

**An NGI Collaboration Tool
to support NOAA Modeling Groups**
FY 2003 Proposal to the NOAA HPCC Program

August 19, 2003

| [Title Page](#) | [Proposed Project](#) | [Budget Page](#) |

Principal Investigator: **Ants Leetmaa**

Line Organization: OAR
Routing Code: R/GFDL
Address:

GFDL
Forrestal Campus - US Route 1
Princeton NJ 08542

Phone: (609) 452-6502
Fax: 609) 987-5070
E-mail Address: Ants.Lleetmaa@noaa.gov

Dr. Chris Sabine
PMEL
Chris.Sabine@noaa.gov

Dr. Anand Gnanadesikan
GFDL
Anand.Gnanadesikan@noaa.gov

Dr Tsung-Hung Peng
AOML
Tsung-Hung.Peng@noaa.gov

Dr. Donald W. Denbo
PMEL-JISAO
Donald.W.Denbo@noaa.gov

Proposal Theme: **Technologies for Collaboration, Visualization, or Analysis**

Funding Summary: FY 2003 \$ 109,000

Ants Leetmaa
Director
GFDL

An NGI Collaboration Tool **to support NOAA Modeling Groups**

Proposal for FY 2003 HPCC Funding

Prepared by: Ants Leetmaa

Executive Summary:

Research efforts at GFDL involve initiation and participation in modeling consortia that involve research partners from multiple government agencies and academic institutions. These close collaborations involve frequent discussions of specific data and model results. Given the geographic distribution of the investigators and the time and financial costs of frequent face-to-face meetings, a collaboration tool that utilizes the NGI for exploration and annotation of data and model results, enabling colleagues in other locations to see those annotations in real time, is vital.

The ability to interactively graph data from each of the collaborating sites, or from other remote network sites, point to a specific feature in the data while conducting a conference call, and have collaborators immediately see the item of interest and/or point out other features in the data is extremely powerful and beneficial for these collaborations. *A fast NGI network is required for this functionality, and in fact, tools such as this are part of the justification for the existence of the NGI.* Another useful aspect of this tool will be the ability to establish a longer-term topic of discussion on-line, where individual participants may participate in a “networked conference session” to discuss a topic of interest, save the session, work off-line to investigate various aspects of this topic, and then rejoin the session later to post results for comment. So-called “screen scrapers”, such as Microsoft NetMeeting, do not provide the required functionality.

We propose a partnership between three NGI-connected colleagues at GFDL, PMEL and AOML to develop a network collaboration tool to support NOAA modeling groups. Development can be leveraged from earlier PMEL Java efforts, and the Carbon Modeling Consortium collaborations between GFDL, PMEL, and AOML are a perfect test-bed for such a tool. Our three co-investigators are researchers who are actively involved in the Carbon Modeling Consortium, and who are personally excited about using the proposed tool. They will provide hands-on testing and science direction for the tool development. The tool will be made freely available, and GFDL will use it in other collaborative modeling efforts and consortia.

Problem Statement:

Background: NOAA is entering a new era of collaboration among the different environmental laboratories, research organizations and academia. Issues like global climate change cut across traditional disciplinary boundaries and require collaborative studies among groups of investigators that may have offices spread across the country or even across the world.

Researchers at GFDL are participating in several modeling consortia that involve research partners from multiple government agencies and academic institutions, such as the Carbon Modeling Consortium ([CMC](#)) and the NOAA Operational Model Archive and Distribution System ([NOMADS](#)).

The Carbon Modeling Consortium involves formal collaborations between the computer modeling groups at GFDL in Princeton, NJ and the ocean observation groups at PMEL in Seattle, Washington and AOML in Miami, Florida to evaluate and help constrain model predictions of ocean carbon distributions generated by GFDL with direct observations collected by PMEL and AOML. This close collaboration involves frequent discussions of specific data and model results, and forms a perfect testbed for the proposed tool.

