NASA's Earth Observing System Data and Information System 3

The National Aeronautics and Space Administration (NASA) plans to make its Earth Observing System Data and Information System (EOSDIS) the world's most capable and advanced data and information system. Although some aspects of EOSDIS are unique, the program is an example of some of the capabilities and challenges common to advanced remote sensing data and information systems of the future. EOSDIS can also be expected to have influence beyond global change research, serving as a catalyst for advanced computing and data system technologies.[1]

EOSDIS OVERVIEW

As a result of concerns that humanity is having a major, detrimental influence on the global environment, in 1990 the U.S. government initiated the U.S. Global Change Research Program (USGCRP).[2] NASA has played a major role in the USGCRP by orienting its Mission to Planet Earth (MTPE) toward the scientific objectives of the USGCRP. The centerpiece of NASA's MTPE,

[1] As one example, new tools for manipulating scientific imagery could benefit other fields relying on databases of three-dimensional structures, such as crystallography, medical imagery, and computer-aided design.

[2] "The USGCRP was established as a Presidential initiative in the FY 1990 Budget to help develop sound national and international policies related to global environmental issues, particularly global climate change." The USGCRP seeks to "address significant uncertainties in knowledge concerning the natural and human-induced changes now occurring in the Earth's life-sustaining environmental envelope . . The USGCRP is designed to produce a predictive understanding of the Earth system..." Office of Science and Technology Policy, Committee on Earth and Environmental Sciences, *Our Changing Plane/: The FY 1993 U.S. Global Change Research Program* (Washington DC: National Science Foundation 1992), pp. 3-4.

as well as the USGCRP, is the Earth Observing System (EOS). EOS consists of a space-based observing system (figure 3-1), a scientific research program, and its data and information system, EOSDIS.[3]

EOSDIS plays a crucial role in global change research. EOSDIS helps transform heterogeneous remotely sensed and other data into useful information for integrated, interdisciplinary, predictive studies of the Earth's environment. NASA planners expect EOSDIS to provide increasingly effective access to data, as well as extensive data processing and analysis, the tools needed by researchers to transform data into useful information for policy makers.

EOSDIS (figure 3-2) presents NASA with very difficult management and technology challenges. By the first years of the next century, NASA expects EOSDIS to manage over 80 trillion bytes of data per year from EOS satellites alone. Other spacecraft could contribute an additional 80 trillion bytes per year. Processed data from EOSDIS would be well over 300 trillion bytes per year, equaling more than 250 million 1.2 megabyte high-density floppy disks.[4] NASA faces the daunting challenge of making this enormous quantity of data easily usable for a wide variety of users, including 10,000 physical scientists and possibly as many as 200,000 other users, many with little detailed technical knowledge of remote sensing.[5] Furthermore, these data USHS and 'heir needs will change, as will the data system technologies used in the program.[6]

In addition, EOSDIS will administer the scheduling of observations, the calibration of EOS instruments, and the control of EOS spacecraft. To be successful, EOSDIS must effectively incorporate data from a wide range of sources: the EOS satellites, all other NASA Earth remote sensing missions, data from non-NASA space systems, and essential data from atmosphere-, ocean-, and land-based sensors.

NASA has strongly supported the EOSDIS portion of the Mission to Planet Earth since initial planning in the early 1980s. NASA officials believe interdisciplinary global change research demands much more from data systems than the traditional discipline-specific missions of the past. Data management from scientific spacecraft has sometimes suffered inadequate planning and budget neglect. Data systems in NASA programs generally have lower external visibility than accompanying space hardware, and problems in spacecraft and instrument development sometimes have depleted the nonspace portions of program funding. Figure 3-3 shows that most NASA Earth science funding in the 1980s was allocated to spacecraft development. In the 1990s, mission operations and data systems are a much larger proportion of NASA's Earth science budget.

NASA plans to devote more EOS funding to the nonspace segments of the program than to the

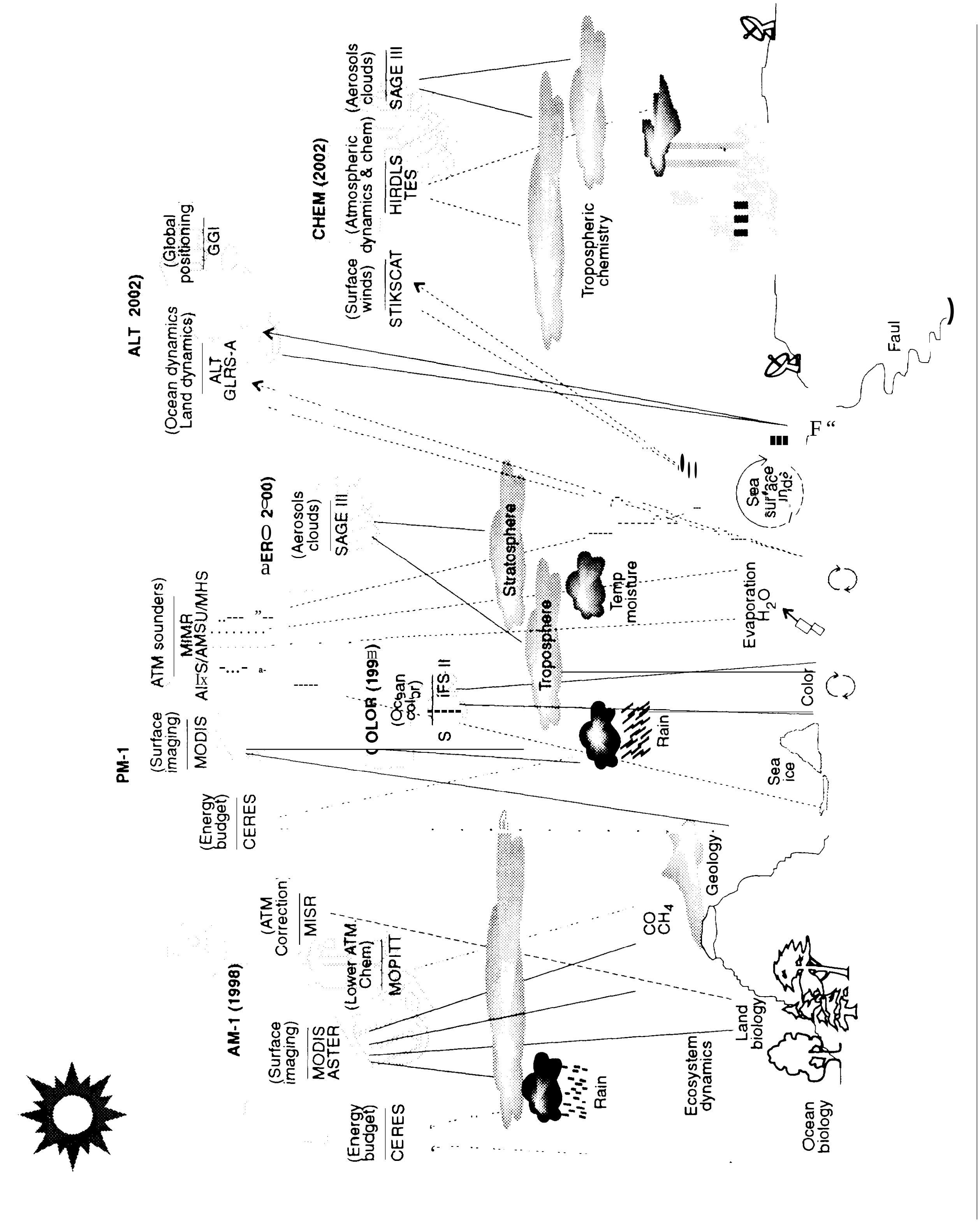
[3] See U.S. Congress, OffIce of Technology Assessment, *The Future ofRemote Sensin~fiom Space: Ci\'i/ian Satellite Systems and Applications.* OTA-ISC-558 (Washington, DC, U.S. Government Printing Office, July 1993) ch. 5.

4 In contrast, in 1990 the amount Of data archived from all NASA missions to date was about 8 trillion bytes, about 2.7% Of what is expected each year from EOSDIS.

