Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 10, 2013

A. Melnikov
Isode Ltd
K. Carlberg
G11

July 9, 2012

Simple Mail Transfer Protocol extension for Message Transfer Priorities

draft-melnikov-smtp-priority-21

Abstract

This memo defines an extension to the SMTP (Simple Mail Transfer Protocol) service whereby messages are given a label to indicate preferential handling, to enable mail handling nodes to take this into account for onward processing.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 10, 2013.

Copyright Notice

Table of Contents

1. Introduction

Where resources for switching or transfer of messages are constrained (e.g., bandwidth, round trip time, transition storage or processing capability) it is desirable to give preferential handling to some messages over others, according to their labeled priority. This is particularly important during emergencies for first responders (Appendix C) and for environments such as military (Appendix A) and aviation (Appendix B) messaging, where messages have high operational significance, and the consequences of extraneous delay can be significant.

In order for an SMTP receiver to be able to relay higher priority messages first, there needs to be a mechanism to communicate (during both Message Submission [RFC6409] and Message Transfer [RFC5321]) the priority of each message. This specification defines this mechanism by specification of an SMTP [RFC5321] extension.

In order to permit end-to-end use of this extension across email infrastructure that does not support it, a companion tunneling mechanism is defined in [PRIORITY-TUNNELING] through use of a new message header field [RFC5322].

This extension provides services to some classes of users in networks with limited available bandwidth or long round trip times, when the actual message transfer over the network can create a significant portion of the overall message delivery time from a sender to a recipient, for example over a satellite or high frequency radio link. It is also useful in case of a Mail Transfer Agent (MTA) queue build-up due to the rate of incoming messages being higher than the rate of outgoing messages. When neither of the two conditions mentioned above is true, the use of the MT-PRIORITY SMTP extension will not result in a better SMTP service to any user. Also note that while this SMTP extension can help in improving delivery speed for higher priority messages, it does not provide any sort of guarantees that for two given messages with priorities M and N (M > N) submitted simultaneously the message with priority M will arrive earlier than the message with priority N. I.e. this extension calls for best effort to provide preferential processing.

Besides the actions taken at the application level it can thus be important to deploy priority or precedence mechanisms offered by the network itself to ensure timely delivery of the emails. Examples would be the use of DiffServ [RFC2474], RSVP [RFC2205] and the work-in-progress effort extension to RSVP that prioritizes reservations.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119] when they appear in ALL CAPS. These words also appear in this document in lower case as plain English words, absent their normative meanings.

The formal syntax use the Augmented Backus-Naur Form (ABNF) [RFC5234] notation including the core rules defined in Appendix B of RFC 5234 [RFC5234].

In examples, "C:" and "S:" indicate lines sent by the client and server respectively. Line breaks that do not start with a new "C:" or "S:" exist for editorial reasons and are not a part of the protocol.

This document uses the term "priority" specifically in relation to the internal treatment of a message by the server: Messages with higher priorities may be given expedited handling, and those with lower priorities may be handled only as resources become available.

3. Definition of the Priority SMTP Extension

The Priority SMTP service extension is defined as follows:

1. The textual name of this extension is "Priority Message Handling".

2. The EHLO keyword value associated with this extension is "MT-PRIORITY".

3. The EHLO keyword has an OPTIONAL parameter which conveys the name of the Priority Assignment Policy (see Section 9.2) used by the server. (See the <mt-priority-ehlo> ABNF non terminal in Section 7 for details of its syntax.) Absence of the parameter means that the server is unwilling to disclose its Priority Assignment Policy. Clients can choose to use the MT-PRIORITY SMTP extension even if they don't recognize a partilar Priority Assignment Policy name advertised by a server.

4. No additional SMTP verbs are defined by this extension.

5. One optional parameter ("MT-PRIORITY") is added to the MAIL FROM command. The value associated with this parameter is a decimal integer number from -9 to 9 (inclusive) indicating the priority of the email message (See Appendix E for more details on why this range was selected.) The syntax of the MT-PRIORITY parameter is

described by the <priority-value> ABNF non-terminal defined in Section 7. Higher numbers mean higher priority.

6. The maximum length of a MAIL FROM command line is increased by 15 octets by the possible addition of a space, the MT-PRIORITY keyword and a priority value.

7. The MT-PRIORITY extension is valid for the submission service [RFC6409] and LMTP [RFC2033].

4. Handling of messages received via SMTP

This section describes how a conforming SMTP server should handle any messages received via SMTP.

4.1. Handling of the MT-PRIORITY parameter by the receiving SMTP server

The following rules apply to SMTP transactions in a server that supports the MT-PRIORITY parameter:

1. If any of the associated <esmtp-value>s (as defined in Section 4.1.2 of [RFC5321]) are not syntactically valid, or if there is more than one MT-PRIORITY parameter in a particular MAIL FROM command, the server MUST return an error, for example "501 syntax error in parameter" (with 5.5.2 Enhanced Status Code [RFC2034] [RFC5248]).

