NAVAL ENVIRONMENTAL OPERATIONAL NOWCASTING SYSTEM (NEONS)

Ted L. Tsui
Naval Oceanographic and Atmospheric Research Laboratory,
Atmospheric Directorate
Monterey, CA 93943-5006, USA

1. INTRODUCTION

A major challenge faced in today's oceanographic and atmospheric data processing (Gardiner et al., 1989; Bergen et al., 1989) is to keep up with the unprecedented rate of change in computer technology and in the methods of geophysical data gathering. Even while the system is being developed, much of the hardware and software becomes obsolete. New sensors are continually being introduced which increase the accuracy of the data and frequency of the observation.

In the shadow of budgetary constraints, designing and developing a database management system becomes an venturesome undertaking. To anticipate the current data processing trends and their significance to the oceanographic and meteorological community (Sadowski, 1989; Soderman and Gibson, 1989; Hibbard and Santek, 1989), one must consider the following factors:

- Increasing volumes of observed data,
- Distributed processing on heterogeneous computer networks,
- Extended databases, and
- Expanded 4-D data assimilation.

In addition, the system must be flexible to accommodate all data types, extensible to include new data types and processing functions, and finally portable to migrate to new hardware and software platforms.

To be cognizant of these trends and challenges, we have developed the Naval Environmental Operational Nowcasting System (NEONS), an environmental database management system. The followings are the system design principles.

1.1 Data Centered Design

NEONS takes the data-perspective, as opposed to the process-oriented, approach to the system design. Similar incidences of environmental data sets are grouped into one generic data type. Only one set of data accessing and managing routines is associated with each generic data type. These utility routines are built with modular construction and used many times to discourage overlapping or duplicating routines.

1.2 Commercial Software

NEONS relies on commercial software as much as possible for several reasons: shortening the applications software development time, keeping the cost down, advancing with computer technology via the commercial market, and availability over multiple hardware platforms.

1.3 Industry Standards

To achieve applications software extensibility and portability, NEONS system software adheres to computer industry standards. Only industry standard commercial utility software packages are integrated with the NEONS.

1.4 Data Butler Concept

The difference between a data server and a data butler is that the latter delivers the "desired" information to users in the format with which they are familiar. In other words, some intelligent data preprocessing are performed before each delivery (e.g. calibration, navigation, collocation, units adjustments, format conversions, etc.). "Bad data" are quality controlled before shipping to the users.

1.5 System Development Documentation

NEONS system development has followed the modern software engineering approach. The Detailed Design Document and Interface Document have a periodical update cycle. These documents are available on the NEONS in the soft copy format and can be reviewed anytime by the users.

Studies have repeatedly shown that most of the software cost is in the maintenance phase, including fixes, enhancements and porting. Emphasis on standards, modularity, and data modeling means that it is easier for the system to be flexible, and to be extended and ported to accommodate new environmental data types, new functional requirements, and new hardware platforms.

2. SYSTEM ARCHITECTURE

NEONS layered architecture, shown in Fig. 1, illustrates the basic design philosophy. The system configuration is separated into four tiers; a commercial hardware system at the foundation, industry standard commercial system software above the base, the NEONS system software, and various applications programs at the top.

The current NEONS computer hardware platform is a UNIX minicomputer workstation. However, the NEONS software has been ported to several other UNIX mini-computer workstations of different vendors and a UNIX PC/386 computer. Users can manipulate data via

X-Windows (V11/R4) from other systems on the network. NEONS also has used two different commercial relational Database Management System (DBMS) Software.

The NEONS system software includes four major components: a data handling subsystem, application program interface, data ingest module, and an X-Browser. The data handling subsystem controls the commercial relational DBMS engine and coordinates the data ingest schedule. The application program interface consists of two libraries: FORTRAN and C libraries. These libraries provide the data/program interface subroutines for user application programs. The X-Browser is an X-Windows based data inventory utility. It browses the database for the available data sets on demand, extracts data upon request, converts data to proper format, and delivers to the users. The X-Browser also includes a graphical function which can provide the data coverage diagram for the users. Fig. 2 shows the connectivity among these functional areas.

3. DATA HANDLING

The NEONS data handling subsystem manages a wide variety of environmental data from diverse sources. NEONS provides a flexible extensible data organization that supports new data types and new methods of deriving environmental information, and governs automatic data management including ingestion, extraction, deletion, and off-line storage transfer.

