NGI/CE/04

Leveraging Net100 to Optimize Bulk NGI Data Transfers

FY 2003 Proposal to the NOAA HPCC Program

August 19, 2002

| Title Page | Proposed Project | Budget Page |

Principal Investigator:	**John P. Hernandez**
Line Organization:	OAR
Routing Code:	R/OM62
Address:	NOAA/OAR/OED 325 Broadway Boulder, CO 80305
Phone:	(303) 497-6392
Fax:	(303) 497-6005
E-mail Address:	John.Hernandez@noaa.gov

Other Investigator 1
Tony.Wu@noaa.gov

Other Investigator 2
John.O.Kinsfather@noaa.gov

Proposal Theme:	**NGI**
Funding Summary:	FY 2003 $ 28,500

John P Hernandez
Network Specialist
NOAA Boulder NOC

Jerry Janssen
Network Manager
NOAA Boulder NOC

Kent Groninger
Director of Boulder Operations
NOAA/OAR/OED

Leveraging Net100 to Optimize Bulk NGI Data Transfers

Proposal for FY 2003 HPCC Funding

Prepared by: John P. Hernandez

Executive Summary:

This proposal aims to increase the speed and efficiency by which NOAA scientists are able to transfer large amounts of data over the NGI. Previous HPCC funding for NESDIS (John O. Kinsfather 2001) was used to gain a better understanding of the current limitations in reliable bulk network throughput, and this proposal intends to directly address the shortcomings revealed by their work.

Despite the fact that Boulder and other NOAA campuses have robust underlying NGI network capabilities with bandwidth in excess of 100Mb/s, it is clear that most computer applications cannot take full advantage of these resources without the expert attention of network engineers. The Net100 project, sponsored by Pittsburgh Supercomputing Center, National Center for Atmospheric Research, National Center for Supercomputing Applications, and Lawrence Berkeley National Lab, is interested in closing this so-called "wizard gap" and enabling researchers and scientists to make full use of available network resources.

The investigators will use the suite of Net100 tools, which include Web100, Netlogger, and Iperf, to demonstrate that there is significant room for improvement in NOAA's current utilization of the Abilene NGI research network. An improvement in bulk network transfer speeds will enable scientists to collaborate with more transparency and ultimately save time and money. For example, the network will replace the traditional manual exchange of information on magnetic tapes.

Problem Statement:

Network engineers are often asked, "Why does this ftp file transfer utilize so little of our supposedly plentiful NGI bandwidth?" Many of the usual suspects, such as raw disk (IO) performance and congestion of shared network segments are quickly disappearing as culprits due to advancements in the computer hardware industry. When these bottlenecks are eliminated, one inadequacy that remains is the network software.

It is now well understood that a limiting factor in bulk network throughput over distances of several hundred kilometers or farther involves the design of the ubiquitous Transmission Control Protocol (TCP). This protocol is used almost exclusively on IP networks to "copy" files from one node to another.

Our computers' network software, even on large servers, is still largely tuned for the highly congested and lower performing networks of the past. Over long distances such as those between Boulder and Washington D.C., networks suffer from the inherent and unavoidable effects of latency (delay) due to the speed of light (in fiber optics) and packet routing decisions. From the point of view of TCP, this latency is construed as network congestion, and the network

software reacts as best it can by throttling back transmission speeds, resulting in poor overall performance.

In a recent discussion of Abilene's statistics at the I2/NLANR Joint Techs conference in Boulder, it was pointed out that bulk TCP transfers (defined as more than 2 MB of data transferred over a single connection) averaged an unimpressive 2 Mb/s. This inefficiency directly affects NOAA's ability to ingest large datasets and export valuable data to our customers.

One of the conclusions reached by John O. Kinsfather in his investigation entitled "Optimizing Massive NOAA Data Transfers on the NGI" was that current TCP implementations are a significant barrier to accomplishing their stated goals.

Proposed Solution:

A team of computer scientists has been hard at work developing a solution to this problem, which involves dynamic fine-tuning and "tweaking" of TCP parameters in order to achieve higher sustained throughput for a variety underlying networks. This effort is knows as the Web100 project. More information can be obtained at http://www.web100.org.

The investigators will implement Web100 and other Net100 techniques in an effort to minimize the TCP barrier to high bulk data transfer rates. Our proposed method is detailed below.

- Obtain two dedicated data transfer test stations with memory adequate to hold large datasets provided by NGDC. One station will be located at the Boulder NOC, the other at the Silver Spring NOC. These stations will be high-end commodity PC's with Linux installed and Gigabit Ethernet network adapters.

- Install Iperf throughput testing software on test stations. Test and baseline raw end-to-end performance by sending unreliable, "connectionless" data streams using the User Datagram Protocol (UDP) between data transfer stations in Boulder and Silver Spring. Also baseline bare TCP performance using plain ftp.

- Install Web100 kernel and libraries and Netlogger performance measuring software on test stations.

- Learn to leverage Web100's "knobs and dials" to increase bulk TCP performance, and continue to experiment with TCP modifications until the potential of the underlying NGI network is realized.

- Implement the final solution on NGDC equipment and coordinate deployment with their remote collaborators.

This solution is preferable to existing workarounds such as splitting up and transferring data sets in parallel using high data compression techniques. Such workarounds require extensive pre-processing which cut into potential benefits, and they often result in degradation or reduction in the resolution of data sets. Although it is tempting to "throw more bandwidth at the problem," it is clear that this approach cannot overcome current software bottlenecks, and nothing will be gained from the perspective of an individual researcher's goal.

Analysis:

The investigators will fully document the effects of each successive change on the performance of the sample data transfer. Our stated goal is to utilize 80% or greater of the available end-to-end network bandwidth in a single TCP session. When well coordinated to avoid resource conflicts, this increased utilization will have a very positive impact on NOAA research and NOAA's data archiving efforts. Other future benefits include improved network backups, reliable multicasting, and richer multi-media web content.

Based on the achievement of this goal, for example, NOAA's M.O.U. with NASA to provide archiving of EOS data can begin to take form. Another direct application is the NOS Continuous Observing Recording System (CORS), which will provide large archives of GPS data to NGDC. Such efficient transfer of massive datasets cannot be realized unless we begin to explore the means to provide researchers with increased network throughput for individual transfers.

Because today's networks allow engineers to give priority to certain types of traffic, we can allow for much larger bursts of traffic for individual bulk transfers, while at the same time guarantee the availability of network resources for other applications.

Performance Measures:

Milestones

Month 3 – Set up test systems and load data sets

Month 5 – Obtain baseline performance metrics

Month 7 – Gain an understanding of the technical aspects of theWeb100 suite

Month 9 – Determine the best use of Net100 tools to achieve the desired goal

Month 10 – Produce methodology and future outlook reports

Month 12 – Deploy final solution on NGDC and remote servers

Deliverables

- A proof-of-concept system that performs bulk TCP transfers with significantly improved performance and minimal engineer oversight
- A detailed methodology for achieving similar results on other NOAA systems, particularly for NGDC
- An outlook for the improvement of network software on operating systems of the future, and valuable feedback for the Net100/Web100 projects