2. When inserting a Received header field as specified in Section 4.4 of [RFC5321], the compliant MTA/MSA SHOULD include the "PRIORITY" clause whose syntax is specified in Section 7.

3. The received MT-PRIORITY parameter value SHOULD be logged as part of any logging of message transactions.

4. If the sending SMTP client specified the MT-PRIORITY parameter to the MAIL FROM command, then the value of this parameter is the message priority.

5. If no priority has been determined by the above, the server may use its normal policies to set the message's priority. By default, each message has priority 0.

The SMTP server MUST NOT allow "upgraded" (positive) priorities from untrusted (e.g. unauthenticated) or unauthorized sources. (One example of an "unauthorized source" might be an SMTP sender which successfully authenticated using SMTP AUTH, but which is not explicitly authorized to use the SMTP MT-PRIORITY service. In case of MTA-to-MTA transfer such authorization will usually be done as a

bilateral agreement between two domains to honour priorities from each other.) The server MAY, however, allow an untrusted source to lower its own message's priorities -- consider, for example, an email marketer that voluntarily sends its marketing messages at a negative priority.

The SMTP server MAY also alter the message priority (to lower or to raise it) in order to enforce some other site policy. (Note that this also includes the case when the priority is not explicitly specified.) For example, an MSA might have a mapping table that assigns priorities to messages based on authentication credentials.

If the SMTP server changes (lowers or raises) the priority of a message, it SHOULD use the X.7.TBD3 Enhanced Status Code [RFC2034] in its response to the MAIL FROM or in the final response to the DATA (or similar) command. The human readable text part after the status code contains the new priority, followed by SP (ASCII space) and explanatory human readable text.

Alternatively an SMTP server, which is an MSA, MAY reject a message based on the determined priority. In such cases, the MSA SHOULD use 450 or 550 reply code. The corresponding Enhanced Status Code MUST be X.7.TBD1 [RFC2034] if the determined priority level is below the lowest priority currently acceptable for the receiving SMTP server. Note that this condition might be temporary. In some environments, operational policies might permit periods of operation that relay only higher priority messages and reject lower priority ones. Such handling choices need to be specified for that operational environment.

4.2. Relay of messages to other conforming SMTP servers

The following rules govern the behavior of a conforming MTA (in the role of an SMTP/LMTP client), when relaying a message which was received via the SMTP protocol, to an SMTP/LMTP server that supports the MT-PRIORITY extension:

1. A MT-PRIORITY parameter with the value determined by the procedure from Section 4.1 MUST appear in the MAIL FROM command issued when the message is relayed to an MTA/MDA which also supports the MT-PRIORITY extension. (Note that due to site policy this value might be different from the value received from the SMTP client. See Section 4.1 for details. Also note that this value might be different than the priority level at which the MTA actually handles the request, due to the rounding described in Section 5.)

2. Further processing of the MT-PRIORITY parameter is described in Section 5.

4.3. Relay of messages to non-conforming SMTP servers

The following rules govern the behavior of a conforming MTA (in the role of an SMTP/LMTP client), when relaying a message which was received via the SMTP protocol, to an SMTP/LMTP server that does not support the MT-PRIORITY extension:

1. The MTA relays the message without including the MT-PRIORITY parameter in the MAIL FROM command.

[[RFC Editor note: The lonely list item above is numbered for reference from another document, and should be left as a numbered item.]]

4.4. Mailing lists and Aliases

Several types of mechanisms exist to redirect or forward messages to alternative or multiple addresses [RFC5598]. Examples for this are aliases and mailing lists [RFC5321].

If a message is subject to such processing, the Mediator node (Section 2.1 of [RFC5598]), SHOULD retain the MT-PRIORITY parameter value for all expanded and/or translated addresses.

4.5. Gatewaying a message into a foreign environment

The following rules govern the behavior of a conforming MTA, when gatewaying a message that was received via the SMTP protocol, into a foreign (non-SMTP) environment:

1. If the destination environment is unable to provide an equivalent of the MT-PRIORITY parameter, the conforming MTA SHOULD behave as if it is relaying to a non-conformant SMTP server (Section 4.3).

2. If the destination environment is capable of providing an equivalent of the MT-PRIORITY parameter, the conforming MTA SHOULD behave as if it is relaying to a conformant SMTP server (Section 4.2), converting the MT-PRIORITY value to the equivalent in the destination environment.

4.6. Interaction with DSN SMTP Extension

An MTA which needs to generate a delivery report (whether for successful delivery or delayed/failed delivery) for a message it is processing SHOULD use the priority value of the message as the

priority of the generated delivery report. In particular this requirement applies to MTAs that also implements [RFC3461].

