Changes in Sendmail Version 8*

Eric Allman

University of California, Berkeley
Mammoth Project

ABSTRACT

Version 8 of sendmail includes a number of major changes from previous versions. This paper gives a very short history of sendmail, a summary of the major differences between version 5 (the last publically available version) and version 8, and some discussion of future directions.

In 1987, the author stopped major work on sendmail due to other time commitments, only to return to active work in 1991. This paper explores why work resumed and what changes have been made.

Section 1 gives a short history of sendmail through version 5 and the motivation behind working on version 8. Section 2 has a rather detailed description of what has changed between version 5 and version 8. The paper finishes off with some thoughts about what still needs to be done.

1. HISTORY

As discussed elsewhere, [Allman83a, Allman83b, Allman&Amos85] sendmail has existed in various forms since 1980. It was released under the name delivermail in 4BSD and 4.1BSD, and as sendmail in 4.2BSD. It quickly became the dominant mail system for networked UNIX systems.

Prior the release of 4.3BSD in November 1986, the author had left the University for private industry, but continued to do some work on sendmail with activity slowly trailing off until effectively stopping after February 1987. There was minimal support done by many people for several years, until July of 1991 when the original author, who had returned the University, started active work on it again.

There were several reasons for renewed work on sendmail. There was a desire at Berkeley to convert to a subdomained structure so that individuals were identified by their subdomain rather than by their individual workstation; although possible in the old code, there were some problems, and the author was the obvious person to address them. The Computer Systems Research Group (CSRG), the group that produced the Berkeley Software Distributions, was working on 4.4BSD, and wanted an update to the mail system. Bryan Costales was working on a book on sendmail that was being reviewed by the author, which encouraged him to make some revisions. And the author wanted to try to unify some of the disparate versions of sendmail that had been permitted to proliferate.

During the 1987–91 fallow period, many vendors and outside volunteers had produced variants of sendmail. Perhaps the best known is the IDA version [IDA87]. Originally intended to be a new set of configuration files, IDA expanded into a fairly large set of patches for the code. Originally produced in Sweden, IDA development passed to the University of Illinois, and was widely used by the fairly large set of people who prefer to get and compile their own source code rather than use vendor-supplied binaries.

*An earlier version of this paper was printed in the Proceedings of the 1994 AUUG Queensland Summer Technical Conference, Gateway Hotel, Brisbane, March 1994.
In about the same time frame, attempts were made to clean up and extend the Simple Mail Transport Protocol (SMTP) [RFC821]. This involved clarifications of some ambiguities in the protocol, and correction of some problem areas [RFC1123], as well as extensions for additional functionality (dubbed Extended Simple Mail Transport Protocol, or ESMTP) [RFC1425, RFC1426, RFC1427] and a richer set of semantics in the body of messages (the Multipurpose Internet Mail Extensions, a.k.a. MIME) [RFC1521, RFC1344]. Neither the IDA group nor most vendors were modifying sendmail to conform to these new standards. It seemed clear that these were “good things” that should be encouraged. However, since no one was working on a publically available version of sendmail with these updates, they were unlikely to be widely deployed any time in the near future.

There are, of course, other mail transport agents available, such as MMDF zmailer smail and PP. However, none of these seemed to be gaining the prominence of sendmail; it appeared that most companies would not convert to another mail transport agent any time in the forseeable future. However, they might be persuaded to convert to a newer version of sendmail.

All of these convinced the author to work on a updated version of sendmail for public distribution.

The new version of sendmail is referred to as version eight (V8). Versions six and seven were skipped because of an agreement that all files in 4.4BSD would be numbered as “8.1”. Rather than have an external version number that differed from the file version numbers, sendmail just jumped directly to V8.

2. CHANGES IN VERSION EIGHT

The following is a summary of the changes between the last commonly available version of sendmail from Berkeley (5.67) and the latest version (8.6.6).

Many of these are ideas that had been tried in IDA, but many of them were generalized in V8.

2.1. Performance Enhancements

Instead of closing SMTP connections immediately, open connections are cached for possible future use. There is a limit to the number of simultaneous open connections and the idle time of any individual connection.

