Development of a Web Availability Analyzer Software Tool

Ali Peiravi[1], Muhammad Sharaeini[2]

Ferdowsi University of Mashhad
Department of Electrical Engineering, School of Engineering, Mashhad IRAN
Telephone number: (0098) 511-881-5100
[1]Ali_peiravi@yahoo.com, [2]shahraeini@ferdowsi.um.ac.ir

Abstract: In this study, results of the development of a web availability analyzer software tool that has been designed in order to measure internet availability from the end user's perspective are reported. The measured results of the availability of local and international sites along with a comparison of results indicate the successful operation of the software tool. The main objective of this paper is to present the approach used to measure the actual availability of internet sites through the development and use of a Web Availability Analyzer software Tool (WATT). [Journal of American Science 2010; 6(6):89-95]. (ISSN: 1545-1003).

Key words: Web, Internet Availability, Software package

Introduction

The use of the internet has become so wide-spread that it covers almost every aspect of human life today. Acts such as banking, payment of bills, shopping, personal and family affairs such as e-mail and community memberships, etc. are relying on computers and the internet more and more. Therefore, the internet has become very vital in man's economic and social life.

Hence, the various issues related to internet and its performance have become of great interest to researchers. One important issue is internet reliability measured by its availability. There are both theoretical and experimental approaches to compute or measure internet availability. The focus of this research is the steps taken in the development of a software package used for the measurement and statistical analysis of actual availability of web sites on the internet.

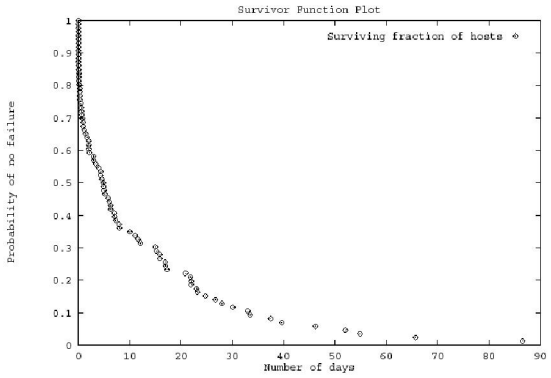

Fig 1 Time to failure extracted from wtmp file adopted from Sriram (1993)

Statistical approaches to measure the availability of the internet have been used since 1991. These approaches include analysis of wtmp file as shown in Fig.1, tattler, ICMP ping, etc.

Tattler System Monitoring Approach

Long (1992) proposed the Tattler network monitoring system as shown in Fig 2.

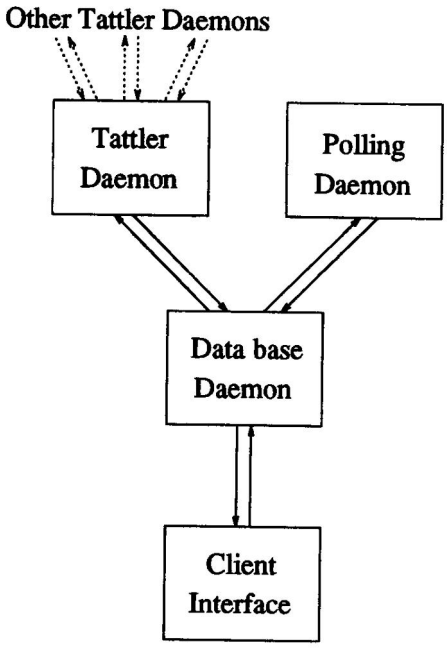

Fig 2 The structure of the Tattler system

Each tattler is composed of a client interface, a polling daemon, a data base daemon and a tattler daemon.

The tattlers are connected to hosts in several locations as shown in Fig 3. The tattlers making up a group are responsible for maintaining a list of hosts to be monitored plus preparing logs. The tattler daemon communicates the logged information to other tattlers so that a consistent log is kept across the network. The drawback of this approach is that the data obtained is invalid when the system itself or its communication lines fail.