For delivery reports (DSNs) received by an MTA for relay, processing rules specified in Section 4.1 apply -- there is no special processing for relayed DSNs. It might seem tempting to try to detect DSNs and process them at elevated priority under the assumption that failure notices need to get through quickly, even or perhaps especially if the DSN came from an untrusted source. But such a policy can create an exposure to fake-DSN attacks by giving untrusted systems a way to inject high-priority messages. Implementation of such a policy also assumes that DSNs can be detected reliability, which may not be the case since some systems use nonstandard DSN formats.

5. The Priority Service Extension

The priorities of messages affect the order the messages are transferred from the client to the server. This is largely independent from the order in which they were originally received by the server.

A message priority is a decimal integer in the range from -9 to 9 (inclusive). SMTP servers compliant with this specification are not required to support all 19 distinct priority levels (i.e. to treat each priority value as a separate priority), but they MUST implement all distinct priority levels specified in the Priority Assignment Policy (see Section 9.2) implemented by the server, i.e. an implementation that only supports N priority levels (where N < 19) will internally round up a syntactically valid priority value that isn't supported to the next higher supported number (or to the highest supported priority, if the value is higher than any supported priority). For example, an implementation can treat priority values below and including -4 as priority -4, priority -3 as priority -2, and all priorities starting from 5 can be treated as priority 6. (See Section 9.2 for implementation/deployment considerations related to Priority Assignment Policy.)

Irrespective of the number of distinct priority levels supported by the SMTP server, when relaying the message to the next hop or delivering it over LMTP, the SMTP server MUST communicate the priority value as determined in Section 4.1.

Note: 19 possible priority levels are defined by this specification for extensibility. For example, a particular implementation or deployment environment might need to provide finer-grained control over message transfer priorities. See Appendix E for more details on why the range from -9 to 9 was selected.

As per the Priority Assignment Policy, some SMTP servers MAY impose additional maximum message size constraints for different message transfer priorities, for example messages with priority 6 might not be larger than 4 Kb. If an SMTP server chooses to reject a message because it is too big for the determined priority, it SHOULD use 552 reply codes, together with the X.3.TBD2 Enhanced Status Code [RFC2034].

Implementation Note: If the SMTP server also supports the SMTP SIZE extension [RFC1870] then an SMTP client can use both SIZE= and MT-PRIORITY= parameters on the MAIL FROM command. This allows the server to perform early rejection of a message in case the message size is too big for the specified priority, thus avoiding wasting bandwidth by transferring the message first and then rejecting it due to its size.

The Priority Service Extension can be combined with DELIVERBY [RFC2852] SMTP service extension, however there is no requirement that both extensions are always implemented together.

5.1. Expedited Transfer

The main service provided by the Priority Message Handling SMTP Service Extension is expedited transfer of emails with a higher priority. Therefore an SMTP client that has more than one email to send at a given time sends those with a higher priority before those with a lower one. Additionally, the retry interval and/or default timeout before non-delivery report is generated MAY be lower (more aggressive) for messages of higher priority. Lower retry intervals/ default timeouts are controlled by the local MTA policy.

Note that as this SMTP extension requires some sort of trust relationship between a sender and a receiver and thus some form of authentication (whether using SMTP AUTH, TLS, IP address whitelist, etc.), so senders using this SMTP extension will not be subject to greylisting [GREYLISTING], unless they are unauthorized to use this SMTP extension, due to an explicit policy decision or a misconfiguration error. But note that in case of connection-level or SMTP HELO/HELO greylisting SMTP AUTH or TLS authentication options are not available to server.

In order to make implementations of this extension easier, this SMTP extension only allows a single priority for all recipients of the same message.

Within a priority level, the MTA uses its normal algorithm (the algorithm used in absence of this SMTP extension) for determining message processing order.

Several possible ways of implementing expedited transfer are described in more details in Appendix D. Note that these sections don't describe all details and pitfalls for each implementation strategy.

5.2. Timely Delivery

An important constraint (usually associated with higher priority levels) in some environments is that messages with high priority values have some delivery time constraints. In some cases, higher priorities mean a shorter maximum time allowed for delivery.

Unextended SMTP does not offer a service for timely delivery, i.e. "deliver this message within X seconds from submission" service. The "Deliver By SMTP Service Extension" (DELIVERBY Extension) defined in [RFC2852] is an example of an SMTP extension providing a service that can be used to implement timely delivery. Note that SMTP DELIVERBY and SMTP MT-PRIORITY extensions are complimentary and can be used together (assuming the SMTP server they are talking to advertises support for both). However note that use of the DELIVERBY extension alone does not guarantee any priority processing. If the client is using both SMTP DELIVERBY and SMTP MT-PRIORITY at the same time, client can consider using smaller DELIVERBY timeouts for higher priority messages.

6. Use of MT-PRIORITY with LMTP

An LMTP server can advertise support for the MT-PRIORITY extension if it supports any combination of the following features:

1. The LMTP server is architected in such a way that it can deliver higher priority messages quicker than lower priority messages.

2. The LMTP server logs that MT-PRIORITY extension was used by the previous SMTP hop.

3. The LMTP server is exposing information about MT-PRIORITY extension to a delivery time filtering engine such as Sieve [RFC5228].

