Bug fixes and changes in 4.2BSD

July 28, 1983

Samuel J. Leffler

Computer Systems Research Group
Department of Electrical Engineering and Computer Science
University of California, Berkeley
Berkeley, California 94720
(415) 642-7780

ABSTRACT

This document describes briefly the changes in the Berkeley system for the VAX between the 4.1BSD distribution of April 1981 and this, its revision of July 1983. It attempts to summarize, without going into great detail, the changes which have been made.

Notable improvements

- The file system organization has been redesigned to provide at least an order of magnitude improvement in disk bandwidth.
- The system now provides full support for the DOD Standard TCP/IP network communication protocols. This support has been integrated into the system in a manner which allows the development and concurrent use of other communication protocols. Hardware support and routing have been isolated from the protocols to allow sharing between varying network architectures. Software support is provided for 10 different hardware devices including 3 different 10 Mb/s Ethernet modules.
- A new set of interprocess communication facilities has replaced the old multiplexed file mechanism. These new facilities allow unrelated processes to exchange messages in either a connection-oriented or connection-less manner. The interprocess communication facilities have been integrated with the networking facilities (described above) to provide a single user interface which may be used in constructing applications which operate on one or more machines.
- A new signal package which closely models the hardware interrupt facilities found on the VAX replaces the old signals and jobs library of 4.1BSD. The new signal package provides for automatic masking of signals, sophisticated signal stack management, and reliable protection of critical regions.
- File names are now almost arbitrary length (up to 255 characters) and a new file type, symbolic link, has been added. Symbolic links provide a ‘‘symbolic referencing’’ mechanism similar to that found in Multics. They are interpolated during pathname expansion and allow users to create links to files and directories which span file systems.
- The system supports advisory locking on files. Files can have ‘‘shared’’ or ‘‘exclusive’’ locks applied by processes. Multiple processes may apply shared locks, but only one process at any time may have an exclusive lock on a file. Further, when an exclusive lock is present on a file, shared locks are disallowed. Locking requests normally block a process until they can be completed, or they may be indicated as ‘‘non-blocking’’ in which case an error is returned if the lock can not be immediately obtained.
- The group identifier notion has been extended to a ‘‘group set’’. When users log in to the system

they are placed in all their groups. Access control is now done based on the group set rather than just a single group id. This has obviated the need for the newgrp command.

- Per-user, per-filesystem disk quotas are now part of the system. Soft and hard limits may be specified on a per user and per filesystem basis to control the number of files and amount of disk space allocated to a user. Users who exceed a soft limit are warned and if, after three login sessions, their disk usage has not dropped below the soft limit, their soft limit is treated as a hard limit. Utilities exist for the creation, maintenance, and reporting of disk quotas.
- System time is now available in microsecond precision and millisecond accuracy. Users are provided with 3 high-resolution timers which may be set up to automatically reload on expiration. The timers operate in real time, user time, and process virtual time (for profiling). All statistics and times returned to users are now given in a standard format with seconds and microseconds separated. This eliminates program dependence on the line clock frequency.
- A new system call to rename files in the same file system has been added. This call eliminates many of the anomalies which could occur in older versions of the system due to lack of atomicity in removing and renaming files.
- A new system call to truncate files to a specific length has been added. This call improves the performance of the Fortran I/O library.
- Swap space configuration has been improved by allowing multiple swap partition of varying sizes to be interleaved. These partitions are sized at boot time to minimize configuration dependencies.
- The Fortran 77 compiler and associated I/O library have undergone extensive changes to improve reliability and performance. Compilation may, optionally, include optimization phases to improve code density and decrease execution time.
- A new symbolic debugger, dbx, replaces the old symbolic debugger sdb. Dbx works on both C and Fortran 77 programs and allows users to set break points and trace execution by source code line numbers, references to memory locations, procedure entry, etc. Dbx allows users to reference structured and local variables using the program's programming language syntax.
- The delivermail program has been replaced by sendmail. Sendmail provides full internetwork routing, domain style naming as defined in the DARPA Request For Comments document #833, and eliminates the compiled in configuration database previously used by delivermail. Further, sendmail uses the DARPA standard Simple Mail Transfer Protocol (SMTP) for mail delivery.
- The system contains a new line printer system. Multiple line printers and spooling queues are supported through a printer database file. Printers on serial lines, raster printing devices, and laser printers are supported through a series of filter programs which interface to the standard line printer ‘‘core programs’’. A line printer control program, lpc, allows printers and printer queues to be manipulated. Spooling to remote printers is supported in a transparent fashion.
- Cu has been replaced by a new program tip. Tip supports a number of auto-call units and allows destination sites to be specified by name rather than phone number. Tip also supports file transfer to non-UNIX machines and can be used with sites which require half-duplex and/or odd-even parity.
- Uucp now supports many auto-call units other than the DN11. Spooling has been reorganized into multiple directories to cut down on system overhead. Several new utilities and shell scripts exist for use in adminstrating uucp traffic. Operation has been greatly improved by numerous bug fixes.

