

Data Management Techniques for NEOCO

The Network for Environmental Observations of the Coastal Oceans

By Margaret McManus

*Ocean Sciences Department
University of California, Santa Cruz*

John Largier

*Scripps Institution of Oceanography
University of California, San Diego*

Eufemia Palomino

Lauren Wilkinson

*Ocean Sciences Department
University of California, Santa Cruz
and*

Libe Washburn

*Department of Geography
University of California,
Santa Barbara*

Over the last two years, a new coastal monitoring system has been developed in California. The Network for Environmental Observation of the Coastal Ocean (NEOCO) addresses the need for long-term high-resolution coastal data. Seven initial observation sites have been established at marine stations affiliated with the University of California.

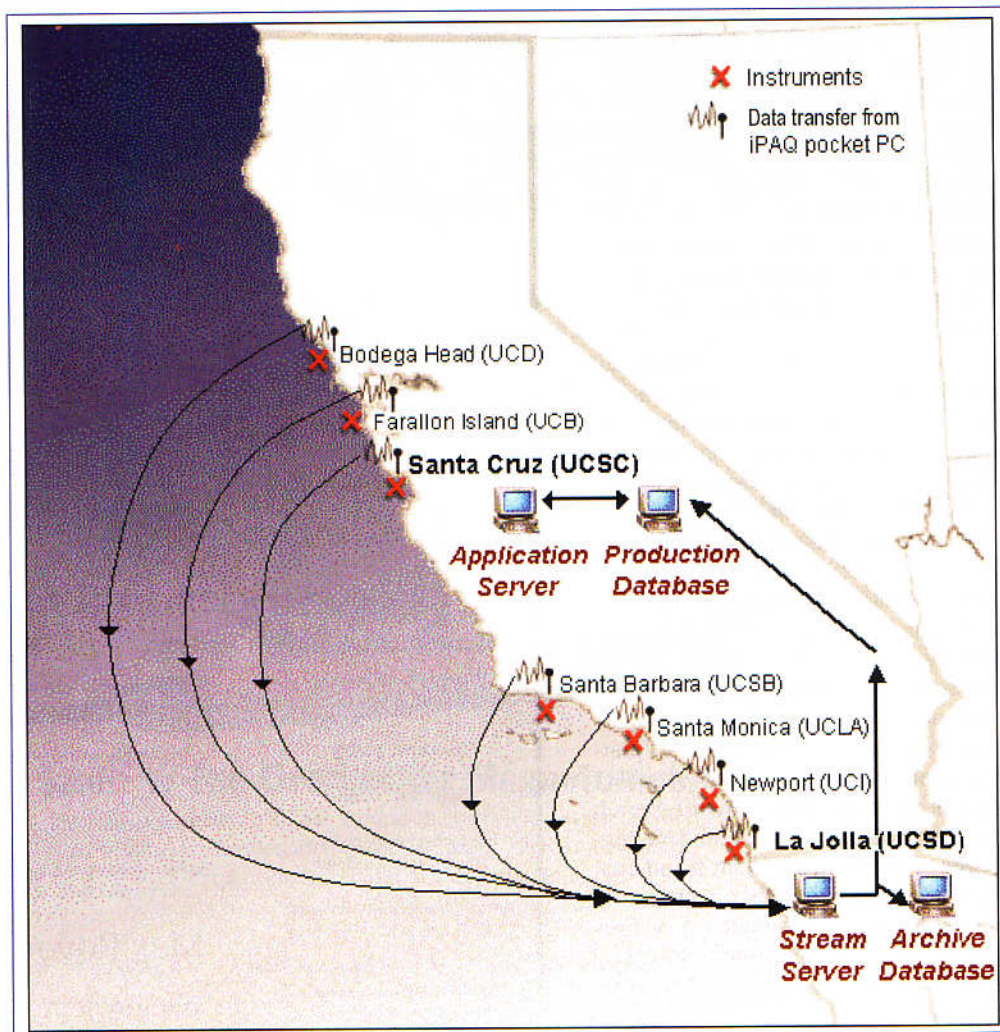
Data streams are obtained continuously using standard instrumentation at each observation site. This spatially distributed network transfers data to a relational database management system with web accessibility and reliable archiving. This article describes the NEOCO infrastructure for managing a continuous influx of oceanographic data and making this data accessible over the Internet.

