes of scene occur. In such a case, the server, upon receipt of a request from the client to switch streams, could search to see if a suitable change of scene will occur within the predetermined period and if not to switch immediately to the new bit stream.

In the present example, upon switching to the higher rate bit stream, the client may have requested a transmission rate of 1500kbit/s corresponding to the rate at which the data will be drawn from the receive buffer by the client. In such a case, the buffer size of 10 seconds should remain constant so long as all of the

transmitted packets are successfully received by the client. However, in the event of congestion on the network, a proportion of packets may fail to arrive at the client. In such a case, the server will be warned of this via the RTCP. If the congestion is sufficiently severe, the server may deduce that the buffer is in danger of emptying which would cause a break in the video displayed by the client to occur. To prevent this, the server may switch back to the lower bit rate stream. Via the notification of how many packets are being lost, the server can deduce how long it will be before the buffer is emptied. For this time, the server can wait to see if a change of scene occurs, and if so, the new stream will be switched to at that point. Note that it would also be possible for the client to simply request that the new stream be switched to upon detecting that its receive buffer is emptying at an unsustainable rate.

Note that the amount of data in the buffer is actually of less significance than the rate of change of the amount of data in the buffer. Thus it is preferably this quantity which either the server or the client measures in order to determine whether to change from one bit stream to another (or whether to add or drop a layer in the first example).

CLAIMS

1. A method of operating a multimedia server (202), said method comprising;

providing a stream of video data representing a video sequence to an output of the multimedia server (202), wherein the output of the media server (202) is connected to a packet based network (203),

measuring a property of the video data in order to determine the occurrence of a scene change in the video sequence,

detecting the available bandwidth on the network (203),

varying the bit-rate of the stream of video data,

wherein the method is characterised in that

variation in the bit-rate of the video data is controlled to occur in response to variations in the capacity of the network (203) and preferentially with a scene change in the video sequence.

2. A method of operating a multimedia server (202) in accordance with claim 1 wherein upon detecting a decrease in the available bandwidth the bit-rate of the video data is decreased.

3. A method of operating a multimedia server (202) in accordance with claim 1 wherein the bit-rate of the video data is periodically increased in order to ensure that all of the available bandwidth is being used.

4. A method of operating a multimedia server 202 in accordance with claims 1 to 3 wherein said stream of video data is comprised of a plurality of hierarchically encoded layers.

5. A method of operating a multimedia server (202) in accordance with claim 4 wherein the bit-rate of the video data is varied by adding or dropping layers from the stream of data transferred to the output port.

6. A method of operating a multimedia server (202) in accordance with claims 4 or 5 wherein the objective quality contribution of each layer to the

overall quality of the video sequence is measured in order to determine which of the layers is to be added or dropped from the stream of data transferred to the output port.

7. A method of operating a multimedia server 202 in accordance with claims 1 to 3 wherein said stream of video data is comprised of a plurality of independently encoded flows each of which encoded at different bit-rates.

8. A method of operating a multimedia server (202) in accordance with claim 7 wherein the bit-rate of the video data is varied by switching between independently encoded flows.

9. A method of operating a multimedia server 202 in accordance with any of the preceding claims in which the source 201 of video data is stored in encoded form in a database.

10. A method of operating a multimedia server 202 in accordance with any of claims 1 to 8 in which the source 201 of video data is encoded in real-time.

11. A method of operating a multimedia server 202 in accordance with any of the preceding claims wherein, upon detecting a variation in the capacity of the network such as to render it appropriate for a change in bit rate to occur, the server waits for a predetermined time to attempt to detect a change of scene, and if such a change of scene is detected within the predetermined time, the change of bit rate is caused to occur at the time of the change of scene, but if a change of scene is not detected within the predetermined time, then the change of bit rate is caused to occur at the expiry of the predetermined time.

12. A method of operating a multimedia server 202 in accordance with any of the preceding claims wherein, upon detecting a variation in the capacity of the network such as to render it appropriate for a change in bit rate to occur, the server determines if there is going to be a change of scene within a predetermined time, and if it is determined that there will be such a change of scene within the predetermined time, the change of bit rate is caused to occur at the time of the change of scene, but if it is determined that there is not going to

be a change of scene within the predetermined time, then the change of bit rate is caused to occur substantially immediately after having made such a determination.

13. The method of claim 12 further including the step of, prior to providing a stream of video data to the output of the media server, processing the video data to determine the positions within the video data in which changes of scenes occur and making this information available for permitting whether a scene change will occur within a predetermined period at any point along the stream of video data.

14. A multimedia server (202) comprising;

a reader (302) for reading video data representing a video sequence from a source (301),

a scene change detector (604) for detecting changes of scene in the video sequence,

a bit-rate controller (607) for controlling the rate at which the video data is transferred from the reader to an output port, wherein the output port is capable of communicating with a client (204) on a packet network (203);

the media server further comprising means (608) for detecting available bandwidth between the multimedia server (202) and client (204) on the packet network,

characterised in that

the bit-rate controller (607) is arranged to vary the bit-rate of the video data transferred to the output port in dependence on the detected bandwidth preferentially in correlation with a scene change detected by the scene change detector (604).

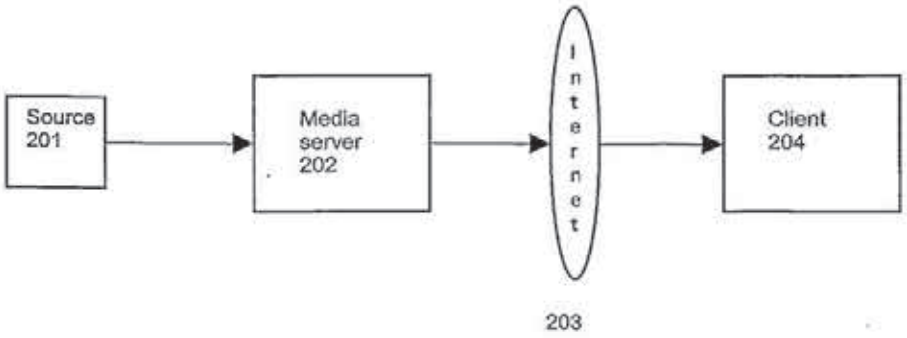


Figure 1

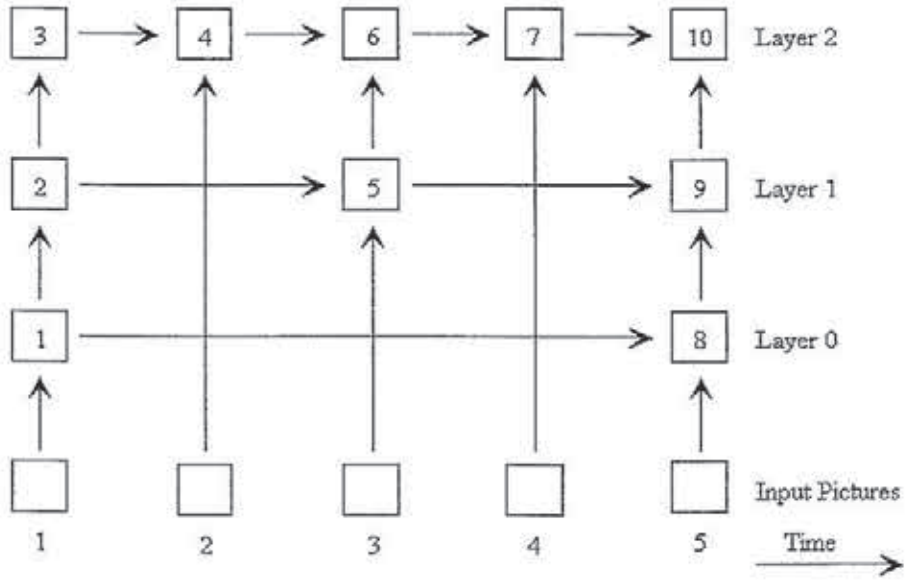


Figure 2

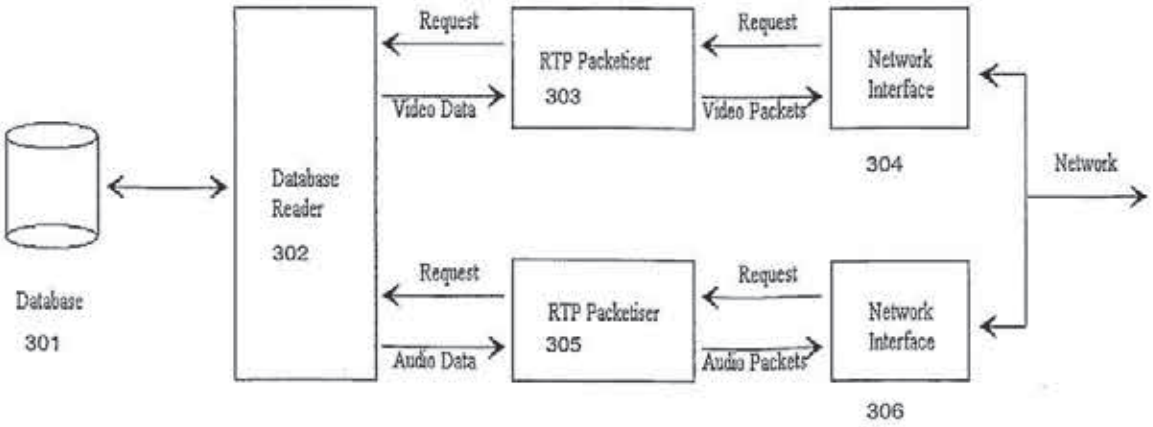


Figure 3

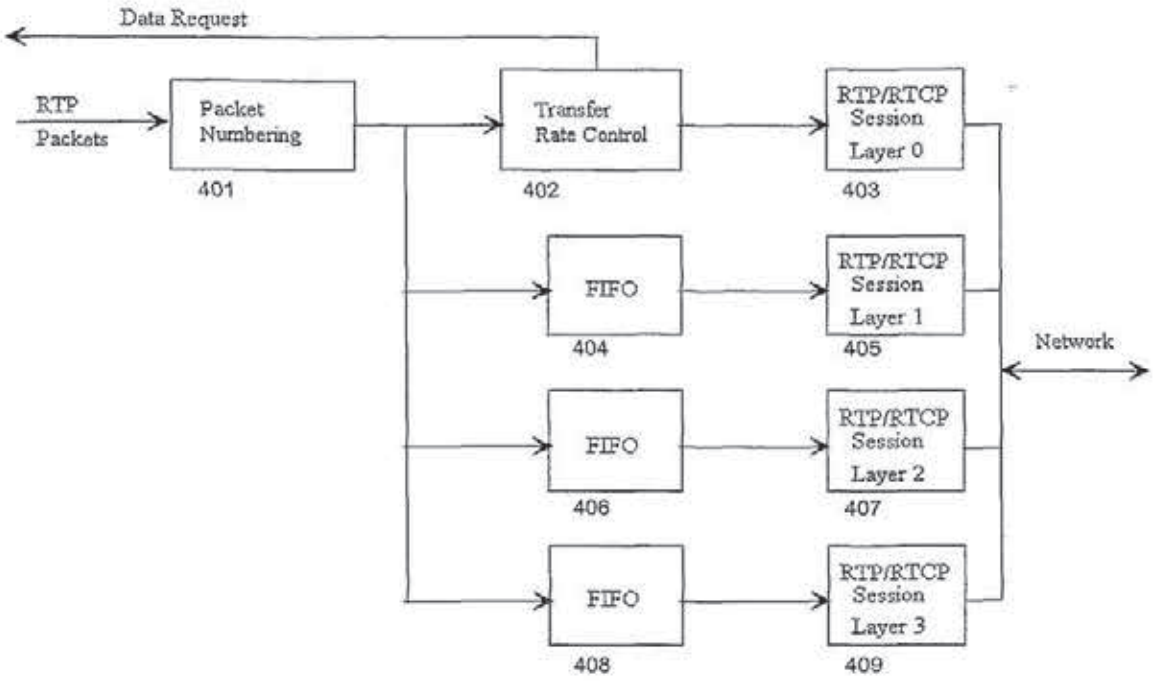


Figure 4

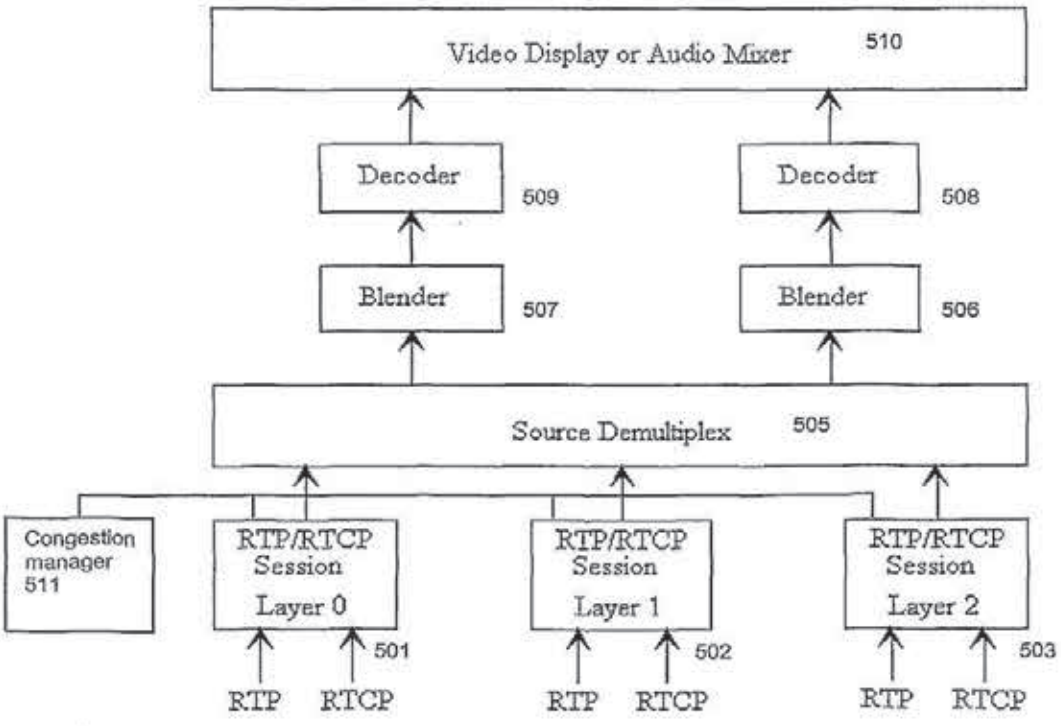


Figure 5

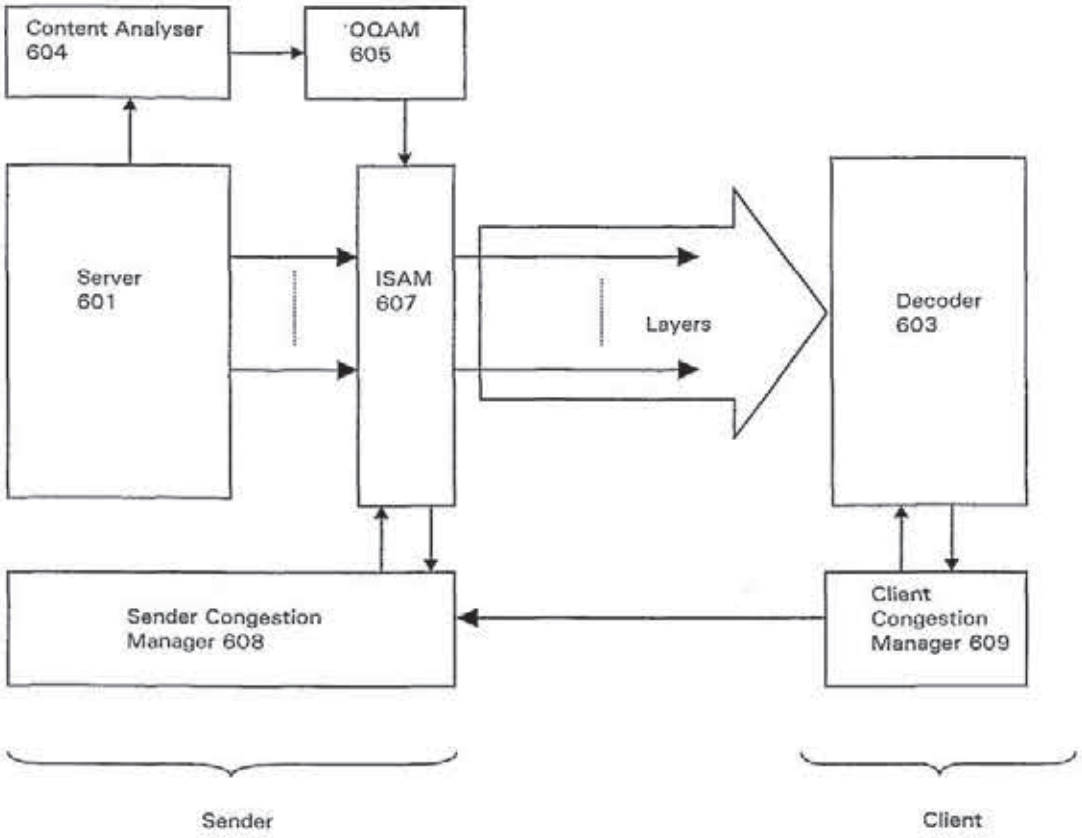

Content Analyser
604
OQAM
605
Server
601
ISAM
607
Layers
Decoder
603
Sender Congestion
Manager 608
Client
Congestion
Manager 609
Sender
Client

Figure 6

INTERNATIONAL SEARCH REPORT

tional Application No
PCT/GB 02/03307

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N5/00 H04N7/24

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 H04N H04L G06F

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, PAJ, IBM-TDB

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6 014 694 A (AHARONI AMIR ET AL) 11 January 2000 (2000-01-11)	1-10,14
Y	abstract column 2, line 29 - line 65 column 3, line 62 -column 4, line 34 column 8, line 54 -column 9, line 35 column 10, line 34 - line 49 column 14, line 34 -column 15, line 6	11-13
X	ANONYMOUS: "Method to Deliver Scalable Video across a Distributed Computer System" IBM TECHNICAL DISCLOSURE BULLETIN, IBM CORP. NEW YORK, US, vol. 5, no. 37, 1 May 1994 (1994-05-01), pages 251-256, XP002079392 ISSN: 0018-8689	1-10,14
A	the whole document	11-13

-/--

[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 September 2002	23/09/2002

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	Fantini, F

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

nal Application No
PCT/GB 02/03307

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6 038 000 A (HURST JR ROBERT NORMAN) 14 March 2000 (2000-03-14)	11-13
A	column 2, line 52 -column 3, line 22 column 4, line 20 - line 49 column 6, line 9 - line 45 column 7, line 1 - line 24 column 13, line 46 -column 14, line 18 column 20, line 13 - line 30	1,14
A	EP 1 079 631 A (SONY UK LTD) 28 February 2001 (2001-02-28) page 3, line 1 - line 4	1-14
A	US 5 990 945 A (DENTENEER THEODORUS J J ET AL) 23 November 1999 (1999-11-23) abstract column 3, line 13 - line 35 column 4, line 42 - line 44 column 5, line 37 - line 39	11-13

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

ıtional Application No
PCT/GB 02/03307

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 6014694	A	11-01-2000	AU	7928998 A	19-01-1999
			EP	0992160 A1	12-04-2000
			WO	9900984 A1	07-01-1999
			JP	2002511216 T	09-04-2002
US 6038000	A	14-03-2000	US	6137834 A	24-10-2000
			EP	0954924 A1	10-11-1999
			JP	2001509354 T	10-07-2001
			WO	9832281 A1	23-07-1998
			EP	0901718 A1	17-03-1999
			JP	2002503402 T	29-01-2002
			KR	2000016034 A	25-03-2000
			WO	9745965 A1	04-12-1997
EP 1079631	A	28-02-2001	GB	2353653 A	28-02-2001
			EP	1079631 A1	28-02-2001
			JP	2001119305 A	27-04-2001
US 5990945	A	23-11-1999	CN	1199532 A	18-11-1998
			EP	0860087 A1	26-08-1998
			WO	9809442 A1	05-03-1998
			JP	11514825 T	14-12-1999

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 January 2003 (09.01.2003)

PCT

(10) International Publication Number
WO 03/003760 A2

(51) International Patent Classification[7]: H04Q 3/00

(21) International Application Number: PCT/US02/16590

(22) International Filing Date: 25 June 2002 (25.06.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/893,364 26 June 2001 (26.06.2001) US

(71) Applicant: **BURST.COM** [US/US]; 613 Fourth Street, Suite 201, Santa Rosa, CA 95404 (US).

(72) Inventor: **ALLEN, Arthur, Douglas**; 1322 Isabelle Avenue, Mountain View, CA 94040 (US).

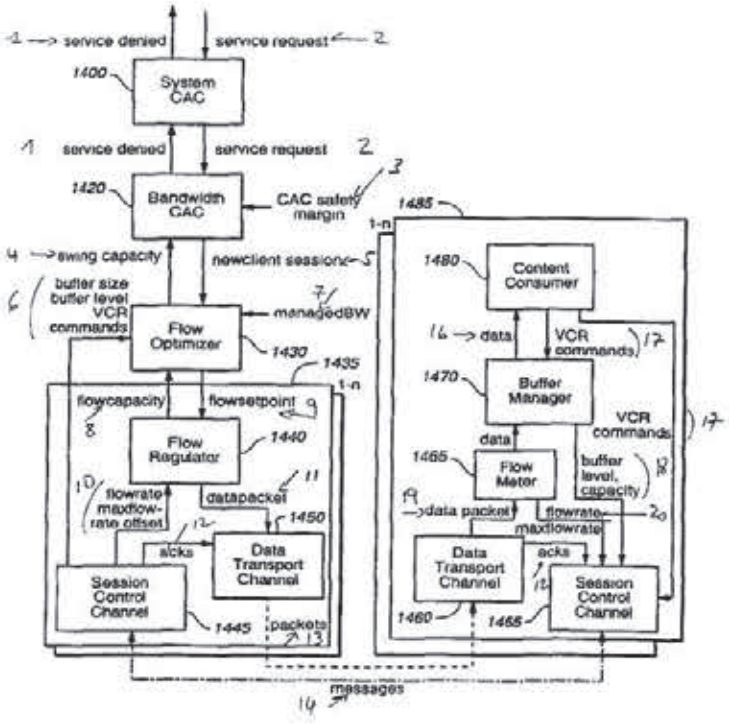
(74) Agents: **SWISS, Gerald, F.** et al.; BURNS, DOANE, SWECKER & MATHIS, L.L.P., P.O. BOX 1404, Alexandria, VA 22313-1404 (US).

(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

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.

(54) Title: METHOD FOR CONNECTION ACCEPTANCE CONTROL AND RAPID DETERMINATION OF OPTIMAL MULTI-MEDIA CONTENT DELIVERY OVER NETWORKS



(57) Abstract: A method of bandwidth allocation for delivery of stored digital content from at least one server device to at least one client device by way of a network is disclosed. The method begins by prescribing a control variable which represents a target flow rate from the server device to each client device. Next, time-varying constraints on the flow rate of the content are determined. A cost function of the control variable for each client is determined. The cost function corresponds to a maximized value of the control variable. Finally, bandwidth is prescribed to each client based upon the value of the control variable maximized by the cost function. In this respect, the method achieves optimal allocation of bandwidth between the server and the respective clients.

WO 03/003760 A2

METHOD FOR CONNECTION ACCEPTANCE CONTROL AND RAPID DETERMINATION OF OPTIMAL MULTI-MEDIA CONTENT DELIVERY OVER NETWORKS

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] This invention relates to the field of delivery of multimedia content over a variety of networks. More specifically, it pertains to multimedia servers which service many clients simultaneously for the delivery of multimedia content which is used and played back at each client. It addresses methods for determining optimal delivery rates to each client and methods for determining whether new clients may be accepted without diminishing the quality of service to existing clients.

Status of the Prior Art

[0002] In the history of multimedia program delivery, some in the industry have long advocated the use of large client-side buffers and faster-than-real-time content delivery over a network as offering the best of all worlds: a jitter-free viewing experience and a cost-effective utilization of the network resources at hand. Few systems, however, go very far in addressing how to schedule clients or a method for accepting new clients. Real-time systems, often known as streaming systems, can schedule new clients in a very simple manner -- if sufficient bandwidth remains for the added real-time stream, then the client may be accepted. However, such systems do not maximize the number of simultaneous clients. On the other hand, faster than real-time delivery, sometimes known as store-and-forward systems, opens up the possibility for more flexible scheduling procedures to control and optimize the number of simultaneous clients while ensuring a high level of quality of service.

[0003] The methods for such call acceptance and flow modulation that have been proposed in the prior art have been largely ad-hoc and also incomplete. These have been ad-hoc in the sense that there has been no guiding rationale for their selection from among many possible and potentially superior alternatives. The methods have also been incomplete insofar as they did not address the question of whether any given incoming request for service should be accepted or denied. Video-on-demand systems, or more generally, any system in which a multimedia server is designed to serve multiple clients over a network to deliver bounded content, can benefit from the use of such flow modulation techniques and call acceptance procedures.

Optimal Content Flow Modulation

[0004] One time-honored way of designing methods of the class required here is to re-cast the problem to be solved as an optimization problem, in which one seeks to maximize a designated value function moment-by-moment, subject to a set of real-world operational constraints which will typically vary over time. Accordingly, given a set of clients and associated sessions, an optimal delivery procedure continuously establishes content flow rates from the content server to each of its clients so as to maximize aggregate value according to the governing value function.

[0005] This approach holds several advantages: 1) optimization problems are well understood, and are tractable by a large and diverse collection of computational methods; 2) if it exists, a global solution that is obtained is arguably optimal by construction, and thus superior or equal to all other.

[0006] The present invention teaches the method of optimizing two particular value functions:

1) total data delivered (maximize throughput).

2) total delivery charges (maximize charges).

[0007] The first value function does not distinguish one customer from another and will deliver as much data as possible from server to clients irrespective of the characteristics of the latter. The second value function favors the service of high paying customers. It can easily be seen that the first function is a special case of the second one whereby all clients are charged equally.

[0008] As will be seen in this disclosure, optimizing for these functions and identifying the necessary constraints requires a new and unique perspective that is specifically designed for the multimedia environment. Moreover, the disclosed methods are specifically designed to account for and accommodate real-world scenarios of today's networks. Consequently many variations of the method are presented to accommodate various scenarios.

SUMMARY OF THE INVENTION

Call/Connection Acceptance Control (CAC)

[0009] A CAC procedure is responsible for deciding whether a candidate for service can be accommodated without jeopardizing sessions already in progress at the present time or at some time in the future; failing that it must decide whether a service request should be queued for a time or rejected.

Flow Modulation

[0010] Flow modulation methods are those portions of the system which manage the communication and data flow between the server and the clients. Collectively, these methods provide the multimedia data to the client and provide the server with the information about the state of the transmission, playback, user status and network status. These parameters are further used by the present invention in the CAC procedures. In fact, as will be shown, the proposed CAC procedures are tightly integrated with the flow modulation methods.

Adaptation to Variations in Network Capacity

[0011] Operational constraints may change over time. For instance, one might elect to vary the total bandwidth available for multimedia content delivery according to the time of day. Alternatively, exogenous data flows on the network may cause unexpected disturbances by usurping available bandwidth and impeding the delivery of data along established session channels. The content delivery strategy of the present invention includes the ability to adapt to scheduled as well as unexpected disturbances so as to minimize unwanted disruptions of services.

Burst Transmissions Provide the Opportunity to Adapt

[0012] The present invention, due to it's faster-than-realtime transmissions (also know as burst transmissions), which are realized by use of high-bandwidth networks and large client cache or intermediate storage, provides an opportunity to adapt to changing network conditions. In contrast real-time (streaming) systems are essentially designed for worst-case scenarios: each client must be assumed to constantly use the complete real-time playback bandwidth. Such a system is unable to adapt to any derivation from this scenario. For example, take the simple case where the total server bandwidth is 100% utilized by all clients playing back the streaming video. Should any network condition change, such as a temporary

decrease in available bandwidth over the network, then one or more clients' playback is interrupted, and the system can not recover from such a condition until the bandwidth is regained. Even worse if a single client presses pause either that unused bandwidth must remain reserved and no more clients can be accepted, or that paused client is pushed out in order to service the new client. In essence little or no CAC procedure may be implemented.

[0013] In contrast the present invention burst transmits portions of a program and immediately 'gets ahead of itself', thus allowing headroom for a myriad of methods to intelligently handle new clients, client interactivity and possible network fluctuations.

[0014] Methods are taught for optimally determining the flow rate to each client. Methods are also taught for accepting or rejecting new clients; these call-acceptance methods are tightly coupled with said flow rate modulation methods. A series of constraint expressions are presented which govern the methods for determining the flow rates and acceptance of new clients. Linear programming techniques are used to optimally solve these expressions. Various embodiments are presented including scenarios for multiple-rate tariffs, and time-of-day bandwidth variations.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] These as well as other features of the present invention will become more apparent upon reference to the drawings wherein:

Figure 1 depicts the flow of control and/or data between the different stations of a content delivery session;

Figure 2 illustrates the Entity Data Model;

Figure 3 geometrically illustrates the linear programming problem statement;

Figure 4a geometrically illustrates an expansion of the flow optimization problem statement;

Figure 4b geometrically illustrates the method for rapid determination of flow rates to maximize flow;

Figure 5 illustrates a method for implementing flow modulation to maximize flow;

Figure 6 illustrates a method for implementing flow modulation for maximized charges;

Figure 7 illustrates typical content flow;

Figure 8 illustrates typical server swing capacity;

Figure 9 illustrates a method for call-acceptance and control (CAC);

Figure 10 illustrates planned constraints on maximum flow;

Figure 11 illustrates a method for call-acceptance and control (CAC) with scheduled flow changes;

Figure 12 illustrates stratification of services;

Figure 13 illustrates a method for call-acceptance and control (CAC) for maximal charge;

Figure 14 is a system block diagram for a system implementing the present optimization of client bandwidth;

Figure 15 is a flowchart showing an algebraic method for calculating the optimization of client bandwidth;

Figure 16 is a flowchart showing an algebraic method for calculating the optimization of client costs; and

Figure 17 illustrates optimized viewing of a live event.

DETAILED DESCRIPTION OF THE INVENTION

Data & Control Flows (Figure 1)

[0016] Figure 1 depicts the flow of control and/or data between the different stations of a content delivery session. As shown a client attempts a connection 100 and manifests itself to the Content Selection subsystem by means of a low bandwidth control channel (not shown). Next the client is authenticated and a selection is made 110, typically with the aid of browser software. If the client is not authenticated, it is dismissed from the system 120. If the client has been authenticated and a program selected for viewing then the rate of service is set at this point 130, perhaps according to the selection that was made, or some contractual stipulation. The client is now placed on the service queue of the CAC subsystem 140. A client that is made to wait too long will eventually balk 150. Assuming this does not occur, the CAC subsystem 140 will eventually allocate a channel to the client and open a session 160. Control now devolves upon the Content Flow Modulator (not shown) which starts the flow of content from server to client 170. Subsequent capacity changes, whether predictable or not, may force an abrupt termination of a session in progress 180. Otherwise the session runs to completion 190.

Entity Data Model (Figure 2, listing, table)

[0017] The entities entering into our discussion are depicted in Figure 2. Client 200 maintains certain data associated with this entity; as shown but not labeled, which includes without limitation, status, id and costOfService. The other entities also each include unlabeled but depicted data. The diagram further depicts the relationship between each entity. As shown, client 200 is assigned a session 240. Client 200 employs a channel 210. Client 200 selects contentSelection 230. Session 240 delivers content through channel 210. Server 220 modulates channel 210. Server 200 contains contentSelection 210. Server 220 accepts, defers or denies client 200 and contentSelection 230 is associated with session 240.

[0018] Furthermore Figure 2 depicts the various one-to-many relationships. Each client 200 employs one channel 210. Client 200 may or may not receive one of channel 210, as depicted by the 0/1 notation. Similarly, client 200 may or may not receive a session 240. However, whenever client 200 does receive a session 240, it will always receive a channel 210 since channel 210 and session 240 are allocated as a pair. One or more (N) of client 200 may select one of contentSelection 230. And server 220 contains one or more (N) of contentSelection 230. Each one of contentSelection 230 is associated with 0-N of session 240. Each session 240 delivers content through one of channel 210. And server 220 modulates one or more (N) of channel 210.

[0019] A more detailed list of each entity of Fgure 2, and each one's associated description, data elements and function calls is listed below. This listing closely resembles that of object-oriented programming. As such, 'methods' represent the ability to obtain or modify data, while 'attributes' represent data which is directly associated with that particular entity. The listing also includes information relating to one embodiment wherein software programming specifics are disclosed, such as

a variable type (double, int and so forth) and more. The present invention is not limited to such an embodiment and other implementations are possible without deviating from the scope and intent of the present invention. The listing, however detailed, is merely illustrative of the data and functions which are used in the equations and methods described herein.

[0020] Consequently, data and functions from this listing, associated with the various entities, will be used in forthcoming equations, flowcharts and methods. The reader is directed to this listing as reference when reading such equations and examining such drawings.

--- start of entity data model detailed listing ---

Model: Untitled 1 (public)

Contains:

client, session, channel, server, contentSelection.

Component: client (public Class/Interface)

Comment:

A client entity stands for a client presently requesting or receiving service.

Methods:

public static **lookup (id: in int) : client**

public **GetId () : const int&**

public **SetId (val : in int&)**

public **GetCostOfService () : const double&**

public **SetCostOfService (val : in double&)**

Attributes:

private **status: client < int >**

Specifies whether or not a client has been allocated a channel and session.

private **id: int**

Integer-valued identifier that is unique to the client (primary key). Can be obtained from a monotonically increasing counter.

private **costOfService: double**

Dollar charge per Mbyte. This value is the same for all customers under flow optimization. Under cost/charge optimization may be an integer value reflective of the rank; the higher the rank the higher the charge.

Has:

public **selected: contentSelection**

public **assigned a: session**

public **employs: channel**

Component: session (public Class/Interface)

Comment:

A session entity holds various state information about the service being received by an associated customer.

public **GetCurrentPosition () : const double&**

public **SetCurrentPosition (val : in double&)**

public **GetPayloadToGo () : const double&**

public **SetPayloadToGo (val : in double&)**

public **GetStatus () : const int&**

public **SetStatus (val : in int&)**

public **GetMinimumFlowRate () : const double&**

public **SetMinimumFlowRate (val : in double&)**

public **GetFlowRateRange () : const double&**

public **SetFlowRateRange (val : in double&)**

public **GetMaxFlowRate () : const double&**

public **SetMaxFlowRate (val : in double&)**

Attributes:

private **playTimeToGo: double**

Indicates the minutes remaining in the viewing experience. Initialized to contentSelectiont.playTime (see below).

private **currentPosition: double**

Pointer into media content from which content is being delivered.

private **payloadToGo: double**

The amount of media content (in Mbytes) as yet undelivered by the server. Does not include any content presently stored in the client-side buffer.

private **status: int**

Indicates whether sesion is active or paused.

private **minimumFlowRate: double**

This is the minimum flow from server to client required to ensure uninterrupted service over the remaining playTime. Has a value of zero if playloadToGo is zero. Given by (payloadToGo * 8)/(playTimeToGo *60)

private **flowRateRange: double**

Specifies the effective range over which the channel content flow serving a session is constained without consideration for interactions with other flows. Equals maxFlowRate - minimumFlowRate

private **maxFlowRate: double**

Effective maximum bound on flow as expressed in formula (8) which must be re-evaluated periodically.

Has:

public **delivers content through: channel**

Component: channel (public Class/Interface)

Comment:

A channel represents the network resources from server to client associated with an ongoing session, encompassing the client-side buffer if any, and its level.

public **GetBufferLevel () : const double&**

public **SetBufferLevel (val : in double&)**

public **GetFlowRate () : const double&**

public **SetFlowRate (val : in double&)**

public **GetMaxFlowRate () : const double&**

public **SetMaxFlowRate (val : in double&)**

Attributes:

private **bufferSize: double**

Capacity of the client-side buffer (or equivalent)

private **bufferLevel: double**

Current buffer level in MBytes of stored content.

private **flowRate: double**

Flow rate through channel specified by the relevant optimizing flow modulator.

private **maxFlowRate: double**

This value represents the maximum possible flow rate from the server to an individual client over its "channel". This value reflects restrictions on flow that pertain to an individual client. It may be determined by factors such as the bandwidth of client's link to the network, or a limit imposed administratively to ensure balanced network utilization.

Component: server (public Class/Interface)

Comment:

Entity representing the media server and its CAC and flow modulation activities.

public **GetFlowRate () : const double&**

public **SetFlowRate (val : in double&)**

public **GetMaxMinFlowRate[] () : const double&**

public **SetMaxMinFlowRate[] (val : in double&)**

Attributes:

private **maxFlowRate: double**

Maximum possible content flow that is allocated to the server by the network.

private **flowRate: double**

Aggregate content flow rate, summed over all sessions and their associated channels.

private **cac_flowSafetyMargin: double**

Tunable safety margin used by the CAC algorithm to protect sessions-in-progress from being affected by changes in available network bandwidth.

private **maxMinFlowRate[]: double**

Applies when N rate tariffs exist. This array holds the maximum floor level for each category of service. The value for the costliest category N is stored in maxMinFlowRate[N-1], and for the least costliest in maxMinFlowRate[0]. It is the relative magnitude of these ascending values that matters, not their absolute value. Thus the actual maximum floor flow rate for category k is given by server.maxFlowRate * (server.maxMinFlowRate[k-1] /server.maxMinFlowRate[N-1]). Similarly, the maximum floor flow rate for category N is server.maxFlowRate.

Has:

public **contains: contentSelection**

public **modulates: channel**

Component: contentSelection (public Class/Interface)

Comment:

Entity represents a video/sound clip or other bounded unit of content. A continuous data feed does not qualify.

Attributes:

private **averagePlayRate: double**

The average rate at which media content is consumed by the client, as computed by dividing the (payload * 8) by the (playTime * 60)

private **playTime: double**

Duration of play of the media content in minutes.

private **payLoad: double**

total size of the content in Mbytes.

Has:

public **is associated with: session**

--- end of entity data model detailed listing ---

The following table summarizes the highlights of the previous detailed description of each entity in Figure 2.

TABLE 1

Entity	**Description**
client 200	Each client is denoted by an associated unique integer $index_{Id}$. The set of active clients is denoted by $\mathbf{S}_{ActiveClients}$. The set of deferred clients is denoted by $\mathbf{S}_{QdClients}$. Incoming clients are expected to select the content they wish to view prior to being queued for dispatch by the CAC sub-system, which requires knowledge of the client's bandwidth requirements, duration of play, and cost of service, all of which may vary according to the selection.

server 220	Servers sit astride a network and can deliver media content through the network to their clients up to a designated maximum flow rate. The server is responsible for accepting or rejecting clients, launching sessions and associated channels for the former, and modulating content flows over all channels in an optimal manner.
channel 210	A channel represents the data path between the server and the client. The channel buffer is typically located near or within the clients viewing station. The flow of content through the channel is set by the flow modulator sub-system.
contentSelection 230	A server will typically act as a repository for media content, which it can deliver to clients upon demand. For our purposes, media content is characterized by its payload and the play duration, which together imply the ***averagePlayRate*** = (payload*8)/(playTime *60). The averagePlayRate is none other than the streaming rate imposed by real-time just-in-time streaming algorithms.
session 240	Every session represents an instance of media content delivery to an associated client over a designated channel. The ***playTimeToGo*** indicates the time remaining before the content is fully played out to the client. The ***payloadToGo*** is the amount of content data as yet undelivered to the channel. A session terminates when this value reaches zero, at which time ***playTimeToGo*** may still be large, according to the capacity, the level of the channel buffer, and the media play rate.

Constraints On Content Flow

[0021] Before referring to more Figures, it is imperative to establish some formulas and problem statements which are used in the methods which follow.

[0022] The flow of content between entities is subject to the following constraints at all times. Buffer levels are always expressed in Mbytes and data rates in Mbits/sec.

(1)$\Sigma_{\text{i } \epsilon \text{ SactiveClients}}$ (client.lookup(i).channel.flowRate) < = server.maxFlowRate
The sum of all channel flows cannot exceed the imposed maximum throughput capacity of the server.

(2)client.lookup(i).channel.flowRate < = client.lookup(i).channel.maxFlowRate for all $\mathbf{i \in \Sigma_{activeClients}}$

The data path from server to client is subject to its own constriction.

(3)client.lookup(i).channel.flowRate < =(client.lookup(i).channel.bufferSize - client.lookup(i).channel.bufferLevel)* 8/60 + client.lookup(i).session.mediaContent.averagePlayRate, for all $\mathbf{i \in \Sigma_{activeClients,}}$

The channel buffer is never allowed to overflow.

(4)client.lookup(i).channel.flowRate < = client.lookup(i).session.payloadToGo * 8/60

for all $i \in \Sigma_{activeClients}$.

Content that does not exist cannot be delivered. (Constraint 1 will ordinarily prevail except at the very end of a session.)

[0023] The constraints listed above are straightforward applications of common sense in relation to the flow of data through constricted channels, out of finite data sources, and into and out of bounded buffers. By contrast, the following constraint, which imposes a minimum channel flow rate instead of a maximum, is less obvious. The minimum value, termed the ***minFlowRate*** is set to the flow rate which, if sustained over the balance of the play time to go (***playTimeToGo***), ensures that all required content will be available when needed -- and no sooner -- until all content is played out. This floor value can be calculated for $i \in \Sigma_{activeClients}$ by the formula:

(5)client.lookup(i).session minFlowRate =(client.lookup(i).session.payloadToGo * 8)/ (client.lookup(i).session.playTimeToGo * 60)

Accordingly, the flow rate on the right-hand-side is termed the just-in-time (JIT) flow rate (f^{JIT}).

Thus:

(6)client.lookup(i).channel.flowRate >= client.lookup(i).session.minFlowRate
for all $i \in \Sigma_{activeClients}$

[0024] The variable constraint bounds (i.e. the values to the right of the inequality symbol) of equations 1 – 4 and 6 are re-evaluated on a periodic basis (e.g. once per second) prior to the execution of the CAC procedure and optimizer. In

particular, the ***minFlowRate*** value starts out at the beginning of a session equal to the streaming rate. By construction the ***minFlowRate*** rate never exceeds this initial value so long as constraint 6 is honored. In fact, constraint 5 implies that the ***minflowRate*** value must be a diminishing function of time that may hold its value for a time but never rises. As seen from equation 6, the actual data rate of the channel, ***flowRate***, is always greater than or equal to the ***minFlowRate***. By design, and virtue of the fact the present invention uses faster-than-realtime transmissions, the system quickly gets ahead of itself and ensures that after initial conditions, the ***minFlowRate*** is always equal to or less than the real-time rate and that it continues to decrease. As we shall see, the CAC procedure exploits this monotonic characteristic of the minimum flow rate over time.

[0025] Constraints 2, 3 and 4 are of like kind, each specifying an upper bound on individual channel flows. Whereas the bound for constraint 2 is typically a constant, the bounds on 3 and 4 will vary over time. Nevertheless, only one of the three bounds is effective at any given time, namely the one with the smallest bound value, given by:

(7) client.lookup(i).session.maxFlowRate = minimum of

1) client.lookup(i).channel.maxFlowRate,

2) (client.lookup(i).channel.bufferSize - client.lookup(i).channel.bufferLevel)*

8/60 + client.lookup(i).session.mediaContent.averagePlayRate,

3) client.lookup(i).session.payloadToGo * 8/60

Consequently, formulas 2, 3, and 4 can be consolidated into a single constraint, the bound for which is computed at every scan to be the smallest bound of associated constraints 2, 3 and 4.

(8) client.lookup(i).channel.flowRate < = client.lookup(i).session.maxFlowRate,

whereby for all $i \in \Sigma_{activeClients}$, maxflowRate is given by equation (7).

At any one time, individual channel flows are constrained over a range, as follows:

(9) client.lookup(i).session.flowRateRange =
client.lookup(i).session.maxFlowRate - client.lookup(i).session.minimumFlowRate

Value Functions

[0026] The value functions introduced previously can be expressed mathematically as linear functions of channel flows, as follows:

Optimizing Throughput (Maximal Flow)

(10) value $= \Sigma_{i \in SactiveClients}$ client.lookup(i).channel.flowRate)

Optimizing Charges (Maximal Charges)

(11) value $= \Sigma_{i \in SactiveClients}$ (client.lookup(i).channel.flowRate * client.lookup(i).costOfService)

Optimization Problem Statement (Figure 3)

[0027] The optimization problem, which in one embodiment is strictly throughput and in another case is charge, can be stated simply as follows:

Find values for

client.lookup(i).channel.flowRate for all $i \in \Sigma_{activeClients}$

constrained by inequalities 1 through 5, such that the value obtained by evaluating expression 10 or 11 assumes a maximum.

[0028] Both of these problem formulations are examples of Linear Programming for which a number of well-known and generally effective computational solutions exist. In linear programming one seeks to optimize a linear cost function of variable **x**

(12) $c^*x = c_1^*x_1 + \ldots + c_n^*x_n$

subject to a set of linear inequality constraints

(13) $A^*x <= b$

where $x^T = (x_1,\ldots,x_n)$, $c = (c_1,\ldots,c_n)$ are the state variable & cost vectors, A is an n-by-m matrix, $b^T = (b_i,\ldots b_m)$ is the constraint vector, and the operator '*' stands for matrix or scalar multiplication.

[0029] Figure 3 is introduced as illustrative of the problem statement and the general methods of the prior art, and is not incorporated as an element of the invention.

[0030] The linear programming problem as well as its solution can best be understood with the aid of geometry. Figure 3, depicting a 2-dimensional Cartesian problem space, inequality constraints (13) define a convex hull H 310 over which a search for an optimum value of $\mathbf{x} = (x_1, x_2)$ is permitted to range. The cost vector **c** 350 defines an infinite family of equal cost lines (hyper-planes) which lie orthogonal to **c**. Three examples of such lines are shown in L_1 360, L_2 365, and L_3 370., each of progressively higher value. The supreme value of the cost function is obtained by sliding along **c** 350 till one can go no further, in this instance toward vertex V_4 340 of hull H 310. Many well-known methods (e.g. the Simplex Method) work roughly in this fashion, exploiting the fact that at least one optimum point must be at a vertex. In particular, the Simplex method algorithm begins by finding a vertex (e.g. V_2 320), and then moves along a sequence of vertices (e.g. V_3 330, V_4 340) improving itself each time until no further improvement is possible & the summit is reached.

[0031] Let us suppose instead that V_3 330 were placed along L_3 370 along with V_4 340. According to prior art methods, V_3 330 and V_4 340 are the two possible solutions, but the equally valuable points in between them are not. As we shall soon see, the problem of throughput optimization (6) falls in this category.

[0032] While vertex V_1 300 does not factor into this description, it is depicted in Figure 3 for completeness.

Flow Modulation

Methods for Maximal Flow

The following sections detail two embodiments to optimize total data flow.

Overview (Figure 4-a)

[0033] Figure 4-a depicts a scenario involving two flows. The convex hull is in this instance bounded by line segments L1, L2, L3, L4 and L5. L6 is a boundary used in a different embodiment, however the present embodiment uses L5 and not L6. Flow f_2 can range over the interval separating line segments L1 from L3, namely f_2^{MIN} and f_2^{MAX}; the range is depicted as f_2^{RANGE}. Flow f_1 can range over the interval between lines L2 and L4, namely f_1^{MIN} and f_1^{MAX}, and depicted as f_1^{RANGE}. The sum of flows f_1 and f_2 is constrained to lie inside of line segment L5 which, by construction, is always orthogonal to the cost vector $\mathbf{C_f}$. Cost vector $\mathbf{C_c}$ is also illustrated but used in a distinct embodiment. In the present embodiment only $\mathbf{C_f}$ is used. In the illustrated example of the present embodiment, the constraint on total flow is set to 5, and is therefore low enough to cause L5 to intersect L3 and L4. This would not have been true had the value chosen had been 10 instead of 5. With L5 safely out of contention, the convex hull would instead be a simple rectangle bounded by L1 through L4, thereby permitting both flows to assume their respective maxima without interference. In practice operational constraints exist intrinsically or are imposed from the outside so as to ensure cost effective sharing of potentially costly network resources.

[0034] Supposing Figure 4 to be correct, the well-known methods would select vertex $V_{3\text{-}5}$, which lies at the intersection of L3 and L5, or $V_{4\text{-}5}$, which lies at the intersection of L4 and L5. These solutions, though optimal, are undesirable for the present invention as they fail to spread available bandwidth over all channels as fairly as would a centrally located interior point of L5. For this reason, two alternative optimization methods are taught, which are adapted to the particular needs of this problem and ensure a fairer allocation of constrained bandwidth among all channels.

Iterative Procedure (Figure 5)

[0035] In order to optimize use of all available bandwidth, the following general method is used, with the details illustrated in Figure 5. This method is a solution for the problem illustrated in Figure 4-a, which geometrically illustrates the optimization problem in the limited case of two flows, f1 and f2. The following description expands the problem to an arbitrary number of clients (and therefore flows) and presents a method for solving this optimization problem.

[0036] Referring to Figure 5, in step 500, values are calculated for the ***session maxFlowRate*** and ***session.minFlowRate*** for each client as per the minimum and maximum constraint bound expressions in equations 6 and 8, respectively. This step correlates to the determination of f_1min, f_1max, f_2min and f_2max from Figure 4-a.

[0037] The difference between these two yields the ***session.flowRateRange*** of each client. Thus:

$$session.flowRateRange = session\ maxFlowRate - session.minimumFlowRate$$

[0038] In step 505, the active clients are sorted in an ascending fashion based upon their **session.*flowRateRange***. As will be shown this critical step facilitates allocation of the remaining server bandwidth as evenly as possible among all active channels, thus maximizing the number of channels that benefit by use of the total server bandwidth. An arbitrary assignment of remaining bandwidth is likely to saturate the server before all channels have been assigned extra bandwidth, thereby favoring certain channels on an ad-hoc basis. This step correlates to keeping the solution in bounds of the space delinated by line segments L1-L5 of Figure 4-a.

[0039] In step 510, each client's channel flow rate is set to the ***session mimimumFlowRate***.

[0040] By doing so, it is ensured that the minimum flow constraint is met for each session and that the minimum flow rate is a non-increasing function of time, which is critical to the proper functioning of the CAC procedure. This portion of the process also ensures that the solution, starting from the vertex in Figure 4-a as defined by the intersection of L1 and L2, moves generally in the direction of vector c_f . All clients are also marked as unprocessed.

[0041] In the next step, 520, ***server.flowRate*** is set to the sum of each active client's **session.flowRate.**

[0042] Next, the following is iterated over all clients in sorted sequence (during any given iteration the selected client is given by its ***id***) by performing steps 530 through 570. In step 530 evaluating the following expressions test for possible server saturation:

delta = (server.maxFlowRate - server.flowRate) / (qty of un-processed clients)
range = client.lookup(***id***).session.maxFlowRate -
client.lookup(***id***).session.flowRate

[0043] If ***range*** is greater then ***delta***, this implies that the server can be saturated in this iteration by allocating ***delta*** to all unprocessed clients (step 540).

[0044] On the other hand, the 'no' path for step 530 implies that the server is not saturated and that the present client (given by ***id***) will saturate first. Accordingly, in 550 the ***delta*** variable is set as follows:

delta = range

[0045] To again correlate this process back to the geometry of Figure 4-a, server saturation is indicated when the solution which is being sought in the direction of vector c_f goes beyond line segment L5.

[0046] Next, the flow rate is incremented for all unprocessed clients by ***delta***, causing client ***id*** to saturate.

[0047] In step 560 the server flow rate is adjusted accordingly:

server.flowRate = server.flowRate + delta * (qty of unprocessed clients)

In step 570 the client given by ***id***, now saturated, is marked as processed.

Algebraic Procedure (Figures 4-b and 15) By Interpolation

[0048] Referring to Figure 4-b, the iterative method described in the previous section begins its search for a maximum from vertex V_{min} along the direction shown by line C_1. In contrast, the present method follows the diagonal line segment L_{diag} of the rectangle bounded by lines L1 through L4, starting at vertex V_{min} and ending at point $V_{3\text{-}4}$.

[0049] Provided it exists, the intersection between L_{diag} and L5, indicated by V_f, is optimal on the same basis as any other point lying along the intersection of L5 with the rectangle bounded by L1 through L4, as previously discussed.

[0050] Whenever L5 does not intersect this rectangle, the optimal solution is given by vertex V3-4 at which f1 and f2 both assume their respective maxima.

[0051] In the first instance, the coordinates for point V_f can be obtained by elementary vector geometry and algebraic manipulation, as follows:

First, we define a number of abbreviations,

f_k^{min} represents client.lookup(k).session.minimum flow rate

f_k^{range} represents client.lookup(k).session.flow rate range

f_{svf}^{max} represents server.maxflowrate

We seek scale factor x such that the vector sum of $(f_1^{min} \ldots f_n^{min}) + x\ (f_1^{range} \ldots f_n^{range})$ intersects an equi-cost hyperplace for capacity c. The point of intersection is obtained by solving equation:

$$(f_1^{min} + \ldots + f_n^{min}) + x(f_1^{range} + \ldots + f_n^{range}) = c$$

Solving for x, we obtain:

$$x = [f_{svr}^{max} - (f_1^{min} + \ldots + f_n^{min})]/(f_1^{range} + \ldots + f_n^{range})$$

where $f_{svr}^{max} - (f_1^{min} + \ldots + f_n^{min})$ represents the unused bandwidth beyond the minimum allocation for each session. For any given session, the optimal flow rate is given by:

$$f_i = f_i^{min} + x\, f_i^{range}, \text{ if } x < 1,$$

or

$$f_i = f_i^{min} + f_i^{range}, \text{ if } x > 1$$

[0052] This solution by interpolation is interesting by virtue of the great efficiency of the calculation, and the fact that no session saturates ($x >= 1$) unless all sessions saturate. In contrast, the previous method entails considerably more computational effort, and tends to saturate sessions with the lowest reserve capacity.

[0053] A block diagram for this algorithm is depicted in Figure 15. Block 1500 establishes the maxim and minima for each flow. For each flow the minimum is subtracted from the maximum to obtain the range. Available Bandwidth is calculated in block 1520, as the difference between the aggregate flow capacity and the sum of flow minima. Factor x is calculated next in block 1540, and tested against 1 in block 1650. Flow rates for each session are computed in block 1560 if x is less than 1, or 1570 in the contrary case.

A Method for Maximal Charge

[0054] The following sections detail one embodiment to optimize the total monetary charges within the system. The second method is algebraic in nature, and thus very efficient.

Overview (Figure 4-a)

[0055] Referring back to Figure 4-a, cost vector C_c lies orthogonal to line L6, which intersects the convex hull at the vertex formed by the intersection of lines L4 and L5, namely $V_{4\text{-}5}$. This cost vector, and the optimal point that it implies, favors flow f1 over flow f2 . In this example, as the cost of service for f1 equals 2, thus exceeding the cost of service of 1 set for f2 . As the number of flows grow to exceed the number of distinct categories of service (and associated costs of service) the unique optimal solution, depicted in Figure 4 for the case where every flow has a distinct cost of service, no longer applies. Once again a plurality of flows within a service category vie for bandwidth which a method should endeavor to distribute evenly. This method is derived from the previous one, and optimizes one cost category after another, starting with the most costly and ending with the least costly, or when all available capacity is allocated.

An Iterative Search Procedure (Figure 6)

[0056] Let the service categories be denoted by k= 1...N, where k also denotes the cost of service.

[0057] Let $C_1..C_N$ be the partition of $S_{activeClients}$ that places all clients with cost of service k in set C_k. Partition sets C_k can be ordered to form sequence SeqC = $C_N...C_1$.

[0058] Figure 6 depicts the method for implementing the method to maximize the cost of service (service charge).

[0059] This method is nearly identical to the iterative procedure used to maximize flow. The principle difference stems from the partitioning of clients according to their category (cost) of service: clients charged most are allocated bandwidth preferentially. This is accomplished by adding another level of iteration around the method of Figure 5. The inner iteration (steps 650 through 680) functions exactly as before, with the difference that its actions are limited to the clients belonging to the given service category ***k*** (i.e. C_k). This difference also holds true of step 640 which sorts category ***k*** clients according to their flow ranges prior to entry in the bandwidth-allocating inner loop. The outer loop proceeds down a sorted sequence of service categories SeqC (generated in step 630), starting with the category generating the greatest revenue to the service provider. Given a fairly static set of service categories, this sort need be performed only when the categories undergo change. Steps 670, 675 and 680 are identical to their counterparts in the method of Figure 5 (i.e. 570, 575 and 580).

[0060] The net effect of this method is preferential allocation of bandwidth according to category of service, and equitable treatment of clients within the same category of service.

An Iterative Algebraic Procedure (Figure 16)

[0061] The algebraic method used to maximize bandwidth (Figures 4b and 15) can be viewed as a special case of cost optimization in which all costs are equal. By construction, all sessions belonging to a given category, starting with the most expensive category, and ending with the least, and updating available bandwidth after each step in sequence, we obtain an iterative algebraic method that scales much better than the preceding method.

[0062] A block diagram for this algorithm is depicted in Figure 16. Block 1600 establishes the maxima and minima for each flow. For each flow, the minimum is subtracted from the maximum to obtain the range. In block 1610, every session flow is assigned the minimum allowed value, which value will apply by default should the iterative procedure run out of bandwidth before all categories can be considered. Block 1620 computes the initial value for available bandwidth, as the difference between the aggregate flow capacity and the sum of flow minima.

[0063] Flow values are obtained for each cost category in order of decreasing cost within blocks 1640 through 1675. Factor x is calculated next in block 1540, and tested against 1 in block 1650. Flow rates for each sessions of the category under consideration are computed in block 1660 if x is less than 1, or 1670 in the contrary case. In the former case there is no more bandwidth to allocated to the sessions in the category under consideration.