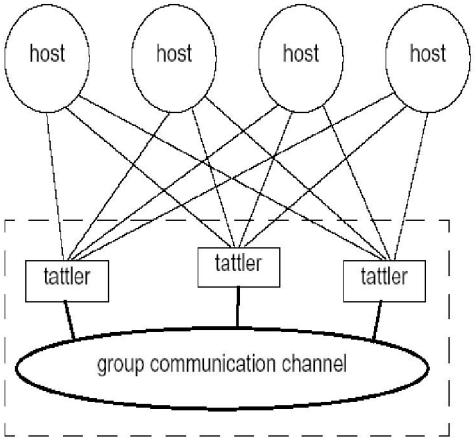

Fig 3 The structure of the Tattler system adopted from Long (1992)

ICMP Ping

ICMP ping may be used to check a remote host for its availability as shown in Fig 4. Local hosts should respond to such ping requests within a few milliseconds, but it may take them longer if there is heavy traffic on the network. Sending an ICMP Echo request and a consequent ICMP timeout may be used to indicate unavailability of the target site as perceived by the requesting system.

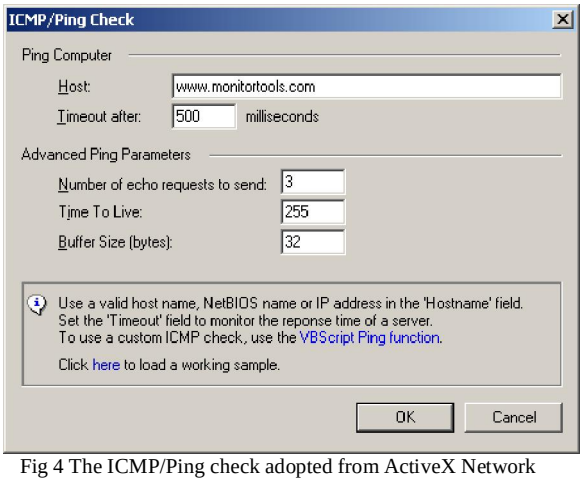

Fig 4 The ICMP/Ping check adopted from ActiveX Network Monitor

The Development of Web Availability Analyzer Tool Software

This software tool was developed using C# computer programming language. This tool is indeed a simulation of a browser that sends its requests to the various web sites and records the results. It analyzes the results at the end of each day. The main screen of the WAAT software tool is presented in Figure 5.

