The Syslog Protocol

Finnbarr P. Murphy

(fpm@fpmurphy.com)

I have used *syslog* for over 30 years now but other than knowing that it uses UDP and port 514, I have never looked at the underlying protocol in any detail.

Syslog is standardized by the IETF in RFC 5424

> This document describes the syslog protocol, which is used to convey event notification messages. This protocol utilizes a layered architecture, which allows the use of any number of transport protocols for transmission of syslog messages. It also provides a message format that allows vendor-specific extensions to be provided in a structured way.

This RFC does not define any transports. They are defined in other documents. One such transport is defined in RFC 5426 and is consistent with the traditional UDP transport.

Here is the ABNF (Augmented Backus–Naur Form) definition for a *syslog* message:

```
SYSLOG-MSG      = HEADER SP STRUCTURED-DATA [SP MSG]

HEADER          = PRI VERSION SP TIMESTAMP SP HOSTNAME
                  SP APP-NAME SP PROCID SP MSGID
PRI             = "< " PRIVAL ">"
PRIVAL          = 1*3DIGIT ; range 0 .. 191
VERSION         = NONZERO-DIGIT 0*2DIGIT
HOSTNAME        = NILVALUE / 1*255PRINTUSASCII

APP-NAME        = NILVALUE / 1*48PRINTUSASCII
PROCID          = NILVALUE / 1*128PRINTUSASCII
MSGID           = NILVALUE / 1*32PRINTUSASCII

TIMESTAMP       = NILVALUE / FULL-DATE "T" FULL-TIME
FULL-DATE       = DATE-FULLYEAR "-" DATE-MONTH "-" DATE-MDAY
DATE-FULLYEAR   = 4DIGIT
DATE-MONTH      = 2DIGIT  ; 01-12
DATE-MDAY       = 2DIGIT  ; 01-28, 01-29, 01-30, 01-31 based on
                          ; month/year
FULL-TIME       = PARTIAL-TIME TIME-OFFSET
PARTIAL-TIME    = TIME-HOUR ":" TIME-MINUTE ":" TIME-SECOND
                  [TIME-SECFRAC]
TIME-HOUR       = 2DIGIT  ; 00-23
TIME-MINUTE     = 2DIGIT  ; 00-59
TIME-SECOND     = 2DIGIT  ; 00-59
TIME-SECFRAC    = "." 1*6DIGIT
TIME-OFFSET     = "Z" / TIME-NUMOFFSET
TIME-NUMOFFSET  = ("+" / "-") TIME-HOUR ":" TIME-MINUTE

STRUCTURED-DATA = NILVALUE / 1*SD-ELEMENT
SD-ELEMENT      = "[" SD-ID *(SP SD-PARAM) "]"
SD-PARAM        = PARAM-NAME "=" %d34 PARAM-VALUE %d34
SD-ID           = SD-NAME
PARAM-NAME      = SD-NAME
PARAM-VALUE     = UTF-8-STRING ; characters '"', '' and
                               ; ']' MUST be escaped.
```

```
      SD-NAME         = 1*32PRINTUSASCII
                        ; except '=', SP, ']', %d34 (")

      MSG             = MSG-ANY / MSG-UTF8
      MSG-ANY         = *OCTET ; not starting with BOM
      MSG-UTF8        = BOM UTF-8-STRING
      BOM             = %xEF.BB.BF
      UTF-8-STRING    = *OCTET ; UTF-8 string as specified
                        ; in RFC 3629

      OCTET           = %d00-255
      SP              = %d32
      PRINTUSASCII    = %d33-126
      NONZERO-DIGIT   = %d49-57
      DIGIT           = %d48 / NONZERO-DIGIT
      NILVALUE        = "-"
```

Syslog message size limits are dictated by the syslog transport in use. There is no upper limit per se. Any transport receiver must be able to accept messages of up to and including 480 octets in length, should be able to accept messages of up to and including 2048 octets in length and may accept messages larger than 2048 octets in length. If a transport receiver receives a message with a length larger than it supports, it should truncate the message or it may discard the message.

I was not aware that facility and severity values are not normative. They are described in the RFC purely for informational purposes. Facility values must be in the range of 0 to 23 inclusive. Severity values must be in the range of 0 to 7 inclusive.

The RFC also defines an optional but useful set of structured data elements. For example the SD-ID *timeQuality* may be used by an originator to describe its notion of system time and should be written if the originator is not properly synchronized with a reliable external time source or if it does not know whether its time zone information is correct.

Interestingly, TCP port 512 is reserved not for *syslog* but for remote shells such as *rsh* and *remsh*.