REINAS: Real-Time Environmental Information Network and Analysis System: Phase V STATUS REPORT*

P.E. Mantey, D.D.E. Long,J.J. Garcia-Luna,
G.G. Langdon, A.T. Pang, D.M. Fernandez,
H.G. Kolsky, E.C. Rosen,C.M. Wittenbrink
(UCSC),
B.R. Gritton (MBARI), W.A. Nuss (NPS)

UCSC-CRL-97-02
January 29, 1997

Baskin Center for
Computer Engineering and Information Sciences
University of California, Santa Cruz
Santa Cruz, CA 95064 U.S.A.

Keywords: Real-time, System Design, Environmental, Sensor, Data Management, Network, Visualization, Monterey Bay, Coastal, REINAS

*Supported by a Grant from the Office of Naval Research, No. N-00014-92-J-1807

Contents

Abstract

The Real-Time Environmental Information Network and Analysis System (REINAS) is a continuing engineering research and development program with the goal of designing, developing and testing an operational prototype system for data acquisition, data management, and visualization. This system is to support the real-time utilization of advanced instrumentation in environmental science. Current applications are in meteorology and oceanography; others envisioned include geophysics, hydrology, air pollution, etc.

The REINAS project has been funded by a Department of Defense University Research Initiative through the Office of Naval Research. Participating institutions include the Naval Postgraduate School (NPS), the Monterey Bay Aquarium Research Institute (MBARI) and the Baskin Center for Computer Engineering and Information Sciences of the University of California, Santa Cruz (UCSC).

The REINAS system has been designed for regional real-time environmental monitoring and analysis. It is a modern system, integrated into the Internet, for conducting interactive real-time coastal air/ocean science. The database design of REINAS is independent of specific database technology and is designed to support operational scientific needs throughout the entire scientific data life-cycle. During the previous phases a survey of available technologies was made, selections of the those to be used in the prototype system were made, and detailed architecture of REINAS and experimentation with subsystems for data collection, data management, processing, and visualization were started.

This report summarizes the total project and documents the status of REINAS in Phase V – the final Experimentation and System Verification Phase.

Acknowledgements

We wish to acknowledge the dedicated work of all the UCSC faculty, students and staff, and especially our partners from MBARI, NPS and other interested groups. It is difficult to name all those who participated in the REINAS project, but the following contributed directly to the work of Phase V in 1996:

- **UCSC Faculty and Staff:**
  Patrick E. Mantey, Daniel M. Fernandez, Harwood G. Kolsky, Glen G. Langdon, Suresh Lodha, J.J. Garcia-Luna, Darrell D.E. Long, Alex T. Pang, Stephen Petersen, Eric C. Rosen, Craig M. Wittenbrink.

- **NPS Faculty and Staff:**
  Wendell A. Nuss, Douglas K. Miller, Richard Lind, Paul Hirshberg, Jeffrey D. Paduan.

- **MBARI Staff:**
  Bruce R. Gritton, Francisco Chavez, Kang Tao.

Students who worked on the REINAS Project during Phase V:

- **System and DBMS Architecture:**
  Eric C. Rosen, Mike Allen, Steven W. Carter, Suzana Djurcilov, Bryan Green, Theodore R. Haining, Carles Pi-Sunyer.

- **Instrumentation:**
  Eric C. Rosen, Ted Dunn, Robert Sheehan, Sume Biak, Amelia Graf, Ted Goodman, Dan Harris, Mike Allen.

- **Instrument Interfaces:**
  Ted Dunn, Jimmy Chan, Scott Lin, Chris Thomas.

- **Networking Research:**
  Chane L. Fullmer, Hans-Peter Dommel, Brian Levine, Clay Shields.

- **Visualization System:**
  Naim Alper, Suzana Djurcilov, Jeffrey J. Furman, Tom H. Goodman, Chris Oates, Michael O'Neil, Elijah C. Saxon, Jeremy Story, Zoe Wood.

- **Data Compression:**
  Paul Kanieski, David Kulp, Bryan Mealy, James Spring, Yi Zhou.

- **Editor:** Harwood G. Kolsky
  January 1997.

1. Introduction

This report[1] presents a review of the history of the REINAS project, its present status and its possible futures. Much of material contained here was presented to the ONR Review Committee in meetings at UCSC and NPS on November 19-20, 1996.

The Real-Time Environmental Information Network and Analysis System (REINAS) was started in 1992 as a multi-year project. REINAS has been funded by a DoD University Research Initiative (URI) through the Office of Naval Research. Participating institutions include the Naval Postgraduate School (NPS), the Monterey Bay Aquarium Research Institute (MBARI) and UCSC's Baskin Center for Computer Engineering and Information Sciences.

1.1 Long-term Goals

The major goal of this research is to achieve faster and more accurate weather prediction. The ocean and atmospheric processes of both operational and research interest at a regional scale, occur on time and space scales that require nearly continuous measurement, as opposed to discrete time interval measurements as used in the past. In addition, measurements and calculations must be made of both mean and fluctuating values in order to understand both mean transports and properties and turbulent fluxes. Observations are often sparse and irregular and must be assimilated into a consistent representation of the phenomena being observed. Techniques are needed to properly collect these data, assimilate them into analysis and/or forecast models, and display both the data and model output using creative visualization software. The quantity of potential data and diversity of possible models require development of sophisticated processing methods for researchers to intelligently sort out the different scales of processes that are active in the environment.

The focus of this effort is to develop the capability for real time collection, assimilation, and display of in-situ and/or remotely sensed environmental data. The research includes methodologies for data collection, compaction and storage as well as appropriate physical analysis and modeling. Integration of database information systems and visualization with direct manipulation interfaces, and development of novel means for visualizing and melding large data sets comprised of diverse data types are major goals.

1.2 Scientific and Technical Objectives

From a scientific perspective, the primary objective of this research is to develop a system to support nowcasting at a regional scale by resolving smaller space/time scales of the phenomena. This requires understanding of the processes being observed, and the corresponding requirements for a system to support their real-time monitoring. REINAS is an engineering research project with the goal of designing, developing and testing an operational prototype system for data acquisition, data management, and visualization to support the real-time utilization of advanced instrumentation in environmental science.

Specific technical objectives for the REINAS system are that it:

- Support interactive user access to real-time and retrospective data,

[1] This chapter by P.E. Mantey

- Provide the performance required to support the most frequently requested products and services,
- Provide access to resources and devices through a common data model supporting rapid system configuration,
- Support dynamic control of devices for real-time interactive scientific investigation,
- Provide fault tolerant data collection, avoiding data loss due to communication link failures, and
- Provide security features that support restricting user access and control with respect to data and equipment.

1.3 Approach

Its partners have designed the REINAS system for regional real-time environmental monitoring and analysis. REINAS is provide on the user's desk top:

- A set of tools to configure and collect data from instruments in the field in real time,
- An integrated problem solving and visualization system supporting individual and collaborative research using both historical and modeled data, and
- A logically consistent distributed data base which stores data independent of file format and which maintains metadata describing where and how data was obtained (i.e. the data base tracks the data "pedigree").

Professionals in the environmental sciences will have the ability to observe, monitor, and analyze regional oceanographic and meteorological phenomena from their desk top. REINAS serves different user groups: Operational forecasters monitor current conditions, view standard data products, synthesize new data views, and issue forecasts and warnings. Modelers analyze new model products and compare them with other models and with past and present conditions. Experimental scientists collaborate with other scientists on-line, observe individual data fields as they are collected, and may modify data collection methods for an on-going experiment. Finally, instrument engineers add new equipment to the system, access metadata describing individual devices and methods of calibration, study maintenance records, and profile sensor quality.

The systems software design for REINAS involves development efforts in several areas. Meteorological and oceanographic instruments attached to the REINAS system are augmented with microcomputers to become "intelligent instruments" and attached as Internet nodes to the system. A network of such instruments has been developed in the Monterey Bay region, employing some new instruments and some existing instruments belonging to the participating institutions: NPS and MBARI. A set of load paths is used to move data from collection devices into the system. A commercial (relational) database system is used as a central component of REINAS, augmented with software to handle the voluminous data from real-time instruments, and to insure access of current data by real-time applications and users. A set of control applications is used to deploy, configure, and steer collection devices. A set of data access tools supports basic system administration. Access to REINAS functionality is provided by a set of Application Programmer Interfaces (APIs). Visualization software supports the integrated display of data via a rich set of visualization tools which couple directly (via an API) to the database.

REINAS contributes to scientific productivity by eliminating the need for each researcher to invent their own situation-specific access routines. The REINAS load paths effectively implement automated, real-time methods to load data into the database, as opposed to the manual loading methods often used in conventional scientific databases. REINAS load paths also perform any data parsing that is required during the first steps of the data load process. The use of a relational database with a special API in REINAS provides a common data access model for several different database products. Thus, researchers who implement a scientific application using the REINAS API effectively have direct access to the data with which they work in common binary format. Unlike file specific applications, REINAS applications should continue to be of value regardless of the specific data feed, database, or machine operating system for which the necessary REINAS software exists.

2. REVIEW OF TASKS COMPLETED

The[1] Real-Time Environmental Information Network and Analysis System (REINAS) was started in 1992 as a five-year project funded by a DoD URI through the Office of Naval Research. It is a real-time spatial-temporal database system, providing data and tools for estimating the "state" of the environment spatially and temporally. The architecture and database design of REINAS is very general purpose, supporting real-time data measurements with a variety of attributes.

The REINAS group is proud of our accomplishments and of our success in meeting the goals and schedule prepared at the initiation of the REINAS project. Table 2.1 shows the overall schedule. Visitors interested in the potential application of REINAS have been surprised to find that, as an academic project, it went through the formal phases of a requirements statement, conceptual design, prototyping, and experimental evaluation. All of these stages were documented in a series of UCSC reports.

2.1 PHASE I: Concept – May 1992

From the start of the REINAS project the members have been disciplined in defining who was being served, what instruments would be supported, and the overall objectives. Also significant effort was spent in the evaluation of relevant technology. One of the early questions concerned the potential of commercially available relational databases for the support environmental data. Our conviction, which was different from the prevailing position of the scientific community at the time, was that commercially available database software could be used, and that it was not necessary to build our own database system. Our experience in REINAS has demonstrated the correctness of that conclusion. As Sequoia 2000 demonstrates, developing our own database system would have been bigger than the entire REINAS effort.

A major result of the REINAS project is the demonstration that commercially available products are appropriate for real-time applications of scientific time-series data from point sources, and from areas. Such products are being continually evolved and developed in the software industry to satisfy a large set of customers, and by developing REINAS using such commercial database software at its center, with our schema and software for loading the database and retrieving data from it, results in an architecture that can exploit developments in the database industry. In REINAS, applications were built on top of the database, using an industry standard (SQL) interface.

2.2 PHASE II: Detailed Requirements – January 1993

From a review of applications and data sources, and requirements for data capture and delivery from the potential suite of instruments, the REINAS team developed a detailed design of the database schema, the network to connect the instruments to the database, and the system configuration for the instrument nodes.

In networking, the industry standard Internet protocols, UDP, TCP/IP and IPMulticast was chosen. This allowed the exploitation of existing Internet connectivity, and the building of new REINAS links, including wireless links, using this standard.

[1]This chapter by P.E. Mantey

REINAS SCHEDULE 1992-1995

- May 1992 – Start of REINAS Project – Phase I
  - Concept Design and Documentation
  - Characterize: Instruments, Data, Users and Uses of REINAS System
  - Create Project Plan
  - Assemble Staff (real start in Fall 1992)
  - Evaluate System Technologies
  - Develop Preliminary Architecture
- January 1993 – Begin Phase II
  - Detailed Requirements Definition
  - Prototype Evaluations of Key Components
  - Refine Architecture
  - Develop Preliminary System Design
- July 1993 – Begin Phase III
  - Detailed System Design
  - Prototype Implementation
  - Development of REINAS Instrument Network
  - Connection of Real Instruments
  - Data Feeds from Other Instruments (MBARI, NPS, NOAA, etc.)
  - Addition of More Instruments
  - Database Design
  - Implementation of Data Load-Paths
  - Advanced Visualization
- July 1994 – Begin Phase IV
  - Real-Time Experimentation / System Verification
  - Visualization Directly from parts of Database
  - Support Mobile (sea and land) Instrument Platforms
  - Connection of Additional (and advanced) Instruments
    - CODARs
    - Wind Profilers
  - Develop Support for Federated Databases
  - Initiate Use of Collaborative Visualization
  - Provide REINAS Real-Time Data to Pt. Lobos ROV operations, buoy support

Table 2.1: **Schedule of the REINAS Project 1992-1995**

2.3 PHASE III: Prototype Implementation – July 1993

2.3.1 Instruments

The summer of 1993 was a very busy time. For an academic project the maximum productivity is in the summer, when faculty and graduate students can focus on research. REINAS acquired its own (Campbell Scientific) met station and installed it on the roof of the Applied Sciences building at UCSC, and configured its hardware and software according to our design. This instrument then provided our first real-time data into a database, and was made available to others. After that demonstration, others began to ask how they could also connect their met stations.

2.3.2 First Database Prototype

The first REINAS prototype database offered support via software called "X-Met", by which users could query meteorological measurements over the Internet. This combination provided a working demonstration of the viability of the REINAS instrument node architecture and the use of the REINAS database. It was maintained until early 1996.