Synchronous collaboration: Given the geographic distribution of the investigators and the time and financial costs of frequent face-to-face meetings, a tool designed to allow a user to explore and annotate data and model results and have colleagues in other locations see those annotations in real time is vital. The ability to interactively introduce and graph data from each of the collaborating sites, or from other remote network sites (e.g. DODS/OPeNDAP sites), point to a specific feature in the data while conducting a conference call, and have collaborators immediately see the item of interest and/or point out other features in the data is extremely powerful and beneficial for these collaborations. *A fast NGI network is required for this functionality, and in fact, tools such as this are part of the justification for the existence of the NGI.*

Asynchronous collaboration: Another useful aspect of this tool will be the ability to establish a longer-term topic of discussion on-line, where individual participants may participate in a “networked conference session” to discuss a topic of interest, save the session, work off-line to investigate various aspects of this topic, and then rejoin the session later to post results for comment. This session could then be saved to document the range of analyses that were investigated and the resulting decision that was made regarding that topic. At a future date, when a question arises about why a particular decision was made regarding that topic or a similar topic of interest is raised, one can log into the specific “conference session” and see the notes and plots that helped resolve that issue.

Relationship to HPCC program objectives: The proposed network collaboration tool directly addresses the HPCC objective to develop “advanced enabling technologies... that include applications sharing, scientific whiteboarding, distributed data integration and session indexing management”, involving “the cooperation of multiple sites”. The proposed tool meets all these requirements and is highly appropriate to stated HPCC objectives. As required, participants in this effort are NGI partners.

A fast NGI network is required for the proposed functionality, and in fact, tools such as this are part of the justification for the existence of the NGI. *Although collaboration tools of the type proposed here have been a stated objective in high performance supercomputing studies for several years, to the best of our knowledge, this will be the first use of such a tool to support modeling efforts in NOAA.*

Scope: Combining the significant scientific credentials of GFDL and the carbon modeling consortium members who are co-investigators with PMEL’s well-recognized expertise in software development will result in a very powerful NGI collaboration tool that should have widespread impact on the modeling community inside and outside NOAA.

Proposed Solution:

Overview: We propose a partnership between GFDL, PMEL and AOML to develop an NGI collaboration tool to support NOAA modeling groups. Our three co-investigators are researchers from GFDL, PMEL and AOML who are actively involved in the Carbon Modeling Consortium. They are familiar with general-purpose, off-the-shelf collaboration software, which lacks the functionality they require, and have seen demonstrations of ncBrowse, a PMEL Java DODS data browsing tool, and the OceanShare collaboration tool developed previously by PMEL for the Fisheries Oceanography community with *in-situ* ocean observations, from which this effort will build. They are personally excited about using the proposed tool, their research organizations are on the NGI, and they will provide hands-on testing and science direction during the tool development. The requirements they have identified are significant and will result in a very powerful collaborative tool to support NOAA modeling groups. Moreover, Dr. Donald W. Denbo, who is the primary designer/developer of these Java components (and is also PMEL’s *in-situ* data manager) will provide the technical leadership for the proposed effort.

Features already available in ncBrowse	Required scientific features (to be added)
<ul style="list-style-type: none"> • Flexible, interactive graphical displays of <i>data and metadata (attributes)</i> from a wide range of netCDF data file conventions. <ul style="list-style-type: none"> ○ Parametric plots (e.g., temp-salinity) ○ Flexible mapping of independent variables to axes. ○ Vector plots , Animations, 3D visualizations ○ Publication quality graphics • Remote netCDF file access with simple HTTP or DODS/OPeNDAP • Supported on Unix, Linux, Windows and Mac OS. Freely available at http://www.epic.noaa.gov/java/ncBrowse • Popular with users. Over 3700 downloads by 1852 unique sites in 49 countries, and an active email discussion group. 	<ul style="list-style-type: none"> • Support ocean/atmospheric model vertical coordinate systems, e.g., sigma, hybrid, and isopycnal. • Extract arbitrary sections from ocean model results, not just east-west and north-south, for comparison with observations. WOCE hydrographic sections do not all lie on east-west or north-south lines. Coastal hydrographic sections are often parallel or perpendicular to the coast. • Support for statistical and graphical comparisons between station data and gridded model results. Statistics should include correlations and Taylor diagrams (comparison of 2-d fields both by pattern and magnitude). • Import and export data to local files. These files should be readable by external analysis tools, e.g., Ferret, Matlab, etc.
Features already available in OceanShare	Required collaboration features (to be added)