Bug fixes and changes

Section 1

adb Support has been added for interpreting kernel data structures on a running system and in post mortem crash dumps created by savecore. A **−k** option causes adb to map addresses according to the system and current process page tables. A new command, $p, can be used to switch between process contexts. Many scripts are available for symbolically displaying kernel data structures, searching for a process' context by process ID, etc. A new document, ''Using ADB to Debug the UNIX Kernel'', supplies hints in the use of adb with system crash dumps.

addbib Is a new utility for creating and extending bibliographic data bases for use with refer.

apply Is a new program which may be used to apply a command to a set of arguments.

ar Has a new key, 'o', for preserving a file's modification time when it is extracted from an archive.

cc Supports the additional symbol information used by dbx. The old symbol information, used by the defunct sdb debugger, is available by specifying the **−go** flag. A new flag, **−pg**, creates executable programs which collect profiling information to be interpreted by the new gprof program. A bug in the C preprocessor, which caused line numbers to be incorrect for macros with formal parameters with embedded newlines has been fixed. The C preprocessor now properly handles hexadecimal constants in ''#if'' constructs and checks for missing ''#endif'' statements.

chfn Now works interactively in changing a user's information field in the password file.

chgrp Is now in section 1 and may be executed by anyone. Users other than the super-user may change group ownership of a file they own to any group in their group access list.

cp Now has a **−r** flag to copy recursively down a file system tree.

csh A bug which caused backquoted commands to wedge the terminal when interrupted has been fixed. Job identifiers are now globbed. A bug which caused the ''wait'' command to uninteruptible in certain cases has been fixed. History may now be saved and restored across terminal sessions with the *savehist* variable. The newgrp command has been deleted due to the new group facilities.

ctags Now handles C **typedefs**.

cu Exists only in the form of a ''compatible front-end'' to the new tip program.

dbx Is a new symbolic debugger replacing sdb. Dbx handles C and Fortran programs.

delivermail
Has been replaced by the new sendmail program.

df Understands the new file system organization and reports all disk space totals in kilobytes.

du Now reports disk usage in kilobytes and uses the new field in the inode structure which contains the actual number of blocks allocated to a file to increase accuracy of calculations.

dump Has been moved to section 8.

error Has been taught about the error message formats of troff.

eyacc A bug which caused the generated parser to not recognize valid null statements has been fixed.

f77 Has undergone major changes.

The i/o library has been extensively tested and debugged. Sequential files are now opened at the **BEGINNING** by default; previously they were opened at the end.

Compilation of data statements has been substantially sped up. Significant new optimization is optionally available (this is still a bit buggy and should be used with caution). Even without optimization, however, single precision computations execute much faster.

The new debugger, dbx, has replaced sdb for debugging Fortran programs; sdb is no longer supported.

Files with ‘‘.F’’ suffixes are preprocessed by the C preprocessor. This allows C-style ‘‘#include’’ and ‘‘#define’’ constructs to be used. The compiler has been modified to print error messages with sensible line numbers. Make also understands the ‘‘.F’’ suffix. Note that when using the C preprocessor, the 72 column convention is not followed.

The −**I** option for specifying short integers has been changed to −**i**. The −**I** option is now used to specify directory search paths for ‘‘#include’’ statements. A −**pg** option for creating executable images which collect profiling information for gprof has been added.

fed Is a font editor of dubious value.

file Now understands symbolic links.

find Has a new −**type** value, ‘l’, for finding symbolic links.

fp Is a new compiler/interpreter for the Functional Programming language. A supporting document is present in Volume 2C of the UNIX Programmer’s Manual.

fpr Is a new program for printing Fortran files with embedded Fortran carriage controls.

fsplit Is a new program for splitting a multi-function Fortran file into individual files.

ftp Is a new program which supports the ARPA standard File Transfer Protocol.

gcore Is a new program which creates a core dump of a running process.

gprof Is a new profiling tool which displays execution time for the dynamic call graph of a program. Gprof works on C, Fortran, and Pascal programs compiled with the −**pg** option. Gprof may also be used in creating a call graph profile for the operating system. A supporting document, ‘‘gprof: A Call Graph Execution Profiler" is included in Volume 2C of the UNIX Programmer’s Manual.

groups Is a new program which displays a user’s group access list.

hostid Is a new program which displays the system’s unique identifier as returned by the new gethostid system call. The super-user uses this program to set the host identifier at boot time.

hostname Is a new program which displays the system’s name as returned by the new gethostname system call. The super-user uses this program to set the host name at boot time.

indent Is a new program for formatting C program source.

install Is a shell script used in installing software.

iostat Now reports kilobytes per second transferred for each disk. This is useful as the unit of information transferred is no longer a constant one kilobytes.

last Now displays the remote host from which a user logged in (when accessing a machine across a network). The pseudo user ‘‘ftp’’ may be specified to find out information about FTP file transfer sessions.

lastcomm Now displays flags for each command indicating if the program dumped core, used PDP-11 mode, executed with a set-user-ID, or was created as the result of a fork (with no following exec).

learn Now has lessons for vi (this is user contributed software which is not part of the standard system).

lint Has a new −**C** flag for creating lint libraries from C source code. Has improved type checking on static variables.

lisp Has been ported to several 68000 UNIX systems, the relevant code is included in the distribution. A new vector data type and a form of ‘‘closure’’ have been added.

ln Has a new flag, −**s**, for creating symbolic links.

login Has been extensively modified for use with the rlogind and telnetd network servers.

lpq Is totally new, see lpr.

lpr And its related programs are totally new. The line printer system supports multiple printers of many different characteristics. A master data base, /etc/printcap, describes both local printers and printers accessable across a network. A document describing the line printer system is now part of Volume 2C of the UNIX Programmer's Manual.

lprm Is totally new, see lpr.

ls Has been rewritten for the new directory format. It understands symbolic links and uses the new inode field which contains the actual number of blocks allocated to a file when the **–s** flag is supplied. Many rarely used options have been deleted.

m4 A bug which caused m4 to dump core when keywords were undefined then redefined has been fixed.

