SEQUOIA 2000

LARGE CAPACITY OBJECT SERVERS

TO SUPPORT GLOBAL CHANGE RESEARCH

September 17, 1991

Principal Investigators:

Michael Stonebraker
University of California
Electronics Research Laboratory
549 Evans Hall
Berkeley, CA 94720
(415) 642-5799
mike@postgres.berkeley.edu

Jeff Dozier
University of California
Center for Remote Sensing and Environmental Optics
1140 Girvetz Hall
Santa Barbara, CA 93106
(805) 893-2309
dozier@crseo.ucsb.edu

STEERING COMMITTEE

John Estes, Director, National Center for Geographic Information and Analysis, Santa Barbara
Edward Frieman, Director, Scripps Institution of Oceanography, San Diego
Clarence Hall, Dean of Physical Sciences, Los Angeles
David Hodges, Dean of Engineering, Berkeley
Sid Karin, Director, San Diego Supercomputer Center
Calvin Moore, Associate Vice President, Academic Affairs, UC Office of the President (Chair)
Richard West, Associate Vice President, Information Systems and Administrative Services,
UC Office of the President

FACULTY INVESTIGATORS

Michael Bailey, San Diego Supercomputer Center, San Diego
Tim Barnett, Scripps Institution of Oceanography, San Diego
Hans-Werner Braun, San Diego Supercomputer Center, San Diego
Michael Buckland, School of Library and Information Studies, Berkeley
Ralph Cicerone, Department of Geosciences, Irvine
Frank Davis, Center for Remote Sensing and Environmental Optics, Santa Barbara
Domenico Ferrari, Computer Science Division, Berkeley
Catherine Gautier, Center for Remote Sensing and Environmental Optics, Santa Barbara

Michael Ghil, Department of Atmospheric Sciences, Los Angeles
Randy Katz, Computer Science Division, Berkeley
Ray Larson, School of Library and Information Studies, Berkeley
C. Roberto Mechoso, Climate Dynamics Center, Los Angeles
David Neelin, Department of Atmospheric Sciences, Los Angeles
John Ousterhout, Computer Science Division, Berkeley
Joseph Pasquale, Computer Science Department, San Diego
David Patterson, Computer Science Division, Berkeley
George Polyzos, Computer Science Department, San Diego
John Roads, Scripps Institution of Oceanography, San Diego
Lawrence Rowe, Computer Science Division, Berkeley
Ray Smith, Center for Remote Sensing and Environmental Optics, Santa Barbara
Richard Somerville, Scripps Institution of Oceanography, San Diego
Richard Turco, Institute of Geophysics and Planetary Physics, Los Angeles

Abstract

Improved data management is crucial to the success of current scientific investigations of Global Change. New modes of research, especially the synergistic interactions between observations and model-based simulations, will require massive amounts of diverse data to be stored, organized, accessed, distributed, visualized, and analyzed. Achieving the goals of the U.S. Global Change Research Program will largely depend on more advanced data management systems that will allow scientists to manipulate large-scale data sets and climate system models.

Refinements in computing — specifically involving storage, networking, distributed file systems, extensible distributed data base management, and visualization — can be applied to a range of Global Change applications through a series of specific investigation scenarios. Computer scientists and environmental researchers at several UC campuses will collaborate to address these challenges. This project complements both NASA's EOS project and UCAR's (University Corporation for Atmospheric Research) Climate Systems Modeling Program in addressing the gigantic data requirements of Earth System Science research before the turn of the century. Therefore, we have named it **Sequoia 2000,** after the giant trees of the Sierra Nevada, the largest organisms on the Earth's land surface.

1. MOTIVATION FOR THE RESEARCH

Among the most important challenges that will confront the scientific and computing communities during the 1990s is the development of models to predict the impact of Global Change on the planet Earth. Among the **Grand Challenges** for computing in the next decade, study of Global Change will require great improvements in monitoring, modeling and predicting the coupled interactions within the components of the Earth's subsystems [CPM91].

The Earth Sciences of meteorology, oceanography, bioclimatology, geochemistry, and hydrology grew up independently of each other. Observational methods, theories, and numerical models were developed separately for each discipline. Beginning about 20 years ago, two forces have favored the growth of a new discipline, **Earth System Science.** One force is the unified perspective that has resulted from space-based sensing of planet Earth during the last two decades. The second is a growing awareness of and apprehension about Global Change caused by human activities. Among these are:

- the "greenhouse effect" associated with increasing concentrations of carbon dioxide, methane, chlorofluorocarbons, and other gases;
- ozone depletion in the stratosphere, most notably near the South Pole, resulting in a significant increase in the ultraviolet radiation reaching the Earth's surface;
- a diminishing supply of water of suitable quality for human uses;
- deforestation and other anthropogenic changes to the Earth's surface, which can affect the carbon budget, patterns of evaporation and precipitation, and other components of the Earth System;
- a pervasive toxification of the biosphere caused by long-term changes in precipitation chemistry and atmospheric chemistry;
- biospheric feedbacks caused by the above stresses and involving changes in photosynthesis, respiration, transpiration, and trace gas exchange, both on the land and in the ocean.

For theorists, Earth System Science means the development of interdisciplinary models that couple elements from such formerly disparate sciences as ecology and meteorology. For those who make observations to acquire data to drive models, Earth System Science means the development of an integrated approach to observing the Earth, particularly from space, bringing diverse instruments to bear on the interdisciplinary problems listed above.

One responsibility of Earth System Scientists is to inform the development of public policy, particularly with respect to costly remedies to control the impact of human enterprise on the global environment. Clearly, human activities accelerate natural rates of change. However, it is difficult to predict the long-term effects of even well-documented changes, because our understanding of variations caused by nature is so poor. Therefore, it is imperative that our predictive capabilities be improved.

Throughout the UC System are many leading scientists who study various components of this Earth System Science. One of these research groups is the Center for Remote Sensing and Environmental Optics (CRSEO) on the Santa Barbara campus. At UCLA, efforts in Earth System Science span four departments (Atmospheric Sciences, Biology, Earth and Space Sciences, and Geography) and the Institute for Geophysics and Planetary Physics. At the core of these efforts is the Climate Dynamics Center. At UC San Diego, the Climate Research Division (CRD) of Scripps Institution of Oceanography focuses on climate variability on time scales from weeks to decades, climate process studies, and modeling for forecasts of regional and transient aspects of Global Change. At UC Irvine a new Geosciences Department has been formed with a focus on Global Change.

Study of Earth System Science requires a data and information system that will provide the infrastructure to enable scientific interaction between researchers of different disciplines, and researchers employing different methodologies; it must be an **information** system, not just a **data** system. For example, it must provide geophysical and biological information, not just raw data from spaceborne instruments or *in situ* sensors. It must also allow researchers to collate and cross-correlate data sets and to access data about the Earth by processing the data from the satellite and aircraft observatories and other selected sources. An additional application is in models of the dynamics, physics and chemistry of climatic subsystems which are accomplished through coupled General Circulation Models (GCMs) which generate huge data sets of output that represent grids of variables denoting atmospheric, oceanic and land surface conditions. The models need to be analyzed and validated by comparison with values generated by other models, as well as with those values actually measured by sensors in the field [PANE91].

1.1. Shortcomings of Current Data Systems

UC Global Change researchers have learned that serious problems in the data systems available to them impede their ability to access needed data and thereby do research [CEES91]. In particular, five major shortcomings in current data systems have been identified:

1) Current storage management system technology is inadequate to store and access the massive amounts of data required.

For instance, when studying changes over time of particular parameters (e.g. snow properties, chlorophyll concentration, sea surface temperature, surface radiation budget) and their roles in physical, chemical or biological balances, enormous data sets are required. A typical observational data set includes:

- topographic data of the region of interest at the finest resolution available;
- the complete time series of high resolution satellite data for the regions of interest;
- higher resolution data from other satellite instruments (e.g. the Landsat Multispectral Scanners (MSS) and Thematic Mappers (TM));
- aircraft data from instruments replicating present or future satellite sensors (e.g Advanced Visible and Infrared Imaging Spectrometer (AVIRIS) and Synthetic Aperture Radar (AIRSAR));
- various collections of surface and atmosphere data (e.g. atmospheric ozone and temperature profiles, streamflow, snow-water equivalence, profiles of snow density and chemistry, shipboard and mooring observations of SST and chlorophyll concentration at sea).

Currently, researchers need access to datasets on the order of one Terabyte; these datasets are growing rapidly.

While it is possible, in theory, to store a Terabyte of data on magnetic disk (at high cost), this approach will not scale as the number of experiments increases and the amount of data per experiment increases also. A much more cost-effective solution would incorporate a multi-level hierarchy that uses not only magnetic disk but also one or more tertiary memory media. Current system software, including file systems and data base systems, offers no support for this type of multi-level storage hierarchy. Moreover, current tertiary memory devices (such as tape and optical disk) are exceedingly slow, and hardware and software must mask these long access delays through sophisticated caching, and increase effective transfer bandwidth by compression techniques and parallel device utilization. None of the necessary support is incorporated in currently available commercial systems.

2) Current I/O and networking technologies do not support the data transfer rates required for browsing and visualization.

Examination of satellite data or output from models of the Earth's processes requires that we **visualize** data sets or model outputs in various ways. A particularly challenging technique is to **fast-forward** satellite data in either the temporal or spatial dimension. The desired effect is similar to that achieved by the TV weather forecasters who show, in a 20-second animated summary, movement of a storm based on a composite sequence of images collected from a weather satellite over a 24-hour period. Time-lapse movies of concentrations of atmospheric ozone over the Antarctic ‘‘ozone hole’’ show interesting spatial-temporal patterns. Comparison of simulations requires browsing and visualization. Time-lapse movies and rapid display of two-dimensional sections through three-dimensional data place severe demands on the whole I/O system to generate data at usable rates.

To do such visualization in real time places severe demands on the I/O system to generate the required data at a usable rate. Additionally, severe networking problems arise when investigators are geographically remote from the I/O server. Not only is a high bandwidth link required that can deliver 20-30 images per second (i.e. up to 600 Mbits/sec), but also the network must guarantee delivery of required data without pauses that would degrade real-time viewing. Current commercial networking technology cannot support such ‘‘guaranteed delivery’’ contracts.

3) Current data base systems are inadequate to store the diverse types of data required.

