Survivable, Real Time Network Services

Defense Advanced Research Projects Agency
Contract F30602-98-1-0225, DARPA Order G409

Quarterly Progress Report
1 Ocxtober 1998 - 31 December 1998

David L. Mills
Electrical Engineering Department
University of Delaware

1. Introduction

This report covers the work done in support of the DARPA Information Technology Office program in computer networking. Contributors to this effort include Prof. David L. Mills and graduate students Qoing Li and Robert Redwinski. The project continues previous research in network time synchronization technology jointly funded by DARPA and US Navy. The technology makes use of the Network Time Protocol (NTP), widely used in the Internet, together with engineered modifications designed to improve accuracy in high speed networks. Specific applications benefiting from this research include multicast topologies, multimedia, real-time conferencing, cryptographic systems, and management of distributed, real-time systems.

This quarterly report is submitted in traditional report form on paper. As the transition to web-based information dissemination of research results continues, almost all status information and progress reporting is now on the web, either on pages belonging to the principal investigator or to his students. Accordingly, this and future progress reports will contain primarily schedule and milestone data; current status and research results are reported on web pages at www.eecis.udel.edu/~mills in the form of papers, technical reports and specific briefings.

2. Network Time Protocol Version 4

Work continues on the Network Time Protocol Version 4. The principal areas of activity include the autokey scheme, compile-time autoconfigure scheme and network simulator.

2.1 Autokey

The autokey scheme provides cryptographically secure server authentication to other servers and clients using public-key signatures and a reverse hashing scheme as described in previous reports and briefings on the web. Previously, a draft proposal was submitted to the IETF as an Internet Draft. The proposal included a packet format for a cryptographic extensions field for NTP packets. This has been implemented in part in the current NTP Version 4 daemon. Recently, this proposal has been amended and submitted as an Internet Draft [11]. Work on integrating this format in the daemon continues.

Future work requires integration with the Secure DNS project. From initial discussions, it appears that simple modifications to the NTP resolver process will be sufficient for this purpose. It also appears that the signature trail implemented by the Secure DNS will be sufficient evidence that the certificates provided can be trusted. Previously, it was assumed that the NTP daemon would have to do that. Additional states to the NTP protocol state machine will be necessary, as well as provisions to import the complement of RSA Laboratories algorithms required.

2.2 Compile-Time Autoconfigure

Maintaining the NTP Version 4 software in the face of the many ports of it to over two dozen combinations of hardware architectures and operating systems has been an ongoing chore. What makes it much worse is that operating system versions seem to have half lives something like one year, so there is always a window spanning several versions which must be supported (fortunately, the VAX and Fuzzball may have gone to software heaven and can be forgotten). Fortunately, volunteer Harlan Stenn, a self employed programming consultant, has been maintaining the gnu autoconfigure system, with which tracks even minute changes in operating system environment and then builds header files which guide the compiler accordingly. With these tools, bringing up new ports and maintaining old ones are much easier.

2.3 New Features

There are two new GPS receiver drivers, one contributed by Trimble Navigation and the other by Motorola, both for their respective new products. There are now 33 drivers for just about every timetelling device now on the market.

A working group has been started to develop a common API for pulse-per-second (PPS) signals. Members include representatives from the Digital, Sun and FreeBSD developer groups. The work has converged on an API suitable for both PPS signals used for time synchronization and for utility timer/counter purposes. Prototype implementations of the API is to be done first for FreeBSD and then for Digital Unix. A preliminary functional implementation has been done in our laboratory for Digital Unix, but is not yet conforment with the API.

3. Network Simulator

Routing instability has always been the most destructive hazard to high survival internetwork services. If a critical network link carrying routing information fails, large portions of the network may be isolated and the entire network may become unstable. Lessons of the early ARPAnet years demonstrated that the design of the routing algorithm itself could contribute to unexpected congestive failures. While modern routing algorithm designs are highly resistant to such failures, there remain significant gaps in understanding the survivability issues when the number of network elements spanned by the routing system becomes very large.

There are a number of available simulators designed for network simulations with relatively small numbers of nodes - from tens to possibly hundreds of nodes. These simulators tend to focus on the detailed interactions between one node or protocol and another, with the remaining nodes acting as routers or generating noise. With very large networks, the interest is on macro behavior and global phenomena and closely watched local behavior is less important than global stability.