5 The most recent estimates from Hughes predict 7, zoo to 16,000 EOSDIS users by 1998, excluding social scientists, libraries, and students. Adding these categories brings the estimated number of users to 76,000 to 200,000 (including a possible 174,000 students). In contrast, today's major supercomputer centers normally serve between 1,000 and 3,000 users. NASA and Hughes currently expect up to 1,000,000 EOSDIS user requests annually. NASA, *EOSDIS: EOS Data and Information System* (Washington DC 1992), p. 25, and Pitt Theme, "Demographics," EOSDIS Progress Review, Dec. 13- I 4, 1993, Landover, MD.

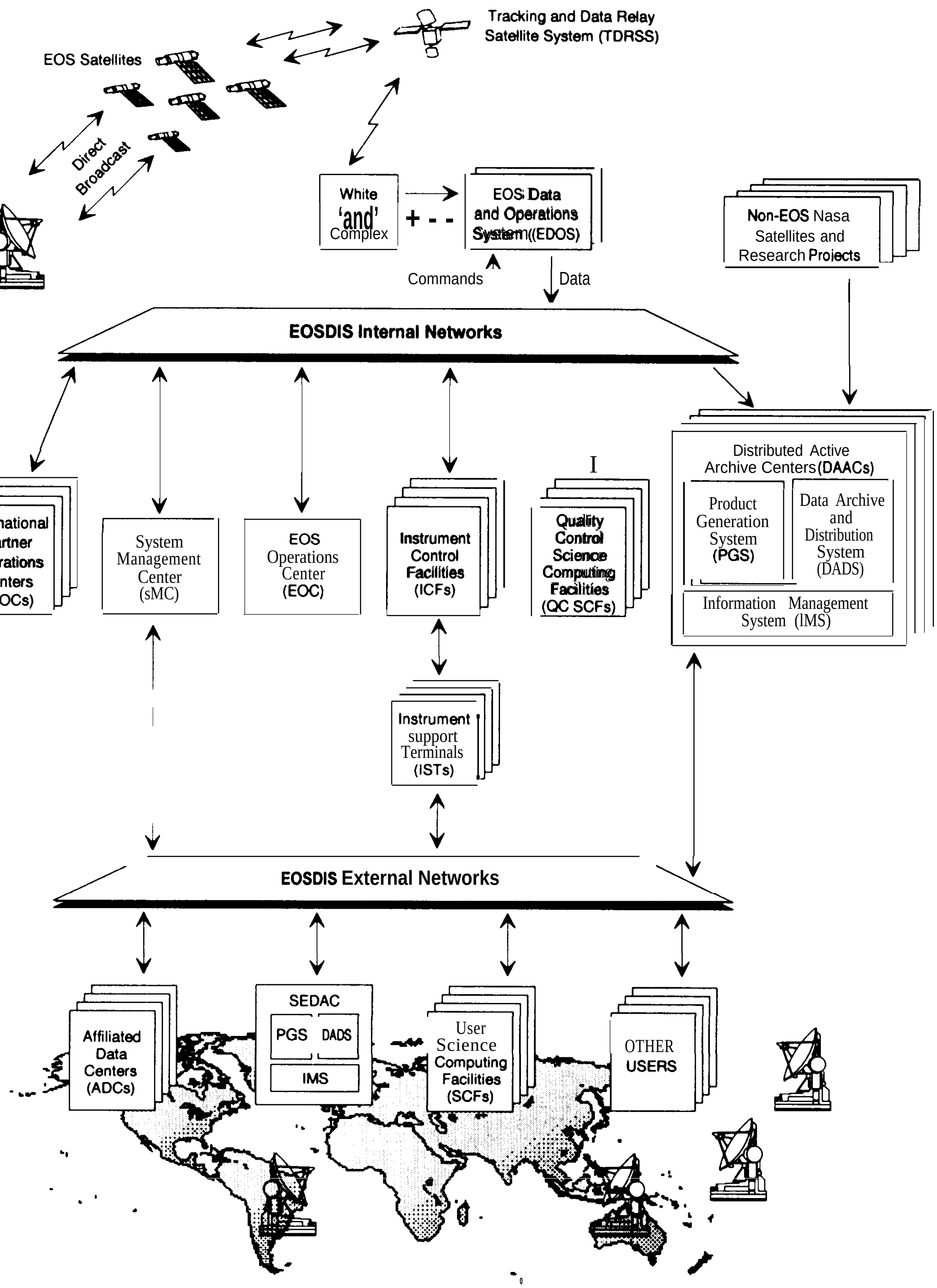
6 Recent experience with Europe's ERS-1 satellite underlines the importance of data systems in the success of research missions involving remote sensing. Although ERS is a single satellite with much smaller data flows than those planned for EOS, ESA has had difficulty processing some of the detailed data researchers need. "ERS- 1 Gives Europeans New Views of Oceans," *Science, vol. 260*, June 18, 1993, pp. 1742- 1743; see also ch. 5.

FIGURE 3-1: NASA's Planned Earth Observing System Sensors and Satellites



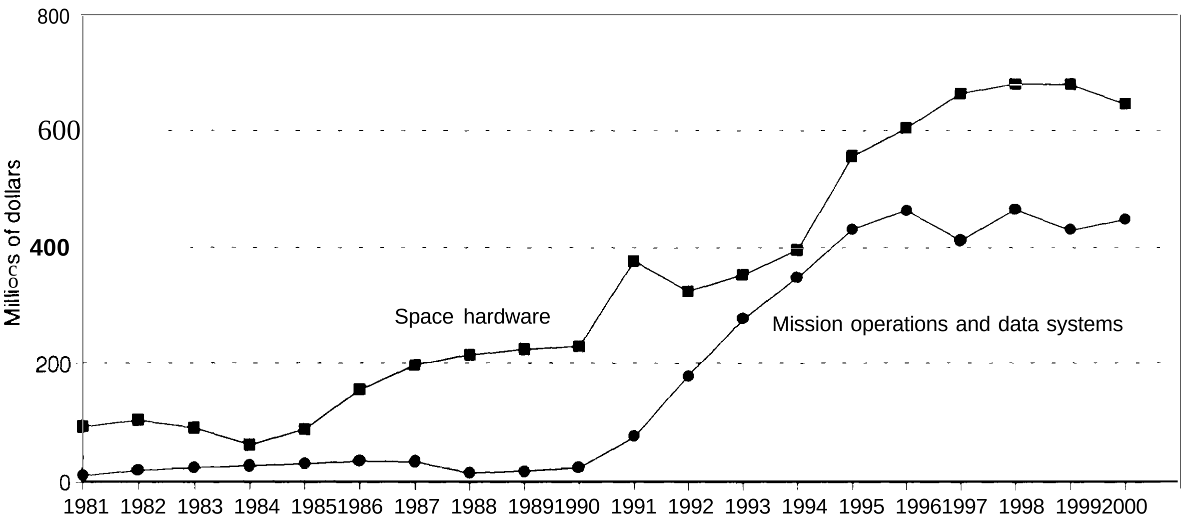
SOURCE: National Aeronautics and Space Administration, 994.