The second met station to be configured to the REINAS design at UCSC's Long Marine Laboratory. After that data began to be fed to REINAS from a variety of other sources including MBARI, NOAA and NPS. There was a continuing process of adding more instruments, developing the database schema, designing, and implementation.

In parallel with the system work, the visualization group was building their first prototypes. These have been extensively demonstrated and documented, and provided users and visualization researchers with early tools which formed the basis for a designing the visualization support for REINAS.

2.3.3 Instrument Load Paths

Another major concern at this stage of the project was the development of the "load paths", which bring the data into the database. Our dream was for some automated procedures – a "load-path generator" whereby one could write high-level specifications to be compiled into a generic or specific database load path. Our latest (third) REINAS prototype database system has tools that greatly simplify the creation of load paths.

2.3.4 Second Database Prototype

A full-scale implementation of REINAS was begun in the summer of 1993 using the MBARI Scientific Information Model (MSIM) schema, usually called the Gritton schema, used to store environmental data for retrospective analysis at MBARI. The Gritton schema was based on a historical view of the data, e.g. the data gathered on a particular voyage or expedition. This second approach concentrated on data formatting and proved to be very labor-intensive. The implementation was done by several people working in parallel using the Oracle database. The resulting implementation had all the complex schema implemented, but it proved cumbersome for users and programmers. In use the system was unstable and had serious performance problems.

With all its problems the creation of this full prototype implementation of a complete schema was a necessary step to the success of REINAS. Until a working system was available one could not evaluate its limitations effectively and take the next step toward an operational system.

2.4 PHASE IV: Experimentation – July 1994

In Phase IV, REINAS began to supply data in a form that our experimental partners (after two years of patiently, and impatiently, waiting) could use in their research. It was late in the summer of 1994 that data was available in a way that could be used experimentally. REINAS also provided some links from the database to the visualization. In the initial prototype, the system and the visualization were built separately.

2.4.1 Additional Instruments

By the fall of 1994, support for mobile platforms and instruments was being added. The Coastal Ocean Dynamics Application Radars (CODARS) came along about that time, as well as the vertical wind profilers. This provided experience with developing a load path for that data and for putting it all in the database. The CODARs brought a new problem. Not only was there the raw CODAR data, it was necessary to develop from this data information that was useful. The individual instruments gave radial measurements, and from these one had to compute the vectors (of the surface current velocities) and also to extend the database support to store them.

At this same time, ideas for use of federated data bases were developed. The idea is that one could have different databases in different places and in different formats, all of which could appear to users as a single database. This is a difficult systems problem and work continues.

2.4.2 Collaborative Visualization

Phase IV was also the time collaborative visualization work was started. In September 1995, support was demonstrated for interactive collaboration in the task of visualizing data from the REINAS database, with collaborators in physically separate locations. In the demo, people at UCSC collaborated with others at NPS. They had synchronous communication between them over the Internet and the cooperatively visualized the same phenomena, each using different tools. That is, each user chose different visualization tools from those available via the spray-can metaphors that are part of the SPRAY visualization system. This software allowed users to collaborate using different data sets simultaneously well as ones they shared. It enabled users to add their own (personally controlled) data sets, and to combine them over the network with those of others and with shared data. Users at each site could see the visualizations created by any of the users.

2.4.3 World Wide Web

Another development was the SlugVideo cameras, which bring a visual dimension to the REINAS system, with images from the cameras also available to users via the web. To date, over 2.4 million separate visits have been recorded to the prototype video platform interface which also displays real-time conditions from a nearby REINAS weather station.

These web pages support many casual users, but also provide an effective mechanism for data access for many of the collaborators, scientists, and professional meteorologists using REINAS.

2.5 PHASE V: Operational Design – July 1995

In July 95 a major extension and refinement of the system was started. This included connection of additional instruments to REINAS and the support of more operational uses and the development of the database schema. By fall of 1995, REINAS was providing real time data in a quasi-operational context that is believed to be credible for a Navy environment. It was encouraging that National Weather Service in Monterey regularly used the REINAS data that was available and monitored our progress through the UCSC web pages.

Experiments using the second database prototype showed that simple queries of real-time interest were very complex in terms of the data structure used. This raised the question whether a single database could satisfy both real-time and retrospective users. In addition, there were problems with the system being unable to keep up with the input data rates, as well as issues concerning the growth of the historical database. How much real-time data does one really need to save for retrospective use?

2.5.1 One Database Schema or Two?

At this point REINAS again considered the possible need for two database schemas. Since the cost for storage has been declining rapidly, it would be economically feasible to store all the data twice, implementing two different schemas, one for the real-time user and the other for the retrospective user. Four years ago that would have been too expensive, but

REINAS SCHEDULE 1995-1997

- July 1995 – Begin Phase V
  - Refine and Extend System
  - Extend REINAS to Additional and New Instruments (e.g. NEXRAD)
  - Support Operational Users (e.g. air pollution monitoring, forest fire, fisheries)
  - Develop REINAS System for Port to Other Locales
  - Incorporate Video Cameras as Digital Instruments
  - Support Dynamic Network Topologies
  - Continue Evaluation of New Technologies
  - Provide Database Support for Automated Data Quality Monitoring
  - Use Combination of Models and Measurements for State Estimation
  - Support Collaborative Visualization over (high speed) network
- April 1997 – End of URI Funding (Requesting extension to end of 1997)

Table 2.2: Schedule of REINAS Project 1995-1997

now with disk prices as they are it would not be prohibitive. However, there was suggested another solution – to redesign and simplify the schema.

2.5.2 The Third Prototype – 1996

The continued use of the first REINAS database prototype (X-met) on the Internet proved the value of a simplified schema. This encouraged an effort to re-implement REINAS using a simpler architecture (known as the Rosen schema) and to build the supporting software (to put data into and to receive data from the database) by identifying common functions and implementing those functions as "tool kits". The Rosen schema emphasized supporting queries for data based on sensor stream, instrument platform or instrument site. Its implementation also demonstrates better software engineering.

The implementation of the third version of the REINAS database and data system was carried out by a different group than had built the prototype based on the Gritton schema. This group drew from the previous experience, and had a vision of an improved architecture for the data system. It also had as chief architect / programmer, Eric Rosen, who had built X-met, understood user needs and knew the deficiencies of the second prototype.

The method followed was to abandon the entire old source tree and start with a new system by defining and then implementing tool-kits which provided the key system functions. This proved to be a significant improvement. For example, in the original design there was a "real time cache", needed for logging and for performance for storing data until there was time to put it in the database. It turned out to no longer be needed because the regular data paths were now fast enough to keep up.

Overall, this re-implementation was a good case study of software engineering – build a system that has the functionality, then go back and re-implement it to achieve performance, maintainability and expandability. In evaluating the success of our recent software data system efforts, one cannot separate the system from the people who built it. It was the builders of the tool-kit who contributed greatly to the speed and quality of the implementation. We are now going back and reexamining the earlier Gritton schema with the goal of adding back some of the lost features, bringing back more support for metadata, without compromising reliability or performance.

2.6 Significance

2.6.1 A Computer at Each Instrument Site

When the REINAS project started back in 1992 with the idea of putting a computer at each instrument site, most people consulted thought that would be too expensive to use a "whole 386" at an instrument site. That was the top-of-the-line at the time. Today for a few hundred dollars (since there is no need for a display) all the compute power needed can fit in a small package at each instrument site. Today, REINAS can show instrument manufacturers how to make modern instruments that would combine their data-logger functions with a computer and have a very viable Internet instrument. This can be the new paradigm for developing a network of instruments.

The REINAS team continues to monitor technology developments, including those in database technology, video cameras, networking, wireless links, and hardware for future instrument sites.

2.6.2 Coupling Visualization with the Database

Another thing that REINAS is the most proud of is the success of its original proposal to couple visualization with the database. In most systems users take data and store it away in files in some form, and then feed those files to visualization software at a later time. In REINAS, one can look at data as it is coming into the database, using the REINAS direct links between the visualization software and the database system. This scientific work has resulted in many good technical papers and theses.

There are many opportunities related to environmental data capture, management and display, that go far beyond the present REINAS funding, as well as issues in design and performance of distributed systems. With REINAS there is now a laboratory in place that supports both environmental science and computer systems research.

2.6.3 The Two Successes of REINAS

The REINAS system faculty and students have made a big investment over the past four years in building the system to a critical size that can provide a platform for further research.

There are two major successes in REINAS: One is the creation of a laboratory to support real-time meteorology and oceanography. REINAS's implementation in Monterey Bay is unique with such a rich array of instrumentation for air-sea interaction. The data sets are being continuously gathered for real-time use, and also being saved for retrospective scientific analysis.

The other success is as a laboratory for computer system research. UCSC's interests are in visualization, networking, federated databases, and architecture issues related to system performance, reliability, and related topics. There are a large number of papers and research contributions in the computer systems and computer science literature that have come out of REINAS (see the Publications section). There are also a number of important research problems in networked real-time database systems that the REINAS system provides an environment for studying.

2.6.4 A Deliverable Prototype

At the end of 1996 there is a version of REINAS together with design and support tools that can be provided to other users. Instruments can be added to a REINAS system, without requiring the expertise and programming skills necessary during the early stages of the REINAS project.

2.6.5 In Conclusion

Although April 1997 is listed as the end of the project, UCSC plans to manage the existing funds to keep project alive until the end of summer 1997. Work will continue on adding new instruments (e.g. 88-D radar), on system enhancements, on support for distributed implementations of the data system, and on visualization and support for collaboration. Work will also continue on various efforts for combining models with measurements, although most of this is beyond the time-scale of the present project.

Currently, REINAS accepts measurements from over 70 distinct instrument platforms comprised of over 220 sensor streams located at about 50 fixed and mobile sites.

If one looks at the original concept statement it is clear that most of what was originally projected for REINAS has been done. The schedule has been met (which is rare in a software project). The methodology followed fits good software engineering principles for the development of complex software projects, which are: Take the first simple implementation as the concept, build a second one implementing all functions, then use that as the basis for a new design and implementation. The result today is a REINAS operational system that is very usable, portable and expandable.

3. REINAS Instrumentation

3.1 Introduction:

The REINAS[1] system has been designed to accept data from a variety of meteorological and oceanographic instruments. One of the main accomplishments of REINAS during Phase V has been the maintenance of a large number of different instruments on line and the accumulation of real-time data in the REINAS data base for months at a time.

In its early stages, REINAS focused upon managing data from three specific types of instruments: surface meteorological (MET) stations, Coastal Ocean Dynamics Application Radars (CODARs), and vertical wind-profilers. These core instruments provided a rich set of environmental data-streams. MET stations generate multiple parallel streams of time-series data (scalar and vector valued), profilers produce vertical arrays of vector and related scalar data, while CODARs generate two-dimensional arrays of vectors organized either radially or, when combined from multiple sites, on a regular grid bounded by the local coastline.

The current REINAS operational network consists of inputs from:

- Radio links (UHF and 915 spread spectrum)
- Leased lines
- Existing Internet links

The current suite of REINAS instruments or environmental data feeds are:

- MET stations
- Wind Profilers
- HF Radar
- CTD (Conductivity, Temperature, Depth) sensors
- Fixed Buoy Platforms, including MET, CTD, thermistor chain, and ADCP (Acoustic Doppler Current Profiler)
- Video
- Rawinsonde
- GOES satellite images
- Drifting Buoys

New or updated instruments expected:

- Mount Hamilton MET station and Video camera
- University of Michigan HF Radar
- Mount Umunhum NEXRAD

What changed in REINAS during 1996?

- New instances of MET/Wind profiler instruments
  - Mount Ununhum MET and Video camera
  - Point Sur "Rock" MET station

[1] Chapter by Dan Fernandez

  - Point Sur Navy Base MET station
- Database support for loading of sonde, image and video data
- Flexible, simple system for establishing a new instrument load path
- Simple means of loading archived or collected data from files.

Currently, REINAS accepts measurements from over 70 distinct instrument platforms comprised of over 220 sensor streams located at about 50 fixed and mobile sites.

3.2 MET Stations:

The basic type of environmental data is a sensor value, which may be comprised of an aggregate of several scalar values. When sampled temporally, the resulting *time-series* data represents the simplest data-type an environmental database system must be able to manipulate (Figure 3.1). REINAS treats time-series data as a fundamental data-type and handles a variety of different flavors; the most common are *meteorological data* from surface weather stations. Most meteorological sensors, such as those measuring air-temperature, humidity, or solar irradiance, present a time-varying scalar quantity. Others, such as a wind-monitor, present a vector, *i.e.* ⟨*wind-speed*, *wind-direction*⟩.