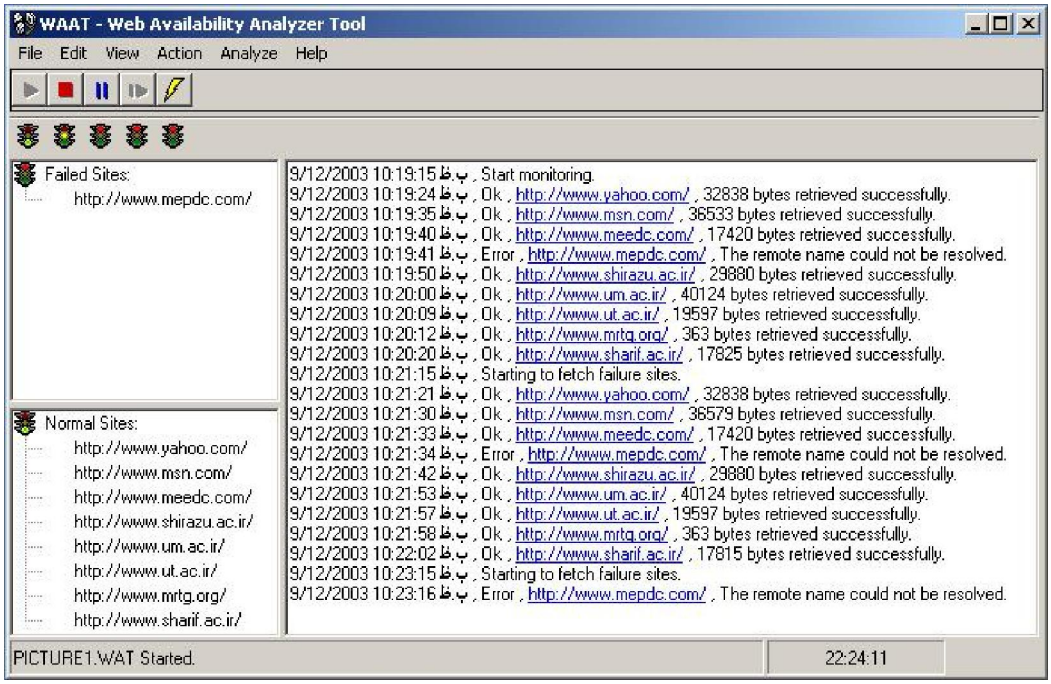

Fig 5 – The main screen of WAAT software

This software tool is composed of the following subsystems:
1) Main Menu
2) Action Bar
3) Internodes Bar
4) Status Windows
5) Log Window
6) Status Panel

The main menu consists of File, Edit, View, Action (Start, Stop, Pause, etc), Resume, Analyze, and Help submenus.

There is an optional internodes bar that shows the status of intermediate nodes being analyzed through the ICMP ECHO protocol. These nodes are analyzed 5 times in each ten minute period. If all the Ping operations are successful, then a green light is turned on. But if this operation is only successful one or two times, then a yellow light is turned on. Finally, a red light will be lit in case there is no response at all to the ping operations. The intermediate nodes are usually routers in the network.

Status Windows

Two status windows are designed in this tool to enable the user to recognize the condition of the current status of the internet sites under study. The upper window shows the functioning sites and the lower window shows the malfunctioning ones. These windows are updated once every time a response is obtained by the statistical engine.

Log Window

The actions taken by the software are constantly reported in the log window. Each line consists of date and time followed by a report of the action taken. This window is updated rapidly since the software package developed is multithread and does not depend on the traffic encountered by the statistical engine.

Status Panel

The name of the current project and its status is recorded on the left hand side of the status window. The time of the system used by the statistical engine is recorded on the right hand side.

Project definition may be started by choosing the New Project submenu in the File menu as shown in Fig 6.

