Bug fixes and changes in 4.1bsd

May 10, 1981

Bill Joy

Computer Systems Research Group

ABSTRACT

This document describes briefly the changes in the Berkeley system for the VAX between the fourth distribution of November 1980 (known originally as 4BSD, and hereinafter referred to as 4.0bsd) and this, its revision of April 1981. It attempts to be summarize, without going into great detail, the changes which have been made.

Notable improvements

- Performance of the system under heavy paging load has been substantially improved by correcting a problem with placement of pre-paged pages. The system now pre-pages more data, greatly benefitting processes which have locality in their behavior. System degradation due to pre-paging has been reduced (even though more pre-paging is done), by placing the pre-paged pages as the bottom of the free-list, rather than in the systems global clock replacement loop.
- Processes which serially reference large amounts of virtual memory may now declare themselves to be sequential (see *vadvise*(2)), causing the paging algorithm to be adjusted to be more appropriate for such behavior.* Processes which are known to need only a small amount of memory, but which tend to accumulate large amounts of memory due to strange page referencing patterns may declare a soft limit on the amount of memory to be used, see *vlimit*(2) or do ‘‘limit memoryuse *n* k’’ in *csh*(1). If the system needs memory, and there are processes which are over their declared limits, the system tends to take pages from these processes first. By saying that these limits are soft, we mean that on a system where physical memory is not scarce, processes with limits can still acquire as much physical memory as they require.
- The system is now compiled from a description of the local machine which includes information about available devices and the maximum number of users to be supported on the system. It is no longer necessary to adjust an armful of constants to re-size the system; normal time-sharing systems are parameterized from a single ‘‘maxusers’’ constant. The system also adjusts the parameters of the paging subsystem based on available memory; it is no longer necessary to adjust these for machines with very small or very large amounts of available memory.
- The system now handles multiple uba’s and mba’s and mixing of devices on mba’s. A number of new devices are supported: RK07’s and TS-11’s which come in the standard 11/750 packages, UNIBUS TM-11 tape emulators (e.g. the EMULEX TC-11 with a KENNEDY 9300 transport). Line printer and C/A/T phototypesetter drivers are included in the system, and modem control for the DH-11 using the DM-11 has been integrated and tested. The new RM-80 Winchester drive from DEC is supported. Notably unsupported are the 11/750 console casettes. A single distribution system boots on all supported hardware configurations by determining the available devices and

---

* Future versions of the system will detect strongly sequential behavior and adapt themselves accordingly, so this call will no longer be necessary.

selecting a root device at boot time.

- The *tar* program has been changed to allow update of magnetic tape archives, adding to the end of the existing tape archive. This uses tape control facilities newly added in this version of the system, but the created tapes are readable on all versions of UNIX. Since blocked tapes are inherently more efficient and fully functional, *tar* now creates blocked tapes by default.
- Shell scripts are now potentially executable. A file which begins with the characters ‘‘#!’’ followed by some amount of white space and the name of an interpreter is executed directly by the system. The standard version 7 shell, /bin/sh, has been changed to recognize a # as a comment character. It is desirable that users change existing shell script so that the first line of each shell script begins with ‘‘#!’’ and the name of the proper shell to execute.
- Two new documents have been added to Volume 2c: a manual for the crt screen package **−lcurses**, and a document describing the **rogue** game.

Major bug fixes

- A number of bugs have been fixed in the FORTRAN compiler F77; these caused some block data routiens and main programs to fail to compile. Other bugs related to use of the symbolic debugger with FORTRAN have been fixed (see **f77** in ‘‘Section 1’’ below.)
- A bug which caused the C compiler to complain ‘‘expression too complicated’’ when an expression contained an embedded structure assignment has been fixed.
- A kernel bug which caused approximately 1/8’th of data collected during program profiling to be lost has been corrected.
- A very confusing bug in the *tset* manual page has been corrected: in the example where it uses the *setenv* command it should rather use the *set* command: ‘‘set noglob; eval ...’’. The alternate form does not work on a number of terminals.
- Due to a inconsistency between the 11/780 and the 11/750 in the interpretation of the **movtuc** instruction, all binaries from the third distribution tape (of January 1980) and the fourth distribution tape (of November 1980) will NOT run correctly on the 11/750. The 4.1bsd distribution tape includes a new version of the standard i/o library which avoids use of this troublesome instruction, and it is recommended that all programs be recompiled from scratch to use this new library to eliminate non-portable UNIX binaries as soon as possible.