Earth System Scientists require access to the following disparate kinds of data for their remote sensing applications:

- Point Data for specific geographic points. *In situ* snow measurements include depth and vertical profiles of density, grain size, temperature, and composition, measured at specific sites and times by researchers traveling on skis. Another example is the chlorophyll concentration obtained from ships or moored sensors at sea.
- Vector Data. Topographic maps are often organized as polygons of constant elevation (i.e. a single datum applying to a region enclosed by a polygon, which is typically represented as a vector of points). This data is often provided by other sources (e.g. USGS or Defense Mapping Agency); it is not generated by Global Change researchers, but has to be linked up to sensor readings. Other vector data include drainage basin boundaries, stream channels, etc.
- Raster Data. Many satellite and aircraft remote sensing instruments produce a regular array of point measurements. The array may be 3-dimensional if multiple measurements are made at each location. This ‘‘image cube’’ (2 spatial plus 1 spectral dimension) is repeated every time the satellite completes an orbit. Such regular array data are called raster data because of their similarity to bitmap image data. The volumes are large. For example, a single frame from the AVIRIS NASA aircraft instrument contains 140 Mbytes.
- Text Data. Global Change researchers have large quantities of textual data including computer programs, descriptions of data sets, descriptions of results of simulations, technical reports, etc. that need to be organized for easy retrieval.

Current commercial relational data base systems (e.g. DB 2, RDB, ORACLE, INGRES, etc.) are not good at managing these kinds of data. During the last several years a variety of next generation DBMSs have been built, including IRIS [WILK90], ORION [KIM90], POSTGRES [STON90], and Starburst [HAAS90]. The more general of these systems appear to be usable, at least to some extent, for point, vector, and text data. However, none are adequate for the full range of needed capabilities.

4) Current visualization software is too primitive to allow Global Change researchers to render the data returned for useful interactive Improved visualization is needed for two purposes in Sequoia 2000:

- visualization of data sets—remote sensing data, in situ data, maps, and model output must be interpreted and compared;
- visualization of the database—input and output to the database management system (queries and answers) would benefit from visualization.

Data sets and model output examined by Global Change researchers include all those described in (3) above. Just as commercial relational database systems are not good at managing those kinds of data, commercial visualization tools and subroutine packages are not good at integrating these diverse kinds of data sets.

Moreover, database management systems depend mostly on textual input and output. In managing geographic information, remote sensing data, and 3D model output, an essential extension to such systems is the ability to query and to examine the database using graphs, maps, and images.

1.2. Objectives of Sequoia 2000

In summary, Global Change researchers require a massive amount of information to be effectively organized in an electronic repository. They also require ad-hoc collections of information to be quickly accessed and transported to their workstations for visualization. The hardware, file system, DBMS, networking, and visualization solutions currently available are totally inadequate to support the needs of this community.

The problems faced by Global Change researchers are faced by other users as well. Most of the **Grand Challenge** problems share these characteristics, i.e. they require large amounts of data, accessed in diverse ways from a remote site quickly, with an electronic repository to enhance collaboration. Moreover, these issues are also broadly applicable to the computing community at large. Consider, for example, an automobile insurance application. Such a company wishes to store police reports, diagrams of each accident site and pictures of damaged autos. Such image data types will cause existing data bases to expand by factors of 1000 or more, and insurance data bases are likely to be measured in Terabytes in the near future. Furthermore, the same networking and access problems will appear, although the queries may be somewhat simpler. Lastly, visualization of accident sites is likely to be similar in complexity to visualization of satellite images.

The purpose of this proposal is to build a four-way partnership to work on these issues. The first element of the partnership is a technical team, primarily computer and information scientists, from several campuses of the University of California. They will attack a specific set of research issues surrounding the above problems as well as build prototype systems to be described.

The second element of the partnership is a collection of Global Change researchers, primarily from the Santa Barbara, Los Angeles, San Diego, and Irvine campuses, whose investigations have substantial data storage and access requirements. These researchers will serve as users of the prototype systems and will provide feedback and guidance to the technical team.

The third element of the partnership is a collection of public agencies who must implement policies affected by Global Change. We have chosen to include the California Department of Water Resources (DWR), the California Air Resources Board (ARB) and the United States Geological Survey (USGS). These agencies are end users of the Global Change data and research being investigated. They are also interested in the technology for use in their own research. The role of each of these agencies will be described in Section 4 along with that of certain private sector organizations.

The fourth element of the partnership is industrial participants, who will provide support and key research participants for the project. Digital Equipment Corporation is a principal partner in this project and has pledged both equipment and monetary support for the project. In addition, TRW and Exabyte have agreed to participate and are actively soliciting additional industrial partners.

We call this proposal **Sequoia 2000,** after the long-lived trees of the Sierra Nevada. Successful research on Global Change will allow humans to better adapt to a changing Earth, and the 2000 designator shows that the project is working on the critical issues facing the planet Earth as we enter the next century.

The Sequoia 2000 research proposal is divided into 7 additional sections. In Section 2 we present the specific Computer Science problems we plan to focus on. Then, in Section 3, we detail goals and milestones for this project that include two prototype object servers, BIGFOOT I and II, and associated user level software. Section 4 continues with the involvement of other partners in this project. In Section 5, we briefly indicate some of our thoughts for a following second phase of this project. Section 6 discusses critical success factors. Section 7 outlines the qualifications of the Sequoia research team. We close in Section 8 with a summary of the proposal.

In addition, there is one appendix which shows investigative scenarios that will be explored by the Global Change research members of the Sequoia team. These are specific contexts in which new technology developed as part of the project will be applied to Global Change research.

2. THE SEQUOIA RESEARCH PROJECT

As noted above, our technical focus is driven by the needs of **Grand Challenge** researchers to visualize selected portions of large object bases containing diverse data from remote sites over long-haul networks. Therefore, we propose a coordinated attack on the remote visualization of large objects using hardware, operating system, networking, data base management, and visualization ideas. In the next subsection we briefly sketch out new approaches in each of these areas.

A large object base contains diverse data sets, programs, documents, and simulation output. To share such objects, a sophisticated electronic repository is required, and in the second subsection we discuss indexing, user interface, and interoperability ideas that we wish to pursue in conjunction with such a repository.

2.1. Remote Visualization of Large Object Bases

2.1.1. Hardware Concepts

The needed system must be able to store many Terabytes of data in a manageable amount of physical space. Even at $1/Mbyte, a Terabyte storage system based on magnetic disk will cost $1,000,000. Since magnetic tape costs about $5/Gbyte, the same Terabyte would cost only $5000! Thus, it is easy to see that a practical massive storage system must be implemented from a combination of storage devices including magnetic disk, optical disk, and magnetic tape. A critical aspect of the storage management system subsystem we propose to construct will be its support for managing a complex hierarchy of diverse storage media [RANA90].

Our research group has pioneered the development of RAID technology, a new way to construct high bandwidth, high availability disk systems based on arrays of small form factor disks [KATZ89]. The bandwidth comes from striping data across many disk actuators and harnessing this inherent parallelism to dramatically improve transfer rates. We are currently constructing a storage system with the ability to sustain 50 Mbyte/second transfers. This controller is being attached to a 1 Gbit/second local area network via a HIPPI channel connect.

We propose to extend these techniques to arrays of small form-factor tape drives. 8mm and 4mm tape systems provide capacity costs that are 10 times less than optical disk [POLL88, TAN89]. A tape jukebox in the 19" form-factor can hold .5 Terabyte in the technology available today, with a doubling expected within the next 1 to 2 years [EXAB90]. These tapes only transfer at the rate of .2 Mbyte/second, but once they are coupled with striping techniques, it should be possible to stage and destage between disk and tape at the rate of 4 Mbytes/second. This is comparable to high speed tape systems with much lower capacity per cartridge.

Besides striping, a second method for improving the transfer rate (and incidentally the capacity) of the storage system is compression [LELE87, MARK91]. An important aspect of the proposed research will be an investigation of where hardware support for compression and decompression should be embedded into the I/O data-path. Coupled with the data transfer rate of striped tape systems, it may be possible to sustain transfers of compressed data from the tape archive approaching 1 Gbyte/sec.

2.1.2. Operating System Ideas

The two of the most difficult problems in managing the storage hierarchy are long access times and low transfer rates of tertiary memory. Several sets of techniques are proposed to address these problems.

The first set of techniques has to do with management of the tape storage to reduce the frequency of tape-load operations. Researchers at Berkeley have recently investigated both read-optimized [SELT90] and write-optimized [ROSE91] file systems for disk storage. Read-optimized storage attempts to place logically sequential blocks in a file physically sequentially on the disk, for fastest retrieval. On the other hand, write-optimized file systems place blocks where it is currently optimal to write them, i.e., under a current disk arm. Write optimization is appropriate when data is unlikely to be read back, or when read patterns match write patterns. We propose to explore how both kinds of file systems could be extended to tertiary memory.

The second set of techniques concerns itself with multi-level storage management: how can a disk array be combined with a tape library to produce a storage system that appears to have the size of the tape library and the performance of the disk array? We will explore techniques for caching and migration, where information is moved between storage level to keep the most-frequently accessed information on disk. Researchers at Berkeley have extensive experience with file caching and migration [SMIT81, KURE87, NELS88]. Although we hope to apply much of this experience to the proposed system, the scale of the system, the performance characteristics of the storage devices, and the access patterns will be different enough to require new techniques.

This investigation will occur in two different contexts. First, Berkeley investigators will explore the above ideas in the context of the BIGFOOT prototypes described below. Second, San Diego Supercomputer Center (SDSC) researchers will explore migration in the context of a production supercomputer. They expect most files to be read sequentially in their entirety, so their approach will be based on migrating whole files rather than physical blocks; Berkeley researchers will likely explore both whole-file and block-based approaches. Also, SDSC researchers will have to contend with a five-level hierarchy and a large collection of on-line users. We propose a collaborative effort between the two groups that will result in enhanced algorithms appropriate to both environments.

2.1.3. Networking Hardware and Software

A common work scenario for Global Change scientists will be visualization at their workstation of time-sequenced images accessed from a large object base over a fast high-bandwidth wide-area network. The data may be produced in real time, or may not (e.g., because of the computational effort required). The visualization will be interactive with users from remote workstations asking for playback, fast-forward, rotation, etc.; this should be possible without necessarily bringing in the entire data set at the outset. This interactivity and the temporal nature of the data's presentation requires a predictable and guaranteed level of performance from the network. Although image sequences require high bandwidth and low delay guarantees, these guarantees are often statistical in nature. Protocols must be developed which support deterministic and statistical real-time guarantees, based on quality of service parameters specified by the user/programmer.

