                    E-mail system for remote sites connected by expensive, intermittent link

Status of this Memo

This memo defines an Experimental Protocol for the Internet community. This memo does not
specify an Internet standard of any kind. Discussion and suggestions for improvement are
requested. Distribution of this memo is unlimited.

Table of Contents

1   Overview

The research fleet (and other remote sites) has grown dependent up effective Internet
communications for operational and maintenance support as well as direct and indirect support of
science. Historically, this need has been met with various communication channels including
PCM modem over satellite, modems over voice channels of INMARSAT, cell phone and other
means. Historically, individual operating institutions have implemented their own systems. Over
time the volume of traffic has grown and the transmission cost have become significant at a time
when the resources to pay the communication bills are harder to find.

The discussion is intended to provide improved e-mail service for remote sites (such as ships in the research fleet) where the transport channel is expensive, slow and intermittent. The procedure provides a number of features that are intended to enhance the functionality while reducing the total volume of traffic moved between shore and remote sites. These features include: shore-side message filtering based on a collection of criteria; according to a profile that can be securely manipulated by the remote user to minimize unwanted messages. It provides address validation prior to message transmission to eliminate bounced messages. Automatic "set-aside" filtering is provided to prevent un-intented transmission of large messages. Per user message size (but not content) is provided. Compression and batching of transfers are supported.

2   Practical Problems to be solved

This effort is motivated by the recognition that more effective use of our communication resources can be achieved by applying a collection pragmatic, standards-based methods.

2.1   High cost of e-mail to/from remote locations

E-mail traffic between the shore-side Internet and ships in the research fleet is largely handled over voice and/or data channels of the INMARSAT system. The per-connect costs are relatively large but more importantly are a highly visible, discrete and recognizable fraction of the annual cost of running a research vessel. In general, these costs are much more visible than the costs of our shore-side Internet connection costs. The costs per unit time and per volume transferred are higher and often represent a much smaller user base.

2.2   Accounting and cost recovery

There has been a significant amount of discussion about how to get the

2.3   Effective use of INMARSAT channels

3   What makes shipboard e-mail so different?

3.1   Expensive Communication channels

3.2   An e-mail dependent user base

3.3   Low bandwidth

3.4   Intermittent connectivity

3.5   Unreliable channels

3.6   Latency of satellite channels

4   Potentially helpful techniques

The purpose of this RFC is to focus discussion on particular problems in emerging portion of the Internet that is only occasionally connected to the Internet such as remote field stations and research ships. No proposed solutions in this document are intended as standards for the Internet. Rather it is hoped that a working consensus and a test-bed implementation will emerge

among those who provide connectivity to such services leading to adoption of effective standards and practice.

4.1   Effective use of mail aliases

Substantial performance improvements can be achieved by creation and user awareness of e-mail distrubution aliases (or list servers) at both ends of the link. The final implementation of this system requires shore-side knowledge of all valid accounts at the remote sites. The same mechanism can be used to provide access to appropriate e-mail lists and replication capabilities.

4.2   Guidance on effective use of e-mail

4.3   Suppression of multiple copies

During the filtering and batching process, multiple recipenets of the same message will share a common text body. This will probably require re-writing some e-mail address headers. [What if any existing e-mail protocols and/or systems implement this now?]

4.4   Batch message transfer

All mesasge traffic that is actually targeted for transfer to or from a remote site will be batched into a small number of data files prior to establishing the link to the remote station. All validation will be done prior to final compression and transfer of the data. During the construction of the batch file(s), log records will be generated covering the remote account (source or destination) and total file size for link accounting purposes.

4.5   Batch compression

Batched traffic will be compressed using a standard, high-efficiency batching application such as that implemented in 'gzip' (RFC: -------). The selected compression scheme will be efficient, operating system independent and robust.

4.6   Interrupted transfer recovery

4.7   Performance monitoring

End to end througput wll be logged with UTC start time and other relevant information (TBR) to provide raw data for characterizing and reporting changes in link performance. Routine traffic reports and exception reports will be geneated and propagated via e-mail.

4.8   Addressee validation

The implementation will prevent the generation of un-deliverable e-mail messages. In common practice today, such messages are transferred to the remote site, found to be undeliverable and sent back ashore. This results in moving the data across the link twice for no useful purpose and exacerbates the already extended turn around time.

To achieve this goal, all manipulation of user accounts and e-mail addresses at the remote site will generate a change-record to be tranmitted to the address validation site ashore. Addressee vaildation filtering will be implemented prior to including each message in the outbound batch to insure that no incorreclty addressed messages are routed to the remote site.

Incorrectly addressed messages that come out of the remote stite toward shore will be trapped and flagged upon returning to the shore-side mail router. The sender (and system postmaster) will be notified of the event, but the body of the offending message will not normally be re-transmitted to the ship.

4.9   Recipient defined message filtering

A method will be provided (perhaps based on 'procmail') to allow individual remote users to establish and modify the filter parameters in effect at the shore side relay site to control what e-mail messages are passed on to them.

A number of different implementation will be expolored that range from no filtering at all, through relatively detailed source/size/content filtering through "the first n-bytes of every message" up to no messages at all.

4.10   Message size filtering

Provision will be made for the system operator to implement a finite size limit for routinely transmitted e-mail messges. Files in excess of the lmit will be set aside and held for some period. Large files will also generate an automatic message to the sender, the recipient and the postmaster on both ends indicting that an over-size file was sent, where it has been temporarily stored and who to contact if the file should be transferred.

4.11   Content caching

A method will be developed to provide for "broadcast" like message transfer to remote sites for various types of information that are of interest to the community. For the research fleet this might cover the funding agency announcements, Lost Bouy Network, UNOLS, DESC, Ridge, RVOC and RVTEC announcements.

At the remote site this content could be maintained on a Web server and distributed via e-mail aliases or list servers.

4.12   Detailed logging

5   Current practices

6   Areas for future collaboration

7   Status Report

In response to the need for maintenance of curren information about the status and progress of various projects in the Internet community, this RFC is issued for the benefit of community members. The information contained in this document is accurate as of the date of publication, but is subject to change. Subsequent RFCs will reflect such changes.

8   References
PPP or UUCP or TCP/IP for connectivity

SMTP, POP3, IMAP for e-mail

HTTP for web form interface to filtering setup

????/

9   Security Considerations

There may be some concerns with regard to filtering e-mail and/or monitoring message traffic size and destination user.

There may be some concern about re-writing address headers but it happens every day now.

10   Author's Address

This document is the work of many contributors. The corresponding author is:

Dale Chayes
Instrument Lab
Lamont-Doherty Earth Observatory of Columbia University
61 Route 9W
Palisades, NY 10964

Phone: 914-365-8434
Fax: 914-359-6940
Email: dale@ldeo.columbia.edu

11   Relation to other standards

Updates: None

Obsoletes: None