  The problem is that when there appears to be a chance that the operands to **movtuc** will overlap, but they in fact do not (because the pattern string to printf is terminated by a null), the 11/750 does not execute the **movtuc** as though a shorter string length were specified which did not imply potential overlap. Thus this instruction which is in the inner loop of *printf* on the 4bsd tape, and which works fine on an 11/780, malfunctions on an 11/750 causing strange and awful behavior.
- (The following change is applicable only to systems bootstrapped from the 4.1bsd release tape; systems upgraded from 4.0bsd may or may not include this change.) The standard i/o library has been made more compatible with the PDP-11 version 7 system and now supports concurrent read and write access to files (see *fopen*(3s) for details). This caused a change in an internal data structure which is visible to compiled programs through the macros in <stdio.h>, and thus necessitates discarding of old **.o** files and recompilation. Note that this standard i/o library now replaces the old **−lNS** library which was in 3bsd but was deleted from 4bsd pending the inclusion of read-write capability in stdio.

Section 1

adb A bug which prevented programs loaded with the **−N** loader option from having breakpoints set on them has been fixed. A bug which prevented backtraces of programs which are stopped at interrupt level (after a signal) has been fixed by teaching adb how to step across the interrupt stack frame. The system now makes available and the $? command in adb interprets more precise codes when SIGILL and SIGFPE exceptions occur. Thus if you divide by zero, you can find out that this was what caused a floating point exception. A bug which prevented format repetition counts (e.g. the 10 in ‘‘0/10X’’) from working has been fixed. A bug related to adb has been fixed that prevented any sensible traceback after a FORTRAN-77 runtime error which trapped an exception or fault (e.g. a floating point divide by 0). A bug which caused the maps to be set improperly when debugging the kernel on a crash dump has been fixed.

New commands have been added to adb to facilitate writing adb procedures stored in files: see the descriptions of $<<, the register <9, and the # operator in the revised adb manual page.

awk Has moved from /usr/bin to /bin because it is needed during bootstrap procedures.

bc Now takes the new C syntax operators: ‘‘=op’’ as well as (for history’s sake) the old syntax ‘‘op=’’.

calendar Several bugs have been fixed in the internals of the calendar program. In addition it is now possible to have a calendar entry consisting of merely a date (with no month specified) and have this entry occur each month.

cc A bug has been fixed in the C optimizer which caused strange things to happen when long variable names were used. The size of the internal ‘‘dimension’’ table has been increased to allow more structure fields to be declared. A problem which caused the compiler to print ‘‘expression causes compiler loop’’ when structure assignment was embedded in an expression context has been fixed. A bug in the code generator tables which affected certain FORTRAN constructs has also been fixed.

checknr Es versteht −**mm**.

csh A cosmetic bug which prevented the distributed source from compiling has been fixed. A bug which caused argument lists including the strings ‘‘{’’ or ‘‘{}’’ to sometimes fail saying ‘‘No match’’ has been fixed. A bug which caused a strange state to result when a command of the form ‘‘(a; b; c)’’ was stopped from the keyboard has been fixed. The shell has been fixed so that it prefers to make stopped jobs the new current jobs after the current job has terminated. The construct ‘‘${TERM}’’ no longer yields an extra ‘‘}’’ when the variable expanded is from the environment.

Previously undocumented was the shell variable *histchars*; it can be set to two characters, and the first is then a replacement for the normal **!** history character and the second for the ˆ quick history substitution character. Thus after ‘‘set histchars=,;’’ the command ‘‘,,’’ repeats the previous command, as ‘‘!!’’ would normally have done.

ctags Now takes an option −**v** to produce output suitable for input to *vgrind*(1).

cu Now honors uucp’s locking protocol for the dialout lines. Several bugs relating to internal buffer overruns have been fixed. An option −**b** has been added which causes the BREAK key to generate breaks (by turning input null characters into breaks sent over the line).

delivermail
Some bugs which caused strange loops when aliases were self-referential have been fixed.