7. Syntax

```
priority-value = (["-"] NZDIGIT) / "0"
                 ; Allowed values are from -9 to 9 inclusive

NZDIGIT = %x31-39
          ; "1"-"9"

CFWS = <defined in RFC 5322>

; New "clause" that can be used in the Received header field
Pri  = CFWS "PRIORITY" FWS priority-value
          ; Complies with the <Additional-Registered-Clauses>
          ; non-terminal syntax from RFC 5321.

mt-priority-ehlo = "MT-PRIORITY" [SP priority-profile]
          ; Complies with the <ehlo-line> ABNF production from RFC 5321.

priority-profile = 1*20 (ALPHA / DIGIT / "-" / "_" / ".")
          ; name of the Priority Assignment Profile advertized in
          ; the MT-PRIORITY EHLO response.

ALPHA = <Defined in RFC 5234>

DIGIT = <Defined in RFC 5234>
```

8. Example

The original submission (from MUA to MSA) might look like shown below. Note that the example is also making use of the DELIVERBY [RFC2852] and DSN [RFC3461] SMTP extensions, even though there is no requirement that these other extensions are to be supported when the MT-PRIORITY SMTP extension is implemented.

```
S: 220 example.com SMTP server here
C: EHLO mua.example.com
S: 250-example.com
S: 250-AUTH STARTTLS
S: 250-AUTH SCRAM-SHA-1 DIGEST-MD5
S: 250-DSN
S: 250-DELIVERBY
S: 250-ENHANCEDSTATUSCODES
S: 250 MT-PRIORITY MIXER
C: AUTH SCRAM-SHA-1
[...authentication exchange...]
S: 235 2.7.0 Authentication successful
C: MAIL FROM:<eljefe@example.com> BY=125;R ENVID=QQ314159
    MT-PRIORITY=3
S: 250 2.1.0 <eljefe@example.com> sender ok
C: RCPT TO:<topbanana@example.net>
S: 250 2.1.5 <topbanana@example.net> recipient ok
C: RCPT TO:<Dana@Ivory.example.net> NOTIFY=SUCCESS,FAILURE
    ORCPT=rfc822;Dana@Ivory.example.net
S: 250 2.1.5 <Dana@Ivory.example.net> recipient ok
C: DATA
S: 354 okay, send message
C:  (message goes here)
C: .
S: 250 2.1.0 message accepted
C: QUIT
S: 221 2.0.0 goodbye
```

In the above example the MUA has specified the priority 3 and the server has accepted it. The server is advertising the MIXER Priority Assignment Policy (the default). Another variant of the initial submission might look like:

```
S: 220 example.com SMTP server here
C: EHLO mua.example.com
S: 250-example.com
S: 250-AUTH STARTTLS
S: 250-AUTH SCRAM-SHA-1 DIGEST-MD5
S: 250-DSN
S: 250-DELIVERBY
S: 250-ENHANCEDSTATUSCODES
S: 250 MT-PRIORITY
C: AUTH SCRAM-SHA-1
[...authentication exchange...]
S: 235 2.7.0 Authentication successful
C: MAIL FROM:<eljefe@example.com> BY=125;R ENVID=QQ314159
S: 250 2.1.0 <eljefe@example.com> sender ok
C: RCPT TO:<topbanana@example.net>
S: 250 2.1.5 <topbanana@example.net> recipient ok
C: RCPT TO:<Dana@Ivory.example.net> NOTIFY=SUCCESS,FAILURE
    ORCPT=rfc822;Dana@Ivory.example.net
S: 250 2.1.5 <Dana@Ivory.example.net> recipient ok
C: DATA
S: 354 okay, send message
C:  (message goes here)
C: .
S: 250 2.7.TBD3 3 is the new priority assigned to the message
C: QUIT
S: 221 2.0.0 goodbye
```

[[RFC Editor, please fix TBD3 in the example above.]] In the above example the MUA has not specified any priority, but the MSA has assigned priority 3 to the message. Also note that the server is unwilling to adverte the Priority Assignment Policy it supports in the EHLO response.

The MSA relays the message to the next MTA.

```
S: 220 example.net SMTP server here
C: EHLO example.com
S: 250-example.net
S: 250-DSN
S: 250-DELIVERBY
S: 250 MT-PRIORITY STANAG4406
C: MAIL FROM:<eljefe@example.com> BY=120;R ENVID=QQ314159
    MT-PRIORITY=3
S: 250 <eljefe@example.com> sender ok
C: RCPT TO:<topbanana@example.net>
S: 250 <topbanana@example.net> recipient ok
C: RCPT TO:<Dana@Ivory.example.net> NOTIFY=SUCCESS,FAILURE
    ORCPT=rfc822;Dana@Ivory.example.net
S: 250 <Dana@Ivory.example.net> recipient ok
C: DATA
S: 354 okay, send message
C:  (message goes here)
C: .
S: 250 message accepted
C: QUIT
S: 221 goodbye
```

The receiving SMTP server advertises support for the "STANAG4406" Priority Assignment Policy which supports 6 priority levels as described in Appendix A. This means that the server will use the priority value 4 internally (the next supported priority higher or equal to 3) and will communicate the priority value 3 when relaying it to the next hop (if necessary).