Bandwidth (as well as storage space) requirements can be reduced by image compression. Clearly, compression will be applied to images before transmission from the object server to a remote user. In the object server, this can be done as part of the I/O system when image representations move to or from storage. However, on the user workstations, decompression must be done along the path from the network interface to the frame buffer.

Mechanisms which support the guaranteed services offered by the network must be integrated with the operating system, particularly the I/O system software, which controls the movement of data between arbitrary I/O devices, such as the network interface, frame buffer, and other real-time devices. The network software and I/O system software must work in a coordinated fashion so that bottlenecks, such as those due to memory copying or crossing of protection boundaries, are avoided. The I/O system software, is one of the least understood aspects of operating system design [PASQ91], especially regarding soft real-time I/O. We intend to explore the relationship between I/O system software and network protocol software, and how various degrees of design integration affect performance. One specific idea we have is the construction of fast in-kernel datapaths between the network and I/O source/sink devices for carrying messages which are to be delivered at a known rate. Since processing modules (e.g., compression/decompression, network protocols) may be composed along these datapaths, a number of problems must be solved, such as how to systematically avoid copying processed messages between modules, or between kernel and user address spaces.

The network, or even the workstation's operating system, can take advantage of the statistical nature of guarantees by conveniently dropping packets when necessary to control congestion and smooth network traffic. This is particularly relevant when one is fast-forwarding through a sequence of images; supporting full resolution might not be possible, and users might be willing to accept a lower resolution picture in return for faster movement. One approach to this problem is hierarchical coding [KARL89], whereby a unit of information such as an image is decomposed into a set of ordered sub-images. A selected subset of these may be re-composed to obtain various levels of resolution of the original image. This gives the receiver the flexibility of making the best use of received sub-images that must be output by some deadline, and gives the network the flexibility of dropping packets containing the least important sub-images when packets must be dropped. One research issue is how to route hierarchically coded packets in a way that provides the network with the maximum flexibility in congestion control, and how to compose them in time at the receiver so that integrity and continuity of presentation are preserved. In particular, the layers at which multiplexing and demultiplexing will be performed should be carefully designed to take full advantage of hierarchical coding.

Remote visualization places severe stress on a wide-area network; this raises open problems in networking technology. One fundamental issue is the choice of the mode of communication. The Asynchronous Transfer Mode (ATM) is emerging as the preferred standard for the Broadband Integrated Services Digital Network (B-ISDN). However, the small ATM cells (53 bytes) into which messages are subdivided may not provide efficient transport when network traffic is dominated by large image transmissions and video streams. On the other hand, FDDI takes the opposite stance, allowing frames up to 4500 bytes long. We will evaluate how packet size and mode of communication affect the applications for the proposed environment. We also propose to investigate efficiency problems that might arise in gateways between FDDI and ATM-based wide-area networks.

A final issue is that the protocols to be executed on the host workstations, the gateways, and the switches (or the switch controllers) will have to include provisions for real-time channel establishment/disestablishment [FERR90a], so that guarantees about the network's performance (throughput, delay, and delay jitter) can be offered to the users who need them [FERR90b]. A related issue is the specification of quality of network service needed by the user. Such a specification must be powerful enough to describe the required guarantees, and yet must be realizable by mechanisms that already exist, or that can be built into the networks of interest.

2.1.4. Data Management Issues

In some environments it is desirable to use a DBMS rather than the file system to manage collections of large objects. Hence, we propose to extend the next-generation DBMS POSTGRES [STON90] to effectively manage Global Change data. There are three avenues of extension that we propose to explore.

First, POSTGRES has been designed to effectively manage point, vector and text data. However, satellite data are series of large multidimensional arrays. Efficient support for such objects must be designed into the system. Not only must the current query language be extended to support the times series array data that is present but also, the queries run by visualizers in **fast-forward** mode must be efficiently evaluated. This will entail substantial research on storage allocation of large arrays and perhaps on controlled use of redundancy. We propose to investigate decomposing a large multidimensional array into **chunklets** that would be stored together. Then, a fast-forward query would require a collection of chunklets to be accessed and then intersected with the viewing region. The optimal size and shape of these chunklets must be studied, as well as the number of redundant decompositions that should be maintained.

Second, POSTGRES has been designed to support data on a combination of secondary and tertiary memory. However, a policy to put historical data onto the archive and current data in secondary storage has been hard-coded into the current implementation. The rationale was that current data would be accessed much more frequently than historical data. While this may be true in many business environments, it will not be the case in Global Change research. Therefore, a more flexible way of dealing with the storage hierarchy must be defined that will allow "worthy" data to migrate to faster storage. Such migration might simply depend on the algorithms of the underlying file system discussed above to manage storage. However, the DBMS understands the logical structure of the data and can make more intelligent partitioning decisions as noted in [STON91].

If both the file system and the DBMS are managing storage, then it is important to investigate the proper interface between DBMS managed storage and operating system managed storage. This issue arises in disk-based environments, and is more severe in an environment which includes tertiary memory. In addition, the query optimizer must be extended to understand the allocation of data between secondary and tertiary memory as well as the allocation of objects to individual media on the archive. Only in this way can media changes be minimized during query processing. Also, processing of large objects must be deferred as long as possible in a query plan, as suggested in [STON91].

The third area where we propose investigations concerns indexing. The conventional DBMS paradigm is to provide **value** indexing. Hence, one can designate one or more fields in a record as **indexed,** and POSTGRES will build the appropriate kind of index on the data in the required fields. Value indexing may be reasonable in traditional applications, but will not work for the type of data needed to support Global Change research. First, researchers need to retrieve images by their content, e.g. to to find all images that contain Lake Tahoe. To perform this search requires indexes on the result of a classification function and not on the raw image. Second, indexing functions for images and text often return a collection of values for which efficient access is desired [LYNC88]. For example, a keyword extraction function might return a set of relevant keywords for a document, and the user desires indexing on all keywords. In this case one desires **instance** indexing on the set of values returned by a function. We propose to look for a more general paradigm that will be able to satisfy all indexing needs of Global Change researchers.

2.1.5. Visualization Workbench

Scientific visualization in the 1990's cannot be restricted to the resources that are available on a single scientist's workstation, or within a single processing system. The visualization environment of the future is one of heterogeneous machines on networks. A single scientific application must have access to the plethora of resources that are available throughout the net from compute servers, hardcopy servers, data storage servers, rendering servers, and realtime data digestors. Visualization must be incorporated into the database management system, so that the **database** can be visualized, in addition to the **data sets** in the database. Input through a "visual query language" will be needed.

Several commercial or public-domain software packages for visualization have useful features for Sequoia 2000, but do not contain the full menu of tools needed for our purposes. In the commercial domain, PV-Wave, Wavefront, Spyglass, and IDL are extensive packages that can be accessed with programming-like commands (much like a fourth-generation language). NCAR graphics and UNIDATA programs (i.e. netCDF, units, and mapping utilities) provide a software library with subroutines that can be incorporated in users' programs. NASA/Goddard's meteorological data display program can be of use to display realtime and archived grided and text datasets. UNIDATA's Scientific Data Management(SDM) system can be utilized for the ingestion and display of realtime weather datasets. In the public domain, a package developed by NCSA has been widely distributed, and the SPAM (Spectral Analysis Manager) package from JPL/Caltech has many routines for visualization and analysis of data from imaging spectrometers, where the spectral dimension of the data is as large as the spatial dimension, with images of more than 200 spectral bands. At UCSB, the Image Processing Workbench (IPW) is a portable package for analysis of remote sensing data. At the University of Colorado, IMagic is extensively used for analysis of data from the NOAA meteorological satellites.

SDSC has developed and implemented a production hardcopy server that allows all members of a network access to a suite of hardcopy media for images—slides, movies, color paper [BAIL91]. This capability is to be expanded to allow automated production of videotapes, from time-sequenced images.

Another major network-based visualization project underway at SDSC is the building of a prototype server for volume visualization. Front-end processes elsewhere on the network can connect to it and submit data files for rendering. Once the rendering is complete, the server will then do one of three things: return the image for display on the originating workstation, store the image for later retrieval, or automatically pass it off to the network hardcopy server. This prototype utility needs to be formalized with more robust software development and a good workstation front-end program. With high-speed networks, it can be incorporated into the data management software so that users could visualize data on remote servers.

The successful completion of the remote volume visualization project will leave Sequoia 2000 with a working skeleton of a general-purpose remote visualization package. Other similar packages can then be produced for Sequoia researchers, including various types of remote rendering systems.

2.2. The Electronic Repository

The electronic repository required by Global Change researchers includes various data sets, simulation output, programs, and documents. For repository objects to be effectively shared, they must be **indexed,** so that others can retrieve them by content. Moreover, effective user interfaces must be built so that a researcher can **browse** the repository for desired information. Lastly, programs in the repository must be able to **interoperate,** i.e. they must operate correctly in other environments than the one in which they were generated. Therefore, in this section we discuss proposed research on indexing paradigms, user interfaces and interoperability of programs.

2.2.1. Indexing Techniques

A large object store is ineffective unless it can be indexed successfully. We must address the issue of how to index the raster data of Global Change researchers. For example, they wish to find all instances of "El Nino" ocean patterns from historical satellite data. This requires indexing a region of spatial data and a region of time according to an imprecise (fuzzy) classification. One approach will be to use existing thesauri of geographical regions and place names that include the cartographic coordinates of the places. Researchers may also need to create their own classifications that can be used to select and partition the data.

We must also index computer programs in the repository. We will take two approaches in this area. First, we will index the documentation that is associated with a program using traditional techniques. In addition, because we have the source code available, we can also index program variables and names of called functions. This will allow retrieval, for example, of all repository programs that include the variable ‘‘drought_level’’ or the ones that call the subsystem SPSS.

Finally, mature techniques exist for indexing and retrieving textual documents based on automatic and manual indexing using keywords, thesaurus terms, and classification schemes. Statistically-based probabilistic match techniques have been developed that present the "best-matching" documents to the user in response to an imprecise query [BELK87, LARS91]. These techniques must be extended to deal with indexing large collections of complete textual documents, rather than just collections of document surrogates (i.e. titles and abstracts).