System Structure

Coastal observation sites have been established at Bodega Head (University of California, Davis), Farallon Island (University of California, Berkeley), Santa Cruz (University of California, Santa Cruz), Santa Barbara (University of California, Santa Barbara), Santa Monica (University of California, Los Angeles), Newport (University of California, Irvine) and La Jolla (University of California, San Diego). The sites at Bodega Head, Santa Cruz and La Jolla are currently operational.

Angeles), Newport (University of California, Irvine) and La Jolla (University of California, San Diego). The sites at Bodega Head, Santa Cruz and La Jolla are currently operational.

Data stream from the NEOCO system showing University of California-affiliated observation sites.



This summer, the Santa Barbara, Santa Monica and Newport sites will be functional. The installation at Farallon Island has been delayed due to permitting issues.

Pier-mounted instrument packages are maintained at each observation site to monitor temperature, density, salinity, conductivity and pressure. Each instrument package has a Sea-Bird SBE16 CTD with temperature, conductivity and pressure sensors. There are plans for installing auxiliary sensors at all sites (e.g., fluorometers and

“Pier-mounted instrument packages are maintained at each observation site to monitor temperature, density, salinity, conductivity and pressure.”

transmissometers). We are currently testing the performance of these auxiliary sensors in the nearshore environment at the La Jolla site. The fluorometers that are being tested are the Seapoint fluorometer and the WET Labs ECO-FLS fluorometer with copper shutter. The transmissometers

being tested are the WET Labs C-Star transmissometer 10-centimeter and 25-centimeter path length. These instruments collect data every four minutes, with 30-second averages. Power is supplied from the pier, with a Sea-Bird battery pack providing backup power.

At each observation site, raw data streams from the instruments are logged to the memory buffer in a local Compaq iPAQ pocket PC. Each iPAQ transmits the data via Ethernet, DSL or modem to a stream server at Scripps Institution of Oceanography. This server autonomously executes several programs to handle the incoming data. Specifically, the server performs four key tasks. First, the server loads a full copy of the raw data into an archive database at Scripps. Second, the stream server performs quality assurance protocols to ensure that incoming data values fall within the expected ranges. Third, raw voltage data streams are processed and converted into standard engineering units using relevant instrument calibrations. Finally, the stream server establishes a connection with the database server at the University of California, Santa Cruz, and automatically loads the processed, filtered data into the primary NEOCO database. This database is an Oracle relational database, implemented and administrated at the Santa Cruz site. An application server, also at the Santa Cruz facility, provides web access to the database.

Backup and recovery are primary considerations in the design of this data transfer system. As with any real-time network, uncontrollable system interruptions must be anticipated. For example, a power outage may bring the network down, or instrument failure may corrupt a data stream. To maintain data integrity during such events, network recovery procedures are built into every step of the data transfer pathway. The field instruments are equipped with backup power supply for up to three days, and spare instruments are available for a rapid swap-out in the event of instrument problems. Data backups are preserved in the iPAQ memory buffers and in the archive database. Additionally, an auto-startup script has been

All trademarks and logos displayed herein are registered and unregistered trademarks of Pelican Products, Inc.

PELICAN™



How Pelican™ can protect your ideas.

- ① Panel Frame Mount
- ② Custom Foam
- ③ Shock proof Watertight Enclosure
- ④ Waterproof Bulkhead Connector Compatible
- ⑤ Custom Hot-stamp Imprinting
- ⑥ Built-in Components Applications

Get the protection your sensitive equipment and instruments need with custom applications from Pelican.

