

AUUGN

Australian Unix systems

User Group Newsletter

Volume 8

Number 5

Includes Winter 1987 Conference Papers

Registered by Australia Post Publication No. NBQ6624

Amazon Ex. 1029
IPR Petition - USP 7,784,058

The Australian UNIX* systems User Group Newsletter

Volume 8 Number 5

October 1987

CONTENTS

AUUG General Information	3
Editorial	4
President's Report	5
Softway Advertisement	6
Adelaide UNIX Users Group Information	7
Winter 1987 Conference Papers	8
UNIX at the Turn of the Century - Abstract	10
The Locking of Critical Regions Under UNIX	11
Fun with Virtual Memory	20
Recent Work on Research UNIX	26
The Shared Library Minefield	27
SunOS Release 4.0	36
Writing Parallel Programs for the Sequent Multiprocessor - Abstract	44
UNIX on the Cray - Abstract	44
A Low Cost, Short Range, Reconfigurable Microwave Data Link	54
Some Aspects of System V Release 3 Networking	75
Optimizing C - Benchmarks and Real Work - Abstract	75
Measuring Database Performance using TPI benchmark - Abstract	76
DBMS - Efficient Implementations for UNIX systems - Abstract	77
What's in a name? (or coping with lots of small files)	78
An Image of the Future - Abstract	84
Plexus XDP System - Background Information	85
Awk-ward yacc ::= lex	90
Mail links from VMS to UNIX	94
The SUN Network File System: Design, Implementation and Experience	96
Shared Libraries in SunOS	112
Virtual Memory Architecture in SunOS	128
Virtual Address Cache in UNIX	143

From the <i>login</i> : Newsletter - Volume 12 Number 5	151
Computer Graphics Workshop	152
POSIX Portability Workshop	152
C++ Workshop	153
Call for Papers: Winter 1988 USENIX Conference	154
Call for Papers: Summer 1988 USENIX Conference	155
Second Distribution of Berkeley PDP-11 Software for UNIX	156
RT PC Distributed Service: File Systems	159
Book Review: UNIX System Security	170
Book Review: troff typesetting for UNIX Systems	172
Work-in-Progress Reports from the Phoenix Conference	174
Management Committee Meeting Minutes - May 1987	180
Management Committee Meeting Minutes - August 1987	184
Annual General Meeting Minutes - August 1987	188
AUUG Membership Categories	191
AUUG Forms	193

Copyright © 1987. AUUGN is the journal of the Australian UNIX* systems User Group. Copying without fee is permitted provided that copies are not made or distributed for commercial advantage and credit to the source must be given. Abstracting with credit is permitted. No other reproduction is permitted without prior permission of the Australian UNIX systems User Group.

* UNIX is a registered trademark of AT&T in the USA and other countries.

SunOS Release 4.0

Rich Burridge

Sun Microsystems Australia,
123 Camberwell Rd,
Hawthorn, VIC 3122.
richb@sunk.oz

ABSTRACT

SunOS Release 4.0 is targeted for availability in the first quarter of calendar year 1988. This paper summarises the planned content. It is being issued early in the recognition of our customers' needs for advance information to accommodate their product planning efforts. All contents and schedules are subject to change without notice. Sun will, to the best of its ability publicise any such changes.

A. Positioning Overview

SunOS Release 4.0 follows the Sun tradition of:

- o promoting a single UNIX standard converging System V and BSD features,
- o providing maximum functionality,
- o incorporating the latest state-of-the-art operating system technology, and
- o supporting computing in heterogeneous environments

This release of SunOS includes many enhancements for the installed base:

- o extended converged UNIX functionality,
- o addressing key systems administration needs, and
- o creating a foundation for the anticipated hardware architectures

The Sun UNIX kernel has been significantly redesigned as another step in the on-going process of incorporating state-of-the-art technology to establish a stable architectural base for the next decade. This base includes work to incorporate the latest virtual memory management techniques, shared libraries, and improved system security

While this type of implementation work is essential to maintain the leading performance and functionality of SunOS, Sun is committed to retaining industry standard interface compatibility. In this manner, implementation changes can be accommodated while maximising code portability to this and future SunOS releases.