It is well-known that standard keyword-based techniques for document retrieval are intrinsically limited [COOP88]. In general, for the best keyword-based techniques to retrieve 90% of the desired documents, only 10% of what is retrieved is what the user desired. To improve on the performance of such systems, some analysis of the content of the texts and of the users' queries is required [CROF89, SMIT89]. Such analysis is difficult on arbitrary text because it requires recourse to knowledge of the domain to which the text pertains, and entails solving hard general problems of natural language processing. However, in narrow enough domains, natural language processing techniques can provide enough analysis to significantly improve performance.

We plan to investigate applying such techniques to the retrieval of textual document in Sequoia 2000. In particular, two techniques to help improve performance are goal analysis and lexical disambiguation. Goal analysis is the process of extracting from the user's query some representation of what the user wants. Even simple goal analysis has been shown to produce modest but significant improvements in retrieval accuracy. Lexical disambiguation is the process of deciding which of the senses of a word is in play in a given use. Having this information can lead to dramatic improvements in performance. In the limited domain of the documents of concern to Sequoia 2000, there is expectation that the techniques we have been developing will be tractable.

2.2.2. User Interfaces

We expect to have a large collection of data sets, documents, images, simulation runs, and programs in the repository, and tools are needed to allow users to **browse** this repository to find objects relevant to their work. We propose to explore tools based on two different paradigms. The first paradigm is to view the repository as a conventional library. Information retrieval techniques can then be applied to run queries on the repository in much the same way that electronic card catalogs support queries to their contents. Unfortunately, this requires that a human **librarian** be available to **catalog** any incoming object to identify relevant **descriptors** and

classifications that serve to ease subsequent searches. We propose to explore full text as well as descriptor and classification searching techniques as our first user interface to the repository.

Our second approach is to use a graphical spatial paradigm. In this case, we would require users to furnish an **icon** which would represent their object. The repository can then be viewed as a collection of icons, for which we would attempt to build an organizing tool that would place them spatially in 2 or 3 dimensions. This tool would support movement through the space of icons, by simply panning geographically. A user who located an icon of interest could then **zoom** on the icon and receive increasing amounts of information about the object. For functions, the tool could capture relevant information from program documentation. For example, the first level might be the documentation banner at the top of the function and the second level might be the call graph of the function. At the finest level of granularity, the entire source code would be presented.

These techniques can be extended to apply to maps in two ways. First, a collection of maps of different resolutions could be organized hierarchically. Zooming into a map would then cause it to be replaced by a higher resolution map of the target area. This paradigm could be further extended to include a **time** dimension, through which a user would be able to "pan" forward or backward in time, thereby obtaining **time travel** with the same interface. Second, icons could be associated with geographic co-ordinates and represent information associated with a point or region. For example, a data set of manually measured snow depths could have an appropriate icon assigned, say a depth gauge. A join query could be run to construct a composite map with the icons overlaid on any particular map. Zooming into an icon would then cause the icon to be replaced by a window with the more detailed data. As our second approach to the repository, we expect to explore this "pan and zoom" paradigm, popularized for military ships by SDMS [HERO80].

2.2.3. Interoperability

Users wish to share programs as well as data sets that are stored in the repository. Hence, a program written by one researcher must be usable by a second researcher. Our approach to this kind of interoperability of programs focuses around data base technology. We expect to encourage users to put their programs as well as their data into our prototype next generation DBMS system, POSTGRES. This software will run on the proposed workstations at the users' sites as well as on the centralized prototype repository discussed in the next section.

Specifically, application programs written by individual research groups should be thought of as **functions** which accept a collection of **arguments** of specific **data types** as input and produce an answer of some specific data type. With this methodology, we would then encourage users to define the data types of all arguments of each function as DBMS data types and register each function with the DBMS. As a result, any other user could **reuse** a function written by someone else by simply using the function in a query language command. The DBMS would be responsible for finding the code for the function, loading it into main memory, finding the arguments, and performing type conversion, if necessary. A powerful implication of this is that functions can be executed at the server as part of query evaluation, potentially reducing the amount of data that needs to be shipped across the network to the workstation in response to some types of queries.

We propose to explore a specific application of this idea at SDSC. SDSC clients employ visualization programs that use several different file formats. As a result, the output of one tool is often in a file format that cannot be read by a second tool. Consequently, visualization tools cannot interoperate. To combat this problem, SDSC has designed an intermediate internal data structure as a way of translating among the various formats, and has written translators from this intermediate form to some 20 different file formats. Also, these conversion routines currently run on 6 different hardware platforms. We propose to extend these libraries into a highly robust visualization software toolkit as well as install it in POSTGRES as a test of the usefulness of the data base paradigm for interoperability.

3. GOAL AND MILESTONES

Our goal is to support the requirement of Global Change researchers to perform real-time visualization of selected subsets of a large object base on a remote workstation over long-haul networks. We organize prototyping activity into a hardware component, an operating system component, a networking component, a DBMS component, a visualization component and a repository component. All these activities are interconnected to meet the above requirements as discussed in the following subsections.

3.1. Hardware Prototypes

We propose to construct two prototype high speed object servers and associated user interface software. The first will be constructed from off-the-shelf hardware and should be operational by June 1992, and we call it **BIGFOOT I.** Although SDSC has an existing tertiary memory system, it is based on technology nearing the end of its evolution, unable to scale to meet long-term needs of researchers (9 track tape and 14 inch disks). We believe that future tertiary storage systems will be built from large numbers of small form-factor devices. The purpose of BIGFOOT I is to gain experience with such devices and to explore striping techniques to augment bandwidth.

BIGFOOT I will store 1-3 Terabytes of information to be specified by the Santa Barbara, Los Angeles and Scripps teams. It will consist of 100 Gbytes of DEC disks backed up by both an optical disk jukebox and several Exabyte 8mm tape jukeboxes. The server for this I/O system will be two DEC 5500 CPUs running the Ultrix operating system. We propose to engineer this system so that members of the Sequoia research team can receive data from BIGFOOT I at 45 Mbits/sec. over a T3 link to be supplied by UC as part of their contribution to the project.

The second prototype will be a "stretch" system that will require innovative software and hardware research, will run in June 1994, and be called **BIGFOOT II.** BIGFOOT II will have the following goals:

- 50 - 100 Terabytes of storage
- 200 Mbits/sec input or output bandwidth
- seamless distribution of the server over two geographically distant sites

It is likely that this system would be built primarily from 4mm tape robots with a RAID-style disk array of 2.25" disk drives serving as a cache. A sophisticated custom I/O controller would be built that would control both devices and perform automatic migration of objects as well as compression. Funding for the construction of this large-scale prototype has been solicited from DARPA. Hence, it is expected that Sequoia 2000 would supply any DEC equipment appropriate for BIGFOOT II, with the remainder provided external to the Sequoia program.

In parallel with these efforts, the San Diego Supercomputer Center (SDSC) will extend the tertiary memory system for their CRAY supercomputer to be software compatible with the BIGFOOT prototypes above. Their system, **PRODUCTION BIGFOOT,** must serve the needs of a demanding collection of existing supercomputer users, and will be scaled as rapidly as the technology permits toward the capabilities of BIGFOOT II. Using this compatibility, the caching strategies and protocols developed in other portions of the Sequoia Project can be tested in the prototype BIGFOOT I and II environments and then stress tested for supercomputer viability in the SDSC production environment.

3.2. Operating System Prototypes

On the 1992 prototype we propose to run Ultrix as noted above. Moreover, we will port the striping driver, operational on RAID I [PATT88], from Sprite [NELS88] to Ultrix to turn the 100 Gbytes of DEC disks into a redundant RAID 5 disk array [PATT88]. Furthermore, we will port the Log Structured File System (LFS) [ROSE91] from Sprite to Ultrix. This software will ensure that high performance will be available for all write operations. Lastly, we will implement a migration package which will allow blocks of files to migrate back and forth from secondary to tertiary memory. It is likely that we will start with the UniTree migration software available from Lawrence Livermore Labs, and extend it to be block-oriented rather than file-oriented. Our design goal is to implement migration software outside the kernel, so it can readily be ported to other environments. However, it is not implausible that we will require portions of the source code for the Ultrix file system. BIGFOOT I will also require Ultrix device drivers for our tertiary memory devices. If they are unavailable from another source, we propose to write these with the assistance of DEC engineers assigned to the project.

For the 1994 prototype, we propose to optimize the file migration code for both the SDSC UniTree environment as well as the BIGFOOT Ultrix environment. Moreover, we plan to build a higher performance distributed computing model than NFS. Finally, we fully expect that high performance sequential I/O out to the network will have to bypass the host CPU, as is done in the design of RAID II. We propose to engineer interfaces to allow I/O streams to proceed directly from an I/O controller to the network with minimal host CPU intervention. This will require access to Ultrix internals.

3.3. Networking Prototypes

The first concern of the networking group will be to produce a networking benchmark in cooperation with the Global Change scientists and the visualization group at SDSC. This is important because there is very little detailed information available for the required networking performance capabilities, and user tolerances and preferences (and extrapolations from other domains, such as real-time voice or video requirements, seem to be risky). This benchmark will then be used to best budget delay tolerances (and loss tolerances, if any) throughout the network.

We expect the Sequoia network to consist of a set of Local Area Networks (LANs) interconnected by 45 Mb/s T3 lines, initially. Given the 100 Mb/s data rate of FDDI LANs and their support for "synchronous" traffic with bounded delay, FDDI seems a natural choice for local distribution. (Some problems experienced with FDDI parameter settings, leading to high overhead and eventually to low data rates, can be effectively controlled in this experimental environment and will be eventually resolved.) However, even though FDDI networks provide the necessary guarantees at the medium access control layer, there are no protocols at the network and transport layers to make effective use of this underlying capability. Moreover, there are no existing protocols that deal with the synthesis of end-to-end guaranteed performance channels from a set of such channels and networks in tandem. A network layer protocol (RTIP) and some ideas that address this problem of providing end-to-end performance guarantees have been described in [LOWE91] following the methodology for real-time virtual circuit establishment provided in [FERR90a]. An important feature of the RTIP is that it is an extension of IP (the Internet network layer protocol), and interoperates with it. We propose to implement RTIP, by extending IP, and to experimentally evaluate its performance.