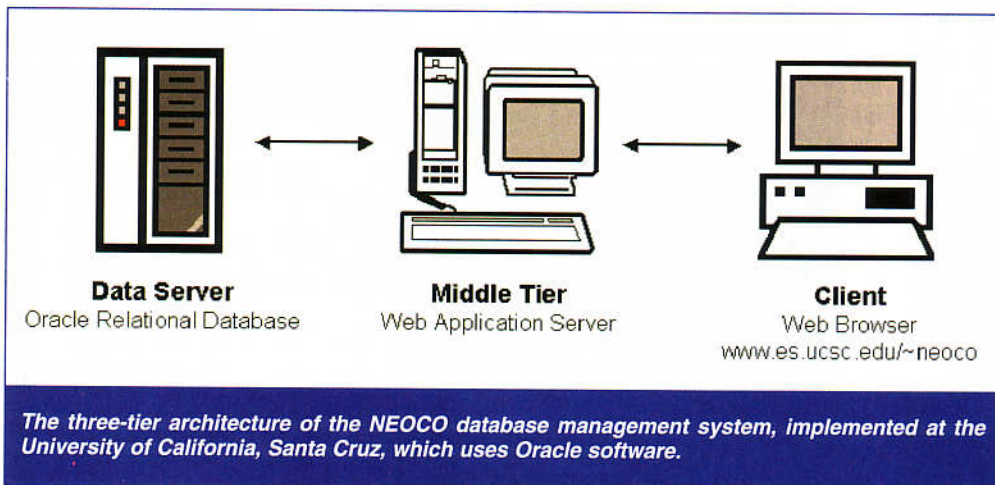
Custom foam inserts secure that perfect, snug fit. Your gear will stay in place, even when the case is knocked around. Pelican's Panel Frame Mounting System for electronic interfaces offers ease of installation with pre-drilled screw holes and stronger attachment with stainless steel retaining screws. Plus, custom decals or hot stamps for your logo on any size case.

The Pelican Protector Case™. 25 sizes, watertight, and crush proof. When you've got a good idea, protect it with a Pelican case.



THE PELICAN UNCONDITIONAL LIFETIME GUARANTEE

23215 Early Ave., Torrance, CA 90505
Tel 310.326.4700 • Fax 310.326.3311 • www.pelican.com



ity for data storage. Separate tables contain the details about observation sites, instrument calibrations and other auxiliary information. This modular design allows new sites or instruments to be added to the NEOCO system without altering the database structure. For example, when a new observation site is established, the new location is recorded in the site's table while the data from that site is written to the master data table. This way, new information can fit easily into the

developed at the University of California, Santa Cruz, to automatically bring the NEOCO database back online after power failure.

The NEOCO database management system at Santa Cruz is implemented in a three-tier architecture.¹ The bottom tier consists of an Oracle 8.1.7 relational database, which stores the data. The design of this database was carefully planned in light of the size and scalability issues unique to a long-term, real-time monitoring system.² Because the database is publicly accessible, many different types of

queries are expected. All logical data are stored in one master table to enable maximum flexibility in query options. Independent data parameters (i.e., temperature, salinity, etc.) are assigned to attributes in a relational table keyed by timestamp and observation site. This main table receives a continuous influx of roughly 10,000 rows of data per month from each of the observation sites. The table is indexed spatially to prevent its large size from slowing query performance. An internal 80-gigabyte hard drive supplies the required disk space capac-

existing database structure, without requiring a new table each time project expansion occurs.

An Oracle 9i 1.0.2.0.1 Application Server functions as a middle-tier interpreter between the NEOCO database and a web client.¹ The application server executes a program developed at the University of California, Santa Cruz, that parses query requests from a web user. This program embeds HTML and dynamic SQL in the Oracle database language PL/SQL.³ When a client's request has been parsed, the application server connects

SEA CON[®]