FIGURE 3-2: EOSDIS Architecture



SOURCE National Aeronautics and Space Administration, 1994

FIGURE 3-3: Budgets for Earth Science Space Hardware and Operations and Data Systems



SOURCE National Aeronatics and Space Administration, 1994

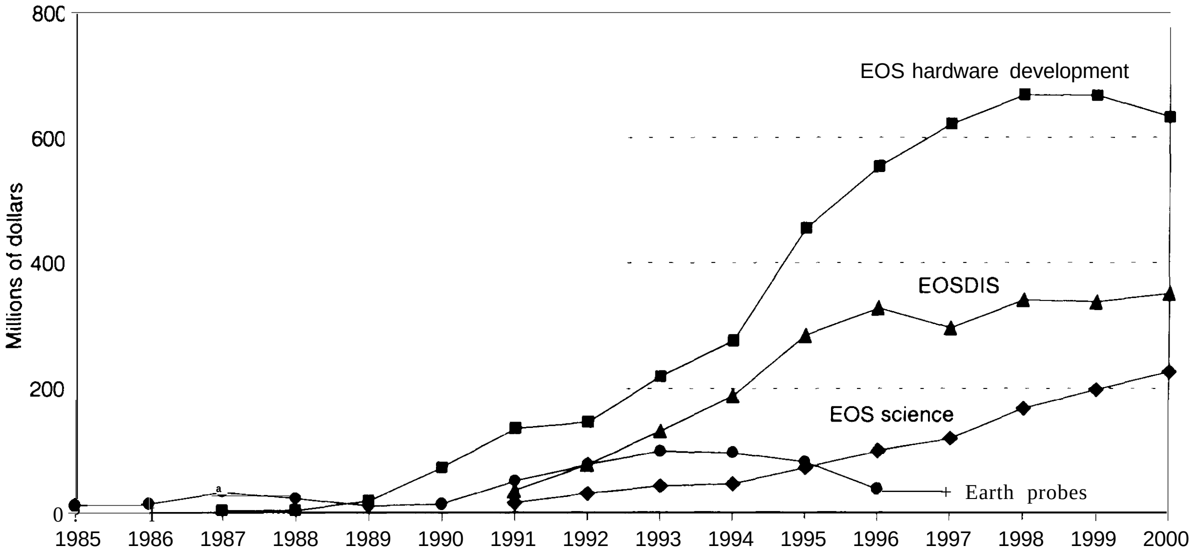
space segments, a major departure from previous space missions (figure 3-4). In the 1990s, about 30 percent of EOS funding will support EOSDIS, totaling about $2.37 billion.[7] NASA expects this massive government investment in EOSDIS to maximize its return on investments in space-based remote sensing hardware.

The previous chapter noted problems data users currently face navigating data archives. If EOSDIS is successful, researchers will spend much less time acquiring data, and will have easy and quick access to vastly increased amounts of data, allowing more time to transform these data into information.

Some other problems in the use of Earth observation data stem from the isolation of research disciplines. Research communities and individual researchers have highly individual views of how data should be organized and used. Disciplinary researchers also approach their research differently, using widely disparate nomenclature and methodologies, sometimes making communication across disciplines difficult. Some researchers claim Earth remote sensing data sometimes suffers from inadequate peer review of production algorithms, scientific quality control, and assessment. If EOSDIS is successful, it will help bridge the gaps among these diverse environmental re-

[7] At a projected 10,000 scientific users, EOSDIS funding in the 1990s would be an expenditure of $240,000 per research user. NASA Modeling, Data, and Information Systems Program Office, February 1993, and NASA *Budget Estimates, FY 1993* and FY 1994.

FIGURE 3-4: NASA EOS Expenditures—Hardware, Data Management, and Science



SOURCE Office of Technology Assessment, 1994

search communities, increasing interaction among disciplines and stimulating new research questions.[8]

Thus, the success of EOSDIS will be critical to the overall success of the Mission to Planet Earth and the USGCRP (box 3-l).

\ Incremental and Evolutionary Design

NASA states it has adopted an "incremental and evolutionary" approach to the development of EOSDIS. Because science and data requirements for studies of the Earth system will change as knowledge and experience grow, while computer technology develops extremely rapidly, EOSDIS must be capable of evolving.

NASA's approach to EOSDIS is a marked deviation from the typical data system development in which scientists and engineers perform a detailed "requirements analysis" for the system, followed by a comprehensive system design and development. Instead, by using an "open" architecture, NASA plans to reduce system costs and increase performance by delaying acquisitions of system components to take advantage of technology growth. This approach should also allow system users to play a role in each new increment of EOSDIS, a "learn-as-you-go" approach. NASA hopes to avoid costly system modifications that would follow delivery of a "monolithic" data system.[9] However, traditional government policies for budgeting, procurement, and contracting are all challenged by the trends of rapid increase in performance and decrease in cost of information technologies and rapidly changing user expecta-

[8] For further information on the possible impacts of information technology on scientific research, see National Research Council, *National Collaboratories: Applying Information Technology for Scientific Research* (National Academy Press, Washington, DC, 1993).

[9] Futher explanation of this development on on the possible impacts of information oach can be found in Taylor, Ramapriyan, & Dozier, "The Development Of the EOS Data and Information System," paper presented at 29th Aerospace Sciences Meeting, Reno Nevada, January 1991, p. 2.

BOX 3-1: Success Criteria for EOSDIS

Criteria to measure long-term success in EOSDIS are not quantifiable, However, NASA management Will consider EOSDIS a "success" to the extent it meets the following descriptive criteria

1 Maximization of number of users and '(intensity" of use of Earth science data.

2 Continuous Improvement in data access and services,

3 User satisfaction expressed in endorsements, political support, integration of EOSDIS into research plans, and willingness to use and contribute to the system.

4 Research results Increasingly robust to invalidtion by previous results or overlooked data

5 Users able to acquire the observations they request.

6 Voluntary provision of researchers' datasets for archiving and use by others,

7 Decrease In lag time from data Ingest to published research results.

8 Decrease in proportion of researcher time spent handling data vs. analyzing them

9 Increased use of EOSDIS data in wide ranging applications.

SOURCE NASA Modeling and Data Information Systems Program Management breif to OTA, Feb. 11, 1993

tions. The **EOSDIS evolutionary design will require steady, continuous funding, extremely close cooperation between NASA and the system contractors, and rapid procurement.**

Distributed Architecture

NASA states it has implemented a "distributed architecture" for EOSDIS, rather than central processing for all Earth observing data. Because expertise in various Earth science disciplines is geographically distributed across the country, NASA has chosen eight of the existing discipline-specific data centers as "Distributed Active Archive Centers" (DAACS); they will serve as geographically distributed "nodes" of the EOSDIS system (figure **3-5).** [10]

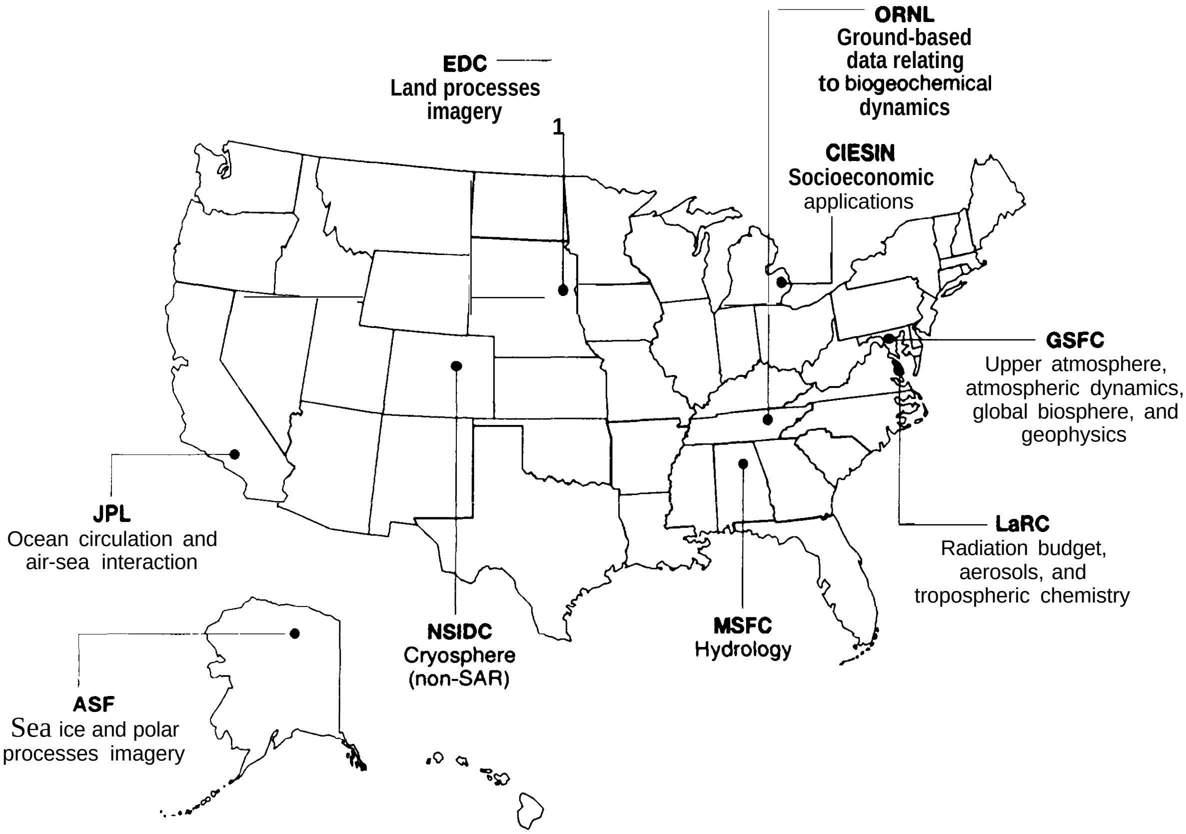
Until recently, these DAACS functioned as relatively independent data centers requiring users to contact each one individually in order to view data stored there. Each center set its own policies and methods for distributing data. By contrast, when EOSDIS is fully operational, users at any DAAC site will have complete access to all data sets anywhere in EOSDIS, regardless of physical location. Box 3-2 discusses DAAC system architecture.

