

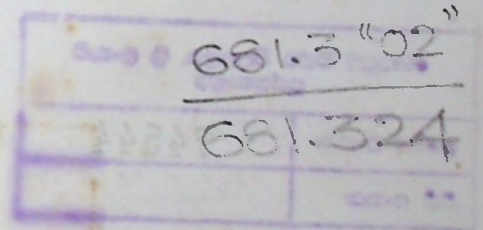
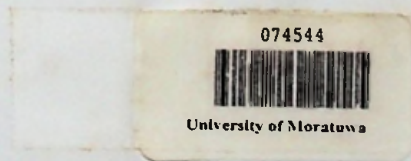
LB/DON/56/02

DCS 03/02

2

Design and Implementation of a Network Traffic Analyzing Utility

This thesis was submitted to the
Department of Computer Science & Engineering
of the
University of Moratuwa
in partial fulfillment of the requirements for the
Degree of Masters of Science



ප්‍රධාන කතෘ
මහත්මා විජයවික්‍රම, ටී ජයවර්ධන
මහාචාර්ය

74544

by

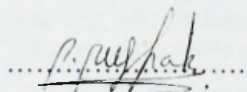
Priyantha Pushpa Kumara
Department of Computer Science & Engineering
University of Moratuwa
Sri Lanka
February 2002

TH

74544

Declaration

I, Priyantha Pushpa Kumara hereby certify that the work included in this thesis has not been submitted in part or whole for any other academic qualification at any institution.



Priyantha Pushpa Kumara

Research Student

UOM Verified Signature

.....

Dr. P G V Dias

Supervisor

Abstract

Developments in Information and Communication Technology (ICT) have created many new applications, which require a large amount of Internet bandwidth for proper application usage. However, Internet bandwidth is an expensive resource especially in this part of the world. Therefore, enterprises that use the Internet for business always need to efficiently use the bandwidth. Internet Service Providers may also need to manage their Internet bandwidth efficiently in order to increase the profit margin of their services.

Knowing the current (and past) usage is an important requirement in estimating the bandwidth requirement of a corporate network. The information about the composition of the use of the bandwidth by applications, user groups, etc. will help an administrator to efficiently manage it. Furthermore, knowing what is happening in the network always helps an administrator to keep the network secure, efficient and reliable all the time.

We studied the requirements of a network administrator by considering LEARN as a test bed for this study. We then assessed, whether the existing network traffic monitoring tools can provide the required information to the administrator. We examined some of the popular network monitoring tools that are widely used in the Internet community for their features and drawbacks. We found that the existing tools for network traffic monitoring are not capable of providing most of the required information by the administrator.

With these results, we identified the features available in those tools, and developed a new tool, LEARNStat, re-using some of the freeware utilities available in order to meet the requirements of a network administrator. We tested LEARNStat in LEARN which is our test-bed, and during the short period of running LEARNStat, we were able to obtain several important results.

In this thesis, we discuss the requirements of a network administrator and how we met those with LEARNStat. We present a brief description of related principles and also discuss the results we obtained with LEARNStat. The thesis also includes possible future enhancements for the LEARNStat.

Acknowledgement

I take pleasure in thanking Dr. Gihan Dias, for supervising this project and using his valuable time to contribute, guide and instruct me in not only revising and writing this thesis but also proposing, commenting and evaluating my work throughout the project. I would also like to express my sincere thanks to Dr. Nalin Wickramarachchi, head of the department of Computer Science & Engineering for his guidance and permission given to me, to use departmental resources for this work. My sincere thanks to Dr. Nalin Ranasinghe, the external supervisor, for his guidance and positive comments for the success of this project.

I am also grateful to the LEARN management committee for allowing me to carry out this project while working in LEARN as well as allowing me to use its resources for this project. Without the sponsorship, this work would not have been possible.

Thanks are also due to Prof. (Mrs.) Ratnayake, Director, Postgraduate Studies, for her guidance throughout the project.

I received valuable comments and guidance from Vishaka Nanayakkara, throughout the project that always encouraged me to successfully complete this project. Without her help, I would not have met Prof. Sven Tafvelin, who gave me valuable guidance as well as a contact for NorduStat. Thank you Vishaka and thank you Prof. Tafvelin.