UNDERWATER ELECTRICAL AND FIBER OPTIC CONNECTORS

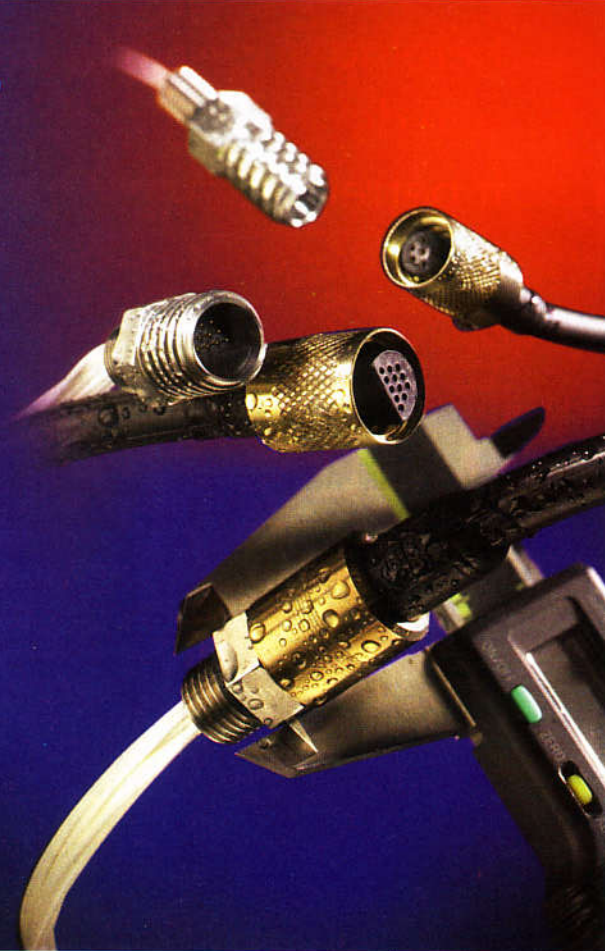
NEW HUMMER SERIES

PATENT PENDING

- Ultra Miniature
- High density
- Low Cost
- Brass or Stainless Steel Bulkhead
- Rated to 10,000 psi
- Contact Configurations (2 to 28)

See the Hummer range on SEA CON[®]'s stand 527 & 529 Oceans 2003 Conference, San Diego, California, 22nd - 26th September

SEA CON[®] Global Production, Callejon Terrazos #8, Local 2-C, Las Brisas 1ra. Seccion, Tijuana, B.C., Mexico C.P. 22610.
Tel: +52 (664) 626-2726 Fax: +52 (664) 686-8922
Email: sales@seacon-global.com • Website: www.seacon-global.com



OFFSHORE OFFSHORE

What
you do
underwater
is your
business . . .

Ultra Short
Baseline Acoustics

- Navigation
- Tracking
- Positioning

Communication
& Control

- Acoustic Release
- Transponders
- Recovery Systems

. . . helping
you do
it better
is ours.



4 Little Brook Road
West Wareham, MA 02576
Tel. 508-291-0960
Fax 508-291-0975
e-mail: sales@ore.com

www.ore.com

CIRCLE NO. 45 ON INQUIRY CARD

a. Specify your query:

Run Query Reset

Query Options
Choose data fields to query:

- Conductivity
- Salinity
- Temperature
- Pressure
- Sigma_t
- Fluorescence
- Chlorophyll Absorption
- Transmittance
- Beam Attenuation Coefficient

Start Time
("MON-DD-YYYY HH24-MI-SS")
(eg. JUN-11-2002 14:00:00): APR-01-2003 08:00:00

End Time
("MON-DD-YYYY HH24-MI-SS")
APR-07-2003 20:00:00

Choose Observation Sites:

- La Jolla, CA
- Newport, CA
- Santa Monica, CA
- Santa Barbara, CA
- Santa Cruz, CA
- Farallon Island, CA
- Bodega Head, CA

b. Results from the NEOCO Database

Observation Site	Date	Salinity (psu)	Temperature (°C)
La Jolla, CA	APR-01-2003 08:03:24	33.265099	15.8236
La Jolla, CA	APR-01-2003 08:07:24	33.265598	15.9164
La Jolla, CA	APR-01-2003 08:11:24	33.264702	15.8988
La Jolla, CA	APR-01-2003 08:15:24	33.2659	16.051399
La Jolla, CA	APR-01-2003 08:19:24	33.264702	16.0222
La Jolla, CA	APR-01-2003 08:23:24	33.266701	16.0865
La Jolla, CA	APR-01-2003 08:27:24	33.267601	16.1861
La Jolla, CA	APR-01-2003 08:31:24	33.2677	16.174601
La Jolla, CA	APR-01-2003 08:35:24	33.266602	16.1166
La Jolla, CA	APR-01-2003 08:39:24	33.266701	16.065399
La Jolla, CA	APR-01-2003 08:43:24	33.265999	16.044399
La Jolla, CA	APR-01-2003 08:47:24	33.265202	16.005199
La Jolla, CA	APR-01-2003 08:51:24	33.265099	15.9957
La Jolla, CA	APR-01-2003 08:55:24	33.265999	16.053699
La Jolla, CA	APR-01-2003 08:59:24	33.264999	16.048
La Jolla, CA	APR-01-2003 09:03:24	33.2644	16.0471
La Jolla, CA	APR-01-2003 09:07:24	33.265202	15.9969
La Jolla, CA	APR-01-2003 09:11:24	33.2649	16.0114
La Jolla, CA	APR-01-2003 09:15:24	33.2645	16.084801

c. La Jolla: Last 7 Days

salinity :: temperature :: pressure :: sigma_t :: chlorophyll :: transmittance :: beam_atten

engineering plots

Salinity (psu)