B. Contents Overview.

The following is a partial list of the new features in 4.0:

Support for Sun-4

Release 4.0 supports the new Sun-4 product family workstations. This involved changes to the kernel, as well as to the compilers.

New Architectural Features

Features:

Description or comments:

- | | |
|---|---|
| 1. Major New Kernel Architecture | <p>The kernel has been significantly restructured to accomodate a new virtual memory management scheme that promotes sharing system resources. Swap space requirements are reduced, system resources are utilised more efficiently, cacheing of frequently accessed data is more efficient, and files can be treated as part of virtual memory making access to large files more convenient and efficient.</p> <p>The new VM system provides sharing on a page-by-page basis and employs a copy-on-write mechanism to create individual copies of pages as they become necessary.</p> <p>Swap area is now a regular file sharing disk resources with other files.</p> |
| 2. Shared Library Facility | <p>Using shared libraries reduces program size and swap space requirements and provides the flexibility to incorporate new versions of libraries automatically, thus facilitating bringing new libraries into system-wide use. <i>libc</i> and user built libraries are sharable in this release and future releases will offer more system libraries converted to shareable format.</p> <p>Shared libraries utilise the new virtual memory system and a revised link editor (<i>ld</i>). The C compiler has been enhanced to generate position independent code (<i>pic</i>) and <i>ld</i> supports either linking code dynamically at run-time or statically at initial program load.</p> <p>Note: Sharing and its associated dynamic binding is the current planned default operation although a small performance penalty may result from run-time linking.</p> |
| 3. Resizable swap area for diskless clients | <p>Resizing client swap space no longer requires taking a server and all its clients off-line or reinstalling the system.</p> <p>Only the client(s) being modified needs to be suspended.</p> |
| 4. Lightweight Processes | <p>Provides the ability for a C program to have multiple concurrent threads of execution. Lightweight processes facilitate developing programs that manage asynchronous events.</p> <p>Facilities provided include messages, monitors, exception handling, and flexible context switching. Both coroutines and preemptive scheduling are possible.</p> |

4.0 provides lightweight processes only at the user application level. Kernel support is not provided.

5. Other Kernel Enhancements
a.

These include:

System V shared memory implementation now supports paging of shared data space.

b.

Network interface tap (NIT) improvements include packet filtering and integration with the streams environment. Packet filtering provides an Etherfind-like capability allowing applications to monitor ethernet traffic selecting, for example, ethernet packets from a certain host or with a specific field setting.

Networking

1. Improved support for diskless clients

Diskless clients are supported via the NFS file system. *nd* and its associated partitions no longer exist. This improves heterogeneity support (non-Sun servers can now support diskless Sun clients) and improves administration since swap area is now a regular file. (Note: Being a regular file also means that client swap areas shares disk resources with other applications running on the client and with other clients.)

The performance goal is to achieve Release 3.2 performance levels in the absence of *nd*. Eliminating *nd* does NOT impact the number of clients a server can handle.

A boot block server will be provided to support Sun-2s with prompts that don't support tftp booting, so no prom change is required by this change.

2. Filesystem re-org

The re-organized filesystem facilitates maintenance of servers by permitting more sharing of binaries, matching diskfull and diskless file layouts, easier mixing of multiple architectures, and easier access to remote executables in general. A single server can more easily support multiple architecture clients and non-Sun servers can support diskless Suns.

The *root* filesystem and */usr/lib* have been re-organised. Shareable files, including architecture dependent files, are now all in */usr/lib* and client private data is all in *root*.

4. YP Enhancements

The YP databases can be updated over the network. This simplifies system administration and, depending on local administrative policy, the system administrator need no longer be solely responsible for incorporating YP map updates.

5. Secure networking features

Secure networking improvements provide more thorough authentication of user identification prior to allowing file access. Security is provided by supporting the exchange of conversation keys and by preventing superusers from masquerading as other users to access secure filesystems across the network.

NFS utilizes secure RPC so that servers can optionally become more secure. Secure operation is selected by specifying the *-secure* option for individual */etc/exports* entries. For secure operation RPC utilises a YP database of public and private conversation keys.

6. NFS performance

NFS benefits significantly from the improved cacheing improvements provided by the new virtual memory system.

Standards

1. System V enhancements
a.

New features include:

Remaining system calls required for Base Level Release 2 functionality: *chown*, *creat*, *fcntl*, *kill*, *mknod*, *open*, and *utime*.

b.