Thanks to Rasika Amarasiri for his guidance and help given to me whenever I had any problems with Perl or HTML.

Thanks to all staff members of the Department of Computer Science & Engineering and network administrators of LEARN member sites who directly or indirectly helped me to successfully develop and test the LEARNStat.

Thanks to Dave Plonka and Håvard Eidnes for their kindness in allowing me to publish some of the graphs that represent their own network statistics in this thesis.

My special thanks to the final year project group, Karunarathna, Sendanayake, Ms. Wickramasinghe, & Ms. Jayawardana, those who first installed and configured NeTraMet and showed me the capabilities of NeTraMet.

Thanks to Mr. G. H. A. Perera for his help in proof reading and correcting the language mistakes.

Finally, the most important people, my wife Preethi and even little Dilsha endured everything with infinite patience and good grace, during this period. Thank you.



Contents

| | |
|---|-----------|
| ABSTRACT | I |
| ACKNOWLEDGEMENT | II |
| CONTENTS | III |
| LIST OF FIGURES..... | V |
| LIST OF TABLES..... | VII |
| ABBREVIATIONS..... | VIII |
| 1. INTRODUCTION | 1 |
| 1.1 OBJECTIVES..... | 1 |
| 1.2 LEARNSTAT | 2 |
| 1.3 OUTLINE OF THE THESIS | 3 |
| 2. LEARN | 4 |
| 2.1 HISTORY..... | 4 |
| 2.2 EXPERIENCE IN LEARN | 6 |
| 2.2.1 <i>Network Monitoring tools being used</i> | 6 |
| 2.2.2 <i>Problems encountered with tools used</i> | 7 |
| 2.2.3 <i>Summary</i> | 8 |
| 3. NETWORK MANAGEMENT..... | 9 |
| 3.1 ISO/OSI NETWORK MANAGEMENT STANDARDS | 9 |
| 3.2 INTERNET NETWORK MANAGEMENT..... | 10 |
| 3.2.1 <i>Simple Network Management Protocol-SNMP</i> | 10 |
| 3.2.2 <i>TCP/IP</i> | 14 |
| 3.3 NETWORK MONITORING | 18 |
| 3.3.1 <i>Which layer to monitor?</i> | 18 |
| 3.3.2 <i>Types of monitors</i> | 18 |
| 3.4 CASE STUDIES..... | 19 |
| 3.4.1 <i>MRTG v2.0 (Multi Router Traffic Grapher)</i> | 19 |
| 3.4.2 <i>FlowScan</i> | 21 |
| 3.4.3 <i>NeTraMet</i> | 24 |
| 3.4.4 <i>NORDstat</i> | 26 |
| 4. LEARNSTAT SYSTEM DESIGN AND ARCHITECTURE | 30 |
| 4.1 SYSTEM OVERVIEW..... | 30 |
| 4.2 LEARNSTAT SYSTEM DESIGN..... | 30 |
| 4.3 SYSTEM ARCHITECTURE..... | 32 |
| 4.4 USER INTERFACE..... | 33 |
| 4.4.1 <i>User Interface Design</i> | 39 |
| 5. LEARNSTAT IMPLEMENTATION..... | 40 |
| 5.1 SYSTEM COMPONENTS | 40 |
| 5.1.1 <i>LEARNStat Regular Module</i> | 40 |
| 5.1.2 <i>LEARNStat On-line Module</i> | 45 |
| 5.2 SYSTEM INTEGRATION DETAILS | 48 |
| 5.2.1 <i>Integrating Locally developed LEARNStat components</i> | 49 |
| 5.2.2 <i>Integrating Re-used components to LEARNStat</i> | 50 |