diff Some bugs related to directory diffs when the −**h** option was used have been fixed; previously the −**h** option was non-functional.

dump An obnoxious bug in the dump program where it would wrongly suggest that further dump tapes were needed after a multi-tape dump had already completed has been fixed. The dump program no longer attempts to time tape rewinds itself, but uses a mechanism which does not depend on the speed of the tape drive to determine when a rewind is complete.

error Has been taught about the error message formats of the PDP-11 C compiler.

expand A internal error check which was incorrect has been fixed.

f77 Has had a bug fixed which caused some block data subroutines and main programs to not compile (the sympton was a diagnostic of the form ‘‘Undefined: LWM%d’’. A number of minor bugs which combined to make sdb function poorly with f77 have been fixed. The result of these bugs were that if file names were longer than 6 characters sdb could not read the symbol table of the resulting FORTRAN binary.

An i/o library bug which caused numbers of the form ‘‘-1.2’’ to convert improperly in list-directed input (reading as ‘‘-0.8’’, taking the mantissa as positive) has been fixed.

ld If you are using the **–r** flag and there remain undefined symbols you no longer get a spurious error return. Now respects modes of existing files and your file-creation mask more.

lisp The functions +, /, -, *, 1+, 1- now expect fixnum arguments, for interpreted functons as well as compiled ones. The source for the VMS and UNIX versions of lisp have been combined and synchronized. The construction of very large integers has been altered to use regular dotted pairs in a non-standard way, so that fewer garbage collections will occur. (It is no longer safe to use ‘‘car’’ and ‘‘cdr’’ to disect large integers). Mixed-mode arithmetic has been made faster. There is a new compiler option (liszt -r) which allows you to create **small** lisp application programs (with a mechanism other than ‘‘dumplisp’’, which tends to create very large binaries). When the result of such a compilation is executed directly, it execs /usr/ucb/lisp and then fast loads the file you asked to be executed. A new flag $ldprint has been added to control the printing of fasl/load messages. New functions (purecopy), (getdata), (putdata), and (removeaddress) have been added. A number of minor bugs have been fixed.

lpr Has been modified to work on both DECWRITER type devices at 1200 baud masquerading as printers and also on real line printers. The queueing algorithm has been made first-come-first-serve (replacing the old random job first algorithm), and a line printer driver is now available in the standard distributed system.

ls A bug has been fixed which caused essentially random numbers to be printed for the size of special files. The size now prints as 0.

make A spelling error has been fixed which had the variable EC initialized to f77 rather than FC=f77.

man A bug which caused it to core dump when reformatting a manual page whose source macros were inaccessible has been fixed.

mt Is a new command which provides easy access to operations such as forward and backward spacing of files and records on magnetic tape.

mv Has been changed to have a option ‘‘–’’ which means that all the following arguments are file names. A bug has been fixed which caused strange behaviour when the **–i** option was used on directories. A bug which prevented renaming of subdirectories to which you had no write access has been fixed.

oldcsh Which was non-functional due to a change in a internal library interface which it had taken advantage of has been fixed to work.

od New options **O**, **X**, and **D** print 32 bit integers in octal hex and decimal format; an option **w** has been added to cause 132 column output rather than the default 80 column oriented output.

pc/pi A number of bugs in the sdb symbol table produced have been fixed. Structured types may now exceed 65k bytes in size, but in that case may not be passed by value nor assigned by simple assignment. Structured types are properly returned from recursive functions. Minor bugs in passing function parameters have been fixed, and there still exist some more minor bugs here. Runtime tests can be now be enabled or disabled on a per routine basis. Runtime checks are now done properly on builtin functions. Warnings (or standards violations) are now produced for extensions to the language. Some inconsistencies in dealing with running out of memory have been fixed.

Some bugs are known: invoking a formal routine from inside a formal invocation of the same routine will mess up. In pc only: linking seperately compiled modules from other than the current directory only works if you give full path names (or the same relative paths) for the compilation and the linking.