3.1 Data Organization

The NEONS conceptual data model organizes all environmental data into three generic data types. Having few generic types simplifies the storage and access functions, and in turn streamlines and reduces the amount of software required. Hence, the user interface software is based on generic categories and not on specific data sources or occurrences. Having fewer generic types also provides data extensibility, since most new data sets introduced to the system fit into the preexisting types. The three generic environmental data types are "Image" type for satellite data, "Grid-Point" type for numerical model output, and "Latitude-Longitude-Time" (LLT) type for observations. The NEONS internal Grid-Point format is very similar to the GRIB format and the LLT internal format is very similar to the BUFR format.

In addition to the three data types, the geographical data are stored separately. The data include coastlines, bathymetry, topography, and land surface type. Geographic data are distinguished from the environmental data because of their independence from time characteristics.

3.2 Data Archives

Commercial archival software is linked with NEONS to create an "unlimited" on-line data storage capability. Some commercial DBMS vendors provide solutions to the data archival function. Some commercial vendors employ the third party vendor product to solve the problem, because their DBMS relational tables are regular UNIX files. All archival software systems for storing UNIX files can be used to perform the data archival functions. Currently, NEONS uses the optical disk library as the archival medium.

4. DATA INGEST

As shown in Fig. 2, data sets of a generic type are individually ingested into the database. Individual format converters check data qualities and units, fill the database description tables, and prepare the data sets for loading. A generic loader associated with each data type loads the data and updates the database inventory.

5. APPLICATION PROGRAM INTERFACE (API)

All application programs access data via the API routines. These routines are grouped along the three generic data types and are designed to provide high-level "logical" access to the environmental data based on geophysical properties such as time, latitude, longitude, and parameter. Thus, Image data are accessed by sensor, region of interest, time, and channel (band). Grid-Point data are accessed by model, parameter, level, time and forecast time. LLT data are accessed by parameter, report type and latitude, longitude, and time range. The NEONS/API layer insulates application programs from storage implementation details such as packing, table names and on-line/off-line data status.

Co-registration of environmental data of different data sets or types is performed in the API layer. Co-registration involves collocation and mapping image, grid-point field, LLT, and climatology data to a user defined grid of arbitrary projection, size or location. Any of the data stored in the database including geographic data can be easily combined and integrated in support of the data assimilation and fusion research.

The NEONS/API provides application programs with a standard interface to the database and promotes the development of reusable code. Each routine is written in C language embedded with industry-standard Structured Query Language (SQL) calls to communicate with the commercial relational DBMS. The API also includes a FORTRAN binding to link with FORTRAN programs.

6. X-BROWSER

6.1 Data Browser

This data browser, providing users with an on-line database inventory facility, is based on the Graphical User Interface design and achieves a user friendly network environment. The X-Browser provides users with an off-line as well as on-line database inventory facility. As in the API, users browse the database contents on the basis of geophysical characteristics. For example, one can query the database across the network in a matter of seconds to find which numerical model output fields within a specific period of time are presently available. Or, one can determine the archival status of satellite images for a region of interest within a time interval. Summary information about each data set such as the time span, location, processing history, size, resolution, ingestion date, status, etc. is presented with each data set (ordered by time).

NEONS X-Window Browser is implemented according to the X-11/R4 Window System standards and MOTIF convention. Since all objects are derived from the MOTIF toolkit, they all have a consistent look and feel. Preserving a common look and feel across different configurations, users can create new applications rapidly with the same tools. This would allow independent developers to link applications that can be easily integrated with the NEONS database. Fig. 3 shows a sample page of the X-Window browser.

6.2 External Data Format Converter

Data sets can be extracted through a converter to convert the database format to external data format. These converted data sets can then be read by a loosely coupled applications program or written to a tape or disk for shipping. Conversion functions for the BUFR, GRIB, netCDF, HDF standard formats are being developed.

7. PERFORMANCE

Data access speed is effectively independent of the total volume data in the database, or the volume of any particular data type such as bathythermographs. This results from two factors. First, the DBMS utilizes the UNIX file system, storing each relational table in one or more than one UNIX files. Thus, adding more tables means creating more files, instead of adding more data to existing ones. Second, the schema is designed so that all the reports for a particular data type, such as bathythermographs, are not stored in the same table. The database administrator constructs as many tables as desired for bathythermograph reports, and controls the number of records put in any table. An indexing scheme ensures that only those tables with qualifying reports are scanned during a query. This data segmentation is transparent to the users.

The following timing tests were performed on a Hewlett-Packard 9000 835 workstation, rated at 14 MIPS and 1.7 MFLOPS. These times are the total CPU usage and include all the NEONS software, commercial DBMS software, and operating-system services. In each case the data was retrieved from the database, unpacked, and placed into program variables.