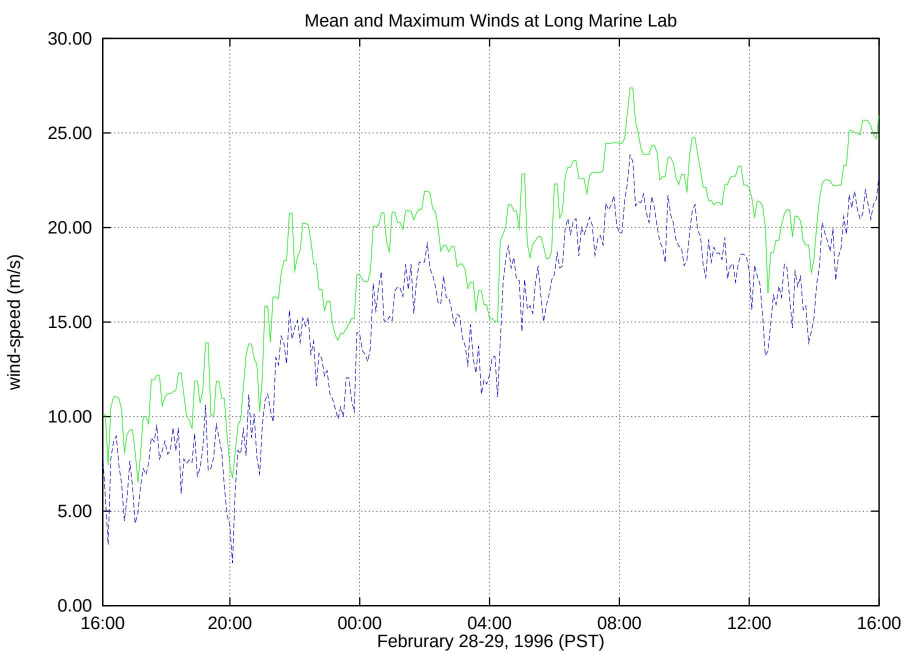


Figure 3.1: An Example of Time-series Data

MET station feeds may be classified as real-time feeds, virtual feeds, and pending feeds. "Real-time" in this case means the data comes directly from the instruments. "Virtual" means the data has been stored in intermediate files and made available to REINAS by file transfer later, usually in summarized form.

3.3 Radar Wind Profilers

Doppler radar technology allows *wind-monitor*-type data to be indirectly measured in a vertical column of the atmosphere, at intervals of hundreds of meters to heights of about four thousand meters. *Wind-profilers* thus produce a vertical profile of a tall column of air. An acoustic doppler current profiler can perform the analogous measurement using sound for a water column. Hence, vertical profiles of wind or water currents, possibly augmented with other indirectly measured variables, and updated routinely over time, are also a basic environmental data stream. Such profiles are derived from doppler spectra, which may also be included in the database.

Having a real-time link to a wind profiler results in moments (from which wind vectors may be computed) that are received approximately each minute. Receipt of the data in this fashion results in the capability to estimate the wind speed and direction vs height on a much more regular basis than the hourly data provided by the NOAA wind profilers.

3.4 High Frequency Radar

CODAR (Coastal Ocean Dynamics Applications Radar) is a version of high-frequency radar which indirectly measures ocean-currents along a collection of radials from the radar site, at fixed range intervals to distances of a few tens of kilometers. When CODAR *radial data* from two or more appropriately positioned CODAR sites are combined, *a field of ocean surface-current vectors* can be generated (Figure 3.2). Each radial or vector may also be accompanied with statistical confidence measures. A set of CODAR vector maps collected over time can be used to produce an animation illustrating ocean surface-current trends.

SeaSondes: Originally two CODAR SeaSondes were established via REINAS between two wind profiler locations in the Monterey Bay area, one at Santa Cruz and the other at Point Pinos. An older CODAR system was used at Moss Landing. Currently the system at Moss Landing has been replaced with a new SeaSonde and the system at Point Pinos has been upgraded. ONR DURIP and NSF grants have been received to purchase a new, enhanced SeaSonde to make a total of three systems.

HF radar systems measure radial ocean surface currents by measuring the Doppler shift associated with radar energy that is scattered off of ocean gravity waves that are between 5 and 50 meters in length. The difference between the Doppler shift of the returned signal and the expected Doppler shift of the ocean gravity wave allows an estimate of the value of the advection of the waves by the ocean surface current over the patch of ocean that the radar observes. Data from the CODAR units are copied to the REINAS system at UC Santa Cruz hourly, where it is then combined to form maps of ocean surface currents with a resolution of about 3 km.

U Michigan System: A new HF Radar system developed under Professor John Vesecky at the University of Michigan is currently being beta-tested at the UC Santa Cruz Long Marine Laboratory. The system utilizes multiple frequencies to get estimates of current shear within the uppermost two meters of the ocean. Once tested, the information produced relating to ocean currents and ocean waves will be fed to REINAS through the standard loader/logger mechanism.

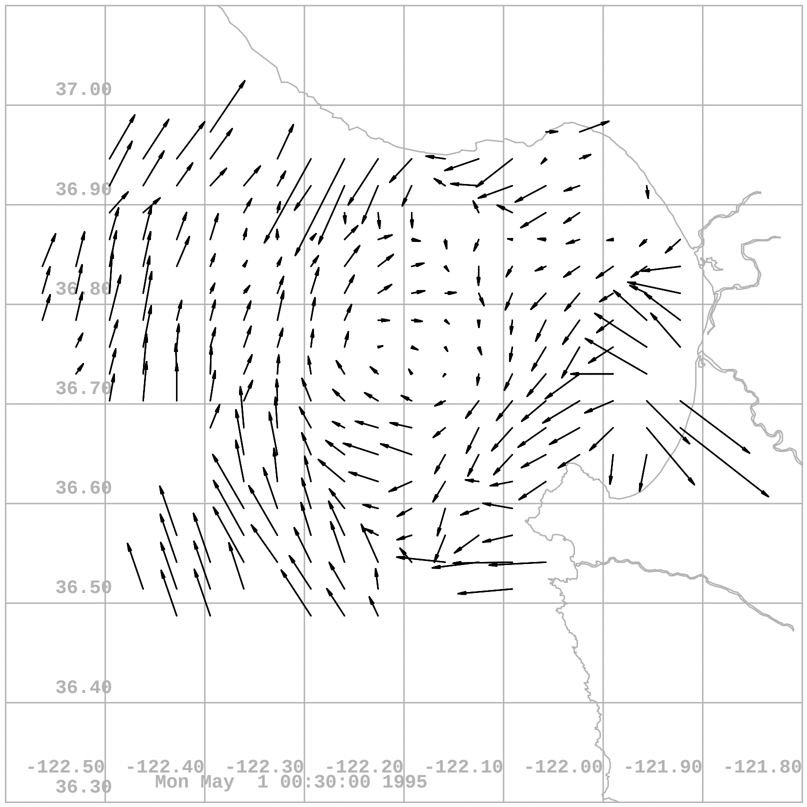


Figure 3.2: An Example of CODAR Vector Data

3.5 Biological Instruments

Other instruments that are part of the current data feed into REINAS include CTD (Conductivity, Temperature, Depth) sensors, ADCP (Acoustic Doppler Current Profiler), thermistor chains, and assorted biochemical sensing instruments, such as those that measure Chlorophyll-A or trace gases. Instruments that make these measurements exist both along the coast (CTD at Granite Canyon) and (more generally) offshore, such as on the MBARI buoys M1 and M2:

- MBARI M1 and M2 buoys (microwave link to MBARI)
  - Standard atlas MET instrumentation
  - Thermistor chains to 250 meter depth
  - CTD at surface
  - Various chemical/biological sensors (e.g. fluorescence)
  - M1 has ADCP
- CTD at Granite Canyon, NMFS (UHF link)

Experiments with satellite data have been done on:

- GOES (data collected from NWS, Monterey)
- Synthetic Aperture Radar (6 images from ERS-1 spring 1995)

3.6 Portable Meteorological Station

The development of "Port-a-Met" was completed the summer of 1994 and has been used for many demonstrations. It is a portable battery-powered MET station and REINAS PC architecture which is linked to REINAS on the ethernet via half-duplex 9600 baud radio modems from Teledesign. "Port-a-Met" was built on a trailer and is easily deployable to a location of interest.

In addition, two MET stations have been deployed on ships. One is currently aboard the vessel R/V Pt. Lobos and is Internet accessible via their existing network shipboard microwave connection. The other has recently been deployed aboard the NOAA research vessel R/V McArthur as well as the R/V Point Sur and was Internet accessible via a UHF radio link from UC Santa Cruz to the MET station aboard the vessel.

3.7 Video

In 1995 a sub-project into the development and implementation of remote video streams as real-time REINAS instruments was initiated. At a system level this included defining a prototype video instrument, building or acquiring the necessary hardware and software, deploying the camera at an appropriate location, and commencing work on integrating the resulting video data into REINAS. The prototype steerable video platform was initially deployed at Long Marine Lab (an existing REINAS instrument site) in March 1995. In August it was moved to a more appropriate site (atop a 10 story beach-front hotel) offering a better view of Monterey Bay. In 1996 a remote video camera was installed on top of Mount Umunhum. Both these images are available on the Internet and have become a very popular feature on the World Wide Web.

The high data-rates achievable with a remote camera instrument (64 kbps to 8 Mbps depending on frame rate, content, and other factors) also motivated a re-examination of the radio modem technology used to network remote sites. A new option augmenting the existing 9 kbps technology was discovered integrated into the REINAS/PC design; this technology provides IP-level data rates exceeding 0.8 Mbps.

3.8 Instrument Load Paths

Modern instruments [2] usually are configured to output data and accept commands through a generic serial interface. Typically, this interface is connected to an automated storage device or dial-up modem. In REINAS the instrument is connected to a local microcomputer which is networked to the Internet. This creates a generic and flexible connection that enhances the utility of these remote instruments. The instrument node connection to REINAS is shown in Figure 3.3.

The data sources existing and planned require REINAS to manipulate scalar and vector data in point, linear, gridded, and higher-dimensional structures. REINAS has already been extended to include other environmental data-streams, including multi-spectral satellite imagery, rawinsondes, and three-dimensional regional and global model outputs. Even with this complexity and variety of standards, the REINAS design only requires two load paths which can be made quite similar:

[2] This section from Steven W. Carter

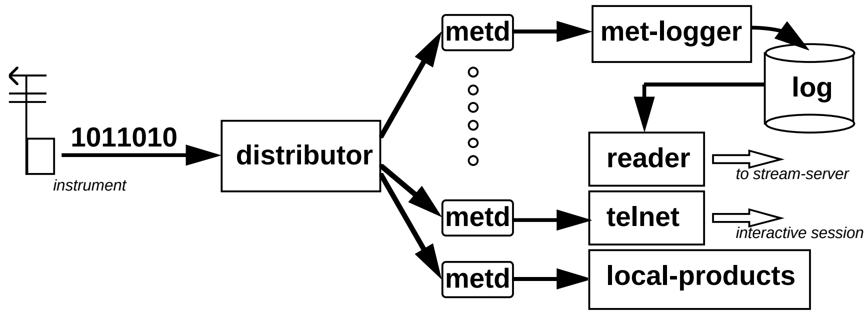


Figure 3.3: Instrument Node Connection to REINAS

- The **real-time path** in which data are made available to the REINAS database as soon as recorded by an instrument. There are five steps in the path:
  1. An instrument records data and forwards that data to a logger process.
  2. The logger process coverts the data into a form readable by the REINAS database and writes that data to a log file.
  3. A reader process reads the data file from the log file and forwards it to the REINAS stream-server using a reliable UDP protocol.
  4. The stream-server writes the data file to a log file.
  5. The REINAS loader reads from the log file and inserts the data into the database.
- The **Archival path** is one in which the data file is made available to the REINAS database after an extended time delay. Archival data may be hours, days, months, or even years old. This path is also used by the system when a data link goes down and is then restored at a later time.

  There are five steps to this path as well:
  1. An instrument records data, and at some later time (e.g. hourly) the data is placed in a file in a known location.
  2. A process parses the file, converts the data into a form readable by the REINAS database, and writes that data to a log file.
  3. Steps 3-5 are the same as for the real-time path.

Note that these data paths are very "safe". The log at the instrument ensures that once data has been recorded (parsed), there is no way for it to be lost until it is transferred over the network to the stream-server. The reliable UDP protocol ensures that no data is lost during transfer. And the log at the database end – which does not acknowledge UDP packets until they have been logged – ensures that no data is lost between reception and insertion into the database. Thus the data are preserved during every step of the path.

4. REINAS Schema

4.1 Goals

Scientific[1] data in general, and the data streams in REINAS, can be broadly categorized into one of two distinct classes: (1) environmental or *sensor-produced* data streams (also referred to as *measurement data* or sensor-streams), and (2) metadata or *data-about-the-data*. REINAS was designed to handle data of both types as well as many relationships between them.

The first REINAS experimental database schema, called "X-Met", offered support by which users could query meteorological measurements over the Internet. This combination provided a working demonstration of the viability of the REINAS instrument node architecture and the use of the REINAS database.

However, the operational needs of REINAS required that the REINAS schema employ a data model able to support a growing, and unknown set of instruments, environmental models, and data sets, and that the operations required to add, activate, maintain, and deactivate these sources of data not be excessively complex.

Since one of the fundamental goals of REINAS is to provide metadata sufficient to make sensor-produced data useful to scientists over time, the schema must preserve appropriate metadata to allow scientists to determine the history or *pedigree* of sensor-produced data when necessary.