9. Deployment Considerations

9.1. Multiple MX records

If multiple DNS MX records are used to specify multiple servers for a domain in section 5 of [RFC5321], it is strongly advised that all of them support the MT-PRIORITY extension and handles priorities in exactly the same way. If one or more servers behave differently in this respect, then it is strongly suggested that none of the servers support the MT-PRIORITY extension. Otherwise, unexpected differences in message delivery speed or even rejections can happen during temporary or permanent failures, which users might perceive as serious reliability issues.

9.2. Priority Assignment Policies

This document allows up to 19 distinct priority values. In a particular operating environment independent originators need to assign priority values according to roughly the same criteria, so

that the same "high priority message" doesn't get associated with the value 3 for one sender and with the value 5 for another, as such messages might end up getting different preferential treatment when it was not the intent.

In order to achieve consistent behaviour in an operating environment, the Priority Assignment Policy (together with possible associated restrictions on maximum message sizes for each priority (if any), default timeouts, etc.) should be documented for the environment. Each SMTP/LMTP server supports a Priority Assignment Policy, whether explicit (advertised in the MT-PRIORITY EHLO response) or implict (not advertised). The default Priority Assignment Policy (assumed by the client when no Priority Assignment Policy name is advertised in the MT-PRIORITY EHLO response) is specified in Appendix B. Two other policies are specified in Appendix A and Appendix C. Additional policies SHOULD be registered with IANA as specified in Section 10.1.

Moreover, all MSAs/MTAs/MDAs within any given Administrative Management Domain has to be configured to use the same Priority Assignement Policy. Otherwise a differently configured MSA/MTA/MDA can expose the whole domain to possible attacks, like injection of hight priority fake-DSN.

When this SMTP extension is deployed across multiple cooperating Administrative Domains, such Administrative Domains need to use the same or at least compatible policies. Again, differences in policies (for example differences in how users are authenticated or differences in how priorities are handled) can expose an Administrative Domain to weaknesses in a partner domain.

10. IANA Considerations

This specification requests IANA to add the MT-PRIORITY SMTP extension to the "SMTP Service Extensions" registry (in http://www.iana.org/assignments/mail-parameters). This extension is suitable for the Submit port.

This specification requests IANA to add the following new Received header field clause to the "Additional-registered-clauses" sub-registry (in http://www.iana.org/assignments/mail-parameters) to help with tracing email messages delivered using the MT-PRIORITY SMTP extension:

Clause name: PRIORITY
Description: Records the value of the MT-PRIORITY parameter specified in the MAIL FROM command
Syntax of the value: See Section 7 of RFCXXXX
Reference: [[anchor12: RFCXXXX]]

This specification requests IANA to add the following Enumerated Status Codes to the "Simple Mail Transfer Protocol (SMTP) Enhanced Status Codes" registry established by [RFC5248] (in http://www.iana.org/assignments/smtp-enhanced-status-codes/smtp-enhanced-status-codes.xml):

1.

Code: X.7.TBD1

Sample Text: Priority Level is too low

Associated basic status code: 450, 550 (other 4XX or 5XX codes are allowed)

Description: The specified priority level is below the lowest priority acceptable for the receiving SMTP server. This condition might be temporary, for example the server is operating in a mode where only higher priority messages are accepted for transfer and delivery, while lower priority messages are rejected.

Reference: RFC XXXX

Submitter: A. Melnikov

Change controller: IESG

2.

Code: X.3.TBD2

Sample Text: Message is too big for the specified priority

Associated basic status code: 552 (other 4XX or 5XX codes are allowed)

Description: The message is too big for the specified priority. This condition might be temporary, for example the server is operating in a mode where only higher priority messages below certain size are accepted for transfer and delivery.

Reference: RFC XXXX

Submitter: A. Melnikov

Change controller: IESG

3.

Code: X.3.TBD3

Sample Text: Requested priority was changed

Associated basic status code: 250 or 251

Description: The message was accepted for relay/delivery, but the requested priority (possibly the implied default) was not honoured. The human readable text after the status code contains the new priority, followed by SP (space) and explanatory human readable text.

Reference: RFC XXXX

Submitter: A. Melnikov

Change controller: IESG

IANA is also requested to create a new IANA registry called "SMTP PRIORITY extension Priority Assignment Policy". Future registrations in this registry are governed by the "Specification Required" [RFC5226] IANA registration policy. Requirements on registrations (to be verified by the Designated Expert) are specified in Section 10.1. Changes to registrations undergo the same process as initial registrations. In cases of significant changes to registrations (other than editorial clarifications) the Designated Expert MAY require registration of a Priority Assignment Policy with a new name instead of updating the existing one.