Mail Now supports mail folders in the style of the Rand MH system. Has been reworked to cooperate with sendmail in understanding the new mail address formats. Allows users to defined message header fields which are not be displayed when a messages is viewed. Many other changes are described in a revised version of the user manual.

make Understands not to unlink directories when interrupted. Understands the new ‘‘.F’’ suffix for Fortran source files which are to be run through the C preprocessor. Has a new predefined macro MFLAGS which contains the flags supplied to make on the command line (useful in creating hierarchies of makefiles).

mkdir Now uses the mkdir system call to run faster.

mt Has a new command, status, which shows the current state of a tape drive.

mv Has been rewritten to use the new rename system call. As a result, multiple directories may now be moved in a single command, the restrictions on having ‘‘..’’ in a pathname are no longer present, and everything runs faster.

net And all related Berknet programs are no longer part of the standard distribution. These programs live on in /usr/src/old for those who can not do without them.

netstat Is a new program which displays network statistics and active connections.

oldcsh No longer exists.

od Has gobs of new formats options.

pagesize Is a new program which prints the system page size for use in constructing portable shell scripts.

passwd Now reliably interlocks with chsh, chfn, and vipw, in guarding against concurrent updates to the password file.

pc/pi **For** loops are now done according to the standard. Files may now be dynamically allocated and disposed. Records and variant records are now aligned to correspond to C structures and unions (this was falsely claimed before). Several obscure bugs involving formal routines have been fixed. Three new library routines support random access file i/o, see /usr/include/pascal for details.

pc **For** loop variables and **with** pointers are now allocated to registers. Separate compilation type checking can now be done without reference to the source file; this permits movement (including distribution) of .o files and creation of libraries. Display entries are saved only when needed (a speed optimization).

pdx Is a new debugger for use with pi. Pdx is invoked automatically by the interpreter if a runtime error is encountered. Future work is planned to extend the new dbx debugger to understand code generated by the Pascal compiler pc.

ps Has been changed to work with the new kernel and is no longer dependent on system page size. All process segment sizes are now shown in kilobytes. Understands that the old ‘‘using new signal facilities’’ bit in the process structure now means ‘‘using old 4.1BSD signal facilities’’.

pwd Now simply calls the *getwd* (3) routine.

rcp Is a new program for copying files across a network. The complete syntax of cp is supported, including recursive directory copying.

refer Has had many bugs fixed in it and the associated −ms macro package support made to work.

reset Now resets all the special characters to the system defaults specified in the include file <sys/ttychars.h>.

rlogin Is a new program for logging in to a machine across a network. Rlogin uses the files /etc/hosts.equiv and .rhosts in the users login directory to allow logins to be performed without a password. Rlogin supports proper handling of ˆS/ˆQ and flushing of output when an interrupt is typed at the terminal. Its ‘˜’ escape sequences are reminiscent of the old cu program (as it is based on the same source code).

rmdir Now uses the rmdir system call to run more efficiently and not require root privileges. Unfortunately, this means arguments which end in one or more ‘‘/’’ characters are no longer legal.

roffbib Is a new program for running off bibiliographic databases.

rsh Is a new program which supports remote command execution across a network.

ruptime Is a new program which displays system status information for clusters of machines attached to a local area network.

rwho Is a new program which displays users logged in on clusters of machines attached to a local area network.

script Has been rewritten to use pseudo-terminals. This allows the C shell job control facilities (among other things) to be used while scripting. A side effect of this change is that scripts now contain everything typed at a terminal.

sdb Has been replaced by dbx; it still lives on in /usr/src/old for those with a personal attachment.

sendbug Is a new command for submitting bug reports on 4.2BSD in a standard format suitable for automatic filing by the bugfiler program.

sh No longer has a newgrp command due to the new groups facilities.

sortbib Is a new command for sorting bibliographic databases.

strip Has been made blindingly fast by using the new truncate system call (thereby eliminating the old method of copying the file).

stty The default system erase, kill, and interrupt characters have been made the DEC standard values of DEL (‘ˆ?’), ‘ˆU’, and ‘ˆC’. This is not expected to gain much popularity, but was done in the interest of compatibility with many other standard operating systems.

su Has been changed to do a ‘‘full login’’ when starting up the subshell. A new flag, **–f**, does a ‘‘fast’’ su for when a system is heavily loaded. Extra arguments supplied to su are now treated as a command line and executed directly instead of creating an interactive shell.

sysline Is a new program for maintaining system status information on terminals which support a ‘‘status line’’; a poor man’s alternative to a window manager (or emacs).

tail Has a larger buffer so that ‘‘tail –r’’ and similar show more.

talk Is a new program which provides a screen-oriented write facility. Users may be ‘‘talked to’’across a network, though satellite response times have indicated overseas conversations are still best done by phone. Can be very obnoxious when engaged in important work.

tar Now allocates its internal buffers dynamically so that the block size can be specified to be very large for streaming tape drives. Also, now avoids many core-core copy operations. Has a new −**C** option for forcing chdir operations in the middle of operation (thereby allowing multiple disjoint subtrees to be easily placed in a single file, each with short relative path-names). Has a new flag, ‘B’, for forcing 20 block records to be read and written; useful in joining two tar commands with a remote shell to transfer large amounts of data across a network.

telnet Is a new program which supports the ARPA standard Telnet protocol.

tip Replaces cu as the standard mechanism for connecting to machines across a phone line or through a hardwired connection. Tip uses a database of system descriptions, supports many different auto-call units, and understands many nuances required to talk to non-UNIX systems. Files may be transferred to and from non-UNIX systems in a simple fashion.

ul A bug which sometimes caused an extra blank line to be printed after reaching end of file has been fixed.

uucp And related programs have been extensively enhanced to support many different auto-call units and multiple spooling directories (among other things). A large number of bugs and performance enhancements have been made.

uusnap Is a new program which gives a snap-shot of the uucp spooling area.

vfontinfo Is a program used to inspect and print information about fonts.

vgrind Now uses a regular expression language to describe formatting. A **–f** flag forces vgrind to act as a filter, generating output suitable for inclusion in troff and/or nroff documents. Language descriptions exist for C, Pascal, Model, C shell, Bourne shell, Ratfor, and Icon programs.

vi A bug which caused the ˆB command to place the cursor on the wrong line has been fixed. A bug which caused vi to believe a file had been modified when an i/o error occurred has been fixed. A bug which allowed ‘‘hardtabs’’ to be set to 0 causing a divide by zero fault has been fixed.

vlp Is a new program for pretty printing Lisp programs.

vmstat Has had one new piece of information added to **–s** summary, the number of fast page reclaims performed. The fields related to paging activity are now all given in kilobytes.

vpr And associated programs for spooling and printing files on Varian and Versatec printers are now shell scripts which use the new line printer support.

vwidth Is a new program for making troff width tables for a font.

wc Is once again identical to the version 7 program. That is, the **–v**, **–t**, **–b**, **–s**, and **–u** flags have been deleted.

whereis Understands the new directory organization for the source code.

which Now understands how to handle aliases.

who Now displays the remote machine from which a user is logged in.