This is of best help during queue processing (since there is the potential of many different messages going to one site), although it can also help when processing MX records which aren’t handled by MX Piggybacking.

If two hosts with different names in a single message happen to have the same set of MX hosts, they can be sent in the same transaction. Version 8 notices this and tries to batch the messages.

For example, if two sites “foo.com” and “bar.com” are both served by UUNET, they will have the same set of MX hosts and will be sent in one transaction. UUNET will then split the message and send it to the two individual hosts.

2.2. RFC 1123 Changes

A number of changes have been made to make sendmail “conditionally compliant” (that is, it satisfies all of the MUST clauses and most but not all of the SHOULD clauses in RFC 1123).

The major areas of change are (numbers are RFC 1123 section numbers):

§5.2.7 Response to RCPT command is fast. Previously, sendmail expanded all aliases as far as it could — this could take a very long time, particularly if there were name server delays. Version 8 only checks for the existence of an alias and does the expansion later. It does still do a DNS lookup if there is an explicit host name in the RCPT command, but this time is bounded.
§5.2.8 Numeric IP addresses are logged in Received: lines. This helps tracing spoofed messages.

§5.2.17 Self domain literal is properly handled. Previously, if someone sent to user@[1.2.3.4], where 1.2.3.4 is your IP address, the mail would probably be rejected with a “configuration error”. Version 8 can handle these addresses.

§5.3.2 Better control over individual timeouts. RFC 821 specified no timeouts. Older versions of sendmail had a single timeout, typically set to two hours. Version 8 allows the configuration file to set timeouts for various SMTP commands individually.

§5.3.3 Error messages are sent as From:<>. This was urged by RFC 821 and reiterated by RFC 1123, but older versions of sendmail never really did it properly. Version 8 does. However, some systems cannot handle this perfectly legal address; if necessary, you can create a special mailer that uses the ‘g’ flag to disable this.

§5.3.3 Error messages are never sent to <>. Previously, sendmail was happy to send responses-to-responses which sometimes resulted in responses-to-responses-to-responses which resulted in .... you get the idea.

§5.3.3 Route-addrs (the ugly “<@hosta,@hostb:user@hostc>” syntax) are pruned. RFC 821 urged the use of this bletcherous syntax. RFC 1123 has seen the light and officially deprecates them, further urging that you eliminate all but “user@hostc” should you receive one of these things. Version 8 is slightly more generous than the standards suggest; instead of stripping off all the route addressees, it only strips hosts off up to the one before the last one known to DNS, thus allowing you to have pseudo-hosts such as foo.BITNET. The ‘R’ option will turn this off.

The areas in which sendmail is not “unconditionally compliant” are:

§5.2.6 Sendmail does do header munging.

§5.2.10 Sendmail doesn’t always use the exact SMTP message text from RFC 821. This is a rather silly requirement.

§5.3.1.1 Sendmail doesn’t guarantee only one connect for each host on queue runs. Connection caching gives you most of this, but it does not provide a guarantee.

§5.3.1.1 Sendmail doesn’t always provide an adequate limit on concurrency. That is, there can be several independent sendmails running at once. My feeling is that doing an absolute limit would be a mistake (it might result in lost mail). However, if you use the XLA contributed software, most of this will be guaranteed (but I don’t guarantee the guarantee).

2.3. Extended SMTP Support

Version 8 includes both sending and receiving support for Extended SMTP support as defined by RFC 1425 (basic) and RFC 1427 (SIZE); and limited support for RFC 1426 (BODY). The body support is minimal because the “8BITMIME” body type is not currently advertised. Although such a body type will be accepted, it will not be correctly converted to 7 bits if speaking to a non-8-bit-MIME aware SMTP server.