We also need to extend TCP to make use of RTIP. In addition, we plan to modify it to allow applications to specify that they don't require error control (and thus, alleviating the need for software checksum computations, retransmissions, and timer management, which represent major parts of the overall packet delay [CLAR89]). Notice also that, for real-time traffic, a transport layer relying on RTIP has little functionality to add, and thus can be made very lightweight. Therefore, we will consider implementing a transport layer protocol from scratch. This work will be complemented by I/O system software restructuring to specifically include in-kernel data paths as discussed earlier. Processing modules such as the network protocols will be uniformly incorporated in the I/O system. Note that this will require Ultrix source code for the kernel and device drivers.

In parallel with the effort for real-time protocol development we plan to investigate the traffic matrix requirements of the Sequoia sites in order to design the topology of the network. This is required at present since we expect that applications will require a substantial portion of the full T3 bandwidth. As high-speed link availability increases in the future, this will be less critical and will be entirely subsumed by the real-time channel establishment phase of RTIP. These efforts will culminate with a demonstration of continuous satellite image transfer between BIGFOOT I and remote workstations.

For the 1994 prototype, we will investigate alternative approaches to providing the required network performance based on more subtle forms of resource reservations (e.g., by introducing traffic priorities in conjunction with input rate enforcement mechanisms such as the "leaky bucket"). This technique addresses the question of the delay introduced by the virtual-circuit set-up, but is more weak in terms of expressed performance guarantees. We plan to implement PIP, a network layer protocol based on this technique, as an alternative to RTIP and evaluate it against the networking benchmark and RTIP.

In addition, we will develop software to perform hierarchical coding and will incorporate into the networking code the functionality to deal with the additional information available on the packets. For example, with PIP, this would simply mean different priority classifications for the various coding components; non-essential components could then be dropped if required, exactly as current IP specification allows. This will enable us to demonstrate fast-forwarding capability through continuous images. A final but important effort will be the comparison of the functionality and performance provided by the Sequoia networking approach and the BLANCA and/or CASA gigabit networks.

3.4. DBMS Prototypes

We will extend the prototype DBMS POSTGRES to effectively manage data base spread between secondary and tertiary memory. This will require changes to the POSTGRES storage model as well as extensions to the query optimizer to effectively make query execution plans for the environment.

POSTGRES works well on point and vector data; however, we will have to find and implement solutions to efficiently manipulate raster data. We propose to implement the **chunklet** solution described earlier and explore its performance. Furthermore, as the electronic repository gets constructed, we propose to extend the POSTGRES indexing paradigm to meet its needs. The initial indexing system that we propose to adapt for this purpose is contained in the CHESIRE information retrieval package [LARS91].

Lastly, we propose to integrate a set of user interface routines which display and manipulate maps to POSTGRES. This system, GRASS, will give Sequoia researchers a convenient set of user interface routines. These extensions will be realized through a succession of releases of the POSTGRES DBMS over the course of the project.

3.5. Visualization Prototypes

Available visualization packages (primarily PV-Wave, NCSA, SPAM, IPW, and IMagic, AVS, GEMPAK, and SDM) will be examined for their useful and unique features, and the SDSC work on translation between image formats will be extended to allow us to use a variety of visualization tools on the same sets of data, without necessarily writing yet another visualization package ourselves. We will focus on fabricating a volume visualization system, starting with existing packages. This system will allow convenient specification of the volume to be rendered, perform rendering in a device-independent way, and run on hardware from the SDSC CRAY Y-MP to users' workstations and the BIGFOOT server. We will also extend our visualization hardcopy system to include the automated production of videotapes.

3.6. Electronic Repository

We also plan to bring up a prototype electronic repository. This will be a large POSTGRES data base of documents, geographic data, programs, and data sets relevant to Global Change. Into this repository we will load as many of the Santa Barbara, Los Angeles and Scripps data sets as is possible. Furthermore, we will import as much of their suite of application programs as possible into POSTGRES as user-defined functions. We will also bring up POSTGRES on an SDSC machine, and define our visualization code as data types and functions as a stress test of our interoperability ideas.

The repository will also include a collection of UC technical reports on Global Change. Some of these are available in electronic form but others will be scanned in from paper copy. Further, we will solicit Global Change documents from other sources for our storage and dissemination. We propose to provide a reviewing service which will generate textual evaluations of submitted works. Hence, a user will be able to ask for the ‘‘good’’ papers on ozone depletion, rather than for all of them. Our methodology will be to define a POSTGRES schema for documents which recognizes that documents have a variety of formats (e.g. Troff, Framemaker, ...).

To this end, we will include any **converters** that we could obtain; this will allow documents stored in one format to be retrieved in another. Moreover, we will add sophisticated indexing discussed in the previous section as well as user interface code.

Furthermore, the Division of Library Automation (DLA) in the Office of the President will work with the project to link our repository into the mainstream library infrastructure nationally through the development of server using the NISO z39.50 information retrieval protocol [LYNC91]; this would be implemented as an application layered on top of POSTGRES. This server would support z39.50 client access to Sequoia technical reports and perhaps some of the other kinds of Sequoia objects. This project will build on previous DLA work on z39.50 done in cooperation with DEC.

Lastly, we will build at least two prototype browsing interfaces for the repository, one based on a traditional information retrieval paradigm, and one based on iconic representation of objects.

3.7. Technology Transfer

There are several ways in which we propose to enhance technology transfer of our ideas. First, we will continue the tradition started by SPUR and XPRS of holding periodic retreats to discuss research results with industrial partners. Second, we propose to publish a technical report series and a periodic newsletter using our electronic repository as a distribution mechanism. Third, we expect to distribute our software to others who might wish to run it. Technology transfer through prototype software has been a successful distribution medium, and we expect to continue this tradition.

4. Involvement of Other Organizations

Several other organizations have already agreed to collaborate on the proposed research. These include:

DEC Colorado Springs Research Laboratory
DEC San Francisco Research Laboratory
TRW
Exabyte Corp.
National Center for Atmospheric Research
State of California Air Resources Board
State of California Department of Water Resources
United States Geological Survey
University of Colorado

In this section we briefly describe the role of each of these collaborators.

The DEC Colorado Springs Research Laboratory is interested in setting up an electronic information repository for that facility. They will donate the time of one engineer to work on this project. Our plan is to have them run the same object server software that we run, and to bring up all the DEC technical reports that are publicly available.

Dr. Jim Gray, Director of the DEC San Francisco Research Laboratory, has agreed to be a consultant to the project and to spend up to one day a week in technical discussions.

TRW Corp. does real time reconnaissance projects for various agencies of the U.S. Government. As such, they are keenly aware of the need to integrate image, sensor and other kinds of data. They wish to participate in the project as observers and possible users.

Exabyte Corp. is interested in cooperating with the project and has donated an 8mm tape jukebox with a capacity of 590 Gbytes. As noted earlier, we will integrate this storage device into BIGFOOT I. Ms. Kelley Scharf of Exabyte will be the liaison person to this project.

The National Center for Atmospheric Research (NCAR) and the University of Colorado (CU) are co-investigators on the AVHRR project, producing other data sets besides the snow-cover maps. The Department of Aerospace Engineering at CU will retrieve the satellite data through their antennae, and will send them to BIGFOOT for archiving. Mr. Paul Rotar of NCAR and Dr. William Emery of the University of Colorado will be the liaisons.

Within NOAA's National Meteorological Center, Dr. Eugenia Kalnay is head of the Development Division and Principal Investigator of NOAA's Climate Data Analysis System, for which Professor Ghil is a scientific consultant, and of their joint Re-Analysis Project with NCAR. She has agreed to pledge a planned 35-year homogeneous, four-dimensional climate data set to span 1958-1993. This data set will be used in the UCLA group's work on assimilation.

The State of California Air Resources Board is charged with monitoring air, water, and snow pollution in California. Their role in this project is mainly through their precipitation chemistry measurements, particularly at seven sites within the Sierra Nevada at which snow chemistry measurements are made throughout the winter. Dr. Kathy Tonnessen from the ARB staff will be their liaison.

The role of the Department of Water Resources is to protect, conserve, develop, and manage California's water. DWR is responsible for supplying water for personal use, irrigation, industry, recreation, power generation, and fish and wildlife. DWR also encourages flood plain management; assures public safety by managing, maintaining, and operating flood control facilities under its jurisdiction; and provides flood warning information as flood threats develop. One aspect of this water management is controlling the water level in all the state reservoirs. Consequently, they are interested in the research of the Santa Barbara team, which is trying to predict snow melt rates and resulting runoff.

Another major program of DWR is the monitoring of agricultural land use and crop pattern changes in California. These data are used by DWR to estimate evapotranspiration, develop ground and surface water models, and to prepare a water balance for the state. Because of these applications, DWR is interested in the development of a GIS to survey cropping patterns annually in the agricultural sector, using remote sensing data.

For these reasons, DWR wishes to become users of the geographic information system being constructed. To this end, they will install compatible DEC equipment at their site in Sacramento and high speed connections to the UC network, so they can access BIGFOOT. DWR has pledged to contribute Mr. Gary Darling from their staff to be a liaison to the project. He will work with the project up to one day a week to ensure that project goals remain aligned with DWR needs.

The United States Geological Survey will cooperate with the project in several ways. USGS is the producer and archiver of a multitude of data sets required for Global Change research including AVHRR data, Landsat data, aerial photography, digital elevation models, and digital map data. As such they are interested in the applications of database, mass storage, and network communications technologies to enhance the access and use of earth science data sets by not only the Global Change research community (both within and outside USGS), but by the general public as well. In addition to using the system as a test bed for their research applications, USGS will facilitate access of other project scientists to the appropriate data holdings. Dr. Stephen Guptill will serve as liaison to the project to advise on the integration of geographic data and geographic information system functionality within an extensible database system and ensure that the resulting project remains aligned with USGS needs. Dr. William Miller, USGS, will implement a prototype application utilizing geologic data on the GRASS/POSTGRES system and serve as a user of the test bed. Additional applications of the test bed for multiple resolution imagery data sets as well as digital map data are envisioned.