Section 2.

The most important change in section 2 is that the documentation has been significantly improved. Manual page entries now indicate the possible error codes which may be returned and how to interpret them. The introduction to section 2 now includes a glossary of terms used throughout the section. The terminology and formatting have been made consistent. Many manual pages now have ‘‘NOTES’’ or ‘‘CAVEATS’’ providing useful information heretofore left out for the sake of brevity. As always the manual pages are still for the programmer; they are terse and extremely concise. The ‘‘4.2BSD System Manual’’ is likewise concise, but a bit more verbose in providing an overall picture of the system facilities.

With regard to changes in the facilities, these fall into three major categories: interprocess communication, signals, and file system related calls. The interprocess communication facilities center around the *socket* mechanism described in the ‘‘A 4.2BSD Interprocess Communication Primer’’. The new signals do not have an accompanying document, so the manual pages should be studied carefully. The new file system calls pretty much stand on their own, with a late section of the document ‘‘A Fast File System for UNIX’’ supplying a quick overview of the most important new file system facilities. Finally, it should be noted that the job control facilities introduced in 4.1BSD have been adopted as a standard part of 4.2BSD.

No special distinction is given to these calls (in 4.1BSD they were earmarked ''2J'').

Many of the new system calls have both a ''set'' and a ''get'' form. Only the ''get'' forms are indicated below. Consult the manual for details on the ''set'' form.

intro Has been updated to reflect the new list of possible error codes. Now includes a glossary of terminology used in section 2.

access Now has symbolic definitions for the *mode* parameter defined in <*sys/file.h*>.

bind Is a new interprocess communication system call for binding names to sockets.

connect Is a new interprocess communication system call for establishing a connection between two sockets.

creat Has been obsoleted by the new *open* interface.

fchmod Is a new system call which does a *chmod* operation given a file descriptor; useful in conjunction with the new advisory locking facilities.

fchown Is a new system call which does a *chown* operation given a file descriptor; useful in conjunction with the new advisory locking facilities.

fcntl Is a new system call which is useful in controlling how i/o is performed on a file descriptor (non-blocking i/o, signal drive i/o). This interface is compatible with the System III fcntl interface.

flock Is a new system call for manipulating advisory locks on files. Locks may be shared or exclusive and locking operations may be indicated as being non-blocking, in which case a process is not blocked if the requested lock is currently in use.

fstat Now returns a larger stat buffer; see below under stat.

fsync Is a new system call for synchronizing a file's in-core state with that on disk. Its intended use is in building transaction oriented facilities.

ftruncate Is a new system call which does a *truncate* operation given a file descriptor; useful in conjunction with the new advisory locking facilities.

getdtablesize Is a new system call which returns the size of the descriptor table.

getgroups Is a new system call which returns the group access list for the caller.

gethostid Is a new system call which returns the unique (hopefully) identifier for the current host.

gethostname Is a new system call which returns the name of the current host.

getitimer Is a new system call which gets the current value for an interval timer.

getpagesize Is a new system call which returns the system page size.

getpriority Is a new system call which returns the current scheduling priority for a specific process, a group of processes, or all processes owned by a user. In the latter two cases, the priority returned is the highest (lowest numerical value) enjoyed by any of the specified processes.

getrlimit Is a new system call which returns information about a resource limit. The getrlimit and setrlimit calls replace the old vlimit call from 4.1BSD.

getrusage Is a new system call which returns information about resource utilitization of a child process or the caller. This call replaces the vtimes call of 4.1BSD.

getsockopt Is a new interprocess communication system call which returns the current options present on a socket.

gettimeofday Is a new system call which returns the current Greenwich date and time, and the current timezone in which the machine is operating. Time is returned in seconds and microseconds since January 1, 1970.

ioctl Has been changed to encode the size of parameters and whether they are to be copied in, out, or in and out of the user address space in the *request*. The symbolic names for the various ioctl requests remain the same, only the numeric values have changed. A number of new ioctls exist for use with sockets and the network facilities. The old LINTRUP request has been replaced by a call to fcntl and the SIGIO signal.

killpg Has now been made a system call; in 4.1BSD it was a library routine.

listen Is a new interprocess communication system call used to indicate a socket will be used to listen for incoming connection requests.

lseek Now has symbolic definitions for its *whence* parameter defined in <*sys/file.h*>.

mkdir Is a new system call which creates a directory.

mpx The multiplexed file facilities are no longer part of the system. They have been replaced by the socket, and related, system calls.

open Is different, now taking an optional third parameter and supporting file creation, automatic truncation, automatic append on write, and ‘‘exclusive’’ opens. The open interface has been made compatible with System III with the exception that non-blocking opens on terminal lines requiring carrier are not supported.

profil Now returns statistical information based on a 100 hz clock rate.

quota Is a new system call which is part of the disk quota facilities. Quota is used to manipulate disk quotas for a specific user, as well as perform certain random chores such as syncing quotas to disk.

read Now automatically restarts when a read on a terminal is interrupted by a signal before any data is read.