- Grid-Point data: read 50 global NOGAPS model grid fields, each containing 144x288 = 41,472 points in 13 seconds.

- LLT data: read 20,000 WMO surface land reports, each containing 27 environmental parameters in 15 seconds.

- Image data: read a satellite image containing 1024x1024 pixels by 4 bands (6, 8, 10, & 12 bits) in 10 seconds.

The NEONS database has been implemented on various UNIX mini-computer workstations and a UNIX PC/386 computer. Work has also been started on porting the software to 64-bit word supercomputer. Timing tests done on these various machines have shown that database performance is closely related to the processor speed. A closer examination of time spent doing database activities shows the majority of time is spent in the Central Processing Unit (CPU), and a much smaller percentage in disk I/O. Of the CPU time, about one-half is devoted to unpacking the primary data from the packed binary GRIB and BUFR internal storage formats. These CPU-to-I/O ratios depend, of course, on the machine processor speed, degree of vectorization or parallelism, and I/O bandwidth. Our studies have shown that, for currently popular workstations operating at about 20 MIPS, the overall database activity is strongly CPU bound. This implies that database performance will improve significantly as processors speed up. The commercial DBMS performance is also improving steadily. A recent upgrade of the commercial DBMS used at NOARL resulted in a 35% overall performance improvement. New DBMS features such as shared libraries and shared memory are also yielding significant performance improvements in multi-user environments.

8. SUMMARY

The NEONS developed at the Naval Oceanographic and Atmospheric Research Laboratory (NOARL), Atmospheric Directorate, Monterey, CA is a set of generic database management interface software which interacts between user's application programs and a commercial relational database management system. The NEONS data handling schema provides a basis for handling multitudes of existing and future meteorological and oceanographic data types. The generic approach insulates the applications programs from the database itself. Applications program data access routines can remain unchanged while the database system go through the computer hardware and software upgrades.

The NEONS implementation has shown that bulky and complex environmental data can be readily handled by an appropriate commercial relational database management system. The timing performance tests showed that much of the total I/O time is spent in the CPU. As the speed of microprocessors increases, and the performance of commercial DBMS packages improves, and the environmental data volume and complexity increases, the benefits of NEONS' data handling approach will be even more pronounced.

9. ACKNOWLEDGEMENTS

The Naval Environmental Operational Nowcasting System (NEONS) project is sponsored by the Oceanographer of the Navy (OP-096) through the Space and Naval Warfare Systems Command program office (PMW-165), Washington D.C., Program Element 35111N. NOARL contribution No. 92:015:440. Approved for public release; distribution is unlimited.

NEONS is developed by a highly dedicated team. The database handling subsystem was built by Andy Jurkevics, Jim Clark, Rick Titus, Maureen Thompson, and Scott Christensen. The database X-Browser was developed by Christina Shaw and Chuck Stein. Graphics utilities were integrated together by Don Schertz and Don Ramsey. System administration has been ably provided by Allan Caughey and Neil Withers. We are indebted to Dr. Eve Schwartz, Mr. Bruce Gritton and Mr. Duncan Ross for their original idea of developing the system and their efforts in conducting the system requirements survey.

10. REFERENCES

Bergen, W., B. Gardiner, and M. Mathewson, 1989: "Past development, present trends and future evolution of meteorological work-stations", Preprints, AMS 5th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim, CA, 297-299.

Gardiner, B., W. Bergen, and M. Mathewson, 1989: "New design concepts for the development of meteorological workstations", Preprints, AMS 5th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim, CA, 6-9.

Hibbard, W.L. and D.A. Santek, 1989: "Visualizing large data sets", Preprints, AMS 5th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim, CA, 172-174.

Sadowski, A., 1989: "The key to the future of meteorological systems", Preprints, AMS 5th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim, CA, 69-76.

Soderman, D. and J.K. Gibson, 1989: "The binary revolution and its implications for international and national systems", Preprints, AMS 5th Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, Anaheim, CA, 58-62.

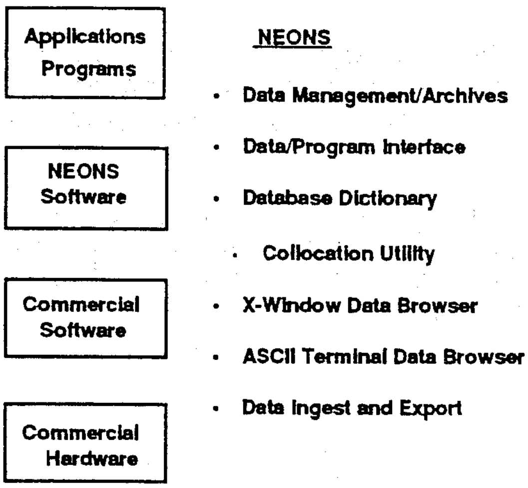

Fig. 1: NEONS Layered Architecture.

