RFC Formats
How hard could it be?

The first RFCs

TABLE OF CONTENTS.

I Transmission features
1. Transmission checking
2. HOST(A) to HOST(B) links.

II Functional software specifications.
1. User program - DEL language
2. Network program
3. Transmission Handler

III Link establishment procedure.
1. General procedure
2. Example

RFC 8

Remark:

Checksum calculation:

The last 16 bits of every message sent by a HOST is a checksum. This checksum is computed on the whole message including any marking, but excluding the 32 bit leader and any padding. To compute the checksum:

1. Consider the message to be padded with zeroes to a length of 8640 bits.
2. Section the 8640 bits into six 1440-bit segments, S_0, $S_1 \ldots S_5$.
3. Section each 1440-bit segment S into 90 16-bit elements, t_0, $t_1 \ldots t_{89}$.
4. Define a function $\oplus$, which takes two 16-bit elements as inputs and outputs a 16-bit element. This function is defined by

$$t_m \oplus t_n = t_m \oplus t_n, \text{ if } t_m + t_n < 2^{16}$$

$$t_m \oplus t_n = t_m \oplus t_n - 2^{16} + 1, \text{ if } t_m + t_n \geq 2^{16}$$

5. For each 1440-bit segment S_i compute $C_i = K(S_i)$, where

$$K(S) = t_0 \oplus t_1 \oplus \ldots\ldots t_{89}$$

6. Compute $C = C_0 \oplus C_1 \oplus C_1 \oplus C_2 \oplus C_2 \oplus C_2 \oplus C_2 \ldots\ldots \oplus C_5$

(Notice that $C_1 \oplus C_1$ is just C_1 rotated left one bit)

The number C is the checksum. The reason the C_i are rotated by i bits is to detect packet transposition.

RFC 11

Network Working Group
Request for Comments: 503
NIC: 15747
References: 349, 433

N. Neigus
J. Postel

13 April 1973

Socket Number List

This is a second pass at the specification of the socket numbers. More sockets have been made available for experimental functions, some new sockets have been assigned for network standard functions, host specific functions and experimental functions, and the list of available sockets for server hosts has been updated.

Following are the current assignments for socket numbers used for public functions. Note that a socket number is a 40 bit quantity, the first 8 bits being the host specification, the next 24 bits being site specific, and the last 8 bits being user specified -- these last 8 bits are called the AEN [it stands for, cleverly enough, Another Eightbit Number].

For the assignments here the value in the host field ranges across al hosts (unless otherwise specified), the value in the 24 bit field is zero, and the AEN is indicated here.

General Assignments:

Sockets	Assignment
0 - 63	Network Wide Standard Functions
64 - 127	Host Specific Functions
128 - 223	Reserved for Future Use
224 - 255	Any Experimental Function

File and document formats

- **Line printer aka “plain ASCII”**
- **Postscript and PDF experiments**
- **Nroff**
- **Xml2rfc**

Line printer

- **Fixed pitch, formatted for 80x66 pages**
- **No formatting other than line and page breaks**

```
Internet Research Task Force (IRTF)                          J. Levine
Request for Comments: 5782                          Taughannock Networks
Category: Informational                                     February 2010
ISSN: 2070-1721


                    DNS Blacklists and Whitelists

Abstract

   The rise of spam and other anti-social behavior on the Internet has
   led to the creation of shared blacklists and whitelists of IP
   addresses or domains.  The DNS has become the de-facto standard
   method of distributing these blacklists and whitelists.  This memo
   documents the structure and usage of DNS-based blacklists and
   whitelists, and the protocol used to query them.

Levine                       Informational                    [Page 1]
```

Postscript and PDF

- **Done more or less as experiments**
- **Look mostly OK, but not archival quality**

source routing. [12] [13]

9 Summary

Along with the emergence of very high speed applications and media, resource management has become a critical issue in the Research Internet and internets in general. A fundamental characteristic of the resource management problem is allowing administratively ADs to interconnect while retaining control over resource usage. However, we have lacked a careful articulation of the types of resource management policies that need to be supported. This paper addresses policy requirements for the Research Internet. After justifying our assumptions regarding AD topology we present a taxonomy and examples of policies that must be supported by a PR protocol.

Nroff

- **The clay tablets of document formatters**
- **Can be highly structured, but almost never is**
- **Mostly used to do simple formatting**

```
.ce
An efficient method to publish ranges of IP addresses in the DNS
.ce
\%draft-levine-iprangepub-02

.in 3
.ti 0
Abstract

.fi
The DNS has long been used to publish lists of IPv4 address ranges in
blacklists and whitelists.  The size of the IPv6 address space makes
the \%entry-per-IP approach used for IPv4 lists impractical.  A new
```

Xml2rfc

- **Simple XML profile**
- **Includes metadata**
  - **(unlike all previous formats)**
- **Good: it's XML**
- **Bad: it's XML**

```
<middle>
 <section title='Introduction'>
   <t>For many years, the Domain Name
     System<xref target="RFC1034"/>
   <xref target="RFC1035"/> has
     been the Internet's de facto
     distributed database. Blacklists
     and whitelists of IPv4 addresses
     have been published in the DNS using
     a simple system adapted from
     rDNS<xref target='RFC5782'/>.
```

What's good about the current RFC format?

- Very stable over 30 years
- Easy to read on most computers
- Compact
- Easy to grep

What's bad about the current RFC format?

- No non-ASCII author names
- No non-ASCII anything else
- No graphics beyond ASCII art
- Not reflowable for small screens
- Arguably ugly and hard to read

What else is bad?