A truly distributed system approach reduces the problems associated with failures at a *central* or *controlling* site. NASA expects the specialization and competition inherent in a distributed architecture to result in much better overall service to the EOSDIS user community, and avoid problems inherent in centralized, "bureaucratic control" of the system. [11]

[10] EOSDIS is working C] osely with the NOAA data centers, both to broaden access to NOAA data through EOSDIS, and to acquire operational, real-time or near-real-time data required for some EOSDIS data products. The special agreements between NASA and NOAA designate the NOAA data centers as '"Affiliated Data Centers" in EOSDIS. Affiliated Data Centers are not as closely linked to EOSDIS development as the DAACs and do not receive funding from NASA. However, officials of both NASA and NOAA expect to be able to access data easily from each other's systems.

[11] An early description of the rationale behind distributed architecture can be found in Dozier and Ramapriyan, "Planning for the EOS Data and Information System (EOSDIS)," 1990.

FIGURE 3-5: Geographically Distributed Architecture



ASF = Alaska Synthetic Aperture Radar Facility, CIESIN = Consortuim for International Earth Science Information Network; EDC = Earth Resources Observation Systems Data Center, GSFC = Goddard Space Flight Center, JPL = Jet Propulsion Laboratory; LaRC = Langley Research Center: MSFC = Marshall Space Flight Center, NSIDC = National Snow and Ice Data Center, ORNL = Oak Ridge National Laboratory.

SOURCE National Aeronauhcs and Space Administration, 1993

Technical requirements call for a distributed system as well, as NASA expects the computational power required to produce high-level EOSDIS data products to be immense. No single system could provide this performance. Instead, EOSDIS requires multiple systems, each with different characteristics. Also, the sheer number of expected users projected for EOSDIS would present a formidable service task if all users were using one site or had to interact with EOSDIS through a single site or system.

I EOSDIS Status

EOSDIS implementation includes three comprehensive contracts with information systems firms:

- The **EOSDIS Core System (ECS)** provides command and control of EOS spacecraft, science data processing, data archive and distribution, and communications, networking, and systems management. NASA selected Hughes Applied Information Systems as prime con-

BOX 3-2: EOSDIS DAAC Systems

Three systems Will operate at each DAAC

1) *Product Generation System (PGS);* the Product Generation System at each DAAC WIll convert raw data signals into standard sets of Earth science data, using data processing software developed by the scien - tific user community

2) *Data Archiva/and Distribution System (DADS);* the Data Archival and Dlstribution System at each DAAC wiIl serve as the archive and dlstribution mechanism for EOS and interdiscplinary data products, as well as essential ancillary data such as radiometric and geometric calibrtions, metadata, command history, algorithms documentation, and correlative data from EOS and non-EOS Sensors

3) *information Management System (/MS)* The Information Management System ıs the user Interface for EOSDIS The IMS at each DAAC will give users access to all data throughout EOSDIS, as well as help in locating and ordering data, through convenient, easy user Interfaces for both novices and experts The IMS will use simple search critera such as instrument name, product name, time of collection, and spatial location, as well as cross-instrument and cross-disciplinary searches

SOURCE National Aeronautics and Space Administration , 1993

tractor on September 29, 1992; contract negotiations were completed March 30, 1993. The combined cost and fee has been set at $766 million for the contract period 1993 through 2003.[12]

- The **EOS Data and Operations System (EDOS)** will capture data from EOS spacecraft, provide systematic corrections processing, distribute preprocessed data to the DAACS, and archive these data. NASA has selected TRW for this contract; cost and fee are under negotiation.
- The **Independent Validation & Verification (IV&V)** contract will provide for the testing and verification of the performance and capabilities of all elements of the EOS ground system and their integration with the EOS flight system. This contract will also provide for "acceptance testing" of all EOSDIS contractor deliveries. NASA has selected Intermetrics Corp. for this contract, at a cost and fee of $64 million.
- NASA and the EOSDIS contractors will build EOSDIS in a series of versions. The first is Version O, which is providing interoperability among the 8 DAACs and connections with other Earth science data systems. NASA expects Version O to be fully operational as an integrated "virtual" system in July 1994 (box 3-3).

The strategy and schedule for delivering subsequent EOSDIS "versions" is undergoing major revision, because NASA and Hughes are reevaluating overal1 planning for EOSDIS .As a result of advice from the EOS Investigators Working Group in October 1993, and the National Academy of Sciences report of January 1994,[13] NASA and Hughes have

[12] NASA stated Hughes's proposed cost and fee for the program was $685 million in September 1992, while Hughes' and TRW's original bids were much lower and rejected as unrealistic estimates. Subcontractors to Hughes are : Electronic Data Systems (new technology evaluation), Loral AeroSys (flight operations Applied Research Corp. (algorithm tookits) the Center for Space and Advanced Technology (research), and NYMA (independent verification I iaison).

[13] National Research Council, Space Studies Board, *Panel to Review EOSDIS Plans, Final Report* (Washington, DC: National Academy Press, 1994).

BOX 3-3: EOSDIS Version 0

Version O is a working prototype of EOSDIS with some operational elements. However, Version O will not have all the functions, reliability, and performance of subsequent versions. Planning and preliminary design of Version O began in the summer of 1990, and development of Version O as an integrated system began in January 1991,

NASA has focused the bulk of Version O prototyping on achieving system interoperability among the DAACS, NASA's philosophy for Version O is to allow individual DAAC systems to develop at their own rate, and focus on providing interconnections among those systems, Version O is improving user access to the DAACS by providing an overall view of the data available from the various DAAC systems, establishing common systemwide services, e.g., user assistance and support, problem resolution, and request and tracking statistics, and providing a single point from which any user can search and order data from *any* archive.[1]

- NASA Climate Data System
- NASA Ocean Data System
- Cryospheric Data Management System
- Alaska Synthetic Aperture Radar Facility
- Global Land Information System
- NASA Pilot Land Data System
- NASA Crystal Dynamics Data Information System
- Trace Gas Dynamics Data Information System

NASA and the EOSDIS contractors plan to transfer knowledge and experience from Version O into subsequent Versions of EOSDIS However, Version O is not a true working prototype of EOSDIS in many respects. Version O is a relatwely small effort, compared to subsequent versions of EOSDIS, Thus, Version O projects cannot address some of the technically critical areas of EOSDIS, nor is the effort substantial enough to allow users to assess some important EOSDIS functions. Nonetheless, the Version O effort has already successfully achieved user involvement, Interoperablllty among heterogeneous and distributed data systems, and a cooperative development environment that enables NASA to use DAAC expertise and data system experience,

[1] Such interoperability is likely to be an expensive task, since computer and network systems are increasingly complex, usually requiring specialists to enable applications to operate properly

SOURCE Office of Technology Assessment, 1994, and Judy Feldman, "Building on Version O', EOSDIS Progress Review, Dec 13-14, 1993, Landover, MD.