Sendmail tries to speak ESMTP if you have the ‘a’ flag set in the flags for the mailer descriptor, or if the other end advertises the fact that it speaks ESMTP. This is a non-standard advertisement: sendmail announces “ESMTP spoken here” during the initial connection message, and client sendmails search for this message. This creates some problems for some PC-based mailers, which do not understand two-line greeting messages as required by RFC 821.
2.4. Eight-Bit Clean

Previous versions of sendmail used the 0200 bit for quoting. This version avoids that use. However, you can set option ‘7’ to get seven bit stripping for compatibility with RFC 821, which is a 7-bit protocol. This option says “strip to 7 bits on input”.

Individual mailers can still produce seven bit output using the ‘7’ mailer flag. This flag says “strip to 7 bits on output”.

2.5. User Database

The User Database (UDB) is an as-yet experimental attempt to provide unified large-site name support. We are installing it at Berkeley; future versions may show significant modifications. Briefly, UDB contains a database that is intended to contain all the per-user information for your workgroup, such as people’s full names, their .plan information, their outgoing mail name, and their mail drop.

The user database allows you to map both incoming and outgoing addresses, much like IDA. However, the interface is still better with IDA; in particular, the alias file with incoming/outgoing marks provides better locality of information.

2.6. Improved BIND Support

The BIND support, particularly for MX records, had a number of annoying “features” which have been removed in this release. In particular, these more tightly bind (pun intended) the name server to sendmail, so that the name server resolution rules are incorporated directly into sendmail.

The major change has been that the $[ ... $] operator didn’t fully qualify names that were in DNS as A or MX records. Version 8 does this qualification.

This has proven to be an annoyance in Sun shops, who often still run without BIND support. However, it is really critical that this be supported, since MX records are mandatory. In SunOS you can choose either MX support or NIS support, but not both. This is fixed in Solaris, and some sendmail support to allow this in SunOS should be forthcoming in a future release.

2.7. Keyed Files

Generalized keyed files is an idea taken directly from IDA sendmail (albeit with a completely different implementation). They can be useful on large sites.

Version 8 includes the following built-in map classes:

- dbm: Support for the ndbm(3) library.
- hash: Support for the “Hash” type from the new Berkeley db(3) library. This library provides substantially better database support than ndbm(3), including in-memory caching, arbitrarily long keys and values, and better disk utilization.
- btree: Support for the “B-Tree” type from the new Berkeley db(3) library. B-Trees provide better clustering than Hashed files if you are fetching lots of records that have similar keys, such as searching a dictionary for words beginning with “detr”.
- nis: Support for NIS (a.k.a. YP) maps. NIS+ is not supported in this version.
- host: Support for DNS lookups.
- dequote: A “pseudo-map” (that is, once that does not have any external data) that allows a configuration file to break apart a quoted string in the address. This is necessary primarily for DECnet addresses, which often have quoted addresses that need to be unwrapped on gateways.
2.8. Multi-Word Classes & Macros in Classes

Classes can now be multiple words. For example,

CShofmann.CS.Berkeley.EDU

allows you to match the entire string "hofmann.CS.Berkeley.EDU" using the single construct "$=S".

Class definitions are now allowed to include macros — for example:

Cw$k

is legal.

2.9. IDENT Protocol Support

The IDENT protocol as defined in RFC 1413 [RFC1413] is supported. However, many systems have a TCP/IP bug that renders this useless, and the feature must be turned off. Roughly, if one of these system receives a “No route to host” message (ICMP message ICMP_UNREACH_HOST) on any connection, all connections to that host are closed. Some firewalls return this error if you try to connect to the IDENT port, so you can’t receive email from these hosts on these systems. It’s possible that if the firewall used a more specific message (such as ICMP_UNREACH_PROTOCOL, ICMP_UNREACH_PORT or ICMP_UNREACH_NET_PROHIB) it would work, but this hasn’t been verified.

IDENT protocol support cannot be used on 4.3BSD, Apollo DomainOS, Apple A/UX, ConvexOS, Data General DG/UX, HP-UX, Sequent Dynix, or Ultrix 4.x, x ≤ 3. It seems to work on 4.4BSD, IBM AIX 3.x, OSF/1, SGI IRIX, Solaris, SunOS, and Ultrix 4.4.