Call Acceptance Control (CAC)

CAC for Maximal Flow Overview (Figures 7-8)

[0064] The CAC procedure applicable to this flow optimization relies on the essential step of accepting a new client if and only if the added load induced thereby does not compromise service to existing clients or the new one. This critical step could not be accomplished without the close integration with previously- described flow-modulation methods of Figures 5, 6, 15 and 16.

[0065] According to the previous discussion, the minimum flow rate is the minimum sustained flow rate that guarantees that the associated viewer will not be subject to interruptions in service due to a shortfall of content from the media server. It follows that whenever data is being delivered at a rate in excess of the minimum flow rate, a downward adjustment toward the minimum level could be accommodated as needed to surrender bandwidth to any newcomer.

[0066] Figure 7 depicts content flow over a channel for the course of a typical session, and also how data is delivered under real-time streaming D. The amount of content delivered is the same in either case, but the manner of delivery differs considerably. A session is launched at time 0 as the network is lightly loaded, and the optimizer sets an accordingly high flow rate. Another client emerges at the end of interval 700, causing a downward adjustment to the flow rate over interval B, as available bandwidth is shared between two sessions. During both of these intervals the minimum flow rate 720 drops quickly, as data accumulates in the client's media buffers. At the end of interval B, a massive influx of clients necessitates that flow be dropped to the minimum flow rate, which now lies substantially below the streaming rate D and is held until all data is delivered at the end of interval C. Note that the minimum flow rate, shown as element 720, diminishes monotonically over time.

[0067] The ***server swing capacity*** is defined as the difference between the maximum capacity of the server and the total minimum flow rates for all active clients. Therefore:

(14) swingCapacity =

server.maxFlowRate - $\Sigma_{\text{i e SactiveClients}}$ (client.lookup(i).session.minFlowRate)

[0068] Given the monotonic decreasing nature of session minimum flow rates, server swing capacity can readily be seen to be a *monotonic increasing* function of time over the intervals separating client admissions, at which point it undergoes a drop as a new load is taken on. This all-important characteristic implies the following:

Any client admitted for service based on the present value of swing capacity is guaranteed to have sufficient bandwidth at its disposal over the entire future course of the session.

[0069] Figure 8 depicts the server swing capacity 800 over the course of the sessions illustrated in Figure 7. Swing capacity rises quickly over intervals A & B as data is delivered at high flow rates over the network. It holds steady over interval C when all channels flow at their minimum rate then jumps at the end of C before resuming its monotonic rise once again.

Procedure (Figure 9)

[0070] In this procedure, which executes on a periodic basis, queued clients awaiting bandwidth are scanned in FIFO order. For each one, the required bandwidth is computed as per the client's prior content selection. If the available swing capacity (reduced by a safety margin) exceeds the amount required, then the client is activated and swing capacity is adjusted accordingly. Otherwise, two possible cases are considered: 1) under the *FirstFit* embodiment, the procedure

continues scanning clients to the end of the queue, activating clients whose requirements can be met; 2) under the *FIFO* embodiment, the procedure ends with the first candidate client whose requirements cannot be met.

[0071] In step 900, available server swing capacity is evaluated according to the formula

swingCapacity =

server.maxFlowRate - $\Sigma_{i \in \text{SactiveClients}}$ (client.get(i).session.minimumFlowRate)

The bandwidth requirement for client *id* in Step 920 is obtained as follows:

required_bandwidth = client.lookup(*id*).contentSelection.averagePlayRate

The predicate evaluated in Step 940 is given by the expression:

(required_bandwidth < = swingCapacity - server.cac_flowSafetyMargin)

[0072] In step 950, client activation entails allocation of a session and a channel, and insertion in the set of active clients eligible for bandwidth allocation by the optimal flow modulator.

[0073] In step 960 the swing capacity is diminished by the amount reserved for the activated client:

swingCapacity = swingCapacity - required_bandwidth;

Responding to Variations in Network Capacity (Maximal Flow)

[0074] In the CAC procedure for maximal flow, a safety margin was introduced, namely *server.cac_flowSafetyMargin*, to provide the means for ensuring that the server's swing capacity will never fall below a minimal threshold value.

[0075] According to this procedure, the following inequality always holds true:
(15) swingCapacity > = server.cac_flowSafetyMargin

[0076] In the previous discussion, a server's swing capacity provided the basis for determining whether or not a prospective client should be allocated bandwidth.

[0077] By holding *server.cac_flowSafetyMargin* in reserve, the CAC algorithm forces delivery of content at faster than real-time rates among accepted clients, even under the heaviest load. The net effect is to apply upward pressure on client side buffer levels, thus promoting continuous accumulation of content and a jitter-free viewing experience once a session is under way.

[0078] In another embodiment, server swing capacity can also be interpreted as specifying the *maximum* amount by which the *server.maxFlowRate* constraint can be dropped without affecting service, should such an adjustment prove necessary due, for instance, to an influx of exogenous network traffic that diminishes the amount available for multi-media services. Parameter *server.cac_flowSafetyMargin* can thus be set so as to *guarantee a minimum capacity to tighten the constraint on maximum server flow* in response to unexpected load changes that affect the server's ability to service its existing clients as well as new ones.

Anticipating Scheduled Variations in Network Capacity (Maximal Flow)

Overview (Figure 10)

[0079] Figure 10 depicts how the constraint on maximum flow might be allowed to vary according to the time of day, day of the week, and so forth, in expectation of time-varying traffic flow levels extrapolated from past experience, traffic flow models, etc. Maximum flow rate 1000 can be seen to vary based upon the time of

day. In practice, defining the right-hand-side of inequality constraint 1 as a time-dependent expression can impose such time-varying capacities. According to the previous description, the optimizer, which executes on a periodic basis, will automatically seek new flow levels for every active session as the constraint varies. There is, however, no guarantee that an acceptable operating point will be found for all sessions (i.e. one that respects the minimal and maximum constraints on session channel flow). One such example is the case where the server is loaded to the limit and total capacity is curtailed in excess of the aforementioned safety margin. Should such a situation arise the only recourse may well be the termination of a number of established sessions (i.e. load shedding).

[0080] The goal is to eliminate service disruptions of this sort by allowing the CAC procedure to look ahead into the future, and accept new clients only if these can be accommodated without any compromise in service in the midst of *previously anticipated changes* in available network capacity. The following CAC procedure generalizes the previous one: before accepting a client, the test on swing capacity is repeated over a sequence of time segments that cover the proposed viewing period.

Definitions -

Let:

(16) t_end(i) = client.lookup(i).session.playTimeToGo + t_now

Let server.maxFlowRate(t) be server flow capacity as a function of time, as exemplified in Figure 10.

[0081] Let $Seq_T(t\_now)$ = advancing sequence of future times, lead by t_now, when server.maxFlowRate(t) undergoes a step change. For instance, at 9:15 in Figure 10 this sequence reads as follows: 9:15, 9:30, 11:30, 13:30, 6:30, 7:30.

[0082] The server swing capacity at a future time t is computed according to the capacity and worst-case client flows at time t.

(17) swingCapacity(t) = server.maxFlowRate(t) -

$$\Sigma_{t\_end(i) > t} \text{(client.lookup(i).session.minFlowRate)}$$

[0083] It is noteworthy that the worst-case client flows at time t are expressed in terms of the present minimum flow rates, which cannot increase over time, but might hold steady. Finally, a predicate is defined that tests whether a prospective customer will cause swing capacity to be exceeded at some time t, as follows:

```
(18)  boolean client_fits(i ,t) {

        if(client.lookup(i).contentSelection.averagePlayRate < =

            swingCapacity(t) - server.cac_flowSafetyMargin)

          return true;

      else return false;
        }
```

Procedure (Figure 11)

[0084] This procedure is an adaptation of the first, which has been extended to consider swing capacity at times in the future when capacity undergoes scheduled changes. Before accepting a client, its minimal bandwidth requirement (which by construction of the flow modulator will never increase over time) is checked against *projected swing capacity* at points in time when total available capacity undergoes scheduled step changes, provided these times fall within the proposed content viewing period. A candidate is activated only if all tests succeed.

[0085] Step 1100 builds a sequence of time values (SeqT) at which step capacity changes are scheduled to occur. The first element of this sequence is t_now, representing the present.

[0086] Beyond step 1100 the queue of waiting clients is scanned in FIFO order, yielding a candidate designated by ***id*** at each iteration.

[0087] The bandwidth requirement for client ***id*** in Step 1120 is obtained as follows:

required_bandwidth = client.lookup(***id***).contentSelection.averagePlayRate

[0088] The worst-case end time for content flow to ***id*** is obtained according to the content selected, as follows:

t_end = t_now + client.lookup(***id***).selected.playTime

[0089] Steps 1130 through 1150 are executed within an iteration for each time point t in SeqT falling between t_now and t_end. This iteration is ended in step 1130 if t exceeds the time window of interest, or in step 1150 if the supply of scheduled capacity changes is exhausted.

[0090] For each time value, step 1140 compares required bandwidth to projected swing capacity.

Projected swing capacity at time t is:

$$\text{swingCapacity}(t) = \text{server.maxFlowRate}(t) - \sum_{t\_end(i) > t} (\text{client.lookup}(i).\text{session.minimumFlowRate})$$

[0091] Note that only active clients whose t_end times occur after t are considered in the sum of minimum flow rates.

[0092] The predicate expression used in step 1140 at time t is thus:

(required_bandwidth < = swingCapacity(t) - server.cac_flowSafetyMargin)

Step 1160 performs the same actions as step 660 in the previous CAC flowchart

[0093] The first CAC process is a special case of the present one, in which the set of step change times to server.maxFlowRate is empty (i.e. server.maxFlowRate is constant), and $Seq_T(t\_now) == t\_now$.

[0094] In preparing $Seq_T(t\_now)$, one need only consider future times that will pass before the longest possible content is played out if started at t_now. In order to sidestep problems associated with rollover (at midnight, year 2000, etc.), time is best expressed as a monotonically increasing value (e.g. seconds since Jan 1 1990).

CAC for Maximal Charges

Overview (Figure 13)

[0095] Previously, a method for flow modulation was presented that maximizes charges to clients with active sessions. The CAC embodiment presented previously was not sufficient as it does not consider the cost of service as a basis for connection acceptance. As a result, it may turn away higher paying customers while granting service to lower paying ones, thereby defeating the purpose for this particular embodiment. Therefore, another embodiment is defined which offers the following features:

1. Awaiting clients are serviced in order of their respective service categories, higher paying clients first.

2. Once accepted, a client is guaranteed to receive acceptable service irrespective of its service category.

3. Under heavy load conditions higher paying customers are more likely to be accepted than lower paying ones.

4. Lower paying customers will no be starved for service when higher paying ones enjoy a surplus.

5. Available bandwidth is not thrown away needlessly while clients are being denied service.

[0096] The first objective is easily met by dividing the client queue into as many bands as there are service categories, resulting in a banded queue. Bands are ordered within the queue according to their service categories, with the costliest category in front. As prospective clients arrive and make their selection they are placed in their respective queue band according to their service category (which may be set contractually, according to content selection, etc.).

[0097] Our second objective is met by employing a procedure patterned after those presented previously & offering the same guarantee. Toward our third and fourth objectives we propose dividing total available bandwidth in as many strata as there are service categories. As depicted in Figure 12, two service categories are shown, Premium and Basic, each entailing an associated cost of service. A prospective client is accepted only if there is sufficient swing capacity available within its given service category. The swing capacity for a given category is given by the smaller of 1) the difference between its maximum floor flow rate (corresponding to the top of the stratum for the service category) and the sum of the minimum rates of all active sessions in its category or below, and 2) available swing capacity overall. Finally, our fifth objective is met by allowing the flow optimizer to function freely subject to its operational constraints. The imposed ceilings on call acceptance by category relate to *minimum flow rates*, which merely impose a floor on *actual flow rates*. For example, basic clients might well consume all available bandwidth 300 in the absence of any premium customers, yet could be throttled back toward their floor flow rates (which together cannot exceed 200 in this example) at any time should any premium customer suddenly demand service. In contrast, premium customers could consume the entire 300

bandwidth. As lower paying customers appear these would be admitted to the degree that their quota on minimum flow is not exceeded (i.e. 200) and the availability of swing capacity on the system.

Procedure (Figure 13)

[0098] The present procedure requires a number of ancillary definitions, which follow:

Let the service categories be denoted by k= 1...N, where k also denotes the cost of service.

Let server.maxMinFlowRate[k-1] be the top of the stratum for service category k. Note that server.maxMinFlowRate [N-1] = server.maxFlowRate.

Let S_K be the set of active client indices with a service category *equal to or less than* k. Note that S1 is contained in S2, S2 is contained in S3, and so forth, and that $S_N = S_{activeClients}$.

Let swingCapacity(k) denote available swing capacity for service category k. By construction:

(19) swingCapacity(k) = minimum of:
(server. maxMinFlowRate [k-1] - $\Sigma_{i \in Sk}$(client.lookup(i).session.minFlowRate)),
(server.maxFlowRate - $\Sigma_{i \in SactiveClients}$ (client.lookup(i).session.minFlowRate)))

[0099] Referring to Figure 13, this method is used for CAC when multiple rate tariffs are in effect, and there is a desire to maximize economic returns to the service provider while offering acceptable service to all.

[00100] All waiting clients are scanned in FIFO sequence. The actions taken in Steps 1320 and 1360 are identical to those described in connection with earlier CAC flowcharts.

[00101] Step 1340 evaluates a predicate expression that tests whether the required bandwidth can be accommodated without exceeding the lesser of 1) swing capacity available to the client's category of service, and 2) total available swing across all categories of service. The latter factor could be determinative if all available bandwidth were allocated to high paying customers, leaving lower paying ones such as the proposed client unable to draw from their unfilled quota.

Let us suppose that candidate client ***id*** belongs to rate category ***k***.

We define the swing capacity available in rate category ***k*** as:

swingCapacity(***k***) = least of:

(server.maxMinFlowRate[k-1] - $\Sigma_{\text{i in Sk}}$ (client.lookup(i).session.minimumFlowRate))

and

(server.maxFlowRate - $\Sigma_{\text{i} \in \text{SactiveClients}}$ (client.lookup(i).session.minimumFlowRate))

The predicate expression invoked by step 1340 can now be written as follows:

(required_bandwidth < = swingCapacity(***k***) - server.cac_flowSafetyMargin)

[00102] This algorithm processes queued clients in band sequence, and within every band in FIFO order. If the predicate evaluates to true, the client is activated. Otherwise two possible cases are considered: 1) under the *FirstFit* embodiment, the procedure continues scanning clients to the end of the banded queue, activating clients whose requirements can be met; 2) under the *FIFO* embodiment, the procedure ends with the first candidate client whose requirements cannot be met. Many other variations on these two embodiments might also be considered.

Anticipating Scheduled Variations in Network Capacity (Maximal Charge)

Overview

[00103] The procedure applicable to optimization of delivery charges is obtained by blending elements of the CAC method depicted in Figure 13 into the method depicted in Figure 11, which applies without change. To understand how this might work it may be useful to visualize a version of Figure 10 stratified along its length in the manner of Figure 8. As the maximum flow level undergoes a step change, so too do the widths of its constituent strata in equal proportion.

Procedure

[00104] As previously mentioned, the CAC method (Figure 11) applies to this circumstance also, provided we alter the definition of two routines, (17) and (18), upon which that procedure relies, yielding (20) and (21), and adopt the banded queue organization outlined in the previous section.

[00105] The server swing capacity at a future time t is computed according to the capacity and worst-case client flows at time t.

(20) swingCapacity(k, t) = minimum of (

(server.maxFlowRate(t) * (server.maxMinFlowRate[k-1] /server.maxMinFlowRate[N-1])

$\Sigma_{\text{i e Sk \& (t\_end(i) > t)}}$ (client.lookup(i).session.minFlowRate)),

(server.maxFlowRate - $\Sigma_{\text{i e SactiveClients \& (t\_end(i) > t)}}$ (client.lookup(i).session.minFlowRate)))

Finally, we define a predicate that tests whether a prospective customer will cause swing capacity to be exceeded at some time t, as follows:

```
(21) boolean client_fits(i ,t) {
        k = client.lookup(i).costOfService;
        if(client.lookup(i).contentSelection.averagePlayRate <=
                        swingCapacity(k,t) - server.cac_flowSafetyMargin)
                return true;
        else return false;
    }
```

System Description

[00106] Referring to Figure 14, a block diagram illustrating a preferred embodiment for a system implementing the methods presented herein is shown.

[00107] Block 1485 depicts plurality of network-based client computers that receive data from a server over the course of a session. Every client sessions has an associated session block 1435 within the server by means which a flow rate of content from client to server is effected and regulated to a given flow rate set-point. The flow optimizer 1430 manages bandwidth utilization across all sessions that have been admitted by the system call admission control (CAC) and the bandwidth call admission control blocks, labeled 1400 and 1420 respectively. Specifically, the flow optimizer 1430 modulates the flow rate set-point of every active session so as to optimize aggregate flow or, more generally, cost.

Call Admission Control

[00108] Whenever a client contacts a server requesting service, the server must determine whether or not to admit the client. This admission function is performed in two separate phases. The system call admission control block 1400 considers whether or not sufficient internal resources exist to support the request.

Toward this end, the system CAC 1400 estimates the needs of the prospective client with respect to a set of limited system resources such as I/O bandwidth, cache memory, open files, threads, etc. A request is denied if any required resource cannot be obtained. In the contrary case, the request is passed on to the Bandwidth CAC, which tests if the client fits and admits or rejects the client accordingly. The swing capacity is computed by the flow optimizer 1430 on a periodic basis according to equations (14) and (17) as appropriate. The managed bandwidth must reflect the true capacity of the attached network and the adapters leading thereto. This parameter should not be set arbitrarily high to prevent rejection of client requests. Doing so will cause the system to admit calls that cannot be properly handled, thereby diminishing the quality of the viewing experience.

[00109] Having admitted a client, block 1430 creates a client session block 1435, which then begins to converse with its client-resident counterpart, 1480) during their launch phase. Thereafter, the flow optimizer 1430 assumes control, and modulates flow over the session until either all content is delivered, or a failure to deliver is detected, or one of a number of VCR-like stream control events intervene (e.g. pause, rewind, fast-forward).

Server and Client Session Control Blocks

[00110] Data Transport Channel 1450 is responsible for transporting data packets originating from the flow regulator 1440 within the server to the former's peer entity 1460 within the client. The mode of communication between the peer entities may be connection-oriented (e.g. TCP, which delivers a byte stream in order, without duplication or loss) or datagram-oriented (e.g. UDP). Best-efforts datagram service is typically enhanced with acknowledgements to facilitate rapid recovery from occasional packet loss, out-of-sequence delivery, and duplication.

[00111] Peered session control channels 1445 and 1465 permit configuration and status data, together with miscellaneous commands, to be sent from client to server and vice-versa. A reliable connection-oriented transport service is typically employed for this purpose (e.g. TCP). Data and commands alike are multiplexed over this channel in the service of other entities within the clients or server. Thus, a datagram-oriented data transport channel 1460 within a client 1485 might employ the session control channel to ferry selective acknowledgements to its peer in the server. Similarly, the flow meter 1465 within the client 1485 forwards the measured flow rate to the associated flow regulator 1440 within the server.

[00112] The flow regulator 1440 obtains content data from a cache or disk, at the current offset into the content file, which it forwards as packets to the data transfer channel 1450. The size and/or pacing of the packets are determined by the flow regulator 1440 in order to achieve the flow rate set-point imposed by the flow optimizer 1430. The flow regulator 1440 is also responsible for determining the channel flow capacity to the client, which is the least of a configured value obtained from the client and a measured value. Toward this end, during the launch phase, the flow regulator 1440 might send content data packets to the client flow meter 1475 in a series of packet trains at progressively higher peak flow rates ending with the configured maximum flow rate, if any. The highest measured flow rate reported by the client flow meter 1475 is accepted as the flow rate capacity, which is posted for use by the flow optimizer 1430. Subsequently, the flow regulator 1440 will compare the flow rate set-point with the flow rates reported by the flow meter 1465, and down grade capacity as needed should the measured flow rate fall consistently fall below the requested flow rate. Whenever such a downgrading occurs, the flow regulator 1440 will test for subsequent

relaxation of a transient constriction by means of the aforementioned series of packet trains in which the peak packet flow rate exceeds the delivered flow rate.

[00113] The buffer manager 1470 accepts data from the flow meter 1465, which data is retained until it is consumed or passed over the content consumer 1480. The buffer will typically be implemented as a ring buffer in memory or on disk. The write cursor will lag the read cursor by a no less than a specifiable time interval so as to permit rewinds of acceptable scope without requiring selective data re-retrieval from the server.

Flow Optimizer 1430

[00114] The flow optimizer 1430 assumes control over a session once it has been launched, and its channel capacity has been determined. Subsequently, it modulates the flow rate set-point of every session such as to optimize aggregate flow or charges across all active sessions, subject to a variety of min/max constraints on sessions flows, and a maximum constraint on aggregate session flow, as previously discussed.

[00115] The flow optimizer 1430 continues to modulate session flow until all content has been delivered, or one of a number of session anomalies intervene, as follows:

[00116] A downgraded session flow capacity drops below the minimum session flow rate. The session is dropped by the optimizer 1430 as a solution involving this channel does not exist. The user is forced to apply for re-admission, with the option of accepting slower than real-time delivery at the present encoding rate, or accepting a degraded viewing experience at a reduced encoding rate better adapted to the presently available bandwidth.

[00117] A pause VCR command is received from the client. Flow to the client can continue until the client-side buffer is topped off, after which time flow drops to zero. The session is permanently dropped should the pause persists too long. Until that time, the CAC safety margin is temporarily increased by the jit (Just-In-Time) flow rate of the paused session. In this way the session can be resumed without the need for re-admission and the jit bandwidth is made available to currently sessions so long as the pause persists.

[00118] A VCR Jump-Forward command is received from the client. This action is problematic as it causes the just-in-time flow rate (calculated at the new forward offset) to increase, in violation of the assumptions of the CAC algorithm. It must be noted that this violation will occur even if the content that is sought has already been delivered to the client-side buffer. One workable policy is to force the client to apply for re-admission at the forward offset and with a suitably modified buffer level.

[00119] A VCR Jump-Backward (Rewind) command is received. This action may entail the need for re-transmission of content, as previously discussed. The just-in-time flow rate will decrease as per the new position and adjusted buffer contents.

Tunable Minimum Constraint

[00120] A parameterized family of minimum flow constraints can be defined as follows:

(22)

wherein: $f_k(t)$ = Target content flow rate for session k;

$f_k^{JIT}(t)$ = Just-in-time flow rate for session k;

(RHS of Equation (5))

β = Burst tuning factor (positive or zero); and

$f_k^{min}(t)$ = Minimum (reserved) flow rate for session k.

[00121] This family converges on $f_k(t) > = f_k^{JIT}(t)$ when β approaches 0. A β value of .1, for example, ensures a minimum flow rate that exceeds the JIT flow rate by 10%. Thus, non-zero values of β force early delivery of content to a player's media buffer, with beneficial effects on the viewing experience.

Alternative Call Admission Control

[00122] The traditional CAC algorithm compares the flow rate requirement of a prospective client with the unused bandwidth (obtained by subtracting the aggregate flow target from the server flow capacity) and grants service if unused bandwidth is not exceeded. This procedure may break down as the optimizer causes all available bandwidth to be used if it can, thereby reducing unused bandwidth to zero in the best case, even when a new client could be accommodated with a healthy margin to spare.

[00123] Nevertheless, a simple variation on this idea can be used. Available bandwidths can be defined as the difference between server flow capacity and the sum of minimum flow rates for each session computed according to (22) as follows:

(23) $f_{available} = f_{Svr}^{max} - (f_1^{min} + \ldots + f_n^{min})$

wherein: $f_{available}$ = Available bandwidth

f_{Svr}^{max} = Server aggregate flow capacity; and

f_n^{min} = Minimum (reserved) flow rate for session n.

[00124] Service is granted if the computed minimum flow rate of the prospective client is less than available bandwidth even when unused bandwidth is zero.

Toward a Tunable QOS

[00125] In the preceding discussion it is assumed that all clients share the same value of β. One might also consider a scheme whereby a client receives a β value according to the class/cost of service that client has signed up for: the greater the cost (and expected quality of service, or QOS) the greater the value of β.

[00126] Instead of maximizing charges by the methods described above, a variation on the algorithm to maximize flow is considered whereby the reserve flow computation for every session takes into account the individual β value of the associated client. In this fashion a single algorithm can serve either end. The optimizer first ensures that every session receives its required minimum bandwidth (which would now typically depend on β and elapsed session time) before allocating any excess bandwidth more or less evenly to all by interpolation. (According to Figure 15.)

[00127] The scheme is simpler computationally than the iterative interpolation method shown in Figure 16; it is also arguably more fair and stable with respect tot he allocation of available bandwidth.

Deadline-Driven Content Distribution

[00128] In delivering viewable or audible content to a player, the optimization algorithm imposes a minimum flow rate that ensures that all content is delivered by the implied deadline, which is the last moment of a play session. Notwithstanding, the optimizer attempts to deliver content earlier according to bandwidth availability within the server taken as a whole. A similar, though far simpler scenario arises relative to the distribution of content from an optimized server to a target device, whenever the delivery must be accomplished by a designated time, or equivalently, within a given elapsed time, and when overlapped non-blocking consumption of the content at a fixed rate is not required.

[00129] One example concerns the application where a session bundle comprises a session associated with the featured program together with sessions for every advertisement scheduled to interrupt the feature program at a designated time. Each session is endowed with a separate media buffer, which, in the case of ads, would typically be large enough to accommodate the ad in full, thereby permitting a seamless transition to the ad at the appointed time.

[00130] In adapting the optimizer to the needs of content distribution we must revisit the formulation of the flow constraints, as follows:

1. Media buffers are not of a concern, let alone buffer overflow, and it can be assumed that all content received is saved away in a file on disk, or in memory awaiting further use.

2. For these same reasons, the value of β may be zero, at which the minimum flow rate equals the just-in-time rate obtained by dividing the content as yet undelivered by the time till deadline:

(24) $f_k^{JIT}(t_k) = L_k^{TOGO}/(T_k - t_k)$

wherein: L_k^{TOGO} = Content to go for session k, i.e., as yet undelivered by the server at time t, (client.lookup (k).session.payloadToGo).

Given the arbitrary format of the file to be distributed, T no longer signifies the "duration of play" at the average play rate (averagePlayRate); rather, T represents the elapsed time to the deadline measured at the moment the delivery session was launched, and t represents the elapsed time since session launch.

Session Bundles

[00131] A session bundle is a set of sessions that flow along a shared channel, and are thus subject to an aggregate flow constraint equal to the capacity of the shared channel. One important application of session bundles is connected with ad insertion.

[00132] It is not enough to treat each of the constituent sessions as independent with respect to the flow optimizer, for this practice might well result in session flow targets that do not exceed channel capacity when considered singly, yet exceed this same shared capacity when summed together.

[00133] A solution to this problem involves two distinct steps:

1. The sessions bundle is treated as a virtual session by the server optimizer, subject to the minimum flow constraint obtained by summation of the individual session minima, and a maximum constraint that is the least of (a) the sum of the least of constraints (3) and (4) for each session, and (b) the shared

channel capacity client.lookup(k).channel.maxFlowRate (f_k^{cap}). The flow optimizer then generates an aggregate flow target for all channels virtual sessions by interpolation, using:

(25) $\text{Cost}(t) = f_1(t) + .. + f_n(t)$

[00134] The cost vector in this instance is the n-dimensional unit vector, where n is the number of active clients.

(26) $c = (1, 1, \ldots, 1)$

[00135] Now it is possible to have an achievable aggregate flow target for the session bundle k that we must apportion optimally among the constituent sessions. Fortunately for us, our interpolation procedure of Figure 15 can be applied again to the sessions in our bundle, but where the aggregate flow target generated above replaces the aggregate flow constraint f_{svr}^{max}

Specifically:

For every flow f_{ki}, within session bundle k, the optimal flow rate is obtained by interpolation between minimum and maximum session flows:

(27) $f_{ki}^{opt} = f_{ki}^{min} + \alpha * (f_{ki}^{max} - f_{ki}^{min})$

(28) $\alpha = [(f_K^{bundle} - f_{k1}^{min} + \ldots f_{kn}^{min})]/[(f_{k1}^{max} - f_{k2}^{min}) + + (F_{kn}^{max} - f_{kn}^{min})]$

and the value of f_k^{bundle} is obtained from a higher level optimization in which session bundle k is treated as a virtual session for which:

(29) $f_k^{min} = f_{k1}^{min} + \ldots + f_{kn}^{min}$

(30) $f_k^{max} = \min(f_{k1}^{max} + .. + f_{kn}^{max}, f_k^{cap})$

It must be noted that session bundles are typically dynamic in nature, outliving many if not all of their constituent sessions. Before a session is added to the bundle two CAC decisions must be reached:

(1) establish bandwidth availability within the bundle; and

(2) bandwidth availability within the server.

Relationship to Adaptive Rate Control

[00136] The flow optimization methods should be viewed as constituting the lowest and perhaps only the first of many supervisory control layers to sit astride all streaming sessions, which comprise their own hierarchical stack of control functions; rate adaptive control sitting above a flow rate regulator which relies on a transport entity.

[00137] Channel capacity is known to vary widely over time. A capacity drop to below the minimum reserve flow rate is problematic as it causes the latter to rise in violation of its monotonicity assumption. In a server with available bandwidth to spare such a rise can be accommodated by the flow optimizer (by bumping the session's minimum flow reservation to the higher value) provided the flow deficit can be made up in the near term by draining pre-delivered content from the media buffer. Should either of these resources approach depletion, stream degradation to a lesser bit-rate or frame rate may be the only viable option short of a pause. Whether up or down, any rate change is a change to *fe*, which must be preceded by change to session's reserve flow rate allocation by the optimizer.

[00138] As can be seen, the optimizer and rate adapter must be integrated with care. One possible approach would be to modify a rate adapter such that it renegotiates its proposed *fe* value with the flow optimizer prior to effecting the rate change it deems appropriate based on information provided by the flow regulator. Rate increases might well be subject to deferral under heavy load conditions, awaiting bandwidth availability. The opposite is true of rate flow decreases: these would be forestalled as long as possible by the optimizer on a system that is not fully loaded, as previously discussed.

Live Events

[00139] Referring to Figure 17, the method of the present application can be extended to the optimization of "live", time-shifted "near-live" events, and cascaded delivery, wherein one server delivers content to another while the latter accepts clients for the same content. The extension assumes the following characteristics of the server and the live performance:

1. The performance is of a fixed duration T known in advance, encoded at an average rate *fe*, with a projected content payload of *fe*T*;

2. The origin server at the point of injection delays the stream by a small dwell time T^{DWELL} on the order of seconds (e.g. 10). Notwithstanding, the delayed content is available for early delivery to mirrored servers and players with the aim of providing these with a modicum of isolation from network jitter affecting the timely delivery of the live video frames.

3. Throughout the live event, captured content is stored in a file, thereby permitting late joiners to view the "live event" starting from any point between the present and the event start time;

4. At the conclusion of the live event, the encoded content, now fully captured to a file, is available for viewing over the balance of the admission window (e.g., 60 minutes in Figure 17).

[00140] The optimisation of live events and video on demand (VOD) differ principally in one respect, namely in the formulation of the server content underflow constraint.

For VOD, this constraint is

(31)

$L_k^{TOGO}(t)$ is initialised to the content size (*fe*T*) at that start of play, and decremented as content is delivered to the client. Δt is the time interval between successive periodic optimizations (e.g. 1 second). In practice this constraint rides

well above other operative maximum flow constraints (e.g. buffer overflow, channel capacity) until the very last calculation of the session flow target by the optimiser.

[00141] For a live stream, this constraint must additionally reflect the fact that content is streaming into the server (where it builds up), even as the server streams it out toward clients, and that the latter activity cannot possibly overtake the former.

[00142] Referring to Figure 17, (which illustrates how optimised live streams behave) line L1 depicts the build-up of freshly captured content to the server. Line L2, which is delayed (time-shifted) relative to L1 by the dwell time T^{DWELL}, represents the live performance as cast and viewable by clients. One such client is admitted at time A and joins the live session at the live moment B. The dashed curved line linking points B and C shows a plausible build-up of content in the client's media buffer in advance of play, as the viewer continues its advance along line L2. Past point C the top of the media buffer tracks the live capture stream and thenceforth advances along line L1. In this instance, the time shift between the live performance as cast (L2) and the live performance as captured (L1) imposes a maximum on the amount of content that could ever accumulate in a client media buffer, irrespective of its actual capacity.

[00143] Another client, admitted at time $t=10$, is interested in viewing the live stream from the beginning without missing anything. As before, the dashed line joining 10 and D depicts a plausible build-up of content in the client's media buffer in advance of play, as the viewer continues its advance along line L3, which is time-shifted by $(10+T^{DWELL})$ minutes relative to L1. The greater time shift (now between L1 and L3) permits more extensive pre-buffering of "near live" content.

[00144] The present constraints do not impose suitable flow limits at and beyond points C and D when the client delivery streams finally links up with the live capture stream, in the live moment. Consequently the optimiser may propose flow targets that cannot be realized.

[00145] Accordingly, the following modification for use in connection with live or cascaded streams is made:

(32) $f_k(t) <= [(L_k^{TOGO}(t) - Live^{TOGO}(t)] / \Delta t$

for a live event delivered at flow rate *fe*, we have

(33) $Live^{TOGO}(t) = \max(fe*(T - T^{DWELL} - t), 0)$

[00146] $Live^{TOGO}(t)$ is the span between projected content size *fe* **T* (level line L0 in Figure 17) and the content capture level line L1, whereas $L_k^{TOGO}(t)$ is the gap between L0 (i.e., *fe* * *T*) and the dashed curved line representing the content level delivered to the media buffer. The latter must always exceed the former. As the two approach one another this constraint takes effect, causing $L_k^{TOGO}(t)$ to always exceed $Live^{TOGO}(t)$ as it decreases at rate *fe*.

[00147] Viewers joining late and willing to see the live performance with some delay relative to the live case enjoy higher potential levels of isolation from network disturbances, and thus QOS, by virtue of their greater effective buffer capacity. By the same token, such viewers afford the optimizer greater latitude to optimize the flow of VOD and "live" content across load fluctuations, thereby enhancing server capacity. Viewers that wish to experience the live event in "real-time" nevertheless benefit by a modestly enhanced QOS conferred by a dwell time's worth of content in their buffers.

Summary

[00148] A method and system for call/connection acceptance and flow modulation for network delivery of video/audio programming is thus provided. Although several embodiments have been illustrated and described, it will be apparent to those skilled in the art that various changes and modifications may be made without departing from the spirit of the invention or the scope of the claims.

CLAIMS:

1. A method of bandwidth allocation for delivery of stored digital content from at least one server device to at least one client device by way of a network, the method comprising the steps of:

a) prescribing a control variable which represents a target flow rate from the server device to each client device;

b) determining time-varying constraints on the flow rate of the content;

c) determining a cost function of the control variables for all clients; and

d) prescribing bandwidth to all clients based upon the value of the control variables that maximize the cost function.

2. The method of Claim 1 wherein step (d) comprises performing periodic computations to update the value of the control variable such that the bandwidth can be continuously allocated to each client.

3. The method of Claim 2 wherein a new client is accepted by:

i) determining an admission capacity of the bandwidth;

ii) admitting a prospective client if the clients minimum allowed value of the control variable is less than the admission capacity; and

iii) wherein a client admitted for service is guaranteed to have sufficient content flow over the entire session.

4. The method of Claim 3 wherein:

The admission capacity equals a server swing capacity reduced by a predetermined safety margin which may be zero, the swing capacity equaling the difference between the server flow capacity and the sum of the minimum allowed flow rate for all clients.

5. The method of Claim 4 wherein the minimum allowed value of the control variable is initialized to the average consumption rate.

6. The method of Claim 4 wherein the minimum allowed value of the control variable is initialized to the average consumption rate multiplied by a factor greater than one.

7. The method of Claim 4 wherein the server flow capacity varies in a step-wise function over time according to a predetermined schedule and the method of accepting a new client further comprises:

i) determining a sequence of future step changes of server flow capacity;

ii) determining at each time in the sequence of future step changes the value for a future worst swing case capacity, the worst case swing capacity being obtained from the server flow capacity at the time by subtracting an extrapolated present minimum allowed flow rate for all active clients that are potentially active in the future;

iii) admitting a prospective client if a minimum allowed value of the control variable is less than the capacity obtained from the sequence and the present capacity;

iv) whereby a new client can be admitted without compromising the service to existing clients in the presence of previously scheduled server flow capacity changes and all clients accumulate content in their buffers to the extent of their respective buffer capacity.

8. The method according to Claim 7 wherein the method of accepting a new client further comprises:

i) partitioning the available bandwidth into partitions representing a prescribed service charge;

ii) determining a maximum-minimum capacity for a service charge to be the maximum allowed sum of all minimum allowed flow rates for all clients incurring an equal or lesser service charge, the maximum-minimum capacity equaling the sum of bandwidth allocations for all service charges less than or equal to the service charge;

iii) determinating a service-charge-specific swing capacity for a given service charge as the difference between the maximum-minimum capacity for the service charge and the sum of all minimum allowed flow rates for all clients incurring the service charge or a lesser service charge;

iv) determining the server swing capacity to be the server flow capacity minus a margin and minus the sum of the minimum allowed flow rate for all clients; and

v) determining the service-charge-specific capacity to be the least of the server swing capacity and the charge specific swing capacity.

9. The method of Claim 8 wherein the method of accepting a new client further comprises:

i) determining the sequence of future times of step changes to server flow capacity over a proposed consumption period;

ii) determining for each time in the sequence of future times, the future worst-case value for the service-charge-specific capacity, the worst case value being obtained by using the extrapolated present minimum allowed flow rate of all presently active clients that are potentially active at the time in the future; and

iii) admitting a client incurring a given service charge if the minimum flow rate of the client is less than the capacity obtained from the sequence and the present service-charge-specific capacity.

10. The method of Claim 3 wherein step (b) comprises determining time-varying constraints wherein:

i) an aggregate flow rate for all clients does not exceed a predetermined server flow capacity;

ii) the flow rate from the server to the client does not exceed a maximum allowed flow rate for the client;

iii) the flow rate to the client will never cause a buffer of the client to overflow;

iv) the flow rate to the client stops when the content is exhausted; and

v) the flow rate from the sever will never be less than the client's minimum allowed flow rate which may vary over time and may be zero.

11. The method of Claim 10 wherein the maximum allowed flow rate to the client is given by the minimum of one selected from the following:

i) a client flow rate ceiling;

ii) the flow rate required to fill the client buffer before the next periodic computation; and

iii) the flow rate required to complete the delivery of the content before the next periodic computation.

12. The method of Claim 11 wherein:

i) the minimum allowed client flow rate cannot increase over time, but may decrease such that the initial value of the control rate is no less than the average consumption rate;

ii) the control variable for the client is always greater than or equal to the current minimum allowed client flow rate; and

iii) the client buffer will never underflow until the content is fully consumed at the average consumption rate.

13. The method of Claim 12 wherein the minimum client flow rate is the product of the constant content flow rate from the client to the server that causes the last piece of content to be delivered at the last possible moment as the client buffer is being drained at the average consumption rate, with a factor greater or equal to one.

14. The method of Claim 13 wherein the minimum client flow rate is initialized to the average consumption rate.

15. The method of Claim 13 wherein the minimum client flow rate is initialized to the average consumption rate multiplied by a factor greater than one.

16. The method of Claim 13 further comprising the step of calculating the minimum client flow rate periodically.

17. The method of Claim 1 wherein the cost function represents an aggregate flow rate and is the sum of the flow rates for all clients.

18. The method of Claim 1 wherein the cost function represents a service charge and is the sum of the client flow rates multiplied by the client's cost of service.

19. The method of Claim 18 wherein:

step (a) comprises:

i) determining a maximum allowed flow rate and a minimum allowed flow rate for each client;

ii) determining a flow rate range for each client as the difference between the maximum allowed flow rate and the minimum allowed flow rate; and

iii) initializing a current flow rate for each client as the minimum allowed flow rate and summing the flow rate into the total server flow rate; and

step (d) comprises:

i) computing remaining server bandwidth as the difference between the maximum server flow capacity and the total server flow rate; and

ii) allocating remaining server bandwidth to remaining clients wherein all clients with a given cost of service receive bandwidth equally to the extent of their respective ranges, and no client with a given cost of service being allocated bandwidth unless all clients with a higher cost of service have received their maximum possible allocation of bandwidth.

20. The method of Claim 17 wherein step (a) prescribing the control variable comprises:

i) forming a hyperplane of control variables that aggregate to the maximum allowed aggregate flow rate;

ii) computing a first vector of minimum allowed client flow rates;

iii) computing a second vector of maximum allowed client flow rates;

iv) determining a multi-dimensional shape of which an interior diagonal spans the first and second vectors; and

v) finding a solution of control variables at the point of intersection between the diagonal and the hyperplane.

21. The method of Claim 20 wherein step (v) comprises finding a solution of control variables at the second vector if there is no point of intersection between the diagonal and the hyperplane.

22. The method of Claim 17 wherein step (a) comprises:

i) computing the sum of the minimum allowed flow rates to all active clients;

ii) computing the sum of the flow rate ranges for all active clients;

iii) computing the difference between the maximum allowed aggregate flow rate and the sum of the minimum;

iv) computing a factor as the ratio of the difference over the sum of the ranges;

v) if the factor exceeds 1, the control variable is set to the maximum allowed client flow rate; and

vi) if the factor is less than 1, computing the control variable by multiplying the client range by the factor and then adding the result to the minimum allowed client flow rate.

23. The method of Claim 17 wherein:

step (a) comprises:

i) determining the maximum allowed flow rate and minimum allowed flow rate for each client;

ii) determining the flow rate range for each client as the difference between the maximum flow rate and the minimum flow rate; and

iii) initializing the control variable for each client as said minimum allowed flow rate and summing the flow rate into a total server flow rate; and

step (d) comprises:

i) computing remaining server bandwidth as the difference between the maximum server flow capacity and the total server flow rate; and

ii) allocating remaining server bandwidth to remaining clients to the extent of their respective ranges.

24. The method of Claim 23 wherein the step of allocating remaining server bandwidth further comprises the clients receiving the bandwidth equally to the extent of their respective ranges.

25. The method of Claim 17 wherein the maximized value of the control variable can be determined by:

i) determining a maximimum and minimum flow rate for each client;

ii) determining a range between the maximimum and mininmum flow rates in order to find a flow rate range;

iii) determining available bandwidth by finding the difference between an aggregate flow capacity and a sum of the minimum flow rate;

iv) determining a flow variable by dividing the unused bandwidth by the flow rate range;

v) prescribing the control variable to be the minimum flow rate added to the flow rate range if the flow varible is less than one or prescribing the control variable to be the minimum flow rate added to the flow rate range corrected by the flow variable if the flow favorable is greater than one such that the control variable is easily calculated for each client.

26. The method of Claim 17 wherein at least one session is a session bundle and:

1) the minimum flow constraint is computed as the sum of the minimum flow rate constraints for all bundled sessions;

2) the maximum flow constraint is the least of 1) the sum of the maximum individual flow rate constraints of each session and 2) the flow capacity of the channel shared among the bundled sessions; and

3) the target flow rate is computed for the session is apportioned among the bundled constituent sessions.

27. The method of Claim 26 wherein the bundled session is an advertisement.

28. A system for allocating bandwidth between a server device and at least one client device, the system comprising:

a call acceptance module operative to receive an incoming request for service;

a flow regulator configured to deliver content at a modulated target flow rate, the content being delivered between the server device and a respective client device when a call is accepted by the call acceptance module; and

a flow optimizer configured to modulate the target flow rate of the flow regulator in order to optimize the aggregate flow of content.

29. The system of Claim 28 wherein the flow optimizer is configured to modulate the flow rate of the flow regulator in order to optimize charges for the content.

30. The system of Claim 29 wherein the flow regulator is configured to deliver the modulated flow rate of content in response to the control variable.

31. The system of Claim 28 wherein the flow optimizer is configured to determine a control variable which corresponds to an optimized flow rate.

32. The system of Claim 31 wherein the flow optimizer is configured to generate a cost function of the control variable that corresponds to a maximized value of the control variable.

33. The system of Claim 28 wherein the flow regulator is configured to deliver the modulated flow rate of content in response to the control variable.

34. The system of Claim 28 wherein the flow optimizer is configured to determine a maximized value of the control variable by:

i) determining a maximum and minimum flow rate for each client;

ii) determining a range between the maximum and minimum flow rates in order to find a flow rate range;

iii) determining unused bandwidth by finding the difference between an aggregate flow capacity and a sum of the minimum flow rate;

iv) determining a flow variable by dividing the unused bandwidth by the flow rate range;

v) prescribing the control variable to be the minimum flow rate added to the flow rate range if the flow variable is less than one or prescribing the control variable to be the minimum flow rate added to the flow rate range corrected by the flow variable if the flow favorable is greater than one such that the control variable is easily calculated for each client.

35. The system of Claim 33 wherein the call acceptance module is configured to accept a call based upon the value of the control variable.

36. A method of bandwidth allocation for delivery of stored digital content from at least one server device to at least one client device by way of a network, the method comprising the steps of:

a) prescribing a control variable which represents a target flow rate from the server to each client device based upon the amount of buffer of the client;

b) determining time-varying constraints on the flow rate of the content;

c) determining a cost function of the control variables for all clients in response to the size of the client's buffer; and

d) prescribing bandwidth to all clients based upon the value of the control variables that maximized the cost function.

37. A method for apportioning a given total flow rate (bandwidth) among a plurality of sessions for clients, the method comprising the steps of:

a) computing a sum of minimum flow rate constraints for each session;

b) computing a sum of flow rate ranges for all active clients;

c) computing a difference between the given total flow rate and a sum of the flow rate minima;

d) computing a factor as the ratio of the difference over the sum of the flow rate ranges;

e) if the factor exceeds 1, setting the flow rate to the maximum allowed client flow rate; and

f) if the factor is less than 1, computing the flow rate by multiplying the client range by the factor and then adding the result to the minimum allowed client flow rate.

38. The method of Claim 37 wherein the clients are distinguished by their cost of service and bandwidth is apportioned to clients iteratively, the amount apportioned to each client in a cost category is the amount remaining from prior allocations to clients in higher cost categories.

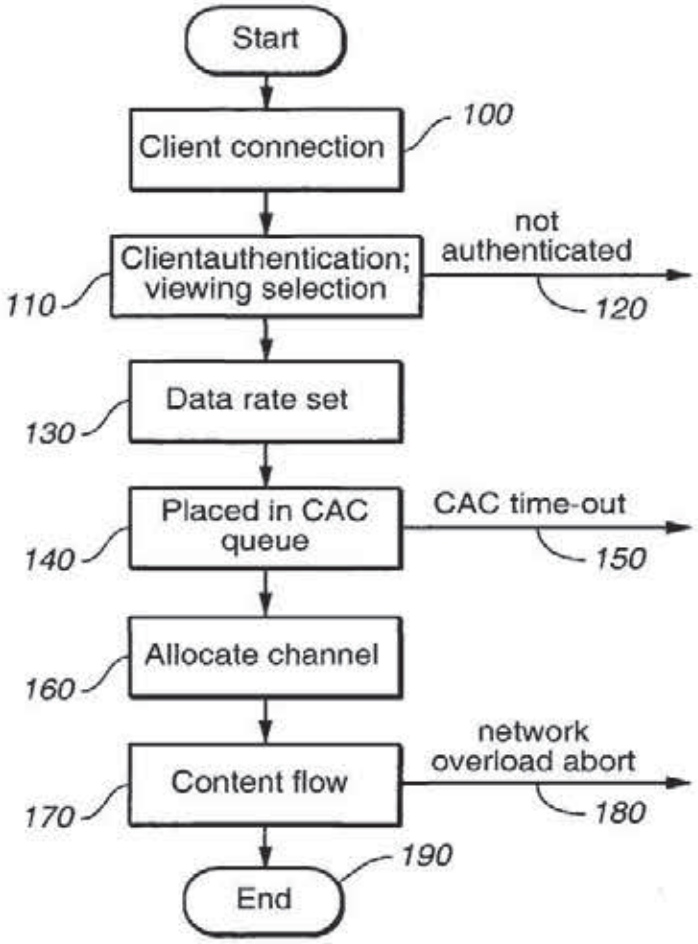


FIG._1

Ex. 1018-1677

Ex. 1018-1678

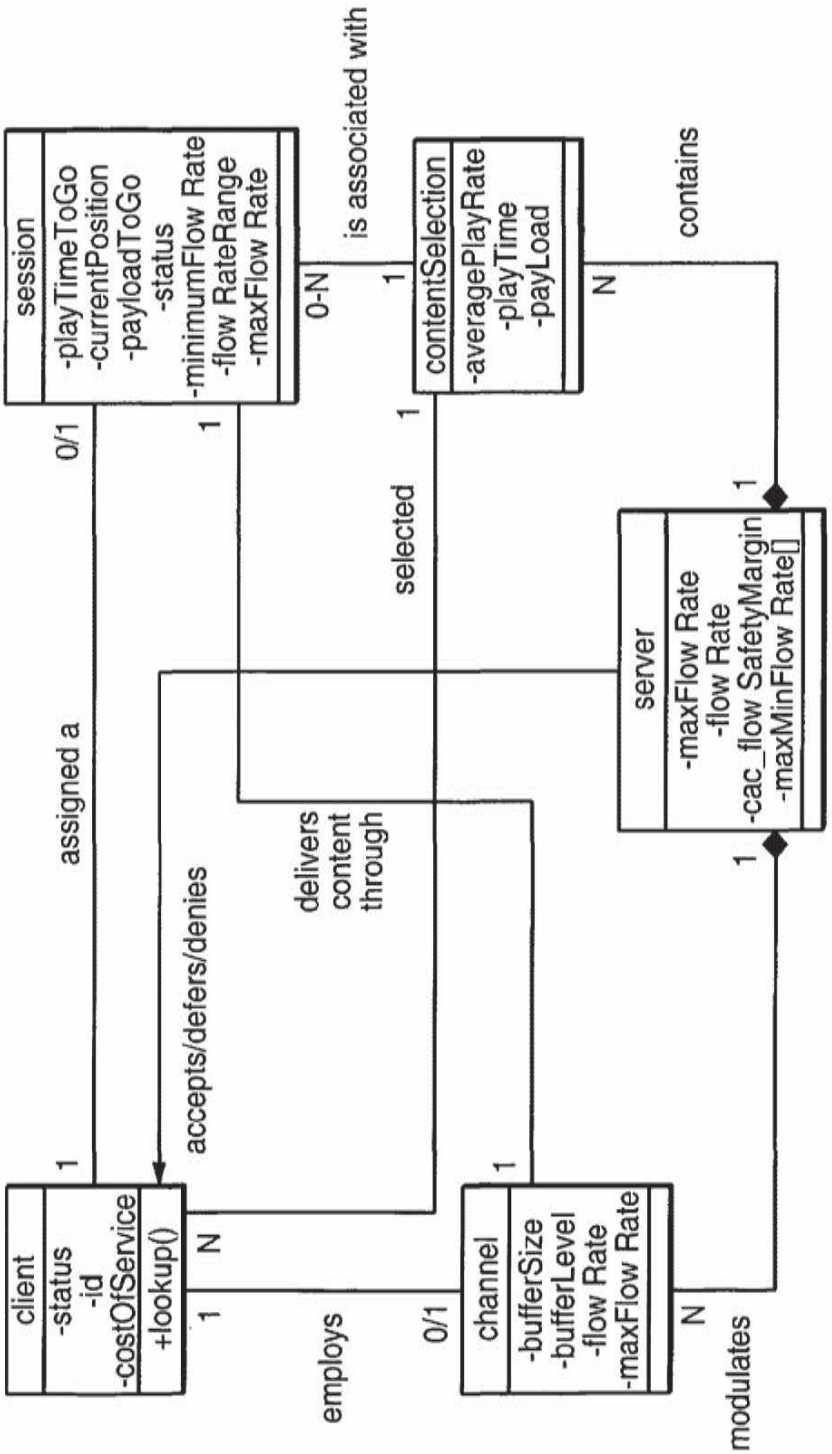


FIG._2

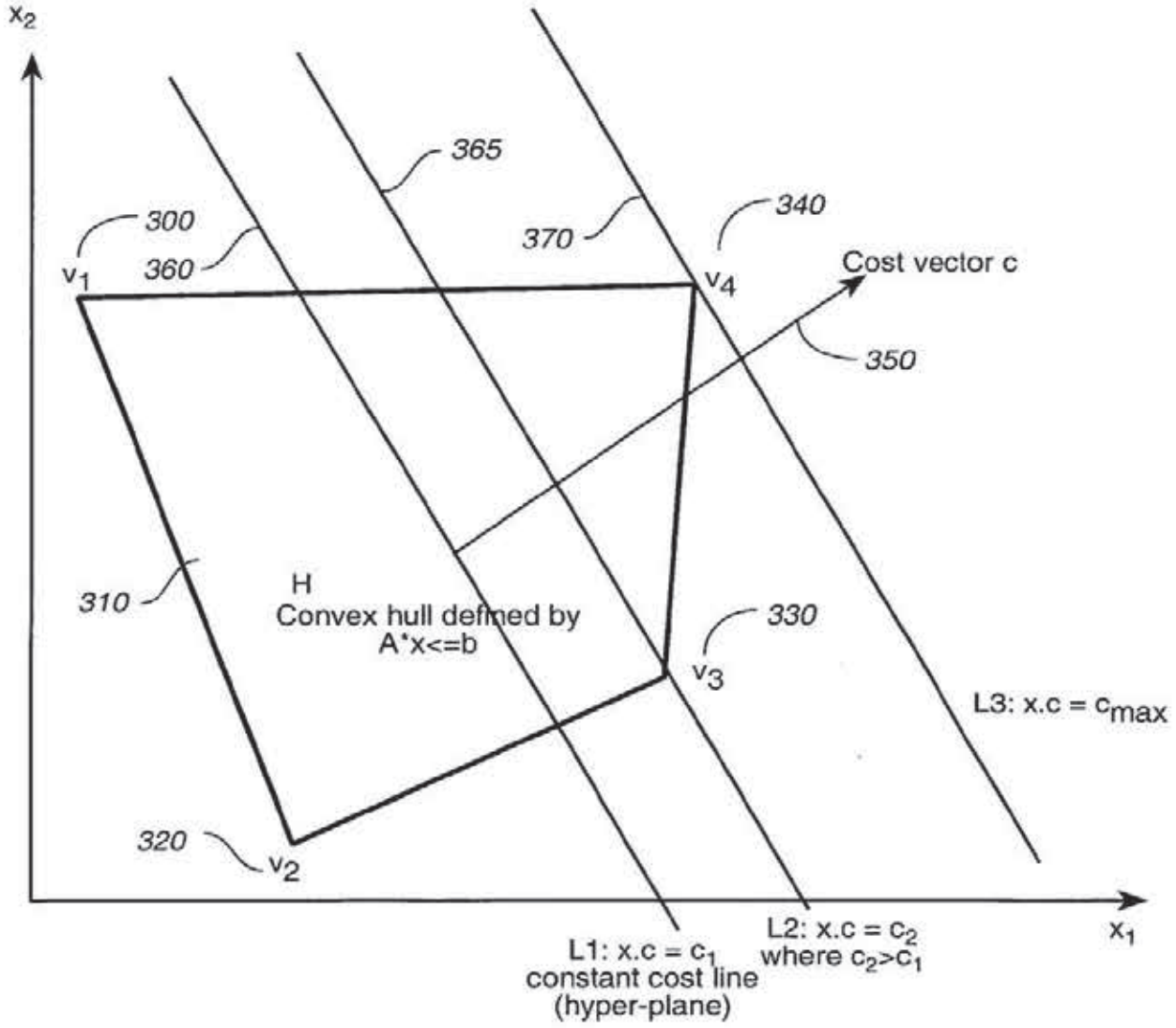

x2
x1
365
300
v1
360
370
340
v4
Cost vector c
350
310
H
Convex hull defined by
A*x<=b
330
v3
L3: x.c = cmax
320
v2
L1: x.c = c1
constant cost line
(hyper-plane)
L2: x.c = c2
where c2>c1