10.1. Requirements on Priority Assignment Policy registrations

Priority Assignment Policy registrations with IANA are accompanied by a policy specification document, that MUST specify the following information:

1. The Priority Assignment Policy name, which is a case-insensitive string of 1 to 20 US-ASCII characters to be advertised as the MT-PRIORITY EHLO parameter. Allowed characters are: ALPHA, DIGIT, "-", "_" and "."

2. Number of distinct priority levels supported by all servers implementing the policy and their respective values.

3. For each supported priority level: default retry timeouts (how often to retry sending a message if there is a temporary error to transfer/deliver it). The policy specification can also explicitly define such information as implementation and/or deployment specific.

4. For each supported priority level: default expiration timeouts (how long to attempt transfer/delivery before the message expires and causes a non delivery report to be generated). The policy specification can also explicitly define such information as implementation and/or deployment specific. Note that a client can override such default when it uses additional SMTP extensions (such as the one mentioned in Section 5.2).

5. Maximum message size associated with each priority level. The policy specification can also explicitly define such information as implementation and/or deployment specific.

6. Any requirements/restrictions on what kind of SMTP client authentication is required in order for a SMTP server implementing this policy to accept priority values specified by a SMTP client. For example this can limit which SASL [RFC4422] authentication mechanisms are to be used, require TLS, etc.

7. Any other information that might affect processing of messages with different priorities.

8. Note that the policy specification document is not allowed to redefine the allowed range of priorities specified in Section 5 and other aspects of handling of different priorities, unless explicitly specified by this document.

10.2. Initial Priority Assignment Policy registrations

IANA is requested to register the following initial values in the "SMTP PRIORITY extension Priority Assignment Policy" registry:

Initial Priority Assignment Policy registrations

Policy Name	Reference	Comment
MIXER	Appendix B of RFCXXXX	Default policy
STANAG4406	Appendix A of RFCXXXX	
NSEP	Appendix C of RFCXXXX	

11. Security Considerations

Message Submission Agents ought to only accept message transfer priorities from users (or only certain groups of such users) who are authenticated and authorized in some way that's acceptable to the MSA. As part of this policy, they can also restrict maximum priority values that different groups of users can request, and can override the priority values specified by MUAs.

Similarly, MTAs ought to only accept message transfer priorities from senders (or only certain groups of such senders) who are authenticated and authorized in some way that's acceptable to the MTA. As part of this policy, they can also restrict maximum priority values that different groups of senders can request, and can override the priority values specified by them.

In the absence of the policy enforcement mentioned above an SMTP server (whether an MSA or an MTA) implementing this SMTP extension might be susceptible to a Denial of Service attack. For example, malicious clients (MUAs/MSAs/MTAs) can try to abuse this feature by always requesting Priority 9.

12. References

12.1. Normative References

[RFC2033] Myers, J., "Local Mail Transfer Protocol", RFC 2033, October 1996.

[RFC2034] Freed, N., "SMTP Service Extension for Returning Enhanced Error Codes", RFC 2034, October 1996.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3461] Moore, K., "Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)", RFC 3461, January 2003.

[RFC5321] Klensin, J., "Simple Mail Transfer Protocol", RFC 5321, October 2008.

[RFC5322] Resnick, P., Ed., "Internet Message Format", RFC 5322, October 2008.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for

Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

[RFC5248] Hansen, T. and J. Klensin, "A Registry for SMTP Enhanced Mail System Status Codes", BCP 138, RFC 5248, June 2008.

[RFC6409] Gellens, R. and J. Klensin, "Message Submission for Mail", STD 72, RFC 6409, November 2011.

12.2. Informative References

[RFC5598] Crocker, D., "Internet Mail Architecture", RFC 5598, July 2009.

[RFC1845] Crocker, D. and N. Freed, "SMTP Service Extension for Checkpoint/Restart", RFC 1845, September 1995.

[RFC1870] Klensin, J., Freed, N., and K. Moore, "SMTP Service Extension for Message Size Declaration", STD 10, RFC 1870, November 1995.

[RFC2156] Kille, S., "MIXER (Mime Internet X.400 Enhanced Relay): Mapping between X.400 and RFC 822/ MIME", RFC 2156, January 1998.

[RFC2205] Braden, B., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[RFC2852] Newman, D., "Deliver By SMTP Service Extension", RFC 2852, June 2000.

[RFC4190] Carlberg, K., Brown, I., and C. Beard, "Framework for Supporting Emergency Telecommunications Service (ETS) in IP Telephony", RFC 4190, November 2005.

[RFC4412] Schulzrinne, H. and J. Polk, "Communications Resource Priority for the Session Initiation Protocol (SIP)", RFC 4412, February 2006.

[RFC4422] Melnikov, A. and K. Zeilenga, "Simple Authentication and Security Layer (SASL)", RFC 4422, June 2006.

