Mail Servers with Embedded Data Compression Mechanisms

Alka Nand
Paging Systems Group
Motorola Inc.
Fort Worth , TX 76137
alka_nand@ftw.paging.mot.com

Tong Lai Yu
Department of Computer Sc.
California State University
San Bernardino, CA 92407
tongyu@csci.csusb.edu

Typically, e-mail messages are moved across the Internet using the Simple Mail Transfer Protocol (SMTP) which utilizes the connection-oriented Transmission Control Protocol (TCP) to establish connections between two mail servers. The POP3 (Post Office Protocol) is used to retrieve mails for individual users from a server. We designed and implemented e-mail servers that contain embedded data compression mechanisms; the SMTP protocol is extended to allow for the mail client and server to negotiate compression which is transparent to the users and the new servers are backward-compatible with traditional mail servers. Currently, the LZSS compression algorithm is used to carry out the data compression. Different kinds of mail data were used to test the E-mail system. Textual data, binary data, and graphical data were transported across the Internet using the designed E-Mail system. Several Windows NT hosts were identified for this experiment. These hosts were connected with Internet. On each of the hosts, two E-Mail systems were installed - the implemented E-Mail system with embedded compression, and a standard E-Mail system. Mail data was sent from one host to another. First the mail data were transferred from one host to another using the standard E-Mail system on the receiving end, and the designed E-Mail system on the transmitting end. The uncompressed mail data was transmitted in this case. The designed E-Mail system displays the time taken to transmit the mail data using 'time' system call. This time was noted. Then the designed E-Mail system with embedded data compression was started at both the hosts and the same mail data were transported between the same hosts. The compressed mail data were transmitted this time. The time taken in the transfer was again noted.

File Name	**File Type**	**Size**	**Avg Time on Traditional E-Mail System (1)**	**AvgTime on Designed E-Mail System (2)**	**Transmission Time Ratio (2) / (1)**
mbox.cpp	Text File	45k	63 sec	41 sec	65%
msrvr.exe	Executable	449k	527 sec	379 sec	72%
thesis.doc	MS-Word	288k	317 sec	190 sec	60%
rfc.txt	Text File	76k	81 sec	38 sec	47%
excite.htm	HTML Text	13k	45 sec	25 sec	55%
alska.htm	HTML Text	26k	160 sec	91 sec	57%
ftr.dll	Binary DLL	256k	276 sec	220 sec	80%
res.001	Binary	317k	332 sec	235 sec	71%
back.pcx	Graphics	65k	68 sec	38 sec	56%
tt25.res	Binary	1,085k	1345 sec	672 sec	50%
Average		***260k***	***309 sec***	***185 sec***	***59%***