Further, to accommodate the needs of environmental science, research, and forecasting, REINAS must support multiple data classes (numbers, sets of numbers, text, images, video, sound), structured versus unstructured data, multiple aggregation schemes for observations (point data, profile data, two and three-dimensional fields), and multiple sources (in-situ, remotely sensed, mobile platforms, historical files and databases).

4.2 The MSIM Schema

The initial major design to meet these requirements was based on the MBARI Scientific Information Model (MSIM) schema, usually called the Gritton Schema, which was used to store environmental data for retrospective analysis at MBARI.

Since REINAS seeks to store metadata about instrument configuration and maintenance as well as store data required for retrospective analysis, MSIM was expanded to include a richer description of some metadata concepts. These design changes resulted in a schema of approximately 66 metadata tables plus containers for scientific data, with a rich and varied set of data abstractions for describing both data and metadata.

Because of this schema's size and complexity, query performance was poor (especially for site-based queries or queries for real-time measurements) and query development was a difficult and time-consuming task for application programmers. The required level of detail in the metadata also made maintaining the schema very difficult, essentially requiring the attention of a full time database administrator willing and able to deal with metadata issues.

---

[1] Chapter from Eric Rosen and Ted Dunn

4.3 The REINAS Lite Schema

As a result of these difficulties, the MSIM schema design was replaced in favor of one that placed greater focus on operational convenience and real-time performance. This schema, originally termed the "REINAS Lite" schema or Rosen Schema, seeks to best encapsulate the operations of creating and activating new instruments, and reduces the number of metadata tables, identifying the following six as the essential metadata abstractions: Sites, Instruments, Instrument Types, Sensors, Sensor Types, and Deployments. The present REINAS database architecture is shown in Figure 4.1.

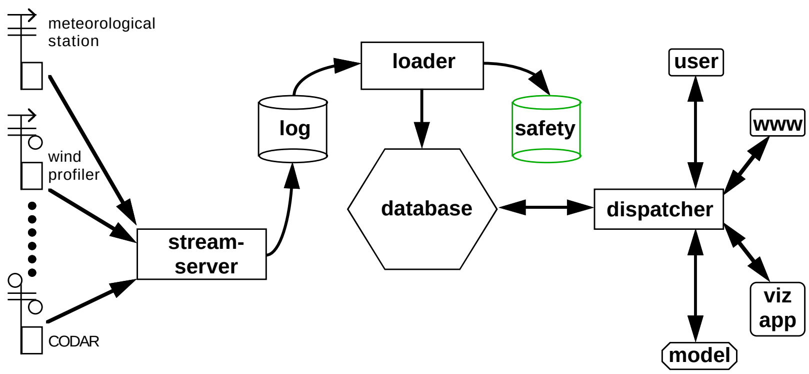


Figure 4.1: REINAS System Data Flow, Database Node View

To further improve the performance of queries for real-time, site, instrument, and sensor based measurements, the REINAS Lite schema also changed the basic model for organizing scientific data to favor the class of queries most often submitted to the database. This results in a schema which is significantly easier to maintain, and provides much improved performance, but which also sacrifices much of the metadata richness which REINAS seeks to preserve.

The supporting software for the Rosen schema that puts data into and receives data from the database was done by identifying common functions and implementing those functions as "tool kits". The main software programs and modules of the schema are given below:

REINAS-Lite Software Modules

- **Programs**
  - **metd –distributor**
  - **met-logger – generic metd station logger**
  - **reader – relays log to stream-server**
  - **stream-server – logs reader streams**
  - **loader – loads stream log into database**
  - **dispatcher – database query interface**

- **Modules**
  - **mo – "Measurement Object" encapsulation toolkit**
  - **network – REINAS networking toolkit, includes security tools**
  - **debug-logger – status message toolkit**

- **time – time manipulation toolkit**
- **parse – parsing and related TCP I/O utilities**
- **squawk – generic SQL/database interface**
- **log-manager – REINAS "RSL" facility, shared memory toolkit**
- **molog – MO logging interface**

4.4 Major Tools from the REINAS Lite Reference Manual

4.4.1 Measurement-Object (MO) Tool

```
SYNOPSIS
     The MO type and its associated utility functions implement a
     data  encapsulation  allowing measurements to be transported
     with REINAS Lite in an  instrument  and  device  independent
     manner.
DESCRIPTION
     Measurement Objects (MO) allow measurement data to be encap-
     sulated  in a device and instrument independent manner.  MOs
     are primarily used on the instrument side of the REINAS Lite
     system.  In fact, instrument loggers such as the met-logger,
     as well as the standard system modules: the reader,  stream-
     server,  and  loader, can be viewed as modules which operate
     primarily on measurement objects.

     An instrument ``logger'', such as the met-logger,  interacts
     either  with  a  generalized  instrument daemon (such as the
     metd  for  meteorological  stations)   or   the   instrument
     directly,  packaging  the  instrument  data into one or more
     measurement objects of appropriates  types,  and  logs  this
     stream of objects to disk.

     A reader reads this log of measurement objects and  forwards
     it  via  rudp_recv()  to  the  stream-server,  which  merges
     streams from multiple reader's into a single stream log. The
     stream-server  and  remote  readers  act  as  transportation
     agents for MOs, and do not  directly  manipulate  MO  struc-
     tures.

     Finally, the loader reads the stream log, parses the MOs and
     determines  into  which table in its associated database the
     measurement data should be inserted.  The  loader  does  not
     concern  itself with inserting data for a particular instru-
     ment type -- such as meteorological or  CODAR  radial  data;
     rather,  it  knows  only how to insert specific instances of
     each of the base measurement object types.
```

4.4.2 Dispatcher Tool

```
SYNOPSIS
     dispatcher [ -options ]
     telnet dispatcher-host 2801
DESCRIPTION
     The dispatcher provides a generic query interface to  ``REI-
     NAS  Lite''.   It  provides access to canned queries and SQL
     passthrough.
```

4.4.3 File2db Tool

```
SYNOPSIS
     file2db [ options ] [ file(s) ]
DESCRIPTION
     A program which parses archive files so the data may be
     inserted into the REINAS Lite database.
     It  parses  data archive files (not to be confused with data
     log files), and it either writes the data it finds to a  log
     file or sends the data to the stream-server.
```

4.4.4 Loader Tool

```
SYNOPSIS
     loader [ -options ]
DESCRIPTION
     The loader reads MO packets from a RSL(2) log manager stream
     log file, converts to local byte order, interprets each, and
     insert one or more records into the  REINAS  Lite  database.
     Normally, a colocated stream-server process receives the MOs
     and logs them to the stream log for the loader to interpret.
```

4.4.5 Net-Logger Tool

```
SYNOPSIS
     met-logger [options] [<metd-host>]
DESCRIPTION
     Met-logger - REINAS Lite Meteorological Station Logger
     The met-logger attaches to a REINAS metd  on  the  indicated
     host, registers, determines what measurements are available,
     maps these measurements to REINAS Lite  sensor  streams  via
     the   associated  met-distributor  configuration  file,  and
     enters stream mode, logging metd measurements to  an  RSL(2)
     log  manager  log file.  If no host is specified, met-logger
     attempts to connect to a metd on localhost.

     The met-logger also serves as the prototypical  REINAS  Lite
     instrument ``logger'' application.
```

4.4.6 Reader Tool

```
SYNOPSIS
     reader [-options] [host]
DESCRIPTION
     The reader reads  individual  records  from  an  RSL(2)  log
     manager  log  file  and  forwards  these  records,  without
     interpretation or conversion, to a remote stream-server run-
     ning on host, or the localhost if no remote host argument is
     provided.
```

4.4.7 REINAS System Logging Toolkit

```
SYNOPSIS
     The RSL package provides facilities for  reliably  creating,
     opening,  closing, reading, and writing arbitrary data using
     a circular log abstraction. State information regarding  the
     start and end of valid data is stored within the log header,
     making data recovery possible in the event of a crash.
DESCRIPTION
     RSL provides simple functions for  creating  and  using  log
     files  which  behave  like  circular buffers and are able to
     survive crashes of the processes reading and writing to  the
     logs  in  most cases.  The size of a log is fixed throughout
     its lifetime and determined at creation; hence, an  RSL  log
     can be filled to capacity if data is written to the log fas-
     ter than it is read.

     The log files are divided into two parts: a  header  portion
     and a data portion.
```

4.4.8 Stream-Server Tool

```
SYNOPSIS
     stream-server [ -options ]
DESCRIPTION
     The stream-server receives MO  packets  from  remote  reader
     processes  and inserts these into a local RSL(2) log manager
     stream log file, without conversion  or  interpretation.   A
     separate loader process reads the stream log, interprets the
     MO packets, and inserts one or more records into the  REINAS
     Lite database.
```

The Figure 4.3 shows the REINAS Home Page as it appears on the Internet. Figure 4.4 shows the instruments currently attached to REINAS and their status. When this page is opened the instruments are queried on the Internet as to their status at the moment.

4.5 The Next Proposed Schema

Motivated[2] by metadata omissions in the REINAS Lite schema, a project was started to go back and reexamine the earlier Gritton schema with the goal of adding back some of the lost features, such as bringing back more support for metadata, without compromising reliability or performance.

The new schema, depicted in Figure 4.2, benefits from each of the earlier design experiences. It seeks to remain close to the simplicity of the REINAS Lite schema and borrows many of that schema's core concepts. At the same time, it also adds more of the metadata facilities for recording configuration management and establishing data pedigree that the first schema possessed.

The latest REINAS schema uses the simplifying assumption that any data collected by meteorological or oceanographic activities associated with REINAS can be described by metadata that falls into one of six broad categories:

1. Data collected and stored as a part of REINAS (temperature, humidity, and wind, for example) have a known *Measurement Type* (such as scalar, vector, or grid) and that such measurement types can be re-used to describe new sources of data as they are added to REINAS. Additional *Data Type* information can be associated with a measurement type to provide details about specific environmental parameters.
2. The sources of REINAS data streams have associated with them a geographic location or *Site* which can be described as a small number of subtypes, such as points, lines, circles, rectangles, or time-dependent arbitrary data sets.
3. The sensors, instruments, data sets, calibration processes, and numerical models employed in the earth and marine sciences can be commonly described as *Data Sources* for purposes of identifying owner, manufacturer, and supplier (in the case of hardware) or process id, program name, and author (in the case of software). Classes of such sources have a common *Data Source Type* and a common data type.

[2]This section by Theodore Haining

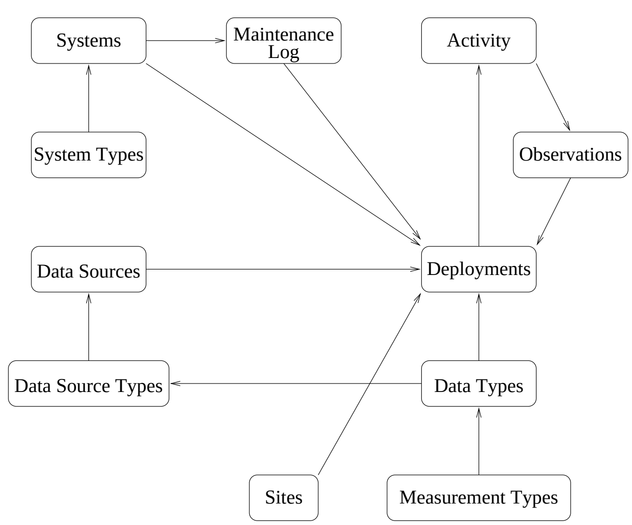
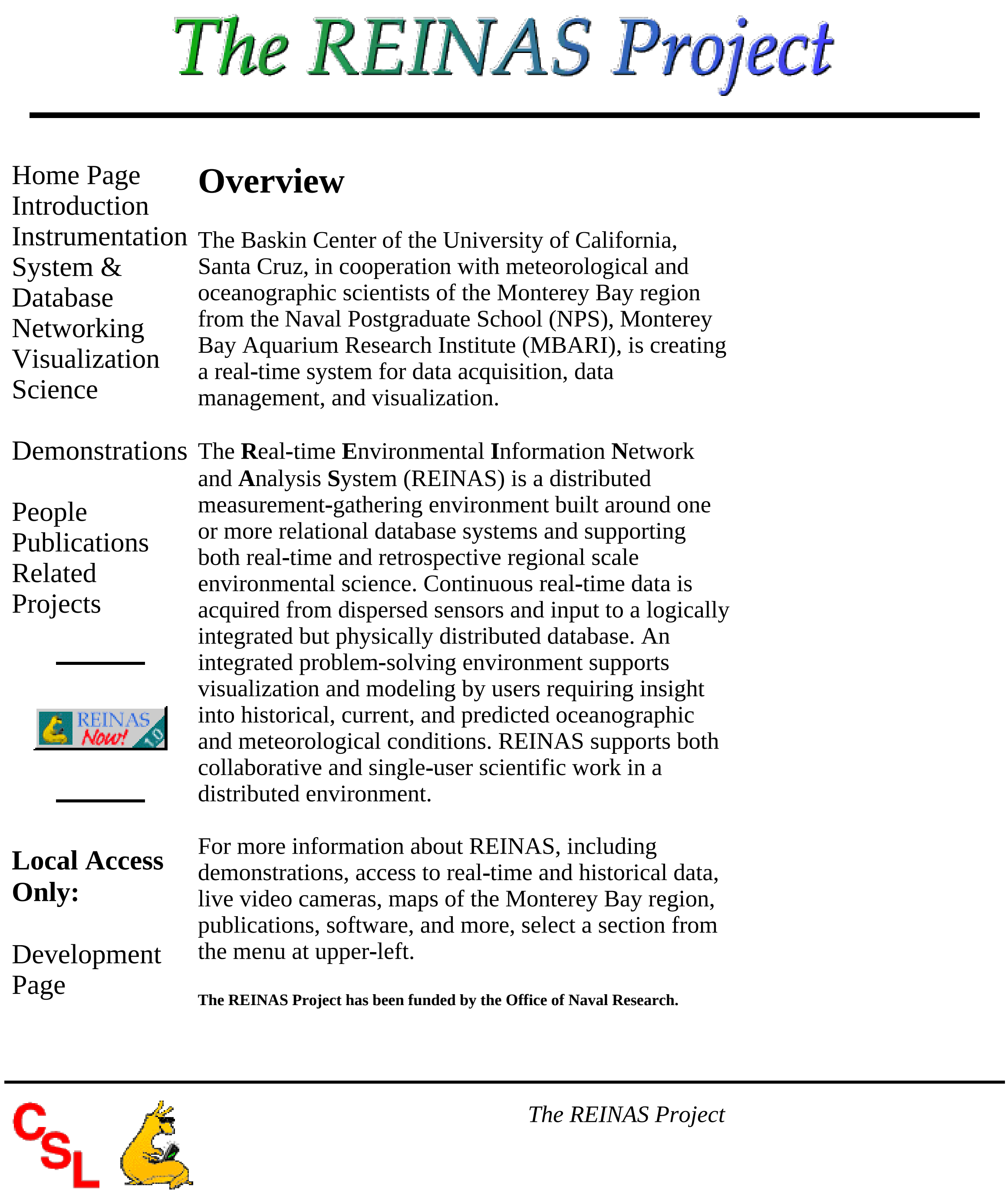


