

# **E-MAIL OVER LOW BANDWIDTH NETWORKS**



University of Moratuwa, Sri Lanka.  
**WASALA W.A.L.I**  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

**This dissertation was submitted to the department of Computer Science and Engineering of University of Moratuwa in partial fulfillment of the requirements for the Degree of Master of Science in Computer Science.**

**Department of Computer Science and Engineering**

**University of Moratuwa**

**December 2008**

The work presented in this dissertation in part or whole has not been submitted for the fulfillment of any other academic qualification at any institution.



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

.....  
W.A.L.I Wasala  
Candidate

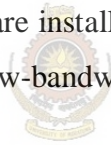
.....  
Prof. Gihan Dias  
Supervisor

## Abstract

Many Sri Lankans still use low bandwidth connections (e.g. dial-up, GPRS) to internet due to the high cost and limited geographical coverage of broadband internet (e.g. ADSL, 3G) services. Email access over this type of network is painful due to the large amount of time taken to download messages, especially when emails with large attachments and images are received. Bandwidth is also wasted for downloading spam and less important emails.

The objective of this research is to minimize the delays of reading email over a low bandwidth link. It also aims to reduce the load on the low bandwidth connection, thereby improving the performance for other users and applications.

This thesis presents a proxy based solution for accessing email using the IMAP protocol, thereby minimizing changes to the email server and client. Two IMAP proxy servers are installed between the mail server and the client, upstream and downstream of the low-bandwidth link.



University of Moratuwa, Sri Lanka  
Electronic Theses & Dissertations  
www.lib.mrt.ac.lk

A user defines priorities for email, and specifies whether all, part or none of each incoming message should be pre-fetched to the local email client.

The upstream IMAP Proxy processes the IMAP responses according to the set priorities, compresses them using zlib library and sends them to the downstream IMAP proxy which is in the client computer. The downstream IMAP proxy in turn decompresses the emails and sends them to the email client. Finally, the mailbox of the email client is populated according to the priorities set by the user. In this process instead of downloading all full emails to the client, a mix comprising of full emails, the 1<sup>st</sup> KB of emails and headers only are downloaded saving bandwidth.

Bandwidth is further saved by compressing image files.

A GUI application is provided for the user to configure the filter rules.

The performance of the solution was measured and shows that although the system imposes an average delay of 0.18s for full text emails and 0.33s delay for emails with JPEG attachments on a high-bandwidth network, it significantly improves both response time and bandwidth utilization on a low bandwidth network. (Average 25s reduction of response times for full text emails and 45s for emails with JPEG attachments. An average of 73% of the bandwidth is saved for full text emails and for emails with JPEG images, average of 83% of the bandwidth is saved.) Moreover email filtering and prioritizing also helps improving the response times of downloading emails. (It only takes about 0.40s and 0.39s to download 1<sup>st</sup> KB and headers of a non important email respectively.)



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## Acknowledgement

I would like to thank Prof. Gihan Dias, the project supervisor for the invaluable supervision and guidance given throughout the research project. Especially for guiding towards project objectives and suggesting alternative methods which helped me a lot in implementing the solution. I would also thank Prof. Gihan Dias for nominating me for the international workshop on *Optimization Technologies for Low Bandwidth Networks* held in ICTP, Trieste, Italy. It enabled me to gain valuable knowledge about open source technologies that can be used to manage bandwidth which contributed a lot to the research. My gratitude also goes to Dr. Enrique Canessa - ICTP and INASP for giving financial assistance and selecting me for the workshop.

I am greatly indebted to Ms. Vishaka Nanayakkara for her guidance on how to do research projects and the comments during progress reviews. I am also grateful to all the other staff members and my colleagues for their comments and help for the project.

I should also mention the invaluable support given to me by Mr. Dinesh Saparamadu – CEO of hSenid Software International in doing my research.

I would also like to thank Mr. Sanath Jayalath from School of Computing, University of Colombo for the great help in proof reading my Thesis.

Finally gratitude goes to my parents, spouse and brother for the tremendous support given throughout the project.