| | |
|---|-----------|
| 6. OUTCOME OF LEARNSTAT | 52 |
| 6.1 INTERESTING RESULTS | 52 |
| 6.1.1 <i>Discovery of unknown servers</i> | 52 |
| 6.1.2 <i>Discovery of unauthorized cache users</i> | 52 |
| 6.1.3 <i>Discovery of a Loop-Hole in Router Access-lists</i> | 53 |
| 6.1.4 <i>Monitoring Bandwidth Utilization by Applications</i> | 54 |
| 6.2 LEARNSTAT SYSTEM VALIDATION & VERIFICATION | 56 |
| 6.2.1 <i>LEARNStat Unit Testing</i> | 56 |
| 6.2.2 <i>LEARNStat Integration Testing</i> | 57 |
| 6.2.3 <i>LEARNStat System Testing</i> | 57 |
| 6.3 LEARNSTAT RESULT VALIDATION | 57 |
| 6.3.1 <i>Validation of LEARNStat Graphs</i> | 58 |
| 6.3.2 <i>Validation of non-graphical outputs</i> | 59 |
| 6.3.3 <i>Validation using known operations</i> | 61 |
| 6.4 PROBLEMS ENCOUNTERED | 62 |
| 6.4.1 <i>Problems in General</i> | 62 |
| 6.4.2 <i>Problems Encountered in Regular module</i> | 63 |
| 6.4.3 <i>Problems Encountered in On-line Module</i> | 66 |
| 7. CONCLUSION AND RECOMMENDATIONS..... | 68 |
| 7.1 MERITS AND DRAWBACKS OF LEARNSTAT | 68 |
| 7.1.1 <i>Merits of LEARNStat</i> | 68 |
| 7.1.2 <i>Drawbacks of LEARNStat</i> | 68 |
| 7.2 FUTURE ENHANCEMENTS POSSIBLE TO LEARNSTAT | 69 |
| 7.2.1 <i>A new Rules Syntax</i> | 69 |
| 7.2.2 <i>A new Algorithm to match Alarm Rules</i> | 70 |
| 7.2.3 <i>Sending urgent alarms</i> | 71 |
| 7.2.4 <i>Alarms on traffic volumes</i> | 71 |
| 7.2.5 <i>A GUI for Administrators</i> | 71 |
| 7.2.6 <i>Data Validation</i> | 72 |
| 7.2.7 <i>Code re-writing</i> | 72 |
| 7.3 LEARNSTAT IN SUMMARY..... | 72 |
| APPENDIX A..... | 73 |
| A.1 INSTALLING AND CONFIGURING RE-USED COMPONENTS | 73 |
| A.1.1 <i>Installing and Configuring NeTraMet</i> | 73 |
| A.1.2 <i>Configuring NetFlow on Cisco routers</i> | 74 |
| A.1.3 <i>Installing and Configuring RRDTools</i> | 74 |
| A.2 INSTALLING AND CONFIGURING LOCALLY DEVELOPED COMPONENTS..... | 74 |
| A.2.1 <i>Installation procedure</i> | 74 |
| A.2.2 <i>Configuration procedure</i> | 74 |
| APPENDIX B..... | 79 |
| B.1 LEARNSTAT DATA FORMATS | 79 |
| B.1 IDENTIFYING DATA FLOWS | 80 |
| B.2 FILE FORMATS | 80 |
| B.2.1 <i>Configuration File Format</i> | 80 |
| B.2.2 <i>Rule File Format</i> | 82 |
| REFERENCES | 85 |