Implemented a major shift in orientation, focusing on responsiveness to user needs and a more open, distributed, and evolutionary architecture. Many of the details and results of this reorientation are not yet clear, but the reaction from the Earth science community is generally positive.[14]

[14] The former plan was for Version 1 to provide a fully functioning science data processing segment of EOSDIS (processing, archiving, and distribution). This Version would be much more capable than Version O, appearing completely integrated to users. NASA and Hughes planned its initial release for 1995, with a fully operational system in 1997.

Version 2 would have provided for full EOSDIS data system capacity and flight operations for the EOS-AM spacecraft launch in 1998. It would have been followed by Version 3, which would have supported data collection and operation of other EOS flights, in 2001.

Instead, Hughes plans 8 releases of the ECS, through 4 overlapping release cycles, beginning with the first full releases in September 1996. Hughes expects this "dual track" approach to improve incorporation of operational feedback as well as feedback from incremental development activities. "Release Schedules", John Gainsborough, EOSDIS Progress Review, Dec. 13-14, 1993, Landover, MD.

Impact of EOS Restructuring

The overall EOS program has undergone major restructuring since its initial congressional approval in 1990, resulting in significant reductions in scope and capabilities. In the summer of 1991, the EOS External Engineering Review Committee, organized at the request of the Office of Management and Budget and the National Space Council, restructured the EOS program. The result was a smaller more focused program of about $11 billion through fiscal year 2000 (down from the previous $17 billion estimate). The External Engineering Review Committee focused on distributing EOS instruments, and the reduced EOS instrument requirements, onto a larger number of smaller spacecraft to provide increased budgetary and technical resilience. The Restructuring Committee did not examine EOSDIS, but the delay in deploying some EOS instruments allowed EOSDIS to be smaller than originally planned, with a reduced budget.[15] NASA stated it made this adjustment in EOSDIS without altering the basic architecture or the evolutionary design of EOSDIS.[16]

In the fall of 1992, the restructured EOS program was further reduced by an internal NASA review to fit within an $8 billion budget envelope through fiscal year 2000. Again, NASA reduced the overall EOSDIS budget roughly commensurate to the overall program reduction. NASA reduced the planned suite of data products available at launch of EOS-AM 1 from 600 to approximately 160 data products.[17] Other changes included deferring the migration of existing data sets into Version O in cases where the data are already available through an existing operational system, deletion of the HIRIS science computing facility, and a major reduction in program reserves.[18]

As a result of these changes, EOSDIS is smaller than originally envisioned and program resources are substantially reduced. However, resilience in meeting future challenges has also been reduced, although many goals regarding data delivery remain unchanged. EOSDIS remains complex and demanding, raising a number of technical and programmatic challenges, discussed below.

EOSDIS TECHNOLOGY CHALLENGES

EOSDIS will be only as capable as the information systems technology on which it relies. Overall data rates and volumes will be unprecedented. EOS instruments will require very precise calibration, and data will require extensive validation to be useful. The DAACS will need to reprocess previously acquired data periodically, to accommodate updated processing algorithms. EOSDIS will have many simultaneous users, many of whom will require access to interactive databases. Data analysis and visualization[19] will be highly sophisticated and complex. **Although EOSDIS faces these and other technical challenges, there appear to be no technical obstacles to an operational EOSDIS that NASA could not overcome with sufficient funding and infrastructure.** In

15 The EOSDIS estimated budget dropped from $3.900 billion to $2.141 billion, a change of 45% compared to an overall EOS program budget change of 31%. U.S. Congress, General Accounting Office, *NASA: Changes to the Scope, Schedule, and Estimated Cost of the Earth Observing System,* NSIAD-92-223 (Gaithersburg, MD: U.S. General Accounting Office, July 1992), p. 18.

16 N h a t i e *Earth Observing System,* Mar. 9, 1992, p. 13.

17 NASA Earth Science and Applications Division, "Earth Observing System (EOS) Status," briefing to OTA, Washington DC, Nov. 6, 1992.

18 National Aeronautics and Space Administration, "Adapting the Earth Observing System to the Projected $8 Billion Budget: Recommendations from the EOS Investigators," Washington DC, October 1992, p. 39. The House Science, Space, and Technology Committee Report to accompany the NASA FY 1994 and FY 1995 Authorization Bill, Report 103-123, notes EOSDIS reserves were cut by $550 million, or 60 percent, in the rescoping (p. 46).

19 Using computer-generated pictures to represent data instead of a list of numbers, and viewing time sequences to track temporal change.

other words, data processing and communication capacity and speeds have evolved to the point that a system approximating an EOSDIS in some respects, but with less capability, could conceivably be built with existing hardware. Developing a highly capable EOSDIS, however, will present a challenge to NASA.

Simply "keeping up with" rapidly advancing technology will bean important challenge. Rapid new technology insertion will be essential for the system to retain its value to researchers. If users have independent access to data processing systems significantly more advanced than EOSDIS, many researchers would eventually use EOSDIS simply to download data into their own computing systems. This would defeat one of the primary goals of EOSDIS—interoperability among a variety of researchers and disciplines. On the other hand, reliance on emerging technologies that are not field proven would threaten system operability if they were to fail in full-scale implementation.

~ Data Storage and Access Technology

The data storage systems for EOSDIS will be extremely demanding. The archives will last for at least the lifetime of the EOS satellites,[20] and will be interactive with users, in contrast to the more traditional (and simpler) view of archives as a repository in which to store data for occasional use.[21] In short, as a result of the high **demands for** data storage that will be placed on EOSDIS, performance and cost of storage media may need to be much improved over current technology.[22]

Data mass storage costs are falling rapidly, but they remain a major expense for a system like EOSDIS. Most industry experts expect no breakthroughs in the cost of data storage in the next several years, although the development of optical storage systems should continue to bring storage costs down. Perhaps what is more important is that storage and access performance is not improving as rapidly as storage capacity. Searching the vast Earth science datasets for specific features will require major improvements in the ability to access specific data sets.

The advanced mass storage systems attempted in recent years have experienced serious problems.[23] Since the EOSDIS program philosophy is to procure hardware as late as possible to take advantage of falling cost and improved performance, EOSDIS may rely on advanced mass storage systems that have not yet proven commercially reliable. Maintaining a flexible system development strategy to accommodate rapid technological change successfully will be important.

NASA has not decided how much and which kinds of EOSDIS data will be directly available online to users but expects to use a hierarchy of data storage, in which small, often-used data sets are rapidly accessible, and very large datasets that are rarely used are stored offline. EOSDIS offline archives will be vast, and standards and media considerations (tape, cartridge, optical storage technologies) for offline archives should not be overlooked. The offline media will need to be easily serviced, and have reliable backups. Technology advancements and funds to archive and service

20 NASA EOSDIS planning currently extends only through the 15-year EOS mission lifetime.

21 Several Earth science data systems are already *improving* on this model. For example, the EROS Data Center, operated by the U.S. Geological Survey and included in EOSDIS as a DAAC, operates a large active archive for Landsat and AVHRR data, featuring online search. Some smaller data sets at NASA centers, such as the Coastal Zone Color Scanner System and the NASA Climate Data System, also allow these capabilities.

22 The EOSDIS project management team at Goddard Space Flight Center has estimated that a roughly thousand-fold increase in NASA Earth science data volume will occur during the 1990s. National Aeronautics and Space Administration, "Presentation to the EOSDIS Team on Ground Infrastructure Interfaces, Formats, and Directions," paper presented at IPD/EOSDIS seminar, Mar. 12, 1992.

23 U.S. Congress, General Accounting Office, *Space Data: NASA's Future Data Volumes Create Formidable Challenges,* IMTEC-91-24 (Gaithersburg, MD: General Accounting Office, April 1991).