# Table of Contents

ABSTRACT .....	II
ACKNOWLEDGEMENT .....	IV
TABLE OF CONTENTS.....	V
LIST OF FIGURES .....	VII
LIST OF TABLES.....	IX
<b>1. INTRODUCTION .....</b>	<b>1</b>
1.1 MOTIVATION.....	1
1.2 OBJECTIVE .....	3
1.3 TARGET GROUP.....	4
1.4 SOLUTION .....	4
1.5 PERFORMANCE OF THE SOLUTION .....	4
1.6 ORGANIZATION OF THESIS .....	5
<b>2. RELATED WORK.....</b>	<b>6</b>
2.1 GENERAL CONSIDERATIONS FOR BANDWIDTH MANAGEMENT .....	6
2.2 RFC 4978 - THE IMAP COMPRESS EXTENSION [6].....	7
2.3 ON-THE-FLY INTER-PROXY DATA COMPRESSION FOR WEB ACCESS [2] .....	8
2.4 A MARKET-BASED APPROACH TO CONTROL WEB BANDWIDTH USAGE [4].....	10
2.4.1 System Components: .....	10
2.4.2 System Algorithm .....	11
2.5 A GENERAL PURPOSE PROXY FILTERING MECHANISM APPLIED TO THE MOBILE ENVIRONMENT [3].....	11
2.6 TEXT / GRAPHICS AND IMAGE TRANSMISSION OVER BANDLIMITED LOSSY LINKS [9] .....	13
2.6.1 Primitive Based Approach.....	14
2.6.2 Improved Bitmap using virtual frame buffer.....	14
2.6.3 Hybrid method .....	14
2.7 A UNIFIED HEADER COMPRESSION FRAMEWORK FOR LOW-BANDWIDTH LINKS [10] .....	15
2.8 USING PREDICTIVE PREFETCHING TO IMPROVE WORLD WIDE WEB LATENCY [11] .....	17
2.9 INVESTIGATION OF A PREFETCH MODEL FOR LOW BANDWIDTH NETWORKS[12].....	18
<b>3. TECHNOLOGIES USED.....</b>	<b>21</b>
3.1 INTERNET MESSAGE ACCESS PROTOCOL (IMAP) .....	21
3.1.1 IMAP Vs POP.....	22
3.1.2 IMAP Protocol Architecture.....	22
3.1.3 Bandwidth Optimizations of IMAP.....	26
3.2 DATA COMPRESSION .....	29
3.2.1 Lossless Compression .....	29
3.2.1.1 Statistical Modeling Algorithms (Dictionary based) for TEXT .....	30
3.2.1.2 Entropy Encoding Algorithms.....	32
3.2.1.3 DEFLATE Algorithm .....	35
3.2.1.3 Zlib[8].....	37
3.2.2 Lossy Compression .....	38
3.2.2.1 JPEG.....	38
3.3 TWISTED FRAMEWORK .....	44
3.4 PYTHON MIME HANDLING PACKAGE -EMAIL .....	46
<b>4. SYSTEM ARCHITECTURE.....</b>	<b>48</b>
4.1 APPROACH.....	48
4.1.1 SMTP Server.....	50
4.1.2 Mail Filter (Tagged Message Delivery Agent - TMDA).....	51
4.1.3 IMAP Server .....	51
4.1.4 IMAP Proxy server/client .....	52
4.1.5 Email client.....	52
4.1.6 User Interface (Rule Generator).....	53
4.2 SYSTEM OVERVIEW .....	53
4.2.1 Priorities and Prefetching .....	54

4.3 SYSTEM ALGORITHM.....	55
4.3.1 Email Filtering with Tagged Message Delivery Agent (TMDA).....	55
4.3.1.1 Filter Rule File .....	58
4.3.1.2 FilterParser.py .....	60
4.3.1.3 Filter Rule Generator - UI.....	62
4.3.2 Email Fetching.....	64
4.3.2.1 Client side prefetching operation when email client is started and connected to network.....	64
4.3.2.2 IMAP Proxy Server prefetching operation when email client is started and connected to network .....	65
4.3.2.3 Client Side Operation when user needs to read a particular email .....	68
4.3.2.4 IMAP Proxy Server Operation when user needs to read a particular email .....	70
4.3.2.5 Modified BINC IMAP Server .....	71
4.3.3 IMAP Proxy Server.....	71
4.3.4 IMAP Proxy Client.....	74
4.3.5 Email client - Thunderbird .....	75
<b>5. SYSTEM INSTALLATION AND CONFIGURATION.....</b>	<b>76</b>
5.1 PREREQUISITES .....	76
5.1.1 Server Requirements .....	76
5.1.2 Email Client Requirements.....	77
5.2 TMDA INSTALLATION AND CONFIGURATION .....	77
5.3 BINC IMAP IMPLEMENTATION AND CONFIGURATION.....	77
5.4 IMAP PROXY SERVER INSTALLATION .....	78
5.4 IMAP PROXY CLIENT INSTALLATION.....	78
5.5 THUNDERBIRD INSTALLATION AND CONFIGURATION.....	78
<b>6. RESULTS .....</b>	<b>80</b>
6.1 EXPERIMENTAL SETUP .....	80
6.2 RESULTS .....	83
6.2.1 Bandwidth Utilization.....	83
6.2.2 Response Time .....	88
6.2.3 Load on server and client.....	93
<b>7. CONCLUSION AND FUTURE WORK.....</b>	<b>95</b>
7.1 OTHER MERITS .....	96
7.2 FUTURE WORK.....	96
<b>APPENDIX 1: A TYPICAL IMAP SESSION .....</b>	<b>97</b>
<b>APPENDIX 2: TMDA FILTERPARSER.PATCH.....</b>	<b>99</b>
<b>APPENDIX 3: THUNDERBIRD NSMSGDBVIEW.PATCH .....</b>	<b>102</b>
<b>APPENDIX 4: BINC IMAP OPERATOR-FETCH.PATCH.....</b>	<b>102</b>
<b>REFERENCES .....</b>	<b>103</b>
<b>ADDITIONAL READING .....</b>	<b>105</b>