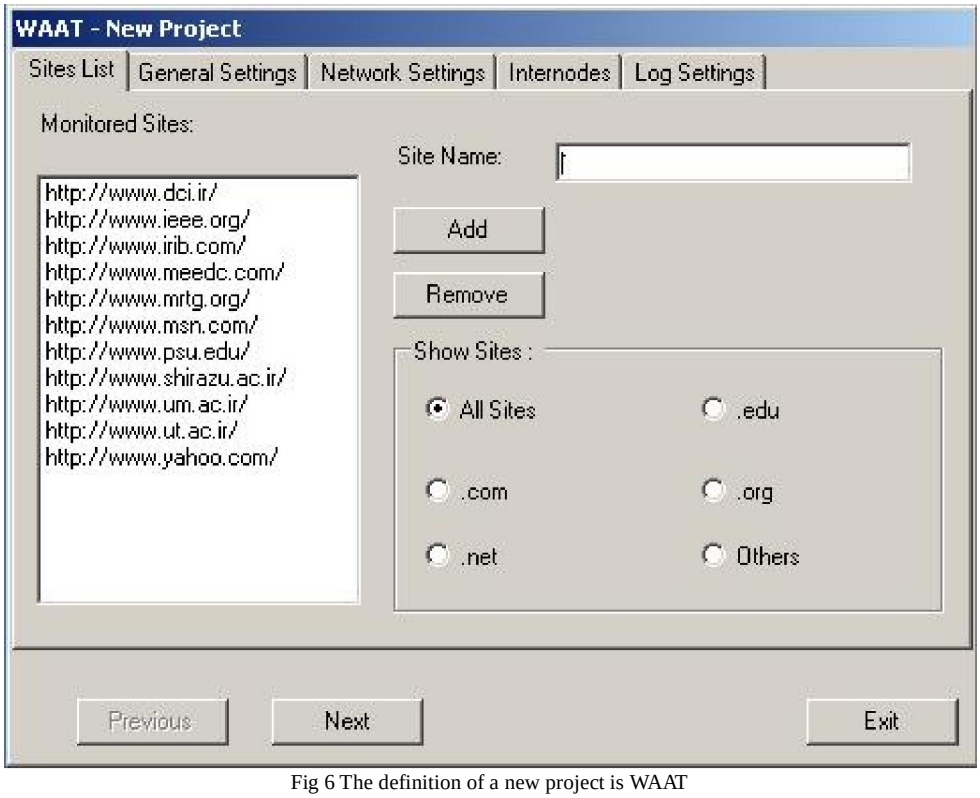

Fig 6 The definition of a new project is WAAT

This window consists of the following subsections:
1- Sites List
2- General Settings
3- Network Settings
4- Internodes
5- Log Settings

The general settings window may be used to

define the parameters that indicate the period between retrials for a malfunctioning site, the period between retrials for a functioning site and the number of successful trials to check a site to determine that it is available as shown in Fig. 7.

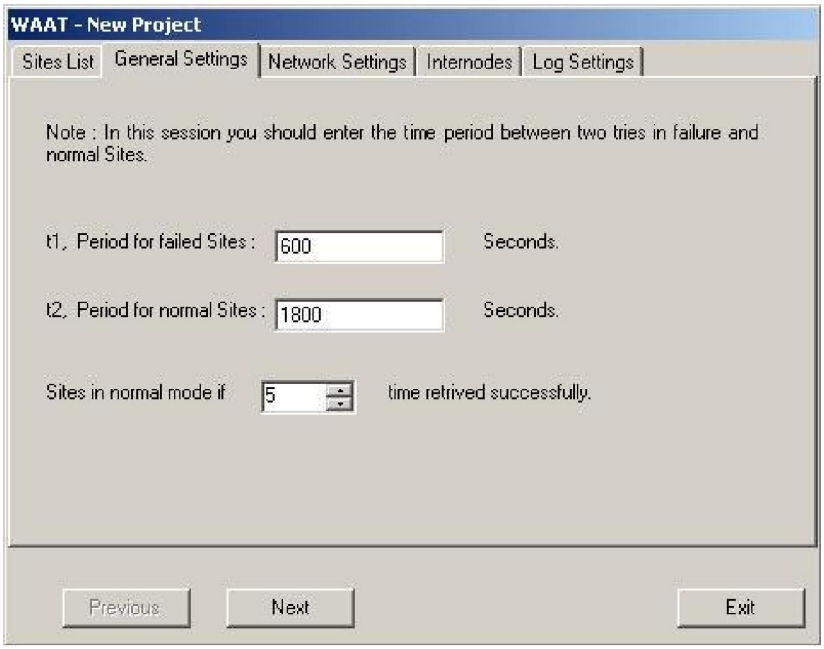

Fig 7 The General Settings window in WAAT

The Network Settings window may be used to set the output port used by the software in case there are several network adaptors in the system. The DNS server is also specified here so that the software does not use its TCP/IP settings as shown in Fig 8.