FIG._3

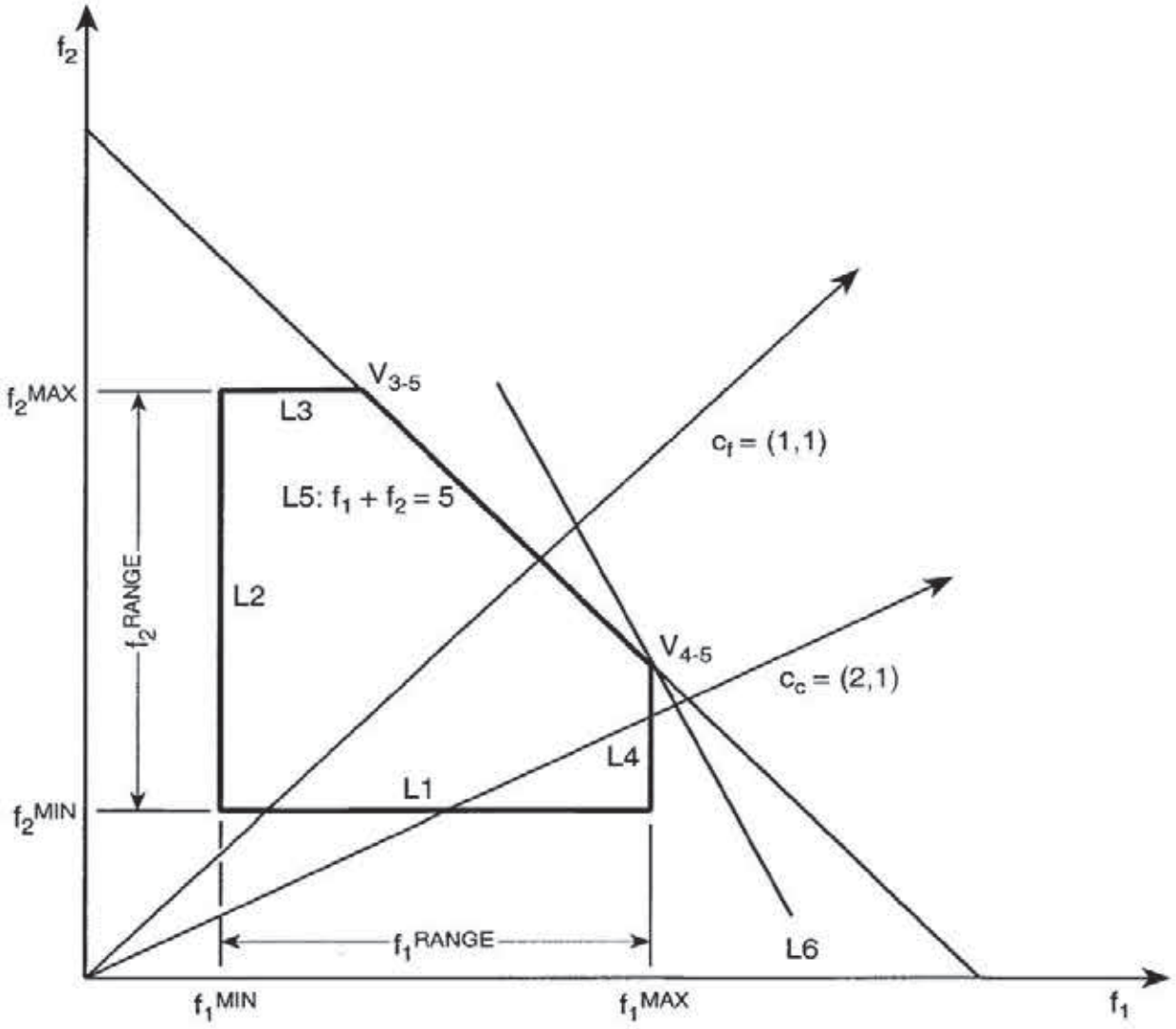

f2
f2MAX
f2MIN
f2RANGE
V3-5
L3
L5: f1 + f2 = 5
L2
V4-5
cf = (1,1)
cc = (2,1)
L4
L1
L6
f1RANGE
f1MIN
f1MAX
f1

FIG._4a

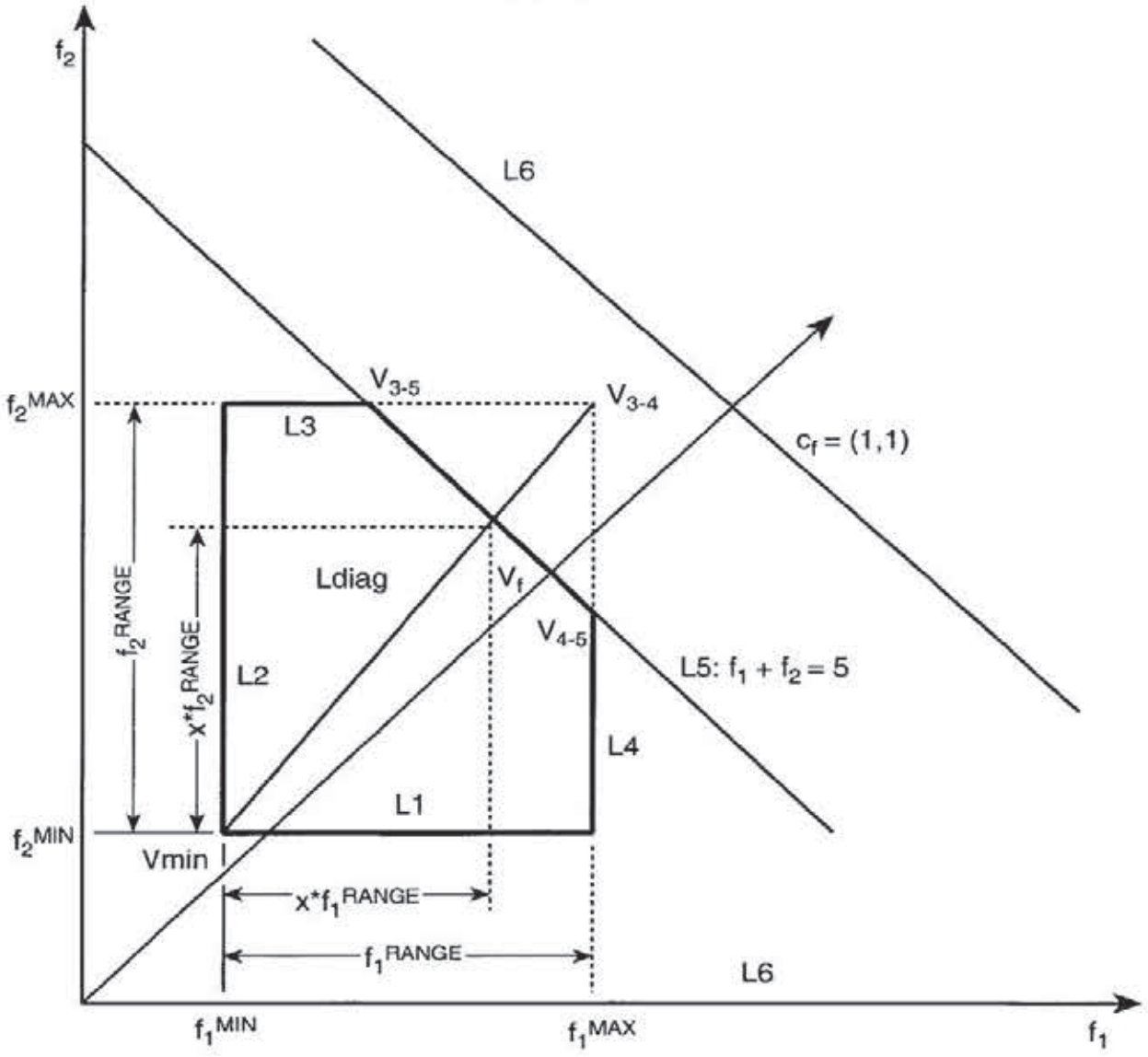

f2
L6
V3-5
V3-4
f2MAX
L3
cf = (1,1)
Ldiag
Vf
f2RANGE
x*f2RANGE
V4-5
L5: f1 + f2 = 5
L2
L4
L1
f2MIN
Vmin
x*f1RANGE
f1RANGE
L6
f1MIN
f1MAX
f1

FIG._4b

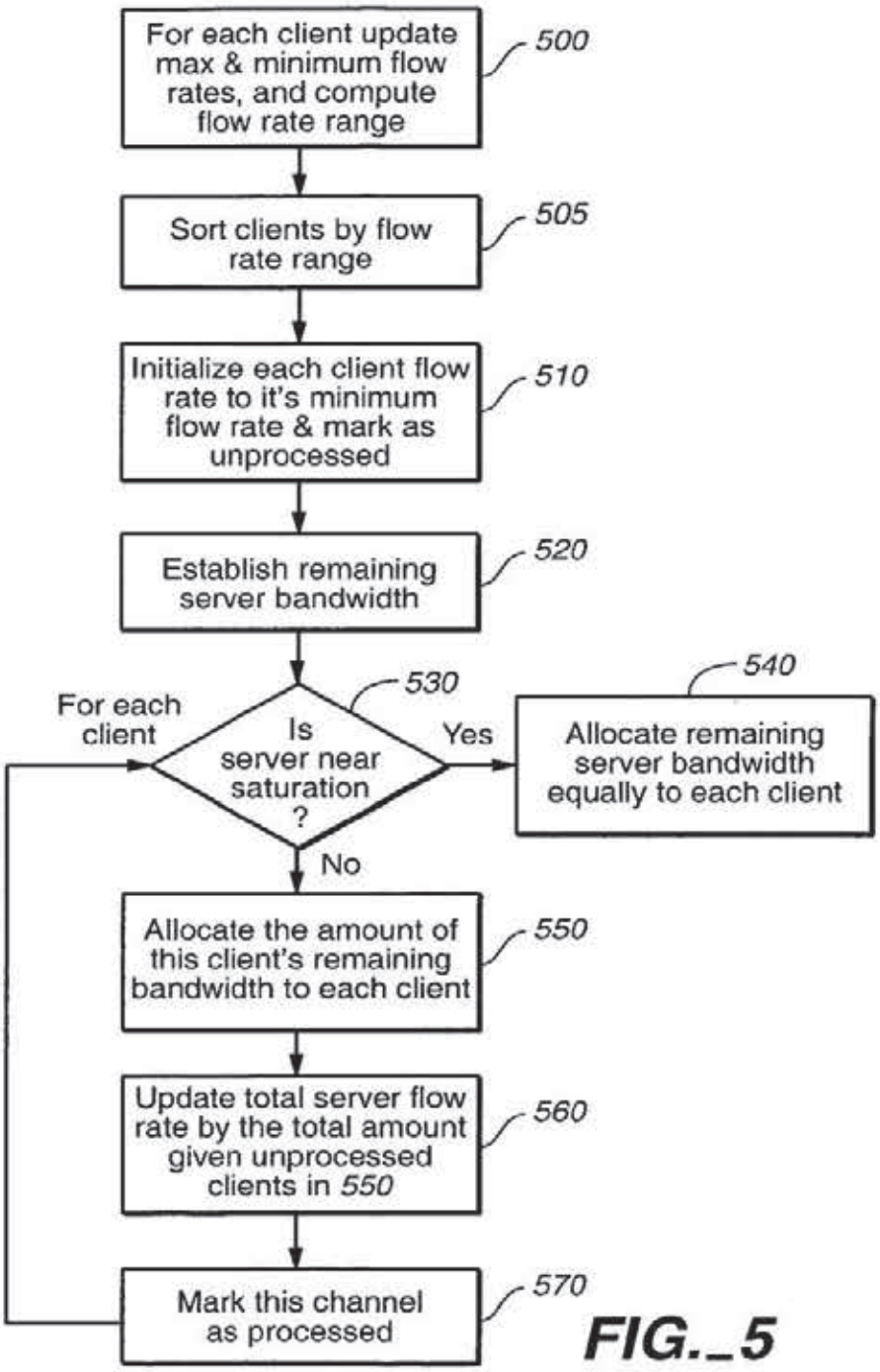

For each client update max & minimum flow rates, and compute flow rate range
500
Sort clients by flow rate range
505
Initialize each client flow rate to it's minimum flow rate & mark as unprocessed
510
Establish remaining server bandwidth
520
For each client
Is server near saturation ?
530
Yes
Allocate remaining server bandwidth equally to each client
540
No
Allocate the amount of this client's remaining bandwidth to each client
550
Update total server flow rate by the total amount given unprocessed clients in 550
560
Mark this channel as processed
570

FIG._5

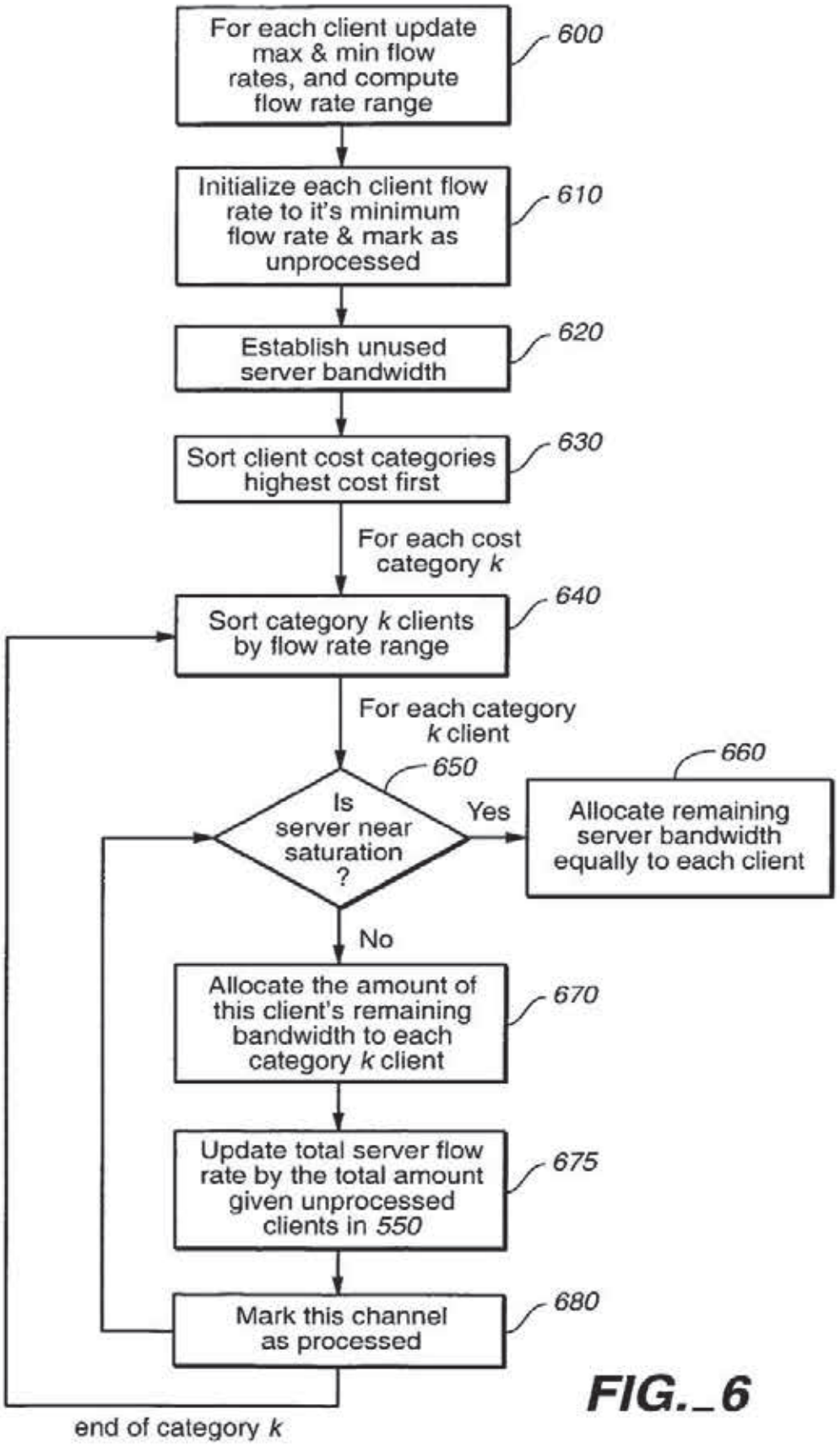

For each client update max & min flow rates, and compute flow rate range
600
Initialize each client flow rate to it's minimum flow rate & mark as unprocessed
610
Establish unused server bandwidth
620
Sort client cost categories highest cost first
630
For each cost category k
Sort category k clients by flow rate range
640
For each category k client
650
Is server near saturation ?
Yes
660
Allocate remaining server bandwidth equally to each client
No
Allocate the amount of this client's remaining bandwidth to each category k client
670
Update total server flow rate by the total amount given unprocessed clients in 550
675
Mark this channel as processed
680
end of category k

FIG._6

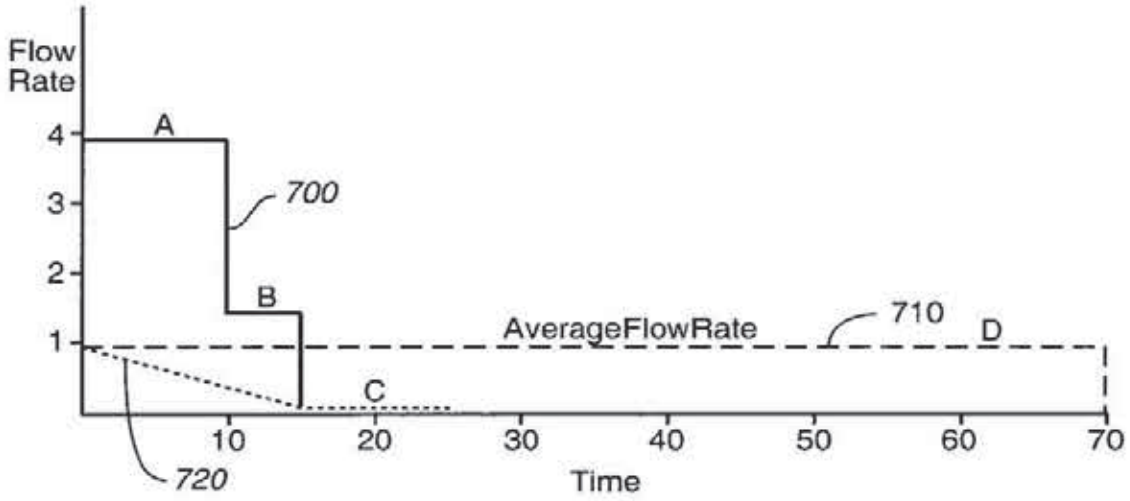

Flow Rate
4
3
2
1
A
700
B
C
AverageFlowRate
710
D
720
10
20
30
40
50
60
70
Time

FIG._7

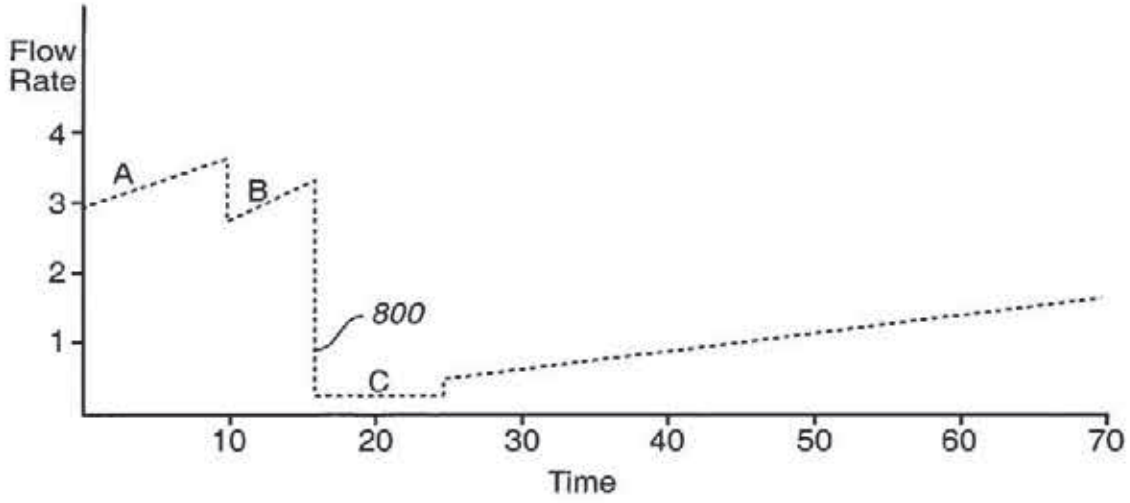

Flow Rate
4
3
2
1
A
B
800
C
10
20
30
40
50
60
70
Time

FIG._8

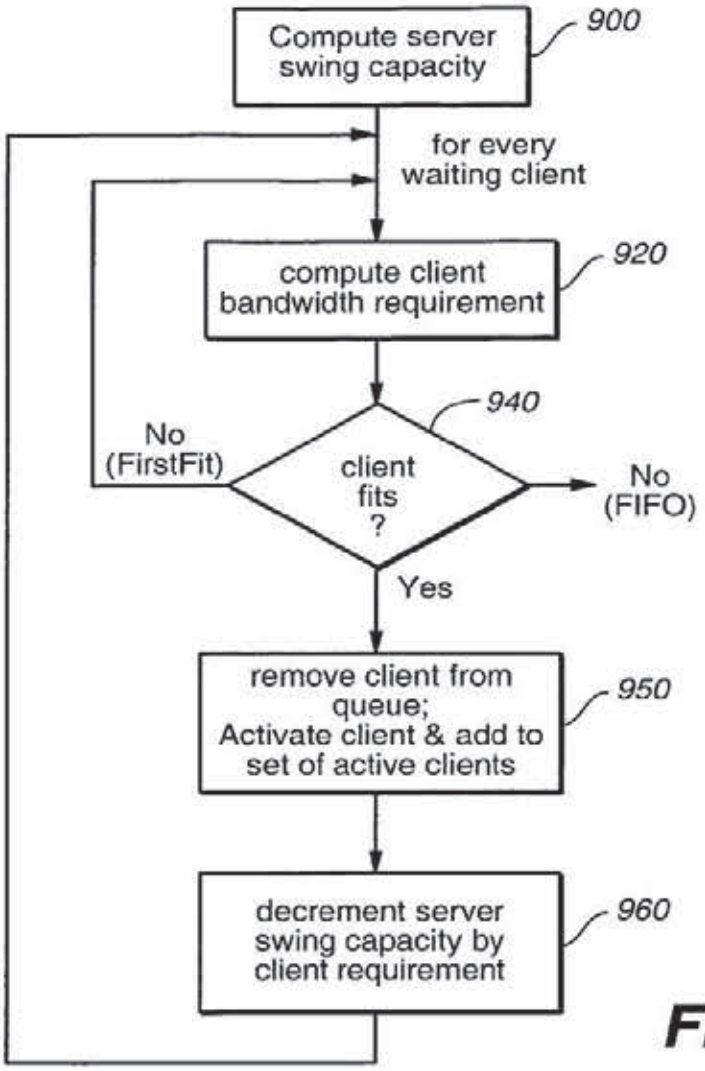

Compute server swing capacity
900
for every waiting client
compute client bandwidth requirement
920
940
No (FirstFit)
client fits ?
No (FIFO)
Yes
remove client from queue; Activate client & add to set of active clients
950
decrement server swing capacity by client requirement
960

FIG._9

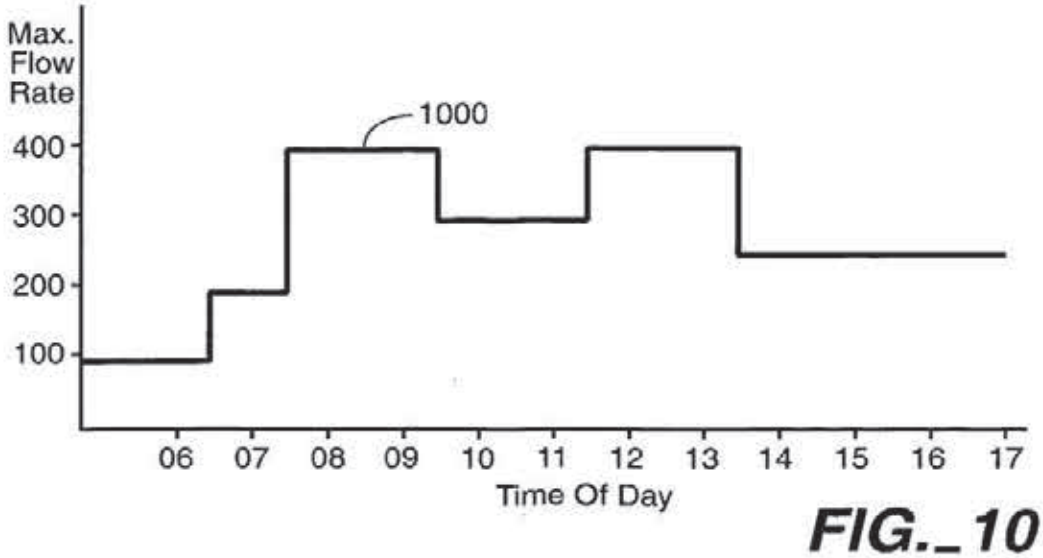

Max. Flow Rate
1000
400
300
200
100
06
07
08
09
10
11
12
13
14
15
16
17
Time Of Day

FIG._10

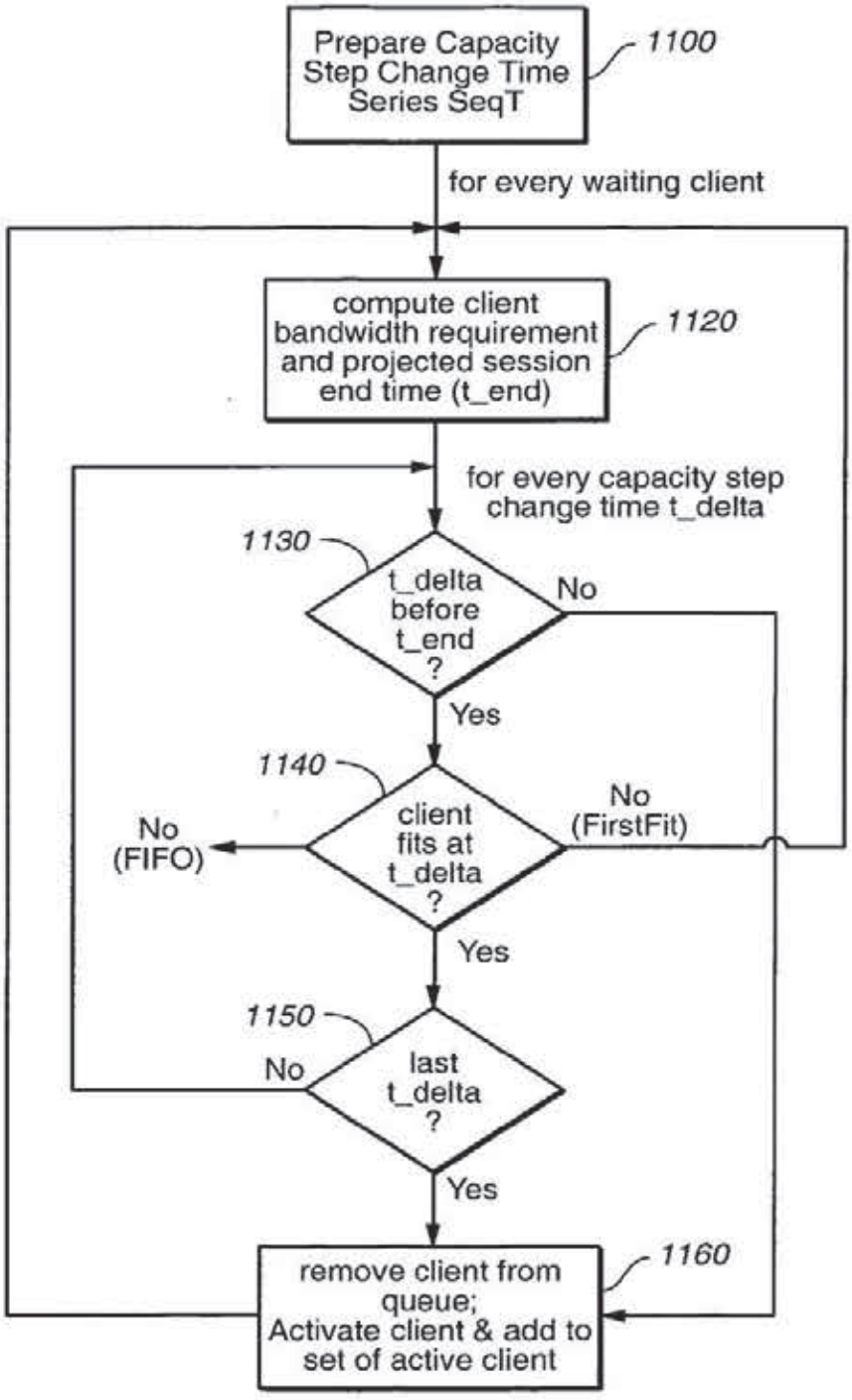

Prepare Capacity Step Change Time Series SeqT
1100
for every waiting client
compute client bandwidth requirement and projected session end time (t_end)
1120
for every capacity step change time t_delta
1130
t_delta before t_end ?
No
Yes
1140
client fits at t_delta ?
No (FIFO)
No (FirstFit)
Yes
1150
last t_delta ?
No
Yes
remove client from queue; Activate client & add to set of active client
1160

FIG._11

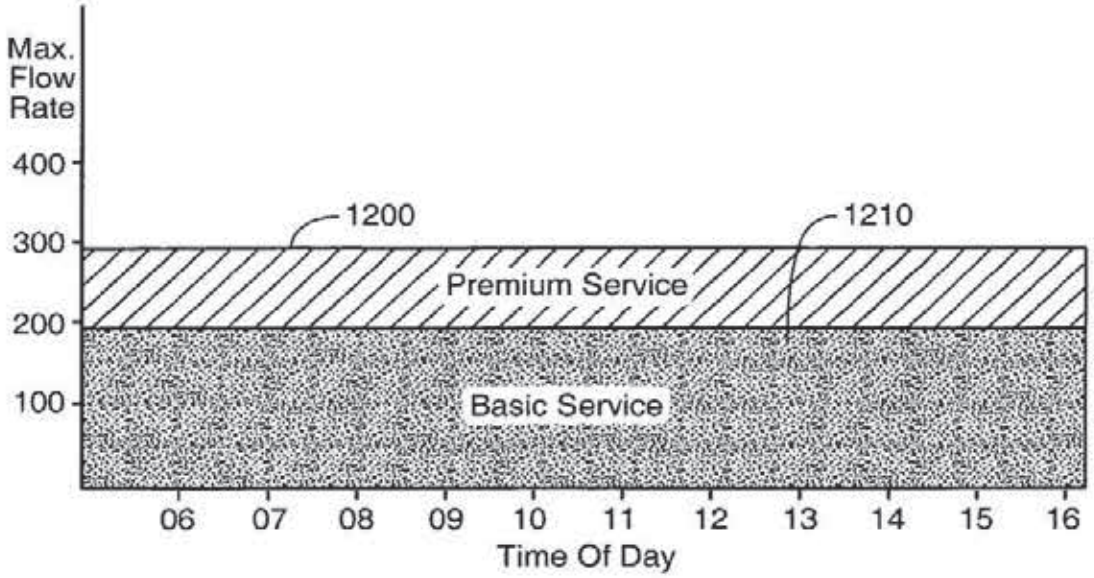

Max.
Flow
Rate
400
300
200
100
1200
1210
Premium Service
Basic Service
06
07
08
09
10
11
12
13
14
15
16
Time Of Day

FIG._12

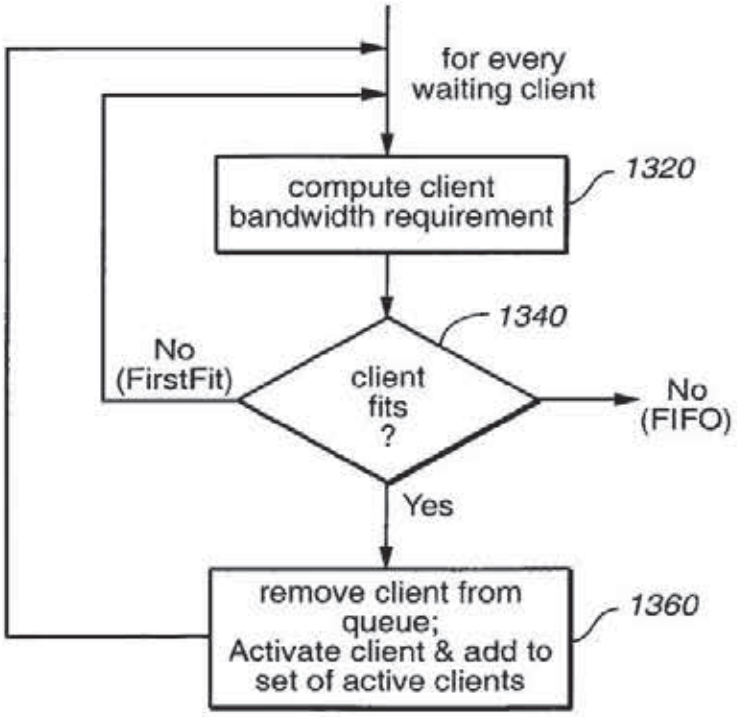

for every
waiting client
compute client
bandwidth requirement
1320
1340
No
(FirstFit)
client
fits
?
No
(FIFO)
Yes
remove client from
queue;
Activate client & add to
set of active clients
1360

FIG._13

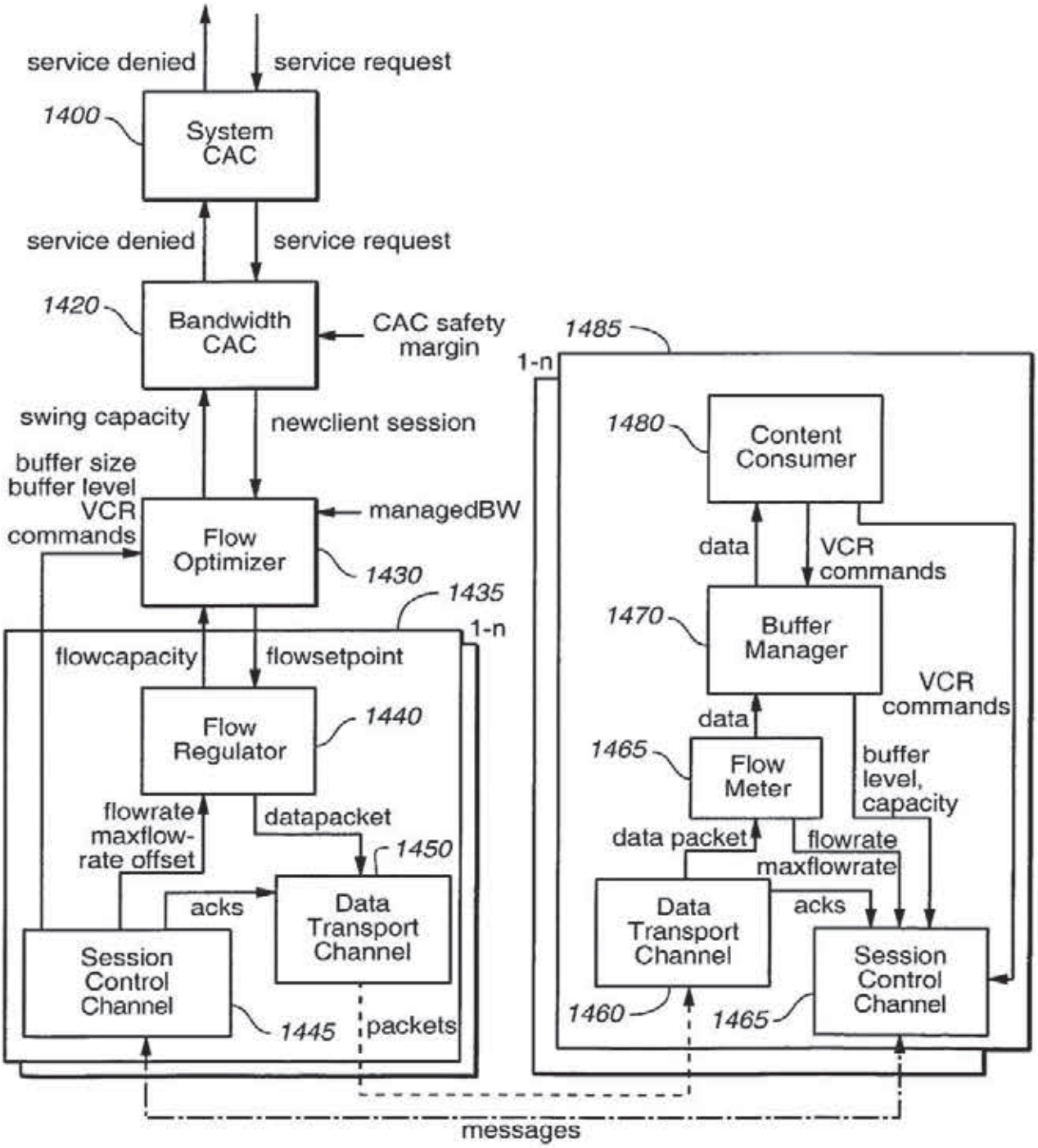

service denied
service request
1400
System CAC
service denied
service request
1420
Bandwidth CAC
CAC safety margin
swing capacity
newclient session
buffer size
buffer level
VCR commands
Flow Optimizer
managedBW
1430
1435
1-n
flowcapacity
flowsetpoint
Flow Regulator
1440
flowrate maxflow-rate offset
datapacket
1450
acks
Data Transport Channel
Session Control Channel
1445
packets
1485
1-n
1480
Content Consumer
data
VCR commands
1470
Buffer Manager
VCR commands
data
1465
Flow Meter
buffer level, capacity
data packet
flowrate maxflowrate
Data Transport Channel
acks
Session Control Channel
1460
1465
messages

FIG._14

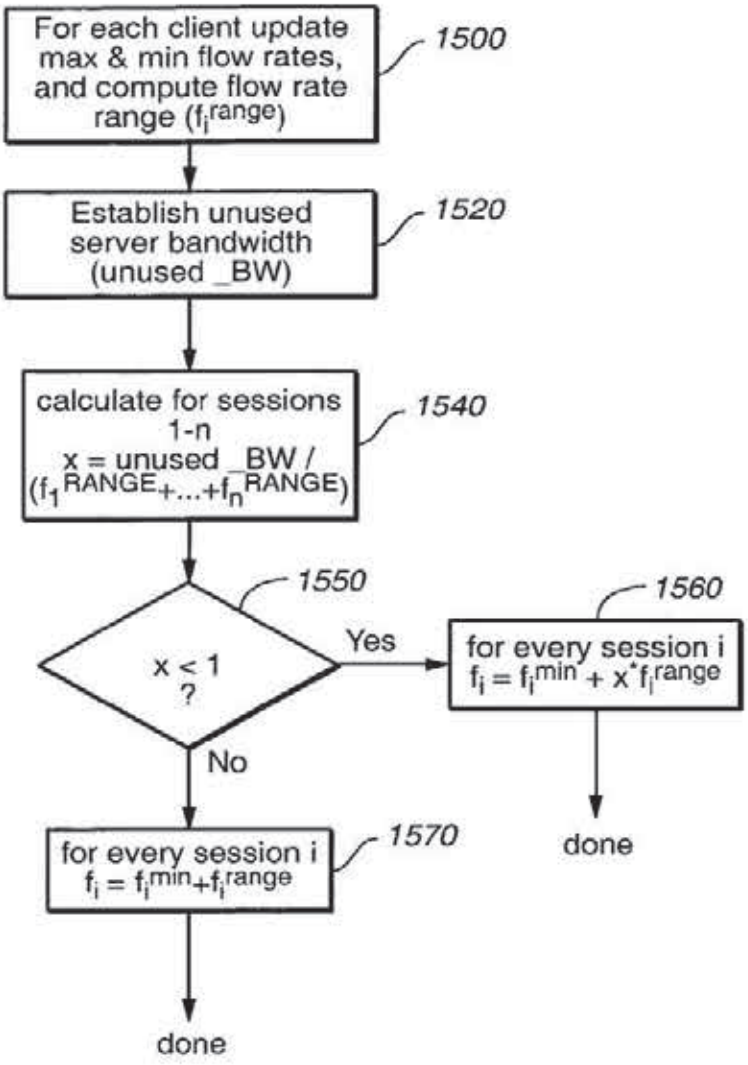

For each client update max & min flow rates, and compute flow rate range (f_i^{range})
1500
Establish unused server bandwidth (unused _BW)
1520
calculate for sessions 1-n
x = unused _BW / (f_1^{RANGE}+...+f_n^{RANGE})
1540
1550
x < 1 ?
Yes
1560
for every session i
$f_i = f_i^{min} + x^* f_i^{range}$
done
No
for every session i
$f_i = f_i^{min} + f_i^{range}$
1570
done

FIG._15

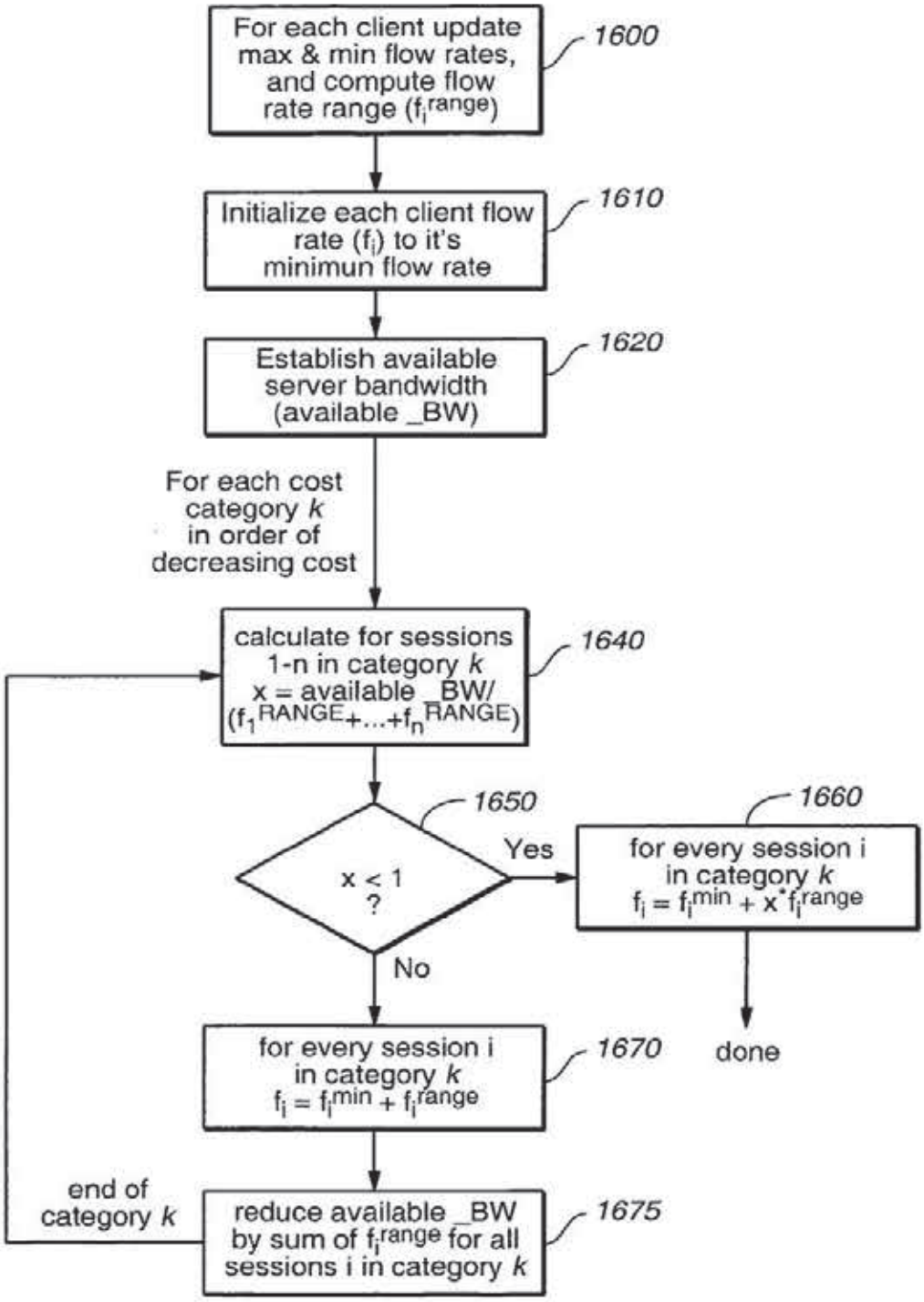


FIG._16

Ex. 1018-1690

Ex. 1018-1691

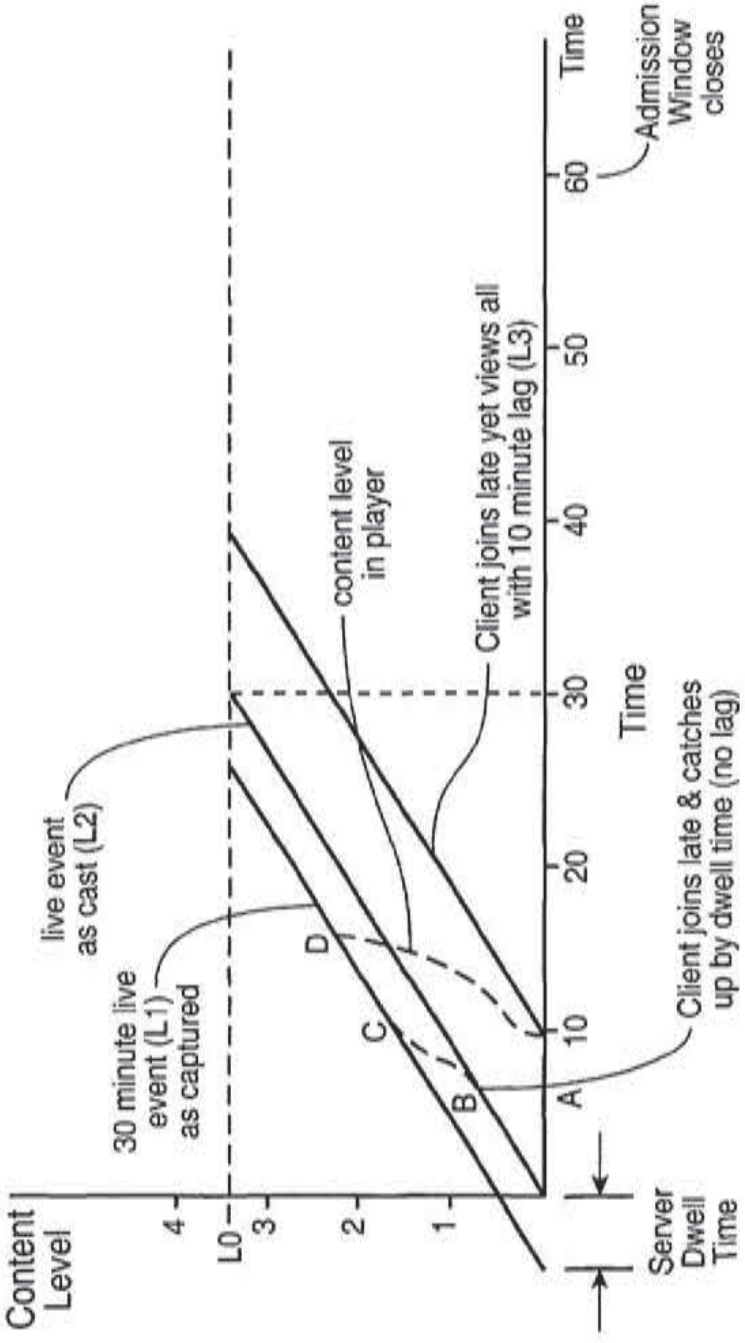


FIG._17

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 January 2006 (26.01.2006)

PCT

(10) International Publication Number
WO 2006/010113 A2

(51) International Patent Classification:
G06F 15/173 (2006.01)

(21) International Application Number:
PCT/US2005/024517

(22) International Filing Date: 11 July 2005 (11.07.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/586,952 9 July 2004 (09.07.2004) US

(71) Applicant *(for all designated States except US)*: **NETWORK FOUNDATION TECHNOLOGIES, LLC** [US/US]; 2215 Cooktown Road, Ruston, LA 71270 (US).

(72) Inventors: **O'NEAL, Mike**; 2215 Cooktown Road, Ruston, LA 71270 (US). **TALTON, John P.**; 2215 Cooktown Road, Ruston, LA 71270 (US).

(74) Agents: **LANDA, Adam B.** et al.; Greenberg Traurig, MetLife Building, 200 Park Avenue, New York, NY 10166 (US).

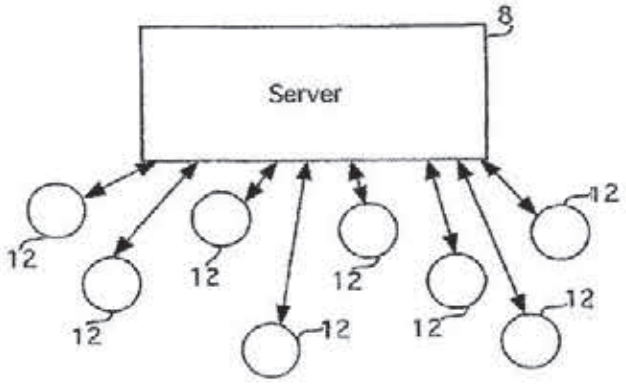
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK,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

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.

(54) Title: SYSTEMS FOR DISTRIBUTING DATA OVER A COMPUTER NETWORK AND METHODS FOR ARRANGING NODES FOR DISTRIBUTION OF DATA OVER A COMPUTER NETWORK



PRIOR ART

(57) Abstract: Various embodiments of the present invention relate to a system for distributing data (e.g., content data) over a computer network and a method of arranging receiver nodes in a computer network such that the capacity of a server is effectively increased (e.g., the capacity of a server may be effectively multiplied many times over; the capacity of the server may be effectively increased exponentially). In one embodiment the present invention may take advantage of the excess capacity many receiver nodes possess, and may use such receiver nodes as repeaters. The distribution system may include node(s) having database(s) which indicate ancestor(s) and/or descendant(s) of the node so that reconfiguration of the distribution network may be accomplished without burdening the system's primary server. An embodiment of the present invention may include a process for configuring a computer information distribution network having a primary server node and user nodes docked in a cascaded relationship, and reconfiguring the network in the event that a user node departs from the network. In one example (which example is intended to be illustrative and not restrictive), the process may include the steps of providing a new user node (or connection requesting user node) with a connection address list of nodes within the network, having the new user node (or connection requesting user node) go to (or attempt to go to) the node at the top of the connection address list, determine whether that node is still part of the distribution network, and connect thereto if it is, and if it is not, to go to (or attempt to go to) the next node on the connection address list. In another example (which example is intended to be illustrative and not restrictive), when a user node departs from the distribution network, a propagation signal may be transmitted to the nodes below it in the network, causing them to move up in the network in a predetermined order. In another example (which example is intended to be illustrative and not restrictive), the present invention may provide a decentralized approach which provides, to each new user node (or connection requesting user node) a path back to the root server.

WO 2006/010113 A2

SYSTEMS FOR DISTRIBUTING DATA OVER A COMPUTER NETWORK AND METHODS FOR ARRANGING NODES FOR DISTRIBUTION OF DATA OVER A COMPUTER NETWORK

FIELD OF THE INVENTION

This patent application is related to and claims the benefit of the filing date of U.S. Provisional Patent Application Serial No. 60/586,952 for "Systems for Distributing Data Over A Computer Network And Methods For Arranging Nodes For Distribution Of Data Over A Computer Network", filed July 9, 2004.

Various embodiments of the present invention relate to systems for distributing data (e.g., content data) over a computer network and methods of arranging nodes for distribution of data (e.g., content data) over a computer network. In one example (which example is intended to be illustrative and not restrictive), the systems and methods of the present invention may be applied to the distribution of streaming audiovisual data over the Internet.

For the purposes of the present application the term "node" (e.g., as used in the phrase a first node communicates with a second node) is intended to refer to a computer system (or sometimes just "computer") configured to send and/or receive data over a computer network (e.g., a computer network including the Internet, a local-area-network, a wide-area-network, a wireless network). Since a "node" in a computer network would not exist but for a computer system being available on such node, the terms "node" and "computer system" (or sometimes just "computer") may be used interchangeably (i.e., the term "node" should be understood to include an available computer system (or sometime just "computer").

Further, for the purposes of the present application the term "upstream" (e.g., as used in the phrase a first computer system sends data via an upstream connection to a second computer system) is intended to refer to the communication path between a first computer system and the Internet when the first computer system is sending data to a second computer system via the Internet.

Further still, for the purposes of the present application the term "downstream" (e.g., as used in the phrase a first computer system receives data via a downstream connection from a second computer system) is intended to refer to the communication path between a first computer system and the Internet when the first computer system is receiving data from a second computer system via the Internet.

Further still, for the purposes of the present application the term "uptree" (e.g., as

used in the phrase a first node sends data via an uptree connection to a second node) is intended to refer to the network topology communication path between a first node and a second node when a first node is sending data to a second node which is higher up on the network topology tree (that is, closer to the root server).

Further still, for the purposes of the present application the term "downtree" (e.g., as used in the phrase a first node sends data via a downtree connection to a second node) is intended to refer to the network topology communication path between a first node and a second node when a first node is sending data to a second node which is lower down on the network topology tree (that is, closer to the leaves).

Further still, for the purposes of the present application the term "docked" (e.g., as used in the phrase a child node docks with a parent node) is intended to refer to forming a connection between nodes via which data may flow in at least one direction (e.g., at least uptree or downtree).

Further still, for the purposes of the present application the term "apparent available capacity" (e.g., as used in the phrase apparent available capacity to transmit content data) is intended to refer to a nominal capacity (e.g., a node's presumed capacity based upon a design specification but not necessarily actually tested in the network).

BACKGROUND OF THE INVENTION

In a computer network such as the Internet, each node in the network has an address. A computer system resident at a particular address may have sufficient bandwidth or capacity to receive data from, and to transmit data to, many other computer systems at other addresses. An example of such a computer system is a server, many commercial versions of which can simultaneously exchange data with thousands of other computer systems.

A computer system at another location may have only sufficient bandwidth to effectively exchange data with only one other computer system. An example of such a system is an end user's personal computer connected to the Internet by a very low speed dialup modem. However, even typical personal computers connected to the Internet by higher speed dialup modems may have sufficient bandwidth that they may exchange data essentially simultaneously in an effective manner with several other computer systems. Moreover, an end user's personal computer system may have even greater bandwidth when connected to the Internet by ISDN lines, DSL (e.g., ADSL) lines, cable modems, T1 lines or even higher capacity links. As discussed more fully below, various embodiments of the present invention

may take advantage of the availability of such higher capacity end user systems (e.g., those computer systems capable of essentially simultaneously exchanging data with multiple computer systems).

In a typical situation, as shown in FIG. 1, a content provider distributes its data by making the data available on a server node 8 simultaneously to a plurality of users at user nodes 12 (of note, the terms "server", "root server" and "primary server" may be used interchangeably throughout the present application to refer to the same device (i.e., the highest level parent node in a given network). The double-headed arrows show the two-way communication between each end user's computer system and the server. Essentially the content provider's server transmits a separate stream of signals to each receiver node. To accommodate additional users, the content provider would typically either add equipment to increase capacity or it would engage a mirror site to accomplish essentially the same result as adding equipment. The capacities of the end user computers is of virtually no consequence in such a system.

Another system for distributing data is the Napster™ music file exchange system provided by Napster, Inc. of Redwood City, Calif. A schematic of the Napster™ music file exchange system (as its operation is presently understood) is illustrated in FIG. 2.

More particularly, it is believed that in the Napster™ system a copy of the music data is not kept on the server. The server 9 instead maintains a database relating to the various music files on the computers of users who are logged onto the server 9. When a first user 12a sees that a desired music file is available from a second logged on user 12b, the first user causes his computer to query the server 9 for the second user's node address and a connection is made between the first and second user's computers through which the first user's computer notifies the second user's computer of the desired file and the second user's computer responds by transmitting a copy of the desired music file directly to the first user's computer. It is further believed that a first user attempting to download a particular file from a second user must start completely over again if the second user cancels its transmission or goes off line during the data transfer.

In another area, engineers have developed what is known as "streaming media." In summary, streaming media is a series of packets (e.g., of compressed data), each packet representing moving images and/or audio.

To help understand streaming media in more detail, it is helpful to review the traditional Internet distribution method. Each node (whether it is a server node or a user node) in a computer network has a unique identification (sometimes referred to as an "IP" address)

associated with it. On the Internet, the unique address may be referred to as a Uniform Resource Locator ("URL"). A user desiring to obtain data from a particular server enters that server's URL into the user's browser program. The browser program causes a connection request signal to be sent over the Internet to the server. If the server has the capacity to accept the connection, the connection is made between the server and the user node (files requested by the user are typically transmitted by the server in full to the user node and the browser program may store the files in buffer memory and display the content on the user's computer system monitor - some files may be more permanently stored in the computer system's memory for later viewing or playing.) The connection with the server is typically terminated once the files have been received at the user node (or the connection may be terminated a short time thereafter). Either way, the connection is usually of a very short time duration.

With streaming media, the contact between the server and user nodes is essentially continuous. When a connection between a server node and user node is made and streaming media is requested, the server sends streaming media packets of data to the user node. A streaming media player installed on the user's computer system (e.g., software, such as RealMedia™ from RealNetworks, Inc. of Seattle, Wash.,) causes the data to be stored in buffer memory. The player decompresses the data and begins playing the moving images and/or audio represented by the streaming media data on the user's computer system. As the data from a packet is played, the buffer containing that packet is emptied and becomes available to receive a new packet of data. As a result, the memory assets of a user's computer are not overly taxed. Continuous action content, such as, for example, the display of recorded motion picture films, videos or television shows may be distributed and played in essentially "real time," and live events, such as, for example, concerts, football games, court trials, and political debates may be transmitted and viewed essentially "live" (with only the brief delays needed for compression of the data being made available on the server, transmission from the server to the user node, and decompression and play on the user's computer system preventing a user from seeing the event at the exact same moment in time as a person actually at the event). And, when the systems are working as designed, the server node and user node may stay connected to each other until all the packets of data representing the content have been transmitted.

SUMMARY OF THE INVENTION

Various embodiments of the present invention relate to a system for distributing data (e.g., content data) over a computer network and a method of arranging receiver nodes in a computer network such that the capacity of a server is effectively increased (e.g., the capacity of a server may be effectively multiplied many times over; the capacity of the server may be effectively increased exponentially). In one embodiment the present invention may take advantage of the excess capacity many receiver nodes possess, and may use such receiver nodes as repeaters. The distribution system may include node(s) having database(s) which indicate ancestor(s) and/or descendant(s) of the node so that reconfiguration of the distribution network may be accomplished without burdening the system's primary server. An embodiment of the present invention may include a process for configuring a computer information distribution network having a primary server node and user nodes docked (i.e., connected) in a cascaded relationship, and reconfiguring the network in the event that a user node departs from the network. In one example (which example is intended to be illustrative and not restrictive), the process may include the steps of providing a new user node (or connection requesting user node) with a connection address list of nodes within the network, having the new user node (or connection requesting user node) go to (or attempt to go to) the node at the top of the connection address list, determine whether that node is still part of the distribution network, and connect thereto if it is, and if it is not, to go to (or attempt to go to) the next node on the connection address list. In another example (which example is intended to be illustrative and not restrictive), when a user node departs from the distribution network, a propagation signal may be transmitted to the nodes below it in the network, causing them to move up in the network in a predetermined order. In another example (which example is intended to be illustrative and not restrictive), the present invention may provide a decentralized approach which provides, to each new user node (or connection requesting user node) a path back to the root server.