5. PHASE 2 OF SEQUOIA 2000

This proposal contains a three year technical plan. We view this as only the start of a more extensive research effort; if Phase 1 is successful, we will propose to embark on a Phase 2 effort to last several additional years. The goals of Phase 2 would be:

support for massive distribution of data
support for heterogeneity
real-time collaboration
more sophisticated user interfaces

In point of fact, there will be data sets relevant to Global Change on thousands of computer systems worldwide. We must extend our electronic data repository to **seamlessly** access this worldwide information base. We plan initial work on distributed file systems and distributed data bases in Phase 1 which will lay the groundwork for a greatly enlarged effort in this direction in Phase 2. The goal will be to support high performance access to distributed data sets in a **seamless** fashion, so that a user need not know the location of desired data. This same seamless quality must also be present in DBMS services, and the query optimizer must be extended to support distributed queries in a **location transparent** way.

Access to this multiplicity of data sets generates a requirement to seamlessly **merge** data sets. For example, both DWR and the U.S. Bureau of Reclamation collect data on the depth to groundwater at selected sites in the San Joaquin Valley of California. These measurements are taken at different times at different locations with different measurement techniques. One would like to request the expected depth to groundwater at a specific site, which was not co-located at a DWR or Reclamation site. To estimate the "best" answer to this query, the two data sets above must be merged. Although they contain no common information, a composite data set must be created based on statistical or expert system techniques about the component data sets. We expect to work on generic techniques to support merging of such data sets.

During Phase 1 we will concentrate on the electronic repository as the most effective way of supporting collaboration. However, during Phase 2 we will focus on real-time interactive collaboration between small groups of investigators. This will include real-time interactive video and implementation of shared windows on geographically remote workstations.

We expect the most profound impact of this project to be in the software area. In industry, there is already some development of distributed object repositories, and we anticipate that these industrial trends will be influenced by Sequoia 2000. Additionally, the underlying technology (e.g., RAID) will have an impact broader than its immediate applications. In short, we expect that our repository, visualization, migration, and guaranteed network services will find their place in future mainstream computer systems in the same way that X windows and Kerberos from the Athena project migrated into commercial offerings.

Lastly, we expect to provide much more sophisticated interfaces for visualization and for browsing of the electronic information repository.

In Phase 2, we expect to involve a much larger community of Global Change researchers, including major roles for groups at the Scripps Institute of Oceanography and the Department of Geosciences at Irvine. As such, we expect the California Space Institute at San Diego and the San Diego Supercomputer Center to play a more central role in supporting this larger group of researchers. In addition, the DEC-supported computing installation at the UCSB Map and Imagery Library may play a role in making BIGFOOT services available to less computationally-sophisticated science users.

6. CRITICAL SUCCESS FACTORS

There is only one aspect of this project that merits discussion in this section, namely the technology from which we expect to build BIGFOOT II. If we attempted today to reach a storage system of 100 Terabytes, we would require 160 Exabyte jukeboxes. For both cost and footprint considerations, this is an infeasible storage system. Therefore, within the next two years, we will require at least a factor of four increase in density and decrease in cost for BIGFOOT II to achieve its size goal. If this does not materialize, then BIGFOOT II will have to be scaled back in capacity to a feasible size.

7. QUALIFICATIONS OF THE INVESTIGATORS

A strength of this proposal is that it combines scientists who need management and storage of massive amounts of data with computer and information scientists who want to build the hardware and software to make it happen. The computer scientists involved are those who established Berkeley's reputation as one of the world's leading system's research institutions. This group did the pioneering work on the following ideas (ranging from hardware to applications):

Reduced Instruction Set Computers (RISC)
Redundant Arrays of Inexpensive Disks (RAID)
Portable, Open Operating Systems
Relational Databases
Computer Aided Design Tools for Very Large Scale Intergrated Circuits

and demonstrated the value of the innovations by building working systems in each of these areas. Also involved in this proposal are leading information scientists from the UC Berkeley School of Library and Information Studies. Both Michael Buckland and Ray Larson were substantially involved in the development of MELVYL, the UC-wide on-line library catalog.

Joining this technical team is a collection of world-class earth scientists from four UC campuses. They are led by Jeff Dozier of UCSB, who is currently the Project Scientist for NASA's Earth Observation System. He is currently on part-time leave from the University, and he will be assisted by Catherine Gautier from CRSEO, UCSB.

The combination of new opportunities in technology, users with real problems to solve, and systems builders with an established track record make this an unusually exciting project. We expect the most profound impact of this project to be in the software area. In industry, there is already some development of distributed object repositories, and we anticipate that these industrial trends will be influenced by Sequoia 2000. Additionally, the underlying technology (e.g., RAID) will have an impact broader than its immediate applications. In short, we expect that our repository, visualization, migration, and guaranteed network services will find their place in future mainstream computer systems in the same way that X windows and Kerberos from the Athena project migrated into commercial offerings.

APPENDIX I

INVOLVEMENT BY GLOBAL CHANGE RESEARCHERS

The Global Change researchers participating in Sequoia 2000 are exploring many of the key questions in Earth System Science. Jeff Dozier (fresh water and snow cover, Santa Barbara) studies water and nutrient balances of snow-covered alpine watersheds. Ray Smith (ocean productivity and Antarctic ecosystems, Santa Barbara) models primary production in the ocean, and has recently returned from an expedition to the Antarctic Ocean that examined the effects on phytoplankton and the aquatic ecosystem of increased ultraviolet radiation due to atmospheric ozone depletion (the ‘‘ozone hole’’). Catherine Gautier (radiation balance and large-scale hydrology, Santa Barbara) uses large-scale meteorological satellite images to estimate the Earth’s radiation balance and the global fresh water flux at the ocean surface and investigate how they are related to clouds and how they may change as a result of global warming. Frank Davis (terrestrial ecosystems, Santa Barbara) uses high-resolution Landsat Thematic Mapper and SPOT data, along with digital elevation data, to study plant ecosystems. Michael Ghil (theoretical and observational climate dynamics, Los Angeles) studies the climatic system at time scales from days to millions of years, and assimilates satellite data into coupled atmosphere-ocean GCMs. Carlos Mechoso (coupled ocean-atmosphere modeling and stratosphere dynamics, Los Angeles) models the global atmosphere and frontal systems with an emphasis on the stratosphere of the Southern Hemisphere. David Neelin (interannual climate variability and the El Nio/Southern Oscillation, Los Angeles) investigates the oscillations of the tropical oceans and global atmosphere. Richard Turco (atmospheric chemistry and tracer modeling, Los Angeles) models heterogeneous atmospheric chemistry and tracer transport with an emphasis on the Antarctic Ozone Hole. Richard Somerville (cloud-radiation interactions, San Diego) examines the effects of clouds on radiation and their parameterizations in GCMs. Tim Barnett (air-sea interactions and climate predictions, San Diego) studies air-sea interactions and their role in climate prediction. John Roads (atmospheric diagnostics and predictability, San Diego) performs diagnostic studies of the atmosphere and its predictability.

In this section we indicate four scenarios that present a cross section of the investigations that these researchers intend to pursue. Each of these scenarios represents an area in which technology developed in the Sequoia 2000 project will provide major improvements in the quality of the research that Global Change scientists can pursue.

Analysis of High-Resolution Land Data

For the snow-cover mapping project and other studies of land processes in the western U.S., AVHRR data will be acquired at the University of Colorado, Boulder, for every orbit within the field-of-view of their local antenna. The data will be sent to UCSB, where they will be analyzed by the snow-mapping algorithms [DOZ81] and then integrated into the BIGFOOT system. The data base will include information about the algorithm generation, calibration coefficients, etc. so that long time series can be checked for consistency. In addition, reduced-volume products will be compressed and sent to cooperating investigators in universities and laboratories throughout the western U.S. The snow-cover maps from the 1-km-resolution AVHRR data will be periodically validated with higher resolution data, from Landsat Thematic Mapper (7 spectral bands, 30 m resolution) [DOZ89], SPOT (4 bands, 10 and 20 m resolution), and AVIRIS (224 bands, 20 m resolution) [DOZ90a]. Algorithms using the higher-resolution data need to be combined with digital elevation models, available from the U.S. Geological Survey. Radiation calculations on these data are used to drive energy-balance snowmelt models [DOZ90b]. To extend these studies to other parts of the world, it will be necessary to archive and manage a global 1 km data set from AVHRR. NASA is assembling such a data set, to cover one year, and will make it available to us. These data will be assembled from various ground-recording stations and local tape-recorder coverage on the satellite.

Analysis of High-Resolution Visible Data for Photosynthetically Available Radiation (PAR) Computations

Another project of CRSEO researchers is the study of the spatial and temporal variability of Photosynthetically Active Radiation (PAR) as it pertains to primary production over both land and oceans. Biomass production represents a key element in understanding the processes controlling biogeochemical cycling on regional to global scales, and, ultimately, the fate of atmospheric carbon dioxide. Over the ocean, it has recently been shown [MORE78, PLAT88] that there is a stable relationship between PAR and the energy stored by phytoplankton photosynthesis when normalized to integrated water column pigment biomass. That is, bio-optical models can be used to accurately predict oceanic primary production given a knowledge of PAR and column-integrated pigment biomass [BIDI87, SMIT87a, SMIT87b, SMIT91]. Pigment biomass can be estimated spatially using satellite ocean color sensors (CZCS, SeaWIFS, MODIS-T on EOS) and temporally using moored optical sensors [SMIT87b, SMIT91]. PAR can be obtained as discussed below so that these combined data sets can be used to accurately predict oceanic primary production on space and time scales that have been impossible thus far. Over land, similar but more complex models can be applied to relate indices of vegetation to vegetation amount and activity. Green leaves are the only Earth-surface materials that strongly absorb visible wavelengths while reflecting strongly in the near-infrared wavelengths longer than 0.7 μm [CURR81]. Spectral reflectances of other materials, such as soil and dead litter generally increase smoothly with wavelength, resulting in lower reflectance in near-infrared wavelengths. There is an approximately linear relationship between amount of photosynthetically available radiation (APAR) and the near-infrared/visible reflectance ratio [MONT77, HATF84, SELL86, CHOU87]. Spectral reflectance ratios are closely related to instantaneous rates of activity (photosynthesis, transpiration). Their relationship to structural properties of the canopy such as biomass and leaf area index is monotonic but non-linear. Thus, time integrals of reflectance ratios could be more directly related to primary productivity than values from single dates, as has been demonstrated in studies that have related the Normalized Difference Vegetation Index (NDVI) derived from multi-temporal AVHRR imagery to geographic, seasonal, and interannual variations in vegetation cover and, in turn, to global carbon fluxes, net primary productivity at biome levels, and temperature and precipitation fluctuations [JUST85, TUCK86, MALO90]. The relationship of spectral vegetation indices to surface biophysical conditions is complicated by atmospheric effects, solar geometry, topography, vegetation structure and canopy shadows, and variable soil reflectance. Models of radiative transfer in plant canopies are under development to strengthen the relationship between reflectance data and canopy conditions [FRAN91]. Alternative vegetation indices that account for variations in soil background also appear promising [HUET88, MAJO90]. Combining satellite imagery with digital maps of topography, soils, and land management can be used to greatly improve the relationship between spectral properties and surface conditions [DAVI90, DAVI91].