BOX 3-4: EOSDIS Networks

The EOSDIS network system consists of four networks:

1, Ecom (dedicated network providing real-time, high reliability and secure communications between ground and spacecraft)

2 ECS Internal Network (dedicated EOSDIS network providing communication among EOS Principal Investigators, DAACS, and the External Network)

3. External Network (NASA Science Internet) (shared network providing communications among EOSDIS and users who are not Principal Investigators, including CIESIN and the Affiliated Data Centers)

4, Version O Network (dedicated network for prototyping)

Although more demanding than previous Earth remote sensing satellite communication systems, NASA does not expect serious difficulty implementing the Ecom network. NASA expects more challenges Implementing the EOSDIS Internal Network, but the "External Network, " which will essentially connect EOSDIS with the outside world beyond NASA, will offer the most difficult technical obstacles

The External Network, using the services of NASA's Science Internet (coordinated by Ames Research Center), supports several protocols and is interoperable with the NSF Internet, This network con-.sIsts of T1 transmission lines (1.5 million bits per second (Mbps)) connecting 27 regional networks and over 100 lower capacity circuits to research sites (although the tail circuits have much lower bandwidths) The NASA Science Internet reaches approximately 2500 end users. In 1994, NASA plans to upgrade the NASA Science Internet to T3 technology (45 Mbps). Eventually, NASA expects half of the NASA Science Internet (NSI) users to be EOSDIS users,

NASA planned to employ the UNIX operating system, HDF format, and the communications-related standards of the Consultive Committee on Standard Data Services where appropriate. However, a more "open" system is now planned,

SOURCE National Aeronautics and Space Adminiarration, 1994

the tremendous amounts of EOSDIS data, often overlooked in the past, maybe critical to EOSDIS long-term success. Recent surveys show substantial increases in estimates of EOSDIS processing and storage/distribution requirements.[24]

Data Communication

Networks perform the crucial tasks of linking researchers to EOSDIS and integrating the EOSDIS user community through cooperative research (box 3-4). The mode of delivery of data to EOSDIS researchers and the uses of the system will be determined by network capacity. Fortunately, network performance continues to increase as system costs decrease.

EOSDIS users are likely to request increasingly greater online access to increasing volumes of data. NASA has designed EOSDIS to deliver large data sets through EOSDIS networks, in contrast to routine data delivery through physical media (tapes, CD-ROMs, optical disks, etc.).[25] Current input/output and networking technologies cannot support this increased on-line data demand, nor the expected data rates required for

24"CostiPerfonmmce," Joe Guzek, EOSDIS Progress Review, Dec. 13-14, 1993, Landover, MD.

25 National Aeronautics and Space Administration, "Adapting the Earth Observing System to the Projected $8 Billion Budget: Recommendations from the EOS Investigators," Washington DC, October 1992, p. 9. On the other hand, if a researcher acquires a large datasct through a network, he or she still must store it on a physical medium.

browsing and visualization of EOSDIS data. Very high data rate workstation network interfaces also will be required for doing research using EOSDIS.[26]

External EOSDIS users will vary greatly in their sophistication; most will connect to EOSDIS through the future equivalent of today's personal computer and modem. NASA does not plan to provide a level of service to the larger global change community and other users equal to that available to NASA Investigators. However, NASA plans to provide the maximum of services to users who do not possess highly sophisticated workstations. This will place a considerable burden on EOSDIS design. **Because providing the maximum benefit from the public investment in EOSDIS may require broad access, Congress may wish to examine the potential of providing EOSDIS services to a broad community of users.** On the other hand, it is not feasible for EOSDIS to provide full service capabilities to casual users.

The National Research and Education Network (NREN) is currently of great value to the EOSDIS program in distributing data widely. However, NREN must be an operational system to be of use for EOSDIS; development of NREN is in the early stages, and the question of its status as an operational system has not been decided. It is also undecided whether NREN will be free for researchers, or if the system will require tariffs similar to the national telephone system (the EOSDIS program has not budgeted the funds to pay for NREN service for Earth science researchers). Finally, access to NREN would need to be widely distributed, serving the broader academic community outside the networks operated by the National Science Foundation, the Department of Energy, and NASA. It is unclear whether NREN will achieve such extensive distribution. For these reasons, EOSDIS planners are avoiding dependence on NREN.

DATA MANAGEMENT TECHNOLOGY (ORGANIZATION AND ACCESS)

Advanced techniques for indexing data for storage and access must be developed by NASA, the EOSDIS contractors, and the computer industry. Current relational data management technology, developed for use in commercial applications, is improving significantly in performance, and is accommodating some new types of data. However, relational databases have difficulty accommodating searches of spatial data sets and many other data processing and display requirements of EOSDIS. Relational data management software is most appropriate for manipulating small records of highly similar text or numeric data. Earth science data records are enormous, temporal, highly varied, and contain many more dimensions (time, latitude, longitude, spectral value, etc.) than most data records. Current commercial relational database systems and data processing software cannot efficiently work with these diverse types of data (e.g., point, vector, raster, text) .27 Version O (box 3-3) focuses on satellite data, but EOSDIS must incorporate non-satellite datasets[28] and their special requirements, complicating data management.

The interdisciplinary nature of global change research requires the capability to view the same data in different ways. It also requires common, and broader, access of data among different disciplines. In order to give data maximum utility, the EOSDIS program may have to support basic re-

26 Note that access to EOSDIS capabilities will be quite limited without advanced equipment. For example, at present, using a typical 9,600 baud modem, a single typical browse image of approximately 1.4 Mbytes would require approximately an hour to transmit.

27 See ch. 2 for a discussion of different data types.

28 Data acquired for ground-based facilities and from aircraft provide essential calibration for satellite data, and provide much essential data impassible to acquire from orbit. See U.S. Congress, Office of Technology Assessment, OTA-l SC-538, *The Future of Remote Sensing from Space: Civilian Satellite Systems and Applications* (Washington, DC: U.S. Government Printing Office, July 1993), ch.5 and app. B.

search into data management software tailored to scientific needs.

Data Processing, Analysis, and Assimilation Technology Issues

Much data acquired by satellite instruments go unused as a result of the time needed to process them on conventional computers, particularly to compute images. The tasks of visualization and assimilation of EOSDIS data into climate models are critical steps in the transformation of data into inform at ion.

The algorithms used by scientists to transform digital remote sensing data into information will undergo revision as knowledge grows. **Because changes in processing algorithms could leave small errors larger than any change in the global environment, rapid reprocessing of years of older data must be possible to maintain a continuous record of comparable data for research use.** Given the high spectral and spatial resolution of EOS instruments, and massive data volumes, this will be a formidable, continuous task.[29] Updated algorithms, which can have more than a million lines of code, must be transferred from the scientists to the DAAC Product Generation System for execution. Transporting and integrating these complex algorithms to generate "bug-free" products is a serious technical challenge.

Effective visualization technology will be an important challenge for the program, requiring significant *advances* in data processing technologies (e. g.. researchers can be expected to eventually use virtual reality to enter into a dynamic model).[30]

EOSDIS will generate higher level data sets by assimilating applicable observations into global climate and other models, which then will generate new data sets based on these models. These data sets will need to be of much higher quality than those currently produced for numerical weather prediction, and will be much more complex, since they will require assimilation of many more types of data (including non-EOS spacecraft; measurements from the ground, ocean, and air; and non-U. S. data sources). The computational requirements to produce these datasets will eventual] y go far beyond current practice in Earth science. Mass storage, network bandwidth, and processing power of computers will need to be greatly expanded for use of EOS data in future global climate models.[31]

EOSDIS will produce assimilated data sets on computers distributed geographically, which in some cases will use different system architectures. Computer languages and other technologies to allow these high-level analyses on massively parallel computer architectures are not yet well-developed. Standards are only just emerging in distributed systems management.

Computer processing power and network performance are increasing rapidly, while costs are decreasing. However, based on experience with other spacecraft projects, scientists typically underestimate the computer resources required to process their algorithms. This is considered the top risk in the entire EOSDIS Product Generation System, according to a recent risk assessment.[32] Although EOSDIS processing requirements appear great now, they could very well become several times greater when the system is actually implemented. Box 3-5 provides a summary of technical challenges in EOSDIS.

EOSDIS PROGRAM CHALLENGES

Overcoming technical obstacles will be important to the success of EOSDIS, but managerial, insti-

[29] Ibid., p. 38.

[30] In this manner, the scientist can virtually "become" a particle flowing through the model.

[31] Major reanalysis programs are now beginning at several data centers, including Goddard.

[32] Computer Sciences Corp., *Earth Observing System Data and Information System (EOSDIS) Product Generation System (PGS) Risk Analysis and Mitigation Strategies* (February 1993), p. 47.