List of Figures

| | | |
|-------------|---|----|
| FIGURE 2.1 | LEARN PHASE 1 NETWORK TOPOLOGY | 5 |
| FIGURE 2.2 | LEARN TOPOLOGY EXPECTED AT THE EARLY 2002 | 6 |
| FIGURE 2.3 | AN MRTG GRAPH SHOWING INTERNET BANDWIDTH USAGE | 7 |
| FIGURE 2.4 | AN MRTG GRAPH SHOWING A LOCAL LINK USAGE | 7 |
| FIGURE 2.5 | AN MRTG GRAPH SHOWING THE USAGE OF ONE OF THE LEARN BACKBONE LINKS | 8 |
| FIGURE 3.1 | OSI NETWORK MANAGEMENT ARCHITECTURE | 9 |
| FIGURE 3.2 | SNMP MANAGEMENT MODEL | 10 |
| FIGURE 3.3 | A PART OF ASN.1 OBJECT NAMING TREE | 12 |
| FIGURE 3.4 | TCP/IP PROTOCOL SUITE..... | 15 |
| FIGURE 3.5 | ADDING HEADERS AND TRAILERS TO APPLICATION DATA AS IT GOES DOWN THE TCP/IP PROTOCOL STACK..... | 15 |
| FIGURE 3.6 | IP DATAGRAM FORMAT..... | 16 |
| FIGURE 3.7 | IP ADDRESS AND NETMASK | 16 |
| FIGURE 3.8 | MRTG GRAPHS WITH DIFFERENT DATA RESOLUTIONS | 19 |
| FIGURE 3.9 | MRTG USES SNMP QUERIES TO GATHER DATA FROM A REMOTE ROUTER | 20 |
| FIGURE 3.10 | A GRAPH GENERATED BY RRDTOOLS..... | 21 |
| FIGURE 3.11 | FLOWSCAN GRAPH SHOWING WELL-KNOWN SERVICES | 22 |
| FIGURE 3.12 | NETFLOW VERSION 5 PDU FORMAT..... | 23 |
| FIGURE 3.13 | INTERACTION BETWEEN METERS AND COLLECTOR & MANAGER | 24 |
| FIGURE 3.14 | A SAMPLE FLOW DATA FILE..... | 25 |
| FIGURE 3.15 | NORDSTAT GRAPH SHOWING TRAFFIC RATE IN KBITS/S | 26 |
| FIGURE 3.16 | A NETWORK MAP PRODUCED BY NORDSTAT..... | 27 |
| FIGURE 3.17 | HOURLY AVERAGE GRAPH BY NORDSTAT..... | 27 |
| FIGURE 3.18 | SNAPSHOT OF TABULAR INFORMATION PRESENTED BY NORDSTAT..... | 28 |
| FIGURE 3.19 | GRAPH OBTAINED BY CLICKING THE FIRST LINK IN THE TABLE (FIGURE 3.18)..... | 28 |
| FIGURE 4.1 | LEARNSTAT SOFTWARE COMPONENTS..... | 31 |
| FIGURE 4.2 | LEARNSTAT LAYERED ARCHITECTURE | 32 |
| FIGURE 4.3 | SCREEN SHOT OF LEARNSTAT FRONT PAGE | 33 |
| FIGURE 4.4 | LEARNSTAT USAGE GRAPHS..... | 34 |
| FIGURE 4.5 | LEARNSTAT PORT SUMMARY SCREENS | 35 |
| FIGURE 4.6 | LEARNSTAT NETWORK USERS SCREENS..... | 36 |
| FIGURE 4.7 | SERVERS LISTED BY SERVICES THEY RUN | 37 |
| FIGURE 4.8 | RRGRAPHER SCREENS..... | 38 |
| FIGURE 4.9 | LEARNSTAT FRONT PAGE WITH AN ALARM | 38 |
| FIGURE 4.10 | LEARNSTAT USER INTERFACE SUMMARY | 39 |
| FIGURE 5.1 | THE LEARNSTAT.PL ALGORITHM IN BRIEF..... | 41 |
| FIGURE 5.2 | SUMMARY OF MAKEDB.PL PROGRAM LOGIC | 42 |
| FIGURE 5.3 | LOGIC OF MAKEGRAPHS.PL..... | 43 |
| FIGURE 5.4 | LOGIC OF SERVERS.PL..... | 43 |
| FIGURE 5.5 | SIMPLE LOGIC OF ALARMS.PL | 44 |
| FIGURE 5.6 | LEARNSTAT ON-LINE MODULE COMPONENTS AND THEIR INTERACTION WITH HTML PAGES..... | 47 |
| FIGURE 5.7 | A PART OF A SAMPLE HTML TEMPLATE FILE..... | 48 |
| FIGURE 5.8 | LEARNSTAT COMPONENTS AFTER INTEGRATION | 49 |
| FIGURE 5.9 | A SAMPLE BATCH JOB LISTING..... | 50 |
| FIGURE 5.10 | THE UPDATED BATCH JOB LISTING | 50 |
| FIGURE 6.1 | THE ALARM RULES FILE USED FOR BETA TEST | 53 |
| FIGURE 6.2 | ALARM DETAILS SCREEN | 53 |
| FIGURE 6.3 | LEARNSTAT GRAPH SHOWING VIDEO CONFERENCING BANDWIDTH USAGE | 54 |
| FIGURE 6.4 | NEW LINES ADDED TO THE RULES FILE FOR IDENTIFYING VIDEO CONFERENCING DATA FLOWS | 54 |
| FIGURE 6.5. | MIRROR SYNCHRONIZATION BANDWIDTH UTILIZATION..... | 55 |
| FIGURE 6.6 | A COMPARISON OF LEARNSTAT GRAPHS WITH MRTG GRAPHS..... | 59 |
| FIGURE 6.7 | LEARNSTAT DATA VALIDATION QUERY AND A RESULT | 60 |