<ul style="list-style-type: none"> • Work with all collaborators on-line simultaneously (synchronous collaboration). • Prior sessions can be saved; re-entered later (asynchronous collaboration) • Access to local and remote data in a networked collaborative tool environment <ul style="list-style-type: none"> ○ Group access to local and remote data files, graphics ○ Standard collaboration tool features (annotation, whiteboard, etc) • Import of model output and data sources being collected for model initialization, assimilation, and validation. • Secure Document Repository <ul style="list-style-type: none"> ○ Secure location for sharing documents, e.g., proposals, papers in progress, data, images ○ Access control list for individual documents and/or folders 	<ul style="list-style-type: none"> • Locally remove a graphics window from the current session, edit the graphics, and then put the window back into the session. • Delete a graphics window from the session. • Bookmarks. Tagging the session so that new actions added after you have logged out of the collaborative session are indicated once you have re-logged in. • Retain participant activity logs by session. • Session replay: fast forward to bookmark, then single step. • Enable new logins to a secure session after the session creator has logged out. • Restrict a session to specific individuals. • Date/Time stamp all actions. List date-time in the chat window. • Session editing: remove specific or redundant objects • Archive a session. Save/restore a session from JavaSpaces to a local directory.
---	---

Leveraging and Implementation: The network collaboration tool with the functionality the modeling community requires can be developed by integrating Java components developed previously.

- a) *OceanShare*, a collaboration tool focused on in-situ ocean observations
<http://www.epic.noaa.gov/collab/> (HPCC, FY01),
- b) *Plug and Play Java Graphics with SGT*
<http://www.epic.noaa.gov/java/sgt/> (HPCC, FY02),
- c) *3-D Visualizations with a Networking netCDF Data File Browser: ncBrowse*,
<http://www.epic.noaa.gov/java/ncBrowse/>, (HPCC, FY02) and
- d) *DODS dataset exploration using a portable Java application*,
<http://www.epic.noaa.gov/java/ncBrowse/> (ESDIM FY01)

The OceanShare collaboration tool addressed a different set of functional requirements, but has some of the collaboration features identified as requirements for the modeling community. We will continue to use [Jini](#) and [JavaSpaces](#), a simple unified mechanism for dynamic communication, coordination, and sharing of objects between Java™ technology-based network resources, to provide the basic collaborative framework and the SGT Scientific Graphics Toolkit to provide the interactive Java graphics. NcBrowse has some of the graphics functions required.

Matching Funds: All partners in this proposal are providing PI salaries at no cost. This represents a significant contribution, since the scientist PI's will be actively involved in hands-on testing, using the tool as it is developed, and providing scientific direction.

Analysis:

Collaborative tools provide an enabling technology for the scientist. The proposed NGI collaboration tool will immediately improve the ability of scientist to effectively communicate complex ideas related to NOAA's carbon modeling effort. While this proposal is focused on a single modeling group, the tool developed here will be easily transferable with little or no modifications to other modeling groups having similar collaborative requirements. Core pieces of the technology used to create the tool are provided as part of the Java™ technology platform or from NOAA sponsored development with a long track record (SGT and ncBrowse have been publicly available since 2000 and OceanShare, the collaborative tool, since 1999).

Alternatives to using JavaSpaces and OceanShare for the collaborative framework include Microsoft's NetMeeting or other off the shelf or "screen scraper" products. We are aware of no other product which meets the requirements as described in this proposal. NetMeeting was rejected because of several factors. NetMeeting is not platform independent, NetMeeting only runs on Windows 95/98/NT/2000/XP. A large number of researchers use UNIX workstations and Macintosh computers as their desktop system. NetMeeting does not allow off-line (asynchronous) collaborations and collaborative sessions can not be saved and restored. NetMeeting, since it shares an application on a single, it does not allow remote users to access their local data files or save to a local disk.

Performance Measures:

The following milestones reflect an iterative software design process where the requirements-design-development cycle may be repeated many times.

Milestones

- Month 1 – Collect requirements
- Month 3 – Design the tool from requirements
- Month 6 – Initial release
- Month 7 – Collect user feedback
- Month 8 – Modify design from feedback
- Month 12 – Deliver tool with an install package
- Month 12 – Present results at NOAA Tech and/or professional society meeting

Deliverables

NGI collaborative tool for the Carbon Modeling Consortium. The tool will be tailored to meet the needs of the Carbon Modeling Consortium, enabling scientists to load model results and observations into the tool and provide a set of tools to administrate collaborative sessions. Once completed, the tool will be made freely available, and utilized in support of other GFDL modeling efforts and consortia.