Figure 4.2: The Proposed REINAS Schema

4. For maintenance logging and configuration purposes, such data sources are part of, logically connected by, or physically co-located in a *System*, which itself is an instance of a specific *System Type*.
5. The human activities associated with the collection and storage of data in REINAS can be concisely described as an *Activity* with a person or group responsible for the data collection, and a set of free-form *Observations* used to annotate the data collected.
6. A Data Source is operated as a *Deployment* which has a data type, is part of a system, is located at a site, and is part of an activity. The data from each deployment is stored in a specified data table.

By expanding each of these categories using at most three or four database relations, it is possible to create a schema of sufficient complexity to describe environmental and scientific data in general while allowing the meaning of relational links and location of metadata to remain understandable by those operating the system.



The REINAS Project

Home Page
Introduction
Instrumentation
System & Database
Networking
Visualization
Science

Demonstrations

People
Publications
Related Projects

REINAS Now! 1.0

Local Access Only:

Development Page

Overview

The Baskin Center of the University of California, Santa Cruz, in cooperation with meteorological and oceanographic scientists of the Monterey Bay region from the Naval Postgraduate School (NPS), Monterey Bay Aquarium Research Institute (MBARI), is creating a real-time system for data acquisition, data management, and visualization.

The **R**eal-time **E**nvironmental **I**nformation **N**etwork and **A**nalysis **S**ystem (REINAS) is a distributed measurement-gathering environment built around one or more relational database systems and supporting both real-time and retrospective regional scale environmental science. Continuous real-time data is acquired from dispersed sensors and input to a logically integrated but physically distributed database. An integrated problem-solving environment supports visualization and modeling by users requiring insight into historical, current, and predicted oceanographic and meteorological conditions. REINAS supports both collaborative and single-user scientific work in a distributed environment.

For more information about REINAS, including demonstrations, access to real-time and historical data, live video cameras, maps of the Monterey Bay region, publications, software, and more, select a section from the menu at upper-left.

The REINAS Project has been funded by the Office of Naval Research.

CSL

The REINAS Project

Figure 4.3: REINAS Web Home Page

The REINAS Project

Whatsup

Please be patient as remote met stations are being called.

(typically 5 seconds to get data from each reachable station
and 60 seconds to confirm a station is unreachable)
Current time is Tue Jan 28 13:40:20 PST 1997

STATIONS	metd	met_distributor	localproducts	met-logger	reader
fiddler.cse.ucsc.edu 20 days, 20:48 load 0.14, 0.05, 0.02 /usr 93%	alive 1.21 1997/01/17 00:53:38	4-days 19:16:10 1.22 1997/01/24 01:02:20	4-days 19:16:05 1.15 1996/04/18 07:37:33	4-days 19:16:00 $Id: met-logger.c,v 1.7 1996/10/09 03:05:45 eric Exp $	13-days 0:5 $Id: reader. 1.9 1996/10/ 19:16:19 er Exp $
hydra.mbari.org 343 days, 50 mins load 0.01, 0.03, 0.22 /home 99%	alive 1.19 1996/02/15 02:47:05	273-days 14:40:24 1.19 1995/12/15 03:13:05	273-days 14:40:18 1.14 1996/04/18 05:12:34	111-days 19:32:33 $Id: met-logger.c,v 1.6 1996/10/04 07:02:43 eric Exp $	111-days 19:32:29 $Id: reader. 1.9 1996/10/ 19:16:19 er Exp $
mumble.cse.ucsc.edu 54 days, 23:41 load 0.00, 0.01, 0.04 / 64%	alive 1.19 1996/02/15 02:47:05	54-days 23:40:27 1.19 1995/12/15 03:13:05	54-days 23:40:21 1.14 1996/04/18 05:12:34	54-days 23:40:16 $Id: met-logger.c,v 1.6 1996/10/04 07:02:43 eric Exp $	54-days 23:40:17 $Id: reader. 1.9 1996/10/ 19:16:19 er Exp $

Figure 4.4: REINAS "WHAT'S UP?" Web Page

5. REINAS Application Interfaces

Data[1] residing within REINAS can be retrieved and used in a variety of ways. Although queries for current conditions tend to dominate, fairly complex queries are also possible. For some purposes, a *Structured Query Language* (SQL) interface for submitting queries is sufficient. In general, however, users and user-level applications require a more abstract interface for requesting and receiving data from REINAS.

5.1 Query Types

Present experience is that the most popular type of query issued by users is a request for current or real-time sensor values. Often these queries pertain to a specific instrument site and hence they have been labelled *snapshot queries*. Snapshot queries may return one or multiple sensor values.

It is also possible to query over a broader set of environmental data. For example, one could ask for a plot of the previous hour's mean ocean surface current speeds at every CODAR vector grid-point when the air-temperature at the Monterey Bay Aquarium dropped below a previously recorded minimum air-temperature for the region. Such complex queries may involve manipulating and returning large amounts of scientific data.

REINAS queries may involve querying non-scientific data, and the ability to support such queries has been one of the main goals of the project. The ability to find the ten year old calibration of a particular temperature probe may be essential for a researcher studying long term trends. Such queries return metadata, data about the collection process itself, which helps in the assessment of the validity or uncertainty of the data, and also makes it possible to compare data collected in different times, with different processes, and with different instruments.

5.2 Application Programmer Interface (API)

Collections of data residing in a well-designed, well-maintained database have little value unless they can be easily retrieved by user-level applications. It is also impossible to predetermine the scope of useful applications that may operate on such data collections. As a result, a predefined interface capable of supporting an extensible set of query types was required to insure the utility of data contained within REINAS.

Initially, only one such *API* existed within REINAS. A low level *REINAS-API* was constructed essentially to provide a generic framework for issuing SQL-level operations against the database. It was quickly realized that only a fairly specialized group of individuals could effectively use this API. Its primary drawback was that the application programmer had to understand not only SQL but also the REINAS schema. This requirement would be unreasonable in a system where the typical application-programmers are environmental scientists with little knowledge of database technology.

To provide better application-development support to users, a higher level API that buffers the client application programmer from the REINAS schema and SQL semantics was designed and implemented within the dispatcher through a set of predefined or "canned" queries.

[1] This section by Craig Wittenbrink

The set of predefined queries includes requests to retrieve sensor data in real-time or retrospectively, and at a given time or over a time-interval. Sites, instruments, instrument types, sensors, sensor types, and deployments can all be listed, selected, or inverted from a corresponding identifier. The geographical position of a site, instrument, or deployed sensor can also be mapped from any related entity.

Other more advanced canned queries are also available. The dispatcher also includes mechanism that allow SQL queries to be "passed through" it to the database, with the results provided to client application using the same standard interface with which canned query results are returned.

5.3 RSObjects

The dispatcher API has proven to be general, flexible, and simple enough to be used by environmental scientists who are not intimately familiar with the REINAS schema or relational database concepts. However, an additional higher-level API has also been developed within the REINAS effort to facilitate the transport of generic environmental data between a database system like REINAS and powerful multi-dimensional visualization applications which must manage and manipulate large amounts of scientific data of varying structure and size.

RSObjects is an object-oriented API implemented within the C programming language that has been developed to provide such an interface to C and FORTRAN programmers. RSObjects was developed following several other API development efforts with earlier prototype systems.

These APIs were difficult to extend, failed to isolate client applications from changes within the dispatcher and database, and were unable to support large objects (among other failures). In order to mitigate these problems, RSObjects provides a dialogue model of interaction where the application asks the database what it may request. In this way the applications maintain backwards compatibility, while also providing opportunity for querying future instruments and views without requiring recompilation.

6. REINAS Visualization Effort

6.1 Background

The visualization group[1] of REINAS has developed several data retrieval and display tools including graphical monitoring tools for environmental data, standard and customizeable forecast products, and advanced visualization tools for retrospective analysis use. These tools were packaged into a single program called "SPRAY". This was later enhanced to support collaborative visualization (CSPRAY – collaborative SPRAY). Another related work that grew out of this effort is in uncertainty visualization.

We have since evaluated and redesigned, and are currently re-implementing the visualization system for REINAS. Among the major changes are: use of OpenGL to support a wider array of platforms for running the visualization system; restructuring of the program into a set of tools callable anytime instead of a set of functionality within a given mode of operation (e.g. monitoring, forecast, analysis). We are carrying the strengths of SPRAY (collaborative, modular, extensible) forward into the current design – SLVG.

6.2 Status of PET SLVG

The REINAS visualization software is a modular, window and platform independent, application built with C++, OpenGL, and xforms. Internally, we refer to it as PET SLVG. PET is an acronym for Products, Elements, and Tools – reflecting how the SLVG program is organized and how it can distributed. As far as casual users are concerned, they can request visualization products of model forecast data and measurement data through SLVG. SLVG in turn obtains the actual data through an application programmer's interface (API) either directly from the REINAS database, or locally from (hierarchical data format) HDF files.

More technical users will discover that products are simply one or more visualization tools with datasets bound to them. They can generate customized products by activating and manipulating these tools directly. Even more sophisticated users will discover that tools are in turn made of one or more simpler elements, and that one can create new tools by simply hooking up these elements together. Because of this hierarchical and modular structure of SLVG, we can distribute SLVG as a kernel set plus a subset of elements and tools.

We have a semi-operational version of SLVG with the following features:

Global to SLVG session:

1. *Support multiple graphics windows.* Each window may have its own space/time of interest, and associated set of tools. Tools and datasets may be shared across multiple windows. The graphics in a window, with its set of tools and datasets, constitute a visualization product. In addition, windows will also be the mechanism for doing remote collaborative visualization.
2. *Region/time selector.* The region of interest is selected from a spherical or a flat world projection of the earth's coastline. This is also where the start/end time of interest as well as the time resolutions are specified. Information from the region/time selector are

[1]This chapter by Alex Pang

subsequently used to narrow the queries for environmental parameters (e.g. pressure, temperature, ...) from the database.

3. *Dataset manager.* Provides uniform interface for tools to access data that have been loaded from files or database queries. File loading have support for HDF/netCDF. Database queries have support for (a) query by station (e.g. give me information from met station at Long Marine Laboratory), (b) query by environmental parameter (e.g. give me all the temperature data from the region/time of interest), and (c) query by instrument (e.g. give me all the wind profiler data for the region/time of interest).
4. *Animation controller.* This utility provides time synchronized playback of data from multiple sources (e.g. met stations, GOES data, profilers). Time resolution for playback are also specified here. Various filter elements are available for handling disparity between time resolution of datasets and playback option.

Specific to each window:

1. *Axes mapper.* The axis within each window can be mapped to different parameters such as pressure, temperature, etc. Currently, we support the more traditional LAT/LONG/Z. Arbitrary mapping requires support for scattered data and re-sorting (see research topics).
2. *Customization.* Each window can be customized, saved, and edited. Information about tools, datasets, including window placement, etc. are included in the saved information. Each named window is a visualization product.
3. *Multi-element tools.* Support for single element and multi-element tools. Each tool has its own set of graphical user interface (GUI). Among the tools available in our current release are:
   (a) Line-of-sight tool for determining visibility/occlusions between two geographic locations.
   (b) Balloon tool for displaying multiple sonde data. A similar tool can be used for displaying floating buoy as soon as we get data for it.
   (c) Coastline tool for displaying higher resolution coastline data over selected region.
   (d) GOES tool for displaying GOES7, GOES-7 (1km), and GOES-9 (4km) images at various display resolutions.
   (e) Interpolation element for spatial/temporal interpolation of environmental data.
   (f) Quake tool for displaying fault lines and earthquake events – another application from geophysics.
   (g) Station tool for displaying information from multiple met stations and buoy sites.
   (h) Terrain tool for displaying topography/bathymetry of selected region. Variation of this tool can be used to display aircraft data as soon as we get into in/out of database.
   (i) Vector plot tool for displaying vector information such as those from codar and wind profilers. Tool provide option for different types of glyphs: arrows, barbs, and uncertainty glyphs.