Of note, various embodiments of the present invention may be applied to the transmission (e.g., the "appointment" transmission) of audiovisual content such as, for example (which example is intended to be illustrative and not restrictive), live or pre-recorded concerts, football games, court trials, political debates, motion picture films, videos or television shows. Such transmission of audiovisual content may be made with acceptable levels of quality, wherein "acceptable levels of quality" may vary depending upon the end users and the type of transmission. Further, such transmission of audiovisual content may be

Ex. 1018-1697

carried out using streaming media transmissions intended to reach large audiences, in much the way that television shows transmitted over television cable and broadcast media reach large audiences (in this regard, the present invention may enable a server to transmit streaming media to the large number of users which would be essentially simultaneously logging on to view a particular audiovisual presentation under this television-type "large audience" transmission).

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic drawing of a prior art computer information distribution network;

FIG. 2 is a schematic drawing of another prior art computer information distribution network;

FIG. 3 is a schematic drawing of an embodiment of a computer information distribution network formed pursuant to the present invention;

FIG. 4 is a schematic drawing of another embodiment of a computer information distribution network formed pursuant to the present invention;

FIG. 5 is a schematic drawing of another embodiment of a computer information distribution network formed pursuant to the present invention;

FIG. 6 is a schematic drawing of a particular topology of a computer information distribution network formed pursuant to an embodiment of the present invention;

FIG. 7 is a schematic drawing of a particular topology of the computer information distribution network formed pursuant to an embodiment of the present invention as shown in FIG. 6 (after the occurrence of an event);

FIG. 8 is a schematic drawing of another topology of a computer information distribution network formed pursuant to an embodiment of the present invention;

FIG. 9 is a schematic drawing of the topology of the computer information distribution network formed pursuant to an embodiment of the present invention as shown in FIG. 8 (after the occurrence of an event);

FIG. 10 is a flow diagram of an embodiment of the present invention showing a Server Connection Routine which may be performed when a prospective child node seeks to join the distribution network;

FIG. 11 is a flow diagram of an embodiment of the present invention showing a Prospective Child Node's Request to Server for a Connection Routine;

FIGS. 12A-12E are schematic drawings showing varying topologies of the computer information distribution network formed pursuant to an embodiment of the present invention under several circumstances;

FIG. 13 is a block diagram of an embodiment of the present invention showing the memory blocks into which the software used may partition a user node's memory;

FIG. 14 is a flow diagram of an embodiment of the present invention showing a Prospective Child Node's Connection Routine - the routine which a new user node (or connection requesting user node) may go through in attempting to connect to a distribution chain (or tree) after receiving a connection address list;

FIG. 15 is a flow diagram of an embodiment of the present invention showing a Prospective Child Node's Connection Routine With Return to Server Subroutine;

FIG. 16 is a flow diagram of an embodiment of the present invention illustrating a Prospective Parent Node's Connection Routine;

FIG. 17 is a flow diagram of an embodiment of the present invention illustrating a Server's Connection Instruction Routine;

FIG. 18 is a flow diagram of an embodiment of the present invention illustrating a Fully Occupied Parent's Connection Instruction Routine;

FIG. 19 is a flow diagram of an embodiment of the present invention illustrating a Multi-Node Selection Subroutine;

FIGS. 20A and 20B are together a flow diagram of an embodiment of the present invention illustrating a Universal Connection Address Routine;

FIG. 21 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention before a new node will be added using the Universal Connection Address Routine;

FIG. 22 is a schematic drawing of a topology of the computer information distribution network of an embodiment of the present invention when a new node is added using the Universal Connection Address Routine;

FIG. 23 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention before a reconfiguration event;

FIG. 24 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention shown in FIG. 23 after a reconfiguration event;

Ex. 1018-1699

FIG. 25 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention, slightly different from the topology shown in FIG. 23, before a reconfiguration event;

FIG. 26 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention shown in FIG. 25 after a reconfiguration event;

FIG. 27 is a flow diagram of an embodiment of the present invention showing a Child Node's Propagation Routine;

FIG. 28 is a schematic drawing of another topology of a computer information distribution network of an embodiment of the present invention before a reconfiguration event;

FIG. 29 is a schematic drawing of a topology of a computer information distribution network of an embodiment of the present invention shown in FIG. 28 after a reconfiguration event;

FIG. 30 is a schematic drawing of another topology of a computer information distribution network of an embodiment of the present invention before a "complaint" regarding communications;

FIG. 31 is a flow diagram of an embodiment of the present invention showing a Grandparent's Complaint Response Routine.

FIGS. 32, 33A-33C, 34A-34C, 35A, 35B and 36 relate to "Registration Process" examples according to the present invention;

FIGS. 37 and 38 relate to a "Firewall Request / Firewall Status" example according to the present invention;

FIGS. 39A, 39B and 40 relate to a "Propagate Path / Down" example according to the present invention;

FIGS. 41A, 41B and 42 relate to a "Propagate NTM / Up" example according to the present invention;

FIGS. 43 and 44 relate to a "Banned List" example according to the present invention;

FIG. 45 relates to a "Node Interlacing" example according to the present invention; and

FIGS. 46A-46C relate to various distribution network topology examples according to the present invention.

Ex. 1018-1700

Among those benefits and improvements that have been disclosed, other objects and advantages of this invention will become apparent from the following description taken in conjunction with the accompanying figures. The figures constitute a part of this specification and include illustrative embodiments of the present invention and illustrate various objects and features thereof (of note, the same reference numeral may be used to identify identical elements throughout the figures).

DETAILED DESCRIPTION OF THE INVENTION

Detailed embodiments of the present invention are disclosed herein; however, it is to be understood that the disclosed embodiments are merely illustrative of the invention that may be embodied in various forms. In addition, each of the examples given in connection with the various embodiments of the invention are intended to be illustrative, and not restrictive. Further, the figures are not necessarily to scale, some features may be exaggerated to show details of particular components. Therefore, specific structural and functional details disclosed herein are not to be interpreted as limiting, but merely as a representative basis for teaching one skilled in the art to variously employ the present invention.

Referring now to FIG. 3, an illustration of a linear propagation arrangement 10 of a computer network according to an embodiment of the present invention is shown. In this embodiment, the primary server node (or simply, server) 11 provides content data (e.g., streaming media) to user nodes 12 connected directly to it (sometimes referred to as "first level user nodes"). Each first level user node 12 has a second level user node 13 connected to it and each second level user node 13 has a third level user node 14 connected to it. The computer system at each first level user node 12 passes a copy of the content data received from server node 11 to the computer system at the second level user node 13 attached to such first level user node 12. The computer system at each second level user node 13 in turn passes the content data onto the computer system at the fourth level user node 14 attached to it.

As more fully discussed below, under this embodiment the computer systems at the server and user nodes may have distribution software installed in them which enables the nodes to be arranged as shown and for the computer systems to receive and re-transmit data.

The cascadingly connected arrangement of nodes (i.e., first level nodes are connected to the server, second level nodes are connected to first level nodes, third level nodes are connected to second level nodes and so on) shown in FIG. 3 takes advantage of the bandwidth available in certain nodes to essentially simultaneously receive and transmit data.

In a linear propagation arrangement, the effective distribution capacity of a server is in essence multiplied by the number of levels of nodes linked together. In the example of FIG. 3 (which example is intended to be illustrative and not restrictive), the distribution capacity of the server node is increased from 8 user nodes to 24 in just three levels.

Of note, many user nodes may have at least sufficient bandwidth (e.g., upstream bandwidth as well as downstream bandwidth) to receive data from one node and to re-transmit streams of data essentially simultaneously to two or more other nodes. This capacity could be used in another embodiment to set up a computer network with a cascadingly connected exponential propagation arrangement 16 (as shown in FIG. 4). As the name implies, an exponential propagation arrangement effectively increases the distribution capacity of a server exponentially. For example (which example is intended to be illustrative and not restrictive), with just three levels of user nodes, each having the capacity to retransmit two data streams, the distribution capacity of the server in FIG. 4 is increased from 8 user nodes to 56.

In another embodiment a distribution network may also be set up as a cascadingly connected hybrid linear/exponential arrangement 18, such as shown in FIG. 5.

Of note, the effective distribution capacity grows more quickly in a hybrid linear/exponential arrangement with each new level than does the distribution capacity in a linear propagation arrangement, and less quickly than does the distribution capacity in a pure exponential arrangement. However, any of these arrangements allows a server's distribution capacity to be greatly increased (e.g., with little or no additional investment in equipment for the server). Further, any other desired hybrid arrangement may be used.

Referring again to FIGS. 1-5, all of the connections between nodes are illustrated with arrows on each end. This is intended to signify that data flows in both directions between connected nodes. For example (which example is intended to be illustrative and not restrictive), a user node connected to a server may transmit data to the server indicating the identity of the user node, what the user node wants and/or other data, while the server node may transmits data confirming its identity and containing information and other content to the user node. In the remaining drawings such arrows may not be shown for the sake of simplicity. Further, it is noted that the bulk of data which is transmitted may be content (e.g., web pages, music files, streaming media), and such content will be understood to flow downtree from node to node in the drawings (i.e., in the direction from the root server to a first level node, to a second level node, to a third level node, etc.).

Ex. 1018-1702

As can be understood from the above, an exponential propagation arrangement may create the most distribution capacity. However, while many (and perhaps even most) user nodes may have sufficient bandwidth to re-transmit acceptable quality multiple copies of the data, a number of user nodes may not have sufficient bandwidth to re-transmit more than one copy of the data with acceptable quality, and some user nodes may not have sufficient bandwidth to re-transmit even a single acceptable copy. Under these conditions, a distribution system employing the present invention may be configured, for example, as a hybrid propagation network. In such a system, a personal computer acting as a server node may reach many (e.g., hundreds, thousands or more) user nodes, even if the server node itself has capacity to transmit content data directly to only one other node.

In the following discussion it will be assumed that the system employing the present invention will take advantage of the capability of many user nodes to simultaneously re-transmit up to two copies of data (e.g., content data). However, it should be understood that the invention could take advantage of user nodes capable of simultaneous re-transmission of even higher numbers of copies.

In one embodiment the length of the distribution chains (i.e., the number of levels of user nodes linked through each other to the server) may be kept as small as possible to reduce the probability that the nodes at the ends of the chains will suffer a discontinuity of service. Thus, under this embodiment, in order to maximize the number of user nodes which may be connected to the server while trying to keep the length of the distribution chains as small as possible, the user nodes having the greatest bandwidth may be placed as far up uptree as possible (i.e., be placed as close as possible to the server, where "close" refers to network node level and not necessarily to geographical proximity).

In this regard, if an unreliable user node is placed near or at the end of a chain, few or no other user nodes would be affected by the unreliable user node's leaving the network. In contrast, if an unreliable user node were placed at or near the beginning of a chain (i.e., be far uptree), many other nodes would be affected by the unreliable node's departure. Because various embodiments of the invention may be used for appointment viewing of streaming media, in which an essentially continuous connection to an audiovisual content source may essentially be a necessity, it may be important (e.g., for this steaming media use) that the most reliable user nodes be placed high up in the chain (i.e., high uptree).

A user node may be deemed unreliable for any of a number of reasons. One example (which example is intended to be illustrative and not restrictive), is that the user node is connected to the Internet by lines having intermittent failures. Another example (which

example is intended to be illustrative and not restrictive), is that the user is merely sampling the content available on the network and intentionally disconnects the user node from the distribution network after discerning that he or she has no interest in the content.

Thus, under an embodiment of the present invention the user nodes may be positioned in the most advantageous positions, taking into account the dynamic nature of the distribution network, in which many user nodes may enter and leave the distribution network throughout the server's transmission of a streaming media show. In addition, the invention may help to preserve the viewing experience of users at user nodes positioned even at the end of a distribution chain (the server and/or user nodes may be enabled to perform the operations required to set up and maintain the distribution network by having data distribution software installed therein).

Referring now to a "User Node Arrival" example (which example is intended to be illustrative and not restrictive), it can be seen from FIGS. 3-5 that the distribution chains can be viewed as a plurality of family trees, each family tree being rooted to the server 11 through one of the first level nodes 12 (each node in the distribution network may have distribution software loaded in it which enables the node to perform the functions described below; before any new node may join the distribution network, such new node may also have such software loaded in it).

Of note, in the bulk of the examples discussed below, the network topologies have a branch factor of two (i.e., no user node is assigned more than two child nodes to be connected directly to it). A network topology with a branch factor of two may be referred to as a binary tree topology. It should be understood, of course, that the teachings set forth herein may be extended to network topologies having other branch factors, for example (which example is intended to be illustrative and not restrictive), branch factors of three, four or more.

In any case, FIG. 6 is a schematic drawing of a distribution network topology useful for describing the current "User Node Arrival" example. This FIG. 6 shows a distribution chain or family tree rooted to the server 11 through user node A, a first level user node 12 (the dashed lines represent connections from the server to other first level nodes or from node A to another user node). User node A could be thought of as a child node of the server and as a parent node for other user nodes connected directly to it. User node B, a second level user node 13, could be thought of as A's child. User nodes C and D, third level user nodes 14, may be thought of as B's children and A's grandchildren (and each other's siblings). User nodes E and F, fourth level user nodes 15, may be thought of as C's children (and each other's

siblings). User node G, also a fourth level user node 15, may be thought of as D's child. And user nodes E, F and G may be thought of as B's grandchildren and A's great grandchildren.

In the present example, whenever a new user node (or connection requesting user node) 19, such as node X in FIG. 6, seeks connection to the distribution network, it will first make a temporary connection to the server node (or a connection server, not shown) in order to begin the process for connecting to the distribution system. The server (or connection server) will discern from the user node a bandwidth rating (discussed below) appropriate to that node and, depending upon the available capacity of the server and any existing distribution chains, the server will either assign the new user node to a spot directly connected to the server or will provide the new user node with a connection path through a tree to the server.

FIG. 10 is a flow diagram associated with this example showing a Server's Connection Routine which is performed when a prospective child node seeks to join the distribution network. In step 101 the server performs the Server's Connection Instruction Routine (discussed below), in which the server determines what connection instructions to give to the new user node (or connection requesting user node). The server then goes to step 102 where it determines whether, as a result of the Server's Connection Instruction Routine, the prospective child node is being instructed to dock with the server. If the prospective child node is being instructed to dock with the server, then the server goes to step 103 in which the server would allow the new user node to dock with it, and the server would begin transmitting data (e.g., streaming media) directly to the new user node.

Of note, as referred to above, two different servers could be used - one to perform the server's connection routine and the other to transmit data (e.g., streaming media) - since both servers would be performing server functions, they will hereinafter sometimes be considered a single server for purposes of the description herein.

Of further note, while the prospective child node may be a new user node or a connection requesting user node, to simplify the following discussion such prospective child node may simply be referred to as a "a new user node".

In any case, if the new user node is not being instructed to dock directly with the server, then the server goes to step 104 in which it provides the new user node with an address connection list and disconnects the new user node from the server.

Essentially contemporaneously with the performance by the server of the Server's Connection Routine, the new user node is performing the Prospective Child Node's Request to Server for a Connection Routine. FIG. 11 is a flow diagram associated with this example

illustrating the Prospective Child Node's Request to Server for a Connection Routine. Upon making the temporary connection to the server, the new user node goes to step 111 in which it provides bandwidth rating and performance rating information to the server. It then proceeds to step 112 in which it receives connection instructions from the server. Then the new user node proceeds to step 113 to determine whether it has been instructed to dock directly with the server. If the answer is "yes," then the new user node proceeds to step 114 in which it erases any information which may have been in its ancestor database and, if the distribution software has a Return to Server Subroutine in it, resets the connection attempt counter to zero. The new user node then proceeds to step 115 in which it docks with the server and begins receiving data (e.g., streaming media). If the new user node is not being instructed to dock directly with the server, then the new user node goes to step 116 in which it receives the new connection address list from the server and loads such list into the user node's ancestor database and begins the Prospective Child Node's Connection Routine (discussed below). If the distribution software has a Return to Server Subroutine in it, the connection attempt counter is reset to one in step 116.

In the examples discussed in connection with FIGS. 6-9 and 12A-12C (which examples are intended to be illustrative and not restrictive), the server has determined that the new user node X will not be allowed to connect directly to the server. Also, for the purposes of these examples, all of the user nodes are presumed to be equally capable of simultaneously re-transmitting two copies of the content data and that the tree rooted through node A is the most appropriate tree through which node X should be connected to the server. In one embodiment the server may rely on chain length in determining to which particular user node, already in the distribution network, that node X should connect.

Referring again to the example of FIG. 6, the topology of the distribution network based on the most recent information available to the server at the moment that node X seeks to join the distribution network is shown. Path D-B-A-S (where "S" is the server) represents the shortest available path from an end of a chain back to the server (e.g., from a node level point of view and not necessarily from a geographical proximity point of view), and the server 11 gives that path information, or connection address list, to node X during step 104 of FIG. 10 (and node X receives such list during step 116 of FIG. 11). That is, node X will be given a connection address list with the URLs and/or IP addresses of each of nodes D, B, A and S. The distribution software in node X causes the path information to be stored in the ancestor portion (or ancestor database) 132 of node X's topology database 131 shown in FIG. 13. The ancestor database may include an ancestor list, a list of node X's ancestors' addresses

from node X's parent back to the server node. (FIG. 13 shows an example block diagram (which example is intended to be illustrative and not restrictive) showing the memory blocks into which the software used in connection with the present invention may partition a user node's memory 130.) Node X then attempts to contact node D first, the user node most distant from server 11 in the path. Note that when the server provides the D-B-A-S connection address to node X, the server is giving what it "believes" to be the complete path information going all the way back to the server. That is, subsequent to the most recent generation of the topological data, node D may have departed from the network, as may have one or more of its ancestors, resulting in a reconfiguration (discussed below) of at least a portion of the tree of which D was a part.

Referring now to FIG. 14 is a flow diagram associated with this example showing the Prospective Child Node's Connection Routine, the routine which a new user node, here node X, will go through in attempting to connect to a distribution chain (or tree) after receiving a connection address list (which node X has stored in the ancestor portion of its topology database) from the server or from a prospective parent node or during a reconfiguration event. In step 141, node X attempts to contact the first node on the connection address list. The first node, and only node, on the connection address list could be the server itself. Here node D is the first node on the list. Node X then proceeds to step 142 and determines whether the first node on the connection address list is still on-line and still part of the distribution network (in one example, if no response is received within a predetermined period of time, from the first node on the connection address list, the answer to the query in step 142 will be deemed to be no). If node D is on-line and still part of the distribution network, node X proceeds to step 143 in which node X inquires whether node D has room for node X. This inquiry may need to be made because the distribution network may have gone through a reconfiguration event resulting in node D's not having sufficient capacity to provide a copy of the content data to node X. If node D has room for node X, then node X proceeds to step 144 in which it connects (or docks) with node D and begins receiving content data from it. This is depicted in FIG. 7. Note that node X is now one of several level four nodes 15.

On the other hand, in step 142, if node D is not on-line (e.g., no response is received from node D within a predetermined period of time) or if node D is on-line but is no longer part of the distribution system (e.g., subsequent to the server's obtaining its topology data the user of the system at node D either caused his or her computer system to go off-line or to leave the distribution system, or there was a line failure causing the computer system to go off line), as depicted in FIG. 8, then node X goes to step 145 in which it deletes the first

address from the connection address list in node X's ancestor database (and, in one example, and for a purpose which will become clear when discussing reconfiguration events below, sets its reconfiguration flag buffer 136 to the lowest propagation rating). At this time node B, in the present example, becomes the first node on the connection address list. Then node X goes back to step 141 and repeats the routine described above. Note that because of node D's leaving the distribution network, a reconfiguration event was triggered which resulted in node G changing from a fourth level node 15 to a third level node 14.

In step 143, if the prospective parent node has no room for the new node (e.g., the capacity of the prospective parent node is fully occupied), the new node goes to step 146, in which it receives a new connection address list from the prospective parent. The prospective parent may then perform a Fully Occupied Parent's Connection Instruction Routine, discussed below in connection with FIG. 18, wherein it creates the new connection address list based on topology data obtained from its progeny. That new list may include the path back to the server through the prospective parent just in case, as discussed above, there are user node departures along the path.

Of note, by having a prospective parent node prepare the new connection address list as described in this example, the burden on the server is reduced and is distributed among the user nodes.

Still referring to the example depicted in FIG. 8, when node X gets to step 143, node B will respond that it has no room and node X will proceed to step 146. When node X performs step 146, the new connection address list it receives from node B will be G-B-A-S. Then node X proceeds to step 141 and repeats the routine from that point on. When the routine is performed on the topology shown in FIG. 8, step 143 will result in node G responding that it has room for node X. Node X will then perform step 144 and be connected to the distribution network through node G as shown in FIG. 9. Here, node X becomes a fourth level node 15.

The distribution software may include a Return to Server Subroutine, comprised of steps 151, 152 and 153 as shown in FIG. 15, as part of the Prospective Child Node's Connection Routine. This subroutine reduces the risk that a prospective child node would enter into an endless cycle of fruitless attempts to dock with nodes in a particular tree. If the answer to the query in step 143 is "no," then node X goes to step 151 in which it increments by one the connection attempt counter 135 in node X's memory. Then node X goes to step 152 in which it determines whether the connection attempt limit has been reached. In one example (which example is intended to be illustrative and not restrictive), the limit may be

preset at any number greater than one and may depend upon what the designer of a particular distribution network determines would be a reasonable amount of time for a node to attempt to make a connection on a particular tree, or a branch of a tree, before that node should be given an opportunity to obtain a completely new connection address list directly from the server. If the connection attempt limit has not been reached, then node X proceeds with step 146 as discussed above. If the connection attempt limit has been reached, then node X goes to step 153, in which it goes back to the server and begins the connection routine again as discussed above in connection with FIG. 6. If docking with a parent node is successful, then after step 144 node X performs step 154 in which the connection attempt counter is set to zero.

Referring now to FIG. 16, a flow diagram associated with this example illustrating the Prospective Parent Node's Connection Routine is shown. Using the example illustrated in FIGS. 6, 8 and 9, when node X queries node B in step 142 during the Prospective Child Node's Connection Routine, node B may begin performing the Prospective Parent Node's Connection Routine. In step 161, in response to node X's query, node B determines whether it is part of the distribution system node X seeks to join. If node B were not part of the distribution network, it would respond in the negative to node X's query and node B would be finished with the Prospective Parent Node's Connection Routine. In the example of FIG. 8, the answer is "yes" and node B proceeds to step 162.

In step 162 node B determines whether it has room for a new node. If the answer were "yes," node B would proceed to step 163 where it would allow node X to dock with it, and node B would begin transmitting to node X data (e.g., streaming media) originating from the server. In the example illustrated in FIG. 8, the answer is "no," and node B, acting as an instructing node, goes to step 164 where it performs the Fully Occupied Parent's Connection Instruction Routine (discussed below) and provides the prospective new child node (here node X) with a new connection address list. As noted above, the new connection address list may include the path back to the server through the node B (the prospective parent in this example) in the event that there are user node departures along the path, which departures may include node B.

After performing step 164 the temporary connection between node B and node X is terminated, and node X is sent on its way. In the example of FIGS. 8 and 9, the new connection address list is G-B-A-S. When node X approaches node G, node G performs the Prospective Parent Node's Connection Routine discussed above. In the example illustrated in FIGS. 8 and 9, node G allows node X to dock with it.

Referring now to a "Distribution Network Construction" example (which example is intended to be illustrative and not restrictive), it is noted that under this example the length of the distribution chains (i.e., the number of levels of user nodes linked through each other to the root server) is aimed to be kept as small as possible.

Assuming that all new user nodes have the same re-transmission capability (or disregarding the different re-transmission capabilities which different new user nodes may have), the user nodes would be distributed in this example through the first level until all the direct connection slots to the server were filled. Then as new user nodes sought connection to the distribution network, they would be assigned to connections with the first level nodes 12 until all the slots available on the first level nodes were filled. This procedure would be repeated with respect to each level as more and more new user nodes attempted to connect to the distribution network. In other words, the server, acting as an instructing node, would perform a Server's Connection Instruction Routine in which one step is determining whether there is room on the server for a new user node and, if so, the server would instruct the new user node to connect directly to the server. If there were no room on the server, then the server would perform the step of consulting its network topology database and devising a connection address list having the fewest number of connection links to a prospective parent user node. After performing the Server's Connection Instruction Routine, the server would either allow the new user node to dock directly with the server or send the new user node on its way to the first prospective parent user node on the connection address list.

In this example (which example is intended to be illustrative and not restrictive), a partially occupied potential parent node in a particular level (i.e., a prospective parent node already having a child but still having an available slot for an additional child node) may be preferred over unoccupied (i.e., childless) potential parent nodes on the same level. This may help to keep the number of interior nodes (i.e., the number of nodes re-transmitting data) to a minimum.

Alternatively, in another example (which example is intended to be illustrative and not restrictive), nodes that have zero children may be preferred over nodes that already have one child. While it is true that this may increase the number of repeat nodes, what the inventors have found is that by filling the frontier in an "interlaced" fashion, connecting nodes build up their buffers more quickly (allowing a reduction in the incidence of stream interruptions).

Referring now to FIGS. 12A-12C, these FIGS. illustrate, in this example, what happens when a partially occupied parent node is preferred over a parent childless node as a

destination address for a new user node (assuming all other factors are equal). FIG. 12A is a schematic diagram associated with this example showing a topology wherein nodes C and D, both third level nodes 14, have remaining capacity for one or more child nodes. Server 11 has the choice of sending new user node X to either node C, as shown in FIG. 12B, or node D, as shown in FIG. 12C, without increasing the length of the longest chain in the distribution network. However, user nodes are free to leave the distribution network at any time. With the topology shown in FIG. 12C, if either of nodes C or D leaves the network, there would be an effect on a child node. With the topology shown in FIG. 12B, there is a significant chance that if one of nodes C and D leaves the network, there would be no impact on any child nodes. That is, if it is node D that leaves the network, there is no child which would be made an orphan.

Of note, in discussing the examples illustrated in FIGS. 6-9 and 12A-12C, it has been presumed that all of the user nodes are equally capable of simultaneously re-transmitting two copies of the content data. However, the user nodes actually joining a distribution network will have various capabilities for re-transmitting data. Some will have no reliable capability because of hardware limitations. Other user nodes will have high nominal capability because of their hardware and type of Internet connection, but may have low reliability because of poor line conditions or the vagaries of the desires of the humans actually using the user nodes. For example, a user node having a T1 or faster connection to the Internet has a significantly large bandwidth, but is not reliable if the user (or viewer) is merely sampling the content available on the distribution network.

As noted above, in one example the user nodes having the greatest reliable capability should be placed as high up in a distribution chain as possible (i.e., as far uptree as possible) because they would have the ability to support the greatest number of child nodes, grandchild nodes and so on.

Thus, to differentiate the reliable capabilities of user nodes, a number of factors may be considered. Examples of such factors (which examples are intended to be illustrative and not restrictive), include following:

- One is time, that is, the number of seconds (or other units of time) since the node made its most recent connection to the network. Everything else being equal, a user node which has been continuously connected to the distribution network for a long period of time may be considered under this example to be likely more reliable

Ex. 1018-1711

(either because of line conditions or user interest) than a user node which has been continuously connected to the network for a short period of time.

- Another factor is bandwidth rating, which may be determined, for example, by actual testing of the user node when it first attempts to connect to the server or a parent node (e.g., at initial connection time) or determined by the nominal bandwidth as determined by the type of connection made by the user node to the Internet. In one example (which example is intended to be illustrative and not restrictive), we will describe ratings based on nominal bandwidth as follows:

    - A user node with a 56 Kbits/sec dialup modem connection to the Internet is essentially useless for re-transmission of content data because of its small bandwidth. In this example, such a node is assigned a bandwidth rating of zero (0.0).

    - A user node with a cable modem or DSL (e.g., ADSL) connection to the Internet may be given, in this example, a bandwidth rating of one (1.0), because it is a common type of node and has a nominal upstream transmission bandwidth of 128 Kbits/sec, which is large enough to potentially re-transmit two acceptable quality copies of the content data it receives (assuming a 50Kbit/sec content data stream). Such capability fits well into a binary tree topology.

    - A full rate ISDN modem connection nominally has an upstream bandwidth of 56 Kbits/sec and a downstream bandwidth of 56 Kbits/sec, which would potentially support acceptable quality re-transmission of a single copy of the content data stream (assuming a 50Kbit/sec content data stream). For this reason, a user node with a full rate ISDN modem connection to the Internet may be given, in this example, a bandwidth rating of one-half (0.5), or half the rating of a user node connected to the Internet by a DSL (e.g., ADSL) or cable modem connection.

    - User nodes with T1 or greater links to the Internet should be able to support more (e.g. at least twice as many) streams as DSL (e.g., ADSL) or cable modems, and therefore may be given, in this example, a bandwidth rating of two (2.0). In the event that in a distribution network parent nodes

may be assigned more than two child nodes directly connected thereto, bandwidth ratings greater than 2.0 may be assigned under this example to Internet connections having greater bandwidth than T1 connections.

- In another example (which example is intended to be illustrative and not restrictive), other bandwidth networks (e.g., 50 Kbit/sec and 100 Kbit/sec) may be supported.
- In another example, the bandwidth needed for a node to be "repeat capable" may depend on the particular network to which it is attached (upon initial connection the connecting node may first test its upstream bandwidth and then report that bandwidth to the network server; the server may determine whether the node is "repeat capable" or not.
- In another example (which example is intended to be illustrative and not restrictive), a node may be deemed "repeat capable" if the node is not firewalled (e.g., port 1138 is open) and the node has sufficient bandwidth to retransmit two copies of the incoming stream.
- A third factor is performance. A user node's performance rating may, in one example, be zero (0.0) if it is disconnected as a result of a Grandparent's Complaint Response Routine (discussed below in connection with FIG. 31). Otherwise, the user node's performance rating, in this example, may one (1.0).

Further, in this example (which example is intended to be illustrative and not restrictive), a user node's utility rating may determined by multiplying connection time by performance rating by bandwidth rating. That is, in this example:

- Utility Rating = Time x Performance Rating x Bandwidth Rating.

Referring once again to FIG. 13, it is seen that in this example information regarding a user node's time connected to the network, bandwidth rating, performance rating, utility rating and potential re-transmission rating (discussed below) may be stored in the user node's elapsed time buffer 137, bandwidth rating buffer 138, performance rating buffer 139, utility rating buffer 121 and potential re-transmission rating buffer 122, respectively.

In another example (which example is intended to be illustrative and not restrictive), nodes entering the network may be judged to be either "repeat capable" or "non-repeat capable" (wherein non-repeat capable nodes may be called "bumpable" nodes). Repeat

Ex. 1018-1713

capability may be based on: (1) Upstream bandwidth (e.g., tested at initial connection); and (2) the firewall status, opened or closed, of a node. In one specific example, if a node is either firewalled or has upstream bandwidth less than (approx.) 2.5 times the video streaming rate, that node will be deemed bumpable. All nodes joining the network, whether bumpable or non-bumpable, may be placed as close to the server as possible. However, for the purpose of placement (as described for example in the Universal Connection Address Routine of Figure 20), a repeat capable node can bump a bumpable node when being placed in the network.

Referring now to FIG. 17 a flow diagram associated with this example illustrating the Server's Connection Instruction Routine is shown. This routine need not necessarily rely on the new user node's utility rating. This routine may rely on the potential re-transmission rating of a user node, which is arrived at by multiplying the user node's performance rating by its bandwidth rating. That is, in this example:

- Potential Re-transmission Rating = Performance Rating x Bandwidth Rating.

Thus, in this example, the server would want to put those user nodes with the highest potential re-transmission rating as close to the server in a distribution chain as possible (i.e., as high uptree as possible) because they have the greatest likelihood of being able to re-transmit one or more copies of content data. On the other hand, a potential re-transmission rating of zero in this example indicates that a user node has no ability to (or little expected reliability in) re-transmitting even one copy of content data (the server in this example would want to put a user node with a zero rating as far as reasonably possible from the server in a distribution chain (i.e., as low downtree as possible)). For the purpose of ease of discussion, in the flow diagram example of FIG. 17 illustrating the Server's Connection Instruction Routine, the server is concerned about whether the potential re-transmission rating is zero (i.e., either one or both of the performance rating and bandwidth rating is zero) or greater than zero (i.e., both the performance rating and the bandwidth rating are greater than zero).

In step 171 the server node interrogates the new user node (or connection requesting node) for its bandwidth rating and performance rating. If the new user node is really new to the distribution network, or if it has returned to the server because all of the user node's ancestor nodes have disappeared from the network, the new user node's performance memory will contain, in this example, a performance rating of 1.0 (e.g., the default rating). However, if the new user node has been dispatched to the server for a new connection to the network

Ex. 1018-1714

because the new user node had failed to provide content data to one of its child nodes, then, in one example, its performance memory will contain a performance rating of zero.

In step 172, the server determines, in this example, whether the potential re-transmission rate of the connection requesting node is greater than zero (i.e., whether both the bandwidth rating and the performance rating are greater than zero, or, if only the bandwidth rating is considered, whether the bandwidth rating is greater than zero (i.e., the connection requesting node is a high-bandwidth node)). If the answer is "yes," then the server goes to step 173 in which the server determines whether it has room for the new user node. If the answer to the query in step 173 is "yes," then the server goes to step 174 in which it instructs the new user node to connect directly to the server. Then the server goes to step 102 in the Server's Connection Routine (see FIG. 10).

If the answer to the query in step 173 is "no" (i.e., the server does not have the capacity to serve the new user node directly), then the server goes to step 175. In step 175, the server, acting as an instructing node, consults its network topology database and devises a connection address list having the fewest number of connection links to a prospective parent node. That is, the server checks its information regarding the nodes in the level closest to the server (i.e., the highest uptree) to determine whether there are any potential parent nodes with space available for the new user node. If there are no potential parent nodes with space available, then the database is checked regarding nodes in the level one link further from the server, and so on until a level is found having at least one potential parent node with space available for the new user node. That is, the server determines which parent node with unused capacity for a child node is closest to the server (i.e., in the highest level, with the first level being the highest), and devises the connection address list from such prospective parent node to the server. The server then goes to step 102 (shown in FIG. 10).

In another embodiment of the invention, the server could skip step 172 and go directly to step 173 (or it could go to step 172 and, if the answer to the query in step 172 is "no" (i.e., either one or both of the bandwidth rating and the performance rating are zero, or, if only the bandwidth rating is considered, whether the bandwidth rating is zero (i.e., the connection requesting node is a low-bandwidth node)), the server could go to step 175). However, this could prevent the server from loading up the highest levels of the distribution chains with nodes capable of re-transmitting at least one acceptable copy of the content data (in contrast to if the server does perform step 172 and does goes to step 176 if the answer to the query in step 172 is "no.").

In step 176 the server consults its network topology database and devises a connection address list having the greatest number of connection links to a prospective parent node. That is, the server checks its information regarding the nodes in the level furthest from the server (i.e., the lowest downtree) to determine whether there are any potential parent nodes with space available for the new user node. If there are no potential parent nodes with space available, then the database is checked regarding nodes in the level one link closer to the server, and so on until a level is found having at least one potential parent node with space available for the new user node. In this manner, user nodes having limited or no reliable re-transmission capability may be started off as far from the server as possible and will have a reduced effect on the overall capacity of the distribution network.

As indicated above, and as discussed more fully below, one or more reconfiguration events may have transpired since the server's topology database was last updated. As a result, the first prospective parent node which is actually present on the distribution network for the new user node to contact may not have room for the new user node. By way of example (which example is intended to be illustrative and not restrictive), when node X tries to join the distribution network having the topology shown in FIG. 12A, the server provides node X with the following connection address list: C-B-A-S. If node C had disappeared from the network between the last update of the server's topology database and node X's attempting to contact node C, then node E, by virtue of a reconfiguration event, would be connected, in this example, to node B as shown in FIG. 12D. Then node X, in performing the Prospective Child Node's Connection Routine discussed in connection with FIGS. 14 and 15, would contact node B. Node B, in the Prospective Parent Node's Connection Routine, discussed in connection with FIG. 16, would have to answer the query of step 162 in the negative and go to step 164, in which it performs the Fully Occupied Parent's Connection Instruction Routine.

Referring now to FIG. 18, a flow diagram associated with this example illustrating that routine is shown. It is similar to the Server's Connection Instruction Routine. Since the fully occupied parent node has already determined that it has no room for the new user node (or connection requesting user node), the Fully Occupied Parent's Connection Instruction Routine does not need to include a step in which a determination is made regarding whether there is room for the new user node (or connection requesting node). In the Fully Occupied Parent's Connection Instruction Routine the fully occupied parent node, acting as an instructing node, first performs step 181 in which it interrogates the new user node for its bandwidth rating and performance rating. In step 182, the fully occupied parent node determines whether the potential re-transmission rate of the new user node is greater than

zero. (If only the bandwidth rating is considered, then it determines whether the bandwidth rating is greater than zero (i.e., the connection requesting node is a high-bandwidth node).) If the answer is "yes," then the fully occupied parent node goes to step 183 in which the fully occupied parent node consults its topology database, which contains the latest information available to that node regarding the links from the fully occupied parent node back to the server and regarding the fully occupied parent node's own progeny (i.e., its children, grandchildren etc.,) and devises a new connection address list having the fewest number of connection links to a new prospective parent node. That is, the fully occupied parent node checks its information regarding the nodes in the level closest to the fully occupied parent node (but not closer to the server) to determine whether there are any potential parent nodes with space available for the new user node. If there are no potential parent nodes with space available, then the database is checked regarding nodes in the level one link further from the fully occupied parent node (and further from the server), and so on until a level is found having at least one potential parent node with space available for the new user node. That is, the fully occupied parent node determines which new prospective parent node with unused capacity for a child node is closest to the fully occupied parent node, and devises the connection address list from such new prospective parent node through the fully occupied parent node and on to the server. The fully occupied parent node then goes to step 184 in which it provides the new user node with the new connection address list and disconnects the new user node from the fully occupied parent.

If the answer to the query in step 182 is no (i.e., either (i) one of the bandwidth rating and performance rating is zero or (ii) if only the bandwidth rating is considered, the bandwidth rating is zero (i.e., the connection requesting node is a low-bandwidth node)), the fully occupied parent node goes to step 185. In step 185 the fully occupied parent node consults its network topology database and devises a connection address list having the greatest number of connection links to a new prospective parent node. That is, the fully occupied parent node checks its information regarding the nodes in the level furthest from it (and farther from the server than it is) to determine whether there are any potential parent nodes with space available for the new user node. If there are no potential parent nodes with space available, then the database is checked regarding nodes in the level one link closer to the fully occupied parent node, and so on until a level is found having at least one potential parent node with space available for the new user node. As discussed above, this helps assure that new user nodes having limited or no reliable re-transmission capability are started off as far from the server as possible. After the fully occupied parent node devises the connection

address list from such new prospective parent node through the fully occupied parent node and on to the server, the fully occupied parent node then goes to step 184 where it performs as discussed above.

In another embodiment of the invention, the distribution software could be designed such that a fully occupied parent performs an abbreviated Fully Occupied Parent's Connection Instruction Routine, in which steps 181, 182 and 185 are not performed. That is, it could be presumed that the server has done the major portion of the work needed to determine where the new user node should be placed and that the fully occupied parent user node need only redirect the new user node to the closest available new prospective parent. In such event, only steps 183 and 184 would be performed.

In the example discussed above in which node C had disappeared from the network when new user node X had been given, by the server, connection address C-B-A-S, and in which node B is a fully occupied parent node as shown in FIG. 12D, node B would appear to have the choice of devising either connection address list D-B-A-S or E-B-A-S regardless of whether the full or abbreviated Fully Occupied Parent's Connection Instruction Routine were performed.

In this regard, to determine which connection address list to choose when there are two or more prospective parent nodes in a particular level having space for a new user node, the distribution software could have an additional subroutine as part of steps 175, 176, 183 and 185. An example of this subroutine, called the Multi-Node Selection Subroutine is illustrated in FIG. 19.

In step 191 the server or fully occupied parent node deciding where to send a new user node determines whether any of the potential new parent nodes is partially occupied. As discussed earlier, in one example a partially occupied potential parent node may be preferred over an unoccupied potential parent node. In this case, if any of the potential parent nodes is partially occupied, then the server or fully occupied parent node goes to step 192. In step 192 the partially occupied prospective parent node with the highest utility rating is selected as the new prospective parent node. If there were only a single partially occupied potential parent node, then that node is selected.

If in step 191 it is determined that there are no partially occupied potential parent nodes, then the server or fully occupied parent node goes to step 193. In step 193 the unoccupied prospective parent node with the highest utility rating is selected as the new prospective parent node.

In another example (which example is intended to be illustrative and not restrictive),the software engineer could have step 193 follow an affirmative response to the query in step 191 and step 192 follow a negative response; in such event, unoccupied prospective parent nodes would be selected ahead of partially occupied prospective parent nodes (all other things being equal, it may be advantageous to select nodes that have zero children over nodes that already have one child; while it is true that this may increase the number of repeat nodes, what the inventors have found is that by filling the frontier in an "interlaced" fashion, connecting nodes build up their buffers more quickly (allowing a reduction in the incidence of stream interruptions).

After whichever of steps 192 and 193 is completed, the server or fully occupied parent node returns to the step from which it entered the Multi-Node Selection Subroutine (i.e., step 175, 176, 183 or 185), and completes that step.

So, in the example shown in FIG. 12D, where node C left the network thereby leaving node B with the chore of devising a new connection address list for node X, node B would perform the Fully Occupied Parent's Connection Instruction Routine. Regardless of the bandwidth and performance ratings of node X, node B would be choosing between nodes D and E in the third level. In step 191 node B would determine that neither D nor E is partially occupied, and therefore node B would go to step 193. Assuming that nodes E and D have equal bandwidth and performance ratings and that node D was connected to the network longer than node E, node D would be selected because it would have the higher utility rating since it was connected to the network longer than node E. Node B would then go to step 194 and then return to the step from which it entered the Multi-Node Selection Subroutine. When node B returns to step 183 or 185, it completes that step and moves on to step 184. In that step, node B provides new user node X with new connection address list D-B-A-S and node X connects to the distribution network as shown in FIG. 12E.

Referring now to an "Alternative Distribution Network Construction" example (which example is intended to be illustrative and not restrictive), it is noted that under this example the length of the distribution chains (i.e., the number of levels of user nodes linked through each other to the server) is aimed to be kept as small as possible.

Also under this example, user nodes having the highest bandwidth capabilities are aimed to be kept closer to the server (e.g., in order to allow the greatest expansion of the distribution system). However, it is possible that zero bandwidth rated nodes may nevertheless appear relatively far uptree (thereby stunting the growth of that chain). The following method may be used in constructing the distribution network both by servers and

by prospective parents which are actually completely occupied, either of which may be thought of as an instructing node (that is, software enabling the routines discussed below could be installed on servers and user nodes alike).

More particularly, in the distribution system of this example, each child node reports directly to (or is tested by) its parent node with respect to information relating to the child node's bandwidth rating, performance rating, potential re-transmission rating and utility rating. In turn, each parent node reports all the information it has obtained regarding its child nodes on to its own parent node (a parent node also reports to each of its child nodes the address list from that parent back to the server, which list forms what may be referred to as the ancestor portion of the topology database – in addition, a parent node reports to each of its child nodes the addresses of their siblings). The reports occur during what may be referred to as a utility rating event. Utility rating events may occur, for example, on a predetermined schedule (e.g., utility rating events may occur as frequently as every few seconds). As a result of all the reporting, each node stores in its topology database the topology of the tree (including all its branches) extending below that node, and the server stores in its topology database the topology of all the trees extending below it. This may be referred to as the descendant portion (or descendant database) 133 of the topology database (see FIG. 13). The descendant database of a particular node may include a descendant list, a list of the addresses of all the nodes cascadingly connected below that particular node. Included in the topology database information may be the utility ratings of the nodes below the node in which that particular topology database resides.

After each reporting event in this example, each parent node (including the server), acting as an instructing node, devises two lists of prospective (or recommended) parent nodes. In this example, the first list, or Primary Recommended Parent List ("PRPL"), stored in the Primary Recommended Parent List buffer 123 (see FIG. 13), lists all the nodes in the descendant portion of that node's topology database which have bandwidth available to support another child node (in one example (which example is intended to be illustrative and not restrictive), in a binary tree system, all nodes in the descendant portion of the topology database having (i) a bandwidth rating of at least one and (ii) less than two child nodes would be listed). They would be listed with those node's which are closest to the node in which that particular topology database resides at the top of the list, and those nodes which are in the same level would be ranked such that the node with the highest utility rating would be listed first, the node with the second highest utility rating would be listed second and so on. By way of example (which example is intended to be illustrative and not restrictive), the PRPL of a

second level node would list a third level node with available bandwidth ahead of a fourth level node with available bandwidth even if the fourth level node's utility rating were higher than that of the third level node.

The second list in this example, or Secondary Recommended Parent List ("SRPL"), stored in the Secondary Recommended Parent List buffer 124 (see FIG. 13), lists all the nodes in the descendant portion of that node's topology database which have the ability to re-transmit content data to child nodes but are fully occupied, and at least one of its child nodes is incapable of re-transmitting content data to another child node (in one example (which example is intended to be illustrative and not restrictive), in a binary tree system, all nodes in the descendant portion of the topology database having (i) a bandwidth rating of at least one and (ii) at least one child node having a bandwidth rating less than one (i.e., being incapable of re-transmitting content data to two child nodes) would be listed). Like the nodes in the PRPL, the nodes in the SRPL would be listed with those node's which are closest to the node in which that particular topology database resides at the top of the list, and those nodes which are in the same level would be ranked such that the node with the highest utility rating would be listed first, the node with the second highest utility rating would be listed second and so on.

Of note, in this example the SRPL lists those parent nodes having the growth of their branches (i.e., their further progeny) blocked or partially blocked by a low-bandwidth child node. This may lead to an unbalanced growth of the distribution system, and a limitation on the total capacity of the system.

Of further note, to the extent that a node (including a server) has room for another child node or is the parent of a low bandwidth node, it may be listed on its own PRPL or SRPL.

Referring now to FIGS. 20A and 20B, an example flow diagram illustrating what may be referred to as the Universal Connection Address Routine is shown. As indicated above, a server or a fully occupied prospective parent node receiving a connection request may be referred to as an "instructing node." When a new node (here node N) approaches an instructing node, the instructing node performs step 201, it receives a connection request. It then goes to step 202 in which it interrogates or tests node N to determine whether it is a low-bandwidth node ("low-bandwidth" may mean, for example, either or both: (1) a node that is incapable of broadcasting two copies of its incoming stream on to its children; and (2) a node that is incapable of rebroadcasting its stream because it is firewalled). In the binary tree system of one example, a low-bandwidth node is a node with a bandwidth rating less than

one. If node N is a low bandwidth node, the instructing node proceeds to step 203 in which the instructing node determines whether there are any available prospective parent nodes which are not fully occupied. Sometimes the distribution network may be fully occupied. If it is, the instructing node's PRPL will be empty. If it is empty, the response to the query in step 203 would be yes. Then the instructing node goes to step 209 in which the new node is sent back to the server to start the connection process from the beginning. If the response to the query in step 203 is no, then the instructing node goes to step 204 in which it selects a prospective (or recommended) parent node for node N. The instructing node then moves on to step 205 in which it consults its topology database and devises a connection address list from the recommended parent node back to the server, and provides such connection address list to node N. (If the instructing node is a user node, then the connection address list leads back to the server through the instructing node.) At this point, node N performs the Prospective Child Node's Connection Routine discussed above in connection with FIG. 14.

As part of step 205, the instructing node inserts node N into its topology database as a child of the recommended parent node. It does this because other new nodes may seek connection instructions prior to the next utility rating event (i.e., before reports providing updated information regarding the topology of the distribution network are received by the instructing node), and such new nodes should not be sent to a prospective parent which the instructing node could know is fully occupied. In this regard, the instructing node then goes to step 206 in which it checks its topology database to determine whether the recommended parent, with node N presumptively docked to it as a child node, is fully occupied (in the example here of a binary tree network, whether it has two child nodes). If the answer is no, then the instructing node has finished this routine. If the answer is yes, then the instructing node goes to step 207 in which it removes the recommended parent from the PRPL and inserts it into the SRPL, and the instructing node has finished the routine.

If the answer to the query in step 202 is no (e.g., node N is not a low-bandwidth node; in the binary tree network example it is a high-bandwidth node capable of re-transmitting content data to two child nodes), the instructing node moves on to step 208. There it determines whether both the PRPL and SRPL are empty (which may occur under certain circumstances, such as, for example, when the number of levels in a distribution system is capped and at least all the nodes on all but the last level are fully occupied with high-bandwidth nodes). If so, the instructing node goes to step 209 in which the new node is sent back to the server to start the connection process from the beginning. If the response to the query in step 208 is no, then the instructing node goes to step 210 in which it selects a

prospective (or recommended) parent node for node N from the PRPL and an alternate recommended parent node from the SRPL (if either the PRPL or SRPL is empty, the instructing node will make a "nil" selection from that list – the instructing node knows from step 208 that at least one of the lists will not be empty). The instructing node then goes to step 211 in which it determines whether the alternate recommended parent node is closer to the server (i.e., higher uptree) than the recommended parent node derived from the PRPL. If the alternate recommended parent node is on the same level as, or on a lower level than the recommended parent node derived from the PRPL (or if the selection from the SRPL is nil), then the answer to the query in step 211 is no.

In such event, the instructing node goes to step 212 in which it consults its topology database and devises a connection address list from the recommended parent node back to the server, and provides such connection address list to node N (if the instructing node is a user node, then the connection address list leads back to the server through the instructing node).

As in step 205 discussed above, at this point, node N performs the Prospective Child Node's Connection Routine discussed above in connection with FIG. 14, and in step 212, the instructing node inserts node N into its topology database as a child of the recommended parent node.

The instructing node moves to step 213 in which it adds node N to an appropriate position on the PRPL. It does this because, as a result of step 202, it knows that the node N is capable of re-transmitting content data to its own child nodes.

The instructing node then goes to step 214 in which it checks its topology database to determine whether the recommended parent, with node N presumptively docked to it as a child node, is fully occupied (in the example here of a binary tree network, whether it has two child nodes). If the answer is no, then the instructing node has finished this routine. If the answer is yes, then the instructing node goes to step 215 in which it removes the recommended parent from the PRPL because it is now deemed to not be an available prospective parent node.

Then the instructing node goes to step 216 in which it consults its topology database to determine whether any of the recommended parent node's children is a low-bandwidth node (in this example, knowing that node N is not a low-bandwidth node and knowing that the recommended parent node has two child nodes, the question is whether the child node other than node N is a low-bandwidth node.) If the answer is no (i.e., all the recommended parent node's children are high-bandwidth nodes), then the instructing node has finished the routine.

Ex. 1018-1723

If the answer is yes, then the growth of the recommended parent node's line of progeny is partially blocked by a low-bandwidth child node. The instructing node moves on to step 217 in which it adds the recommended parent to the SRPL.

If the answer to the query in step 211 is yes (i.e., the alternate recommended parent (selected from the SRPL) is: (i) on a higher level than the recommended parent (selected from the PRPL) or (ii) the selection from the PRPL is nil), then the instructing node moves to step 218. In that step the instructing node consults its topology database to determine: (i) which of the alternate recommended parent node's child nodes is a low-bandwidth node or (ii) if they both are low-bandwidth nodes, which of the child nodes has been connected to the system a shorter period of time. In one example (which example is intended to be illustrative and not restrictive), the instructing node may send a disconnect signal to that child node with instructions to return to the server to start the connection process from the beginning (as a new user node (or connection requesting user node)). In another example (which example is intended to be illustrative and not restrictive), the bumped node may climb its path to its grandparent which gives it a new connection path – rather than returning to the root server (doing so typically means that the incoming node that bumped the non-repeat-capable node ends up being that node's new parent).

The instructing node moves on to step 219 in which it consults its topology database and devises a connection address list from the alternate recommended parent node back to the server, and provides such connection address list to node N (if the instructing node is a user node, then the connection address list leads back to the server through the instructing node).

As in steps 205 and 212 discussed above, at this point, node N performs the Prospective Child Node's Connection Routine discussed above in connection with FIG. 14, and in step 219, the instructing node inserts node N into its topology database as a child of the alternate recommended parent node.

The instructing node moves to step 220 in which it adds node N to an appropriate position on the PRPL. It does this because, as a result of step 202, it knows that the node N is capable of re-transmitting content data to its own child nodes.

The instructing node then goes to step 221 in which it checks its topology database to determine whether all the child nodes of the alternate recommended parent are high-bandwidth nodes. If the answer is no, then the instructing node has finished this routine. If the answer is yes, then the instructing node goes to step 222 in which it removes the alternate recommended parent from the SRPL because it is now deemed to not be an available

alternative prospective parent node since the growth of its progeny line is not even partially blocked by one of its own children. At this point the instructing node has finished the routine.

With the routines discussed in connection with the above examples, the distribution network will be built with each new node assigned to the shortest tree (or chain), and those with the fewest number of links between it and the server. However, low-bandwidth nodes, which would tend to block the balanced growth of the distribution network, would be displaced by high-bandwidth nodes and moved to the edges of the network where they would have reduced effect on the growth of the network.

Referring now to FIG. 21, this FIG. may be used as an example to illustrate how the Universal Connection Address Routine works. It shows primary server 11 with node A as first level node 12, its child nodes B and C as second level nodes 13, their child nodes D, E, F and G as third level nodes 14, node D and G's child nodes H, I, J and K as fourth level nodes 15, and node I's child nodes L and M as fifth level nodes 17. Assume for this example that all available docking links directly to the server 11 are occupied by high-bandwidth nodes, and that all trees are at full capacity other than that rooted to the server 11 through first level node A. The utility rating for each node is set forth in FIG. 21 under its letter designation. Low-bandwidth nodes (in this example, those nodes having a bandwidth rating less than one) are shown with a bandwidth rating of "Ω," which indicates in this example that such nodes will not have any child nodes assigned to them. Nodes A, B, C, D, E, G, I, J and K are high-bandwidth nodes (or repeater nodes (i.e., they are capable of re-transmitting the content data they receive from their respective parent nodes to their respective child nodes, if any)).

A new user node (or other connection requesting user node), here referred to as "node N", may be directed to node A as a result of various reconfiguration events. When node A receives a connection request from node N, node A will, in this example, either consult both of its PRPL and SRPL buffers, if node N is a high-bandwidth node, or consult just its PRPL buffer (if node N is a low-bandwidth node).

Nodes F, H, L, and M will not appear on any list in this example since they are low-bandwidth nodes. Nodes B, C, D, G, I and A itself will not appear on the PRPL in this example since these nodes are fully occupied. However, nodes C, D, and I will appear on the SRPL in this example because they each have at least one low bandwidth child node.

Using the rules discussed above (i.e., ranking prospective parent nodes by level and within levels by utility rating) the PRPL would be as follows in this example:

- PRPL

- E (a third level node with available capacity)
  - J (the fourth level node with available capacity having the highest utility rating)
- K (a fourth level node with available capacity)

The SRPL in this example would be as follows:

- SRPL
  - C (the fully occupied second level node with a low-bandwidth child node)
  - D (the fully occupied third level node with a low-bandwidth child node)
  - I (the fully occupied fourth level node with a low-bandwidth child node)

If node N is a low-bandwidth node, node A will give, in this example, node N the following as its new connection address list:

- E-B-A-S.

Since node E would not have two child nodes, it would remain on the PRPL.

If node N is a high-bandwidth node, node A will compare (step 211 of FIG. 20B) the first node on the PRPL (the recommended parent node) with the first node on the SRPL (the alternate recommended parent node). Here, node C, the first node on the SRPL is a higher level node than node E, the first node on the PRPL. So, node A will send a disconnect signal (step 218 of FIG. 20B) to node F, node C's low-bandwidth child node. Then it will provide node N with the following new connection address list and add node N to the PRPL (step 219 of FIG. 20B):

- C-A-S.

Since node C would now have two high-bandwidth child nodes (nodes N and G), node C would be removed from the SRPL (step 222 of FIG. 20B).

Further, in this example wherein node N is a high-bandwidth node, when node F returns to the server 11 (or, in another example, goes to Node A) for a new connection, the

Ex. 1018-1726

server (or Node A) will also use the Universal Connection Address Routine (Node F may go to Node A rather than the server because (in one example) when a node is "kicked" it may climb its internal connection path rather than jumping directly to the root server). Since node F is a low-bandwidth node, the server (or Node A) will give node F the following as its new connection address list:

- E-B-A-S.

FIG. 22 illustrates the new topology. As can be seen, absent intervening events, low-bandwidth node F will end up moving down from the third level to the fourth level and the bandwidth capacity of the third level will increase from six child nodes to eight (its maximum in this binary tree example). The potential bandwidth capacity of the fourth level would also increase, from ten child nodes to twelve.

Referring now to a "Distribution Network Reconfiguration" example (which example is intended to be illustrative and not restrictive), it is noted that under this example the user nodes are free to exit the distribution network at will. And, of course, user nodes may exit the distribution network due to equipment failures between the user node and the communication network (the Internet in the examples discussed herein). This was first discussed in connection with FIGS. 10A-E.

Under this example the present invention may handle the reconfiguration required by a user node departure by having certain user nodes in the same tree as the departing node perform a Propagation Routine. The results of the propagation routine of this example are illustrated in FIGS. 23 and 24. There a tree rooted to the server through first level user node P is illustrated. Node P has two child nodes, second level nodes Q and X. Through node X, P has two grandchild nodes, third level nodes A and B. Based on the relative utility ratings of nodes Q and X, P has sent signals to its children instructing them to set the color of the flag represented by the information in their respective reconfiguration flag buffers to "green" and "red," respectively. The use of colors as designators is merely discretionary. In reality they represent relative ratings of siblings for purposes of determining their roles in a reconfiguration event, and may be referred to as "propagation ratings." In addition, instead of colors, numbers (e.g., 1 and 2) could be used, as well as any of many other nomenclature schemes to describe the grades of propagation ratings. The "red" "green" ratings discussed herein shall be deemed representative of all such schemes. The number of grades of propagation ratings assigned by a parent node may be equal to the number of children each

Ex. 1018-1727

parent node has. For example, for a branch factor "n" system, the maximum number of grades of propagation ratings may be "n." Since the distribution network in the examples discussed herein is a binary tree distribution network, a parent node will be required to assign at most up to two grades of propagation ratings.