[RFC5228] Guenther, P. and T. Showalter, "Sieve: An Email Filtering Language", RFC 5228, January 2008.

[PRIORITY-TUNNELING] Melnikov, A. and K. Carlberg, "Tunneling of SMTP Message Transfer Priorities", draft-melnikov-smtp-priority-tunneling-00 (work in progress), 2012.

[ACP123] CCEB, "Common Messaging strategy and procedures", ACP 123, May 2009.

[STANAG-4406] NATO, "STANAG 4406 Edition 2: Military Message Handling System", STANAG 4406, March 2005.

[GREYLISTING] Kucherawy, M. and D. Crocker, "Email Greylisting: An Applicability Statement for SMTP", draft-ietf-appsawg-greylisting (work in progress), April 2012.

[RFC4125] Le Faucheur, F. and W. Lai, "Maximum Allocation Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering", RFC 4125, June 2005.

[RFC4127] Le Faucheur, F., "Russian Dolls Bandwidth Constraints Model for Diffserv-aware MPLS Traffic Engineering", RFC 4127, June 2005.

[RFC6401] Le Faucheur, F., Polk, J., and K. Carlberg, "RSVP Extensions for Admission Priority", RFC 6401, October 2011.

Appendix A. Priority Assignment Policy for Military Messaging

Military Messaging as specified in ACP 123 [ACP123] (also specified in STANAG 4406 [STANAG-4406]) defines 6 priority ("precedence") values. While ACP 123/STANAG 4406 allow for 32 different priority levels (16 levels are reserved for NATO and additional 16 are reserved for national use), only 6 are in use in practice. This section specified the Priority Assignment Policy for Military Messaging and how the MT-PRIORITY parameter can be mapped when

gatewaying between an SMTP and a ACP 123/STANAG 4406 environments.

Where SMTP is used to support military messaging, the following mappings SHOULD be used.

Recommended mapping of MT-PRIORITY values for MMHS

Priority value	MMHS Precedence name
-4	Deferred
-2	Routine
0	Priority
2	Immediate
4	Flash
6	Override

Table 1

The Priority Assignment Policy registration for Military Messaging is as follows:

1. The Priority Assignment Policy name is "STANAG4406".

2. Number of distinct priority levels: 6, as specified in the table above.

3. Default retry timeouts for each priority level are implementation and/or deployment specific.

4. Default expiration timeouts for each priority level are implementation and/or deployment specific.

5. Maximum message size associated with each priority level: implementation and/or deployment specific.

6. No restrictions on what kind of SMTP client authentication is required.

Appendix B. Priority Assignment Policy for MIXER

MIXER [RFC2156] defines the Priority header field with 3 values. This section specified the Priority Assignment Policy for MIXER and how the MT-PRIORITY parameter can be mapped when used with MIXER.

Where SMTP is used to support MIXER messaging, the following mappings SHOULD be used.

Recommended mapping of MT-PRIORITY values for MIXER

MIXER Priority value	MT-PRIORITY value
non-urgent	-4
normal	0
urgent	4

Table 2

The Priority Assignment Policy registration for MIXER is as follows:

1. The Priority Assignment Policy name is "MIXER".

2. Number of distinct priority levels: 3, as specified in the table above.

3. Default retry timeouts for each priority level are implementation and/or deployment specific.

4. Default expiration timeouts for each priority level are implementation and/or deployment specific.

5. Maximum message size associated with each priority level: implementation and/or deployment specific.

6. No restrictions on what kind of SMTP client authentication is required.

Appendix C. **Priority Assignment Policy for National Security /** Emergency Preparedness (NS/EP)

There are several forms of communication systems used during an emergency or disaster. The most well known form involves the many-to-one model of the general public contacting a public safety access point via 911/999/112 calls through the public telephone network. Typically, these calls do not require authorization, nor do they invoke any prioritization.

Another form of emergency communications involves a set of authorized users or nodes that use prioritized services to help established and continue communication given limited available resources. [RFC4190] includes descriptions of several systems that have been developed to support National Security / Emergency Preparedness (NS/EP). These deployed systems require a form of authentication and have focused on prioritization of telephony based services. They have also been

designed as a binary form (on/off) of signaled priority communications.

[RFC4412] includes examples of a more expansive view of NS/EP communications in which priority migrates from a single on/off bit value to one that comprises 5 priority values. This is shown in the cases of the ETS and WPS Namespaces. Given a lack of pre-existing NS/EP values assigned for email, we follow the paradigm of the ETS and WPS Namespaces and recommend 5 ascending values shown in the table below.

Priority value	Relational Order
-2	Lowest Priority
0	----------
2	----------
4	----------
6	Highest Priority

The Priority Assignment Policy registration for NS/EP is as follows:

1. The Priority Assignment Policy name is "NSEP".

2. Number of distinct priority levels: 5, as specified in the table above.

3. Default retry timeouts for each priority level are implementation and/or deployment specific.

4. Default expiration timeouts for each priority level are implementation and/or deployment specific.

5. Maximum message size associated with each priority level: implementation and/or deployment specific.

6. No restrictions on what kind of SMTP client authentication is required.

Appendix D. **Possible implementation strategies**

This appendix suggests some implementation strategies to implement the SMTP extension defined in this document. The list is not exhaustive.