Data sets:

1. Sonde data – deciphered and available through database. Awaiting regular feed.

2. ETA model data – deciphered and available through HDF files. Awaiting database load, and regular feed from NPS.
3. CODAR data – derived codar with uncertainty available, but not on regular basis, thru database.
4. Wind profiler data – available thru database. Uncertainty information still need to be derived. Regular feed available.
5. Met station data – available thru database on regular basis.
6. GOES-9 – available thru HDF files and database on regular basis.
7. Static data sets available as HDF files: terrain, coastline, and fault/quake.
8. Aircraft data – deciphered, available through HDF file. Awaiting 2D grid tool for display.
9. Drifting buoy – awaiting data from NPS.

6.3 Research Topics/Issues for Visualization

There are several outstanding research problems that still need to be addressed for REINAS visualization. Below is a brief description of each one:

1. Scattered data sets. Efficient handling/organization of data to support rapid axes mapping. Point location within scattered data sets, etc. Also important for doing volumetric modeling or interpolation of sparse volumetric data.
2. Data assimilation / ensemble forecasting (AASERT) – resolving differences between model forecasts and sensor readings. Resolving differences among model forecasts.
3. Uncertainty visualization. Methods for presenting uncertainty in vector fields beyond uncertainty glyphs; resolving differences in nested models, etc.
4. Tensor visualization. Deformation methods for displaying tensor forces in environment (e.g. velocity gradients).
5. Collaborative visualization and interactive steering – allow real- time visualization/control for remote users.

7. Network Communication

7.1 Multimedia Collaboration

Our[1] results over the past year include new protocols for multicast routing in the Internet, end-to-end protocols for reliable multicasting over the Internet, and protocols for floor control in distributed applications.

7.1.1 Multicast Routing

We have demonstrated that the Core Based Tree (CBT) protocol can produce undetected loops and fail to construct multicast routing trees. To address the limitations of CBT, we have proposed the Ordered Core based Tree protocol (OCBT). As a result of our interaction with the Internet community, the mechanisms we developed for OCBT are being adopted for the new version of CBT that will be published as an Internet Draft.

7.1.2 Reliable Multicasting

We have demonstrated that end-to-end reliable multicast protocols that organize the set of receivers into a tree are far more scalable and support higher maximum throughput than any other type of reliable multicast protocol proposed to date, such as sender initiated, receiver initiated, or ring-based protocols. Based on this result, we have designed Lorax, which is the first protocol for end-to-end reliable concurrent multicasting that uses a single acknowledgment tree. Our plan is to specify Lorax in detail, analyze its performance by simulation and analytical models, implement it, and work with other research groups in the specification of standards that incorporate our research results.

We have filed an invention disclosure covering the novel methods used in Lorax for the construction of acknowledgment trees.

7.1.3 Floor Control

Currently, no methodology exists to engineer multiparty collaboration software which encompasses solutions to the above problems. Although some specific protocols for floor control have been proposed, no explicit description or analysis of such protocols exists in the literature. In a wider context, floor control, as it is described in the references given below, yields solutions for establishing a cooperative infrastructure to bind together co-workers and computational agents. Our research is primarily targeted at improving on this gap and provide a more precise understanding of the process and effects of cooperation with specific tools, as well as an implementation of a test suite of tools based on the analytic insights.

However, it is not sufficient to provide only the middleware for orchestrating collaborative work. Since groupware merges two "worlds", the local data space and the spaces of remote co-workers, access to both worlds must be possible seamlessly, yet allowing for fine-grained security. User-interfaces must integrate both types of spaces in a way that allows to keep and enhance the local focus through remote work. Current "WIMP" interfaces let workspaces easily be cluttered with many windows and icons as "shallow-spaced" representations of the desktop-metaphor, not indicating the work status and semantics within.

[1] This chapter by J.J. Garcia-Luna

We have developed a general model and semantics for floor-controlled cooperation, based on a model of turn-taking from psycholinguistics. This provides a link between user-level cooperative behavior and the technical rendition within network layers. We have developed, validated, and analyzed new protocols to handle the variety of media and data-streams in multimedia-supported conferencing and collaboration, allowing for fine-grained sharing of objects and data. Because interaction is different for each type of multimedia (text, video, audio etc.), we look at the characteristics of each medium to derive properties that allow to design a generic protocol suite to cover diverse media in an adaptive manner.

We have developed a new protocol for floor control that can be shown to be the most efficient approach proposed to date for floor control over an internet. A variant of our approach to floor control has already been implemented for collaborative visualization, and we are currently implementing several applications to demonstrate floor control of distributed applications.

7.2 Wireless Networking

Today's internetwork technology has been extremely successful in linking huge numbers of computers and users; however, this technology is oriented toward computer interconnection in relatively stable operational environments, which cannot adequately support many of the emerging civilian and military uses and interconnection of networks. A multihop packet-radio network is an ideal technology to establish an "instant communication infrastructure" in disaster areas resulting from flood, earthquake, hurricane, or fire, support U.S. military doctrine, and extend the global communication infrastructure to the wireless mobile environment. Achieving multimedia communication on the move and instant information infrastructures presents a challenge, because of the many differences between wireline and wireless networks, the characteristics of portable devices (e.g., power levels, size), and the dynamics of large mobile environments in the battlefield and urban areas. Furthermore, commercial solutions based on base stations are inadequate to support military applications and many civilian applications, because in such applications no nodes may be static as in cellular systems.

The innovative protocols that we have developed to augment the Internet to the wireless mobile environment include:

1. Floor acquisition multiple access (FAMA) protocols that eliminate the hidden- terminal problems of multihop CSMA networks and support multimedia transport efficiently by scheduling the use of one or multiple channels distributedly.
2. Multipath routing protocols that provide multiple loop-free paths to any destination, regardless of the network dynamics, and which integrate position location and other collateral information.
3. Destination-oriented congestion control protocols that take advantage of transmission agility and support multimedia real-time applications over mobile WINGs.
4. Multicast protocols in which data distribution is not preceded by a time-consuming setup of a multicast tree, which may have to be modified after setup anyhow due to node mobility.

8. Environmental Science Value of REINAS

8.1 Design and Accomplishments

From the environmental science[1] point of view the REINAS system was designed to support (1) Real-time weather forecasting, (2) Retrospective meteorological and oceanographic research, and (3) Modeling and data assimilation on mesoscale to enhance forecasting and research.

To provide this support REINAS has made available: (1) A Database which can provide a complete set of observations, imagery, models, (2) Ready access to any date and time which facilitates rapid probing of data, and (3) Visualization tools which have been developed to meet the special display needs in environmental science. How well does the REINAS design and functioning actually support these areas?

8.2 Real-Time Forecasting

8.2.1 Goals/Needs

The basic need of forecaster is to develop a complete 3-D/4-D view of the current state of the atmosphere. Observations and images must be used to assess and validate mesoscale numerical model forecasts and analyses. The forecaster must build a conceptual model of the current mesoscale weather events in order to predict what the future state of the atmosphere will be.

8.2.2 How does REINAS meet these goals/needs?

The REINAS database and real-time data collection facilities do keep the user up to date with complete real-time observations. The original met-query tool is used to make local analyses. Mesonet data has been brought together with satellite images. The REINAS visualization tools (SLVG) have allowed the display of all types of observational, satellite, and model data in a single 3-D/4-D view. Although not complete, SLVG has great potential application because a variety of operational model data can be explored with SLVG.

The REINAS complete database allows a user to probe the full set of observations on the fly to better define atmospheric structure. The access to variety of data is fairly uniform through the dispatcher interface. There are multiple display tools available to look at observations in common formats.

Cutting edge meteorology requires the use of widely varied observing systems such as profilers, radars, high-frequency surface observations, as well as traditional surface and upper-air observations. Before REINAS these systems were separate displays in operational settings and were only brought together after considerable effort in research settings. REINAS solves this problem by providing a unified database for all types of observations.

[1] This section from presentations by Wendell Nuss, NPS

8.3 Retrospective Research

8.3.1 Goals/Needs

In retrospective research the researcher is engaged in hypothesis posing and testing. The need is for methods to probe data in a variety of ways. Complete data sets are needed for unbiased view of the atmosphere. Quality-controlled data is required to be certain of results.

8.3.2 How does REINAS meet these goals/needs?

REINAS provides complete data sets available in real-time. The same database can provide archives which allow the exploration of many data sets. The advantages are the data set is the same no matter what time is selected. A common framework of data base helps eliminate struggle with incompatible research data sets.

The querying or probing of the data set can be tailored to the hypotheses that a researcher may have. The REINAS database is essentially a laboratory where experiments can be tried easily. One hypothesis is easily tested quickly and then another can be tried without the great effort of reformatting data.

The quality of observations, while not strictly controlled, can be accounted for in REINAS database design through the metadata features. High frequency (full resolution) observations can easily be processed in a variety of ways. Long-time series can be used to check biases and consistency with other observations. And quality data, once derived, can be saved in the same database.

8.4 Modeling and Data Assimilation

8.4.1 Goals/Needs

A modeling/data assimilation system being developed as part of REINAS to support: An objective quality control of mesonet observations for nowcasting and short-term forecasting use of REINAS, and a high quality, near real-time analyses to support visualization and in-depth scientific investigation.

The model being used at NPG is the Navy Operation Regional Atmospheric Prediction System (NORAPS). The model is run in triply-nested mode with 45 km, 15 km and 5 km meshes with 36 sigma levels. The initial conditions are gathered through periodic update using multiquadric (MQ) interpolation to blend observations and first guess. A 24 hour forecast cycle is run once per day.

8.4.2 Data Assimilation Approach

An MQ interpolation used to blend first guess (1 degree NOGAPS analyses or NORAPS forecast) with mesonet observations. The blending presently consists of filling observation gaps (gap size is set for grid resolution) with first guess points, which are treated like degraded observations.

OPPORTUNITIES FOR REINAS

- **Organizations**
  - Bay Area Mesonet Initiative (BAMI)
  - West Coast Forecast System (USWRP)
  - Scripps and Coastal System
- **Science Opportunities**
  - Sea breeze studies
  - Coastal topographic interaction
  - Mesoscale quantitative precipitation forecasting
  - Model validation
  - Data assimilation
  - Mesoscale conceptual models and forecasting techniques

Table 8.1: **Opportunities for Growing REINAS**

Uncertainty weighting is assigned to each observation and first guess point used in the analysis. Presently, observation weights are same for all observations. Model weights are based on RMS deviation of model from all observations and size of gap (tends to 0 for large gaps). Interpolation is done using MQ equation with smoothing. Interpolation is 2-D univariate at 16 pressure levels.

8.4.3 Present Data Assimilation Plans

There are plans to change the model to MM5 and eventually to COAMPS due to personnel changes and need for non-hydrostatic model for small domain. We plan to run 10 km outer grid feed by 10 km meso-eta with 1-3 km inner grid over Monterey Bay region. Constraints are being added to the analysis step by using a least-squares, weak constraint approach.

REINAS observations are presently being used primarily for 2-D surface analysis for test and development of system. The real-time product has been put out on the web. Observation statistics are being collected relative to the eta model to begin to develop better quality control.

Select periods being used for 3-D analysis: (1) A student cruise with 4 soundings in Monterey Bay region during July 1996, and (2) the summer 1994 period when additional NOAA profilers were available. The analysis is in height coordinate so that slope measurements can influence free atmosphere near-by.

There are many other opportunities for REINAS in the environmental sciences (see Table 8.1

9. AASERTS

9.1 Data Compression of Scientific Real-time End-use Data for Remotely-Connected Interactive Visualization Workstations

AASERT grant number Award No. N00014-92-J-1807 Dates: 5/1/93-4/30/95 with a no-cost extension to 4/30/97

Scientific Objective: The objective is to investigate ways to compress data going out of the REINAS database on the way to the end-user. The main interest is the interactive scientific visualization application.

Approach: The approach in data compression is to obtain test images or test data, and investigate how well algorithms perform on the data.

Scientific Results: Between September 1, 1995 and August 31, 1996. A graduate student, Jim Spring, was supported including the summer part-time, and an undergraduate was supported part-time during the summer only. Jim Spring studied the lossless compression of multispectral images. A technique due to May and Spenser for predictive coding with spectral-spatial correlation was investigated. By quantizing the prediction error, the method is also useful for near-lossless compression. The results showed that the type of spectral-spatial correlation of May and Spencer was sometimes inferior to ordinary intra-frame prediction. This led to an investigation of algorithms to dynamically switch the prediction technique.