In the example shown in FIG. 23, node P, during the most recent utility rating measurement event, discerned that node Q has a higher utility rating than node X, and therefore P has assigned node Q a green rating, represented by the solid-lined circle surrounding the letter Q in FIG. 23. P has assigned node X a red rating, represented by the dashed-lined circle surrounding the letter X in FIG. 23. In a similar manner, node X has assigned green and red ratings to third level nodes A and B, respectively. In this example (which example is intended to be illustrative and not restrictive), at the same time that a parent node assigns a propagation rating to a child node, it may also provide to such child node the address of the child node's sibling, if there is any. A child node may store information about its sibling (or siblings) in the sibling portion (or sibling database) 134 of its topology database (see FIG. 13). The sibling database includes a sibling list, a list of the addresses of a node's siblings (the data relating to the siblings' addresses may also contain information regarding the propagation ratings of the siblings (e.g., in the event that the distribution system has a branch factor greater than two)). In the example shown in FIG. 23, nodes Q and X know that they are each other's siblings and nodes A and B know that they are each other's siblings.

In another example (which example is intended to be illustrative and not restrictive), nodes do not store information about their siblings. In other words, in this example, a node generally does not know who its sibling is. Instead, cross connections between a red node and its green sibling are implemented via a "priority join" message sent to a red child that includes the IP address of the green sibling as the first entry in the connection path list.

In the event that node X were to leave the distribution network, nodes A and B would of course stop receiving content data from node X. Nodes A and B would consult their topology databases for the address of their grandparent, and each would interrogate node P for instructions. Since node X is out of the distribution network, node P would send out a reconfiguration event signal (sometimes referred to as a "propagation signal") to the nodes which had been node X's children (i.e., the nodes which were node P's grandchildren through node X) and node P would send an upgrade propagation rating signal to its remaining child, here node Q.

Ex. 1018-1728

In response to the upgrade propagation rating signal, node P's remaining child, node Q, would set its reconfiguration flag buffer 136 (see FIG. 13) to the next highest available propagation rating grade. Since in the example illustrated in FIG. 23, node Q's reconfiguration flag buffer already is set for the highest propagation rating grade (here green), node Q could either do nothing or reset its propagation rating in response to the upgrade propagation rating signal. The result would be the same, node Q's propagation rating would remain green. If node Q's propagation rating were red, then it would set its propagation rating to green in response to the upgrade propagation rating signal. Node Q would do nothing else in response to the upgrade propagation rating signal (note that as a matter of design choice, the software engineer could have the parent of the departed node (here node P is the parent of departed node X) send no upgrade propagation rating signal to its remaining child node (here node Q) if its remaining child node already is at the highest propagation rating).

In this example, the recipients of the propagation signal (i.e., the children of the missing node) would respond thereto as follows. First they would check their own respective propagation ratings. If the propagation signal's recipient has the highest propagation rating grade (here green), it would reset its propagation rating to the lowest rating grade (here red); re-transmit the propagation signal to its own child nodes (if there are any); disconnect the child node which did not have the highest propagation rating prior to its receipt of the propagation signal if the propagation signal's recipient with the green rating has more than one child node; allow its sibling to dock with it as a child node for the purpose of transmitting content data to that child node; and dock (or remained docked), for purposes of receiving content data, with the node sending the propagation signal to it (in systems having a branch factor of more than two, the propagation signal recipient whose propagation rating had been the highest would disconnect its child nodes which did not have the highest propagation rating, prior to the receipt of the propagation signal just sent, and it would allow its (i.e., the formerly green node's) siblings to dock with it as child nodes).

If the propagation signal's recipient has other than the highest propagation rating (i.e., just prior to the receipt of the propagation signal), it would upgrade its propagation rating to the next highest rating grade; dock with its sibling which had the highest rating grade; and begin receiving content data from it. If the propagation signal's recipient has other than the highest propagation rating, it does not re-transmit the propagation signal to its own child nodes.

In the example illustrated in FIG. 23, nodes A and B receive the propagation signal from node P. Since node A has the highest propagation rating grade, here green, it: (i) sets its

Ex. 1018-1729

reconfiguration flag buffer 136 (see FIG. 13) so that it has the lowest propagation rating grade, here red; and (ii) docks with node P (becoming a second level node 13) to begin receiving content data from node P. Node B changes its propagation rating from red to the next higher propagation rating (and since this example is a binary tree (or branch factor two) system, the next higher rating is the highest, green) and docks with node A to receive content data. The resulting topology is shown in FIG. 24. Note that node B remains a third level node 14.

FIGS. 25 and 26 illustrate what happens in this example when the departing node is green. In FIG. 25, node X is the departing node and it is green. When node P sends the upgrade propagation rating signal to node Q, it changes its propagation rating from red to green. In other respects the reconfiguration event proceeds essentially as described in the paragraph immediately above, and results in the topology shown in FIG. 26 (which is the same as the topology in FIG. 24).

FIG. 27 is a flow diagram associated with an example showing the Child Node's Propagation Routine, the routine which is followed by a child node upon receiving a propagation signal. Of note, this example is provided for illustration only, and is not restrictive, and other examples are of course possible (e.g., the color of a node may assigned by its parent such that nodes assign and know the color of their children but not of themselves).

In any case, still referring to FIG. 27, the child node first performs step 271 wherein it determines whether it has received such a signal. If it has not, then it does nothing. If it has received a propagation signal, it proceeds to step 272 wherein it checks its propagation rating grade in reconfiguration flag buffer 136 (see FIG. 13) . If its propagation rating is at the highest grade (i.e., it is a "green" node in this example), then it proceeds to step 273 where it sets its reconfiguration flag buffer to the lowest propagation rating grade. It then proceeds to step 274 in which it re-transmits the propagation signal to its own child nodes, and which results in the undocking of all its child nodes except for the one with the highest propagation rating. The child node (i.e., the child node referred to in the second sentence of this paragraph) then performs step 275 in which it determines whether it is already docked with the node which sent the propagation signal. If it is not (i.e., it received the propagation signal from its grandparent), then it proceeds to step 276. In that step it: (i) devises a new connection address list, which is the ancestor portion of its topology database with the current nominal parent node removed, resulting in the grandparent node becoming the first node on the connection address list; and (ii) then performs the Prospective Child Node's Connection

Routine (i.e., it goes to step 141 discussed above in connection with FIGS. 14 and 15). The Prospective Child Node's Connection Routine is performed because some likelihood exists that even the grandparent node may have departed from the distribution system between the moment it sent out the propagation signal and the moment that its highest rated grandchild from its missing child attempted to dock with it.

If the answer to the query in step 275 is in the affirmative (i.e., the child node is receiving a propagation signal which has been re-transmitted by its parent), then the child node does nothing more (or as illustrated in FIG. 27, it performs step 277 which is to remain docked to its parent node).

If the answer to the query in step 272 is in the negative (i.e., it does not have the highest propagation rating (it is a "red" node in this example)), then it proceeds to step 278 in which it: (i) sets its reconfiguration flag buffer so that its propagation rating is the next higher grade (in this binary tree system the rating is upgraded from red to green); (ii) undocks from the parent with which it was docked before receiving the propagation signal (if it was docked with a parent before receiving such signal); and (iii) either (a) waits a predetermined period of time during which the node's sibling which was green prior to their receipt of the propagation signal should have re-transmitted the propagation signal to its own child nodes (thereby causing any red child nodes to undock from it) or (b) confirms that the node's sibling which was green prior to their receipt of the propagation signal actual did re-transmit the propagation signal to its child nodes (if it has any). Then the child node performs step 279 in which it: (i) devises a new connection address list, which is the ancestor portion of its topology database with its sibling node having the highest propagation rating placed in front of the previous parent as the first node on the connection address list; and (ii) then performs the Prospective Child Node's Connection Routine (i.e., it goes to step 141 discussed above in connection with FIGS. 14 and 15). The Prospective Child Node's Connection Routine is performed because some likelihood exists that the sibling may have departed from the distribution system between the moment that the propagation signal had been sent to the child node and the moment that the child node attempted to dock with its sibling.

In another example (which example is intended to be illustrative and not restrictive), part (iii) of step 278 may be handled using the "priority join" mechanism. More particularly, in this example, a priority join from a node X cannot be refused by a target node T – it must accept (dock) the incoming node. If the target node T already had two children G & R, the target node T instructs its red child R to connect (via "priority join") to the green node G. The node T then disconnects from its red child R immediately prior to accepting the incoming

node X. The bumped node R then sends a priority join to its own green sibling A as the target. These actions cause the reconfiguration events to cascade until the edge of the tree is reached (of note, the algorithm described in this paragraph may have essentially the same effect as illustrated in Figures 28 and 29 – under certain circumstances the algorithm described in this paragraph may be somewhat more reliable).

FIG. 28 illustrates an example distribution system comprising node P as a first level node 12; nodes Q and X as second level nodes 13; nodes A and B as third level nodes 14; nodes C, D, E an F as fourth level nodes 15; nodes G and H as fifth level nodes 17; nodes I and J as sixth level nodes 20; nodes K and L as seventh level nodes 21 and nodes M and N as eighth level nodes 22. Nodes Q, B, C E, H, I, L and M have red propagation ratings, as symbolized by the dashed circles, and the remaining nodes have green propagation ratings as symbolized by the solid circles. As a result of node X's departing from the system, node P sends out an upgrade propagation rating signal to node Q and a propagation signal to the children of node P's departed child node X (i.e., nodes A and B).

In another example (which example is intended to be illustrative and not restrictive), during a graceful depart of node x, node x may send out messages to its parent and children immediately prior to its departure – rather than the parent of node x initiating these messages as described above.

In any case, continuing the example of FIG. 28, node Q changes its propagation rating to green. Further, in this example nodes A and B begin the Child Node's Propagation Routine. With respect to node A, it answers the query of step 271 in the affirmative. Since its rating is green, it also answers the query of step 272 in the affirmative. In step 273 node A changes the setting of its reconfiguration flag buffer to show a red propagation rating. In step 274 it re-transmits the propagation signal to its child nodes C and D. It answers the query of step 275 in the negative because it is not docked for purposes of receiving content data with node P. Node A then goes to step 276 wherein it consults the ancestor portion of its topology database and creates a new connection address list starting with its grandparent. The new connection address list is P-S. Then node A performs the Prospective Child Node's Connection Routine, starting with step 141. (See FIG. 14.) Assuming that no other intervening events have occurred, node A will successfully dock with node P.

Node B answers the query of step 271 in the affirmative and, because its propagation rating was red when it received the propagation signal, it answers the query of step 272 in the negative. From there it goes to step 278 in which it changes its propagation rating to green. Then in step 279 node B consults the ancestor portion of its topology database and creates a

new connection address list with its sibling, node A, placed in front of node X as the first address on the new connection address list. The new connection address list is A-X-P-S. Then node B performs the Prospective Child Node's Connection Routine, starting with step 141. Assuming that no other intervening events have occurred, node B will successfully dock with node A (if node A has disappeared, node B would attempt to dock with node X, but since it is not there, it would move on to node P.)

Node A, which had the higher propagation rating of the two child nodes of departed node X, moves up one level to become a second level node 13. Since it is a "new" entry to the second level, its initial utility rating and its propagation rating will be lower than that of node Q. As time goes by, at subsequent utility rating measurement events node A's utility rating (and hence its propagation rating) may become higher than that of node Q.

Node B, which had the lower propagation rating of the two child nodes of departed node X, does not re-transmit the propagation signal to its child nodes (nodes E and F), and they will follow B wherever it goes in the distribution system. In the example discussed here, node B becomes the child of node A while remaining third level node 14. At least initially, it will be node A's child with the higher utility rating and propagation rating.

As a result of node A's performing step 274, its child nodes C and D (which were fourth level nodes 15 in FIG. 28 when node X was still in the system) receive a propagation signal. Since node C has a red propagation rating, it, like node B, will remain in its level, change its propagation rating to green and dock with its sibling, with the result being that node D becomes its parent.

Since node D had a green propagation rating when it received the propagation signal, it answers the queries of steps 271 and 272 in the affirmative and changes its propagation rating to red in step 274. It answers the query of step 275 in the affirmative, and remains docked with node A. As a result, node D moves up a level and becomes an third level node 14 (with, at least until the next utility rating event, a lower utility rating and propagation rating than its new sibling, node B).

This process proceeds down the tree, with each child of a node which moves up a level doing one of the following under this example: (i) if it is a node having a green propagation rating, it remains docked with its parent, thereby itself moving up a level, changes its propagation rating to red and re-transmits the propagation signal to its child nodes; or (ii) if it is a node having a red propagation rating, it docks with the node which was its sibling, thereby staying in the same level, and changes its propagation rating to green.

Ex. 1018-1733

The resulting topology after the reconfiguration event caused by node X's departure is illustrated in FIG. 29.

Of note, it is believed that reconfiguration using the Child Node's Propagation Routine of this example when a node departs the system does not necessarily result, in the long run, in significantly more inter-node reconnection events than any other reconfiguration method (and may result in fewer reconnection events). Moreover, it is believed that reconfiguration using the Child Node's Propagation Routine of this example helps assure that the more reliable nodes are promoted up the distribution network even if many reconfiguration events occur close in time to each other in a particular tree.

Referring now to a specific example of pseudo-code (which example is intended to be illustrative and not restrictive), it is noted that a propagation routine may be set forth as follows (wherein node X is the node departing from the distribution system):

```
Departs(X)
Begin
 If (X.Parent.GreenChild == X)  /* the departing node is green*/
 {
  X.Parent.GreenChild := X.Parent.RedChild
  X.Parent.RedChild := X.GreenChild
 }
 Else /* the departing node is already red */
 {
  X.Parent.RedChild :=X.GreenChild
 }
 Propagate(X.GreenChild, X.RedChild)
End
```

Referring now to another specific example of pseudo-code (which example is intended to be illustrative and not restrictive), it is noted that a propagation routine may be set forth as follows (wherein node promotion during the network reconfiguration event is precipitated by node X's departure):

```
Propagate (X, RedSibling)
Begin
 OriginalGreenChild := X.GreenChild
```

```
OriginalRedChild := X.RedChild
X.RedChild := X.GreenChild
X.GreenChild := RedSibling
If (OriginalGreenChild <> null)
{
  Propagate(OriginalGreenChild, OriginalRedChild)
}
End
```

Referring now to a "Malfunctioning Nodes" example (which example describes a voting system between grandchildren and grandparents about a potentially malfunctioning parent and is intended to be illustrative and not restrictive), it is noted that what happens when a node leaves the distribution network was discussed in the above example (e.g., when a node leaves the network, its parent will, at the next utility rating event, report that it has room for a new node, except when the parent has been contacted about the missing node by its grandchild node or nodes, in which event, a reconfiguration event will proceed as described above). Sometimes, however, a node intends to remain in the network, but there is a failure of communication.

In this regard, each node's topology database may include information regarding that node's ancestors and descendants. In the event that a child node stops receiving signals from its parent node, the child may, in one example, send a "complaint" message to its grandparent. The grandparent may check whether the parent is still there. If it is not, then the grandparent may send a propagation signal to the child nodes of the missing parent.

Further, if the grandparent detects that the parent node is actually still there, then the grandparent node may send a disconnect signal to the parent node (e.g., sending it back to the server to begin the connection process again). This may occur, for example, when one of the two following conditions is exists: (i) the child node is the only child node of the parent, or (ii) the child and its sibling(s) are complaining to the grandparent.

In this example, the grandparent would also send a propagation signal to the child nodes of the disconnected parent, and a reconfiguration event would occur.

However, if the child node has siblings and they are not sending complaint signals to the grandparent, then the grandparent may assume that the problem is with the complaining child node. Thus, grandparent may send a disconnect signal to the complaining child node (e.g., sending it back to the server to begin the connection process again). If the complaining

child node had its own child nodes, they would contact the departed child node's parent to complain, starting a reconfiguration event (in another example (which example is intended to be illustrative and not restrictive), the grandparent would send out signals to the malfunctioning parent (instructing it to depart) and to the children (instructing the green child to climb its path to its grandparent and instructing the red child to cross connect to the green child).

The foregoing example (which example is intended to be illustrative and not restrictive) can be described in connection with FIG. 30 (which depicts a tree in a binary system having a primary server 11, in which node A, a first level node 12, has two child nodes B and B' (which are second level nodes 13); node B has two child nodes C and C' (which are third level nodes 14); node C has two child nodes D and D' (which are fourth level nodes 15); and node D has two child nodes E and E' (which are fifth level nodes 17)) and FIG. 31 (which is a flow diagram showing an example Grandparent's Complaint Response Routine).

Who does node C complain to in this example when node C no longer gets satisfactory service from node B? Node C complains to its grandparent, node A. If node C does not hear back from node A within a predefined amount of time (e.g., if node A has left the network), node C will then exit the network and immediately reconnect as a new node by going to the primary server S for assignment.

What does node A do in response to a complaint from node C in this example? When node A receives a complaint from grandchild node C about node C's parent node B (i.e., node A's child node B), node A will choose to either remove its child node B or its grandchild node C. A will make this determination based on whether node C alone, or both node C and its sibling node C' are experiencing problems with node B, together with the knowledge of whether node A is continuing to receive satisfactory service reports from node B. If node A is continuing to get satisfactory service reports from node B and there is no indication that node C' (the sibling of node C) is experiencing problems with node B, then node A will assign the "fault" to node C and issue a disconnect order for its removal. At this point the green child of node C (i.e., node D or D') will move up a level connecting to parent B, and the red child of node C (the other of node D or D') will connect as a child of its former sibling. The reconfiguration event will then propagate as discussed above.

If, on the other hand, node A is not getting satisfactory service reports in this example from node B and/or a complaint from node C's sibling arrives within a `window` of node C's complaint, then node A will assign the "fault" to node B and issue a disconnect order for its

Ex. 1018-1736

removal. At this point the green child of node B (i.e., node C or C') will move up a level connecting to grandparent node A, and the red child of node B (the other of node C or C'), will connect as a child of its former sibling. The reconfiguration event will then propagate as discussed above.

An exception to the above is the case where node C is the only child of node B. Under these circumstances node B may be disconnected in this example by node A based solely on the recommendation of node C. In another example, the system will not disconnect a node with one child based solely on the complaint of that child (in this case there may be insufficient data to accurately assign blame - disconnecting the node being complained against and promoting the complainer may have the potential to promote a node that is experiencing problems).

Referring now to FIG. 31 (showing a flowchart associated with this example), node A receives a complaint from node C about node B in step 311. Node A then goes to step 312 in which it checks its topology database to determine whether node C is the only child of node B. If the answer is no (as shown in FIG. 30), then node A goes to step 313 in which node A determines whether there is a communication problem between it and node B. If the answer is no, then node A proceeds to step 314 in which it determines whether it has received a similar complaint about node B from node C' within a predetermined period of time of node A's having received the complaint from node C. If the answer to that query is no, then node A goes to step 315 in which it sends a disconnect signal to node C. At that point node A has completed the Grandparent's Complaint Response Routine (in one example a propagation signal does not necessarily have to sent by node A to node C's child nodes - if any exist, they will complain to node B which, upon discerning that node C is no longer in the distribution network, will send a propagation signal to such complaining nodes).

If the response to any of the queries in steps 312, 313 and 314 is yes in this example, then node A proceeds to step 316 in which it disconnects node B. Node A then proceeds to step 317 in which it sends a propagation signal to its grandchild nodes (here nodes C and C') and a reconfiguration event will occur as described above.

The method of this example described above takes a "2 out of 3" approach on the question of node removal. In this regard, it is believed that many interior nodes in a binary tree distribution network constructed pursuant the present invention will have direct connections to three other nodes: their parent, their green child, and their red child. The connections between a node and its three "neighbor" nodes can be thought of as three separate communications channels. In this example a node is removed from the network

Ex. 1018-1737

when there are indications of failure or inadequate performance on two of these three channels. When there are indications that two of the channels are working normally, the complaining node is presumed to be "unhealthy" and is removed.

In the case where an interior node has only a single child, the "complaint" of that child to its grandparent is sufficient, in this example, to remove the parent node - even when the node to be removed is communicating perfectly well with it's own parent. In other words, given the communication chain A-B-C where node A is the parent of node B and node B is the parent of node C, and given that node C has no siblings, then a complaint of node C to node A will cause node A to remove node B in this example regardless of the fact that node A and node B are not having communication problems (this course of action may help ensure that a potentially unreliable node does not move up the hierarchy of nodes, even at the cost of occasionally sending a "healthy" node back to the edge of the distribution chain based on uncorroborated complaints of its child). Of course, as mentioned above, there are examples where the system will not disconnect a node with one child based solely on the complaint of that child.

In another example of voting (which example is intended to be illustrative and not restrictive), instead of (or in conjunction with) nodes voting on the behavior of other nodes, nodes may continually grade their own performance and if they detect that there is insufficient bandwidth on two or more of their links (links to their parent and their children) the node will perform a graceful depart from the network (initiating a reconfiguration as described elsewhere) and then the node will either climb its path to an ancestor node (or go directly to the server) to request a connection path to the edge of the tree. This system allows nodes to "step aside" when they detect connection / bandwidth related issues.

Referring now to a "Stepped Delay of Playing Content Data" example (which example is intended to be illustrative and not restrictive) it is noted that in order to have all the user nodes experience the playing of the content data at approximately the same time, regardless of the level in which the node resides, the content data buffer 125 shown in FIG. 13 of a user node may be larger the higher uptree it is in a distribution chain (i.e., the closer it is to the root server). The larger the buffer, the greater the amount of time between a particular event (or from the node's receiving the content data), and the node's actually playing the content data in its content data buffer. To the extent that the content data buffer of a node is sized to vary the time that the playing of the content data is started, the content data buffer is a delay. Alternatively, the delay may be a timer or a delay buffer 126 as shown in

Ex. 1018-1738

FIG. 13, or a combination of elements. The delay's purpose may be to help assure that all users experience the play of the content data approximately simultaneously.

For a distribution network having n levels of user nodes cascadingly connected to a server node, where n is a number greater than one (and therefore the distribution network comprises the server node and first through nth level user nodes), the period of time (i.e., the delay time) between:

(i) the moment content data is received by a node or, from a predetermined moment such as, for example, a reporting event or utility rating event (the occurrence of which may be based on time periods as measured by a clock outside of the distribution system and would occur approximately simultaneously for all nodes); and

(ii) the playing of the content data by the node;

is greater for a node the higher uptree the node is on the distribution chain. That is, the delays in first level nodes create greater delay times than do the delays in second level nodes, the delays in second level nodes create greater delay times than the delays in third level nodes and so on. Described mathematically, in an n level system, where x is a number from and including 2 to and including n, the delay time created by a delay in an (x-1) level node is greater than the delay time created by a delay in an x level node.

By way of example (which example is intended to be illustrative and not restrictive), first level nodes could have a built in effective delay time in playing content data of one-hundred-twenty seconds, second level nodes could have a built in effective delay time of one-hundred-five seconds, third level nodes could have a built in effective delay time of ninety seconds, fourth level nodes could have a built in effective delay time of seventy-five seconds, fifth level nodes could have a built in effective delay time of sixty seconds, sixth level nodes could have a built in effective delay time of forty-five seconds, seventh level nodes could have a built in effective delay time of thirty seconds, and eighth level nodes could have a built in effective delay time of fifteen seconds.

The delays may take advantage of the fact that transmission of the packets forming the content data from one node to another may take less time than the actual playing of such packets. This may allow for essentially simultaneous playing of the content data throughout all the levels of the distribution network. The delays may also allow nodes to handle reconnection and reconfiguration events without any significant loss of content data.

In another example (which example is intended to be illustrative and not restrictive), the system may not endeavor to have buffers closer to the server contain more data than buffers of nodes that are further away. Instead, every node may have a buffer that can hold,

for example, approximately two minutes worth of A/V content. The system may endeavor to fill every node's buffer as rapidly as possible, regardless of the node's distance from the server, in order to give each node the maximum amount of time possible to deal with network reconfiguration and uptree node failure events.

Of note, nodes may receive a "current play time index" when they connect to the server. This value may be used to keep all nodes at approximately the same point in the stream – to have all nodes playing approximately the same moment of video at the same time – by telling them, essentially, begin playing A/V content at this time index. The play time index may be, for example, approximately two minutes behind "real time" – allowing for nodes to buffer locally up to two minutes of content (the present invention may permit drift between nodes within the buffer (e.g., two minute buffer) or the present invention may reduce or eliminate this drift to keep all nodes playing closer to "lock step").

Turning now to a general description of an example process for Node X to enter a network including server Node S, Node X must contact Node S to initiate the process. For illustrative purposes, in the following example (which example is intended to be illustrative and not restrictive) each node in the network may have no more than two connections, and the root server (i.e., the highest node in the tree) shall be referred to as Node S.

The process of this example begins with Node X making a request of Node S to join the network.

If Node S has an open connection (e.g., including one or more connected bumpable non-repeat-capable nodes which may be replaced due to an incoming connection request from a repeat-capable node), Node S can accept Node X's request to join the network, and Node X becomes a child of Node S.

If, on the other hand, Node S does not have an open connection, Node S may reject Node X's request, and may provide to Node X a node-by-node path from Node S to Node P, which Node S understands to be in the network, and under which, as determined by Node S, Node X may be able to join the network.

The process of this example continues with Node X making a request of Node P to join the network.

If Node P does not respond – regardless of why Node P does not respond – Node X may request connection from the next higher node in the path. For the purpose of this illustration, thereafter the next higher node is now considered Node P. Node X may make such a request of successive Node Ps, until Node X receives a response.

Ex. 1018-1740

If a particular Node P has an open connection, that Node P can accept Node X's request to join the network, and Node X becomes a child of that particular Node P.

If, on the other hand, a particular Node P does not have an open connection, that particular Node P may reject Node X's request, and may provide to Node X a node-by-node path from that particular Node P to Node P', which the instructing Node P understands to be in the network, and under which, as determined by the instructing Node P, Node X may be able to join the network.

If Node X receives a further rejection, Node X may then treat the Node P' as a Node P and continue the process.

In general, under this example the process is continued until Node X finds a connection.

Referring now more specifically to a "Registration Process" example involving two arbitrary Nodes X and N (neither of which is necessarily a server), FIG. 32 shows Node X sending a request to Node N for acceptance into the distribution network tree (of which Node N is part).

In this example, Node N (which may be thought of as an instructing node) may generally return one of three codes (or messages): Accept, Path, or Reject. FIGS. 33A-33C show example messaging between Node X and Node N under this example. More particularly:

- FIG. 33A shows Node X sending a Register message to Node N and Node N responding to the Register message with an Accept message
- FIG. 33B shows Node X sending a Register message to Node N and Node N responding to the Register message with a Path message (e.g., identifying a path from Node N to a new instructing node)
- FIG. 33C shows Node X sending a Register message to Node N and Node N responding to the Register message with a Reject message (i.e., rejecting Node X's request for a connection to Node N).

In another example, the registration process may be extended to include some system maintenance or configuration activities. Such maintenance or configuration activities may include "flood testing" (discussed in detail below), identification of a bad protocol, and/or Firewall Status (discussed in detail below).

In this regard, FIG. 34A shows the messaging (in connection with registration with a flood test) between Node X, a node requesting to join the network, Node S, a server node

Ex. 1018-1741

within the network, and Node R (e.g., a random (or predetermined) node that may, but not need be, within the network).

Further, FIG. 34B shows the messaging (in connection with registration with a bad protocol version) between Node X and Node S, a server node within the network. Of note, such a bad protocol message may be sent by the server to a connection requesting node to inform that node that it is running an obsolete version of the client software and must upgrade to the current version before being allowed to join the network.

Further still, FIG. 34C shows the messaging (in connection with registration with standard response) between Node X and Node S, a server node within the network (wherein such standard response includes, for example, Accept / Reject / Path).

Referring now in more detail to the "flood test" example shown in FIG. 34A, this example process begins, as above, with Node X making a request to Node S to join the network. Node S then sends Node X a Meta message (such a Meta message may be sent by the server to a connection requesting node to inform that node of certain characteristics of the network such as, for example: the streaming bandwidth rate used by the network, the server's public key for authenticating that secure packets originated from this server, and any other information the architects of the network might deem necessary or desirable for nodes to know when connecting to the network). Node S may also require that Node X undergo performance or flood testing before being accepted or placed within the network. In such an event, Node S responds to Node X's request to join the network with an instruction that Node X undergo a flood test (e.g., of Node X's upstream channel). Such an instruction may be accompanied by an address of Node R, the illustrative flood test target. Note that any node within or without the network may conduct the flood test. One purpose of the flood test may be to permit Node S to be aware of Node X's upstream capability before determining where (or if) to assign Node X in the network.

Where Node S requires that Node X undergo performance or flood testing before being accepted or placed within the network, Node S provides a message so informing Node X, shown as "DoFlood" in FIG 34A. The DoFlood message may provide to Node X the address of the flood test target, but the flood test target could be Node S itself, or a predetermined flood test target.

In response to the DoFlood message, Node X sends an InitFlood message to the flood test target. In one embodiment, Node X would await an acknowledgement (not shown) of its InitFlood message, but in another embodiment, awaiting an acknowledgement is not

necessary. Node X would then begin transmitting FloodData messages, as shown in FIG 34A.

Since, in this example, FloodData messages are transmitted for a purpose of determining Node X's upstream bandwidth, in one embodiment, Node X would transmit an amount of FloodData to the flood test target as shown in FIG 34A. In various embodiments, the maximum amount of FloodData may be limited to a fixed number of FloodData messages, to a fixed amount of data transmitted within FloodData messages, or to transmission of FloodData or FloodData messages over a fixed period of time.

To protect the flood test target from being inundated with FloodData (e.g., where Node X has a very high upstream bandwidth capability), Node X may cap (e.g., be so directed by Node S) the FloodData transmission to, e.g., 100 Kbits/sec, 200 Kbits/sec, or some other data rate that would be a desired data transmission rate for a node on the particular network of interest.

Thus, by receiving the FloodData messages, the flood test target can determine, for example, an upstream capability rating of Node X (provided that the flood test target has the same or more bandwidth in the downstream channel compared to the upstream capability of Node X.

In another example, both the target and the sender will know the bandwidth of the sender at the conclusion of the flood test. More particularly, the present invention may use TCP/IP (rather than UDP), and all packets sent may receive an acknowledgement. In this example, Node X, the node sending out the flood packets, sends out the next flood packet only after receipt of the first packet has been acknowledged by the target of the flood test. Thus, the number of flood data packets Node X is able to send per unit time (along with the size of those packets) may allow Node X to calculate its upstream bandwidth.

In another example, the target of the "flood test" may be any node including the server but excluding any firewalled or low-bandwidth nodes.

After the "flood test", Node X may then connect with Node S (or be instructed to go to another node if Node S can not/will not accept Node X). Of note, if Node X fails the "flood test", Node X may be sent far downtree or may be refused entry into the network entirely. Further, if node N is accepted under these circumstances it will be deemed a "bumpable" non-repeat-capable node.

Referring now in more detail to the "Firewall Status" example shown in FIG. 34C, this example process begins, as above, with Node X making a request to Node S to join the network. Node S then send Node X a Meta message. A firewall status message is a message

sent over a new connection from (in the case of Figure 34C) Node S to Node X. If Node X receives this message, over a communications channel initiated by Node S, Node X will know that it is not firewalled. Failure to receive a firewall status message when a node is expecting such a message allows that node to conclude that it is in fact firewalled. Thus, if Node X is expecting to receive a firewall status message over new connection initiated by some other node and that message does not arrive, Node X can conclude (with a high degree of likelihood of being correct) that it is firewalled. Also, since (in one example) the present invention uses the TCP/IP messaging protocol the sender, Node S, will also know whether the firewall status message was successfully received by Node X and thus Node S will also know whether Node X is or is not firewalled (one or more other nodes may also receive data regarding the firewall status).

After receiving the firewall status report (not shown) Node S will then respond with a standard response (accept, path, reject) as discussed above.

To now give a specific example of the use of both the "flood test" and the "firewall status" mechanisms during registration of a new node, FIG. 35A shows a state of the network before Node X joins the network and FIG. 35B shows a state of the network after Node X joins the network. Further, FIG. 36 shows the messaging among Node X, Node S, Node B and Node R. More particularly:

- Node X sends a Register message to Node S and Node S responds with a Meta message (providing Node X with basic information about the network) and a DoFlood message (which requests Node X to determine its bandwidth and report back to Node S when Node X has that information). Node S knows to send a "DoFlood" to Node X because the initial Register message sent by Node X to Node S contains an indication that Node X does not yet know its current upstream bandwidth capacity. The DoFlood message will contain the address of a target node (Node R), selected by Node S, that Node X should flood against.
- Node X sends an InitFlood message to Node R and then "floods" Node R with data for some period of time (e.g., 15 seconds). At the conclusion of this test Node X will be able to compute its upstream bandwidth based on the number and size of data packets it was able to successfully transmit to Node R.

Ex. 1018-1744

- Node X sends another Register message to Node S and Node S responds with a Meta message. Contained within this second Register message is the result of Node X's "flood test" which gives an indication of Node X's upstream bandwidth capacity.
- Node S now tests whether external nodes can initiate communication with Node X by sending Node X a Firewall Status message over a new connection. If this message is successfully received by Node X (and acknowledgement received by Node S), both Nodes X and S will know that Node X is not firewalled (on a network communication port).
- Node S sends a Path message to Node X. This message will include a list of nodes that X should attempt connection with, beginning with the preferred target node (e.g., Node B). As described elsewhere in the application, the choice of target node, and hence the contents of the path, may be influenced by both the current configuration of the network and whether or not Node X is repeat-capable (based on the results of the "flood test" and "firewall test").
- Node X sends a Register message to Node B and Node B responds with an Accept message.

Referring now to a "Firewall Request / Firewall Status" example involving two arbitrary Nodes A and B (neither of which is necessarily a server), FIG. 37 shows a portion of a network including Node A and Node B. In this example (as shown in FIG. 38), a Firewall Request message is sent by Node B to request that the recipient of the message, Node A, test Node B's Firewall Status. The receiving Node A may send (e.g., over a new connection), the Firewall Status message to Node B. If Node B timely receives and interprets the Firewall Status message it can be certain that it is a non-firewalled node. If Node B does not timely receive a Firewall Status message from Node A then Node B may assume that it is firewalled.

In another example (which example is intended to be illustrative and not restrictive), if Node A fails to connect to Node B or to send the Firewall Status message no further action is needed on the part of Node A.

Node B may also augment its assumption about its Firewall Status by using a three strikes (i.e., three attempts) method before considering itself to be firewalled (of course, three strikes is provided as an example only, and any desired number of "strikes" or attempts may be used).

In another example (which example is intended to be illustrative and not restrictive), the multiple attempts may be carried out within a predetermined time "window" (after which receiving no Firewall Status message Node B may consider itself firewalled).

In another example (which example is intended to be illustrative and not restrictive), Firewall Requests may be sent out by all nodes (e.g., at regular time intervals) in order to maintain a current and accurate picture of their Firewall Status.

In another example (which example is intended to be illustrative and not restrictive), Firewall Requests may be sent out by certain node(s) (e.g., at regular time intervals) in order to maintain a current and accurate picture of their Firewall Status.

In another example (which example is intended to be illustrative and not restrictive), the Firewall Status message may be sent by the server on initial registration (e.g., without the need for a Firewall Request message and both the sender (the server) and the receiving node (the node that is registering) may use the successful delivery of the Firewall Status message as an indicator of the connecting node's Firewall Status.

Referring now to a "Propagate Path / Down" example (which example is intended to be illustrative and not restrictive), FIG. 39A shows a portion of a network including Node A and Node B. Further, FIG. 39B shows an example of nodes in a tree sending Propagate Path / Down messages downtree to their children (e.g., at regular time intervals).

In one example (which example is intended to be illustrative and not restrictive), the Propagate Path / Down message may include topology information containing a path back to the root server and the time interval for sending such message may be calculated on a node-by-node and/or level-by-level basis relative to the amount of data to be sent in the Propagate Path / Down message (wherein the amount of data in the Propagate Path / Down message is related to a particular node's (or level's) downtree position relative to the root server). Note, that although this time interval may not be very important for Propagate Path / Down, since the amount of data for each successive downtree node (or level) may increase relatively slowly (e.g., linearly), the time period may be of critical importance for Propagate NTM / Up (which is discussed below), since the amount of data for each successive uptree node (or level) may increase relatively quickly (e.g., exponentially). Of note, the acronym "NTM" refers to the Network Topology Model.

In any case, FIG. 40 shows the messaging between Node A and Node B in connection with a Propagate Path / Down message (i.e., Node A sends a Propagate Path message downtree to Node B).

Ex. 1018-1746

Referring now to a "Propagate NTM / Up" example (which example is intended to be illustrative and not restrictive), FIG. 41A shows a portion of a network including Node A and Node B. Further, FIG. 41B shows an example of nodes in a tree sending Propagate NTM / Up messages uptree to their parents (e.g., at regular time intervals).

In one example (which example is intended to be illustrative and not restrictive), the Propagate NTM / Up message may include information containing the topology of the subtree(s) below a particular node and the time interval for sending such message may be calculated on a node-by-node and/or level-by-level basis relative to the amount of data being sent in the Propagate NTM / Up message (wherein the amount of data in the Propagate NTM / Up message is related to a particular node's (or level's) uptree position relative to tree height (e.g., the height of a particular node above its lowest leaf level)). Note that in this example (which example is intended to be illustrative and not restrictive), the NTM data that the node sends up is rooted at itself and contains both its children.

In any case, FIG. 42 shows the messaging between Node A and Node B in connection with a Propagate NTM / Up message (i.e., Node B sends a Propagate NTM message uptree to Node A).

In another embodiment, the following process may be employed when a node enters the network. More particularly, when the server adds a node X to the system (gives it a connection path to node Y) the server will add node X to its network topology model as a child of node Y. In other words, the server will assume that the connection will take place normally with no errors. The reason why doing so may be critical to the functioning of the system in this embodiment is that if the server failed to account for the fact that node X is now (probably) a child of node Y, using the "keep the tree balanced" approach embodied in the Universal Connection Address Routine, all new nodes connecting to the system would be assigned as children of node Y up until the point where a propagate up arrives indicating that node X was successfully added to the tree. If the tree is large and node Y is many levels removed from the server it could be some time (e.g., a minute or more, in very large trees several minutes) before confirmation arrives that node X truly connected to node Y as its child.

Now, in addition to simply placing node X into the server's NTM (Network Topology Model) and being done with the matter, in this embodiment the system must account for the fact that it may take a number of propagate NTM/up cycles for the true status of node X to be reported back to the server. Thus, for some number of propagate cycles (or some amount of time) the server may compare the newly propagated (updated) NTM it is receiving from its

Ex. 1018-1747

children to the list of "recently added nodes". If a recently added node is not in the propagated NTM it will be re-added to the NTM as a child of node Y. At some point node X will no longer be considered "recently added" and the server will cease adding node X back into its NTM. In other words, the server will assume that, for whatever reason, node X never successfully connected to the network.

In another embodiment, the following process may be employed when a node gracefully exits the system. More particularly, when a node (node X) leaves the network by performing a graceful exit (e.g., the user selects "exit" or clicks the "x" or issues some other command to close the application) the node will send signals to its parent and child nodes (if any) immediately prior to actually shutting down. Upon receipt of such a "depart" message, the parent of the departing node (node P) will modify its internal NTM (Network Topology Model) to a configuration that corresponds to the predicted shape of the network after the departing node exits. In this embodiment, node P will assume that its NTM is accurate and that the entire cascade of reconfiguration events initiated by the departure of node X will happen normally. As such, node P will now expect the former green child of node X (node XG -- if such a child exists) to connect – essentially taking node X's place in the network.

At this point, node P's NTM represents a prediction of the (anticipated) state of the network rather than the actual reported configuration. In order to indicate that node P does not (yet) have a physical connection to node XG, node P will mark node XG as "soft" in its NTM – meaning node XG is expected to take up this position in the network but has not yet physically done so. Node P will report this "predicted" NTM (along with node XG's status as a soft node) up during its propagate NTM/Up cycles, until node P physically accepts a child connection in node XG's "reserved" spot and receives an actual report of the shape of the downtree NTM from that child, via propagate NTM/ups.

Of note, the above actions may be taken in order to have the propagated NTM's reflect, as accurately as possible, the topology of the network.

Of further note, because of the distributed nature of a network according to the present invention, a complete and accurate picture of the current shape of the network at any instance in time can never be guaranteed. In fact, though the root server has the "widest" view of the network, without the heuristics described herein, that view would be the "most dated" of any node in the network. Thus, heuristics such as those just described may be vital to producing NTM's that are as close to complete and accurate as possible (e.g., so that the Universal Connection Address Routine can position incoming nodes so as to keep the distribution tree as nearly balanced as possible).

Referring now to a "Banned List" example (which example is intended to be illustrative and not restrictive), it is noted that such a "banned list" (or "banished list") may be used for keeping certain nodes off of the network, removing certain nodes from the network, and/or keeping certain nodes as far downtree as possible (e.g., in the case of a binary tree network keeping certain nodes as "red" nodes).

More particularly, FIG. 43 shows a portion of a network including Node A, Node B, Node C and Node D and FIG. 44 shows the messaging between Node A, Node B, Node C and Node D in connection with Banned List messages (i.e., in this FIG. 44, Node A sends a Banned List message to Node B and Node B sends a Banned List message to Nodes C and D).

In one example (which example is intended to be illustrative and not restrictive), nodes receiving this message should re-transmit the message downtree to all children in an unaltered fashion.

In another example (which example is intended to be illustrative and not restrictive), the Banned List messages may include IP lists and/or URL's.

In another example (which example is intended to be illustrative and not restrictive), the Banned List messages may originate at the server.

In another example (which example is intended to be illustrative and not restrictive), Banned List messages may be sent and/or originated by any desired node(s).

In another example (which example is intended to be illustrative and not restrictive), Banned List messages and may be sent using a Secure Packet format.

In another example (which example is intended to be illustrative and not restrictive), node(s) may store the content of the message (e.g., the banned list itself) for future use.

In another example (which example is intended to be illustrative and not restrictive), the server (or any other desired node(s)) may send out and/or originate the Banned List messages at regular time intervals.

Referring now to a "Bumpable Node" embodiment, it is noted that this "bumpable node" mechanism may be used to "bump" (or send downtree) a particular node (e.g., when a new node is added). More particularly, in one example (which example is intended to be illustrative and not restrictive), a server may instruct Node X to connect with Node P and Node P may know to "bump" its child, Node C, by: (a) instructing Node C to disconnect from Node P; (b) allowing Node X to connect to Node P; and (c) instructing Node C to re-connect as a child of Node X. Of note, such a "bumping" operation may be carried out, for example (which example is intended to be illustrative and not restrictive), because the

bumped node is firewalled and/or because the bumped node has failed an upstream and/or downstream bandwidth test (which test may be carried out by the server or by any other node(s)).

Of note, as used in this application, the terms "bumpable nodes" or "bumpable non-repeat-capable nodes" or "non-repeat-capable nodes" are nodes which either: (a) have insufficient bandwidth to act as repeaters in the network or (b) are firewalled.

Of further note, in one example the more inclusive "is bumpable" test may replace the "is low bandwidth" test in the Universal Connection Address Routine of Figure 20.

Referring now to a "Load Manager" embodiment, it is noted that this "load manager" mechanism may be used to rate limit (i.e., slow or stop) the addition of new nodes to the network. Such rate limiting may be useful because the larger the network grows, the more nodes can join per unit of time.

In one example (which example is intended to be illustrative and not restrictive), the rate limited addition of new nodes may be carried out by a server (or connection server) and may be based upon the length of time that it takes for the buffers associated with one or more nodes (or levels) of the existing network to fill (or reach a certain value). In a more specific example, the rate limited addition of new nodes may be based upon the length of time that it takes for the buffers associated with potential parents of the new nodes to fill (or reach a certain value).

In another example (which example is intended to be illustrative and not restrictive), the rate limited addition of new nodes may be carried out by any desired node(s) (e.g., one or more user nodes and/or a combination of one or more user nodes and a server (or connection server) node).

Referring now to a "Node Interlacing" embodiment, it is noted that this "node interlacing" mechanism may be used to add new nodes to the network by interlacing the new nodes on a level-by-level basis as shown in FIG. 45.

More particularly, as seen in this FIG. 45, new nodes may be added in the order indicated (as can be seen below, this example (which example is intended to be illustrative and not restrictive) relates to a binary tree network). That is, Node 1 may first be added as a child of the server followed by Node 2 as a child of the server – this will fully populate the first level; Node 3 may next be added as a child of Node 1, followed by Node 4 as a child of Node 2, followed by Node 5 as a child of Node 1, followed by Node 6 as a child of Node 2 – this will fully populate the second level; Node 7 may next be added as a child of Node 3, followed by Node 8 as a child of Node 5, followed by Node 9 as a child of Node 4, followed

by Node 10 as a child of Node 6, followed by Node 11 as a child of Node 3, followed by Node 12 as a child of Node 5, followed by Node 13 as a child of Node 4, followed by Node 14 as a child of Node 6 – this will fully populate the third level, etc.

Of note, such node interlacing may be used to efficiently fill (or bring to a certain value) the buffers associated with one or more nodes (or levels) of the existing network while new nodes are added.

Of further note, such node interlacing may help prevent adding a second child to a parent until the first child's buffer is full (or brought to a certain value).

Of still further note, such node interlacing may help to: (a) prevent taxing the parent's downtree channel; and/or (b) prevent two children being left with partially filled buffers in the event that their parent departs.

Of still further note, such node interlacing may be used, for example, in conjunction with or as an alternative to the load balancing mechanism discussed above.

In another embodiment the present invention may be used when a plurality of computer systems are connected (e.g., to the Internet) by one or more of an ISDN line, a DSL (e.g., ADSL) line, a cable modem, a T1 line and/or an even higher capacity link.

In another embodiment a distribution network for the distribution of data from a server node to a plurality of user nodes connected to the server node, wherein each of the user nodes has associated therewith a buffer is provided, which distribution network comprises: a mechanism for rate limiting the addition of new user nodes to the distribution network, wherein the rate limited addition of new user nodes to the distribution network is based at least in part upon the length of time that it takes for at least one buffer associated with at least one user node to reach a predetermined percentage of capacity, wherein the predetermined percentage of capacity is greater than zero.

In another embodiment a distribution network for the distribution of data from a server node to a plurality of user nodes connected to the server node, wherein each of the user nodes is in one of a plurality of levels located downtree from the server node is provided, which distribution network comprises: a mechanism for adding new user nodes to the distribution network in an interlaced manner, wherein the interlacing provides that each sibling user node in a given level of the distribution network is associated with just one child before any siblings in the given level are associated with two children.

The following embodiments will now describe a number of additional example algorithms which may be utilized in whole or in part to implement the present invention:

In one embodiment a process for docking a connection requesting user node with a distribution network is provided, said process including the following steps:

(a) having an instructing node receive a connection request from said connection requesting user node;

(b) forming an instructing node topology database indicating, at a point in time:

(i) which, if any, user nodes are docked with said instructing node as child nodes; which, if any, user nodes are docked with each of said child nodes as grandchild nodes of said instructing node; which, if any, user nodes are docked with each of said grandchild nodes as great-grandchild nodes of said server node; and so on; and

(ii) said docked user nodes respective bandwidth capacities;

(c) selecting from a group of nodes including the instructing node and said docked user nodes a recommended parent node for said connection requesting user node, wherein said recommended parent node has apparent available capacity to transmit content data to said connection requesting user node and is at least as close (i.e., in terms of network topology) to said instructing node as any other docked user node having apparent available capacity to transmit content data to said connection requesting user node (of note, placement may also or alternatively depend, as described elsewhere in the present application, in whole or in part on whether the connection requesting node is repeat-capable or not (i.e., whether it can support children)); and

(d) providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said recommended parent node back to said instructing node.

In one example, the process may further comprise the following steps:

(e) having said connection requesting user node go to (or attempt to go to) the node at the top of said connection address list;

(f) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(g) if the node at top of said connection address list is not part of the distribution network, deleting such node from said connection address list and repeating steps (e) and (f) with respect to the next node at the top of said connection address list;

(h) if the node at the top of said connection address list is part of said distribution network, having said connection requesting user node dock with said node at the top of said connection address list.

In another example, step (d) may further comprise the following step:

(i) if said instructing node is not a server node, including in said connection address list a list of nodes cascadingly connected with each other from the instructing node back to said server node.

In another example, the process may further comprise the following steps:

(e) having said connection requesting user node go to (or attempt to go to) the node at the top of said connection address list;

(f) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(g) if the node at the top of said connection address list is not part of the distribution network, deleting such node from said connection address list and repeating steps (e) and (f) with respect to the next node at the top of said connection address list;

(h) if the node at the top of said connection address list is part of said distribution network, determining whether said node at the top of said connection address list has available capacity to transmit content data to said connection requesting user node; and

(i) if the node at the top of said connection address list is part of said distribution network and has available capacity to transmit content data to said connection requesting user node, having said connection requesting user node dock with the node at the top of said connection address list.

In another example, the process may further comprise the following steps:

(j) if the node at the top of connection address list is part of said distribution network and said node at the top of said connection address list does not have available capacity to transmit content data to said connection requesting user node, repeating steps (a)-(d), with said node at the top of connection address list being said instructing node and a new connection address list being the product of steps (a)-(d).

In another example, the process may further comprise the steps of counting the times that step (j) is performed and, when step (j) is performed a predetermined number of times, having said connection requesting user node repeat steps (a)-(d) with said primary server being the instructing node.

In another embodiment a process for docking a connection requesting user node with a distribution network is provided, said process including the following steps:

(a) having an instructing node receive a connection request from said connection requesting user node;

(b) forming an instructing node topology database indicating, at a point in time:

(i) which, if any, user nodes are docked with said instructing node as child nodes; which, if any, user nodes are docked with each of said child nodes as grandchild nodes of said instructing node; which, if any, user nodes are docked with each of said grandchild nodes as great-grandchild nodes of said server node; and so on; and

(ii) said docked user nodes' respective bandwidth capacities;

(c) assigning utility ratings to each of the docked user nodes, said utility ratings being a function of at least the bandwidth capacity of the docked user node and the elapsed time that the docked user node has been docked with the distribution network;

(d) selecting from a group of nodes including the instructing node and said docked user nodes a recommended parent node for said connection requesting user node, wherein said recommended parent node has apparent available capacity to transmit content data to said connection requesting user node and is at least as close (i.e., in terms of network topology) to said instructing node as any other docked user node having apparent available capacity to transmit content data to said connection requesting user node, and when the closest (i.e., in terms of network topology) docked user node with apparent available capacity to transmit content data to said connection requesting user node is equidistant (i.e., in terms of network topology) to said instructing node with at least one other docked user node with apparent available capacity to transmit content data to said connection requesting user node, said recommended parent node is the docked user node having the highest utility rating among such equidistant (i.e., in terms of network topology) docked user nodes with apparent available capacity; and

(e) providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said recommended parent node back to said instructing node and, if said instructing node is not a primary server node, back to said primary server node.

In another embodiment a process for docking a connection requesting user node with a distribution network is provided, said process including the following steps:

(a) having an instructing node receive a connection request from said connection requesting user node;

(b) forming an instructing node topology database indicating, at a point in time:

(i) which, if any, user nodes are docked with said instructing node as child nodes; which, if any, user nodes are docked with each of said child nodes as grandchild nodes of said instructing node; which, if any, user nodes are docked with

each of said grandchild nodes as great-grandchild nodes of said instructing node; and so on; and

(ii) said docked user nodes' respective bandwidth capacities, including a designation of whether said respective docked user node's bandwidth capacity is below a predetermined threshold (e.g., said respective docked user node is a low-bandwidth node) or at least said predetermined threshold (e.g., said respective docked user node is a high-bandwidth node);

(c) forming a primary recommended parent node list ("PRPL") comprised of the instructing node (if it has available capacity to transmit content data to said connection requesting user node) plus those docked user nodes having apparent available capacity to transmit content data to said connection requesting user node, with said instructing node being placed first on the PRPL (if it is on the PRPL) and said docked user nodes having apparent available capacity to transmit content date to said connection requesting user node being ranked with those docked nodes which are closer (i.e., in terms of network topology) to the instructing node being ranked higher than those docked nodes which are further away (i.e., in terms of network topology) and with equidistant (i.e., in terms of network topology) docked user nodes being ranked such that those docked nodes which have the higher utility ratings being ranked higher than docked nodes having lower utility ratings;

(d) forming a secondary recommended parent node list ("SRPL") comprised of the instructing node (if it has no available capacity to transmit content data to said connection requesting user node but does have at least one low-bandwidth node docked directly to it) plus said docked user nodes which (i) are high-bandwidth nodes with no available capacity to transmit content data to said connection requesting user node and (ii) have at least one low-bandwidth node docked directly to it, with said instructing node being placed first on the SRPL (if it is on the SRPL) and said docked user nodes on the SRPL being ranked with those docked nodes which are closer (i.e., in terms of network topology) to the instructing node being ranked higher than those docked nodes which are further away (i.e., in terms of network topology) and with equidistant (i.e., in terms of network topology) docked user nodes being ranked such that those docked nodes which have the higher utility ratings being ranked higher than docked nodes having lower utility ratings;

(e) determining whether the connection requesting user node is a low-bandwidth node;

(f) if the connection requesting user node is a low-bandwidth node:

(i) selecting the highest ranked node on the PRPL as a recommended parent node; and

(ii) providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said recommended parent node back to said instructing node and, if said instructing node is not a primary server node, back to said primary server node; and

(g) if the connection requesting user node is a high-bandwidth node:

(i) selecting the highest ranked node on the PRPL as a recommended parent node;

(ii) selecting the highest ranked node on the SRPL as an alternate recommended parent node;

(iii) determining whether the alternate recommended parent node is closer (i.e., in terms of network topology) to the server node than the recommended parent node;

(iv) if the alternate recommended parent node is not closer (i.e., in terms of network topology) to the server node than the recommended parent node, providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said recommended parent node back to said instructing node and, if said instructing node is not the primary server node, back to said primary server node;

(v) if the alternate recommended parent node is closer (i.e., in terms of network topology) to the server node than the recommended parent node, (1) forcing the disconnection of a low-bandwidth node from the alternate recommended parent node and (2) providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said alternate recommended parent node back to said instructing node and, if said instructing node is not the primary server node, back to said primary server node.

In one example, step (f)(ii) may include the following steps:

(1) adding said connection requesting user node to the topology database as a child of the recommended parent node; and

(2) if said recommended parent node would have no apparent available capacity to transmit content data to an additional node with said connection requesting user node docked with it, deleting the recommended parent node from the PRPL and adding it to the SRPL.

In another example, step (g)(iv) may include the following steps:

(1) adding said connection requesting user node to the topology database as a child of the recommended parent node;

(2) if said recommended parent node would have no apparent available capacity to transmit content data to an additional node with said connection requesting user node docked with it, deleting the recommended parent node from the PRPL; and

(3) if said recommended parent node is deleted from the PRPL but has at least one low-bandwidth node docked directly to it, adding the recommended parent node to the SRPL.

In another example, step (g)(v) may include the following steps:

(3) adding said connection requesting user node to the topology database as a child of the alternate recommended parent node; and

4) if said alternate recommended parent node would have no apparent available capacity to transmit content data to an additional node with said connection requesting user node docked with it, deleting the alternate recommended parent node from the SRPL.

In another example, the process may further comprise the following steps:

(h) having said connection requesting user node go to (or attempt to go to) the node at the top of said connection address list;

(i) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(j) if the node at top of connection address list is not part of the distribution network, deleting such node from connection address list and repeating steps (h) and (i) with respect to the next node at the top of connection address list; and

(k) if the node at top of connection address list is part of said distribution network, having said connection requesting user node dock with said node at the top of connection address list.

In another example, the process may further comprise the following steps:

(h) having said connection requesting user node go to (or attempt to go to) the node at the top of connection address list;

(i) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(j) if the node at top of connection address list is not part of the distribution network, deleting such node from connection address list and repeating steps (h) and (i) with respect to the next node at the top of connection address list;

(k) if the node at the top of connection address list is part of said distribution network, determining whether said node at the top of connection address list has available capacity to transmit content data to said connection requesting user node; and

(l) if the node at the top of said connection address list is part of said distribution network and has available capacity to transmit content data to said connection requesting user node, having said connection requesting user node dock with the node at the top of said connection address list.

In another example, the process may further comprise the following steps:

(m) if the node at the top of said connection address list is part of said distribution network and the node at the top of said connection address list does not have available capacity to transmit content data to said connection requesting user node, repeating steps (a)-(g), with the node at the top of said connection address list being said instructing node and a new connection address list being the product of steps (a)-(g).

In another example, the process may further comprise the steps of counting the times that step (m) is performed and, when step (m) is performed a predetermined number of times, having said connection requesting user node repeat steps (a)-(g) with said primary server being the instructing node.

In another embodiment a process for docking a connection requesting user node with a distribution network is provided, said process including the following steps:

(a) having an instructing node receive a connection request from said connection requesting user node;

(b) determining whether said connection requesting user node's bandwidth capacity is below a predetermined threshold (e.g., said connection requesting user node is a low-bandwidth node) or at least said predetermined threshold (e.g., said connection requesting user node is a high-bandwidth node);

(c) forming an instructing node topology database indicating, at a point in time:

(i) which, if any, user nodes are docked with said instructing node as child nodes; which, if any, user nodes are docked with each of said child nodes as grandchild nodes of said instructing node; which, if any, user nodes are docked with each of said grandchild nodes as great-grandchild nodes of said server node; and so on; and

(ii) said docked user nodes' respective bandwidth capacities;

(d) if said connection requesting user node is a high-bandwidth user node, selecting from a group of nodes including the instructing node and said docked user nodes a recommended parent node for said connection requesting user node, wherein said recommended parent node has apparent available capacity to transmit content data to said connection requesting user node and is at least as close (i.e., in terms of network topology) to said instructing node as any other docked user node having apparent available capacity to transmit content data to said connection requesting user node;

(e) if said connection requesting user node is a low-bandwidth user node, selecting from a group of nodes including the instructing node and said docked user nodes a recommended parent node for said connection requesting user node, wherein said recommended parent node has apparent available capacity to transmit content data to said connection requesting user node and is at least as far (i.e., in terms of network topology) from said instructing node as any other docked user node having apparent available capacity to transmit content data to said connection requesting user node; and

(f) providing a connection address list to said connection requesting user node, said connection address list listing each node in said instructing node's topology database from said recommended parent node back to said instructing node.