| | | |
|-------------|--|----|
| FIGURE 6.8 | LEARNSTAT RESULT VERIFICATION USING A KNOWN OPERATION | 62 |
| FIGURE 6.9. | UDP STREAM OF CACHE-FLOW DATA FROM A ROUTER TO THE METER | 62 |
| FIGURE 6.10 | LEARNSTAT GRAPH THAT SHOWS A DELAY IN GETTING UPDATED | 63 |
| FIGURE 6.11 | LEARNSTAT BATCH JOB SCRIPT WITH TIME STAMPS | 63 |
| FIGURE 6.12 | LEARNSTAT.PL PERFORMANCE VARIATIONS..... | 65 |
| FIGURE 6.13 | A SCRIPT TO RUN LEARNSTAT.PL WITH TWO SETS OF RULES | 66 |
| FIGURE 7.1 | NEW ALARM RULE SYNTAX | 69 |
| FIGURE 7.2 | A NEW ALGORITHM FOR EVALUATING ALARM RULES..... | 70 |
| FIGURE A.1 | ALARM RULES AS SEEN FROM LEARNSTAT ON-LINE MODULE | 78 |
| FIGURE B.1 | LOGIC OF ATTRIBUTE BASED SIMPLE TEST | 81 |
| FIGURE B.2 | CONFIGURATION FILE FORMAT | 82 |
| FIGURE B.3 | FORMAT OF THE LEARNSTAT RULE FILE | 83 |
| FIGURE B.4 | FORMAT OF THE ALARM RULE FILE..... | 84 |

List of Tables

| | | |
|-----------|---|----|
| TABLE 3.1 | SNMP PRIMITIVE DATA TYPES..... | 11 |
| TABLE 3.2 | OBJECTS GROUPS OF MIB-2..... | 12 |
| TABLE 3.3 | IP ADDRESS EXAMPLES..... | 17 |
| TABLE 3.4 | SOME OF THE COMMON TCP/UDP PORTS ASSIGNMENTS..... | 17 |
| TABLE 5.1 | LEARNSTAT REGULAR MODULE COMPONENTS..... | 40 |
| TABLE 5.2 | LEARNSTAT ON-LINE MODULE COMPONENTS..... | 45 |
| TABLE 5.3 | LEARNSTAT PERL MODULE COMPONENTS..... | 50 |
| TABLE 6.1 | LEARNSTAT UNIT TEST RESULTS..... | 56 |
| TABLE A.1 | KEYWORDS USED IN LEARNSTAT CONFIGURATION FILE..... | 75 |
| TABLE B.1 | LEARNSTAT COMPONENTS WITH THEIR INPUTS AND OUTPUTS..... | 79 |



Abbreviations

| | |
|----------|---|
| CIDR | Classless Inter Domain Routing |
| CINTEC | Council for INformation TECnology |
| GUI | Graphical User Interface |
| IANA | Internet Assigned Number Authority |
| ICP | Internet Cache Protocol |
| IETF | Internet Engineering Task Force |
| IOS | Internet Operating System (Cisco uses this name for their router operating systems) |
| ISO | International Standard Organization |
| IX | Internet eXchange |
| LEARN | Lanka Education And Research Network |
| MIB | Management Information Base |
| NARA | National Aquatic Resources Agency |
| NARESA | Natural Resources Energy and Science Authority (Now, this is named National Science Foundation - NSF) |
| NORDUnet | This is the Nordic Internet highway to research and education networks in Denmark, Finland, Iceland, Norway and Sweden and provides the Nordic backbone to the global Internet. |
| PDU | Protocol Data Unit |
| RFC | Request For Comments |
| RPN | Reverse Polish Notation |
| RTFM | Real-time Traffic Flow Measurement |
| SMS | Short Message Service |
| SNMP | Simple Network Management Protocol |
| Sida | Swedish International Development Cooperation Agency |
| SAREC | The research arm of the Sida |