BOX 3-5: Summary of Technical Challenges in EOSDIS

Data Storage

Demands on storage media performance and reliability Will be tremendous m EOSDIS, and data storage system throughput is not keeping pace with improvements processing or communications. Commercial data storage performance may not be successfully adaptable to EOSDIS needs Data storage currently appears to be a "weak link" m EOSDIS.

Data Communications

Demand for online access to larger amounts of data IS increasing, as is the numbers of users, user sosphistication, interagency and inteernational cooperation, data system distribution, and scientific cooperate work through networks. EOSDIS will not succeed If bandwidth and access are limited

Data Management

Effectively searching for data in EOSDIS could be difficult as a result of the quantity and variety of data in the systems Efficiently classifying vast amounts and varieties of data will be challenging, requiring new models of data management that are not yet well developed.

Data Processing, Analysis, and Assimilation

Processing demands will be much greater in EOSDIS than any previous system Software for use on parallel and distributed systems IS difficult to write, and visualization technology is not well developed

SOURCE Off Ice of Technology Assessment, 1994

tutional, and cultural challenges may be even greater.

I The Role of EOSDIS in GCDIS

The Global Change Data and Information System (GCDIS) is meant to allow routine access to all U.S. global change data (box 3-6). Some have called for a stronger NASA role in GCDIS. The National Research Council's Panel to Review EOSDIS Plans, in its April 1992 *Interim Reporf* and its September 1992 letter report, expresses concerns that EOSDIS may become a program ● 'oriented solely to EOS," rather than an integral part of the GCDIS.[33] The NRC Panel believes NASA has the responsibility for "establishing firm and specific plans and budgets for the development and operation of the GCDIS, in conjunction with other agencies''[34]. Thus, the panel desires a national directive to give NASA the lead agency role in the GCDIS, thereby transforming EOSDIS into a prototype for the GCDIS. The House Committee on Science, Space, and Technology largely agrees:

> The National Aeronautics and Space Administration, in coordination with other agencies that belong to the Committee on Earth and Environmental Sciences, shall establish the requirements and architecture for, design, and develop a Global Change Data and Information System that shall serve as the system to process, archive, and distribute data generated by the Global Change Research Program.[35]

[33] "National Research Council Panel to Review EOSDIS Plans: Interim Report," April 1992! p. 11.

[34] [bid., p, 2.

[35] U S.Congress, House Committee on Science, Space, and Technology, *National Aeronautics and Space Administration Authorization Act, Fiscal Years /994 and /995,* H.R. 2200, June 1993, Sec. 109.

BOX 3-6: Global Change Data and Information System

The GCDIS, as conceived by the Interagency Working Group on Data Management for Global Change (IWGDMGC), would provide a single data system for the various federal agencies involved in global change research GCDIS would use a combination of indivdual agency data system assets and a shared Infrastructure to become the primary mechanism for the exchange of data and information among USGCRP participants. GCDIS would include processes for Identifying and generating key inter-agency global change data sets, coordination of data submission procedures for GCDIS centers, standard methods for decribing and documenting data, a common set of archive responsibilities, and uniform order validation, tracking, and billing among agencies Proponents expect GCDIS WIII make data search and access among the various agency data sets much simpler and more effective

Interagency cooperation in GCDIS is currently a collection of extensive, but voluntary, indivdual agency efforts coordinated under the CEES IWGDMGC, GCDIS commonality and Interoperabilty would be made by the agencies in concert. The GCDIS does not have a separately funded budget, but resource requirements for focused program activities are included in USGCRP planning. For GCDIS to be successful, the effort will need to avoid becoming merely a "collector" of individual agency data and Information system plans Ensuring interoperabillty among the data systems of the USGCRP agencies, agreeing on standards for data among agencies so that researchers can easily exchange data, and maintaing high levels of data service among the several agencies will be the most difficult and important issues for GCDIS to resolve

SOURCE Office of Technology Assessment, 1994

NASA has agreed to seek funding to develop the techniques to allow interoperability among agency systems, thus "enabling and not precluding" extension of EOSDIS. All agencies, however, would require substantial additional funding for GCDIS to be implemented as envisioned.

The objectives of the EOSDIS program are already challenging, and NASA's responsibilities for GCDIS are an additional complexity in the program. **However, because NASA is already performing many of the necessary tasks for GCDIS in its EOSDIS program, giving the agency responsibility for GCDIS would be a more efficient use of public funds than assigning GCDIS to another agency. Attempting to add GCDIS requirements to EOSDIS after the latter is built would prove far more costly than planning for them as it is developed.**

Alternative Definitions for DAACS?

NOAA data will be critical for global change research. NOAA is already responsible for collecting and distributing operational and research data for monitoring and predicting the behavior of the atmosphere and oceans. NOAA data centers contain the majority of U.S. Earth remote sensing and in situ environmental data, and NOAA makes these data continuously available for its operational data systems. The National Research Council panel convened to review EOSDIS plans recommended including NOAA data centers as full DAACS (they are currently "Affiliated Data Centers'').[36] However, NOAA officials believe the cost of setting up DAAC, as currently defined, is more than NOAA can afford. Making NOAA data centers, as well as other essential sources of Earth science data, interoperable with EOSDIS should

[36] National Research Council, Space Studies Board, *Panel to Review EOSDIS Plans: Final Report*, op.cit.

be a priority, whether or not they are considered DAACs.[37] The development of alternative definitions of DAACS that prevent disruption of quality service, yet give good data access at minimum added cost, seems essential.

Socioeconomic Data in EOSDIS

For EOSDIS to be effective in meeting the long-term goals of the USGCRP NASA must transform Earth remote sensing data into information useful for nonspecialists (e.g., policy-makers, social scientists, resource managers, etc.). The system should also make potential users aware of, and able to use, available data and information. While EOSDIS and other Earth science data systems have been designed primarily to facilitate physical science-based global change research, the Consortium for International Earth Science Information Network (CIESIN) was founded in 1989 to assist a broader community of users of global change information, with a focus on integrating Earth remote sensing and other global change data with social science data.

CIESIN defines itself as an "international, non-profit consortium of academic, governmental, public, and private organizations that share a mutual goal of understanding global change."[38] Because CIESIN is not housed within any goverment agency, the organization can be more flexible than an agency and can maintain greater institutional neutrality.[39] This flexibility enables CIESIN to work closely with the several government agencies concerned with global environmental change, as well as academia, private companies, and other nongovernmental organizations, encountering fewer bureaucratic impediments.

In its first few years, CIESIN activities focused on assessing the needs and capabilities of users and providers of global change information. While assessment activities continue, CIESIN has begun to design systems to meet those needs. CIESIN is providing "tools and expertise" for data management, statistical analysis and modeling, visualization and imaging, and communications and collaboration.[w]

Although CIESIN intends to produce some new socioeconomic data, and integrate a variety of data from other sources, CIESIN'S strongest role could be as an access point to data and information from diverse sources worldwide. The organization would serve as a global "information cooperative," enabling interdisciplinary links between the natural and human sciences in global environmental change research.

CIESIN'S Socioeconomic Data and Applications Center (SEDAC) is one of the nine data centers in EOSDIS. As the data center responsible for providing access and distribution of interdisciplinary science data sets relating to the human dimensions of global change, SEDAC will make

37 Becoming a DAAC may not mean improved data services. For example, officials of the important CDIAC archive (trace gas and climate data), one of the three main climate change archives at Oak Ridge, do not want it to become a DAAC, and DOE does not want it to be part of the Oak Ridge DAAC. The threat is a fundamental change in operations and possible compromise of current good service. ARM atmospheric radiation data also will not be included in the Oak Ridge DAAC.

38 The founding organizations of CIESIN were the Environmental Research Institute of Michigan, Michigan State University, Saginaw Valley State University, and the University of Michigan. New York's Polytechnic University, Utah State University, and the University of Maryland at College Park were later included in Cl ESIN. CIESIN also works closely with the University of California at Santa Barbara.