In one example, step (f) may further comprise the following step: (i) if said instructing node is not a server node, including in said connection address list a list of nodes cascadingly connected with each other from the instructing node back to said server node.

In another embodiment a process for connecting a connection requesting user node to a computer information distribution network having a primary server node and user nodes docked therewith in a cascaded relationship is provided, said process further comprising the following steps:

(a) providing said connection requesting user node with a connection address list which sets forth a list of user nodes docked in series with each other back to said primary server;

(b) having said connection requesting user node go to (or attempt to go to) the node at top of said connection address list;

(c) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(d) if the node at the top of said connection address list is not part of the distribution network, deleting such node from said connection address list and repeating steps (b) and (c) with respect to the next node at the top of connection address list; and

(e) if the node at the top of said connection address list is part of said distribution network, having said connection requesting user node dock with the node at the top of said connection address list.

In another embodiment a process for connecting a connection requesting user node to a computer information distribution network having a primary server node and user nodes docked therewith in a cascaded relationship is provided, said process further comprising the following steps:

(a) providing said connection requesting user node with a connection address list which sets forth a list of user nodes docked in series with each other back to said primary server;

(b) having said connection requesting user node go to (or attempt to go to) the node at the top of said connection address list;

(c) having said connection requesting user node determine whether the node at the top of said connection address list is part of the distribution network;

(d) if the node at the top of said connection address list is not part of the distribution network, deleting such node from connection address list and repeating steps (b) and (c) with respect to the next node at the top of said connection address list;

(e) if the node at the top of said connection address list is part of said distribution network, determining whether the node at the top of said connection address list has available capacity to transmit content data to said connection requesting user node; and

(f) if the node at the top of said connection address list is part of said distribution network and has available capacity to transmit content data to said connection requesting user node, having said connection requesting user node dock with the node at the top of said connection address list.

In another embodiment relating to a computer information distribution network comprising a primary server node and user nodes docked therewith in a cascaded relationship, wherein each user node may have no more than a predetermined maximum number of child nodes docked with it, a process for reconfiguring the network in the event of a user node's departure therefrom is provided, said process including the following steps:

(a) providing information at a point in time to a node in the distribution network as follows:

(i) the node's propagation rating if it is a user node having a parent node which is a user node, wherein a propagation rating is one of a predetermined number of grades ranging from highest to second highest to third highest and so on, with the number of grades being equal to said predetermined maximum number;

(ii) an ancestor list setting forth the node's ancestor nodes' addresses (if it has any ancestor nodes) back to the primary server node, with the node's parent node being atop the ancestor list;

(iii) a sibling list of the node's sibling nodes' addresses (if it has any sibling nodes) and their respective propagation ratings (of note, in another example (which example is intended to be illustrative and not restrictive), nodes may not have data about their sibling (i.e., nodes do not know who their sibling is); instead, the mechanism used to get a red node to connect to its green sibling may be for the parent of those two nodes (or a proxy for that parent) to send the red child: (1) a connection path with that red child's green sibling at the top of the path; and (2) a command to "priority join" to that green child - when the green child node accepts the incoming "priority join" connection request from its former red sibling, the green node will kick its own red child (if such a child exists) instructing that red child to join as a child of its own green sibling (these actions will cause a chain reaction of reconfigurations which will cascade until the edge of the tree is reached)); and

(iv) a child list of the node's child nodes' addresses (if it has any child nodes) and their respective propagation ratings;

(b) having a first node send a propagation signal to each former child node of a departed node, wherein said first node is an ancestor of said former child node of said departed node;

(c) with respect to each child node receiving the propagation signal which did not have the highest propagation rating prior to its receiving the propagation signal (hereinafter referred to as a "red node"), upon its receiving the propagation signal:

(i) setting the propagation rating of the red node to the next higher grade above the propagation rating of that red node before it received the propagation signal;

(ii) having the red node undock from its parent node (if it was docked with its parent node before it received the propagation signal); and

(iii) having the red node dock with its sibling node which had the highest propagation rating prior to the red node's receiving the propagation signal;

(d) with respect to each child node receiving the propagation signal which had the highest propagation rating prior to its receiving a propagation signal (hereinafter referred to as a "green node"), upon its receiving the propagation signal:

(i) setting the propagation rating of that green node to the lowest grade;

(ii) having the green node retransmit the propagation signal to its child nodes (if it has any); and

(iii) determining whether the green node is docked with the node from which the green node received the propagation signal and if it is, having the green node remain docked with said node from which the green node received the propagation signal, and if it is not, having the green node dock with said node from which the green node received the propagation signal; and

(e) repeating steps (c) and (d) with respect to each user node which receives a retransmitted propagation signal.

In another embodiment relating to a computer information distribution network comprising a primary server node and user nodes docked therewith in a cascaded relationship, wherein each user node may have no more than a predetermined maximum number of child nodes docked with it, a process for reconfiguring the network in the event of a user node's departure therefrom is provided, said process including the following steps:

(a) providing information at a point in time to a node in the distribution network as follows:

(i) the node's propagation rating if it is a user node having a parent node which is a user node, wherein a propagation rating is one of a predetermined number of grades ranging from highest to second highest to third highest and so on, with the number of grades being equal to said predetermined maximum number;

(ii) an ancestor list setting forth the node's ancestor nodes' addresses (if it has any ancestor nodes) back to the primary server node, with the node's parent node being atop the ancestor list;

(iii) a sibling list of the node's sibling nodes' addresses (if it has any sibling nodes) and their respective propagation ratings; and

(iv) a child list of the node's child nodes' addresses (if it has any child nodes) and their respective propagation ratings;

(b) having a first node send a propagation signal to each former child node of a departed node, wherein said first node is an ancestor of said former child node of said departed node;

(c) with respect to each child node receiving the propagation signal which did not have the highest propagation rating prior to its receiving the propagation signal (hereinafter referred to as a "red node"), upon its receiving the propagation signal:

(i) setting the propagation rating of the red node to the next higher grade above the propagation rating of that red node before it received the propagation signal;

(ii) having the red node undock from its parent node (if it was docked with its parent node before it received the propagation signal);

(iii) devising a first connection address list comprising the address of the red node's sibling node which had the had the highest propagation rating prior to the red node's receiving the propagation signal followed by the red node's ancestor list;

(iv) having said red node go to (or attempt to go to) the node at the top of said first connection address list;

(v) having said red node determine whether the node at the top of said first connection address list is part of the distribution network;

(vi) if the node at the top of said first connection address list is not part of the distribution network, deleting such node from said first connection address list and repeating steps (c)(iv) and (c)(v) with respect to the next node at the top of said first connection address list;

(vii) if the node at the top of said first connection address list is part of said distribution network, determining whether the node at the top of said first connection address list has available capacity to transmit content data to said red node; and

(viii) if the node at the top of said first connection address list is part of said distribution network and has available capacity to transmit content data to said red node, having said red node dock with the node at the top of said first connection address list (with regard to these steps (vii) and (viii), in another example (which example is intended to be illustrative and not restrictive), when a red node connects to a green node as part of a network reconfiguration event, it does not matter whether there is "available capacity" or not for the green node to accept its incoming red sibling – the green node has no choice but to accept the node (these actions are facilitated in this example by the red node issuing a "priority join" to the green node; if the green node does not have "available capacity" it has to kick its own red child to make room for the incoming green node));

(d) with respect to each child node receiving the propagation signal which had the highest propagation rating prior to its receiving a propagation signal (hereinafter referred to as a "green node"), upon its receiving the propagation signal:

(i) setting the propagation rating of that green node to the lowest grade;

(ii) having the green node retransmit the propagation signal to its child nodes (if it has any);

(iii) determining whether the green node is docked with the node from which the green node received the propagation signal and if it is, having the green node remain docked with said node from which the green node received the propagation signal, and if it is not, devising a second connection address list starting with said node from which the green node received the propagation signal and followed by that portion of said green node's ancestor list extending back to the primary server from said node from which the green node received the propagation signal;

(iv) having said green node go to (or attempt to go to) the node at the top of said second connection address list;

(v) having said green node determine whether the node at the top of said second connection address list is part of the distribution network;

(vi) if the node at the top of said second connection address list is not part of the distribution network, deleting such node from said second connection address list and repeating steps (d)(iv) and (d)(v) with respect to the next node at the top of said second connection address list;

(vii) if the node at the top of said second connection address list is part of said distribution network, determining whether said node at the top of said second connection address list has available capacity to transmit content data to said green node; and

(viii) if the node at the top of said second connection address list is part of said distribution network and has available capacity to transmit content data to said green node, having said green node dock with the node at the top of said second connection address list; and

(e) repeating steps (c) and (d) with respect to each user node which receives a retransmitted propagation signal.

In one example, step (c)(iii) may include the following additional step: (1) waiting a predetermined period of time so as to allow the red node's sibling with the highest

propagation rating prior to the red node's receipt of the propagation signal (hereinafter the "green node") to retransmit the propagation signal to the green node's child nodes (if any).

In another embodiment relating to a computer information distribution network comprising an instructing node having a fist user node docked with said instructing node to receive content data therefrom and a second user node docked with said first user node to receive content data therefrom, a process for handling a purported communication interruption between said first user node and said second user node is provided, said process including the following steps:

(a) having said instructing node receive a first complaint about said first user node from said second user node;

(b) if the second user node is the only user node docked with said first user node at a point in time, if said instructing node receives another complaint about the first user node from another user node within a predetermined period of time from when said instructing node received said first complaint or if said instructing node experiences a communication problem between it and said first user node during said predetermined period of time: (i) disconnecting said first user node from said instructing node; and

(c) if said instructing node has not experienced a communication problem between it and said first user node during said predetermined period of time and if a plurality of user nodes are docked with said first user node at said point in time and said instructing node has not received another complaint about said first user node from a user node other than said second user node within said predetermined period of time: (i) disconnecting said second user node from said first user node.

In one example, step (b) may include the following step: (ii) sending a signal to the user nodes docked with said first user node indicating the departure of said first user node from said network.

In another example, each user node may have no more than a predetermined maximum number of child nodes docked with it, and wherein step (b) may include the following steps

(ii) providing information at said point in time to each of said second user node, the other user nodes which had been docked with said first user node (if any) and at least one of any user nodes cascadingly connected in the distribution network under said first user node prior to said first user node's being disconnected from said instructing node (if any) as follows:

(1) the node's propagation rating, wherein a propagation rating is one of a predetermined number of grades ranging from highest to second highest to third highest and so on, with the number of grades being equal to said predetermined maximum number;

(2) an ancestor list setting forth the node's ancestor nodes' addresses (if it has any ancestor nodes) back to the instructing node, with the node's parent node being atop the ancestor list;

(3) a sibling list of the node's sibling nodes' addresses (if it has any sibling nodes) and their respective propagation ratings; and

(4) a child list of the node's child nodes' addresses (if it has any child nodes) and their respective propagation ratings;

(iii) having said instructing node send a propagation signal to each of said second user node and the other user nodes which had been docked with said first user node (if any);

(iv) with respect to each user node receiving the propagation signal which did not have the highest propagation rating prior to its receiving the propagation signal (hereinafter referred to as a "red node"), upon its receiving the propagation signal:

(1) setting the propagation rating of the red node to the next higher grade above the propagation rating of that red node before it received the propagation signal;

(2) having the red node undock from its parent node (if it was docked with its parent node before it received the propagation signal); and

(3) having the red node dock with its sibling node which had the highest propagation rating prior to the red node's receiving the propagation signal;

(v) with respect to each user node receiving the propagation signal which had the highest propagation rating prior to its receiving a propagation signal (hereinafter referred to as a "green node"), upon its receiving the propagation signal:

(1) setting the propagation rating of that green node to the lowest grade;

(2) having the green node retransmit the propagation signal to its child nodes (if it has any); and

(3) determining whether the green node is docked with the node from which the green node received the propagation signal and if it is, having the green node remain docked with said node from which the green node received the propagation signal, and if it is not, having the green node dock with said node from which the green node received the propagation signal; and

(vi) repeating steps (iv) and (v) with respect to each user node which receives a retransmitted propagation signal.

In another embodiment a distribution network for the distribution of content data from a server node to user nodes, wherein said user nodes are connected to said server and each other in cascaded relationship is provided, wherein:

(a) at least one of said user nodes is a repeater node connected directly to said server node, wherein said repeater node retransmits content data received by it to a user node docked to it for purpose of receiving content data from said repeater node (hereinafter referred to as a "child node"); and

(b) wherein each repeater node has the ability to provide to a user node which is attempting to dock with said repeater node connection address instructions.

In one example, each repeater node may include a descendant database indicating:

(i) which child nodes, if any, at a point in time, are docked with it so as to receive content data from said repeater node, and

(ii) which user nodes, if any, at said point in time, are purportedly docked with each of said child nodes.

In another example, the connection address instructions may refer said user node which is attempting to dock with said repeater node to a node in said descendant database.

In another example, each user node may include an ancestor database indicating to which node said user node is docked so that said user node may receive content data therefrom (hereinafter referred to as a "parent node"), and to which node, if any, at said point in time, said parent node is docked so that it may receive content data therefrom.

In another example, if said parent node of said user node departs from said distribution network, said user node may contact another node on its ancestor database.

In another example, each child node of a repeater node may include a sibling database indicating which user nodes, if any, are also child nodes of said repeater node.

In another embodiment a distribution network for the distribution of content data from a server node to user nodes is provided, wherein n levels of user nodes are cascadingly connected to said server node, wherein n is a number greater than one, wherein each user node includes a delay which causes the playing of content data by such node to be delayed a period of time (hereinafter "delay time") from a point in time, and wherein delays in higher level nodes create greater delay times than do delays in lower level nodes.

In another embodiment a distribution network for the distribution of content data from a server node to user nodes is provided, wherein n levels of user nodes are cascadingly

connected to said server node, wherein n is a number greater than one, wherein each user node includes a delay which causes the playing of content data by such node to be delayed a period of time (hereinafter "delay time") from a point in time, wherein x is a number from and including 2 to and including n, and wherein delays in (x-1) level nodes create greater delay times than do delays in x level nodes.

In one example, said point in time is an event experienced approximately simultaneously by substantially all user nodes.

In another example, said point in time is when said node receives content data.

While a number of embodiments and examples of the present invention have been described, it is understood that these embodiments and examples are illustrative only, and not restrictive, and that many modifications may become apparent to those of ordinary skill in the art. For example, any desired branch factor may be used (*see*, e.g., FIG. 46A showing a branch factor of 1 (i.e., each parent node may have only one child node); FIG. 46B showing a branch factor of 2 (i.e., each parent node may have one or two children nodes); and FIG. 46C showing a branch factor of 3 (i.e., each parent node may have one, two or three children nodes)). Of course, other branch factors (e.g., 4, 5, 6, etc) may be used and, as discussed above, the distribution network may be a hybrid network utilizing subtrees having different branch factors). Further still, a modified Utility Rating may be determined as follows: Utility Rating = Time (of connection) x Bandwidth Rating. Further still, a node ID may be based upon: (a) an IP address (external/internal/DHCP); (b) a URL; and/or (c) a Port. Further still, the root server may propagate "its" time upon assignment of a connection and the user node may get a "delta" (i.e., difference) from its clock to use (in the event of a reset the user node may re-acquire the time from the server (or from another time-providing source)). Further still, the various time intervals discussed herein may be predetermined time intervals (e.g., .5 sec., 1 sec, 1 min.) and/or the time intervals may be based upon some varying criteria (e.g., flow of data in the network, number of nodes in the network, number of levels in the network, position of a given node in the network). Further still, partial propagation of data (e.g., ancestor data and/or descendent data) from one or more nodes may be permitted. Further still (as discussed above), one or more of the steps described herein may be omitted (and the various steps which are performed may not necessarily need to be carried out in the order described herein (which description was intended to represent a number of examples, and not be restrictive)). Further still, the root server and each of the user nodes may utilize essentially identical software for carrying out the network configuration, reconfiguration and data transfer operations described herein or the root server may have dedicated server software

while the user nodes have different, dedicated user node software. Further still, the entire distribution network (e.g., the root server and all of the user nodes) may have the same branch factor (e.g., a branch factor of 2) or the user nodes may have a first branch factor (e.g., a branch factor of 2) and the root server may have a second branch factor (e.g., the root server may be capable of acting as a parent to more user nodes than 2). Further still, heuristics may be employed to drive network configuration, reconfiguration and/or data flow (e.g., based upon historic network configuration, reconfiguration and/or data flow).

Ex. 1018-1769

CLAIMS

1. A method for propagating a network topology model of a binary tree topology network through the network, the network including a plurality of nodes, including a root node and at least one leaf node, operatively networked together to provide a communications path for data, each of the plurality of nodes except the root node being docked downtree of a parent node, the parent node being one of the plurality of nodes, and each of the plurality of nodes adapted to be docked uptree of a first child node and a second child node, each of the at least one leaf nodes having none of the plurality of nodes docked downtree of and therewith, the method comprising the steps of:

each of the at least one leaf nodes providing a propagate network topology model message containing the address of the providing leaf node to the parent node docked uptree of and with the providing leaf node;

for each of the plurality of nodes receiving the propagate network topology model message from each child node docked downtree of and therewith:

forming a network topology model representing one or more paths from the receiving node to each one of the plurality of nodes downtree of the receiving node;

storing in a memory accessible to the receiving node, the network topology model so formed;

providing to the parent node docked uptree of and with the receiving node a propagate network topology model message containing the network topology model so formed,

thereby propagating a network topology model through the network.

2. The method of claim 1, wherein the communications path for data comprises a downtree communications path.

3. The method of claim 1, wherein the communications path for data comprises a downtree communications path for data and an uptree communications path for data.

Ex. 1018-1770

4. The method of claim 1, wherein the step of each of the at least one leaf nodes providing a propagate network topology model message is carried out a plurality of times.

5. The method of claim 1, wherein the step of each of the at least one leaf nodes providing a propagate network topology model message is carried out periodically.

6. The method of claim 1, wherein a first node being one of the plurality of nodes, but not a leaf node, receives a register message from a second node that is not part of the network, and wherein the first node is docked uptree of a first child node and a second child node, each of the first child node and the second child node being one of the plurality of nodes, the method further comprising the steps of:

searching the network topology model stored in the memory of the first node for a path to an open node, the open node being one of the plurality of nodes that the network topology model shows as not being docked uptree of and with two child nodes;

sending a path message to the second node containing at least an address of the open node.

7. The method of claim 6, wherein the path message comprises a path from the open node to the first node.

8. The method of claim 6, wherein the searching step comprising finding the open node such that a connection to the open node keeps the binary topology network balanced.

9. The method of claim 6, wherein step of searching further comprises the steps of:

searching consecutively further down hierarchical levels of the binary tree represented in the network topology model stored in the memory of the first node:

first to locate a path representing a node not being docked uptree of and with any child nodes, and if no such path is located,

then to locate a path representing a node not being docked uptree of and with two child nodes, and

repeating the searching step until one such path is located; and

designating the located path as a path to the open node.

10. A method for propagating a network topology model of a binary tree topology network through the network, the network including a plurality of nodes, including a root node and at least one leaf node, operatively networked together to provide a communications path for data, each of the plurality of nodes except the root node being docked downtree of a parent node, the parent node being one of the plurality of nodes, and each of the plurality of nodes adapted to be docked uptree of a first child node and a second child node, each of the at least one leaf nodes having none of the plurality of nodes docked downtree of and therewith, the method comprising the steps of:

providing a propagate path message from the root node to each of the plurality of nodes that are docked downtree of and with the root node, the propagate path message indicating an address of the root node;

for each of the plurality of nodes receiving the propagate path message from the root node or from another one of the plurality of nodes:

storing in a memory accessible to the receiving node, the path to the root node stored in the propagate path message;

if the receiving node has at least one of the plurality of nodes docked downtree of and therewith, adding to the propagate path message an address of the receiving node; and

providing the propagate path message to each of the plurality of nodes docked downtree of and with the receiving node, the propagate path message a path to the root node;

thereby propagating a root path of a binary tree topology network through the network.

11. The method of claim 10, wherein the communications path for data comprises a downtree communications path for data.

12. The method of claim 10, wherein the communications path for data comprises a downtree communications path for data and an uptree communications path for data.

13. The method of claim 10, wherein the step of providing a propagate path message is carried out a plurality of times.

14. The method of claim 10, wherein the step of providing a propagate path message is carried out periodically.

15. A method for propagating a network topology model of a binary tree topology network through the network, the network including a plurality of nodes, including a root node and at least one leaf node, operatively networked together to provide a communications path for data, each of the plurality of nodes except the root node being docked downtree of a parent node, the parent node being one of the plurality of nodes, and each of the plurality of nodes adapted to be docked uptree of a first child node and a second child node, each of the at least one leaf nodes having none of the plurality of nodes docked downtree of and therewith, the method comprising the steps of:

each of the at least one leaf nodes, providing a propagate network topology model message containing the address of the providing leaf node to the parent node docked uptree of and with the providing leaf node;

for each of the plurality of nodes receiving the propagate network topology model message from each child node docked downtree of and therewith:

forming a network topology model comprising a path from the receiving node to each one of the plurality of nodes downtree of the receiving node;

storing in a memory accessible to the receiving node, the network topology model so formed;

providing to the parent node docked uptree of and with the receiving node a propagate network topology model message containing the network topology model so formed,

providing a propagate path message from the root node to each of the plurality of nodes that are docked downtree of and with the root node, the propagate path message indicating an address of the root node;

for each of the plurality of nodes receiving propagate path message from the root node or from another one of the plurality of nodes:

storing in a memory accessible to the receiving node, the path to the root node stored in the propagate path message;

if the receiving node has at least one of the plurality of nodes docked downtree of and therewith, adding to the propagate path message an address of the receiving node; and

providing the propagate path message to each of the plurality of nodes docked downtree of and with the root node.

16. The method of claim 15, wherein the communications path for data comprises a downtree communications path for data.

17. The method of claim 15, wherein the communications path for data comprises a downtree communications path for data and an uptree communications path for data.

18. The method of claim 15, wherein the step of providing a propagate network topology model message is carried out a plurality of times.

19. The method of claim 15, wherein the step of providing a propagate network topology model message is carried out periodically.

20. The method of claim 15, wherein the step of providing a propagate path message is carried out a plurality of times.

21. The method of claim 15, wherein the step of providing a propagate path message is carried out periodically.

22. The method of claim 15, wherein a first node being one of the plurality of nodes, but not a leaf node, receives a register message from a second node that is not part of the network, and wherein the first node is docked uptree of a first child node and a second child node, each of the first child node and the second child node being one of the plurality of nodes, the method further comprising the steps of:

searching the network topology model stored in the memory of the first node for a path to an open node, the open node being one of the plurality of nodes that the network topology model shows as not being docked uptree of and with two child nodes;

sending a path message to the second node containing at least an address of the open node.

23. The method of claim 22, wherein the path message comprises a path from the open node to the first node.

24. The method of claim 22, wherein the path message comprises a path from the open node to the root node.

25. The method of claim 22, wherein the searching step comprising finding the open node such that a connection to the open node keeps the binary topology network balanced.

26. The method of claim 22, wherein step of searching further comprises the steps of:

searching consecutively further down hierarchical levels of the binary tree represented in the network topology model stored in the memory of the first node:

first to locate a path representing a node not being docked uptree of and with any child nodes, and if no such path is located,

then to locate a path representing a node not being docked uptree of and with two child nodes, and

repeating the searching step until one such path is located; and

designating the located path as a path to the open node.

27. A method for propagating a network topology model of a binary tree topology network through the network, the network including a plurality of nodes, including a root node and at least one leaf node, operatively networked together to provide a communications path for data, each of the plurality of nodes except the root node being docked downtree of a parent node, the parent node being one of the plurality of nodes, and each of the plurality of nodes adapted to be docked uptree of a first child node and a second child node, each of the at least one leaf nodes having none of the plurality of nodes docked downtree of and therewith, the method comprising the steps of:

providing a banned list message from the root node to each of the plurality of nodes that are docked downtree of and with the root node, the banned list message comprising an identifier for each of a plurality of banned nodes;

for each of the plurality of nodes receiving the banned list message from the root node or from another one of the plurality of nodes:

storing in a memory accessible to the receiving node, the identifiers for each of the plurality of banned nodes; and

providing the banned list message to each of the plurality of nodes docked downtree of and with the receiving node;

thereby propagating the list of banned nodes through the network.

28. The method of claim 27, wherein the communications path for data comprises a downtree communications path for data.

29. The method of claim 27, wherein the communications path for data comprises a downtree communications path for data and an uptree communications path for data.

30. The method of claim27, further comprising the step of rejecting a register message from a new node that is not part of the network if the node is identified on the banned list.

31. A method for propagating a network topology model of a binary tree topology network through the network, the network including a plurality of nodes, including a root node and at least one leaf node, operatively networked together to provide a communications path for data, each of the plurality of nodes except the root node being docked downtree of a parent node, the parent node being one of the plurality of nodes, and each of the plurality of nodes adapted to be docked uptree of a first child node and a second child node, each of the at least one leaf nodes having none of the plurality of nodes docked downtree of and therewith, the method comprising the steps of:

the first node sending a request firewall status message over a first channel to one of the plurality of nodes that is docked with the first node;

the one of the plurality of nodes receiving the request firewall status message sending at least one response message to the first node over a separate channel; and

the first node determining its firewall status based upon whether it receives the response message.

32. The method of claim 31, wherein the communications path for data comprises a downtree communications path for data.

33. The method of claim 31, wherein the communications path for data comprises a downtree communications path for data and an uptree communications path for data.

34. The method of claim 31, wherein the step of sending at least one response message comprises the steps of:

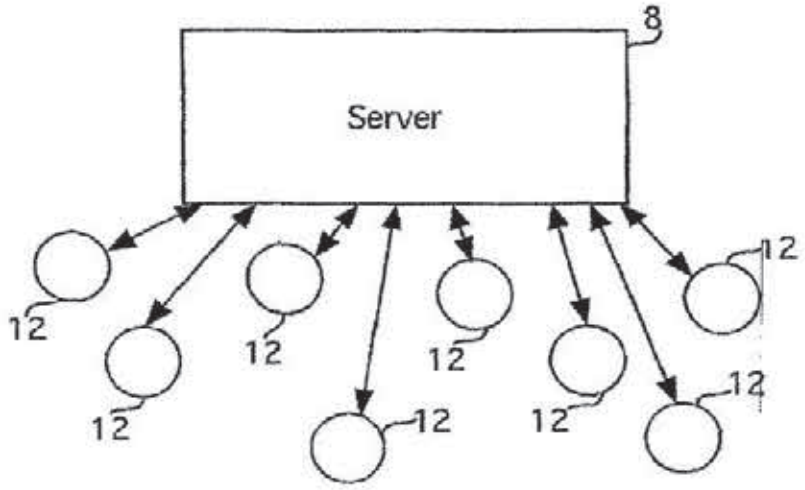
sending a first response message to the first node over a separate channel;

waiting a predefined period for an acknowledgement of the first response message, and if no acknowledgement is received, sending a second response message to the first node over the separate channel;

waiting a predefined period for an acknowledgement of the second response message, and if no acknowledgement is received, sending a third response message to the first node over the separate channel; and

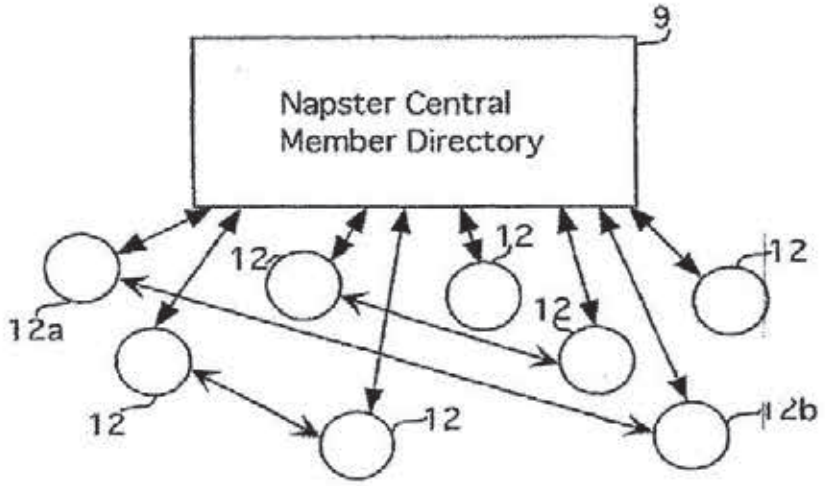
waiting a predefined period for an acknowledgement of the third response message, and if no acknowledgement is received, sending a firewall status on message to the first node over the first channel.

Ex. 1018-1778



PRIOR ART

Fig. 1



PRIOR ART

Fig. 2

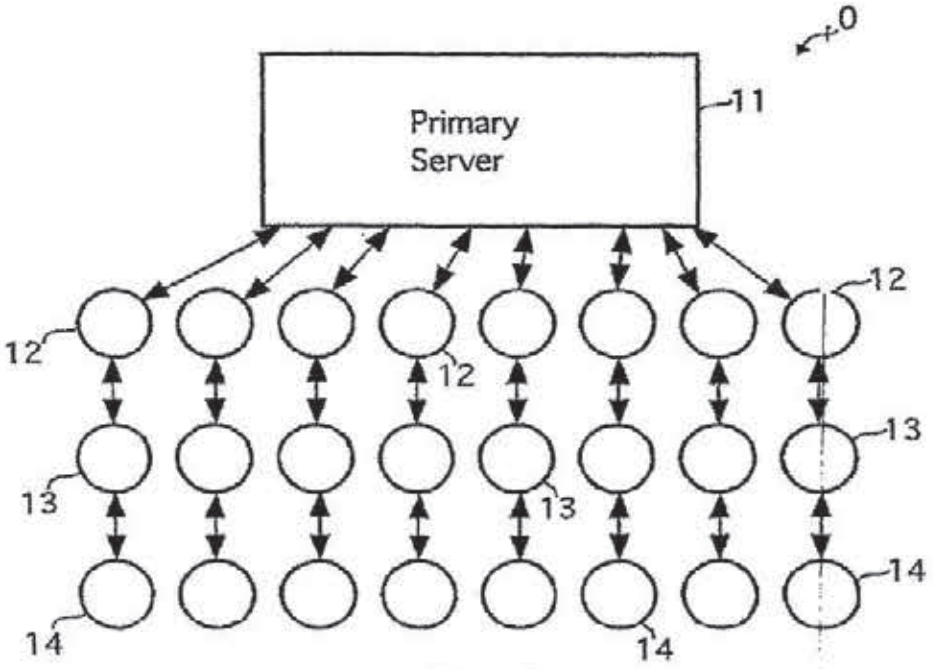


Fig. 3

"Start" 31

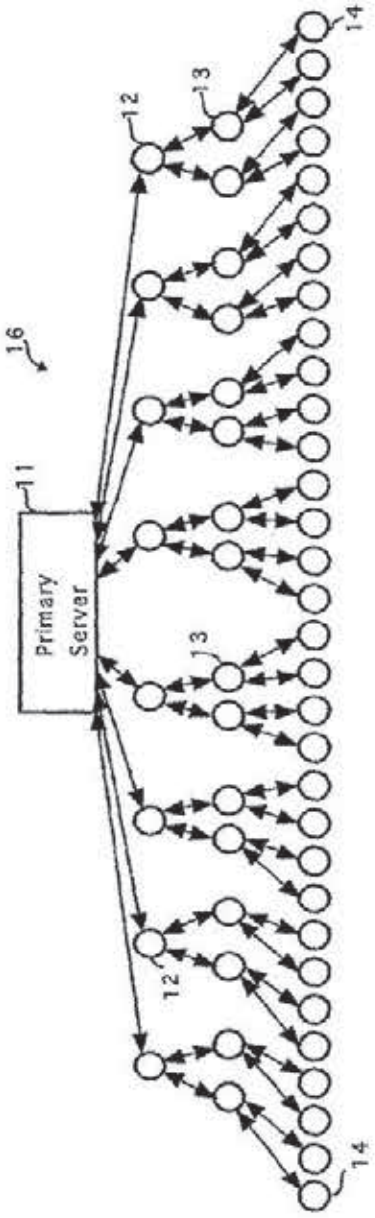

16
11
Primary Server
12
13
14
12'
13
14

Fig. 4

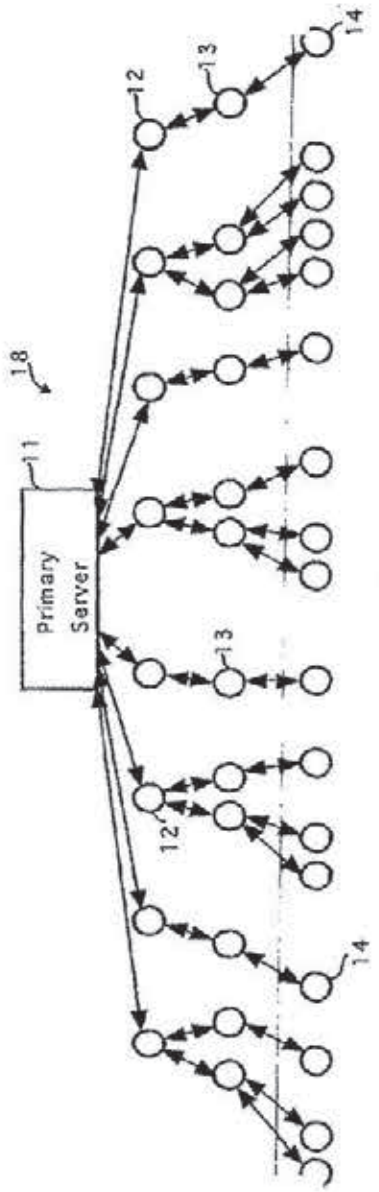

18
11
Primary Server
12
13
14
12'
13
14

Fig. 5

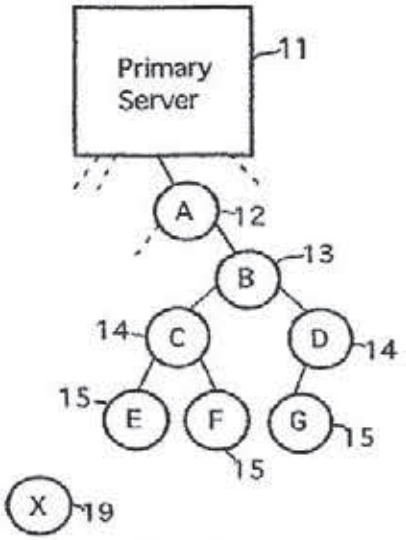


Fig. 6

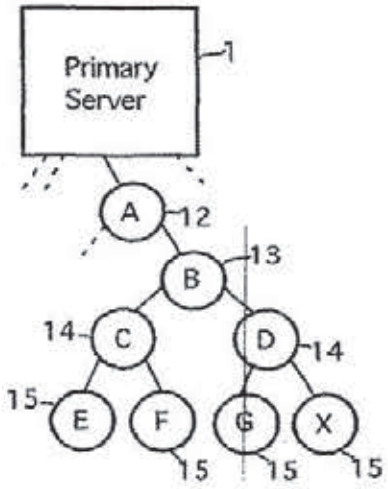


Fig. 7

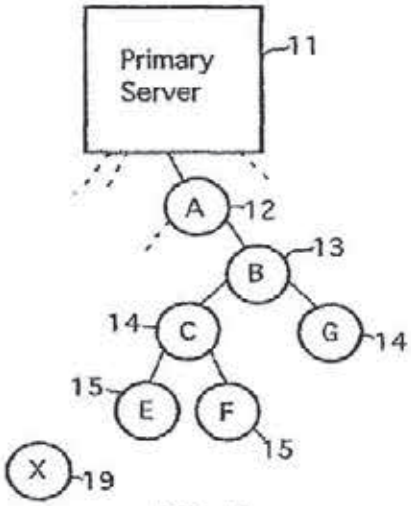


Fig. 8

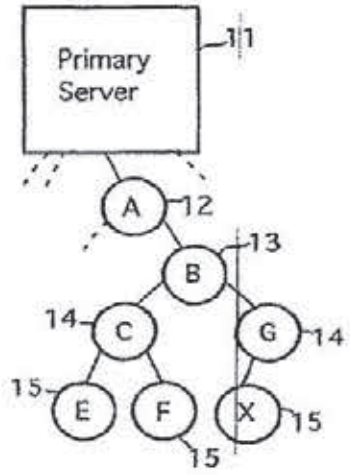


Fig. 9

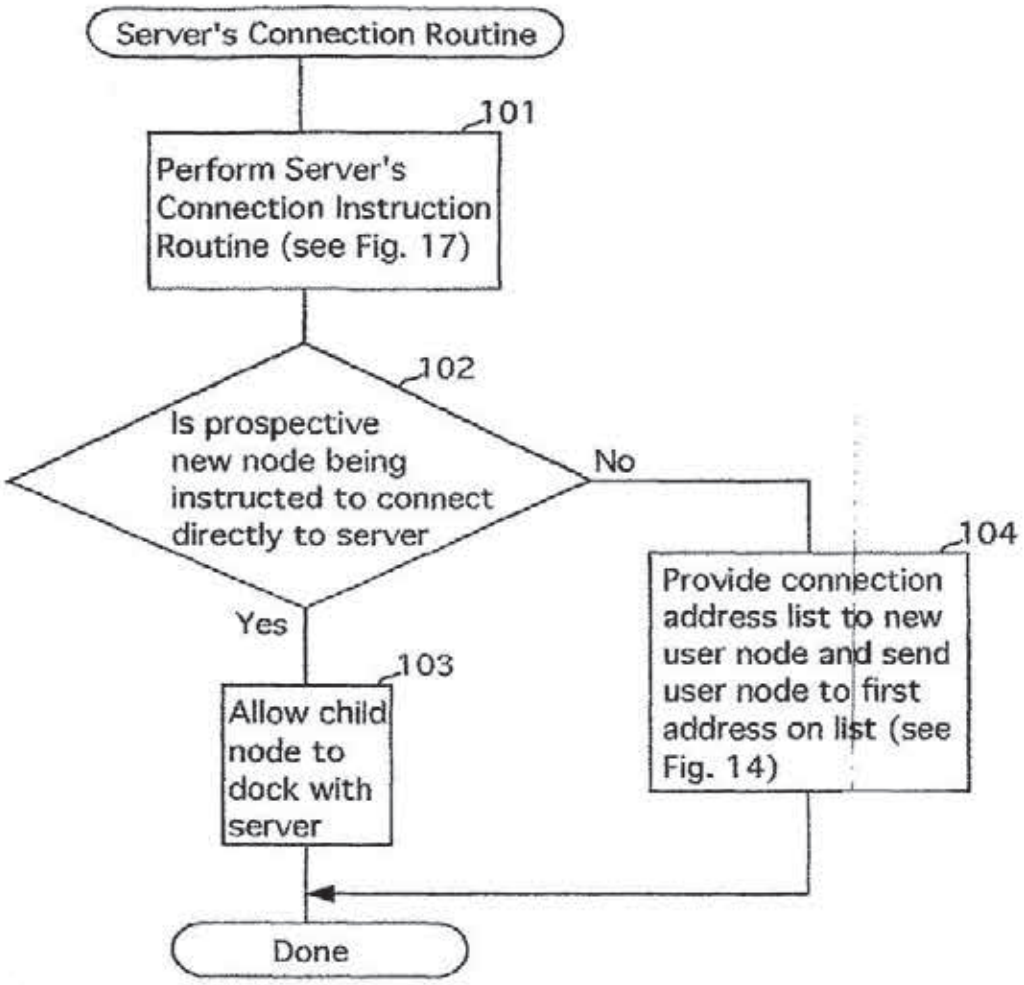

Server's Connection Routine
101
Perform Server's Connection Instruction Routine (see Fig. 17)
102
Is prospective new node being instructed to connect directly to server
No
Yes
103
Allow child node to dock with server
104
Provide connection address list to new user node and send user node to first address on list (see Fig. 14)
Done

Fig. 10

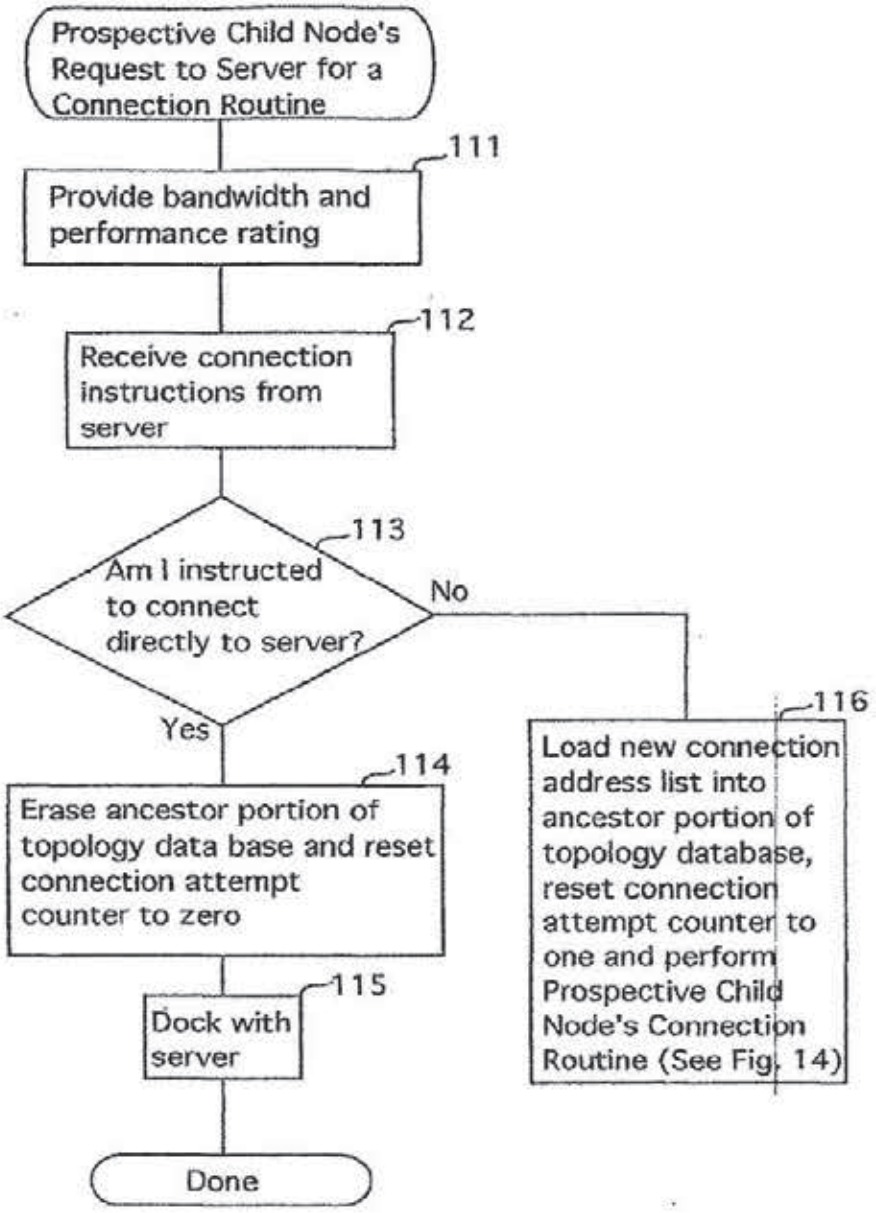

Prospective Child Node's Request to Server for a Connection Routine
111
Provide bandwidth and performance rating
112
Receive connection instructions from server
113
Am I instructed to connect directly to server?
No
Yes
116
Load new connection address list into ancestor portion of topology database, reset connection attempt counter to one and perform Prospective Child Node's Connection Routine (See Fig. 14)
114
Erase ancestor portion of topology data base and reset connection attempt counter to zero
115
Dock with server
Done

Fig. 11

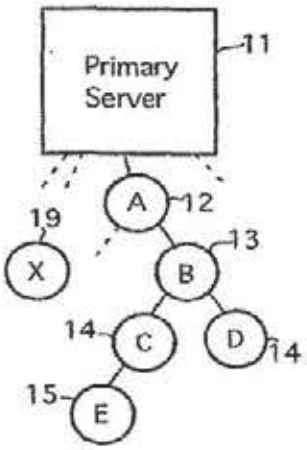


Fig. 12A

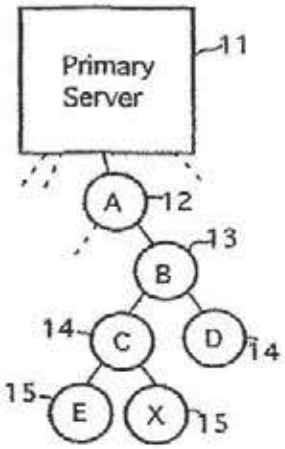


Fig. 12B

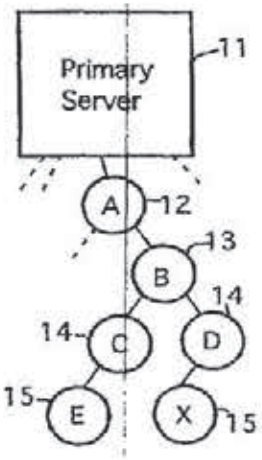


Fig. 12C

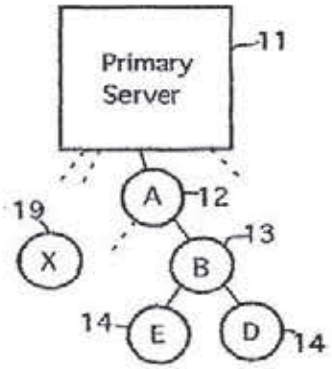


Fig. 12D

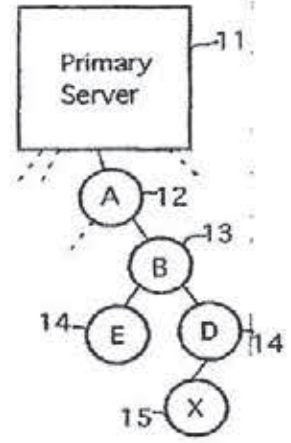


Fig. 12E

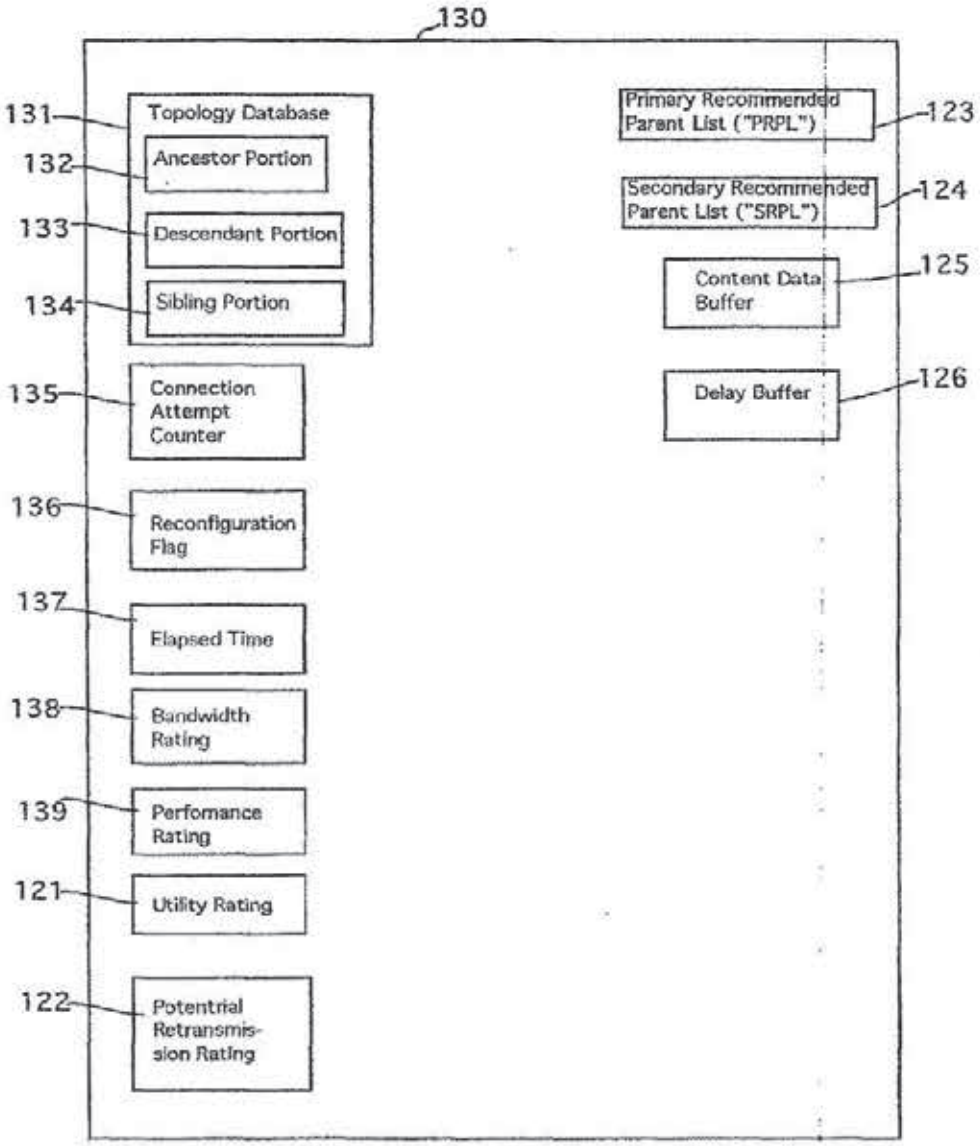

130
131
Topology Database
132
Ancestor Portion
133
Descendant Portion
134
Sibling Portion
135
Connection Attempt Counter
136
Reconfiguration Flag
137
Elapsed Time
138
Bandwidth Rating
139
Performance Rating
121
Utility Rating
122
Potentrial Retransmis-sion Rating
Primary Recommended Parent List ("PRPL")
123
Secondary Recommended Parent List ("SRPL")
124
Content Data Buffer
125
Delay Buffer
126

Fig. 13

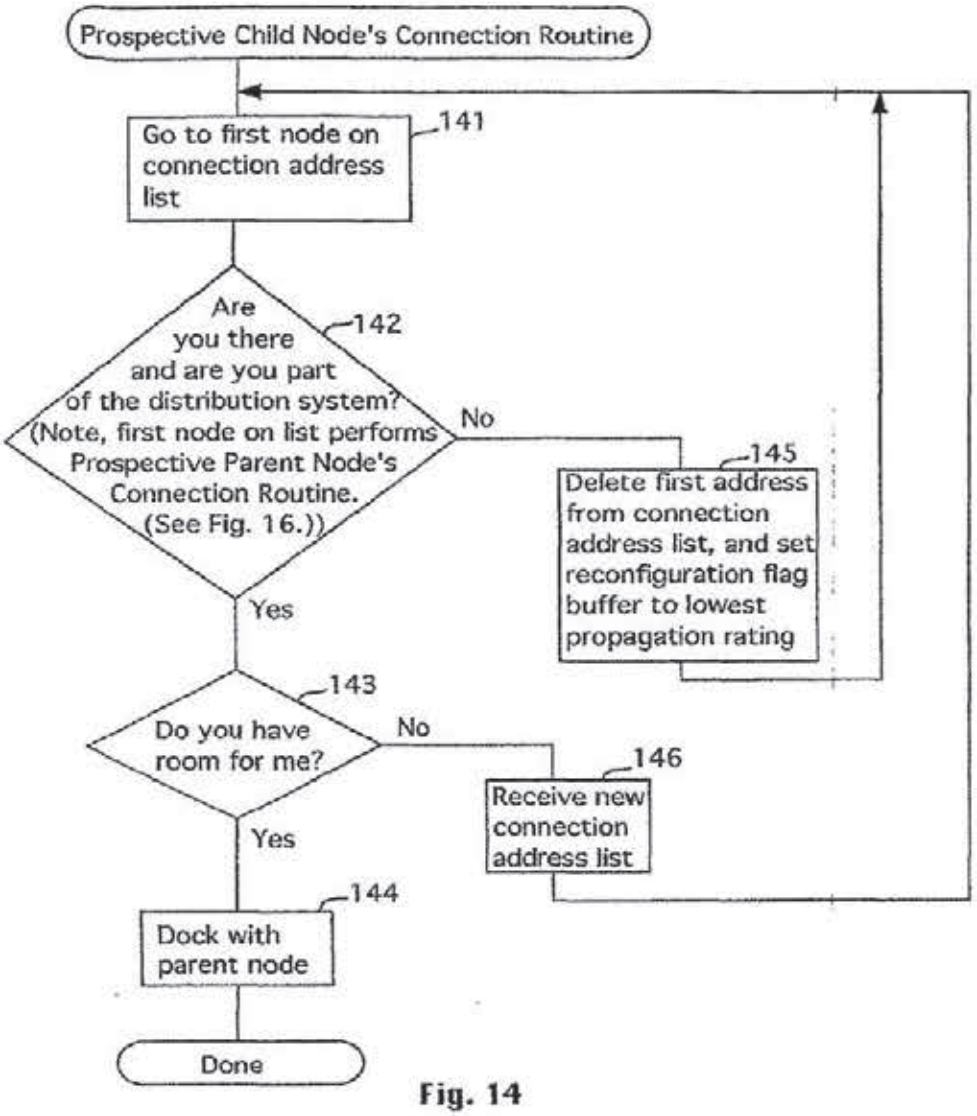

Prospective Child Node's Connection Routine
Go to first node on connection address list
141
Are you there and are you part of the distribution system? (Note, first node on list performs Prospective Parent Node's Connection Routine. (See Fig. 16.))
142
No
Delete first address from connection address list, and set reconfiguration flag buffer to lowest propagation rating
145
Yes
Do you have room for me?
143
No
Receive new connection address list
146
Yes
Dock with parent node
144
Done

Fig. 14

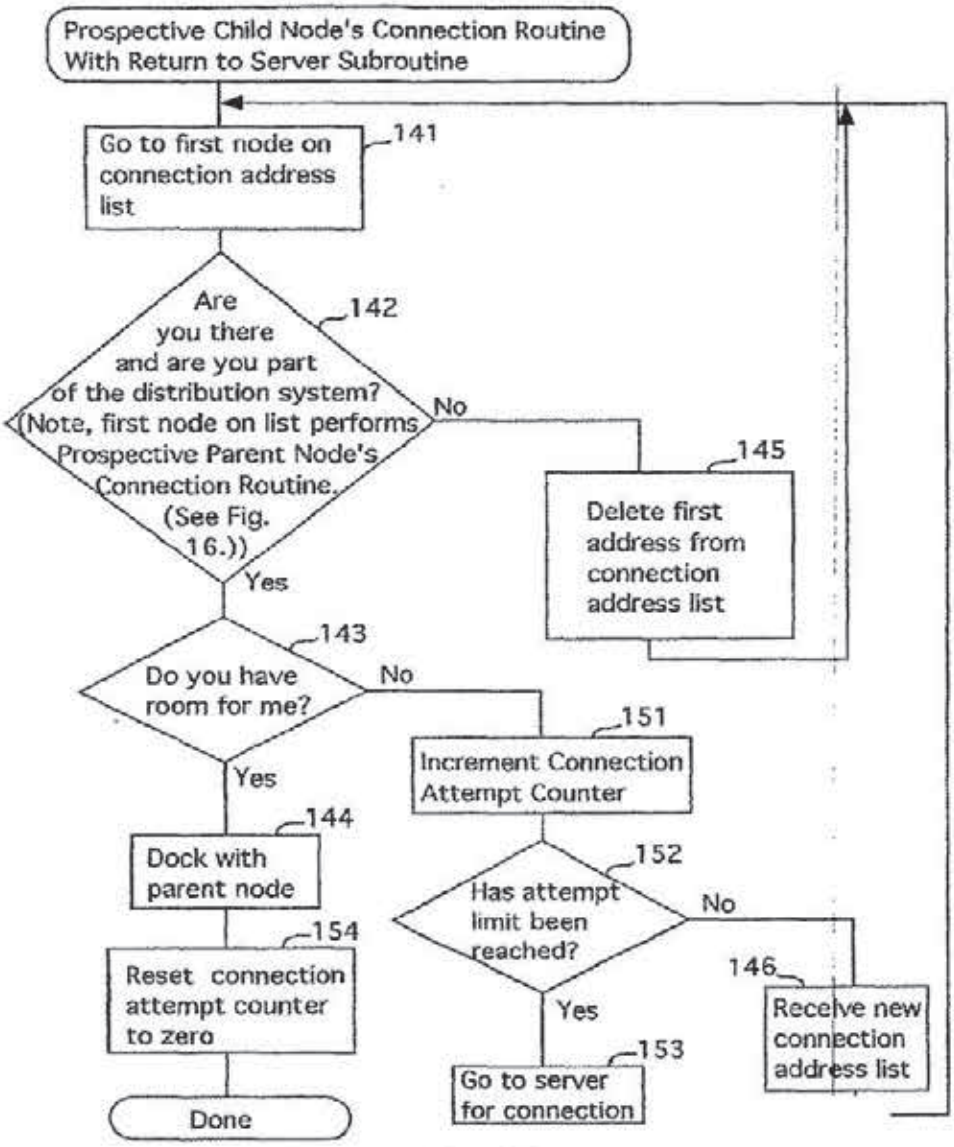

Prospective Child Node's Connection Routine With Return to Server Subroutine
Go to first node on connection address list
141
Are you there and are you part of the distribution system? (Note, first node on list performs Prospective Parent Node's Connection Routine. (See Fig. 16.))
142
No
Yes
Delete first address from connection address list
145
Do you have room for me?
143
No
Yes
Increment Connection Attempt Counter
151
Has attempt limit been reached?
152
No
Yes
Receive new connection address list
146
Go to server for connection
153
Dock with parent node
144
Reset connection attempt counter to zero
154
Done