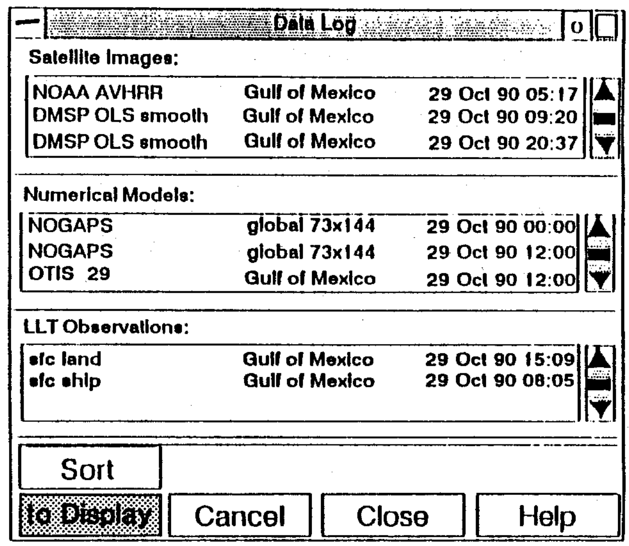

Fig. 3: Sample page of the NEONS X-Window Browser.

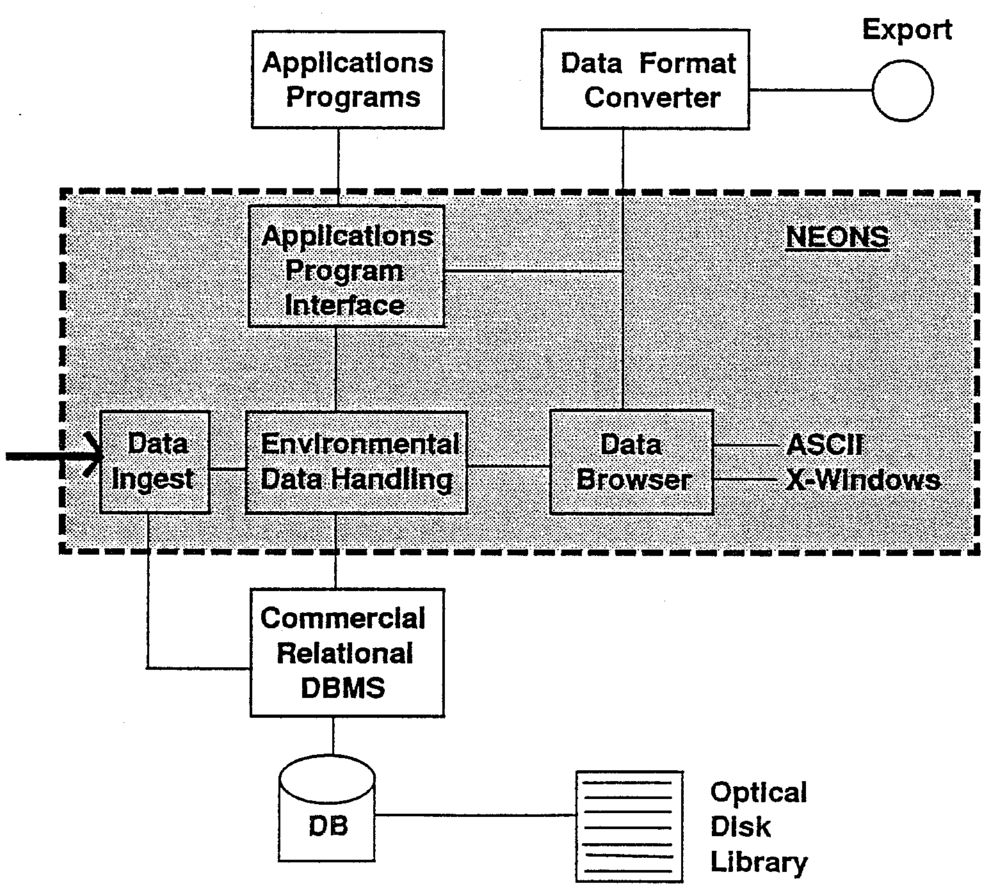

Fig. 2: NEONS functional components.