Complete System V *STREAMS* interface incorporated in the kernel. Using *STREAMS* simplifies writing device drivers by supporting portable protocol modules.

c.

Fully System V and BSD compatible tty interface using *STREAMS*.

d.

System V compatible archive utility (*ar*).

e.

System V batch utility and job scheduler (*at*, *batch*, *cron* and *crontab*).

f.

Access to Sun added-value libraries (e.g., SunView) from System V programs.

2. More 4.3BSD

Most 4.3BSD functionality is provided, including the new upper limit of 64 (vs. 30) on the number of open files per process and a fully 4.3BSD compatible subnet facility.

In general, 4.3BSD provided many bug fixes and performance enhancements. Most of those which are applicable to Sun Workstations and which were not already in SunOS prior to the availability of 4.3BSD or added since then are provided by SunOS Release 4.0. (Note: Sun Consulting has developed an implementation of XNS that includes the Transport and Network layers on top of standard ethernet.)

Peripherals

1. New *setup* utility

The new *setup* utility provides convenience and flexibility. New capabilities include the ability to reuse configuration files for common configurations, to edit existing configuration files to correct errors or support minor configuration differences, and for sophisticated users to specify variants to Sun supplied configuration alternatives.

setup runs off a table interface much like the 3.x terminal version of *setup*; there is no window version.

- | | | |
|----|-----------------------------------|---|
| 2. | On-line disk formatting utility | The disk formatting utility now runs under UNIX and can be used to format any drive not currently mounted. Formatting disks is now much faster and no longer requires a dedicated system. Multiple disks can be formatted simultaneously. |
| 3. | New Mass Storage Systems | Support for as yet to be announced mass-storage systems is provided. |
| 4. | Removal of Interphase 2180 driver | The driver for the Interphase 2180 Disk Controller is no longer included in the SunOS. The removal of obsolete drivers is part of the on-going program to unburden the standard system software from supporting an ever growing number of devices. As drivers become obsolete, the pending removal at the next release is announced. Systems (a few Sun-2/100Us and Sun-2/150Us) with Interphase 2180 Controllers must be upgraded before they can run Release 4.0. |

Other Enhancements

- | | | |
|----|----------------------|---|
| 1. | Secure System Work | <p>Optional security measures provide an auditing log of failed and/or successful attempts to use system resources on both a system wide or individual user basis and identifying who was using the system at the time of attempted security breaches. Improved protection of the password database is also provided.</p> <p>In DoD parlance, Release 4.0 provides "unevaluated C2" level security. This is the groundwork for a future secure system product.</p> |
| 2. | C Global Optimizer | Significant run-time performance for C routines is now possible via the C compiler optimization option which applies the same global optimizing technology that is available in Sun's FORTRAN compiler. |
| 3. | Internationalization | 8-bit integrity will be assured at this release in order to support non-ascii character sets. This is the first step towards an internationalized UNIX that supports local character sets. The 8-bit kernel and tty driver support creating file names with 8-bit characters from within C programs. However, utilities such as <i>sunttools</i> , <i>csh</i> , <i>sh</i> , <i>vi</i> and other tools do not incorporate this 8-bit filenames at this release. |
| 4. | Documentation | <p>New manuals, and improved introductory guides and system administration guides ease the tasks of mastering and administering the system. New documents include: <i>System Services Overview</i> and a Global Index to all system manuals. Documents with major revisions include:</p> <p><i>Installing UNIX and Programming Utilities.</i></p> <p>Repackaging of the documentation offers the convenience of selecting the set of manuals targeted to the specific users needs, e.g. introduction to system use, system reference, system administration, and program development.</p> |

5. **Obsolescence
Mechanism**

Instigation of */usr/old* directory and movement of obsolete modules to that directory. This introduces a warning system to prepare users that the included modules will be removed in the next release. For Release 4.0, the following programs will be moved to */usr/old*:

<i>sun3cvt</i>	Was only needed for 2.0 to 3.0 transition.
<i>compact</i>	Replaced by faster and more efficient but incompatible <i>compress</i> program from 4.3BSD.
<i>eyacc</i>	Was used only to implement Pascal. Has been removed in 4.3BSD.
<i>make</i>	Pre 3.4 version of <i>make</i> .
<i>prmail</i>	Replaced with <i>mail -u</i> by 4.3BSD.
<i>pti</i>	Replaced with <i>troff -a</i> by 4.3BSD

C. Release Specifications

Performance

Release 4.0 is expected to achieve performance levels at least equivalent to that of Release 3.2.