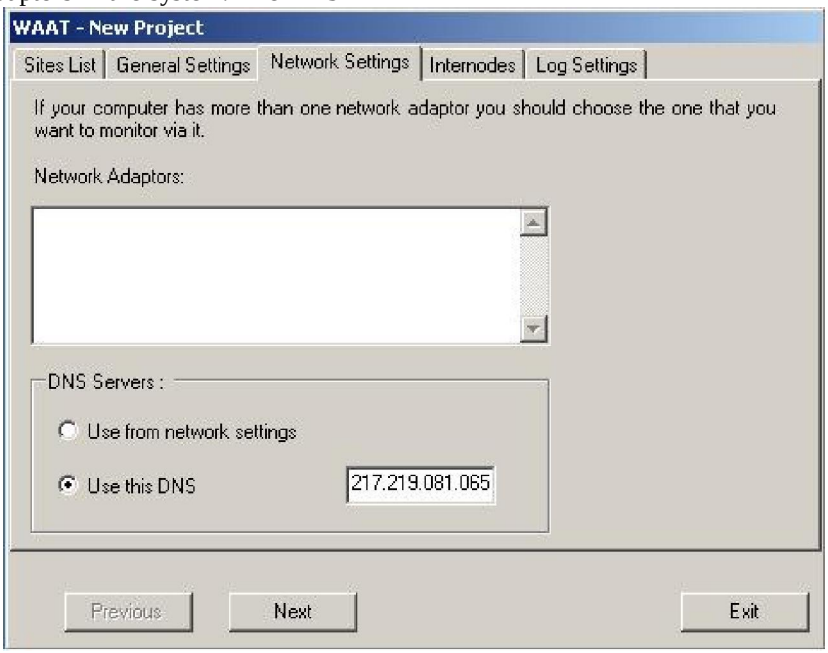

Fig 8 The Network Settings window

The WAAT software uses ping to check intermediate nodes. The maximum number of these nodes may be five and their IPs must be specified in the Internodes window as shown in Fig 9.

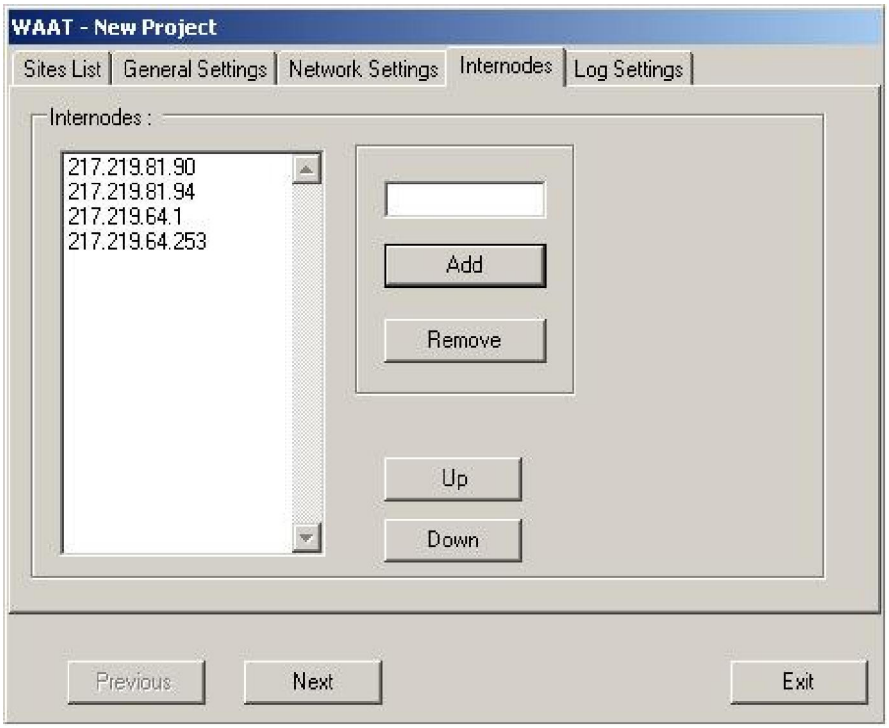

Fig 9 The Internodes window of WAAT

The type of desired log file may be entered into the software through the Log Settings window as shown in Fig 10. Possible options are Text Mode, MS Access Database and MS SQL Server Database. The time of day at which the contents of the log file should be analyzed is also stated in this window.