Part-time during the summer of 1996 an undergraduate, Paul Kanieski, worked on a project assessing the suitability of the forthcoming JPEG lossless and near-lossless predictive coding algorithms for compressing computer-generated versus textured (landscape) images. He extracted sequences of context-dependent prediction errors from contexts of CALIC, a high-compression JPEG lossless extension. This work resulted in Kanieski's senior thesis.

Publications:

Jim Spring, "ISSDI: Improved Spatial Spectral Delta Interleaving for Lossless Compression of Multispectral Images", Master's Thesis, Computer Engineering Department, University of California at Santa Cruz, June 1996.

Jim Spring and Glen Langdon, "Lossless Compression of Multispectral Images with Interband Prediction Error Deltas", accepted for presentation by the Asilomar Conference on Signals, Systems, and Computers, IEEE Computer Society, to be held Nov 4-6, 1996.

Paul Kanieski, "Prediction Error Characterization for Lossless Image Compression", Senior Thesis, Computer Science Department, University of California at Santa Cruz, 1996.

9.2 Integration of Heterogeneous Real-time Data Repositories for Scientific Use

AASERT Award No. N00014-93-1-1038 Dates: 9/1/93 – 5/31/96

Scientific Objective: The original REINAS design proposal described a simple heterogeneous database system. Each organization could maintain it's own copy of the REINAS database if it wished, and each of these "component" databases would be networked together to form a single, logically consistent entity called a logical database network. To make a system such as the REINAS database network useful (especially to users without prior experience with relational databases), one need which must be addressed concerns query processing. The problem is this: The method of planning and processing a query that accesses multiple component data- bases that is most intuitive to a user may actually be very inefficient. Users should not need to deal with planning queries between separate databases themselves. User queries should be transformed into more efficient equivalent forms and processed automatically.

Approach: The development of efficient query processing in multidatabase management systems (MDBMS) requires solutions to many problems in query processing. A substantial amount of research has been published in this area. Much of this work has been directed toward heuristic optimization algorithms. While these algorithms provide practical engineering solutions, they frequently offer little framework for evaluating the quality of their outputs and little understanding of their effect on the performance of the MDBMS under different workloads. As a result, the performance of a MDBMS can vary unpredictably.

We have attempted to develop a theoretical model for operation allocation and operation ordering in query processing. We will use this model to find approximation algorithms whose effect on response time and processing cost can be more accurately described. The use of these algorithms in MDBMS will result more predictable response times for different workloads.

Scientific Results: Between September 1994 and September 1995, we developed a graph theoretic model for query execution. This model uses an information size metric to characterize the problem of locating the query execution with the least cost as a Single-Source Minimum Path Problem. Queries can be optimized in this model by using relational join operations as reducers. We found that while the problem of enumerating query strategies for this problem is computationally impractical in general, it can be solved in polynomial time for chain queries.

Related Work: Many heuristic solutions to allocation and ordering problems in query processing are known to exist. Recent work has focused on a variety of heuristic approaches and for a variety of different environments. There are comprehensive surveys of research in distributed query optimization. The information size metric was chosen because it is a quantity related to both response time and processing cost: the amount of information which must be processed to complete a query. The basic assumption that makes this quantity a optimization metric is this: by reducing the amount of data that the query must process, the amount of time spent manipulating data in main memory (CPU costs), the amount time spent reading from and writing to disk (I/O costs), and the amount of time spent sending data over the network between sites (network costs) will also decrease. The smaller CPU costs, I/O costs, and network costs will reduce both processing cost and the response time.

Work by Orlowska and Zhang presented the information size metric for examining the minimizing transmission cost during query processing. They also showed how this metric can be incorporated in a query execution model to reduce overall communication cost by characterizing the problem as an Integer Linear Program.

Publications: This work has been submitted to the 1995 Annual ACM Conference for the Special Interest Group on the Management of Data (SIGMOD).

9.3 Visualization Tools for Data Assimilation

AASERT Award No. N00014-96-1-0949 Dates: 7/1/96 – 6/30/99

Scientific Objectives: The objective of this AASERT is to investigate different visual methods of presenting data together with auxiliary information, such as deviations, within the context of data assimilation. The availability of tools that incorporate these auxiliary information will positively impact the way scientists analyze model output with batch or continuously assimilated measured data. These tools can also help scientists tune forecast models, and gain a better understanding of the effects of different model parameters. Overall, the proposed work will provide visualization tools to help scientists make better and more reliable forecasts.

Approach: This work is done in close collaboration with Prof. Wendell Nuss from the Meteorology Department at the Naval Postgraduate School. From our discussions during the past 2-3 months, it has been determined that the best strategy that will allow us to quickly get up to speed on this project is to focus on retrospective mesoscale data assimilation at first. That is, we will select a past event (e.g. June 19, 1996) and compare model forecasts against observations for that period. Initially, Prof. Nuss will provide us with model output and assimilation analyses. Our work will focus on visualization tool development to highlight (a) differences between model and observations, and (b) extents and locations where the analysis and subsequent forecast differ.

After an initial suite of tools have been developed, we plan to close the loop by bringing these tools to ongoing data assimilation operations. This exercise will allow us to refine our tools, and also to use the tools to perform sensitivity analyses on models and assimilation techniques (e.g. compare effectiveness of different assimilation methods), with the eventual goal of improving reliability of weather forecasting.

Scientific Results: From September 30, 1996 through the present, we have done some literature search on visualization tools used in data assimilation, decided on the event for performing retrospective data assimilation of this region. We are in the process of transferring model and analyses output in order to test ideas for visualization tools. The first tool that we are planning will highlight places of large differences. For each observation point, we will obtain a difference with model forecasts. These differences will then be ranked, and the top X % mapped to glyphs (or icons) and displayed graphically. This tool will quickly alert the viewer to regions where there is large discrepancies between observations and model output. We plan to complete this tool by November 1996.

Related Work: This AASERT is an offshoot of the parent REINAS grant where the issue of uncertainty visualization was identified. We have subsequently followed this up and obtained funding from NSF to focus solely on uncertainty visualization. Work in this area that may later become applicable to our efforts with this project include: comparison of 3D surface attributes, comparison of direct volume rendering methods, comparison of interpolants for scattered data sets, comparison of human animation data, uncertainty in flow visualization, and use of sonification to sound uncertainty.

Publication: Alex Pang and Suzana Djurcilov, "Visualization Tools for Data Assimilation", accepted in SPIE Conference on Visual Data Exploration and Analysis IV, 1997.

9.4 Instant Infrastructure and Distributed Resource Management

Award No. N00014-95-1-1290 Dates: 8/1/95 – 6/30/98

Currently most mobile computing work is done within the confines of a predefined support network but, rapid deploy-ability and flexibility make mobile computing well suited to provide its own infrastructure and extend beyond these limits. Additionally, such infrastructure dependent models are ill-suited to support mobile computing in remote regions. The need to extend mobile computing to remote regions with no existing network infrastructure requires REINAS to develop techniques that will allow mobile computers to provide their own support network. We have proposed investigating mobile systems that would make greater use of the information available to them in order to construct a rapidly deploy able, highly dynamic, cost-effective instant infrastructure. The nature of mobile links is that bandwidth is limited and latency may be much higher than wired links.

Distributed environments where each node contributes to the total infrastructure have already proven themselves useful. In the environment we have proposed we assume that nodes are geographically distributed, share a common channel of communication and have a relatively large storage capacity. Transmission range and power limitations restrict the connectivity of a node to some subset of other nodes. Since nodes are mobile it is possible for them to move about, dynamically changing the region of network coverage. As they connect with each other, nodes exchange service requests. These requests propagate throughout the network until they reach a node that is able to service them. Although nodes share a communications channel and large local storage, they may be otherwise completely dissimilar.

While the REINAS project focuses on meteorological and oceanographic instruments, our mobile computing research is applicable to many other fields such as: a mobile sales force within a metropolitan region, space exploration, a battle group of ships (possibly integrating air and land support services), a disaster response team, or a firefighting team. As we extend the coverage of REINAS we are faced with a situation where mobile computing has stressed the limits of existing infrastructures. We believe that mobile computers are capable of providing their own infrastructure. By basing this infrastructure on epidemic replication we intend to realize an intrinsically reliable, cost efficient instant infrastructure that ill be rapidly deploy able, provide the flexibility that mobile computing requires and enable disconnected operation. To realize this network we must face challenges such as providing trust between nodes, how to effectively exploit varying characteristics, caching, latency reduction, efficient use of the limit bandwidth, supporting disconnected operation, and distributed resource allocation.

10. REINAS Publications 1994-1997

10.1 REINAS General

P.E. Mantey and D.D.E. Long and J.J. Garcia-Luna and A.T. Pang and H.G. Kolsky and others, "REINAS: Phase 4.1 - Experimentation", Univ of California, Santa Cruz, Santa Cruz, CA 95064, Report UCSC-CRL-94-43, October 1994.

P.E. Mantey and D.D.E. Long and J.J. Garcia-Luna and A.T. Pang and H.G. Kolsky and others, "REINAS: Phase Three - Systems Design", Univ of California, Santa Cruz, Santa Cruz, CA 95064, Report UCSC-CRL-94-08, March 1994.

P.E. Mantey and D.D.E. Long and A.T. Pang and H.G. Kolsky and others, "REINAS: Phase Two - Requirements Definition", Univ of California, Santa Cruz, Santa Cruz, CA 95064, Report UCSC-CRL-93-34, July 1993.

P.E. Mantey and D.D.E. Long and A.T. Pang and H.G. Kolsky and others, "REINAS: Phase One - Concept Statement", Univ of California, Santa Cruz, Santa Cruz, CA 95064, Report UCSC-CRL-93-05, January 1993.

Darrell D.E. Long, Patrick E. Mantey, Craig M. Wittenbrink, Theodore R. Haining and Bruce R. Montague "REINAS: the Real-time Environmental Information Network and Analysis System", Proc. IEEE Computer Society COMPCON, San Francisco, pp. 482-487, March 1995.

D.D.E Long, P.E. Mantey, E.C. Rosen, C.M. Wittenbrink, B.R. Gritton, "REINAS: A Real-time System for Manageing Environmental Data", Proc. 1996 Conf. Software Engineering and Knowledge Engineering (SEKE), June 1996.

Patrick E. Mantey, "State Estimation of Monterey Bay," Proc. 1996 AAAS Annual Meeting and Science Innovation Exposition, Baltimore, p. A41, February 1996.

D.M. Fernandez, "The REINAS System for Real Time and Retrospective Environmental Measurements", (poster presentation), AGU Fall Meeting, San Francisco, CA, December, 1994.

D.M. Fernandez, P.E. Mantey, D.D.E. Long, E.C. Rosen, and C.M. Wittenbrink, "REINAS: Real-time Environmental Information Network", Sea Technology, May 1996.

W.A. Nuss, P.E. Mantey, A.T. Pang, and D.D.E. Long, "The Real-Time Environmental Information Network and Analysis System (REINAS)", 12th Conference on Interactive Information and Procesing Systems (IIPS) for Meteorlogy, Oceanography and Hydrology, Atlanta, February 1996.

10.2 Networking

J.J. Garcia-Luna-Aceves and Shree Murthy, "A Path Finding Algorithm for Loop-Free Routing", accepted for publication in IEEE/ACM Transactions on Networking.

J.J. Garcia-Luna-Aceves and Shree Murthy, "A Loop-Free Path-Finding Algorithm: Specification, Verification and Complexity", Proc. IEEE INFOCOM 1995, Boston, MA, April 1995.

Shree Murthy and J.J. Garcia-Luna-Aceves, "A More Efficient Path-Finding Algorithm", Proc. 28th Asilomar Conference, Pacific Grove, CA, October 31-November 2, 1994.

Shree Murthy and J.J. Garcia-Luna-Aceves, “A Loop-Free Algorithm based on Predecessor Information”, Proc. ICCCN 1994, San Francisco, CA, 1994.

Shree Murthy and J.J. Garcia-Luna-Aceves, “Congestion-Oriented Shortest Multipath Routing”, Proc. IEEE INFOCOM 1996, San Francisco, March 1996.

J.J. Garcia-Luna-Aceves and Jochen Behrens, “Distributed, Scalable Routing Based on Link-State Vectors”, IEEE Journal on Selected Areas in Communications, vol. 13, no. 8, pp. 1383-95, October 1995.

Jochen Behrens and J.J. Garcia-Luna-Aceves, “Distributed, Scalable Routing Based on Link-State Vectors”, Proc. ACM SIGCOMM 94, London, UK, October 1994.

Shree Murthy and J.J. Garcia-Luna-Aceves, “An Efficient Routing Protocol for Wireless Networks”, ACM Mobile Networks and Applications Journal, Special issue on Routing in Mobile Communication Networks, vol. 1, no. 4, 1996.

J. Mathis and J.J. Garcia-Luna-Aceves,“Survivable Multiband Networking”, IEEE MILCOM ’87, Washington, DC, October 1987.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “A Taxonomy of Floor Control for Networked Multimedia Applications”, ACM SIGCHI 95, (submitted), 1995.