Fig. 15

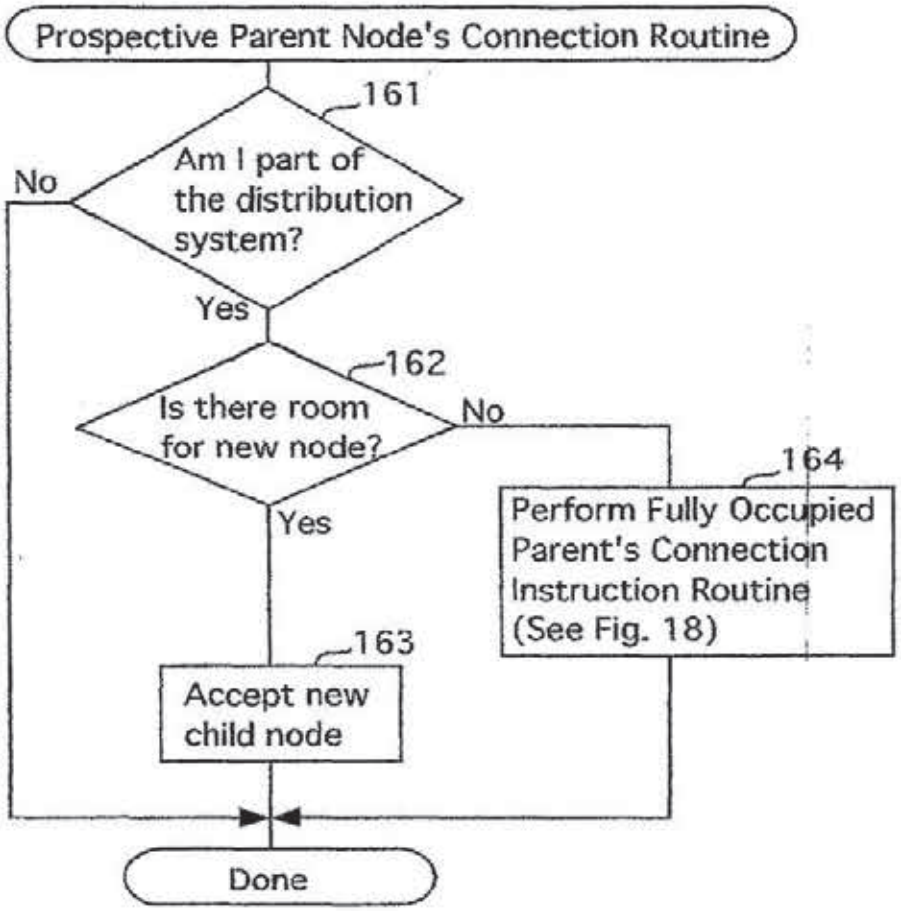

Prospective Parent Node's Connection Routine
161
Am I part of the distribution system?
No
Yes
162
Is there room for new node?
No
Yes
164
Perform Fully Occupied Parent's Connection Instruction Routine (See Fig. 18)
163
Accept new child node
Done

Fig. 16

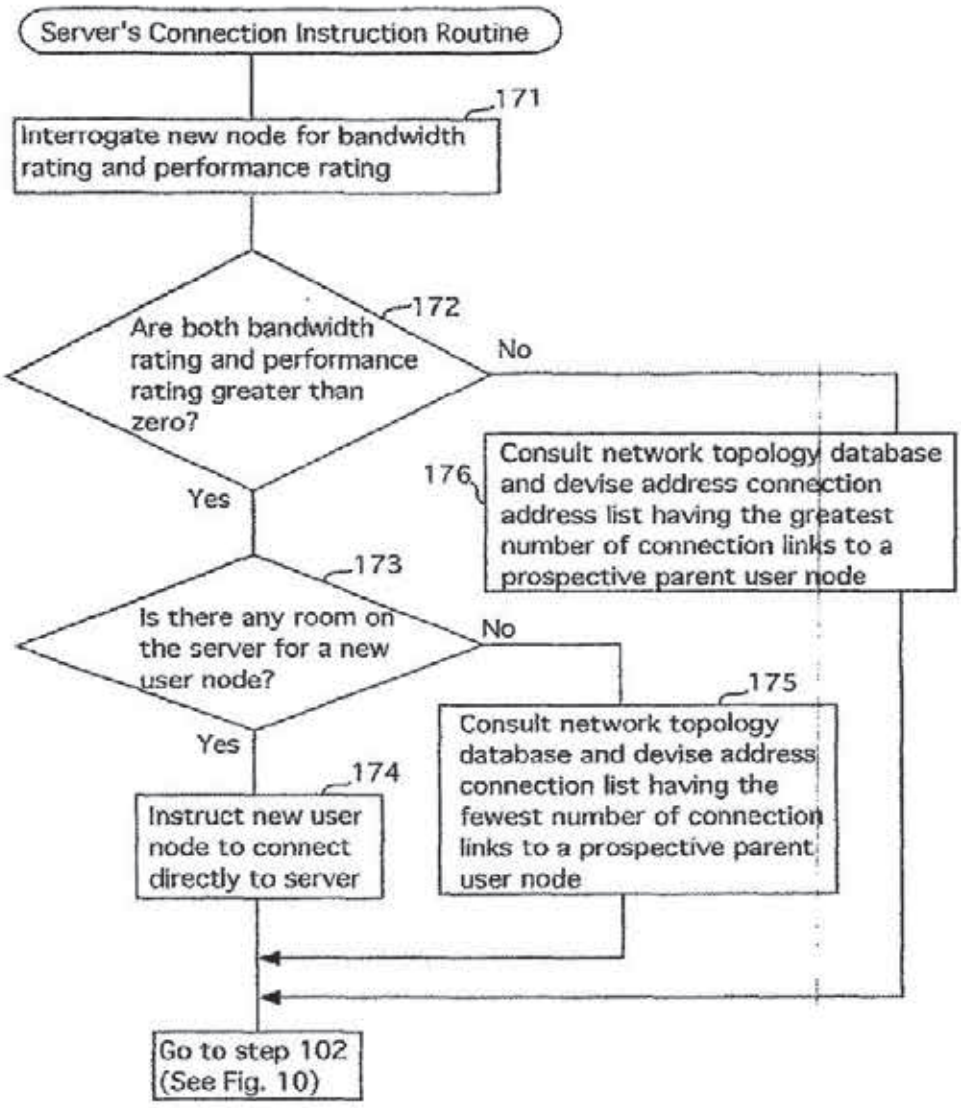

Server's Connection Instruction Routine
171
Interrogate new node for bandwidth rating and performance rating
172
Are both bandwidth rating and performance rating greater than zero?
No
Yes
176
Consult network topology database and devise address connection address list having the greatest number of connection links to a prospective parent user node
173
Is there any room on the server for a new user node?
No
Yes
175
Consult network topology database and devise address connection list having the fewest number of connection links to a prospective parent user node
174
Instruct new user node to connect directly to server
Go to step 102 (See Fig. 10)

Fig. 17

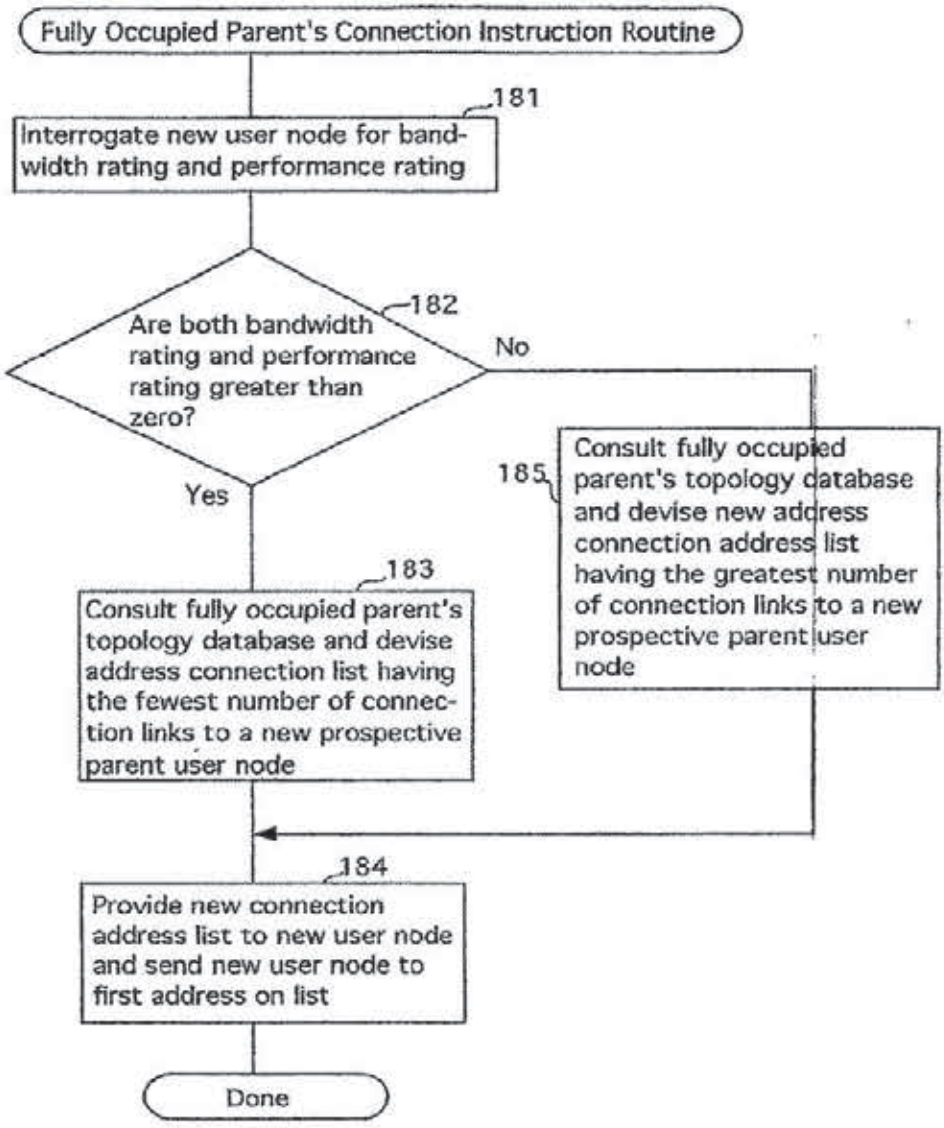

Fully Occupied Parent's Connection Instruction Routine
181
Interrogate new user node for bandwidth rating and performance rating
182
Are both bandwidth rating and performance rating greater than zero?
No
Yes
185
Consult fully occupied parent's topology database and devise new address connection address list having the greatest number of connection links to a new prospective parent user node
183
Consult fully occupied parent's topology database and devise address connection list having the fewest number of connection links to a new prospective parent user node
184
Provide new connection address list to new user node and send new user node to first address on list
Done

Fig. 18

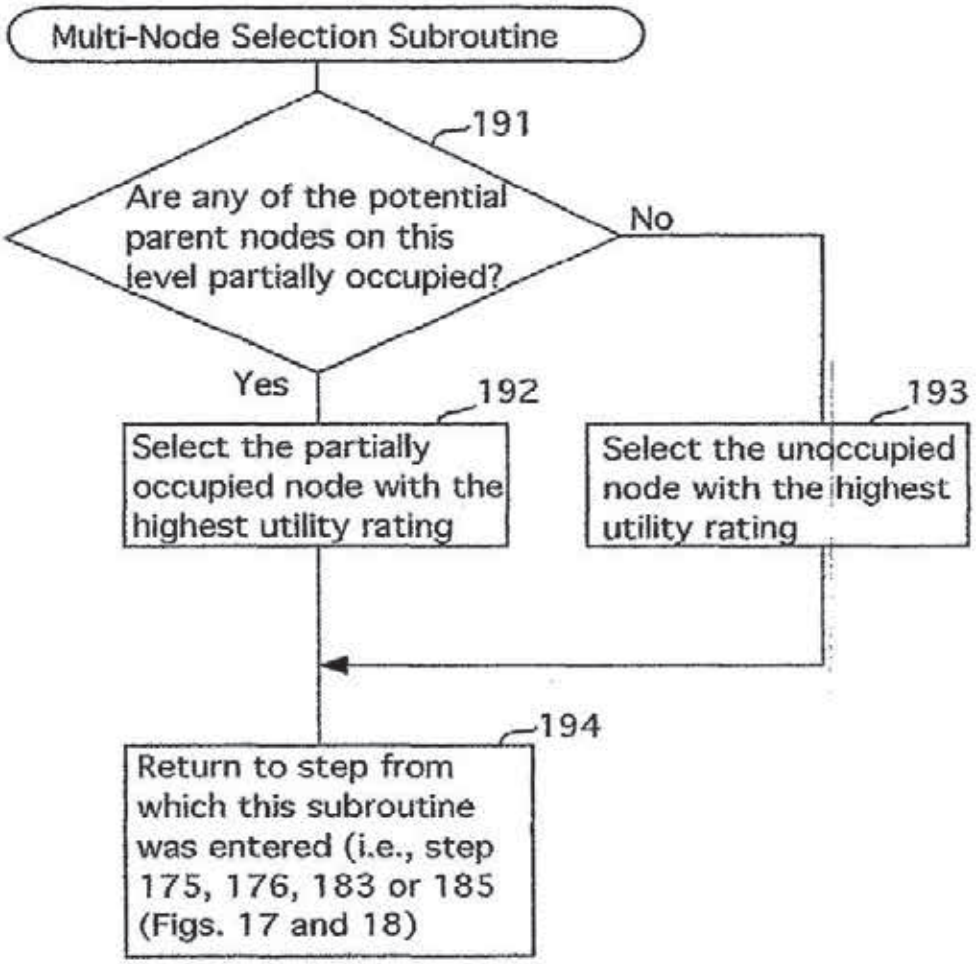

Multi-Node Selection Subroutine
191
Are any of the potential parent nodes on this level partially occupied?
No
Yes
192
Select the partially occupied node with the highest utility rating
193
Select the unoccupied node with the highest utility rating
194
Return to step from which this subroutine was entered (i.e., step 175, 176, 183 or 185 (Figs. 17 and 18)

Fig. 19

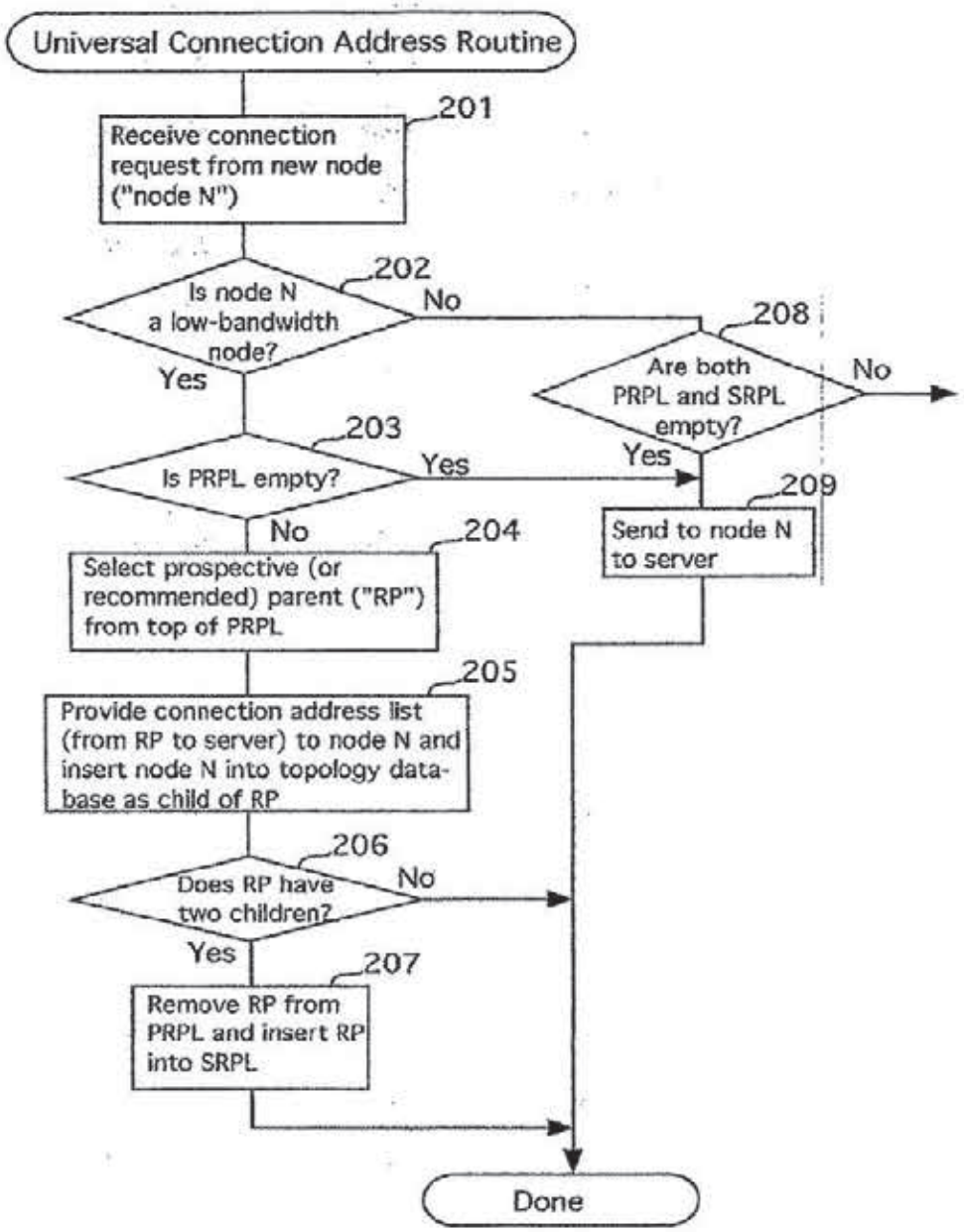

Universal Connection Address Routine
201
Receive connection request from new node ("node N")
202
Is node N a low-bandwidth node?
No
Yes
208
Are both PRPL and SRPL empty?
No
Yes
203
Is PRPL empty?
Yes
209
Send to node N to server
No
204
Select prospective (or recommended) parent ("RP") from top of PRPL
205
Provide connection address list (from RP to server) to node N and insert node N into topology data-base as child of RP
206
Does RP have two children?
No
Yes
207
Remove RP from PRPL and insert RP into SRPL
Done

Fig. 20A

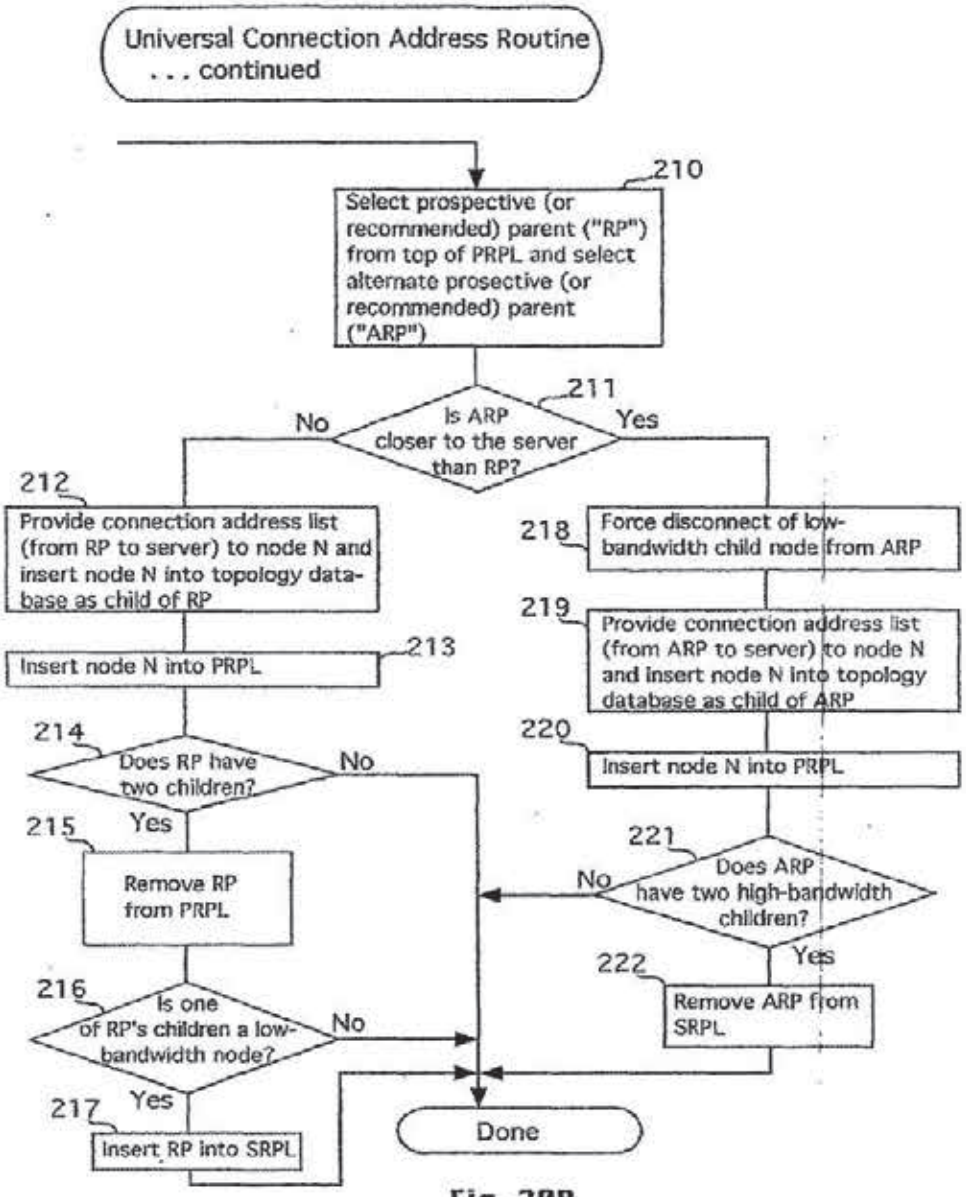

Universal Connection Address Routine
. . . continued
210
Select prospective (or recommended) parent ("RP") from top of PRPL and select alternate prosective (or recommended) parent ("ARP")
211
Is ARP closer to the server than RP?
No
Yes
212
Provide connection address list (from RP to server) to node N and insert node N into topology database as child of RP
213
Insert node N into PRPL
214
Does RP have two children?
No
Yes
215
Remove RP from PRPL
216
Is one of RP's children a low-bandwidth node?
No
Yes
217
Insert RP into SRPL
218
Force disconnect of low-bandwidth child node from ARP
219
Provide connection address list (from ARP to server) to node N and insert node N into topology database as child of ARP
220
Insert node N into PRPL
221
Does ARP have two high-bandwidth children?
No
Yes
222
Remove ARP from SRPL
Done

Fig. 20B

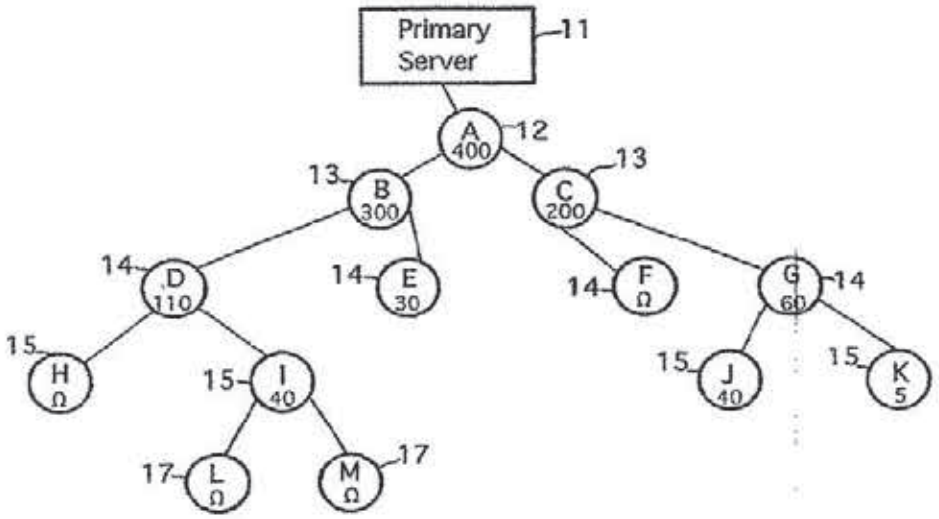

Primary Server
11
A 400
12
B 300
13
C 200
13
D 110
14
E 30
14
F Ω
14
G 60
14
H Ω
15
I 40
15
J 40
15
K 5
15
L Ω
17
M Ω
17

Fig. 21

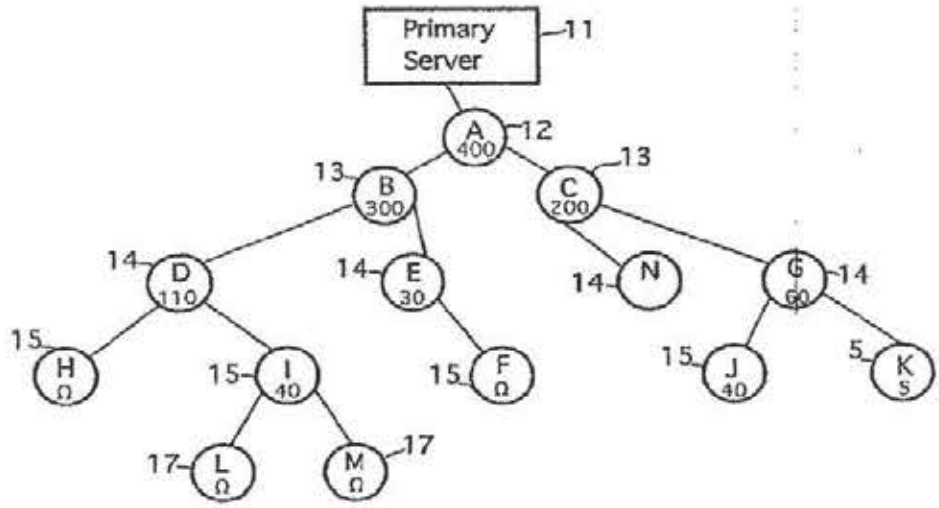

Primary Server
11
A 400
12
B 300
13
C 200
13
D 110
14
E 30
14
N
14
G 60
14
H Ω
15
I 40
15
F Ω
15
J 40
15
K 5
5
L Ω
17
M Ω
17

Fig. 22

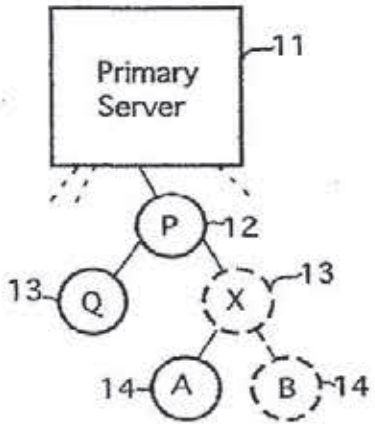


Fig. 23

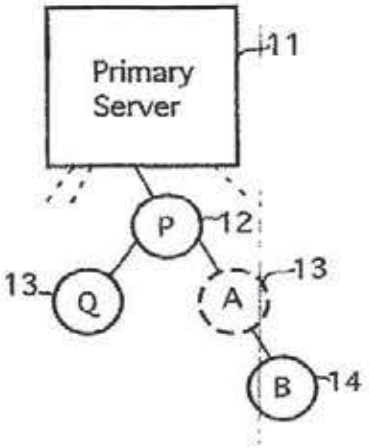


Fig. 24

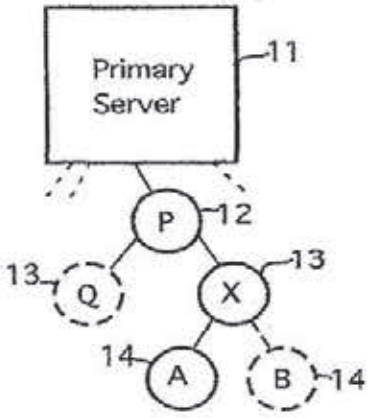


Fig. 25

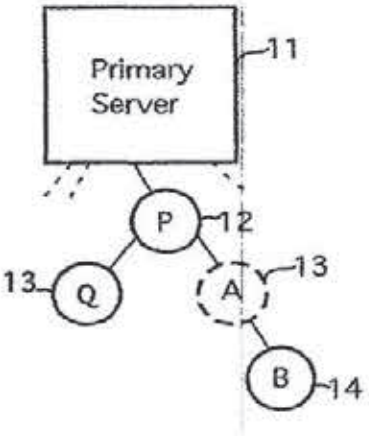


Fig. 26

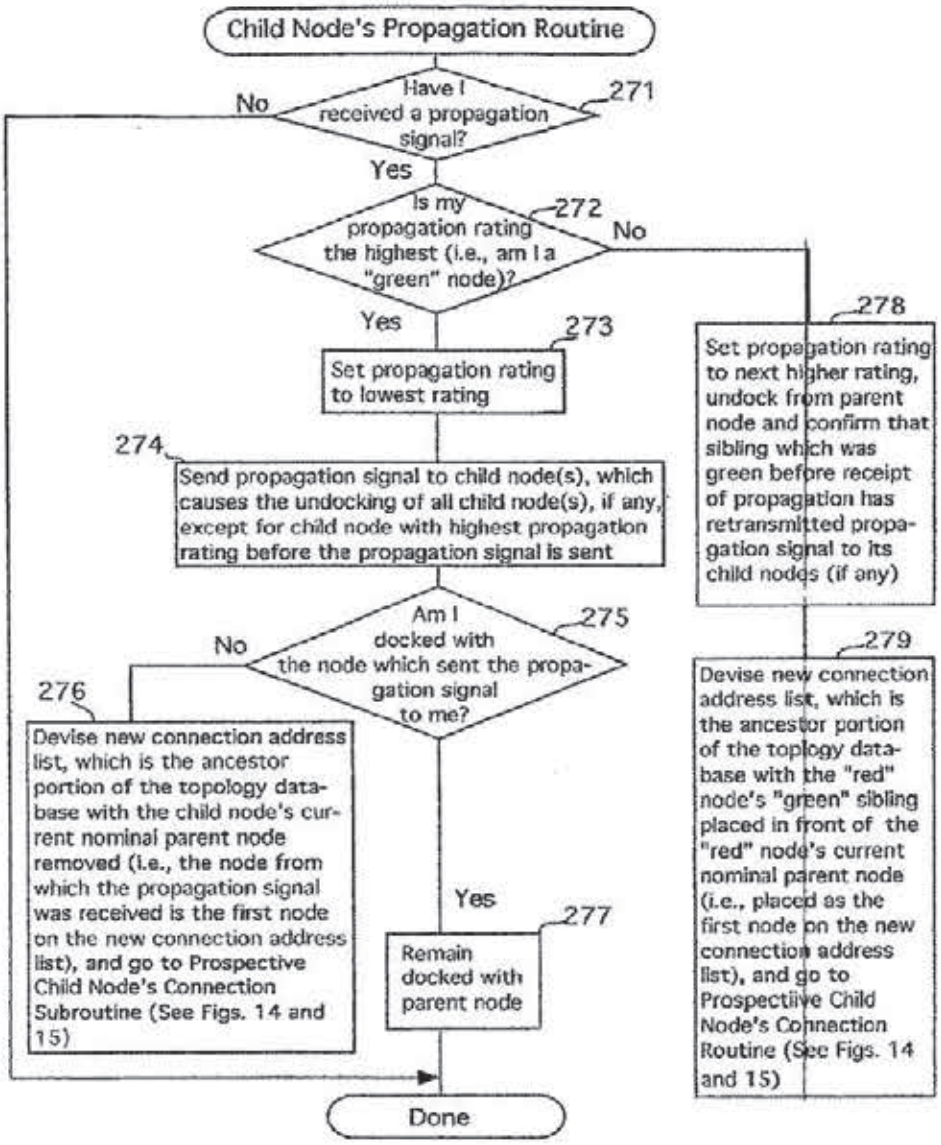

Child Node's Propagation Routine
Have I received a propagation signal?
271
No
Yes
Is my propagation rating the highest (i.e., am I a "green" node)?
272
No
Yes
Set propagation rating to lowest rating
273
Send propagation signal to child node(s), which causes the undocking of all child node(s), if any, except for child node with highest propagation rating before the propagation signal is sent
274
Am I docked with the node which sent the propagation signal to me?
275
No
Yes
Devise new connection address list, which is the ancestor portion of the topology database with the child node's current nominal parent node removed (i.e., the node from which the propagation signal was received is the first node on the new connection address list), and go to Prospective Child Node's Connection Subroutine (See Figs. 14 and 15)
276
Remain docked with parent node
277
Set propagation rating to next higher rating, undock from parent node and confirm that sibling which was green before receipt of propagation has retransmitted propagation signal to its child nodes (if any)
278
Devise new connection address list, which is the ancestor portion of the topology database with the "red" node's "green" sibling placed in front of the "red" node's current nominal parent node (i.e., placed as the first node on the new connection address list), and go to Prospective Child Node's Connection Routine (See Figs. 14 and 15)
279
Done

Fig. 27

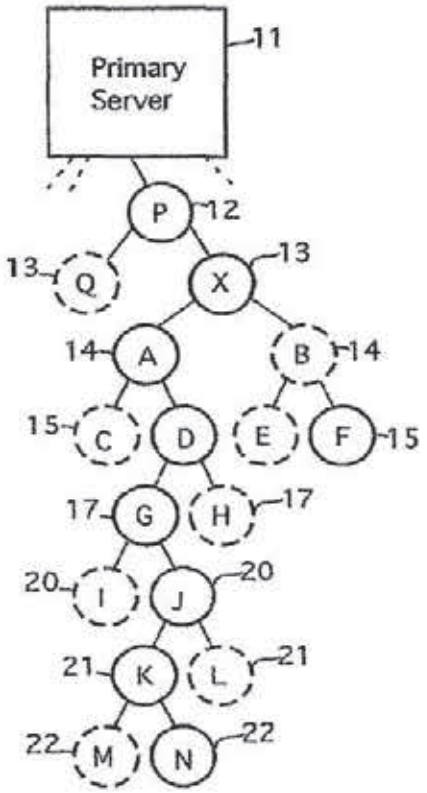

Primary Server
11
P
12
13
Q
X
13
14
A
B
14
15
C
D
E
F
15
17
G
H
17
20
I
J
20
21
K
L
21
22
M
N
22

Fig. 28

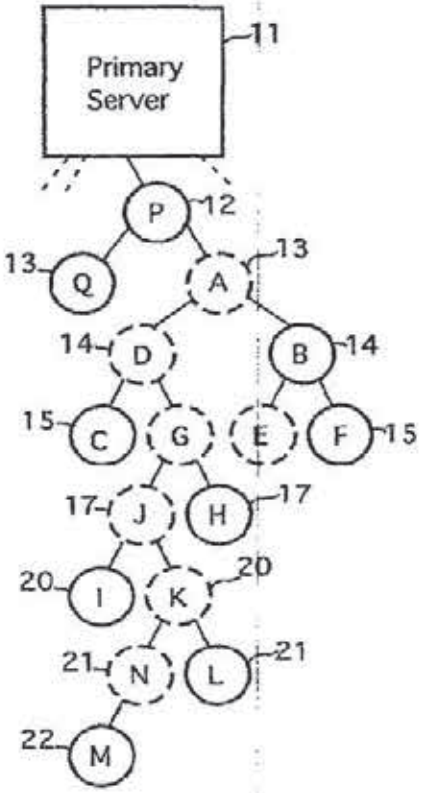

Primary Server
11
P
12
13
Q
A
13
14
D
B
14
15
C
G
E
F
15
17
J
H
17
20
I
K
20
21
N
L
21
22
M

Fig. 29

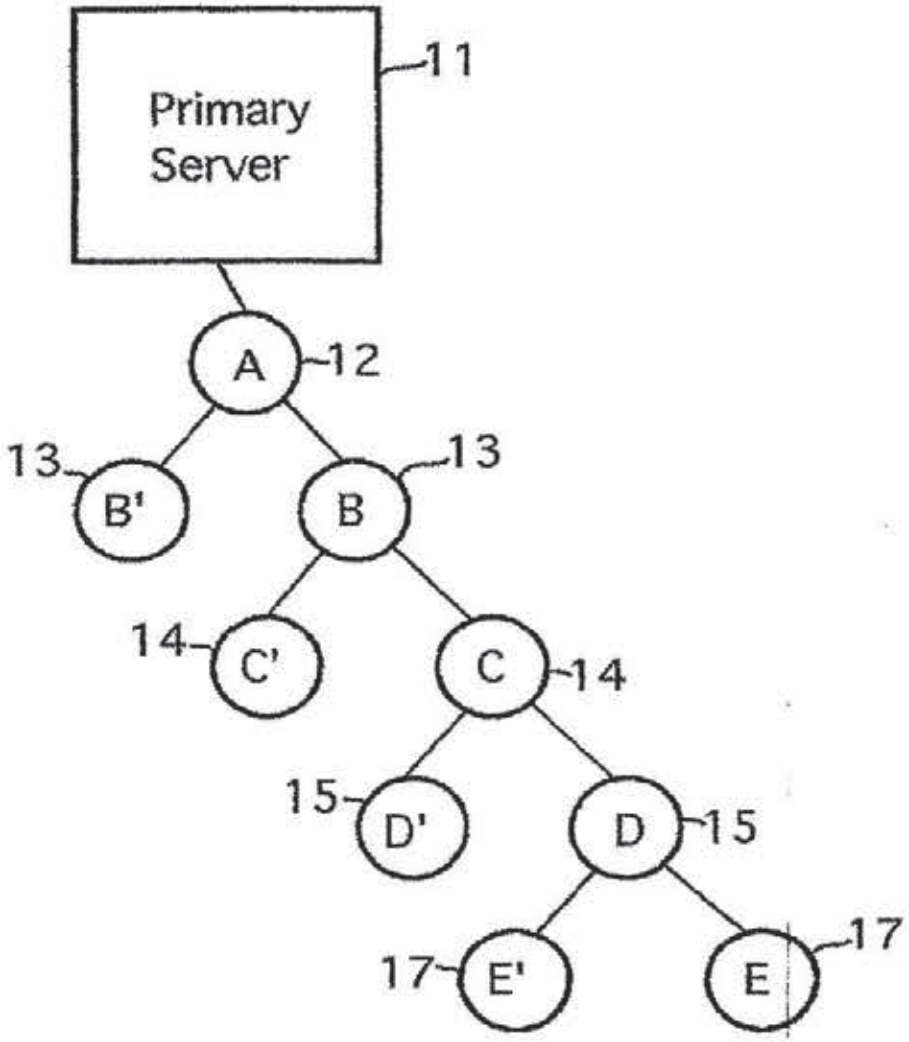


Fig. 30

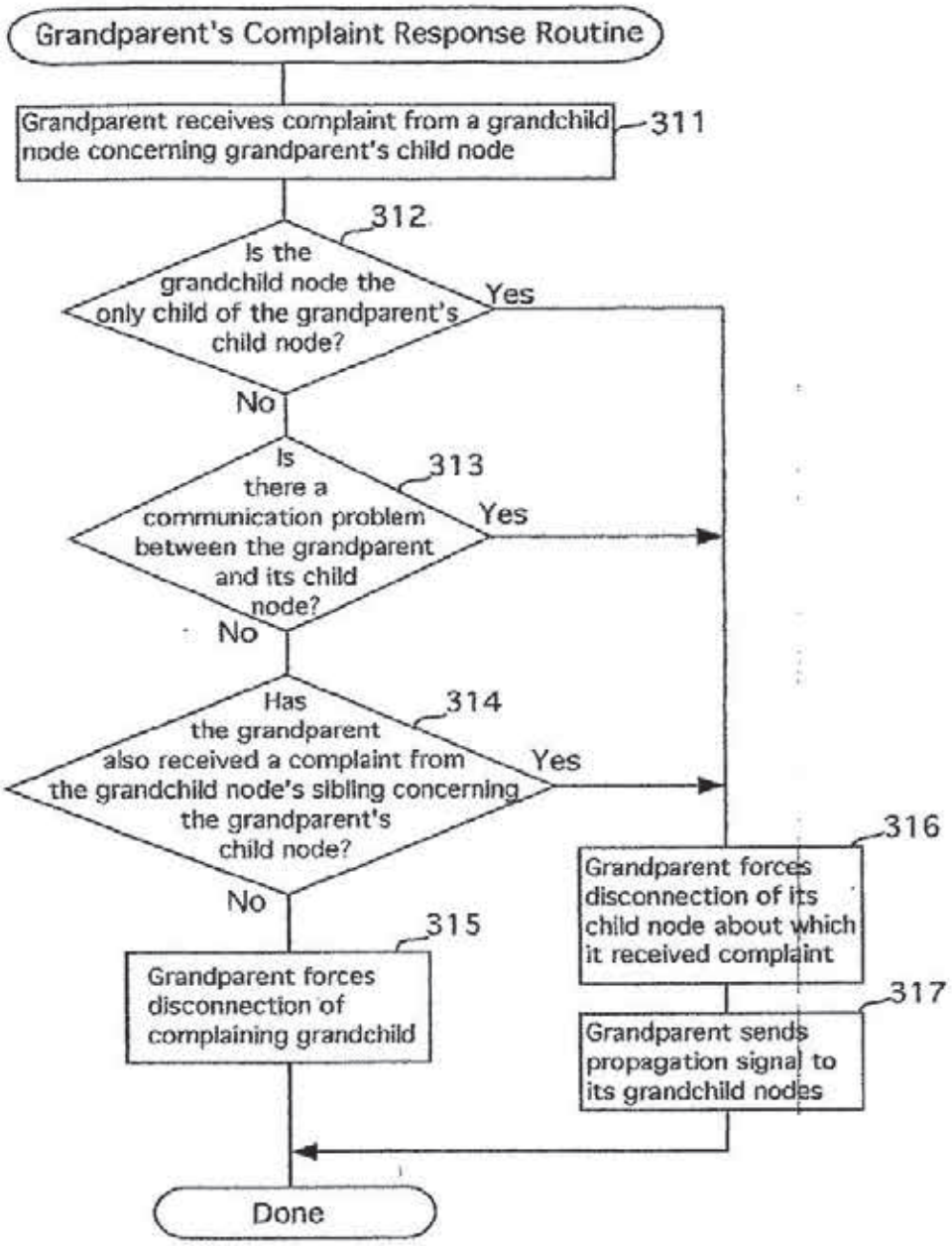

Grandparent's Complaint Response Routine
Grandparent receives complaint from a grandchild node concerning grandparent's child node
311
312
Is the grandchild node the only child of the grandparent's child node?
Yes
No
313
Is there a communication problem between the grandparent and its child node?
Yes
No
314
Has the grandparent also received a complaint from the grandchild node's sibling concerning the grandparent's child node?
Yes
No
315
Grandparent forces disconnection of complaining grandchild
316
Grandparent forces disconnection of its child node about which it received complaint
317
Grandparent sends propagation signal to its grandchild nodes
Done

Fig. 31



Fig. 32

Fig. 33A

Fig. 33B

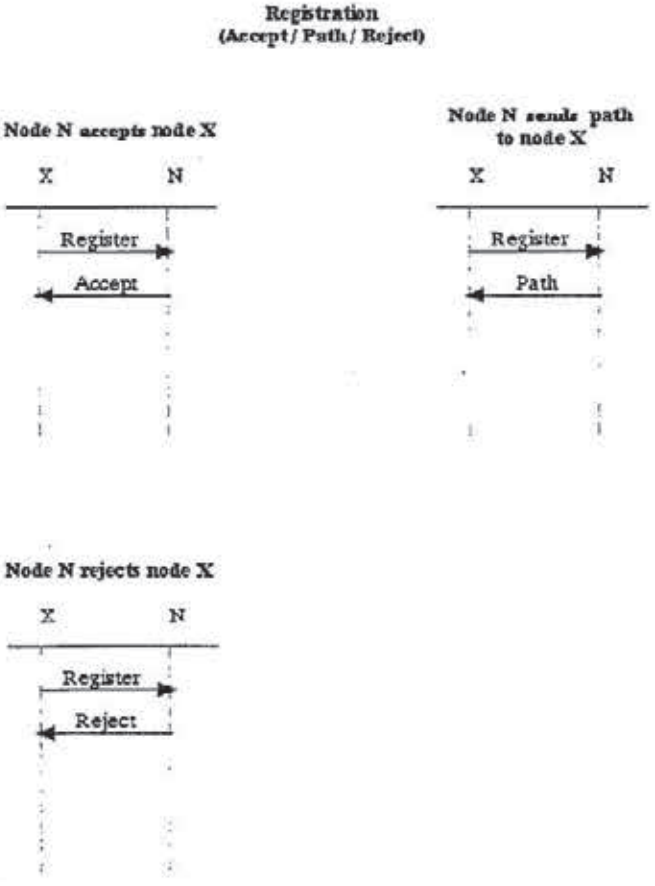


Fig. 33C

Fig. 34A

Fig. 34B

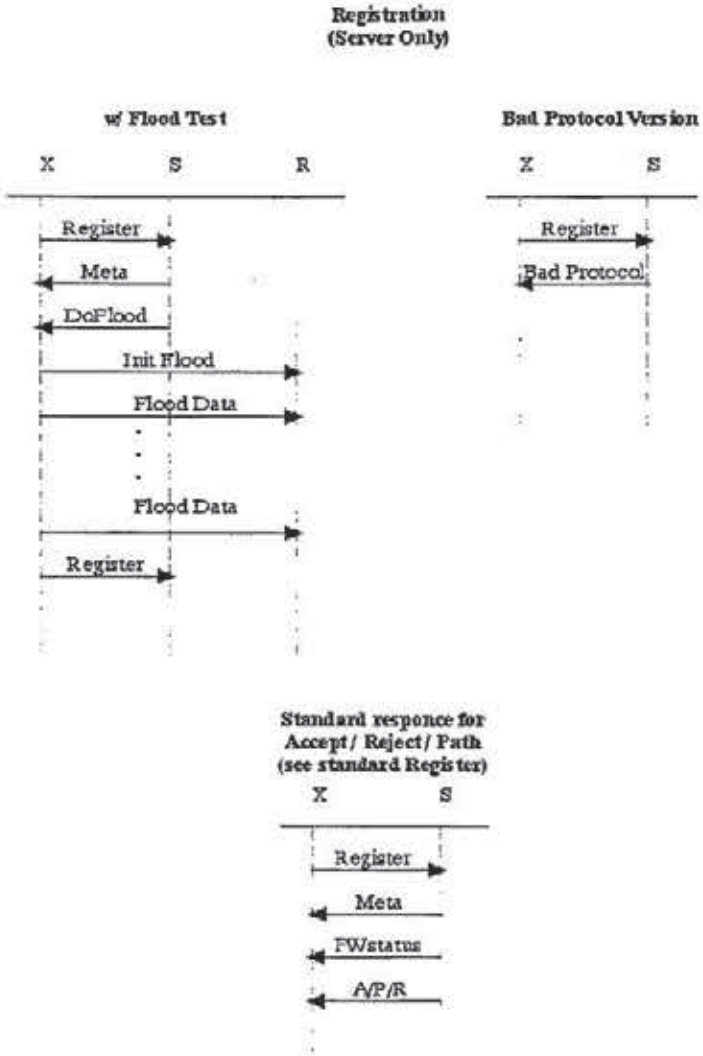


Fig. 34C

Fig. 35A

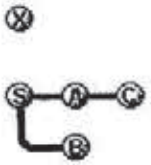
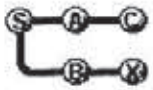


Fig. 35B



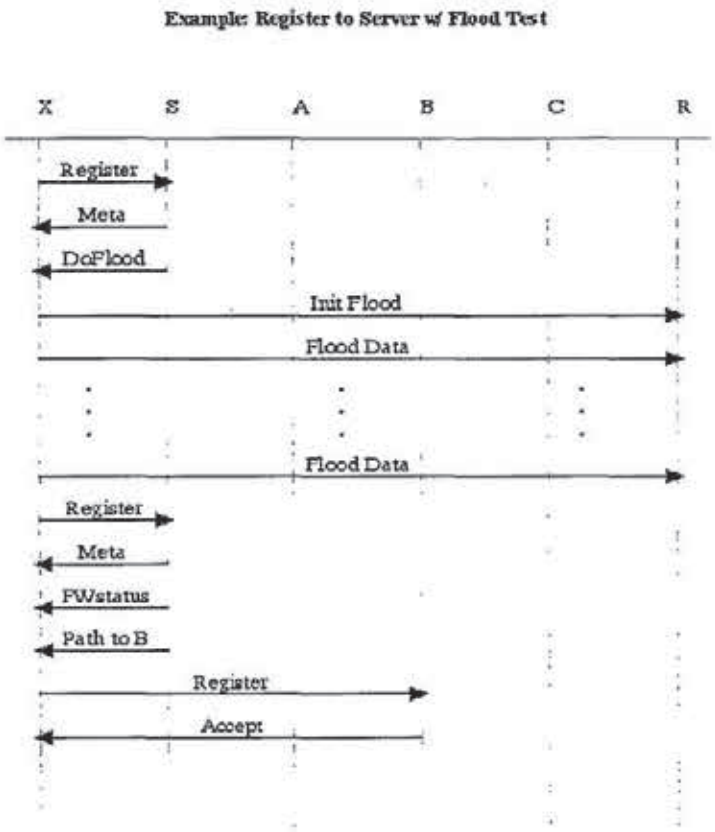


Fig. 36

Fig. 37



Fig. 38

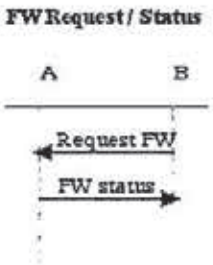


Fig. 39A



Fig. 39B

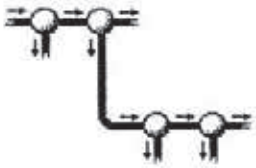


Fig. 40

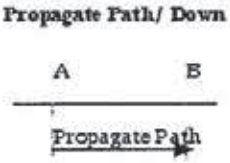


Fig. 41A



Fig. 41B

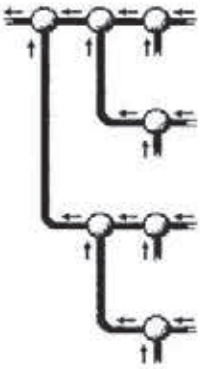


Fig. 42

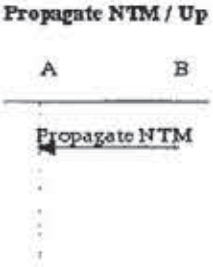


Fig. 43

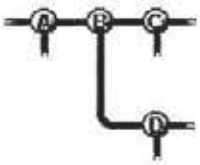


Fig. 44

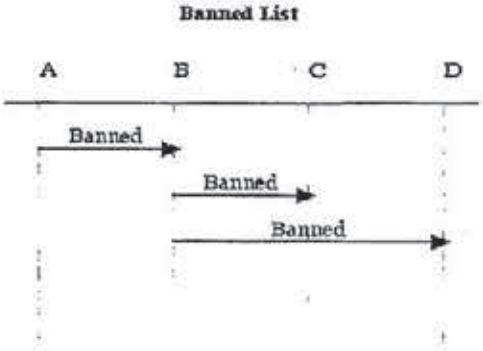


Fig. 43

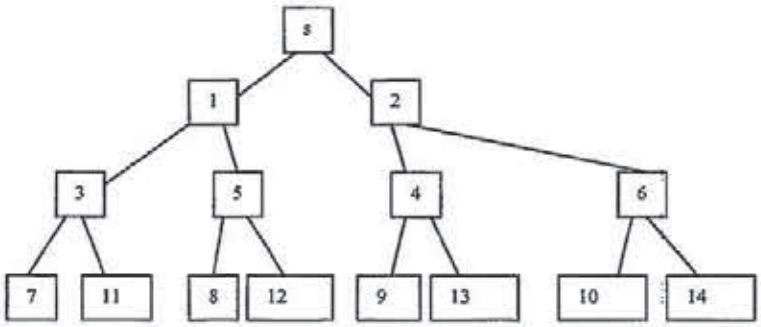


Fig. 46A

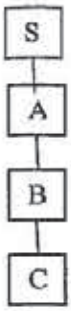


Fig. 46B

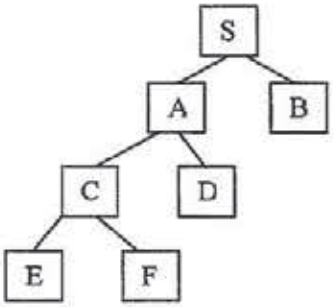
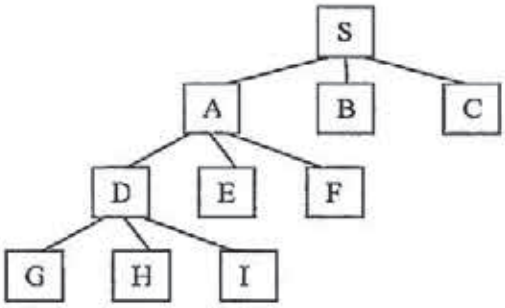


Fig. 46C



PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-092752
(43)Date of publication of application : 06.04.2001

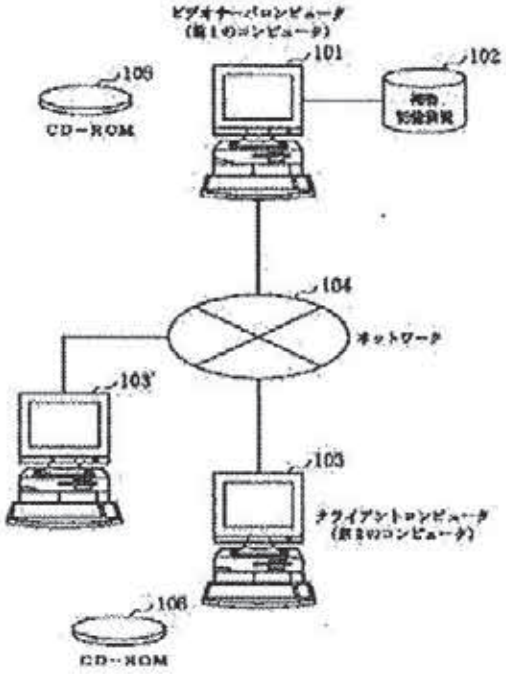
(51)Int.Cl. G06F 13/00
G06F 12/00
H04L 12/56
H04N 7/24

(21)Application number : 11-271422
(22)Date of filing : 24.09.1999

(71)Applicant : HITACHI INFORMATION SYSTEMS LTD
(72)Inventor : SHIOMI YOSHIHIRO
AIKAWA NOBUYUKI
NIGO MASAKI

(54) SYSTEM FOR DISTRIBUTING PICTURE DATA AND RECORDING MEDIUM USED FOR THE SAME

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a picture data distribution system capable of efficiently distributing picture data with appropriate quality in accordance with changes of the network connection environment of a client computer from moment to moment and a recording medium on which its program is recorded.
SOLUTION: This system has a 1st computer (image server computer 101) capable of distributing the picture data and one or more 2nd computers (client computer 103) receiving the picture data, and the 1st computer 101 receives a line situation notification message from the 2nd computer 103 and transmits the picture data (picture data extracted among a plurality of pieces of picture data which are preliminarily stored and have different compressibility or picture data compressed with optimum compressibility every time) which are subjected to compressibility being optimum to the communication speed of a line to which the 2nd computer is connected on the basis of the message.