readv Is a new system call which supports scattering of read data into (possibly) disjoint areas of memory.

readlink Is a new system call for reading the value of a symbolic link.

recv Is a new interprocess communication system call used to receive a message on a connected socket.

recvfrom Is a new interprocess communication system call used to receive a message on a (possibly) unconnected socket.

recvmsg Is a new interprocess communication system call used to receive a message on a (possibly) unconnected socket which may have access rights included. When using on-machine communication, recvmsg and sendmsg may be used to pass file descriptors between processes.

rename Is a new system call which changes the name of an entry in the file system (plain file, directory, character special file, etc.). Rename has an important property in that it guarantees the target will always exist, even if the system crashes in the middle of the operation. Rename only works with source and destination in the same file system.

rmdir Is a new system call for removing a directory.

select Is a new system call (mainly for interprocess communication) which provides facility for synchronous i/o multiplexing. Sets of file descriptors may be queried for readability, writability, and if any exceptional conditions are present (such as out of band data on a socket). An optional timeout may also be supplied in which case the select operation will return after a specified period of time should no descriptor satisfy the requests.

send Is a new interprocess communication system call for sending a message on a connected socket.

sendto Is a new interprocess communication system call for sending a message on a (possibly) unconnected socket.

sendmsg Is a new interprocess communication system call for sending a message on a (possibly) unconnected socket which may included access rights.

setquota Is a new system call for enabling or disabling disk quotas on a file system.

setregid Is a new system call which replaces the 4.1BSD setgid system call. Setregid allows the real and effective group ID’s of a process to be set separately.

setreuid Is a new system call which replaces the 4.1BSD setuid system call. Setreuid allows the real and effective user ID's of a process to be set separately.

shutdown Is a new interprocess communication system call for shutting down part or all of full-duplex connection.

sigblock Is a new system call for blocking signals during a critical section of code.

sigpause Is a new system call for blocking a set of signals and then pausing indefinitely for a signal to arrive.

sigsetmask Is a new system call for setting the set of signals which are currently blocked from delivery to a process.

sigstack Is a new system call for defining an alternate stack on which signals are to be processed.

sigsys Is no longer supported. The new signal facilities are a superset of those which sigsys provided.

sigvec Is the new system call interface for defining signal actions. For each signal (except SIGSTOP and SIGKILL), sigvec allows a ''signal vector'' to be defined. The signal vector is comprised of a handler, a mask of signals to be blocked while the handler executes, and an indication of whether or not the handler should execute on a signal stack defined by a sigstack call. The old signal interface is provided as a library routine with several important caveats. First, signal actions are no longer reset to their default value after a signal is delivered to a process. Second, while a signal handler is executing the signal which is being processed is blocked until the handler returns. To simulate the old signal interface, the user must explicitly reset the signal action to be the default value and unblock the signal being processed.

Four new signals have been added for the interprocess communication and interval timer facilities. SIGIO is delivered to a process when an fcntl call enables signal driven i/o and input is present on a terminal (and a signal handler is defined). SIGURG is delivered when an urgent condition arises on a socket (and a handler is defined). SIGPROF and SIGVTALRM are associated with the ITIMER_PROF and ITIMER_VIRTUAL interval timers, and delivered to a process when such a timer expires (the SIGALRM signal is used for the ITIMER_REAL interval timer). The old SIGTINT signal is replaced by SIGIO.

socket Is a new interprocess communication system call for creating a socket.

socketpair Is a new interprocess communication system call for creating a pair of connected and unnamed sockets.

stat Now returns a larger structure. New fields are present indicating the optimal blocking factor in which i/o should be performed (for disk files the block size of the underlying file system) and the actual number of disk blocks allocated to the file. Inode numbers are now 32-bit quantities. Several spare fields have been allocated for future expansion. These include space for 64-bit file sizes and 64-bit time stamps. Two new file types may be returned, S_IFLNK for symbolic links, and S_IFSOCK for sockets residing in the file system.

swapon Has been renamed from the vswapon call of 4.1BSD.

symlink Is a new system call for creating a symbolic link.

truncate Is a new system call for truncating a file to a specific size.

unlink Should no longer be used for removing directories. Directories should only be created with mkdir and removed with rmdir. Creating hard links to directories can cause disastrous results.

utime Is defunct, replaced by utimes.

utimes Is a new system call which uses the new time format in setting the accessed and updated times on a file.

vfork Is still present, but definitely on its way out. Future plans include implementing fork with a scheme in which pages are initially shared read-only. On the first attempt by a process to write on a page the parent and child would receive separate writable copies of the page.

-- --

vlimit Is no longer supported. Vlimit is replaced by the getrlimit and setrlimit calls.

vread Is no longer supported in the system.

vswapon Has been renamed swapon.

vtimes Is no longer supported. Vtimes is replaced by the getrusage call.

vwrite Is no longer supported in the system.

wait Now is automatically restarted when interrupted by a signal before status could be returned.

wait3 Returns resource usage in a different format than that which was returned in 4.1BSD. This structure is compatible with the new getrusage system call. Wait3 is now automatically restarted when interrupted by a signal before status could be returned.

write Now is automatically restarted when writing on a terminal and interrupted by a signal before any i/o was completed.

writev Is a new version of the write system call which supports gathering of data in (possibly) discontiguous areas of memory

Section 3

The section 3 documentation has been reorganized to group manual entries by library. Introductory sections for each logical and physical library contain lists of the entry points in the library.

A number of routines which had been system calls under 4.1BSD are now user-level library routines in 4.2BSD. These routines have been grouped under section ‘‘3C’’ headings, ‘‘C’’ for compatibility. Further, certain routines present in the standard C run-time library which do not easily categorize as part of one of the standard libraries, have been group under ‘‘3X’’ headings.

curses A number of bug fixes have been incorporated, and the documentation has been revised.

stdio The standard i/o library has been modified to block i/o operations to disk files according to the block size of the underlying file system. This is accomplished using the new *st_blksize* value returned by *fstat*. The resultant performance improvement is significant as the old 1 kilobyte buffer size often resulted in 7 memory-to-memory copy operations by the system on 8 kilobyte block file systems.