Temperature (°C)

Back to Top

Web interface for accessing the NEOCO database: (a) data query interface, (b) example download of data in HTML format, (c) example of seven-day plot automatically generated four times daily.

to the database, executes the appropriate query and ultimately gives the results back to the web client. The client, or end web user, represents the top tier in the data management architecture. The structural details of the database and application server are transparent to the end user. Any public user may query the NEOCO database via a web browser.

Web Interface

The end result of this coastal monitoring network is an extensive repository of data, accessible over the Internet. The NEOCO homepage (www.es.ucsc.edu/~neoco) provides a link to access the database. This link provides web users with three options for viewing the data. The first, most flexible option is to retrieve data in HTML format. Using a simple check-box query form, researchers may request any combination of data parameters, from any specified time frame and any combination of observation sites. For example, a researcher may

wish to view salinity and temperature values from La Jolla over the period of June 13-16, 2002. This query can easily be executed by selecting the appropriate check boxes and clicking the "Run Query" button. The customized subset of data is then extracted from the NEOCO database and returned to the user in HTML. The second option for database access is to download data in ASCII or Excel formats. With this option, researchers fill out a check-box query form to download all data parameters from any combination of observation sites over any specified time frame. Finally, researchers may quickly assess coastal trends by viewing seven-day time-series plots of NEOCO data. These plots are automatically generated four times daily for each site.

Rationale

Of the standard database systems available, a relational database was determined to be the best choice for NEOCO. Relational database technol-

ogy allows the user to extract subsets of the data.² Customized queries can be made to retrieve just those portions of data in which the user is interested. This subsetting capability is ideal for a long-term research facility like NEOCO, which must accommodate a broad spectrum of user interests.

Oracle software was chosen to support NEOCO for several reasons. The size of an Oracle database is unlimited.⁴ Any amount of data may be stored, provided the hardware contains enough disk space. Such a large-scale database is necessitated by the large-scale data flow expected in a coastal monitoring system. NEOCO must manage roughly 10,000 incoming rows of data per site per month. Future project expansion would increase the required database size capacity even further. The NEOCO database has also been designed to include new data types in the future. Oracle databases provide the built-in flexibility to store any type of data, including ASCII, binary data, satellite imagery and video clips.⁴ Finally, Oracle vends software specialized for building and deploying web applications.¹ This software provides the framework for developing and maintaining an online user interface to the database. An Internet-ready platform is essential to make NEOCO data available for students, researchers, legislators, policy makers, managers and the public.

Web access to oceanographic data greatly facilitates environmental literacy. The NEOCO online database offers a public resource, with relevance to coastal water quality, fisheries management, marine conservation and coastal recreation.

With rapid technological advancements in instrumentation, ever-increasing volumes of scientific data can be collected in oceanographic studies. However, accessibility of this data is limited by the rigor of the database. Data management systems need greater capacity and stability to keep pace with growing volumes of data. The database architecture developed for NEOCO provides an operable solution to the challenge of successfully managing a large, continuously expanding influx of data. The NEOCO tools for data management can be applied to other oceanographic projects with large data sets.

Future Work

Future work is planned to enhance the impact of NEOCO through collab-

orations with other institutions and educational outreach activities. Project expansion is anticipated at San Luis Obispo (California Polytechnic State University) and at the Romberg Tiburon Center (San Francisco State University).

In addition, historical daily temperature data from the multi-decadal Shore Stations⁵ program based at Scripps is available from the website. Other institutions are invited to make use of the NEOCO database infrastructure; additional sites will enhance

the value of the network by providing extended spatial coverage.

Summary

A reliable data management infrastructure has been developed for the NEOCO system. Real-time data streams from Bodega Head, Santa Cruz and La Jolla are now available online. The work in progress aims to connect all seven observation sites to the online network during the summer of 2003, and to incorporate other sites along the coast of California. At this

POS MV featuring

TrueHeave

Provides **immunity to long period swell heave errors**

Eliminates the need for line run-ins

Enables data collection through turns



Applanix's TrueHeave technology is just one more reason why POS MV remains the number one choice of the world's most demanding hydrographers. Best of all, this technology is available for both new and existing POS MV systems.