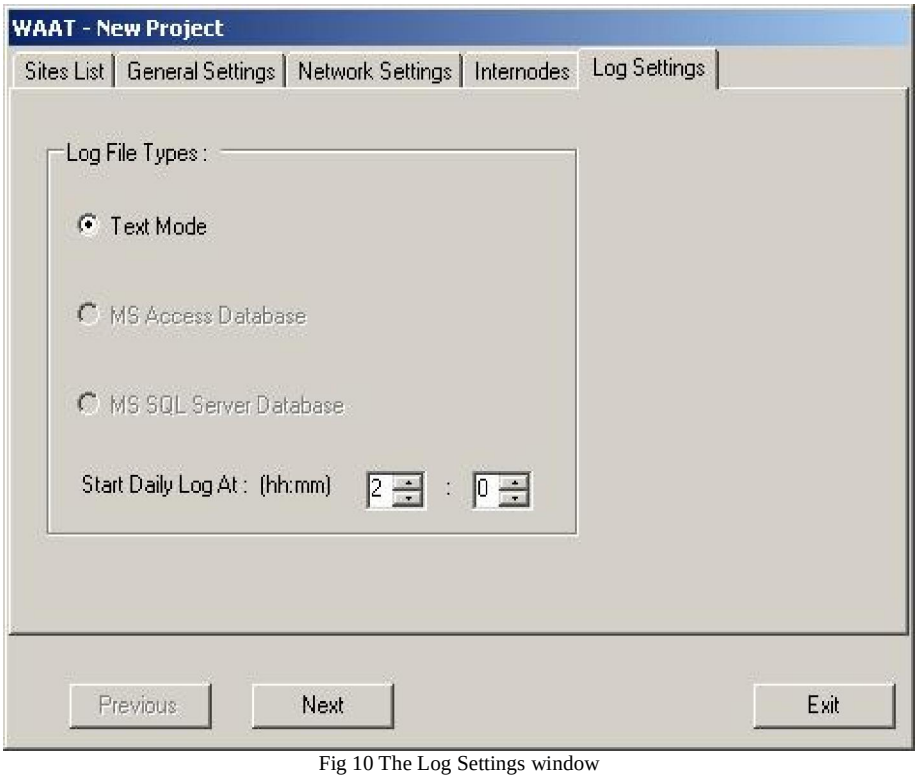

Fig 10 The Log Settings window

A project may be loaded and its data will appear as shown in Fig 11.

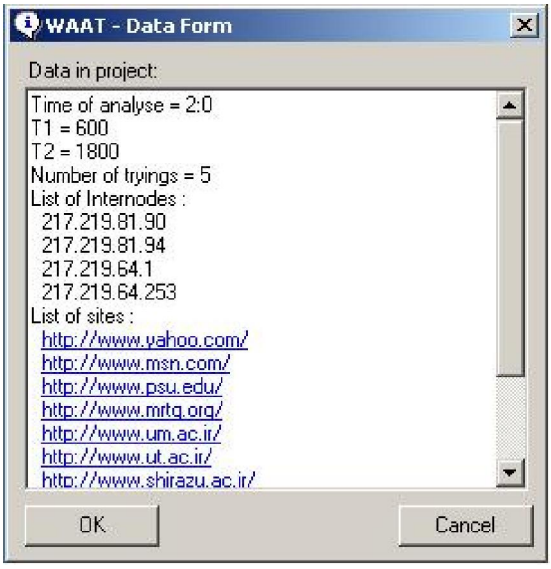

Fig 11 The data form of a loaded project

The software tool developed was tested and it generated a log file for each 24 hours. It was run for 90 days to measure reliability data for 159 Iranian hosts. The hosts chosen included 18 universities, 6 news agencies, 36 internet service providers, 29 government agencies and the rest were other public sites. This mix was chosen so as to obtain an average measure of intrinsic internet availability in Iran. The unavailability data measurement was executed from two different points of connection to the internet to remove any unavailability data related to the facilities of the measurement sites themselves so as to purely obtain the behavior of the hosts under study. We used ping to exclude failures related to intermediate lines and nodes, and thus eliminated any failures due to the internet backbone, too. Analysis of the log files revealed the following results as shown in Table 1.