2.10. Separate Envelope/Header Processing

Since the From: line is passed in separately from the envelope sender, these have both been made visible; the $g macro is set to the envelope sender during processing of mailer argument vectors and the header sender during processing of headers.

It is also possible to specify separate per-mailer envelope and header processing. The SenderRWSet and RecipientRWset arguments for mailers can be specified as “envelope/header” to give different rewritings for envelope versus header addresses.

2.11. Owner-List Propagates to Envelope

When an alias has an associated owner-list name, that alias is used to change the envelope sender address. This will cause downstream errors to be returned to that owner.

Some people find this confusing because the envelope sender is what appears in the first “From:” line in UNIX messages (that is, the line beginning “From<space>” instead of “From:”; the latter is the header from, which does indicate the sender of the message). In previous versions, sendmail has tried to avoid changing the envelope sender for back compatibility with UNIX convention; at this point that back compatibility is creating too many problems, and it is necessary to move forward into the 1980s.

2.12. Command Line Flags

The −B flag has been added to pass in body type information.

The −p flag has been added to pass in protocol information that was previously passed in by defining the $r and $s macros.

The −X flag has been added to allow logging of all protocol in and out of sendmail for debugging. You can set “−X filename” and a complete transcript will be logged in that file. This gets big fast: the option is only for debugging.
The −q flag can limit a queue run to specific recipients, senders, or queue ids using −qRsubstring, −qSsubstring, or −qIsubstring respectively.

2.13. New Configuration Line Types

The ‘T’ (Trusted users) configuration line has been deleted. It will still be accepted but will be ignored.

The ‘K’ line has been added to declare database maps.

The ‘V’ line has been added to declare the configuration version level.

The ‘M’ (mailer) line takes a D= field to specify execution directory.

2.14. New and Extended Options

Several new options have been added, many to support new features, others to allow tuning that was previously available only by recompiling. Briefly:

A The alias file specification can now be a list of alias files. Also, the configuration can specify a class of file. For example, to search the NIS aliases, use “OAnis:mail.aliases”.

b Insist on a minimum number of disk blocks.

C Delivery checkpoint interval. Checkpoint the queue (to avoid duplicate deliveries) every C addresses.

E Default error message. This message (or the contents of the indicated file) are prepended to error messages.

G Enable GECOS matching. If you can’t find a local user name and this option is enabled, do a sequential scan of the passwd file to match against full names. Previously a compile option.

h Maximum hop count. Previously this was compiled in.

I This option has been extended to allow setting of resolver parameters.

j Send errors in MIME-encapsulated format.

K Forward file path. Where to search for .forward files — defaults to $HOME/.forward.

k Connection cache size. The total number of connections that will be kept open at any time.

L Connection cache lifetime. The amount of time any connection will be permitted to sit idle.

l Enable Errors-To: header. These headers violate RFC 1123; this option is included to provide back compatibility with old versions of sendmail.

O Incoming daemon options (e.g., use alternate SMTP port).

p Privacy options. These can be used to make your SMTP server less friendly.

r This option has been extended to allow finer grained control over timeouts. For example, you can set the timeout for SMTP commands individually.

R Don’t prune route-addrs. Normally, if version 8 sees an address like "<@hostA,@hostB:user@hostC>", sendmail will try to strip off as much as it can (up to user@hostC) as suggested by RFC 1123. This option disables that behaviour.

T The “Return To Sender” timeout has been extended to allow specification of a warning message interval, typically something on the order of four hours. If a message cannot be delivered in that interval, a warning message is sent back to the sender but the message continues to be tried.

U User database spec. This is still experimental.
V  Fallback “MX” host. This can be thought of as an MX host that applies to all addresses that has a very high preference value (that is, use it only if everything else fails).
w  If set, assume that if you are the best MX host for a host, you should send directly to that host. This is intended for compatibility with UIUC sendmail, and may have some use on firewalls.
7  Do not run eight bit clean. Technically, you have to assert this option to be RFC 821 compatible.