In order to explore the behavior of routing functions in very large networks, Robert Redwinsky developed a discrete event simulator suitable for networks including many thousands of nodes. There are three components which together make up the simulator system. The first is a random topology generator which generates network graphs using a probabilistic algorithm modelled on principles developed by Waxman. Topology generation is an offline process and concludes with a data file that is input to the simulator itself.

The second component is the simulator itself, which is a conventional deterministic, discrete-event model, but with a very large virtual memory to hold the simulation entities and event queue. The simulator is equipped to generate random failures of one or more links or nodes according to a designated probability model. One of our Sun workstations has been expanded to 640 megabytes of RAM, which at the present stage of development supports a network with up to 3,500 nodes.

The third component is a suite of candidate routing algorithms, including the venerable Bellman-Ford node-state algorithm and the Distance Vector Multicast Routing Algorithm (DVMRP). The simulator implements these algorithms in the same way as on an actual network, including the effects caused by routing tables changing while the routing updates themselves are travelling between nodes, as well as random failures of the nodes and links and lost routing packets.

The present stage of the project will be documented in a Masters thesis written by Mr. Redwinski. At the present stage of development, simulations with up to 3,500 nodes can be supported in a machine with 640 megabytes of memory. As additional memory is now relatively inexpensive, expanding memory to add more nodes may be preferable to intricate redesign of the program to permit this in the present system, should that be necessary.

The eventual goal of the project is to provide useful data in networks i of 10,000 nodes or more and perhaps three times this many links. We plan to explore the behavior in response to various kinds of failure models and failure/recovery rates and determine the character of these failures and how they might be predicted and avoided. We plan to explore other routing algorithms as well, in order to conduct comparative studies of the strengths and weaknesses of the selected algorithms within our experimental framework. Finally, we expect to test the algorithms developed for the autoconfigure project using the simulator.

4. Infrastructure

Our CAIRN/DARTnet router udel.dart.net has been replaced by a 200-MHz Pentium udel.dart.net aka barnstable.udel.edu on load from ISI-East. The Sun SPARC 1 that it replaced has been recommissioned as a utility server for the new ntp.org domain. We recently upgraded all software to the CAIRN 2.5 kernel and reconfigured for changes in LAN connectivity here. In addition, we have commissioned another 200-MHz Pentium hepzibah.udel.edu as a development platform for CAIRN related software and experiments. This machine is now fully integrated in the NFS and NIS domain including all hosts and routers supporting Internet research in our Department. It is our intent to make this machine available to other CAIRN researchers as well.

4.1 NTP Version 4

The latest NTP Version 4 daemon and utility programs have been compiled and installed on both FreeBSD machines. The behavior when the NTP daemons were started on both machines showed severe jitter up to several tens of milliseconds. This was a very surprising finding, since reports from elsewhere suggested that NTP Version 4 in Pentium-class machines with FreeBSD performs as well as in other architectures and operating systems. Initial suspicion centered on the automatic power control (APC), which has been known to destabilize timekeeping on some systems. However, we found APC had already been disabled on both machines.

Somewhere in the confusion that usually surrounds these things, the problem went away and both machines began serious lives as stalwart timekeepers. Unfortunately, we don't know just what the original problem was, since it went away without any specific change on the part of the developers.

One of the chores remaining in the NTP port to FreeBSD is working through the issues remaining in the IRIG driver. An IRIG signal and audio codec is a useful alternative to a PPS signal and serial port. Most GPS receivers and precision timing gear can generate these signals. The IRIG signal is connected directly to the audio port or sound card, which avoids building a level converter and using up a serial port. We are now using an IRIG signal to discipline a Sun-based primary server and find the performance quite satisfactory for most purposes. However, the nominal precision within a few tens of microseconds is far from what is achieved with a PPS signal connected to a serial port or, in the case of a PC, to a parallel port.

4.2 New Domain ntp.org

In order to regularize archive and distribution functions for the NTP community, we have registered the domain name ntp.org with the InterNIC and installed a SPARC 1 to serve as a web server and home directories for the volunteer code developers and testers. This development environment, which was borrowed from the FreeBSD developers community, allows the security policies of the Department and Internet research activities to be isolated from the policies more appropriate for the volunteer corps.