Resource Requirements

Release 4.0 resource requirements (memory and disk) for the system software are expected to be the same as for Release 3.2. Shared libraries have contributed to making the basic resource requirements of Release 4.0 comparable to those of 3.2 in the presence of added capability. The full set of system software and associated files (including text files such as man pages and font files which do not benefit from the space savings provided by shared libraries) will likely require more disk space than consumed by Release 3.x files.

Code Migration and System Administration Changes

Maximizing binary compatibility is a major design objective of Release 4.0. To date no known widespread sources of binary incompatibility for well-written code exist. As incompatibilities arise they will be carefully documented. Source code changes may be required in some binary compatible programs when they are recompiled. The following lists key known compatibility issues and areas subject to visible impact on administrative or program-build procedures:

Features:

Compatibility Issues & Procedural Changes:

New Architectural Features

- | | |
|--------------------------------------|--|
| 1. Major New Kernel Architecture | Programs that read the structure or depend on the format of core files may require revisions. |
| 2. Shared Library Facility | Build procedures for libraries may need to change. Existing programs will work but will require rebuilding to take advantage of shared libraries. Build procedures for programs requiring a non-shared library must explicitly select this option. |
| 3. Resizable swap area | New, easier administrative procedures required. |
| 4. NIT Improvements | Programs utilizing NIT will require source changes and recompilation. |

Networking

- | | | |
|----|---------------------------------------|---|
| 1. | Improved support for diskless clients | New, easier administrative procedures required |
| 2. | Filesystem re-org | New, easier administrative procedures required. |
| 3. | YP Enhancements | Potential changes to administrative procedures. |
| 4. | Secure networking | Pre-4.0 (insecure) yppasswd may not work against a secure 4.0. yp yp server. Secure authentication will prohibit access to setuid programs lacking authorised access. |

Standards

- | | | |
|----|-----------------------------|---|
| 1. | System V enhancements: | |
| | SVID Compliance | Programs using <i>creat</i> , <i>fcntl</i> , <i>kill</i> , <i>mknod</i> , <i>open</i> , and <i>utime</i> may have to be relinked. |
| | System V/BSD tty | Programs using System V <i>termio ioctl</i> calls will perform better if they are relinked. |
| | Batch utility/job scheduler | Slight impact on system administration procedures only. |
| 2. | System V shared memory | Programs using the System V shared memory <i>shmdt</i> library call that were linked on a Release earlier than Release 3.4. should be relinked to avoid a spurious message to the console window. |

Peripherals

- | | | |
|----|-----------------------------------|---|
| 1. | New setup utility | New, easier administrative procedures required. |
| 2. | On-line disk formatting utility | New, easier administrative procedures required. |
| 3. | Removal of Interphase 2180 driver | Systems using Interphase 2180 must be upgraded. |

Other Enhancements

- | | | |
|----|--------------------|--|
| 1. | Secure System Work | Modules using the encrypted password field in <i>/etc/passwd</i> or <i>/etc/group</i> will require modification. |
|----|--------------------|--|

Installation

Release 4.0 is a major release introducing extensive architectural changes. As such, it requires a full system installation. With the new setup utility, system installation and configuration is, however, much more convenient.

D. Acknowledgements.

This paper is mainly based on an internal pre-release report written by Pat Harding, plus a presentation given to Sun personnel at Palm Springs in late July.

E. References

- [GING 87] Gingell, R. A., M. Lee, X. T. Dang, M. S. Weeks, "Shared Libraries in SunOS", Summer Conference Proceedings, Phoenix 1987, USENIX Association, 1987.

- [GING 87] Gingell, R. A., J. P. Moran, W. A. Shannon, "Virtual Memory Architecture in SunOS", Summer Conference Proceedings, Phoenix 1987, USENIX Association, 1987.

- [CHEN 87] Cheng, R., "Virtual Cache in UNIX", Summer Conference Proceedings, Phoenix 1987, USENIX Association, 1987.