PAR can be accurately obtained following the method originally developed by [GAUT80] and refined and validated by [FROU89 and FROU90] from high resolution (0.5 hour and 1 km) and/or degraded resolution (3 h and 32 km) visible satellite observations; the optimal resolution depends on the scale of the processes at hand. For example, phytoplankton varies over a wide range of space/time scales in the ocean [SMIT87a]. High productivity in the coastal upwelling regions is limited to within a few tens of kilometers of the coast and varies over relatively small scales (˜ 1 km.), whereas over the open ocean the spatial variability occurs over much larger scales. We propose to match our sampling scheme to this natural space/time variability performing our computations over a scale of 1 km (within 25 km off the coasts), 4 km (within 250 km off the coasts) and over land, and 32 km in the open ocean.

For this project both geostationary and AVHRR data will be acquired from a variety of sources and loaded into BIGFOOT. One kilometer (1 km) resolution geostationary data will be obtained from the University of Wisconsin where it has been archived since 1979; it will cover latitudes up to 50 degrees. Global Area Coverage (4 km) resolution data will be acquired from NOAA/NESDIS, the national archive center. Low resolution (32 km) geostationary and AVHRR data will be acquired from the International Satellite Cloud Climatology Project. PAR will be computed from BIGFOOT data on an hourly basis within coastal regions and on a 12 hourly basis away from the coasts.

Model Simulations and Data Assimilation

The Climate Dynamics Center runs a suite of atmospheric and oceanic GCMs. Both are grid-point, finite difference models that are constantly being refined. The UCLA atmospheric GCM predicts ground temperature, ozone mixing ratio, potential temperature, horizontal velocity, water vapor mixing ratio, surface pressure, and planetary boundary layer (PBL) depth. The resolution of grid cells is as great as 2.5 degrees longitude by 2 degrees latitude, and the model contains 9, 15 or 17 atmospheric layers. This model has been used extensively to predict tropospheric blocking and stratospheric warming events [MECH85, MECH86] and in studies of the impact of SST anomalies on atmospheric circulation [MECH90]. Roughly, the atmospheric GCM produces 1 Gbyte of output per simulation year at its lowest resolution while the lowest-resolution ocean GCM produces 5 Gbytes per year. Each climate simulation requires tens to hundreds of years, and tens of simulations are required for the study of climate sensitivity to various external factors. As a result, Terabytes of simulation output are generated. Not only must this output be stored for long periods of time for visualization purposes, but also they must be compared with Terabytes of measured data to validate the models. More generally, it is becoming clear to researchers that model predictions can be improved if model output and raw data observations can be intercompared. Data assimilation has been used for more than 30 years in numerical weather prediction to obtain optimal estimates. As a result, we expect to make climate analysis and predictions based both on model output and on observations. Therefore, we need efficient access to both kinds of data sets in a unified environment, and we expect to use BIGFOOT for this purpose. Four-dimensional assimilation results over tens of years from the coupled ocean-atmosphere GCM will be stored, at the highest resolution feasible by the end of the second year of the project. The physics and dynamics of the coupled system will be studied by combining the powerful visualization methods provided by the project with our statistical and dynamical analysis techniques. The same analysis methods, visual and statistical, will be applied to the available raw data sets and to the data sets of model-assimilated data, which will be obtained under separate funding, in collaboration with the university groups and government labs. Inferences about climatic changes and its predictability will be drawn. Both BIGFOOT and the SDSC visualization software will be very helpful in accomplishing the goals of this project.

Climate Process Studies and Modeling for Forecasts of Regional and Transient Aspects of Global Change

At Scripps, Richard Somerville has recently developed improved parameterizations of cloud-radiation interactions, taking into account the effects of cloud microphysics on cloud radiative properties [SOME84, SOME87]. He will incorporate these in global general circulation models (GCMs) which he and his colleagues are now running at the Livermore and Los Alamos National Laboratories to assess the role of cloud/radiation feedbacks in climate change, in modeling and predicting climate variability, and in assessing the predictability of climatic phenomena [SOME87]. This research combines the development of comprehensive numerical models of the climate system and the exploitation of large observational data sets based on both conventional surface-based meteorological measurements and satellite remote sensing capabilities for monitoring the entire planet. Multi-year simulations will be carried out, and the GCM simulations will be compared with the observational data. The specific tasks include surface hydrology budget analysis, studies of the interannual variability of the model climate, and the development of a diagnostic model that includes the same physics as the time-dependent GCM. A continuing goal is to improve the GCM treatment of clouds by comparing the model results with recent observations, particularly satellite remote sensing data [CHER91].

A second line of research is to use the GCM to provide information on the regional aspects of cloud feedback. Nearly all work on this subject to date has focused on the global average climate, but it is probable that cloud feedbacks will be variable from one part of the globe to another, and the three-dimensional resolution of the GCM makes it possible to address this important class of problems. One coupled air-sea climate phenomenon in which recent research has shown strong sensitivity to cloud-radiation interactions is the Indian monsoon [IACO91], and one goal of model improvements is to achieve greater predictive skill in monsoon forecasting.

This research will use BIGFOOT to store the data sets required for input and produced as output. The SDSC visualization software will also be sued to assist the individual researchers.

REFERENCES

[BAIL91] Bailey, M.J., ‘‘Scientific Visualization for a Large, Remotely-Distributed User Community,’’ *Graphicon Conference Proceedings*, pp. 11-26 (February 1991).

[BENG81] Bengtsson, L., et al. (eds.), *Dynamic Meteorology: Data Assimilation Methods*, Springer-Verlag, New York (1981).

[BELK87] Belkin, N. and Croft, W., ‘‘Retrieval Techniques,’’ *Annual Review of Information Science and Technology*, vol. 22, pp. 109-145 (1987).

[BIDI87] Bidigare, R. R., et al., ‘‘Oceanic Primary Production Estimates from Measurements of Spectral Irradiance and Pigment Concentrations,’’ *Global Biogeochemical Cycles*, vol. 1, pp. 171-186 (1987).

[CHER91] Chertock, B., R. Frouin, and R. C. J. Somerville, ‘‘Global monitoring of net solar irradiance at the ocean surface: Climatological variability and the 1982/1983 El Nino,’’ *Journal of Climate*, Vol. 4, in press (1991).

[CEES91] Committee on Earth and Environmental Sciences, *Our Changing Planet: The FY 1992 U.S. Global Change Research Program*, Office of Science and Technology Policy, Washington, D.C. (1991).

[CHOU87] Choudhury, B. J., ‘‘Relationships Between Vegetation Indices, Radiation Absorption and Net Photosynthesis Evaluated by a Sensitivity Analysis,’’ *Remote Sensing of Environment*, vol. 22, pp. 209-233 (1987).

[CLAR89] Clark, D. D., Jacobson, V., Romkey, J., Salwen, H. ‘‘An Analysis to TCP Processing Overhead,’’ *IEEE Communications Magazine*, vol. 27, No. 6, pp. 23-29, June 1989.

[COOP88] Cooper, W., ‘‘Getting Beyond Boole,’’ *Information Processing and Management*, vol. 24, pp. 243-248 (1988).

[CPM91] Committee on Physical, Mathematical and Engineering Sciences, *Grand Challenges: High Performance Computing and Communications*, Office of Science and Technology Policy, Washington, D.C. (1991).

[CROF89] Croft, W., et al., ‘‘Retrieving Documents by Plausible Inference: An Experimental Study,’’ *Information Processing and Management*, vol. 25, pp. 599-614 (1989).

[CURR81] Curran, P. J., ‘‘Multispectral Remote Sensing of Vegetation Amount,’’ *Progress in Physical Geography*, vol. 4, pp. 315-341 (1981).

[DAVI90] Davis, F. W. and Goetz, S., ‘‘Modeling Vegetation-Environment Association Using Digital Satellite and Terrain Data,’’ *Landscape Ecology*, vol. 4, pp. 69-80 (1990).

[DAVI91] Davis, F. W. and Simonett, D. S., ‘‘Remote Sensing and GIS,’’ in *Geographic Information Systems: Principles and Applications*, edited by Maguire, D. et al., Longman Ltd., in press (1991).

[DOZ81] Dozier, J., et al., ‘‘Effect of Grain Size and Snowpack Water Equivalence on Visible and Near-Infrared Satellite Observations of Snow,’’ *Water Resources Research*, vol. 17, pp. 1213-1221 (1981).

[DOZ89] Dozier, J., ‘‘Spectral Signature of Alpine Snow Cover from the Landsat Thematic Mapper,’’ *Remote Sensing of Environment*, vol. 28, pp. 9-22 (1989).

[DOZ90a] Dozier, J., ‘‘Bidirectional Reflectance of Snow from 0.4 to 2.5 μm,’’ *Proceedings IGARSS ’90*, IEEE No. 90CH2825-8, p. 1157 (1990).

[DOZ90b] Dozier, J. and Frew, J., ‘‘Rapid Calculation of Terrain Parameters for Radiation Modeling from Digital Elevation Data,’’ *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, pp. 963-969 (1990).

[EXAB90] Exabyte Corporation, ‘‘EXB-120 Cartridge Handling Subsystem Product Specification,’’ Part No. 510300-002 (1990).

[FERR90a] Ferrari, D. and Verma, D., ‘‘A Scheme for Real-Time Channel Establishment in Wide-Area Networks,’’ *IEEE Journal of Selected Areas in Communications*, vol. 8, pp. 368-379 (1990).