End-of-file marks now ‘‘stick’’. That is, all input requests on a stdio channel after encountering end-of-file will return end-of-file until a *clearerr* call is made. This has implications for programs which use stdio to read from a terminal and do not process end-of-file as a terminating keystroke.

Two new functions may be used to control i/o buffering. The *setlinebuf* routine is used to change **stdout** or **stderr** from block buffered or unbuffered to line buffered. The *setbuffer* routine is an alternate form of *setbuf* which can be used after a stream has been opened, but before it is read or written.

bstring Three new routines, *bcmp*, *bcopy*, and *bzero* have been added to the library. These routines use the VAX string instructions to manipulate binary byte strings of a known size.

ctime Now uses the *gettimeofday* system call and supports time conversion in six different time zones. Daylight savings calculations are also performed in each time zone when appropriate.

isprint Now considers space a printing character; as the manual page has always indicated.

directory Is a new directory interface package which provides a portable interface to reading directories. A version of this library which operates under 4.1BSD is also available.

getpass Now properly handles being unable to open /dev/tty.

getwd Has been moved from the old jobs library to the standard C run-time library. It now returns an error string rather than printing on the standard error when unable to decipher the current working directory.

perror Now uses the writev system call to pass all its arguments to the system in a single system call.

This has profound effects on programs which transmit error messages across a network.

psignal And sys_siglist are routines for printing signal names in an equivalent manner to perror.

qsort Has been greatly sped up by choosing a random element with which to apply its divide and conquer algorithm.

random Is a successor to rand which generates much better random numbers. The old rand routine is still available and most programs have not been switched over to random as doing so would make certain facilities such encrypted mail unable to operate on existing data files.

setjmp And longjmp now save and restore the signal mask so that non-local exit from a signal handler is transparent. The old semantics are available with _setjmp and _longjmp.

net Is a new set of routines for accessing database files for the DARPA Internet. Four databases exist: one for host names, one for network names, one for protocol numbers, and one for network services. The latter returns an Internet port and protocol to be used in accessing a given network service.

An additional collection of routines, all prefaced with ‘‘inet_’’ may be used to manipulate Internet addresses, and interpret and convert between Internet addresses and ASCII representations in the Internet standard ‘‘dot’’ notation.

Finally, routines are available for converting 16 and 32 bit quantities between host and network order (on high-ender machines these routines are defined to be noops).

fstab The routines for manipulating /etc/fstab have been rewritten to return arbitrary length null-terminated strings.

Section 4

The system now supports the 11/730, the new 64Kbit RAM memory controllers for the 11/750 and 11/780, and the second UNIBUS adapter for the 11/750. Several new character and/or block device drivers have been added, as well as support for many hardware devices which are accessible only through the network facilities. Each new piece of hardware supported is listed below.

New manual entries in section 4 have been created to describe communications protocols, and network architectures supported. At present the only network architecture fully supported is the DARPA Internet with the TCP, IP, UDP, and ICMP protocols.

acc A network driver for the ACC LH/DH IMP interface.

ad A driver for the Data Translation A/D converter.

arp The Address Resolution Protocol for dynamically mapping betwee 32-bit DARPA Internet addresses and 48-bit Xerox 10Mb/s Ethernet addresses.

css A network driver for the DEC IMP-11A LH/DH IMP interface.

dmc A network interface driver for the DEC DMC-11/DMR-11 point-to-point communications device.

ec A network interface driver for the 3Com 10Mb/s Ethernet controller.

en A network interface driver for the Xerox 3Mb/s experimental Ethernet controller.

hy A network interface driver for the Network Systems Hyperchannel Adapter.

ik A driver for an Ikonas frame buffer graphics device interface.

il A network interface driver for the Interlan 10Mb/s Ethernet interface.

imp A network interface driver for the standard 1822 interface to an IMP; used in conjunction with either acc or css hardware.

kg A driver for a KL-11/DL-11W used as an alternate real time clock source for gathering kernel statistics and profiling information.

lo A software loopback network interface for protocol testing and performance analysis.

pcl A network interface driver for the DEC PCL-11B communications controller.

ps A driver for an Evans and Sutherland Picture System 2 graphics device connected with a DMA interface.

pty Now includes a simple packet protocol to support flow controlled operation with mechanisms for flushing data to be read and/or written.

rx A driver for the DEC dual RX02 floppy disk unit.

ts Now supports TU80 tape drives.

tu The VAX-11/750 console cassette interface has been made somewhat usable when operating in single-user mode. The device driver employs assembly language pseudo-dma code for the reception of incoming packets from the cassette.

uda Now supports RA81, RA80, and RA60 disk drives.

un A network interface driver for an Ungermann-Bass network interface unit connected to the host via a DR-11W.

up Now supports ECC correction and bad sector handling. Also has improved logic for recognizing many different kinds of disk drives automatically at boot time.

uu A driver for DEC dual TU58 tape cartridges connected via a DL-11W interface.

va The Varian driver has been rewritten so that it may coexist on the same UNIBUS with devices which require exclusive use of the bus; i.e. RK07's.

vv A network interface driver for the Proteon proNET 10Mb/s ring network controller.