Ex. 1018-1810

(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2001－92752
（P2001－92752A）
(43)公開日　平成13年４月６日(2001.4.6)

(51)Int.Cl.⁷		識別記号	FI		テーマコード(参考)
G06F	13/00	354	G06F	13/00	354D 5B082
	12/00	545		12/00	545M 5B089
H04L	12/56		H04L	11/20	102A 5C059
H04N	7/24		H04N	7/13	Z 5K030

審査請求　未請求　請求項の数５　ＯＬ　（全 11 頁）

(21)出願番号　特願平11－271422

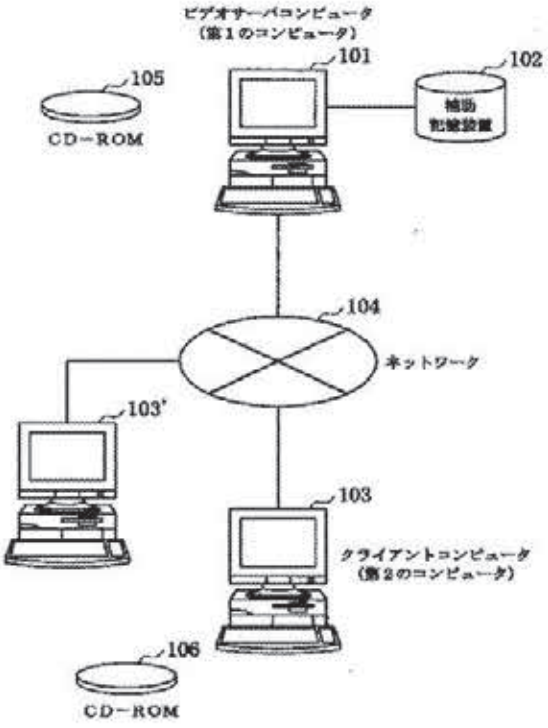
(22)出願日　平成11年９月24日(1999.9.24)

(71)出願人　000152985
株式会社日立情報システムズ
東京都渋谷区道玄坂１丁目16番５号
(72)発明者　塩見　芳弘
東京都渋谷区道玄坂一丁目16番５号　株式会社日立情報システムズ内
(72)発明者　相川　伸之
東京都渋谷区道玄坂一丁目16番５号　株式会社日立情報システムズ内
(74)代理人　100077274
弁理士　磯村　雅俊　（外１名）

最終頁に続く

(54)【発明の名称】　画像データ配信システムおよびそれに用いる記録媒体

(57)【要約】
【課題】　クライアントコンピュータのネットワーク接続環境の時々刻々の変化に応じて、適切な品質の画像データを効率よく配信することが可能な画像データ配信システムおよびそのためのプログラムを記録した記録媒体の提供。
【解決手段】　画像データを配信可能な第１のコンピュータ（画像サーバコンピュータ１０１）とその画像データを受信する１以上の第２のコンピュータ（クライアントコンピュータ１０３）を有し、第１のコンピュータ１０１は、第２のコンピュータ１０３から回線状況通知メッセージを受信し、そのメッセージに基づいて、第２のコンピュータが接続されている回線の通信速度に最適な圧縮を施した画像データ（予め記憶している圧縮率の異なる複数の画像データから抽出した画像データかまたはその都度最適な圧縮率で圧縮した画像データ）を送信する。



【特許請求の範囲】
【請求項１】　画像データをストリーミング配信可能な第１のコンピュータとその画像データを受信する１以上の第２のコンピュータを有する画像データ配信システムにおいて、
前記第１のコンピュータは、
前記第２のコンピュータから回線状況通知メッセージに基づいて、該第２のコンピュータが接続されている回線の通信速度に最適な圧縮を施した画像データを送信する最適画像配信手段を有することを特徴とする画像データ配信システム。
【請求項２】　前記最適画像配信手段は、
送信時間を計測するタイマと、
圧縮率の異なる複数の画像データおよび該画像データ毎に該画像データの内容を識別可能な内部識別アドレスと送信の際の送信タイミングを規定するタイムスタンプとの対応データを保持する記憶手段と、
前記第２のコンピュータからの回線状況通知メッセージに基づいて前記記憶手段から最適な圧縮率の画像データを抽出するとともに、該抽出した画像データの、前記タイマで計測された送信時間に対応する前記内部識別アドレス以降の画像データ内容を配信する手段とを有することを特徴とする請求項１記載の画像データ配信システム。
【請求項３】　前記最適画像送出手段は、
１つの画像データを保持する記憶手段と、
前記第２のコンピュータから所定の時間間隔で受信する回線状況通知メッセージに基づいて、前記記憶手段に保持されている１つの画像データから、該第２のコンピュータが接続されている回線に最適な圧縮率の画像データを動的に作成して配信する手段とを有することを特徴とする請求項１記載の画像データ配信システム。
【請求項４】　前記第１のコンピュータは、画像サーバコンピュータであり、前記第２のコンピュータは、クライアントコンピュータであることを特徴とする請求項１または請求項２記載の画像データ配信システム。
【請求項５】　請求項１または２記載の第１のコンピュータにおける各手段および第２のコンピュータにおける各手段を実現する処理をプログラムコード化して記憶したことを特徴とするコンピュータで読み取り可能な記録媒体。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】本発明は、ネットワーク上でビデオストリーミングデータを送信する場合におけるデータ作成および配信システムに関し、特にデータを要求するクライアントコンピュータのネットワーク接続環境の動的な変化に対応して最適なビデオデータ配信を実現するビデオデータ配信システムおよびそのためのプログラムを記録した記録媒体に関する。
【０００２】
【従来の技術】ネットワーク上における一般的なビデオデータ配信においては、ビデオサーバコンピュータはクライアントコンピュータのデータ送信要求に対し該当するデータのストリーミング配信を行なっている。通常のビデオデータは非常にサイズが大きいため、回線上に大きな帯域幅を必要としていた。従って、従来は、限られたネットワーク内において回線の帯域幅を考慮し適切な圧縮を行うことにより、ビデオデータ配信が行われていた。
【０００３】しかし、近年のインターネット技術の普及に伴い、インターネットを利用したビデオ配信システムが多数考案されている。インターネットを利用した場合、インターネットに接続され世界中に点在する膨大な数のコンピュータがクライアントコンピュータとなり得る。そのため、クライアントコンピュータのネットワーク接続環境はクライアントコンピュータ毎に異なることとなる。
【０００４】従って、例えば高品質（低圧縮率）のビデオデータを配信した場合、帯域幅の狭い通信回線に接続しているクライアントコンピュータは、配信された全てのビデオデータを受信しきれなくなる。この理由により、インターネット上でビデオ配信を行なう場合には、できるだけ多くのクライアントコンピュータで受信可能となるように、帯域幅の狭い通信回線に合わせた低品質（高圧縮率）のビデオデータを配信するようにしていた。この環境においては、高品質のビデオデータの受信可能なクライアントコンピュータであっても品質の低いビデオデータを受信することに甘んじなければならなかった。
【０００５】この問題点に関するものとして、ビデオサーバが、予め記憶装置に格納してある中間形態の動画データを事前に高品質および低品質の複数のデータを準備し、さらにビデオデータを要求するクライアントコンピュータが接続されている回線の通信速度を取得する手段を具備し、取得した通信速度により最適なデータを配信する仕組みが考案されている。
【０００６】
【発明が解決しようとする課題】しかしながら、一般に、クライアントコンピュータのネットワーク接続環境すなわち帯域幅は回線の混雑の度合等により刻々と変化するものであるが、上記従来技術においては、動画データの配信要求のあったクライアント端末のネットワーク性能（伝送速度）を検知し、それに応じた圧縮率の動画データを送信するものであり、ネットワーク性能の時々刻々の変化に応じて送信ビデオデータの品質（データ圧縮率）を変化させるものとはなっていない。
【０００７】本発明の目的は、クライアントコンピュータのネットワーク接続環境すなわち帯域幅の時々刻々の変化に応じて、適切な品質の画像データを効率よく配信

することが可能な画像データ配信システムおよびそのためのプログラムを記録した記録媒体を提供することである。

【０００８】

【課題を解決するための手段】本発明は、上記目的を達成するために、画像データをストリーミング配信可能な第１のコンピュータ（好ましくは画像サーバコンピュータ）とその画像データを受信する１以上の第２のコンピュータ（好ましくはクライアントコンピュータ）を有する画像データ配信システムにおいて、第１のコンピュータは、第２のコンピュータから回線状況通知メッセージに基づいて、第２のコンピュータが接続されている回線の通信速度に最適な圧縮を施した画像データを送信する最適画像配信手段を有することを特徴としている。

【０００９】なお、前記最適画像配信手段は、（ａ）送信時間を計測するタイマと、圧縮率の異なる複数の画像データおよび該画像データ毎に該画像データの内容を識別可能な内部識別アドレスと送信の際の送信タイミングを規定するタイムスタンプとの対応データを保持する記憶手段と、第２のコンピュータからの回線状況通知メッセージに基づいて記憶手段から最適な圧縮率の画像データを抽出するとともに、該抽出した画像データの、タイマで計測された送信時間に対応する内部識別アドレス以降の画像データ内容を配信する手段とを有するか、あるいは、（ｂ）１つの画像データを保持する記憶手段と、第２のコンピュータから所定の時間間隔で受信する回線状況通知メッセージに基づいて、記憶手段に保持されている１つの画像データから、第２のコンピュータが接続されている回線に最適な圧縮率の画像データを動的に作成して配信する手段とを有している。

【００１０】また、本発明の記録媒体は、前記第１のコンピュータ（画像サーバコンピュータ）における各手段および第２のコンピュータ（クライアントコンピュータ）における各手段を実現する処理をプログラムコード化してコンピュータ読み取り可能なＣＤ－ＲＯＭ，ＤＶＤ，ＦＤなどの記録媒体に記録したものである。

【００１１】

【発明の実施の形態】以下、本発明の好ましい実施の形態を説明する。画像データを配信する機能を有するネットワークに接続されている画像配信サーバを具備し、該画像配信サーバは画像データを格納可能な記憶装置を具備し、画像配信サーバにアクセス可能な不特定多数のクライアントコンピュータを具備する構成とする。

【００１２】配信するビデオデータ作成において、ビデオサーバコンピュータは、オリジナルビデオデータを基に、高品質から低品質まで複数のビデオデータを記憶装置に作成し、個々のビデオデータに対し一定時間間隔のタイムスタンプを付加する機能を具備する。

【００１３】また、ビデオデータを受信するクライアントコンピュータでは、ビデオデータ受信中に受信パケットの状態を監視し、損失パケットが一定の基準を超えた場合、クライアント側の通信回線に十分な帯域が確保されていないと判定し、また、一定の基準時間の間に損失パケットが無い場合は、クライアント側の通信回線に十分な帯域が確保されていると判定し、その判定結果を配信中のビデオサーバコンピュータへ送信する手段を具備する。

【００１４】上述の判定結果を受信したビデオサーバは、その判定結果に適切な品質のビデオデータファイルを選択し、現在配信中のビデオデータファイルの切り換えを行う。その際、切り換え前後の２つのビデオデータに付加されているタイムスタンプを参照し、配信を中止する時点でのタイムスタンプと同様のタイムスタンプを持つ時点より切り換え後のビデオデータの配信を開始する機能を具備する。

【００１５】また、上記目的を達成するために、本発明は、ビデオデータを配信する機能を有するネットワークに接続されているビデオサーバコンピュータを具備し、さらにその記憶装置に記憶されているオリジナルビデオデータを送信時に所定の圧縮率で圧縮しながら配信する機能を具備する構成とする。

【００１６】ビデオデータを受信する際において、クライアントコンピュータは、ビデオデータ受信中に受信パケットの状態を監視し、損失パケットが一定の基準を超えた場合クライアント側の通信回線に十分な帯域が確保されていないと判定し、また、一定の基準の時間、損失パケットが無い場合、クライアント側の通信回線に十分な帯域が確保されていると判定し、その判定結果を配信中のビデオサーバコンピュータへ送信する手段を特徴とする機能を具備する。

【００１７】さらに、上述の判定結果を受信したビデオサーバコンピュータは、その判定結果に従い、動的に圧縮率に変更し、その変更された圧縮率で圧縮されたビデオデータを配信する機能を具備する。

【００１８】上記構成を採用することにより、配信されるビデオデータは、受信するクライアントコンピュータのネットワーク接続環境に従い配信するビデオデータの品質を動的に変化させることができる。これにより、クライアントコンピュータは常に最適化されたビデオデータを見掛け上切れ目なく受信することが可能となる。

【００１９】また、本発明の記録媒体は、上記の如きクライアントコンピュータが行う処理やビデオサーバコンピュータが行う処理をプログラムコード化してコンピュータ読み取り可能なＣＤ－ＲＯＭ，ＤＶＤ，ＦＤなどの記録媒体に記録したものである。

【００２０】（実施例の説明）以下、本発明の実施例として、２つの実施例を説明する。図１は、第１の実施例および第２の実施例における共通のハードウエア構成を示すブロック図である。本発明は、圧縮して配信することが可能な全てのデータ（特に画像データ）に対して適

用可能であるが、本実施例では、ビデオデータ（動画データ）を例にとって説明する。

【００２１】図１において、１０１はビデオデータの作成および配信を行なうビデオサーバコンピュータ、１０２はビデオデータを格納する補助記憶装置、１０３および１０３’はクライアントコンピュータ、１０４はビデオサーバコンピュータ１０１とクライアントコンピュータ１０３が接続されるネットワーク環境である。以下の説明では、１つのサーバコンピュータと１つ以上のクライアントコンピュータからなるクライアント・サーバシステムの場合を用いて説明するが、第１のコンピュータから第２のコンピュータへの２つのコンピュータ間でのデータの配信にも適用可能であることはいうまでもない。

【００２２】（第１の実施例）まず、第１の実施例について図２～図６を用いて説明する。図２は、記憶装置に格納する圧縮ビデオデータファイルのファイル構造の説明図である。図２において、２０３は圧縮ビデオデータファイルのヘッダ情報、２０４は本実施例で取り上げる画像データファイル、２０５はＥＯＦ（ＥＯＦ：End of File）を示す。また、２０１は本発明の特徴である圧縮ビデオデータファイルのタイムスタンプであり、２０２は、タイムスタンプ２０１と対応する圧縮ビデオデータファイルの内部識別アドレスである。なお、本実施例で取り上げる圧縮率の異なる他の画像ファイル（圧縮画像ファイルＢおよび圧縮画像ファイルＣ）構造も図２と同様である。内部識別アドレス２０２は、ファイル毎にタイムスタンプ１０２に対応して設けられるものであるが、ファイルの品質（圧縮率）が異なればそれに伴って異なる値を有する。

【００２３】図３は、同一画像ファイルから生成された圧縮率の異なる３種類の圧縮ビデオデータファイルのヘッダ情報の説明図である。同図（ａ）は低圧縮画像ファイルＡのヘッダ情報を、同図（ｂ）は中圧縮画像ファイルＢのヘッダ情報を、同図（ｃ）は高圧縮画像ファイルＣのヘッダ情報を示している。３０１はタイムスタンプ情報、３０２は、インデックスに対応する図２に示すファイルの内部識別アドレスである。

【００２４】図４は、圧縮率の異なる３つの画像ファイルとタイムスタンプと内部識別アドレスの関係を説明するための図である。同図（ａ）は、低圧縮画像ファイルＡとタイムスタンプと内部識別アドレスの関係を、同図（ｂ）は中圧縮画像ファイルＢとタイムスタンプと内部識別アドレスの関係を、同図（ｃ）は高圧縮画像ファイルＣとタイムスタンプと内部識別アドレスの関係を、それぞれ示している。本例では、説明を簡単にするために、同図（ａ）に示す低圧縮画像ファイルＡを１／２に圧縮したものを中圧縮ファイルＢとして同図（ｂ）に示し、さらにこれを１／２に圧縮したものを高圧縮画像ファイルＣとして同図（ｃ）に示すとともに、タイムスタンプ，内部識別アドレスと圧縮画像ファイルＡ～Ｃの内容の関係がわかりやすいように画像ファイル中に「い」、「ろ」、「は」・・・などを便宜的に示した。

【００２５】図５は、ビデオサーバコンピュータ１０１に保持される、通信速度と該通信速度に最適化された圧縮画像ファイルとの対応図である。図５における通信速度を表わす記号（ｍ、ｎ）の大小関係は、ｍ＜ｎとする。同図に示すように、通信速度が低速の場合は（～ｍ）、低品質である高圧縮画像ファイルＣを、通信速度が高速の場合は（ｎ～）、高品質である低圧縮画像ファイルＡを配信するようにする。

【００２６】図６および図７は、第１の実施例の動作を説明するためのフローチャートである。まず、ビデオデータを配信するビデオサーバコンピュータ１０１の動作を、図６を用いて詳細に説明する。

【００２７】図６において、クライアントコンピュータ１０３からのデータ要求に従い、該当する記憶装置内のビデオデータ（システムには初期配信ファイルが設定済とする。本実施例では、初期配信ファイルは「中圧縮画像ファイルＢ」とする）の配信を開始し、かつ配信時間の計測を開始する（ステップ６０１）。

【００２８】クライアントコンピュータ１０３から低品質データの受信要求（後述する図７のステップ７０６参照）があった場合（ステップ６０３）、図５の対応図に従い、現在送信中のファイル（ファイルＢ）より１段階品質の劣る「高圧縮画像ファイルＣ」を送信ファイルとして確定する（ステップ６０４、６０５）。

【００２９】次に現在時点の配信時間を測定する（ステップ６０６）。本実施例では、この時の測定時間が４ｍｓであったと仮定し、図３に示すヘッダ情報より、「ファイルＢ」の４ｍＳの開始内部識別アドレスは「iiii」と確定する（ステップ６０７）。すなわち、画像「い」、「ろ」は配信済みとする。ビデオサーバコンピュータ１０１は即座に配信中の「中圧縮画像ファイルＢ」の配信を中止し、内部識別アドレス「iiii」から「高圧縮画像ファイルＣ」の画像「は」、「に」、・・・の配信を開始する（ステップ６０８）。

【００３０】逆に、クライアントコンピュータ１０３より高品質データの受信要求（後述する図７のステップ７０７参照）があった場合（ステップ６０３：Ｙ）、図５の対応図に従い、現在配信中の「中圧縮画像ファイルＢ」より１段階品質の高い「低圧縮画像ファイルＡ」を次送信ファイルとして確定する（ステップ６０４、６０５）。以下の動作は同様である。すなわち、ビデオサーバコンピュータ１０１は即座に配信中の「中圧縮画像ファイルＢ」の配信を中止し、内部識別アドレス「iiii」から「低圧縮画像ファイルＡ」の画像「は」、「に」、・・・の配信を開始する（ステップ６０８）。

【００３１】なお、現時点での配信データが「高圧縮画像ファイルＣ」であった場合にさらに低品質の高圧縮の

画像ファイルが要求された場合、図５より１段階品質の劣るビデオデータは存在しないことがわかる。また、現時点での配信データが「低圧縮画像ファイルＡ」であった場合にさらに高品質の低圧縮画像ファイルが要求された場合にも、図５より１段階品質の高いビデオデータは存在しないことがわかる。このような場合は、データの存在しない旨のメッセージを送信し、現時点でのデータ配信を続行する（ステップ６０９）。

【００３２】また、ビデオサーバコンピュータ１０１は、ビデオデータの終端すなわち図２に示す「ＥＯＦ」まで配信を行った場合（ステップ６０２：Ｙ）、データの配信を終える。

【００３３】次に、ビデオデータを受信するクライアントコンピュータ１０３の動作を、図７を用いて詳細に説明する。図７において、先ず、ビデオサーバコンピュータ１０１に対するビデオデータ要求により、ビデオサーバから送信されたデータの受信を開始し、同時に損失パケット監視タイマをリセットする（ステップ７０１）。

【００３４】損失パケット数がシステムに既設定済の数を超過した場合（ステップ７０２：Ｙ）、クライアントコンピュータ側の通信回線に十分な帯域が確保されていないと判断し、ビデオサーバに低品質データの要求メッセージを送信する（ステップ７０６）。損失パケット数がシステムに既設定済の数を超過しない場合（ステップ７０２：Ｎ）、損失パケットの発生を常に監視し（ステップ７０３）、損失パケットが発生した場合（ステップ７０３：Ｙ）、損失パケット監視タイマをリセットする（ステップ７０４）。

【００３５】この損失パケット監視タイマが、システムに既設定済の一定時刻を超過した場合、すなわち損失パケットが一定時間発生しなかった場合（ステップ７０３：Ｎ，７０５：Ｙ）、クライアントコンピュータ１０３側の通信回線に十分な帯域が確保されており現在受信中のビデオデータよりも高品質なデータ（低圧縮のデータ）の受信が可能と判断し、ビデオサーバコンピュータ１０１に高品質データの要求メッセージを送信する（ステップ７０７）。ステップ７０８では、ビデオサーバコンピュータ１０１からの配信データに従い今までと異なる品質の配信データを受信し、同時に損失パケット監視タイマをリセットする。

【００３６】ここで、図６のステップ６０９に示すデータ無しのメッセージを受信した場合、そのメッセージに従い、低品質データ無しの場合はステップ７０６の低品質データ要求メッセージ送信を、また、高品質データ無しの場合はステップ７０７の高品質データ要求メッセージ送信を中止する。なお、クライアントコンピュータ１０３はビデオデータの終端をしめす「ＥＯＦ」まで受信した場合（ステップ７０９：Ｙ）、受信動作を終了する。

【００３７】最後に、ビデオデータフォーマットについて詳述する。本実施例における配信ビデオデータのフォーマットは、図２に示すようにヘッダ情報部分（２０３）とデータ部分（２０４）とから構成され、ヘッダ部分には、図３に示すような、タイムスタンプ毎のアドレスが保持されている。ビデオサーバによりビデオデータを圧縮する際に、ヘッダ部分を生成し付加する。なお、図２および図３における２０１と３０１、２０２と３０２は同値である。

【００３８】以上、第１の実施例は、オリジナルなビデオデータを基に高品質（低圧縮）から低品質（高圧縮）まで、品質の異なる複数のビデオデータをビデオサーバの記憶装置に記憶しておき、クライアントコンピュータからの配信要求に対し、クライアントコンピュータにおける受信状態に応じて、最適な品質のビデオデータを動的に切替えながら配信するものである。

【００３９】（第２の実施例）次に説明する第２の実施例は、ビデオサーバに保持するビデオデータは、オリジナルなもの１種類のみとし、サーバから配信するビデオデータの圧縮を、クライアントコンピュータにおける受信状態に応じて動的に最適化するものである。

【００４０】以下、本第２実施例のビデオサーバ側の処理を、図８を用いて詳細に説明する。同図において，まず、クライアントコンピュータからのデータ要求に従い、該当する記憶装置内のオリジナルビデオデータをビデオサーバに取り出し（ステップ８０１）、圧縮を行いながら（システムには初期圧縮レートが設定済とする。）送信する（ステップ８０３）。クライアントコンピュータ１０３より低品質データ要求または高品質データ要求のメッセージ受信があった場合（ステップ８０４）、圧縮レートを変更する（ステップ８０５）。

【００４１】なお、ビデオサーバコンピュータ１０１はビデオデータの終端まで配信した場合（ステップ８０２）、配信動作を終了する。

【００４２】ビデオデータを受信するクライアントコンピュータの動作は、第１の実施例と同様のため詳細な説明は省略する。

【００４３】なお、上記の如き、ビデオサーバコンピュータが行う各処理手順およびクライアントコンピュータが行う各処理手順を、ＣＤ－ＲＯＭ，ＤＶＤ，あるいはＦＤなどの記録媒体に一緒にまたは別々に記録して市場に流通させることができる。利用者は、この記録媒体をビデオサーバコンピュータとクライアントコンピュータにインストールすることにより、本発明を簡単に利用することが可能になる。

【００４４】

【発明の効果】以上詳細に説明したように、本発明によれば、ビデオサーバにより配信されるビデオデータは、受信するクライアントコンピュータのネットワーク接続環境に従い動的に変化させることが実現できる。これにより、クライアントコンピュータは常に最適化されたビ

Ex. 1018-1815

デオデータを見掛け上切れ目なく受信することが可能となる。

【図面の簡単な説明】

【図１】本発明の第１の実施例および第２の実施例における共通のハードウエア構成を示すブロック図である。

【図２】記憶装置に格納する圧縮ビデオデータファイルのファイル構造の説明図である。

【図３】同一画像ファイルから生成された圧縮率の異なる３種類の圧縮ビデオデータファイルのヘッダ情報の説明図である。

【図４】圧縮率の異なる３つの画像ファイルとタイムスタンプと内部識別アドレスの関係を説明するための図である。

【図５】ビデオサーバコンピュータに保持される、通信速度と該通信速度に最適化された圧縮画像ファイルとの対応図である。

【図６】第１の実施例におけるビデオサーバコンピュータの動作を説明するためのフローチャートである。

【図７】第１の実施例におけるクライアントコンピュータの動作を説明するためのフローチャートである。

【図８】第２の実施例におけるビデオサーバコンピュータの動作を説明するためのフローチャートである。

【符号の説明】

１０１：ビデオデータの作成および配信を行なうビデオサーバコンピュータ（第１のコンピュータ）

１０２：ビデオデータを格納する記憶装置

１０３，１０３’：ネットワークと接続されるクライアントコンピュータ（第２のコンピュータ）

１０４：ビデオサーバコンピュータとクライアントコンピュータが接続されるネットワーク環境

１０５：プログラム記録媒体（ＣＤ－ＲＯＭ）

１０６：プログラム記録媒体（ＣＤ－ＲＯＭ）

２０１：圧縮ファイルのタイムスタンプ

２０２：圧縮ファイルのタイムスタンプに対応した内部識別アドレス

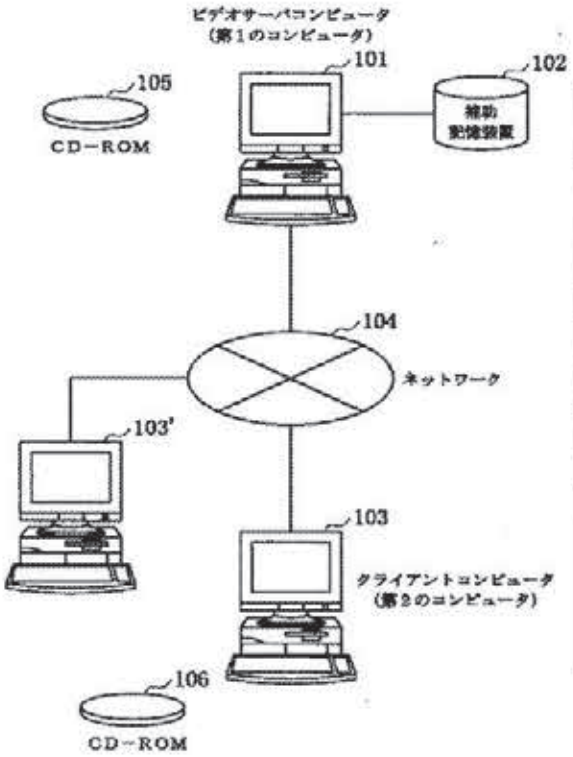
２０３：圧縮ファイルのファイルのヘッダ情報

２０４：本実施例で取り上げるファイルＡのファイルフォーマット

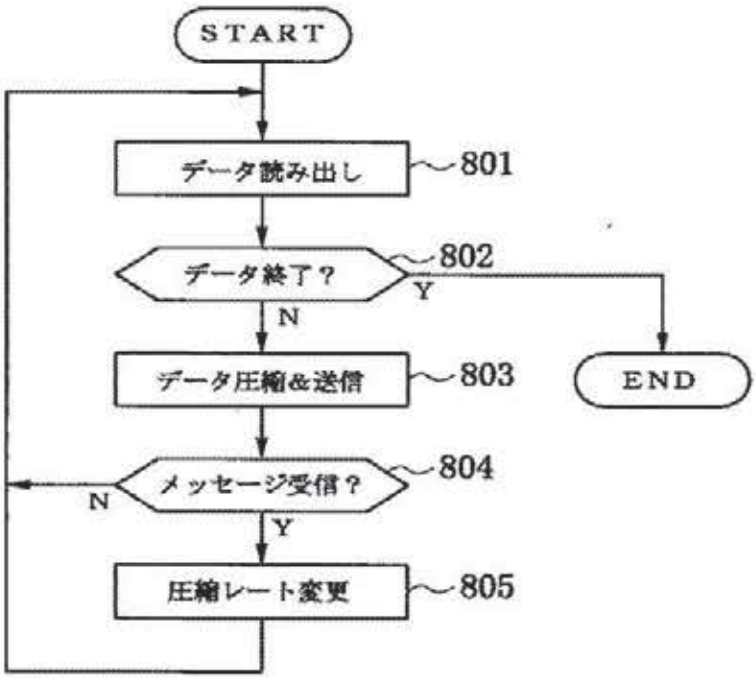
３０１：圧縮ファイルヘッダのタイムスタンプ情報

３０２：圧縮ファイルヘッダのタイムスタンプ情報に対応する内部識別アドレス

【図１】

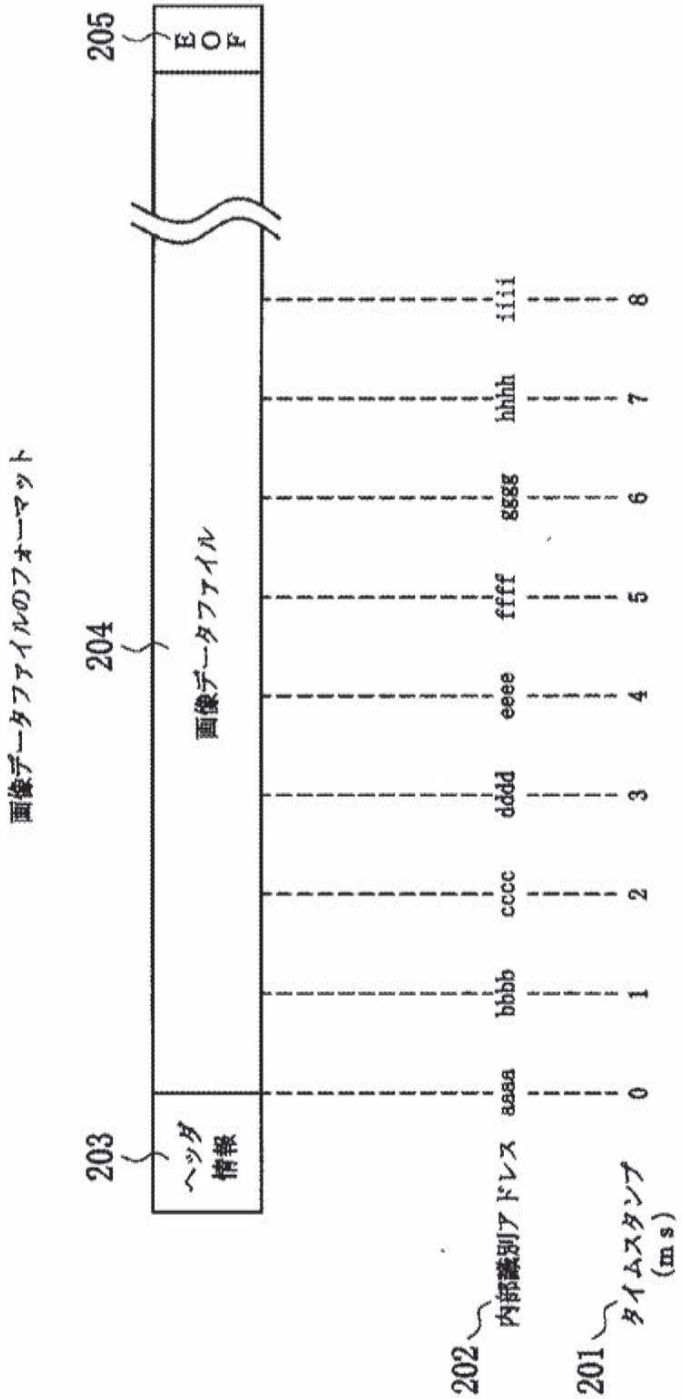


【図８】



Ex. 1018-1817

【図２】



【図３】

ヘッダ情報の構造

301 タイムスタンプ	302 内部識別アドレス
0	a a a a
1	b b b b
2	c c c c
3	d d d d
4	e e e e
5	f f f f
6	g g g g
7	h h h h
8	i i i i
:	:

低圧縮
画像ファイル（A）
（a）

301 タイムスタンプ	302 内部識別アドレス
0	a a a a
1	c c c c
2	e e e e
3	g g g g
4	i i i i
5	k k k k
6	m m m m
7	o o o o
8	q q q q
:	:

中圧縮
画像ファイル（B）
（b）

301 タイムスタンプ	302 内部識別アドレス
0	a a a a
1	e e e e
2	i i i i
3	m m m m
4	q q q q
5	u u u u
6	y y y y
7	:
8	:
:	:

高圧縮
画像ファイル（C）
（c）

【図５】

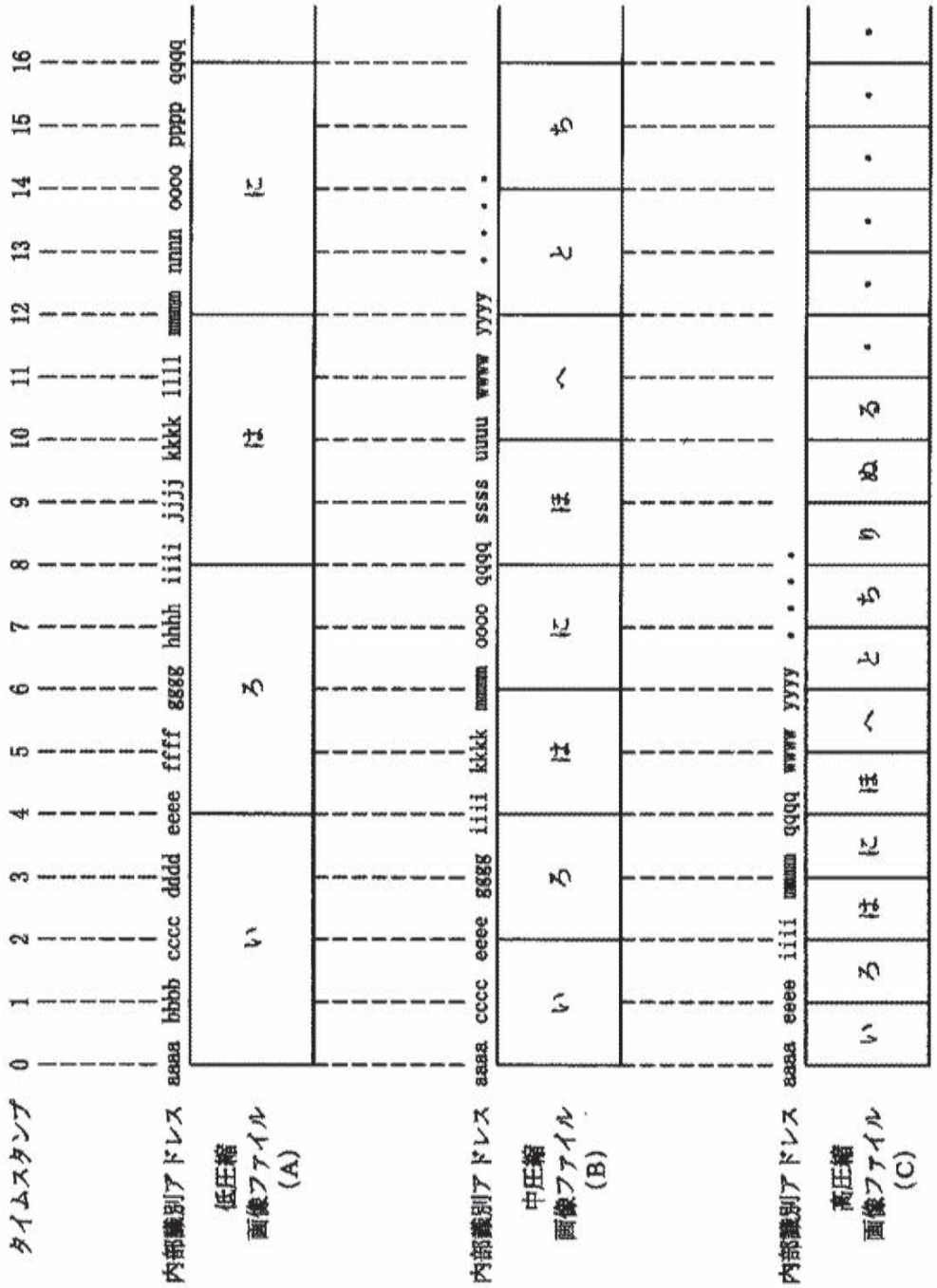
通信速度と対応圧縮ファイル

項番	通信速度（ｂｐｓ）	ファイル名
1	～m	圧縮画像ファイルC
2	m～n	圧縮画像ファイルB
3	n～	圧縮画像ファイルA

通信速度の大小関係　m＜n

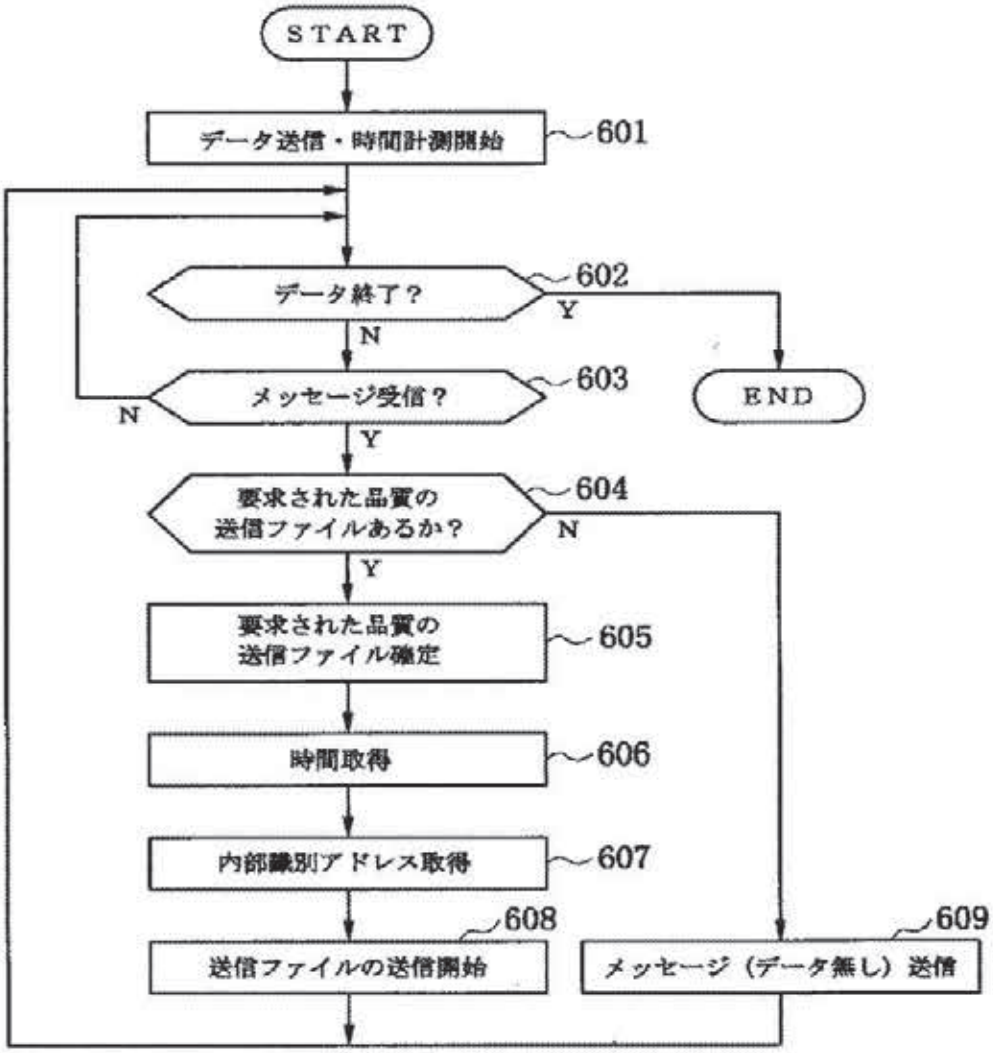
Ex. 1018-1819

【図４】



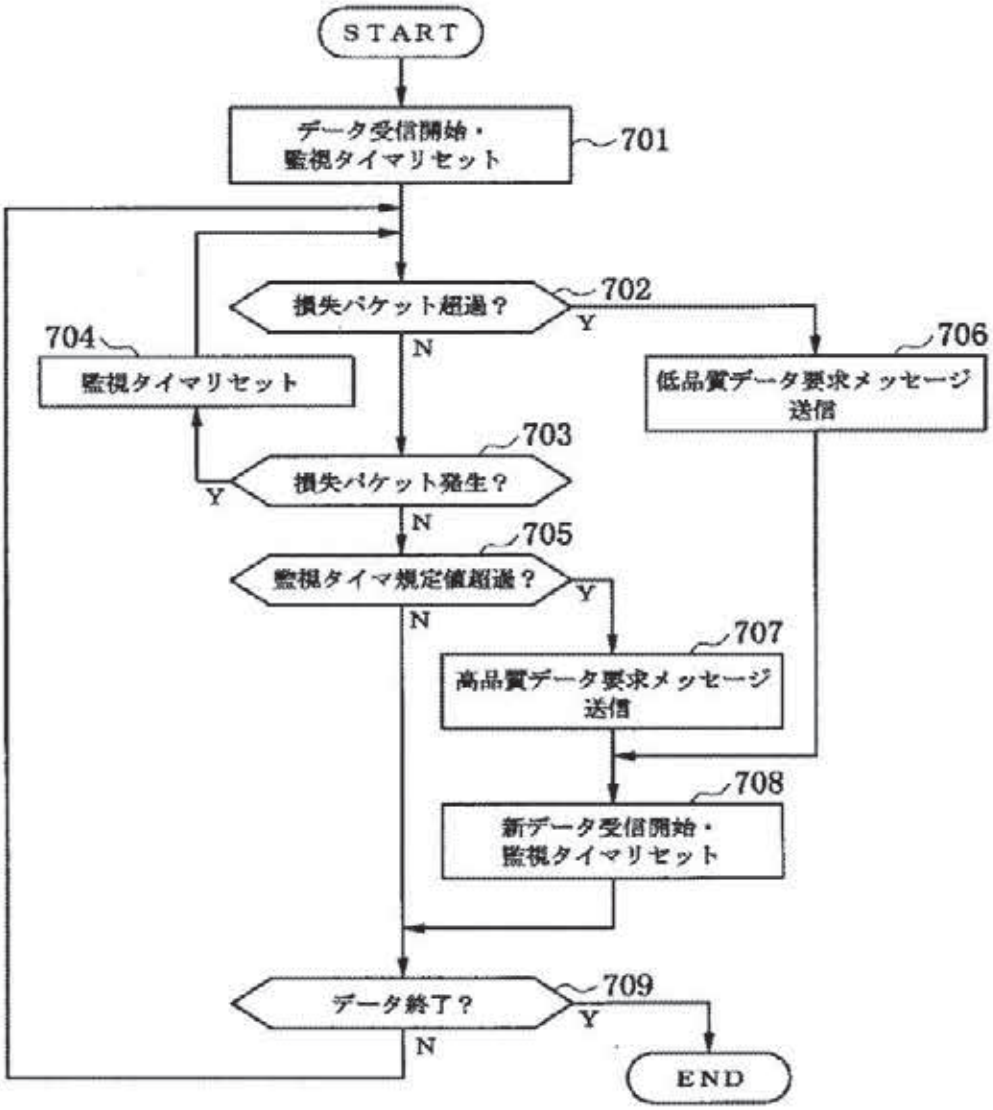
【図６】

(ビデオサーバコンピュータの処理)



【図７】

(クライアントコンピュータの処理)



フロントページの続き

(72)発明者　二郷　正樹
東京都渋谷区道玄坂一丁目16番５号　株式
会社日立情報システムズ内

Ｆターム(参考)　5B082 GA01 HA05
5B089 GA11 JA33 JB03 JB22 KA05
KA12 KB06 KB11 KC28 KC53
KC59 KC60 KH28
5C059 KK34 SS09 TA60 TA71 TC21
UA02 UA38
5K030 GA08 HB02 LA07 LC09 LD13
MB13

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2004-054930
(43)Date of publication of application : 19.02.2004

(51)Int.Cl. G06F 13/42
G06F 13/38
H04N 7/173

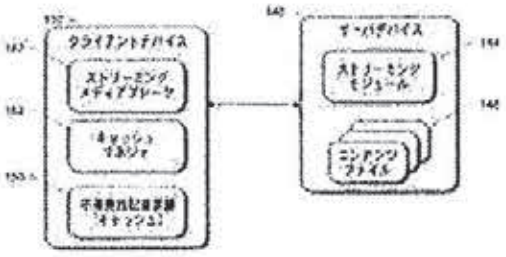
(21)Application number : 2003-180339
(22)Date of filing : 24.06.2003

(71)Applicant : MICROSOFT CORP
(72)Inventor : KLEMETS ANDERS E
BATTERBERRY TROY D
OLIVEIRA EDUARDO P

(30)Priority
Priority number : 2002 179770 Priority date : 24.06.2002 Priority country : US

(54) COMPUTER READABLE MEDIUM, SYSTEM, AND METHOD FOR CLIENT-SIDE CACHING OF STREAMING MEDIA CONTENT

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a technique for client-side caching of a streaming media content.
SOLUTION: Various functions with respect to streaming media content are made available to users. Such functions includes one or more functions for streaming the media content at a rate independent of the encoded bit rate of the content to allow streaming of the content to continue even when a user has selected various shuttle control options (e.g., pause, stop, fast forward, seek, rewind, etc.), a function for allowing the streaming content to be recorded for playback at a later time, and a function for allowing the streaming content to be played back while shifting time.



Ex. 1018-1822

(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2004-54930
(P2004-54930A)
(43)公開日　平成16年2月19日(2004.2.19)

(51)Int.Cl.[7]	F I			テーマコード（参考）
G06F 13/42	G06F	13/42	320C	5B077
G06F 13/38	G06F	13/38	340C	5C064
H04N 7/173	H04N	7/173	640Z	

審査請求 未請求　請求項の数 77　OL　（全 35 頁）

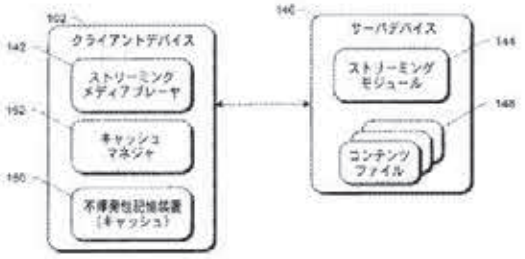
(21)出願番号　特願2003-180339(P2003-180339)
(22)出願日　平成15年6月24日(2003.6.24)
(31)優先権主張番号　10/179,770
(32)優先日　平成14年6月24日(2002.6.24)
(33)優先権主張国　米国(US)

(71)出願人　391055933
マイクロソフト　コーポレイション
ＭＩＣＲＯＳＯＦＴ　ＣＯＲＰＯＲＡＴＩＯＮ
アメリカ合衆国　ワシントン州　９８０５２－６３９９　レッドモンド　ワン　マイクロソフト　ウェイ　（番地なし）
(74)代理人　100077481
弁理士　谷　義一
(74)代理人　100088915
弁理士　阿部　和夫
(72)発明者　アンダース　イー．クレメッツ
アメリカ合衆国　９８１２２　ワシントン州　シアトル　２８　アベニュー　５２３

最終頁に続く

(54)【発明の名称】ストリーミングメディアコンテンツのクライアント側でのキャッシングのためのコンピュータ可読媒体、システムおよび方法

(57)【要約】
【課題】ストリーミングメディアコンテンツのクライアント側でのキャッシングの技術を提供する。
【解決手段】ストリーミングメディアコンテンツに関する各種の機能をユーザが利用できるようにする。そのような機能には、メディアコンテンツをそのコンテンツの符号化ビットレートに依存しないレートでストリーミングすることにより、ユーザが各種のシャトルコントロールオプション（例えば一時停止、停止、早送り、シーク、巻き戻しなど）を選択した際でも、コンテンツのストリーミングを継続することを可能にする、また後の時間に再生するためにストリーミングコンテンツを記録することを可能にする、また時間をずらしてストリーミングコンテンツを再生することを可能にする、のいずれか1つまたは複数が含まれる。
【選択図】　図2



【特許請求の範囲】
【請求項１】
クライアントデバイスの１つまたは複数のプロセッサによって実行されるときに，前記１つまたは複数のプロセッサにより，
サーバデバイスから，ストリーミングメディアコンテンツを該ストリーミングメディアコンテンツの符号化ビットレートに依存しないレートで受信させる
複数の命令が格納されていることを特徴とする１つまたは複数のコンピュータ可読媒体。
【請求項２】
前記複数の命令は，さらに前記１つまたは複数のプロセッサに，
前記受信したストリーミングメディアコンテンツを不揮発性の記憶装置に保存させる
ことを特徴とする請求項１に記載の１つまたは複数のコンピュータ可読媒体。
【請求項３】
前記サーバデバイスはプロキシサーバを含むことを特徴とする請求項１に記載の１つまたは複数のコンピュータ可読媒体。
【請求項４】
前記複数の命令はさらに前記１つまたは複数のプロセッサに，
前記ストリーミングメディアコンテンツについて要求されたレートを指示する速度要求を前記サーバデバイスに送信させ，
前記ストリーミングメディアコンテンツの前記レートの指示を前記サーバデバイスから受信させる
ことを特徴とする請求項１に記載の１つまたは複数のコンピュータ可読媒体。
【請求項５】
前記要求されたレートは値ｘで表され，前記速度要求は，
プラグマ：速度＝ｘ
を含むストリーミング速度メッセージヘッダであることを特徴とする請求項４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項６】
前記要求されたレートは値ｘで表され，前記速度要求は，
速度＝ｘ
を含むストリーミング速度メッセージヘッダであることを特徴とする請求項４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項７】
サーバデバイスで実施される方法であって，
ストリーミングメディアコンテンツに対する要求をクライアントデバイスから受信することと，
前記メディアコンテンツを前記コンテンツの符号化ビットレートに依存しないレートで前記クライアントデバイスにストリーミングすることと
を含むことを特徴とする方法。
【請求項８】
前記ストリーミングメディアコンテンツに対する要求をユーザから受信することと，
前記ユーザの要求に応答して，前記サーバデバイスに前記ストリーミングメディアコンテンツを要求することと
をさらに含むことを特徴とする請求項７に記載の方法。
【請求項９】
前記受信したストリーミングメディアコンテンツをキャッシュすること
をさらに含むことを特徴とする請求項７に記載の方法。
【請求項１０】
前記サーバデバイスは，プロキシサーバを含むことを特徴とする請求項７に記載の方法。
【請求項１１】
前記ストリーミングメディアコンテンツに要求されたレートを指示する速度要求を前記サ

ーバデバイスに送信することと、
前記ストリーミングメディアコンテンツの前記レートの指示を前記サーバデバイスから受信することと
をさらに含むことを特徴とする請求項７に記載の方法。
【請求項１２】
前記要求されたレートは値ｘで表され、前記速度要求は、
プラグマ：速度＝ｘ
を含むストリーミング速度メッセージヘッダであることを特徴とする請求項１１に記載の方法。
【請求項１３】
前記要求されたレートは値ｘで表され、前記速度要求は、速度＝ｘ
を含むストリーミング速度メッセージヘッダであることを特徴とする請求項１１に記載の方法。
【請求項１４】
クライアントデバイスの１つまたは複数のプロセッサによって実行されると、前記１つまたは複数のプロセッサに、
ストリーミングメディアコンテンツの再生を少なくとも一時的に停止する要求を受信させ、
前記クライアントデバイスにおける前記ストリーミングメディアコンテンツの再生を停止させ、
前記ストリーミングメディアコンテンツの再生を停止させる要求に関わらず、前記ストリーミングメディアコンテンツの受信を継続させる
複数の命令が格納されていることを特徴とする１つまたは複数のコンピュータ可読媒体。
【請求項１５】
前記要求は停止コマンドを含むことを特徴とする請求項１４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項１６】
前記要求は一時停止コマンドを含むことを特徴とする請求項１４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項１７】
前記複数の命令はさらに、前記１つまたは複数のプロセッサに前記受信したストリーミングメディアコンテンツをキャッシュさせることを特徴とする請求項１４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項１８】
サーバからストリーミングメディアコンテンツを受信することと、
ユーザからシャトルコントロールオプションを受信することと、
前記サーバに前記シャトルコントロールオプションを指示せずに、前記シャトルコントロールオプションに従って前記ストリーミングメディアコンテンツの再生を変更することと
を含むことを特徴とする方法。
【請求項１９】
前記シャトルコントロールオプションは停止コマンドを含むことを特徴とする請求項１８に記載の方法。
【請求項２０】
前記シャトルコントロールオプションは一時停止コマンドを含むことを特徴とする請求項１８に記載の方法。
【請求項２１】
前記シャトルコントロールオプションは巻き戻しコマンドを含むことを特徴とする請求項１８に記載の方法。
【請求項２２】
前記シャトルコントロールオプションは早送りコマンドを含むことを特徴とする請求項１

８に記載の方法。
【請求項２３】
前記サーバから受信すると前記ストリーミングメディアコンテンツをキャッシュすることをさらに含むことを特徴とする請求項１８に記載の方法。
【請求項２４】
クライアントデバイスの１つまたは複数のプロセッサによって実行されると、前記１つまたは複数のプロセッサに、
サーバデバイスからストリーミングメディアコンテンツを受信させ、
前記受信したストリーミングメディアコンテンツを前記クライアントデバイスの不揮発性記憶装置に保存させる
複数の命令が格納されていることを特徴とする１つまたは複数のコンピュータ可読媒体。
【請求項２５】
前記サーバデバイスはプロキシサーバを含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項２６】
前記不揮発性記憶装置はハードディスクドライブを含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項２７】
前記不揮発性記憶装置は、光ディスクおよび光ディスクドライブを含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項２８】
前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、コンテンツをキャッシュすることが許される旨の指示を前記サーバデバイスから受信した場合にのみ、前記１つまたは複数のプロセッサに、その受信したストリーミングメディアコンテンツを保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項２９】
前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、ストリーミングメディアコンテンツのキャッシングが前記サーバデバイスによってサポートされることを示すサポートヘッダを含むメッセージを前記サーバデバイスから受信した場合にのみ、前記１つまたは複数のプロセッサに、その受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項３０】
前記サポートヘッダは、ストリーミングメディアコンテンツのクライアントでのキャッシングが前記サーバデバイスによってサポートされることを示す機能名を含み、前記機能名は、
com.microsoft.wm.fastcache
を含むことを特徴とする請求項２９に記載の１つまたは複数のコンピュータ可読媒体。
【請求項３１】
前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、キャッシュ制御指示に関連のストリーミングメディアコンテンツをキャッシュすることが前記クライアントデバイスに許されることを示すキャッシュ制御指示を含むメッセージを前記サーバデバイスから受信した場合にのみ、前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。
【請求項３２】
前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを

前記不揮発性記憶装置に保存させる前記命令は、前記ストリーミングメディアコンテンツのキャッシュが許されることが、前記ストリーミングメディアコンテンツの識別子によって示されている場合にのみ、前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３３】

前記識別子はＵＲＬ（Ｕｎｉｆｏｒｍ　ＲｅｓｏｕｒｃｅＬｏｃａｔｏｒ）を含むことを特徴とする請求項３２に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３４】

前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、ユーザがストリーミングメディアコンテンツのクライアントでのキャッシングを使用禁止にしていない場合にのみ、前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３５】

前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、前記クライアントデバイスのメディアプレーヤがストリーミングメディアコンテンツのクライアントでのキャッシングを使用禁止にしていない場合にのみ、前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３６】

前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる前記命令は、前記ストリーミングメディアコンテンツのタイプがキャッシュ可能タイプである場合にのみ、前記１つまたは複数のプロセッサに、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存させる命令を含むことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３７】

前記複数の命令はさらに前記１つまたは複数のプロセッサに、
前記ストリーミングメディアコンテンツに対する事後の要求を受信させ、
前記事後の要求に応答して、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置から取得させる
ことを特徴とする請求項２４に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３８】

前記複数の命令はさらに前記１つまたは複数のプロセッサに、
前記ストリーミングメディアコンテンツの有効期限が切れているかどうかを調べさせ、
前記ストリーミングメディアコンテンツの有効期限が切れていない場合にのみ、前記事後の要求に応答して、前記不揮発性記憶装置から前記ストリーミングメディアコンテンツを取得させる
ことを特徴とする請求項３７に記載の１つまたは複数のコンピュータ可読媒体。

【請求項３９】

前記複数の命令はさらに前記１つまたは複数のプロセッサに、
前記ストリーミングメディアコンテンツの有効期限が切れているかどうかを調べさせ、
前記ストリーミングメディアコンテンツの有効期限が切れている場合は、前記ストリーミングメディアコンテンツを再確認させる
ことを特徴とする請求項３７に記載の１つまたは複数のコンピュータ可読媒体。

【請求項４０】

前記複数の命令はさらに前記１つまたは複数のプロセッサに、

前記不揮発性記憶装置からの前記ストリーミングメディアコンテンツの再生に係わるロギング情報を含むメッセージを生成させ、前記メッセージは、データ構造のメッセージ本体部分に前記ロギングが含まれることを示すロギング情報ヘッダを含むことを特徴とする請求項３７に記載の１つまたは複数のコンピュータ可読媒体。

【請求項４１】

前記ロギング情報ヘッダは、

Ｃｏｎｔｅｎｔ－Ｔｙｐｅ：ａｐｐｌｉｃａｔｉｏｎ／ｘ－ｗｍｓ－ｓｅｎｄｅｖｅｎｔ

を含むことを特徴とする請求項４０に記載の１つまたは複数のコンピュータ可読媒体。

【請求項４２】

不揮発性記憶装置と、

システム上で再生するためにストリーミングメディアコンテンツをサーバデバイスから受信するように構成され、さらに、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置に記憶するように構成されたストリーミングメディアプレーヤコンポーネントとを備えることを特徴とするシステム。

【請求項４３】

前記ストリーミングメディアプレーヤコンポーネントはさらに、ストリーミングメディアコンテンツのキャッシングが前記サーバデバイスによってサポートされることを示すサポートヘッダを含むメッセージを前記サーバデバイスから受信した場合にのみ、前記受信したストリーミングメディアコンテンツを前記不揮発性記憶装置に保存するように構成されることを特徴とする請求項４２に記載のシステム。

【請求項４４】

前記サポートヘッダは、ストリーミングメディアコンテンツのクライアントでのキャッシングが前記サーバデバイスによってサポートされることを示す機能名を含み、前記機能名は、

ｃｏｍ．ｍｉｃｒｏｓｏｆｔ．ｗｍ．ｆａｓｔｃａｃｈｅ

を含むことを特徴とする請求項４３に記載のシステム。

【請求項４５】

前記ストリーミングメディアプレーヤコンポーネントはさらに、前記コンテンツのキャッシングが許されることを示す指示を前記サーバデバイスから受信した場合にのみ、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置に記憶するように構成されることを特徴とする請求項４２に記載のシステム。

【請求項４６】

前記ストリーミングメディアプレーヤコンポーネントはさらに、キャッシュ制御指示に関連のストリーミングメディアコンテンツをキャッシュすることが前記システムに許されることを示すキャッシュ制御指示を含むメッセージを前記サーバデバイスから受信した場合にのみ、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置に記憶するように構成されることを特徴とする請求項４２に記載のシステム。

【請求項４７】

前記ストリーミングメディアプレーヤコンポーネントはさらに、前記ストリーミングメディアコンテンツのキャッシングが許されることが、前記ストリーミングメディアコンテンツの識別子によって示されている場合にのみ、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置に記憶するように構成されることを特徴とする請求項４２に記載のシステム。

【請求項４８】

前記識別子はＵＲＬ（Ｕｎｉｆｏｒｍ　Ｒｅｓｏｕｒｃｅ　Ｌｏｃａｔｏｒ）を含むことを特徴とする請求項４７に記載のシステム。

【請求項４９】

前記ストリーミングメディアプレーヤコンポーネントはさらに、ユーザがストリーミングメディアコンテンツのクライアントでのキャッシングを使用禁止にしていない場合にのみ、前記ストリーミングメディアコンテンツを前記不揮発性記憶装置に記憶するように構成