5. Plans for the Next Quarter

Our plans for the next quarter include continued testing and refinement of the NTP Version 4 protocol model, specification and implementation. Specifically, we plan to resolve the problems with the Unix socket interface mentioned in the previous report, so that the NTP autoconfigure feature is really useful. In addition, we plan to continue the collaboration with Coastek InfoSystems in the design and implementation of the cryptographic certification algorithm. The daemon is to be tested first in the research net, then the DARTnet/CAIRN community. As the extensions are backwards compatible, the new features can be activated and tested in regular operation without impacting current users.

6. Publications

All publications, including journal articles, symposium papers, technical reports and memoranda are now on the web at www.eecis.udel.edu/~mills. Links to the several publication lists are available on that page, as well as links to all project descriptions, status reports and briefings. All publications are available in PostScript and PDF formats. Briefings are available in HTML, PostScript, PDF and Proponent. The project descriptions are cross-indexed so that the various interrelationships are clearly evident. Links to other related projects at Delaware and elsewhere are also included on the various pages. Hopefully, the organization of these pages, which amount to a total of about 300 megabytes of information pages and reference documents, will allow quick access to the latest results and project status in a timely way.

Following is a retrospective list of papers and reports supported wholly or in part on this project and the immediately preceeding project "Scalable, High Speed, Internet Time Synchronization," DARPA Order D012. The complete text of all papers and reports, as well as project briefings, status reports and supporting materials is at www.eecis.udel.edu/~mills.

6.1 Papers

1. Li, Qiong, and D.L. Mills. On the long-range dependence of packet round-trip delays in Internet. *Proc. IEEE International Conference on Communications* (Atlanta GA, June 1998), 1185-1191.

2. Mills, D.L., A. Thyagarajan and B.C. Huffman. Internet timekeeping around the globe. *Proc. Precision Time and Time Interval (PTTI) Applications and Planning Meeting* (Long Beach CA, December 1997), 365-371.

3. Mills, D.L. Authentication scheme for distributed, ubiquitous, real-time protocols. *Proc. Advanced Telecommunications/Information Distribution Research Program (ATIRP) Conference* (College Park MD, January 1997), 293-298.

4. Mills, D.L. The network computer as precision timekeeper. *Proc. Precision Time and Time Interval (PTTI) Applications and Planning Meeting* (Reston VA, December 1996), 96-108.

5. Mills, D.L. Improved algorithms for synchronizing computer network clocks. *IEEE/ACM Trans. Networks 3, 3* (June 1995), 245-254.

6.2 Technical Reports

6. Sethi, A.S., H. Gao, and D.L. Mills. Management of the Network Time Protocol (NTP) with SNMP. Computer and Information Sciences Report 98-09, University of Delaware, November 1997, 32 pp.

7. Mills, D.L. A precision radio clock for WWV transmissions. Electrical Engineering Report 97-8-1, University of Delaware, August 1997, 25 pp.

8. Mills, D.L. Clock discipline algorithms for the Network Time Protocol Version 4. Electrical Engineering Report 97-3-3, University of Delaware, March 1997, 35 pp.

9. Mills, D.L. Proposed authentication enhancements for the Network Time Protocol version 4. Electrical Engineering Report 96-10-3, University of Delaware, October 1996, 36 pp.

10. Mills, D.L. Simple network time protocol (SNTP) version 4 for IPv4, IPv6 and OSI. Network Working Group Report RFC-2030, University of Delaware, October 1996, 18 pp.

6.3 Internet Drafts

11. Mills, D. L., T.S. Glassey and M.E. McNeill. Authentication Scheme Extensions to NTP. Internet Draft draft-mills-ntp-auth-coexist-01.txt, IETF, September, 1998.

12. Mogul, J., D.L. Mills, J. Brittenson, J. Stone, P.-H. Kamp and U. Windl. Pulse-per-Second API for UNIX-like Operating Systems, Version 1.0. Internet Draft draft-mogul-pps-api-02.txt, IETF, June 1998, 25 pp.