Section 5

dir Reflects the new directory format.

disktab Is a new file for maintaining disk geometry information. This is a temporary scheme until the information stored in this file for each disk is recorded on the disk pack itself.

dump Is a superset of that used in 4.1BSD.

fs Reflects the new file system organization.

gettytab Is a new file which idescribes terminal characteristics. Each entry in the file describes one of the possible arguments to the getty program.

hosts Is a database for mapping between host names and DARPA Internet host addresses.

mtab Has been modified to include a ‘‘type’’ field indicating whether the file system is mounted read-only, read-write, or read-write with disk quotas enabled.

networks Is a database for mapping between network names and DARPA standard network numbers.

phones Is a phone number data base for tip.

printcap Is a termcap clone for configuring printers.

protocols Is a database for mapping between protocol names and DARPA Internetwork standard protocol numbers.

remote Is a database of remote hosts for use with tip.

services Is a database in which DARPA Internet services are recorded. The information contained in this file indicates the name of the service, the protocol which is required to access it, and the port number at which a client should connect to utilize the service.

tar Is a new entry describing the format of a tar tape.

utmp Has been augmented to include a remote host from which a login session originates. The wtmp file is also used to record FTP sessions.

vgrindefs Is a file describing how to vgrind programs written in many languages.

Section 6

aardvark Does not work because it requires the ‘‘Dungeon Definition Language’’ processor which is a binary image requiring 4.1BSD compatibility mode; the DDL source is still present.

aliens The aliens have returned home, the game is no longer included in the distribution.

backgammon Is now screen oriented. A new program, teachgammon, instructs the new backgammon player. The old version is now called btlgammon.

canfield Is a new game which plays a brand of the popular game of solitaire. Betting is included, the program cfscores may be used to find out your current debt.

ching Now pipes its output through more. Thus the hacker placates the seekers.

chase No longer exists because the binary does not work under 4.2BSD.

factor Is a rewrite in C of the old version 7 assembly language program which finds the prime factors of a number.

fortune Has yet more adages.

hangman Is now screen oriented.

mille Now plays more intelligently.

primes Is a rewrite in C of the old version 7 assembly language program which finds prime numbers within a specified range.

rogue Has been made more of a scoundrel. The supplementary document ‘‘A Guide to the Dungeons of Doom’’, has been updated as well, and is now part of Volume 2C of the programmer’s manual.

sail Is a new game which simulates sea battles of yore. The manual page is large enough to be a separate document and so has been left in its source directory.

trek The original trek has returned; trekies rejoice.

Section 7

hier Has been updated to reflect the reorganization to the user and system source.

mailaddr Is a new entry describing mail addressing syntax under sendmail (possibly too Berkeley specific).

ms The –ms macros have been extended to allow automatic creation of a table of contents. Support for the refer preprocessor is improved. Several bugs related to multi-column output and floating keeps have been fixed. Extensions to the accent mark string set are available by including the .AM macro. Footnotes can now be automatically numbered (in superscript) by –ms and referenced in the text with a ** string register. The manual page includes a summary of important number and string registers. A new document ‘‘Changes to –ms’’ is included in Volume 2C of the programmer’s manual.

Section 8

Major changes affecting system operations include:

- The system now supports disk quotas. These allow system administrators to control users’ disk space and file allocation on a per-file system basis. Utilities in this section exist for fixing, summarizing, and editing disk quota summary files.
- File systems are now made with a new program, newfs, which acts as front end to the old mkfs program. There no longer is a need to remember disk partition sizes, as newfs gets this information automatically from the /etc/disktab file. In addition, newfs attempts to lay out file systems according to the characteristics of the underlying disk drive (taking into account disk geometry information).
- DEC standard bad block forwarding is now supported on the RP06 and second vendor UNIBUS

storage module disks. The bad144 program can now be used to mark sectors bad on many disks, though inclusion in the bad sector table is still somewhat risky due to requirements in the ordering of entries in the table.

- A new program, format, should be used to initialize all non-DEC storage modules before creating file systems. Format formats the sector headers and creates a bad sector table which is used in normal system operation. Format runs in a standalone mode.
- Getty has been rewritten to use a description file, /etc/gettytab. This allows sites to tailor terminal operation and configuration without making modifications to getty.
- The line printer system is totally new. A program to administer the operation of printers, lpc, is supplied, and printer accounting has been consolidated into a single program, pac.
- The program used to restore files from dump tapes is now called *restore*. This name change was done to reinforce the fact that it is completely rewritten and operates in a very different way than the old restor program. Restore operates on mounted file systems and uses only normal file system operations to restore files. Versions of both dump and restore which operate across a network are included as rdump and rrestore. Dump and restore (and their network oriented counterparts) now perform so efficiently (mostly because of the new file system), that disk to disk backups should no longer be an attractive alternative.

arff No longer asks if you want to clobber the floppy when manipulating archives which are not on the floppy.

bad144 Has been modified to use the /etc/disktab file. Can be used to create bad sector tables for the DEC RP06 and several new Winchester disk drives. Consult the source code for details and use with extreme care.

badsect Has been modified to work with the new file system and now must interact with fsck to perform its duties. Consult the manual page for more information.

bugfiler Is a new program for automatic filing and acknowledgement of bug reports submitted by the sendbug program. Intended to operate with the Rand MH software which is part of the user contributed software. Used at Berkeley to process bug reports on 4.2BSD.

chgrp Has been moved to section 1.

comsat Has been changed to filter the noise lines in message headers when displaying incoming mail. No longer uses a second process watchdog as it uses the more reliable socket facilities instead of the old mpx facilities.

config Has been extensively modified to handle the new root and swap device specification syntax. A new document, ‘‘Configuring 4.2BSD UNIX Systems with Config’’, describes its use, as well as other important information needed in configuring system images; this is part of Volume 2C of the programmer’s manual.

diskpart Is a new program which may be used to generate disk partition tables according to the rules used at Berkeley. Can automatically generate entries required for device drivers and for the /etc/diskpart file. (Does not handle the new DEC DSA style drives properly because it tries to reserve space for the bad sector table.)