Some changes are planned soon: for-loop variables and with pointers will be allocated to registers. The formal routine syntax is currently as in Jensen and Wirth, but will soon be changed to the the syntax of the proposed standards, in which parameters to formals must be given. It is planned that **pc** will soon do compile time constant folding.

pmerge Is a new program which can convert large Pascal programs back into a single file, eliminating the header files and separate source files normally used for separate compilation.

prof A bug which caused prof to loop when a file-static function was profiled has been fixed. A bug in the kernel which caused the results of prof to be inaccurate has also been fixed.

ps Has been changed to work with the new kernel and is no longer dependent on system sizing; thus it is no longer necessary to recompile ps when the system tables are made larger. A bug in the **s** option which caused the computed stack size to be a constant 512 bytes too large has been fixed. The ‘‘STAT’’ field now shows ‘‘>’’ in the second position for programs which are using more memory than their *memoryuse* limit (see *vlimit*(2)). It also shows ‘‘S’’ in the fourth position for processes which have declared themselves to be strongly sequential (see *vadvise*(2)). Under the **v** option, the SWRSS field has been replaced by a new field ‘‘LIM’’ which shows the limit on *memoryuse* if there is a limit, or *xx* if there is not.

px Is now written in C, is much more portable, and shares the runtime library used by the compiler. It is, unfortunately, somewhat slower now.

ranlib No longer considers it a fatal error that an archive member defines no symbols.

reset Has been changed to invoke the command

tset –Q –I

so that all terminal modes will be reset.

rm Now has a **–** option to indicate that all the following parameters are to be treated as file names. It can be used when file names begin with ‘‘**–’’.**

rmdir Now works correctly when arguments end in one or more ‘‘/’’ characters.

sdb Has been fixed to be able to deal with ‘‘interrupt-stack-frames’’, so that it is possible to set breakpoints on and get backtraces from routines called while a signal is being processed. Now handles binaries which are mode 0407 (resulting from ‘‘ld –N’’) properly.

sh Has been changed so that ‘‘#’’ begins a comment. Shell scripts to be executed by /bin/sh should begin ‘‘#! /bin/sh’’ on the first line.

tar Now defaults to blocked tape and can update blocked tapes (finally!). When reading tape tar defaults to the raw tape and figures out what is going on with blocksize so it need not be specified. Thus the only commands which should be needed are the simple ‘‘tar t’’, ‘‘tar c’’, ‘‘tar u’’ and ‘‘tar x’’.

uusend Now uses the **–r** option of *uux* to prevent presenting extreme load to the connected system.

vmstat Has been changed so that the system’s device names are printed for the disks rather than D0, D1, D2, D3. The old format output of the **–i** option is no longer available.

The fields related to paging activity are now all given in pages (of 1024 bytes). In particular, the fields ‘‘pi’’ and ‘‘po’’ which previously showed the number of pagein and pageout events now show pages, and ‘‘sr’’ which previously gave the clock scan rate in revolutions-per-minute now gives the scan rate in pages scanned per second. To provide the extra space needed to present this information, the swap in/out information, which was only marginally interesting, is no longer printed.

w Has been modified so that it no longer needs to be recompiled when the kernel tables change in size.

wall Now checks that the message is not too long for its internal buffer, rather than just overrunning the buffer. It handles running out of processes by pausing, so it is possible to use it when there are more than 25 users logged in and you are not the super-user. It also gives the machine name you are on in the broadcast messages.

who ‘‘Who am I’’ now prints your machine and login name (the first word of its output) in a format suitable for use with uucp.

write Now checks that you do not have write permission disabled and doesn't let you write to someone if they can't write back.

Section 2.

intro Now correctly reflects that argument lists may be upto 10240 characters long.

exec Has been changed so that if a file being executed begins with the magic characters #!, the rest of the line is understood to be the name of an interpreter for the executed file. Previously (and in fact still) the shell did much of this job; it automatically executed itself on a text file with executable mode when the text file's name was typed as a command. Putting the facility into the system gives the following benefits.

1) It makes shell scripts more like real executable files, because they can be the subject of 'exec.'

2) If you do a 'ps' while such a command is running, its real name appears instead of 'sh'. Likewise, accounting is done on the basis of the real name.

3) Shell scripts can be set-user-ID.

4) It is simpler to have alternate shells available; e.g. if you like the Berkeley *csh* there is no question about which shell is to interpret a file.

5) It will allow other interpreters to fit in more smoothly.