2.15. New Mailer Definitions
L=  Set the allowable line length. In V5, the L mailer flag implied a line length limit of 990 characters; this is now settable to an arbitrary value.
F=a  Try to use ESMTP. It will fall back to SMTP if the initial EHLO packet is rejected.
F=b  Ensure a blank line at the end of messages. Useful on the *file* mailer.
F=c  Strip all comments from addresses; this should only be used as a last resort when dealing with cranky mailers.
F=g  Never use the null sender as the envelope sender, even when running SMTP. This violates RFC 1123.
F=7  Strip all output to this mailer to 7 bits.
F=L  Used to set the line limit to 990 bytes for SMTP compatibility. It now does that only if the L= keyletter is not specified. This flag is obsolete and should not be used.

2.16. New or Changed Pre-Defined Macros
$k  UUCP node name from uname(2).
$m  Domain part of our full hostname.
$_  RFC 1413-provided sender address.
$w  Previously was sometimes the full domain name, sometimes just the first word. Now guaranteed to be the first word of the domain name (i.e., the host name).
$j  Previously had to be defined — it is now predefined to be the full domain name, if that can be determined. That is, it is equivalent to $w.$m.

2.17. New and Changed Classes
$s=k  Initialized to contain $k.
$s=w  Now includes “[1.2.3.4]” (where 1.2.3.4 is your IP address) to allow the configuration file to recognize your own IP address.

2.18. New Rewriting Tokens
The $& construct has been adopted from IDA to defer macro evaluation. Normally, macros in rulesets are bound when the rule is first parsed during startup. Some macros change during processing and are uninteresting during startup. However, that macro can be referenced using “$&x” to defer the evaluation of $x until the rule is processed.

The tokens $( and $) have been added to allow specification of map rewriting.

Version 8 allows $@ on the Left Hand Side of an ‘R’ line to match zero tokens. This is intended to be used to match the null input.
2.19. Bigger Defaults

Version 8 allows up to 100 rulesets instead of 30. It is recommended that rulesets 0–9 be reserved for sendmail’s dedicated use in future releases.

The total number of MX records that can be used has been raised to 20.

The number of queued messages that can be handled at one time has been raised from 600 to 1000.

2.20. Different Default Tuning Parameters

Version 8 has changed the default parameters for tuning queue costs to make the number of recipients more important than the size of the message (for small messages). This is reasonable if you are connected with reasonably fast links.

2.21. Auto-Quoting in Addresses

Previously, the “Full Name <email address>” syntax would generate incorrect protocol output if “Full Name” had special characters such as dot. This version puts quotes around such names.

2.22. Symbolic Names On Error Mailer

Several names have been built in to the $@ portion of the $#error mailer. For example:

```$error $@NOHOST $: Host unknown
```

Prints the indicated message and sets the exit status of sendmail to EX_NOHOST.

2.23. New Built-In Mailers

Two new mailers, *file* and *include*, are included to define options when mailing to a file or a :include: file respectively. Previously these were overloaded on the local mailer.

2.24. SMTP VRFY Doesn’t Expand

Previous versions of sendmail treated VRFY and EXPN the same. In this version, VRFY doesn’t expand aliases or follow .forward files.

As an optimization, if you run with your default delivery mode being queue-only, the RCPT command will also not chase aliases and .forward files. It will chase them when it processes the queue. This speeds up RCPT processing.

2.25. [IPC] Mailers Allow Multiple Hosts

When an address resolves to a mailer that has “[IPC]” as its “Path”, the $@ part (host name) can be a colon-separated list of hosts instead of a single hostname. This asks sendmail to search the list for the first entry that is available exactly as though it were an MX record. The intent is to route internal traffic through internal networks without publishing an MX record to the net. MX expansion is still done on the individual items.

2.26. Aliases Extended

The implementation has been merged with maps. Among other things, this supports multiple alias files and NIS-based aliases. For example:

```OA/etc/aliases,nis:mail.aliases
```

will search first the local database “/etc/aliases” followed by the NIS map.
2.27. Portability and Security Enhancements

A number of internal changes have been made to enhance portability.

Several fixes have been made to increase the paranoia factor.