## List of Figures

Figure 1.1: Proxy based solution for web browsing over low bandwidth link .....	2
Figure 2.1: IMAP session for COMPRESS extension with TLS encryption .....	8
Figure 2.2: System Overview [2] .....	9
Figure 2.3: B. Zenel - A general purpose proxy filtering mechanism [3] .....	12
Figure 2.4: Overall Usage Model of the Solution .....	15
Figure 2.5: Example code for field description.....	16
Figure 2.6: Time Durations in [12].....	19
Figure 3.1: IMAP State and Flow diagram.....	24
Figure 3.2: Typical IMAP session between an IMAP server and a client. Here S: is server initiated communication and C: is client initiated communication.....	25
Figure 3.3: BODY[HEADER] fetch of a message.....	27
Figure 3.4: BODY[HEADER.FIELDS] fetch of a message.....	27
Figure 3.5: BODY[TEXT]<partial> fetch of a message .....	27
Figure 3.6: BODYSTRUCTURE fetch of a message .....	27
Figure 3.7: BODY[MIME] fetch of a message.....	28
Figure 3.8: BODY[MIME Part] fetch of a message, above is a Jpeg image .....	28
Figure 3.9: RFC822.SIZE fetch of a message.....	28
Figure 3.10: LZ77 Sliding Window concept .....	30
Figure 3.11: LZW compression algorithm.....	32
Figure 3.12: LZW Decompression Algorithm.....	32
Figure 3.13: Huffman Coding an example.....	33
Figure 3.14: 8x8 block of a particular channel after Step 3.....	40
Figure 3.15: After shifting the gray values .....	41
Figure 3.16: After applying DCT on 8x8 subimage .....	42
Figure 3.17: Typical Quantization Matrix .....	43
Figure 3.18: Matrix after quantization.....	43
Figure 3.19: Zig-Zag Ordering.....	44
Figure 3.20: Typical Network Program which uses Twisted Framework.....	45
Figure 3.21: High level overview of Twisted Framework.....	46
Figure 4.1: Solution Components.....	50
Figure 4.2: System Overview.....	53
Figure 4.3: Mail Filtering Process with TMDA.....	57
Figure 4.4: Filter Rule file (Ignore Line numbers).....	59
Figure 4.5: Filter Rule Generator .....	63
Figure 4.6: Successful Message .....	63
Figure 4.7: System Algorithm – IMAP Proxy Server Side (Part1).....	65
Figure 4.8: System Algorithm – IMAP Proxy Server Side (Part 2).....	66
Figure 4.9: Original Image (Left) and Compressed Image (Right) with JPEG quality =10 by IMAP Proxy Server .....	67
Figure 4.10: Client side operation when a user needs to read an email .....	69
Figure 6.1: 100Mbps link - Amount of data transferred (Sizes of full priority 0 emails) – Baseline system vs. Modified System .....	83
Figure 6.2: 100Mbps link - Amount of data transferred – Baseline system vs. Modified System (Sizes of full priority 0 emails with JPEG attachments - Quality set to 50%) .....	84
Figure 6.3: Transferred Image Sizes – Baseline system vs. Modified System (Sizes of JPEG attachments with Quality set to 50%) .....	85



Figure 6.4: Low bandwidth network (56kbps) - Effective Compression of priority 1 emails with 1st KB .....	86
Figure 6.5: Low bandwidth network (56kbps)- Effective Compression of priority 2 emails with header only .....	87
Figure 6.6: 100Mbps link – Response Times – Baseline system vs. Modified System (Response times for downloading full priority 0 text emails) .....	88
Figure 6.7: Low bandwidth network (56kbps) – Response Times – Baseline system vs. Modified System (Response times for downloading full priority 0 text emails) .....	88
Figure 6.8: 100Mbps link - Response Times – Baseline system vs. Modified System (Response times for downloading full priority 0 emails with JPEG attachments - Quality set to 50%).....	90
Figure 6.9: Low bandwidth network (56kbps) - Response Times – Baseline system vs. Modified System (Response times for downloading full priority 0 emails with JPEG attachments - Quality set to 50%) .....	90
Figure 6.10: Low bandwidth network (56kbps)- Response Time to download 1st KB of priority 1 emails. ....	91
Figure 6.11: Low bandwidth network (56kbps) - Response Time to download header of priority 2 emails. ....	92
Figure 6.12: CPU usage of IMAP Proxy client – Peeks show cpu usage when an email is downloaded.....	93
Figure 6.13: Memory usage of IMAP Proxy client – Two metrics Working Set (Bottom graph) and Private Bytes (Top Graph) are graphed.....	94



University of Moratuwa, Sri Lanka.  
 Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)

## List of Tables

Table 3.1: Input Symbols with known probabilities .....	33
Table 3.2: Arithmetic Coding – First Iteration.....	34
Table 3.3: Arithmetic Coding – Second Iteration .....	35
Table 3.4: Arithmetic Coding – Third Iteration .....	35



University of Moratuwa, Sri Lanka.  
Electronic Theses & Dissertations  
[www.lib.mrt.ac.lk](http://www.lib.mrt.ac.lk)