Contact us today to find out how you can upgrade to Applanix's TrueHeave technology.

APPLANIX
MARINE

Tel: (713) 896-9900
Email: marine@applanix.com

www.applanix.com/trueheave

SEE US AT OCEANS 2003—Booth #409

time, the public is invited to make use of the NEOCO data system.

Co-Authors

Other authors of this article include Steven Morgan from the Bodega Marine Laboratory, University of California, Davis; Keith Stolzenbach from the Institute of the Environment, University of California, Los Angeles; Brett Sanders from the Department of Civil and Environmental Engineering, University of California, Irvine; and Mark Stacey from the Department of

Civil and Environmental Engineering, University of California, Berkeley.

Acknowledgements

Special thanks to the following individuals for their participation in the NEOCO project: J. Shannon Scott and Fred Wright from the University of California, San Diego; Patrick McEnaney, Ed Boring and Gretchen Zwart from the University of California, Santa Cruz; David Menzies from the University of California, Santa Barbara; and Victor

Cho from the University of California, Davis. The NEOCO project is supported by the University of California Marine Council. /st/

References

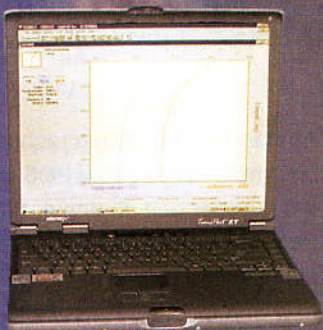
1. Vesterli, S.E., Oracle Web Applications 101. Oracle Press, Osborne/McGraw-Hill CA. 2001.
2. Silberschatz, A., H.F. Korth and S. Sudarshan. Database System Concepts. McGraw-Hill USA. 1997.
3. Feuerstein, S., and B. Pribyl. Oracle PL/SQL Programming, Second Edition. O'Reilly & Assoc. Inc. CA. 1997.
4. Ernst, B., H.R. Rasmussen, U. Schwinn and V. Venkatachalam. Enterprise DBA Part 1A: Architecture and Administration Student Guide. Oracle University, Oracle Corp. Inc. CA.
5. McGowan, J.A., D.R. Cayan and L.M. Dorman, "Climate-Ocean Variability and Ecosystem Response in the Northeast Pacific," *Science* 281, 210-217. 1998.

Visit Sippican at booth B-218 during Oceans 2003 in San Diego.

Portable Digital Data Acquisition

Sippican's MK 21/USB is a self-contained data acquisition system compatible with USB drivers using Windows® 98/2000/XP. Now there's no need to install your MK 21 on an ISA computer; the MK 21/USB will go anywhere your new computer will go. The MK 21/USB comes complete with 19-inch, rack-mount hardware.

The MK 21/USB is compatible with all variants of Sippican expendable probes and launchers.



Sippican, Inc.

Seven Barnabas Road Marion, Massachusetts 02738
TEL (508) 748-1160 FAX (508) 748-3626 www.sippican.com
Sippican is an EEO/AA Employer and an ISO-9001 Certified Company.

Margaret McManus is an assistant professor of ocean sciences at the University of California, Santa Cruz. She holds an M.S. in biological oceanography from Old Dominion University and a Ph.D. in physical oceanography from the Center for Coastal Physical Oceanography at Old Dominion University.

John Largier is an associate research oceanographer in the Integrative Oceanography Division at Scripps Institution of Oceanography, University of California, San Diego. He holds a Ph.D. in oceanography from the University of Cape Town.

Eufemia Palomino earned her master's degree in physical oceanography from Moss Landing Marine Laboratories. She is an associate specialist at the University of California, Santa Cruz.

Lauren Wilkinson was a programmer/analyst at the University of California, Santa Cruz, during her work with the NEOCO project. She is currently earning her master's degree in information management and systems from the University of California, Berkeley. Wilkinson received her B.S. in aquatic biology from Brown University.

Libe Washburn is a professor at the University of California, San Diego. He earned his master's degree in engineering sciences and his doctorate in engineering sciences (fluid mechanics), both from the University of California, San Diego.