To take advantage of this wonderful opportunity, put

```
#! /bin/sh
```

at the left margin of the first line of your shell scripts. Blanks after ! are OK. Use a complete pathname (no search is done). At the moment the whole line is restricted to 32 characters. A single argument may be specified for the interpreter, thus:

```
#! /bin/csh  −f
```

is very useful, avoiding reading *.cshrc* which is often not wanted.

kill Has been changed so that when broadcasting signals universally the signal being broadcast is not sent to the broadcaster. This is used by the *halt*(8), *shutdown*(8), and *reboot*(8) commands.

reboot Has been changed to reflect the now more limited number of option flags available to the system call due to the nature of the console interface on the VAX 11/750.

sigsys Reflects the fact that the system now passes in to a signal handler an indication of what exception really happened when a SIGFPE or SIGILL is generated. The call interface at the point of a signal handler call is now specified.

Section 3

curses A large number of bug fixes have been incorporated, and the documentation has been greatly improved. The library is now much more solid than that originally distributed.

fopen Reflects the new open options which allow simultaneous read-write to files; these extensions are being incorporated to bring compatibility with the standard version 7 PDP-11 UNIX and to replace the (now defunct) **–lNS**.

printf A bug in the documentation which incorrectly stated that a %c format would not put out a null character has been fixed.

Section 4

The system now supports a number of new devices and has capability for detecting the device's presence and interrupt vectors at boot time. The system also has capability to do device location at boot time so that a system description as supplied to *config*(8) can take advantage of available hardware redundancy.

Each configurable device driver in this section now includes a SYNOPSIS giving sample lines to be included in a system configuration. Drivers also indicate the error messages which they produce and describe each error.

ct A driver for a C/A/T phototypesetter interface.

dh Now includes handling of dm-11 modem control hardware. Optional flags to the configuration program can also be given to supply hardware carrier easily for lines where wiring is inadequate.

hp Includes rm80 support.

lp Drives standard line printers, and has been improved to cut the interrupt load by a factor of 20.

mt Describes the standard tape drive interface, which now includes **ioctl** operations for raw magtape allowing such standard operations as forward and backward spacing over files and records, writing of end-of-file marks, rewinds, rewinding and offlining, and sensing of tape status. These operations are used by the *mt*(1) program which can be used to easily manipulate tape files, and by the *tar*(1) program to allow updates to blocked tapes.

tm Supports the UNIBUS TM/11 emulating tape controllers (e.g. EMULEX TC-11 with a KENNEDY 9300 drive).

ts Supports the UNIBUS TS/11 tape drive and transport.

tty Now provides a LDECCTLQ option which causes ˆS and ˆQ to function as necessary for the VT100 to work properly.

up A driver for a EMULEX SC-11 or SC-21 UNIBUS disk controller with storage module (e.g. Ampex 9300) disk drives.

va A varian raster plotter.

Section 5

No changes.

Section 6

chase Is a game where robots try to eat you.

hangman Is now screen oriented.

rain Drops keep falling on my head.

rogue Is a new and very popular screen-oriented game, see *rogue*(6) or ‘‘A Guide to the Dungeons of Doom.’’

worm Is a game where you see how long you can get the worm to grow.

worms Is an amusement where the worms chase each other around the screen.

Section 7

No changes.

Section 8

Major changes affecting system operations include:

* The system is configured from a single per-cpu specification file which includes all locally relevant information: desired devices, size of the maximum user load, timezone information, etc. The system determines at boot time the size of the file system buffer cache based on the amount of physical memory available and sizes the system data structures based on the specification of the maximum number of active users; it is no longer necessary to juggle an armful of constants to enlarge the system. The sizes of system tables are no longer compiled into programs such as **ps** and **w.**

* System error log information is saved across system crashes provided no power failure occurs, as the error log is kept in-core in the last 1024 bytes of memory. Error messages to the console are in a standard format, of the form:
  device: message

  Messages have error registers decoded as bits, so that they are readable, and no-longer refer to devices by their major/minor device numbers, rather using device names. Error messages about file system problems (device full, out of inodes, etc) are referred to with the mounted file system name, and a message is sent directly to a user if his/her process writes to a file system when it is full. All error diagnostics printed by drivers are documented in section 4 of the manual.