In particular, the permissions required for .forward and :include: files have been tightened up considerably. V5 would pretty much read any file it could get to as root, which exposed some security holes. V8 insists that all directories leading up to the .forward or :include: file be searchable ("x" permission) by the controlling user (defined below), that the file itself be readable by the controlling user, and that .forward files be owned by the user who is being forwarded to or root.

The "controlling user" is the user on whose behalf the mail is being delivered. For example, if you mail to "user1" then the controlling user for ˜user1/.forward and any mailers invoked by that .forward file, including :include: files.

Previously, anyone who had a home directory could create a .forward could forward to a program. Now, sendmail checks to make sure that they have an "approved shell", that is, a shell listed in the /etc/shells file.

2.28. Miscellaneous Fixes and Enhancements

A number of small bugs having to do with things like backslash-escaped quotes inside of comments have been fixed.

The fixed size limit on header lines (such as “To:” and “Cc:”) has been eliminated; those buffers are dynamically allocated now.

Sendmail writes a /etc/sendmail.pid file with the current process id and the current invocation flags.

Two people using the same program (e.g., submit) are considered "different" so that duplicate elimination doesn’t delete one of them. For example, two people forwarding their email to |submit will be treated as two recipients.

The mailstats program prints mailer names and gets the location of the sendmail.st file from /etc/sendmail.cf.

Many minor bugs have been fixed, such as handling of backslashes inside of quotes.

A hook has been added to allow rewriting of local addresses after aliasing.

3. FUTURE WORK

The previous section describes sendmail as of version 8.6.6. There is still much to be done. Some high points are described below. This list is by no means exhaustive.

3.1. Full MIME Support

Currently sendmail only supports seven bit MIME messages. Although it can pass eight bit MIME messages, it cannot advertise that fact because the standards say that the mail agent must be able to do 8- to 7-bit conversion to have full 8-bit support. This requires far more extensive modification of the message body than is currently supported.

The best way to do this would be to support the general concept of an external “message filter” that could do arbitrary modifications of the message. This would allow MIME conversion as well as such things as automatic encryption of messages sent over external links. This is probably an extremely non-trivial change.

3.2. Service Switch Abstraction

Most modern systems include some concept of a “service switch” — for example, to look up host names you can try DNS, NIS, NIS+, text tables, NetInfo, or other services in some arbitrary
order. This is currently very clumsy in sendmail, with only limited control of the services provided.

3.3. More Control of Local Addresses

Currently some addresses are declared as “local” and are handled specially — for example, they may have .forward files, may be translated into program calls or file deliveries, and so forth. These should be broken out into separate flags to allow the local system administrator to have more fine-grained control over operations.

3.4. More Run-Time Configuration Options

There are many options that are configured at compile time, such as the method of file locking and the use of the IDENT protocol [RFC1413]. These should be transferred to run time by adding new options.

Similarly, some options are currently overloaded, that is, a single option controls more than one thing. These should probably be broken out into separate options.

This implies that options will change from single characters to words.

3.5. More Configuration Control Over Errors

Currently, the configuration file can generate an error message during parsing. However, it cannot tweak other operations, such as issuing a warning message to the system postmaster. Similarly, some errors should not be triggered if they are in aliases during an alias file rebuild, but should be triggered if that alias is actually used.

3.6. Long Term Host State

Currently, sendmail only remembers host status during a single queue run. This should be converted to long term status stored on disk so it can be shared between instantiations of sendmail. Entries will have to be timestamped so they can time out. This will allow sendmail to implement exponential backoff on queue runs on a per-host basis.

3.7. Connection Control

Modern networks have different types of connectivity than the past. In particular, the rising prominence of dialup IP has created certain challenges for automated servers. It is not uncommon to try to make a connection to a host and have it fail, even though if you tried again it would succeed. The connection management could be a bit cleverer to try to adapt to such situations.

3.8. Other Caching

When you do an MX record lookup, the name server automatically returns the IP addresses of the associated MX servers. This information is currently ignored, and another query is done to get this information. It should be cached to avoid excess name server traffic.

4. REFERENCES

[Allman83a]

[Allman83b]

[Allman&Amos85]