[FERR90b] Ferrari, D., ‘‘Client Requirements for Real-Time Communication Services,’’ *IEEE Communications Magazine,* vol. 28, pp. 65-72 (1990).

[FRAN91] Franklin, J., et al., ‘‘Thematic Mapper Analysis of Tree Cover in Semiarid Woodlands Using a Model of Canopy Shadowing,’’ *Remote Sensing of Environment*, in press (1991).

[FROU89] Frouin, R., D.W. Lingner, C. Gautier, K.S. Baker, and R.C. Smith, ‘‘A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance (PAR) at the ocean surface,’’ *Journal of Geophys. Res-Oceans,* vol. 94, pp. 9731-9742 (1989).

[FROU90] Frouin, R., and C. Gautier, ‘‘Variability of photosynthetically available and total solar irradiance at the surface during FIFE: A satellite description,’’ *Symposium on the First ISLSCP Experiment FIFE*, Feb. 7-9, 1990, Anaheim, CA.

[GAUT80] Gautier, C., G. Diak and S. Masse, ‘‘A Simple physical model to estimate incident solar radiation at the surface from GOES satellite data,’’ *Journal of Applied Meteorology,* vol. 19, pp. 1005-1012 (1980).

[GHIL89] Ghil, M., ‘‘Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework,’’ *Dynamics of Atmospheres and Oceans*, vol. 13, pp. 171-218 (1989).

[HAAS90] Haas, L. et al., ‘‘Starburst Mid-Flight: As the Dust Clears,’’ *IEEE Transactions on Knowledge and Data Engineering* (1990).

[HATF84] Hatfield, J. L., et al., ‘‘Intercepted Photosynthetically Active Radiation Estimated by Spectral Reflectance,’’ *Remote Sensing of Environment,* vol. 14, pp. 65-75 (1984).

[HERO80] Herot, C., ‘‘SDMS: A Spatial Data Base System,’’ *ACM TODS* (1980).

[HUET88] Huete, A. R., ‘‘A Soil-Adjusted Vegetation Index (SAVI),’’ *International Journal of Remote Sensing*, vol. 9, pp. 295-309 (1988).

[IACO91] Iacobellis, S. and R. C. J. Somerville, ‘‘Diagnostic modeling of the Indian monsoon onset: Part 1, Model description and validation, and Part 2: Budget and sensitivity studies,’’ *Journal of the Atmospheric Sciences,* vol. 48, in press (1991).

[JUST85] Justice, C. O., et al., ‘‘Analysis of the Phenology of Global Vegetation Using Meteorological Satellite Data,’’ *International Journal of Remote Sensing,* vol. 6, pp. 1271-1279 (1985).

[KATZ89] Katz, R., et al., ‘‘Disk System Architectures for High Performance Computing,’’ *Proceedings of the IEEE, Special Issue on Supercomputing* (December 1989).

[KARL89] Karlson, G. and Vetterli, M., ‘‘Packet Video and its Integration into the Network Architecture,’’ *IEEE Journal on Selected Areas in Communications,* vol. 8, pp. 380-390 (1990).

[KIM90] Kim, W. et al., ‘‘Architecture of the ORION Next-Generation Database System,’’ *IEEE Transactions on Knowledge and Data Engineering* (March 1990).

[KURE87] Kure, O., ‘‘File Placement and Migration in Distributed Systems,’’ Technical report UCB/CSD 88/413, Computer Science Division, University of California at Berkeley (1988).

[LARS91] Larson, R., ‘‘Classification Clustering, Probabilistic Information Retrieval and the Online Catalog,’’ *Library Quarterly*, vol. 61 (April 1991).

[LELE87] Lelewer, D. and Hirschberg, D., ‘‘Data Compression,’’ *ACM Computing Surveys*, vol. 19 (September 1987).

[LOWE91] Lowery, C., ‘‘Protocols for Providing Performance Guarantees in a Packet Switching Internet,’’ Technical Report TR-91-002, International Computer Science Institute, Berkeley (1991).

[LYNC88] Lynch, C. and Stonebraker, M., ‘‘Extended User-Defined Indexing with Application to Textual Databases,’’ *Proceedings 1988 VLDB Conference*, Los Angeles (1988).

[LYNC91] Lynch, C., ‘‘The z39.50 Information Retrieval Protocol,’’ *Computer Communications Review*, pp. 58-70 (January 1991).

[MAJO90] Major, D. J., et al., ‘‘A Ratio Vegetation Index Adjusted for Soil Brightness,’’ *International Journal of Remote Sensing*, vol. 11, pp. 727-740 (1990).

[MALO90] Malo, A. R. and Nicholson, S. E., ‘‘A Study of Rainfall and Vegetation Dynamics in the African Sahel Using Normalized Difference Vegetation Index,’’ *Journal of Arid Environments*, vol. 19, pp. 1-24 (1990).

[MARK91] Markoff, J., ‘‘A System to Speed Computer Data ,’’ *New York Times*, p. C7 (January 23, 1991).

[MECH85] Mechoso, C. R., et al., ‘‘Numerical Forecasts of Stratospheric Warming Events During the Winter of 1979,’’ *Monthly Weather Review*, vol. 113, pp. 1015-1029 (1985).

[MECH86] Mechoso, C. R., et al., ‘‘Numerical Forecasts of Anomalous Atmospheric Events During the Winter of 1979,’’ *Advances in Geophysics*, vol. 29, pp. 375-413 (1986).

[MECH90] Mechoso, C. R., et al., ‘‘The Impact of Sea Surface Temperature Anomalies on the Rainfall over Northeast Brazil,’’ *Journal of Climatology*, vol. 3, pp. 812-826 (1990).

[MONT77] Monteith, J. L., ‘‘Climate and the Efficiency of Crop Production in Britain,’’ *Philosophical Transactions of the Royal Society of London*, vol. B281, pp. 277-294 (1977).

[MORE78] Morel, A., ‘‘Available, Usable, and Stored Radiant Energy in Relation to Marine Photosynthesis,’’ *Deep Sea Research*, vol. 25, pp. 673-688 (1978).

[NELS88] Nelson, M., et al., ‘‘Caching in the Sprite Network File System,’’ *ACM Transactions on Computer Systems*, vol. 6, pp. 134-154 (1988).

[PANE91] Panel on Model-Assimilated Data Sets, *Model-Assimilated Data Sets for Atmospheric and Oceanic Research*, National Academy Press, Washington, D.C. (1991).

[PASQ91] Pasquale, J. et al., ‘‘Network and Operating System Support for Multimedia Applications,’’ Technical Report CS91-186, University of California, San Diego (1991).

[PATT88] Patterson, D. et al., ‘‘A Case for Redundant Arrays of Inexpensive Disks,’’ *Proceedings 1988 ACM-SIGMOD Conference on Management of Data*, Chicago, IL (1988).

[PLAT88] Platt, T., et al., ‘‘Ocean Primary Production and Available Light: Further Algorithms for Remote Sensing,’’ *Deep Sea Research*, vol. 35, pp. 855-879 (1988).

[POLL88] Pollard, A., ''New Storage Function for Digital Audio Tape,'' *New York Times*, p. C10 (May 25, 1988).

[RANA90] Ranade, S. and Ng, J., *Systems Integration for Write-Once Optical Storage*, Meckler, Westport, CT (1990).

[ROSE91] Rosenblum, M., and Ousterhout, J., ''The Design and Implementation of a Log-Structured File System,'' submitted for publication, February 1991.

[SELL86] Sellers, P. J., ''Relations Between Canopy Reflectance, Photosynthesis and Transpiration: Links Between Optics, Biophysics and Canopy Architecture,'' *Advances in Space Research*, vol. 7, pp. 27-44 (1986).

[SELT90] Seltzer, M. and Stonebraker, M., ''Transaction Support in Read Optimized and Write Optimized File Systems,'' *Proceedings 1990 VLDB Conference*, Brisbane, Australia (1990).

[SMIT81] Smith, A.J., ''Long Term File Migration: Development and Evaluation of Algorithms,'' *Communications of the ACM*, vol. 24, pp. 521-532 (1981).

[SMIT87a] Smith, R. C., et al., ''Optical Characterization of Primary Production Across a Coastal Front,'' *Journal of Marine Biology*, vol. 96, pp. 575-591 (1987).

[SMIT87b] Smith, R. C., et al., ''Multiplatform Sampling (Ship, Aircraft and Satellite) of a Gulf Stream Warm Core Ring,'' *Applied Optics*, vol. 26, pp. 2068-2081 (1987).

[SMIT89] Smith, P. J., et al., ''Knowledge-Based Search Tactics for an Intelligent Intermediary System,'' *ACM Transactions on Information Systems*, vol. 7, pp. 246-270 (July 1989).

[SMIT91] Smith, R. C., et al., ''Optical Variability and Pigment Biomass in the Sargasso Sea as Determined Using Deep Sea Optical Mooring Data,'' *Journal of Geophysical Research*, in press (1991).

[SOME84] Somerville, R. C. J., and L. A. Remer, ''Cloud optical thickness feedbacks in the CO2 climate problem,'' *Journal of Geophysical Research*, Vol. 89, pp. 9668-9672 (1984).

[SOME87] Somerville, R. C. J., and S. Iacobellis, T. Matsuno (ed.), ''Cloud-radiation interactions: Effects of cirrus optical thickness feedbacks. S hort and Medium-Range Numerical Weather Prediction,'' *Meteorological Society of Japan*, pp. 177-185 (1987).

[STON90] Stonebraker, M. et al., ''The Implementation of POSTGRES,'' *IEEE Transactions on Knowledge and Data Engineering* (March 1990).

[STON91] Stonebraker, M., ''Managing Persistent Objects in a Multi-level Store,'' *Proceedings 1990 ACM SIGMOD Conference on Management of Data*, Denver, CO (1990).

[TANA89] Tan, E. and Vermeulen, B., ''Digital Audio Tape for Data Storage,'' *IEEE Spectrum*, vol. 26, pp. 34-38 (1989).

[TUCK86] Tucker, C. J., et al., ''Relationship Between CO_2 variation and Satellite-Derived Vegetation Index,'' *Nature*, vol. 319, pp. 195-199 (1986).

[WILK90] Wilkinson, K. et al., ''The IRIS Architecture and Implementation,'' *IEEE Transactions on Knowledge and Data Engineering* (March 1990).