drtest Is a new standalone program which is useful in testing standalone disk device drivers and for pinpointing bad sectors on a disk.

dump Has been modified for the new file system organization. Mainly due to the new file system, it runs virtually at tape speed. Properly handles locking on the dumpdates file when multiple dumps are performed concurrently on the same machine.

dumpfs Is a new program for dumping out information about a file system such as the block size and disk layout information.

edquota Is a new program for editing user quotas. Operates by invoking your favorite editor on an ASCII representation of the information stored in the binary quota files. Edquota also has a ‘‘replication’’ mode whereby a quota template may be used to create quotas for a group of users.

fastboot Is a new shell script which reboots the system without checking the file systems; should be used with extreme care.

fasthalt Is a new script which is similar to fastboot.

format Is a new standalone program for formatting non-DEC storage modules and creating the appropriate bad sector table on the disk.

fsck Has been changed for the new file system. Fsck is more paranoid then ever in checking the disks, and has been sped up significantly. The accompanying Volume 2C document has been updated to reflect the new file system organization.

ftpd Is the DARPA File Transfer Protocol server program. It supports C shell style globbing of arguments and a large set of the commands in the specification (except the ABORT command!).

gettable Is a new program which can be used in aquiring up to date DARPA Internet host database files.

getty Has been rewritten to use a terminal description database, /etc/gettytab. Consult the manual entries for *getty* (8) and *gettytab* (5) for more information.

icheck Has been modified for the new file system.

init Has been significantly modified to use the new signal facilities. In doing so, several race conditions related to signal delivery have been fixed.

kgmon Is a new program for controlling running systems which have been created with kernel profiling. Using kgmon, profiling can be turned on or off and internal profiling buffers can dumped into a gmon.out file suiitable for interpretation by gprof.

lpc Is a new program controlling line printers and their associated spooling queues. Lpc can be used to enable and disable printers and/or their spooling queues. Lpc can also be used to rearrange existing jobs in a queue.

lpd Has been rewritten and now runs as a ‘‘server’’, using the interprocess communication facilities to service print requests. A supplementary document describing the line printer system is now part of Volume 2C of the programmer’s manual.

MAKEDEV

Is a new shell script which resides in /dev and is used to create special files there. MAKEDEV keeps commands for creating and manipulating local devices in a separate file MAKEDEV.local.

mkfs Has been virtually rewritten for the new file system. The arguments supplied are very different. For the most part, users now use the newfs program when creating file systems. Mkfs now automatically creates the lost+found directory.

mount Now indicates file systems which are mounted read-only or have disk quotas enabled.

newfs Is a new front-end to the mkfs program. Newfs figures out the appropriate parameters to supply to mkfs, invokes it, and then, if necessary, installs the boot blocks necessary to bootstrap UNIX on 11/750’s.

pac Is a new program which can be used to do printer accounting on any printer. It subsumes the vpac program.

quot Now uses the information in the inode of each file to find out how many blocks are allocated to it.

quotacheck

Is a new program which performs consistency checks on disk quota files. Quotacheck is normally run from the /etc/rc.local file after a system is rebooted, though it can also be run on mounted on file systems which are not in use.

quotaon Is a new program which enables disk quotas on file systems. A link to quotaon, named quotaoff, is used to disable disk quotas on file systems.

pstat Has been modified to understand new kernel data structures.

rc Has had system dependent startup commands moved to /etc/rc.local.

rdump Is a new program to dump file systems across a network.

renice Has been rewritten to use the new setpriority system call. As a result, you can now renice users and process groups.

repquota Is a new program which summarizes disk quotas on one or more file systems.

restor No longer exists. A new program, restore, is its successor.

restore Replaces restor. Restore operates on mounted file systems; it contains an interactive mode and can be used to restore files by name. Restore has become almost as flexible to use as tar in retrieving files from tape.

rexecd Is a network server for the *rexec* (3X) library routine. Supports remote command execution where authentication is performed using user accounts and passwords.

rlogind Is a network server for the *rlogin* (1C) command. Supports remote login sessions where authentication is performed using privileged port numbers and two files, /etc/hosts.equiv and .rhosts (in each users home directory).

rmt Is a program used by rrestore and rdump for doing remote tape operations.

route Is a program for manually manipulating network routing tables.

routed Is a routing daemon which uses a variant of the Xerox Routing Information Protocol to automatically maintain up to date routing tables.

rrestore Is a version of restore which works across a network.

rshd Is a server for the *rsh* (1C) command. It supports remote command execution using privileged port numbers and the /etc/hosts.equiv and .rhosts files in users' home directories.

rwhod Is a server which generates and listens for host status information on local networks. The information stored by rwhod is used by the *rwho* (1C) and*ruptime* (1C) programs.

rxformat Is a program for formatting floppy disks (this uses the *rx* device driver, not the console floppy interface).

savecore Has been modified to get many pieces of information from the running system and crash dump to avoid compiled in constants.

sendmail Is a new program replacing delivermail; it provides fully internetwork mail forwarding capabilities. Sendmail uses the DARPA standard SMTP protocol to send and receive mail. Sendmail uses a configuration file to control its operation, eliminating the compiled in description used in delivermail.

setifaddr Is a new program used to set a network interface's address. Calls to this program are normally placed in the /etc/rc.local file to configure the network hardware present on a machine.

syslog Is a server which receives system logging messages. Currently, only the sendmail program uses this server.

telnetd Is a server for the DARPA standard TELNET protocol.

tftpd Is a server for the DARPA Trivial File Transfer Protocol.

trpt Is a program used in debugging TCP. Trpt transliterates protocol trace information recorded by TCP in a circular buffer in kernel memory.

tunefs Is a program for modifying certain parameters in the super block of file systems.

vipw Is no longer a shell script and properly interacts with passwd, chsh, and chfn in locking the password file.