Chane Fullmer and J.J. Garcia-Luna-Aceves, “Floor acquisition multiple access for packet-radio networks”, ACM SIGCOMM 95, Cambridge, MA, September, 1995.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Floor Control for Networked Multimedia Applications”, ACM SIGCOMM’95 Workshop on Middleware, Cambridge, MA, August 28-September 1, 1995.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Floor Control for Multimedia Conferencing and Collaboration”, Multimedia Systems,ACM/Springer, 1996.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Floor Control for Multimedia Conferencing and Collaboration”, (revised) ACM Journal on Multimedia Systems, vol. 5, no. 1, January 1997.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Floor Control for Activity Coordination in Networked Multimedia Applications”, Proc. 2nd Asian-Pacific Conference on Communications (APCC)’95, Osaka, Japan, June 12-16, 1995.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Design Issues for Floor Control Protocols”, Proc. IST/SPIE Symposium on Electronic Imaging: Multimedia Computing and Networking, Vol. 2417, San Jose, CA, Feb 1995.

H.-P. Dommel and J.J. Garcia-Luna-Aceves, “Network Support for Turn-Taking in Multimedia Collaboration”, Proc. SPIE Multimedia Computing and Networking 1997, San Jose, California, February 1997.

Chane Fullmer and J.J. Garcia-Luna-Aceves, “FAMA-PJ: A Channel Access Protocol for Wireless LANs”, Proc. ACM Mobile Computing and Networking ’95, November 14-15, 1995.

Chane Fullmer and J.J. Garcia-Luna-Aceves, “Floor Acquisition Multiple Access for Packet-Radio Networks”, Proc. ACM SIGCOMM 95, Cambridge, MA, August 30-Sept 1, 1995.

Shree Murthy and J.J. Garcia-Luna-Aceves, “Congestion-Oriented Shortest Multipath Routing”, Proc. IEEE INFOCOM 96, San Francisco, March 1996.

C. Shields and J.J. Garcia-Luna-Aceves, “The Ordered Core Based Tree Protocol”, Proc. IEEE INFOCOM 97, Kobe, Japan, April 7–11, 1997.

Shree Murthy and J.J. Garcia-Luna-Aceves, "A Routing Protocol for Packet Radio Networks", Proc. ACM Mobile Computing and Networking, November 14-15, 1995.

Shree Murthy and J.J. Garcia-Luna-Aceves, "Dynamics of a Loop-Free Path-Finding Algorithm", Proc. IEEE Globecom '95, Singapore, November 13-17, 1995.

B.N. Levine, D. Lavo, and J.J. Garcia-Luna-Aceves, "The Case for Concurrent Reliable Multicasting Using Shared Ack Trees", Proc. ACM Multimedia 1996 Boston, MA, November 18–22, 1996.

B.N. Levine and J.J. Garcia-Luna-Aceves, "A Comparison of Known Classes of Reliable Multicast Protocols", Proc. International Conference on Network Protocols (ICNP-96), Columbus, Ohio, Oct 29–Nov 1, 1996.

Chane L. Fullmer, Darrell D.E. Long and Luis-Felipe Cabrera, "Performance Guarantees on ATM Networks", Proc. International Conference on Computers and Communication, Phoenix: IEEE, March 1994, pp. 290-296.

Shree Murthy, Cheng Tang and Darrell D.E. Long, "Performance Guarantees on ATM Networks", Proc. International Conference on Computers and Communication, Phoenix: IEEE, April 1994, pp. 111-114.

C.M. Wittenbrink, Arul Ananthanarayanan, J.J. Garcia-Luna, P.E. Mantey, and Bruce R. Gritton, "Monterey BayNet-ATM", ConneXions, Ole J. Jacobsen, Ed., vol. 10, no. 3, pp. 22-23, March 1996.

10.3 Visualization

Patrick E. Mantey, "A View of Visualization: Its Origins, Development and New Directions", Keynote Address, Conference on Visualization, SPIE Proc. Visual Data Exploration and Analysis, Volume 2178, pp.2–11, February 1994.

Alex T. Pang, "A Syllabus for Scientific Visualization", in Scientific Visualization in Mathematics and Science Teachings, edited by David A. Thomas, AACE, 1995, pp. 261–283.

Alex T. Pang and Michael Clifton, "Metaphors for Visualization", Visualization in Scientific Computing '95, R. Scateni, J. van Wijk, P. Zanarini, editors, Springer, pp. 1-9, 149-150.

Alex T. Pang, "Spray Rendering", IEEE Computer Graphics and Applications, September 1994, pp. 57–63.

Craig Wittenbrink, Alex T. Pang, and Suresh Lodha, "Glyphs for Visualizing Uncertainty in Environmental Vector Fields", IEEE Transactions on Visualization and Computer Graphics, September 1996.

Suresh Lodha, Bob Sheehan, Alex T. Pang, and Craig Wittenbrink, "Visualizing Geometric Uncertainty of Surface Interpolants", Proc. Graphics Interfaces'96, May 1996.

Suresh Lodha, Bob Sheehan, Alex T. Pang, and Craig Wittenbrink, "Visual Comparison of Surfaces", Technical Report UCSC CRL-95-46, University of California, Santa Cruz, 1995.

Suresh Lodha, Bob Sheehan, Alex T. Pang, and Craig Wittenbrink, "Visual Comparison of Scattered Data Interpolants", submitted to Graphics Interface'96.

Craig Wittenbrink, Alex T. Pang, and Suresh Lodha, "Verity Visualization: Visual Mappings", Technical Report UCSC CRL-95-48, University of California, Santa Cruz, 1995.

Alex T. Pang and Craig Wittenbrink, "Spray Rendering as a Modular Visualization Environment", Computer Graphics, Vol.29, No.2, FOCUS: Modular Visualization Environments, Past, Present, and Future, edited by Gordon Cameron, pp.33–36, 1995.

Alex T. Pang and Adam Freeman, "Methods for Comparing 3D Surface Attributes," accepted to SPIE Proc. Visual Data Exploration and Analysis, 1996.

Craig Wittenbrink, "IFS Fractal Interpolation for 2D and 3D Visualization", IEEE Visualization 1995 Conference Proc., 1995, pp. 77–84.

Alex T. Pang, "Visualization Support for Collaborative Spatial Decision-Making", National Center for Geographic Information and Analysis Workshop. Initiative 17: Collaborative Spatial Decision-Making, 1996.

Alex T. Pang, Craig M. Wittenbrink and Tom Goodman, "CSpray: A Collaborative Scientific Visualization Application", Proc. IST/SPIE Symposium on Electronic Imaging: Multimedia Computing and Networking, Vol.2417, pp. 317–326, 1995.

Alex T. Pang, Craig M. Wittenbrink, "Collaborative 3D Visualization with CSpray", IEEE Computer Graphics and Applications, 1996.

Alex T. Pang and Naim Alper, "Bump Mapped Vector Fields", SPIE Proc. Visual Data Exploration and Analysis, Vol.2410, 1995, pp.78–86.

Craig M. Wittenbrink, Elijah Saxon, Jeff J. Furman, Alex T. Pang and Sureh Lodha, "Glyphs for Visualizing Uncertainty in Environmental Vector Fields", Proc. IST/SPIE Symposium on Electronic Imaging: Visual Data Exploration and Analysis, Vol.2410, 1995.

Alex T. Pang and Naim Alper, "Mix and Match: A Construction Kit for Visualization", Proc. IEEE Visualization 1994 Conference, pp.302–309, 1994.

Alex T. Pang, Jeff Furman and Wendell Nuss, "Data Quality Issues in Visualization", SPIE Proc. Visual Data Exploration and Analysis, Vol. 2178, pp.12–23, 1994.

10.4 Instrumentation and Science

Daniel M. Fernandez, J.F. Vesecky, and C.C. Teague, "Measurements of Upper-Ocean Current Shear with High-Frequency Radar", Journal of Geophysical Research, accepted pending revision, February 1995.

Jeffrey D. Paduan and Leslie K. Rosenfeld, "Remotely Sensed Surface Currents in Monterey Bay from Shore-Based HF Radar (CODAR)", Journal of Geophysical Research, submitted.

Fernandez, D.M., "Remote Sensing Ocean Currents in Monterey Bay with OSCR and CODAR", Eastern Pacific Ocean Conference, Stanford Sierra Camp, Fallen Leaf, CA, September 26-29, 1995.

Fernandez, D.M., Vesecky, J.F., and Teague, C.C., "Measurements of Upper-Ocean Surface Currents with High-Frequency Radar", Proc. IEEE Fifth Working Conference on Current Measurement, St. Petersburg, Fl., February, 1995.

D.M. Fernandez and J.F. Vesecky and C.C. Teague, "Measurements of upper-ocean surface currents with high-frequency radar", Proc. IEEE Fifth Working Conference on Current Measurement, St. Petersburg, FL, February, 1995.

Alex T. Pang and Daniel M. Fernandez, "REINAS Instrumentation and Visualization", Proc. Oceans'95, San Diego, California, October 1995, pp. 1892–1899.

Daniel M. Fernandez and Jeffrey D. Paduan, "Simultaneous CODAR and OSCR measurements of Ocean Surface Currents in Monterey Bay", IGARSS'96 Conference, Lincoln, NE, 1996.

P.A. Hirschberg and W.A. Nuss and P. Sirayanone, "Numerical simulation sensitivity of the 9-11 June 1994 California southerly surge to initial conditions", AMS Conference on Coastal Oceanic and Atmospheric Prediction, Atlanta, GA, June, 1996.

W.A. Nuss and D.W. Titley, "Use of Multiquadric Interpolation for Meteorlogical Objective Analysis", SPIE Visual Data Exploration and Analysis, vol. 2178, pp. 12-23, 1994.

10.5 Databases

Darrell D.E. Long, Bruce R. Montague and Luis-Felipe Cabrera,"Swift/RAID: A Distributed RAID System',' Computing Systems, vol. 7, no. 3, pp. 333-359, 1994.

Darrell D.E. Long, Andrew Muir and Richard Golding, "A Longitudinal Study of Internet Host Reliability", Proc. Symposium on Reliable Distributed Systems, Bad Neuenahr, Germany: IEEE, September 1995, pp. 2-9.

Thomas M. Kroeger and Darrell D.E. Long, "Predicting Future File-System Actions rom Prior Events", Proc. 1996 Usenix Winter Technical Conference, San Diego: Usenix, January 1996.

Craig M. Wittenbrink, Eric C. Rosen, Alex T. Pang, Suresh K. Lodha, and Patrick E. Mantey, "Realtime Database Support for Environmental Visualization", Database Issues for Data Visualization, Proc. IEEE Visualization 1995 Workshop, Lecture Notes in Computer Science, vol. 1183, Springer-Verlag 1996.

10.6 Data Compression

C.M. Wittenbrink, G. Fernandez-Ubiergo, and Glen G. Langdon Jr, "Feature Extraction of clouds from GOES satellite data for integrated model measurement visualization", Proc. ISR/SPIE Symposium on Electonic Imaging: Image and Video Processing IV, vol. 2666, pp. 212-222, 1996 (prize paper)

Glen Langdon, Jr., Alex T. Pang, Craig M. Wittenbrink, Eric C. Rosen, William Macy, Bruce R. Montague, Carles Pi-Sunyer, Jim Spring, David Kulp, Dean Long, Bryan Mealy, and Patrick Mantey, "Compression Research on the REINAS Project", NASA, Science Information Management and Data Compression Workshop, Greenbelt, MD, on October 26-27, 1995, pp. 66-73.

Glen Langdon Jr, R. Antonucci, Skip Macy, and Dean Long, "On Compression of Data from a Meteorological Stations", Proc. Data Compression Conference, IEEE Computer Society, Snowbird UT, p 471, 1994.

J. Spring and Glen Langdon Jr, "A Study of Edge-based Lossy Image Compression Algorithms", Proc SPIE, vol 2564, Applications of Digital Image Processing XVIII, San Diego CA, pp 2-10, July 1995.

Glen Langdon Jr. and C. Haidinyak, "Experiments with lossless and virtually lossless image compression algorithms", Proc. SPIE, Still Image Compression, M. Rabbani et al., editors, 2418, 21-27, 1995.

10.7 Theses

Michael Clifton, "Direct Virtual Visualization", Master's Thesis, Computer Science Department, 1995.

Tom Goodman, "Collaborative Scientific Visualization", Master's Thesis, Computer Science Department, 1994.

Brian Neil Levine, "A Comparison of Known Classes of Reliable Multicast Protocols", Master's Thesis, Computer Engineering Department, University of California, Santa Cruz, CA 95064, June 1996.

Clay Shields, "Ordered Core Based Trees", Master's Thesis, Computer Engineering Department, University of California, Santa Cruz, CA 95064, June 1996.

Kyle Smith, "SPRAYIT: An Implementation of Spray Rendering: A Paradigm for Scientific Visualization", Master's Thesis, Computer Science Department, 1993.

Naim Alper, "Mix and Match: A Construction Kit for Scientific Visualization", Ph.D. Thesis, Computer Science Department, 1995.

Shree Murthy, "Routing in Packet-Switched Networks Using Path-Finding Algorithms", Ph.D. Thesis, Computer Engineering Department, University of California, Santa Cruz, CA 95064, September 1996.

Eric C. Rosen, "Evaluating Digital Video Stream Transmission via Packetized Wireless Channels", Ph.D. Thesis, Computer Engineering Department, University of California, Santa Cruz, CA 95064, December 1996.