This appendix and its subsections are Informative.

D.1. Probability

As the name suggests, probability involves increasing the chances of obtaining resources without adversely affecting previously established connections. One example would involve requesting resources set aside for specific priority levels. If these additional resources are exhausted, then the desired connection is denied. Queues, new timers, or combinations thereof can be used to facilitate the higher priority requests, but the key is that mechanisms focus on increasing the probability of message transfer.

D.2. Preemption of sessions or transactions

Preemption is a type of action that focuses only on a comparison of priorities to determine if previously established transactions need to be displaced in favor of higher priority requests. If no additional connection is possible, the client aborts a running session for emails with lower priority no later than directly after the current transaction. The client even can interrupt an active transaction and ought to do so, if other constraints such as delivery time (as specified in the DELIVERBY SMTP extension [RFC2852]) would be violated for the email with higher priority. When interrupting an active transaction, the client ought to take the total message size and the size of the transferred portion of the message being interrupted into consideration. This preliminary termination of sessions or transactions is called preemption.

If preemption of running transactions occurs, the client needs to choose a transaction with the lowest priority currently processed.

If the client has an option (i.e. it is supported by the next hop MTA) to interrupt transactions in a way that it can be restarted at the interruption point later, it ought to deploy it. An example for a mechanism providing such a service is the "SMTP Service Extension for Checkpoint/Restart" defined in [RFC1845].

If a client opts for the preemption of sessions instead of transactions, it needs to preempt the next session that reaches the end of a transaction.

D.3. Resource Allocation Models

Adding prioritization to a design moves the subject away from strictly best effort (and a first-come-first-served model) to one that includes admission control and resource allocation models. Over the years, a variety of work has been done within the IETF in specifying resource allocations models. Examples include the Maximum Allocation Model [RFC4125], the Russian Dolls Model [RFC4127], and

the Priority Bypass Model (Appendix A.3 of [RFC6401]).

While we recognize that these various models have been designed for other protocols (i.e., MPLS and RSVP), an understanding of their design characteristics may be beneficial in considering future implementations of a priority SMTP service.

In cases where the processing of high priority messages by an MTA is not considered negligible and exceeds engineered expectations, then operators managing that MTA may be notified in some form (e.g., pushed alarm, polled status).

Appendix E. Background on Design Choices

This Section provides some background on design choices made during development of the MT-PRIORITY SMTP extension.

The priority applies per message, rather than per recipient in order to keep the protocol simpler, and because of the expectation that it will be uncommon to need different priorities for different recipients on the same message. In cases where that is necessary, it can always be achieved by sending separate messages with the same content, segregating the recipients by desired message priority.

The choice of the priority range -9 to 9 (as opposed to, say, 1 to 6, or 0 to 9) was made after taking the following into consideration:

1. Clearly, having multiple priority levels is the whole point of this extension. Existing implementations of similar functionality in MTAs are already using three levels. One of the use cases motivating this extension requires 6 levels. So at least 6 different values are required.

2. During discussions of this extension, several different use cases were suggested that required differing numbers of priority levels. Defining just the 6 priority levels needed in item 1, above, would limit the extensibility for possible future use cases. Therefore, this document is defining a wider range, which allows implementations and deployments to add higher or lower priority levels and to insert additional priority levels between the recommended set of 6. This avoids the need to further extend this extension just to have a few more priority levels.

3. It seems natural to use 0 for the "normal" or default priority, rather than picking some non-zero number and having the priorities go up or down from there. This way, negative numbers always represent priorities that are lower than normal, with positive numbers as higher priorities.

Appendix F. Acknowledgements

This document copies lots of text from draft-schmeing-smtp-priorities-04.txt and draft-schmeing-smtp-priorities-05.txt. So the authors of this document would like to acknowledge contributions made by the authors of draft-schmeing-smtp-priorities: Michael Schmeing and Jan-Wilhelm Brendecke.

Many thanks for input provided by Steve Kille, David Wilson, John Klensin, Dave Crocker, Graeme Lunt, Alessandro Vesely, Barry Leiba, Bill McQuillan, Murray Kucherawy, SM, Glenn Parsons, Pete Resnick, Chris Newman, Ned Freed and Claudio Allocchio.

Special thanks to Barry Leiba for agreeing to shepherd this document.

Authors' Addresses

Alexey Melnikov
Isode Ltd
5 Castle Business Village
36 Station Road
Hampton, Middlesex TW12 2BX
UK

EMail: Alexey.Melnikov@isode.com

Ken Carlberg
G11
1601 Clarendon Blvd, #203
Arlington, VA 22209
USA

EMail: carlberg@g11.org.uk