Table 1 - The mean and median Availability values

Mean availability	0.865
	±0.007 (50% confidence)
	±0.012 (95% confidence)
	±0.016 (99% confidence)
Median availability	0.934
Number of hosts	159

Long et al. (1995) had reported a similar study with the following results after surveying 1170 hosts that were uniformly distributed over the name space and could respond to RPC polls for 90 days. Their results are shown in Table 2.

Table 2 - The mean and median Availability values

Mean MTTR	0.9260
	±0.002 (50% confidence)
	±0.007 (95% confidence)
	±0.009 (99% confidence)
Median MTTR	0.9723
Number of hosts	1162

Table 3 – A comparison of results for the survey on Iranian hosts and international hosts

Hosts	**MTTF (days)**	**MTTR (days)**	**Availability**
159 Iranian Hosts	21.56	3.375	0.865
1162 International Hosts	15.92	1.201	0.926

Table 3 shows a comparison of the results of the two studies. This indicates an average lower level of reliability for Iranian hosts, compared with international hosts indicating that a lot more work is needed for Iranian hosts to reach the average international availability levels.

Conclusions

The development of a Web Availability Analyzer software Tool is reported along with its real application to measurement of internet availability. The software developed simulates a browser and operates like a web client to make sure that the access to the desired host is truly measured. The program developed is reconfigurable and it performs statistical analysis of logged data to derive the desirable measures of internet reliability. The software developed uses multithreading and is suitable for the development of a tattler system. A comparison of results obtained for the availability of national web sites with that on international sites indicated a lower availability of national sites.

Acknowledgement

I would like to thank the Office of Vice Chancellor of Research and Technology of Ferdowsi University of Mashhad for the Grant Project No. 10448 dated 1388/10/2 to support the study that has resulted in the preparation of this manuscript.

References

[1] Wei Xie, Hairong Sun, Yonghuan Cao, Kishor S. Trivedi, (2002), "Modeling of online service availability perceived by Web users", Technical report, Center for Advanced Computing and Communication (CACC), Duke University, http://citeseer.ist.psu.edu/xie01modeling.html

[2] Darrell D. E. Long, (1992), "A replicated monitoring tool", Proceedings of the second Workshop on the Management of

Replicated Data, pp.96-99, Monterey, Ca., U.S.A.

[3] K.B. Sriram, (1993), "A study of the reliability of hosts in Internet", MS Thesis, The University of California at Santa Cruz, pp.1-55.

[4] Darrell D. E. Long, John L. Carroll, and C. J. Park, (1991), "A study of the reliability of Internet sites", Proceedings of the Tenth Symposium on Reliable Distributed Systems, pp.177–186.

[5] M. Kalyanakrishnan, R. K. Iyer, J. U. Patel, (1999), "Reliability of internet hosts: a case study from the end user's perspective", 6th International Conference on Computer Communications and Networks, pp.47-55.

[6] Darrell Long, Andrew Muir, Richard Golding, (1995), "A longitudinal survey of Internet host reliability", Proceedings of the 14th Symposium on Reliable Distributed Systems.

[7] Peiravi, A., Sharaeini, M, Implementation of a measurement tool for the reliability measurement of the Web server of Ferdowsi University of Mashhad, A paper published in Farsi in the Proceedings of the ICEE2004 (12th Conference on Electrical Engineering, Ferdowsi University of Mashhad, May 11-13, 2004, pp.50-55, 2004.