TTR/CE/22

Technology Transfer for Accelerating Web Access at NSSL in Support of Hurricane Imagery

FY 2004 Proposal to the NOAA HPCC Program

August 11, 2004

| Title Page | Proposed Project | Budget Page |

Principal Investigator: **Alex Hsia**

Line Organization: OAR

Routing Code: R/OM62
Address: NOAA-Boulder NOC
325 Broadway St
MS R/OM62
Boulder, CO 80305

Phone: (303) 497-6351
Fax: (303) 497-6951
E-mail Address: Alex.Hsia@noaa.gov

Jerry.Janssen@noaa.gov
Gary.Skaggs@noaa.gov
Thomas.Renkevens@noaa.gov

Proposal Theme: **HPCC Technology Transfer**

Alex Hsia	Donald Mock
Network Engineer	Executive Director
NOAA-Boulder NOC	NOAA-Boulder OED

Technology Transfer for Accelerating Web Access at NSSL in Support of Hurricane Imagery

Proposal for FY 2004 HPCC Funding

Prepared by: Alex Hsia

Executive Summary:

The Satellite Services Division (SSD) of NESDIS (co-located with NCEP in Camp Springs, MD) processes GOES satellite data into forecast products utilized by NWS, emergency managers, researchers and the general public. The basic product suite of GOES imagery can be accessed at http://www.goes.noaa.gov and at http://www.ssd.noaa.gov. Currently, these are among the most frequently accessed sites in NOAA, particularly during hurricane season when public interest in satellite imagery of these weather phenomena is at its peak.

To address past access problems due to system overloading during hurricane events, SSD invested in a major system upgrade during winter of 2002 and implemented a 3-server, locally load balanced Linux cluster in Camp Springs. Additionally, with the cooperation of the Boulder Network Operation Center (NOC) and the Silver Spring NOC, SSD has implemented a two-site load balancing scheme utilizing the Silver Spring NOC's Cisco Distributed Director to offload network traffic off the D.C. area WAN. While improving the historical situation, this solution has quickly proven to be inadequate for the kind of network loading experienced around hurricane landfall.

The National Severe Storms Lab (NSSL) has agreed to support the web distribution of satellite imagery by hosting a third node at Norman, OK. We propose to implement this third node by utilizing technology investigated and implemented in the past through HPCC funding. (See "Accelerating Retrieval of Web Content" by M. Knezevich and J. Hernandez, funded by HPCC in FY02). The previous investigation led to a hardware implementation for local load balancing at the Boulder NOC.

Problem Statement:

SSD's satellite imagery websites are among the mostly highly visited sites in NOAA and among the most frequently accessed sites anywhere on the Web during hurricane events. They were established some six years ago to provide public access by the public to the most recent satellite imagery. They were never intended to be supported operationally and no funding has ever been forthcoming to convert the web sites to be fully operational. During the hurricane season of 2002, these two sites became inaccessible during hurricane events because they could not handle the hit loading imposed. The problem of access failure quickly came to the attention of senior NOAA and NESDIS management as it became a source of embarrassment for the agency.

In response, SSD conducted a major upgrade to its web system by implementing a three-server web farm with local load balancing among the servers. With the support of the Boulder NOC, we implemented a remote host in Boulder to offload peak traffic. Our recent experience

indicates that this solution is not completely adequate to handle peak user loading around hurricane landfalling. For example, during Hurricane Claudette in July 2003, the combined hit loading on the SSD/Boulder web systems exceeded 8 million hits per day. During the week surrounding hurricane landfall, approximately 2 Terabytes of data were transferred across the network. This, combined with increased NWS traffic, led to a network crisis in Silver Spring as NOAA's link to the commodity Internet (155 Mbps) was overloaded.

This proposal utilizes investigations and implementations supported by the HPCC Program to extend a technology solution elsewhere. Specifically, the investigation funded in FY02 led to a prototype implementation in Boulder, which we now propose to extend to NSSL.

Proposed Solution:

The existing Norman-NOC web server supports the National Severe Storms Laboratory, the National Hurricane Center mirror project (30% load), and the Storm Prediction Center's primary website. Additionally, the NOC is pursuing mirroring the Radar Operations Center and the Warning Decision Training Branch web sites.
The Norman NOC system was upgraded during the past summer. The current system is a Sun Enterprise 220R server, with 2-450 Mhz processors, 2 GB RAM, and 72 GB of on-board disk space. There is a large amount of disk available by way of NFS. As an outgrowth of an HPCC-funded analysis conducted by the NOAA West Network NOC in Seattle, the Norman NOC wishes to participate in a web acceleration project. Norman's part is to investigate the performance gain, management load, and configuration of multiple, low-cost proxies. The Norman NOC will soon have 4 SQUID-enabled proxies serving cached pages from their web server. These will be configured in 'round robin' fashion for DNS service. If ServerA received the last request, then ServerB will receive the next, etc.

While round-robin DNS works well, it does nothing to address the individual loads of each of the proxies, nor does it check to see if the system is even alive before it sends the request to a server or proxy. It blindly refers the request to whichever system is next in its cache. If the server referred to is down, there is no way to dynamically update DNS with that information.

The Norman system would be enhanced by the proposed technology transfer of portions of the recently demonstrated Boulder web site. The needed technology would be in the area of an Application Layer switch. This application layer device would dynamically balance requests to each of the web proxies by traffic volume, not just number. This would increase the efficiency of the proxies. This application layer switch would also be aware of each of the proxies' health, so that requests would not be routed to servers that were down or suffering from a performance problem for whatever reason.

We propose to purchase and to implement a hardware switch for the front end of the NSSL web system utilizing experience gained by Boulder NOC personnel. SSD personnel responsible for managing the SSD website will assist the Norman NOC personnel with implementing an rsync mechanism for periodic update (half hourly) of satellite imagery data.

Equipment acquisition and implementation will be conducted by Gary Skaggs of NSSL with support from Alex Hsia and Jerry Janssen of the Boulder NOC. Development of the data content transfer and updates via rsync will be supported by Ed Ladd (RS Information Systems

and NESDIS/SSD). Overall coordination will be conducted by Ernest Daddio (RS Information Systems, Inc. and NESDIS/SSD) and Tom Renkevens of NESDIS/SSD.

Analysis:

The proposed solution will significantly improve NOAA's ability to serve the public with environmental information. It will go a long way toward solving a problem highly visible to NOAA management.

The selected solution is an optimum one for the current needs. It is a technology already proven in the NOAA environment and leverages investments made by the HPCC Program in the past.

The selected solution will allow the public access to environmental information of particular interest to the public, i.e., satellite imagery. In the past, this information has become unavailable when the public was most interested in it, during hurricane events.

Performance Measures:

Major quantifiable performance measures include:

number of hits sustained per hurricane event

bytes of data transferred during hurricane season

Milestones

Month 1 – Following receipt of funding—Acquire hardware/software
Month 2 – Implement solution at NSSL
Month 3 – Initiation of operation using three-way (Boulder/Camp Springs/Norman) network load balancing.

Deliverables

Provide a list of the final products from this project

Operational 3-way load balancing of SSD satellite data web site

System documentation