- No metadata beyond implicit text scraping
  - Accessibility: Metadata tells screen readers what to read
- No links
  - Within document
  - To other documents (other than refs)
  - Links to sections of documents

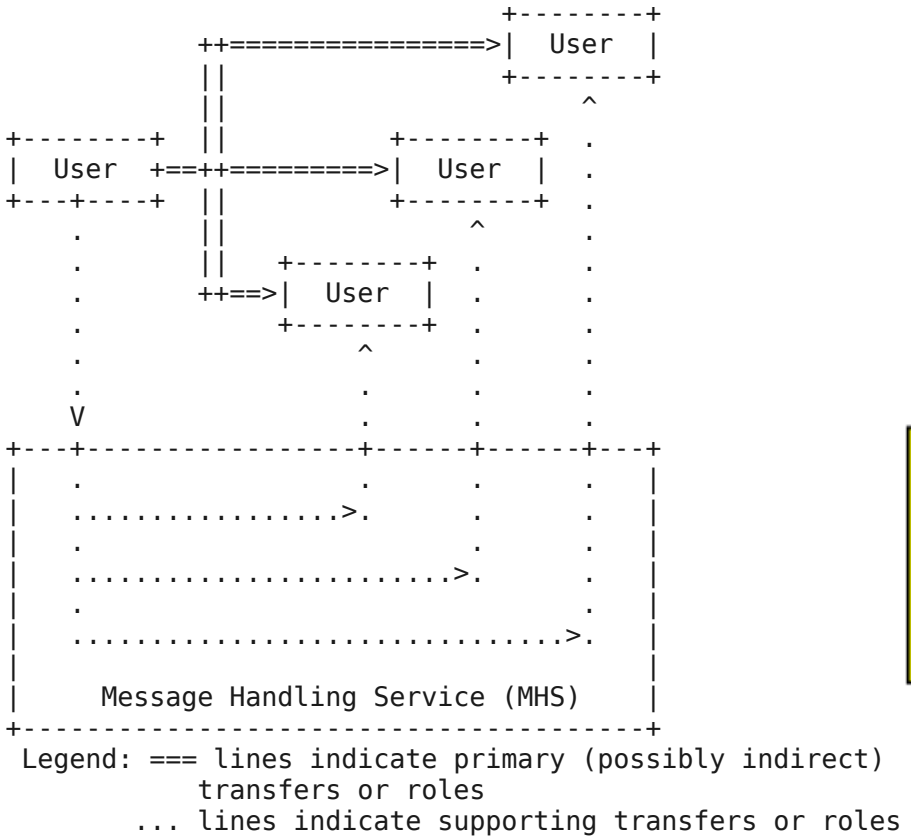
What else is bad?

- **Page breaks interrupt figures and text**

```
--AaB03x
content-disposition: form-data; name="field1"
content-type: text/plain;charset=windows-1250
content-transfer-encoding: quoted-printable

Masinter                    Standards Track                    [Page 3]

RFC 2388                  multipart/form-data                August 1998


Joe owes =80100.
--AaB03x
```

What else is bad?

- Time spent fiddling text and figures for printer image formatting



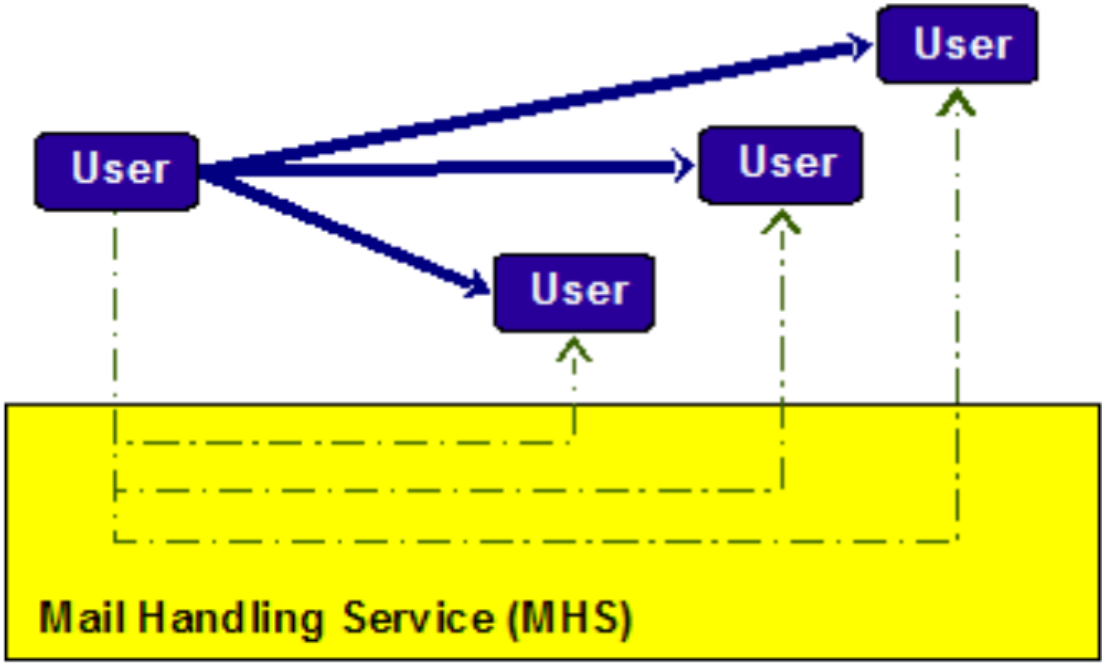


Figure 1: Basic Internet Mail Service Model

RFC 5598

Things we don't want to lose

- **Volunteer authors like their tools**
- **Freely available tools**
- **Lifetime of decades at least**

RSE requirements

- **Universally readable (incl. dumb terminals)**
- **Stable and archival**
- **Easily converted to other formats**
- **Something more flexible than ASCII art**
- **Respects international character set particularly for author names**