* DEC standard bad block forwarding is supported on all DEC disks except the RP06; the code which implements this can be easily moved into the drivers for the UNIBUS storage module disks; this is planned.

* Core dumps are saved after system crashes automatically as the system writes a core image to a portion of the swap area from which it is recovered after reboot. The system then transfers these core images to a specified UNIX directory, saving also the version of the system associated with the crash. The system also maintains an error log **/usr/adm/shutdownlog** which records all system crashes and reboots with the information about the causes.

analyze Now has an option –**u** to print page numbers of the u. areas. A bug which prevented the distributed **analyze** from working has been fixed.

bad144 Is a new program which prints out the DEC-standard bad block information for a disk, or puts it back if it gets clobbered. The name of the program comes from the fact that the program deals with ‘‘DEC standard 144’’ format.

badsect Creates a file in a file system which contains a specified disk sector of that file system. This is a primitive way of dealing with bad sectors, but works on drivers which don’t (yet) have bad block forwarding according to the standard.

getty No longer prints ‘‘Virtual’’ in front of the system identification, since there are no swap systems in sight anymore.

halt Now properly shuts down the system; see **reboot** below for details.

icheck Now properly sets the *tfree* field in the super-block.

init Now interacts with **halt**, **reboot** and **shutdown** to bring about orderly system shutdowns. No longer writes entries in the /usr/adm/wtmp accounting file when a *getty* process dies without a user logging in on a terminal. Prints a warning message (on the console) if a terminal exists, is enabled, but the open returns an error, and waits for terminal to be openable, checking each minute, to prevent thrashing in this case.

last Now distinguishes between system shutdowns and system crashes.

lastcomm A bug causing a core dump if user id’s not present in the password file were encountered has been fixed.

login Now interacts with the shutdown command to forbid logins (except by the super-user) in the last 5 minutes before a shutdown (detected by the presence of the file /etc/nologin, which also contains a message about why the system is shutting down which is presented to the rejected users.) Implements a notion of ‘‘secure terminals’’; if the file /etc/securetty exists, then it should contain a list of the secure terminals where root is allowed to login. If this file does not exist, then root can log in anywhere, as before. If a user tries to login but can't get to their home directory they are left in the root directory and warned, so that they have some chance to send mail about or fix the problem themselves. If a file *.hushlogin* exists in a users home directory the the message of the day and other such information designed for human users will not be printed at login time. This is specifically designed for the user *uucp* which is another computer and just ignores such stuff anyways. Finally a bug has been fixed which cause **login** to core dump if there were tabs in the file /etc/ttytype.

mkfs Now works properly when given a *proto* argument; previously it complained because it didn't know that the format of the file /usr/mdec/uboot had been changed.

pstat Has new flags –**m** and –**g** which interpret the multiplexor data structures.

reboot Now shuts the system down cleanly, stopping all system activity before initiating the reboot procedure. Normally, however, the new **shutdown** command is used to take the system down, notifying the users of the impending downtime. The options –**s** and –**a** options to reboot are no longer available, as it will not be possible to provide them on all VAX cpu's due to the way the console interface works. Instead, you can halt the processor using **halt** which puts the system into a tight loop, and then perform the desired reboot sequence manually at the console. This works on both 11/750 and 11/780 processors.

renice Has been modified for the new kernel and is no longer dependent on system configuration parameters; it no longer needs to be recompiled when system table sizes change.

restor Now properly handles the case where dump files are contained in dump tapes; previously the end of the embedded dump would confuse **restor** and cause the rest of the dump tape to be inaccessible.

sa Now correctly handles an arbitrary number of commands and users. Two new options have been added –**v** and –**f** giving it a googol+2 options. See *sa*(8) for details.

savecore Is a new program which runs after a reboot to save check to see if the system had crashed and, if it had, to save any core image which was successfully written to the paging area. See *savecore*(8) for a full description.

shutdown Is a new command which shuts the system down at a specified time. It broadcasts messages to users warning them as the downtime approaches, records the reason for the reboot in a log file, and cooperates with init in a cleanly shutdown (e.g. during the last 2 minutes of system operation, users attempting to log in are told that the system is going down and the reason for the shutdown but not allowed to log in.)