39 Although CIESIN receives a majority of its funding from NASA (6I.7 percent in fiscal year 1993), CIESIN also receives substantial project funding from DOD (1 7.2 percent), EPA (1 6.3 percent), USDA (3.1 percent), and OSTP (1.7 percent) (fiscal year 1993 figures). Robert Coullahan, Director Washington Operations, CIESIN, personal communication, 1993.

40 CIESIN is also involved in many other projects outside the EOSDIS SEDAC, including software applications, data cataloging, data policy studies, the Global Change Research Information Office (GCRIO) supporting international data exchange, partnerships with federal agencies including the EPA, DOD, and USDA, and international data networks. CIESIN also serves as a training and education center for a diverse audience of current and potential users, teaching users about its application technologies and information products through summer institutes and scientific fellowships.

BOX 3-7: Data Categories in the Socioeconomic Data and Applications Center

The Socioeconomic Data and Applications Center will provide eight general categories of data to its users. The National Research Council has repeatedly identified the first four categories as the highest priorities of the U S research program in the human dimensions of global environmental change; they also are the expicit categories of emphasis for data collection and model development in the USGCRP, The fifth category serves the economics element defined in the FY93 USGCRP, The final three categories serve dlsclpllne-specific studies in the social and health sciences that relate to the human causes and human effects of global environmental change.

1. *Land Use and Land* Cover—land cover describes the land surface in generalized categories, whale land use describes the driving forces behind land cover
2 /ndustrla/Metabolism—the mass flows for key industrial materials, waste emissions, energy, and technical forces that drive the evolution of industrial processes,
3 *Agricultural* Metabolisml—the effects of agriculture and changing agricultural practices.
4. *Population* Dynamucs—demographic data on population and attributes.
5. Economic Activity
6 *Human Attitudes, Preferences,* and behavior-the personal motivations, and thelr sources, among individuals.
7. Social and Political *Structures and* Institions—the organization of human groups and the Influences of such organization on global environmental change
8 *Human and Environmental Health—effects* of global environmental change on the health of humans and the broader environment

SOURCE Consortium for International Earth Science Information Network, 1993

physical science data available for use by social scientists, and vice versa. It is the task of CIESIN and SEDAC to make EOSDIS data easily available to the estimated 100,000 to 200,000 users who are not physical scientists.

SEDAC will also serve as the designated Data and Information System for the Human Dimensions of Global Environmental Change Program of the International Social Sciences Council. In this role, SEDAC will provide international access for social scientists, and other international users, to all of NASA's Mission to Planet Earth data, as well as CIESIN'S socioeconomic data.

Data that serve the more particular research interests of social scientists are already collected and archived by existing data centers, and SEDAC will serve as an information network linking these data centers. Unlike other data centers in EOS DIS, SEDAC will not be a massive data archive, since most archiving of pertinent social science data is already done elsewhere. SEDAC will, however, archive some unique CIESIN-produced data sets. Box 3-7 describes the categories of data to be collected at SEDAC.

CIESIN ISSUES

CIESIN has been vigorously debated by policymakers. CIESIN detractors criticize the program for high costs of facilities, too many programs with insufficient focus, excessive spending for lobbying efforts in Washington DC, inappropriate allocations of funding, and a lack of peer review for funding.[41] CIESIN supporters refute these

[41] As a congressionally initiated program, CIESIN funding was not included in NASA budgets until the FY 1994 request. Instead, funding for CIESIN was inserted each year through congressional "earmarking" in the appropriations process.

claims or believe they have already been resolved, and maintain CIESIN is necessary to redress the lack of priority given to human dynamics research in global change in the USGCRP and EOSDIS .[42]

Congress must decide whether CIESIN funding is justified in comparison to alternative uses of the funding, whether CIESIN is indeed necessary to the success of the USGCRP, and if CIESIN is using resources appropriately. NASA plans to depend heavily on CIESIN for developing the use of global change data beyond the scientific community. **A USGCRP without CIESIN is possible; yet, many of the functions now provided by CIESIN would still need to be supplied by other organizations.**[43] On the other hand, some critics maintain that many of CIESIN'S activities duplicate services provided by other agencies, and will not provide them as effectively as agencies that have already provided such services for decades.

Use of Outside Expertise in EOSDIS

Most observers agree EOSDIS would benefit from increased involvement by data centers outside NASA and technologies developed outside the EOSDIS program.[44] The adaptation of superior technologies or methods used successfully in other systems could greatly enhance overall system capability. For example, recent demonstrations of Version O have elicited criticisms that the system is too narrowly focused. Some data experts argue EOSDIS is being developed as a system for "satellite researchers," while the needs of in situ researchers are not met. Increased consultation with experts at other agencies presumably would result in a more versatile system design.

For example, NOAA holds a majority of all current data related to global change, and pursues some applications likely to be required by EOSDIS. The National Center for Atmospheric Research, with amass storage system of over 36 terabytes (noted by the NRC for its effectiveness), makes extensive use of supercomputers and large data sets to model environmental phenomena such as global warming and depletion of the ozone layer. The University of Wisconsin, with an archive of about 130 terabytes, is the largest archive in the atmospheric sciences. The University Corporation for Atmospheric Research, sponsored by NSF, has developed the nationwide, distributed, real-time Unidata system to facilitate accessing, organizing, storing, analyzing, and displaying Earth science data on-line for educational uses. DOD and the intelligence community have invested heavily in software to transform remotely sensed data into information for national security purposes. In response to criticisms, officials at NASA and Hughes have promised to increase their efforts to examine non-NASA data systems.[45]

---

42 A recent audit from NASA's Office Of the Inspector General (IG) criticizes CIESIN funding and management. The IG recommends reducing NASA funding to space-based data support only. CIESIN supporters contend NASA is obstructing the will of Congress. *Space* News, June 20-26, 1994, p. 1.

43 See U.S. Congress, Office of Technology Assessment, OTA-BP-ISC- 122, *Global Change Research and NASA's Earth Observing System* (Washington, D. C.: U.S. Government Printing Office, November 1993) for a discussion of the need for policy-related data and assessment of the effects of global change.

44 National Research Council, The US *Global Change Research Program: An Analysis of the FY1991 Plans* Washington, DC: National Academy Press, 1990), p. 76. Also, the 1991 EOS Engineering Review Committee, which was mainly preoccupied with restructuring the EOS space hardware implementation strategy, expressed concerns that "EOSDIS makes no provision for bringing non-NASA Global Change Research projects or other investigative teams involved in global change research into the system." Earlier in the program GAO reports also criticized EOSDIS planning for insufficient use of existing database expertise at other federal agencies. "In designing and building its Version O prototypes, NASA has not taken full advantage of experience available at Earth science data and research centers other than the designated DAACs...Previous expert panels, including an internal NASA committee as well as the National Research Council, have noted the value of this experience base and urged NASA to make use of it." U.S. Congress, General Accounting Office, *Earth Observing System: NASA's EOSDIS Development Approach Is* Risky, IMTEC-92-24 (Gaithersburg, MD: General Accounting Office, February 1992), p. 21.

45 "EOSDIS Progress Review: Introduction," D. Butler, J. Dalton, EOSDIS Progress Review, Dec. 13-141993, Landover, MD.

The NRC Panel to Review EOSDIS Plans advocated a much stronger computer science research program for EOSDIS. The U.S. computer science community, Goddard's own in-house computer science experts, and experts at NASA's Ames research center have apparently had very limited input into EOSDIS implementation and operations decisions.[46] ability of EOSDIS ‘o exploit rapid advances of technology may depend on the consistent involvement of computer scientists both within and outside of NASA. NASA has recently devoted some of the EOSDIS budget to computer science and data handling technology development (approximately $20 million over the next few years), and is now soliciting proposals for advanced computer technology work.[47]

Version O and Pathfinder

NASA, NOAA, and USGS have initiated development of "Pathfinder" data sets in EOSDIS Version O to increase the amount of data available to Earth science researchers in the near term. Pathfinder datasets are large data sets collected over a number of years by NOAA environmental operational satellites, DOD DMSP satellites, and Landsat.[48] They are potentially useful to researchers because they span enough years to allow detection of ecological and climate trends. However, Pathfinder datasets require careful reprocessing, since they have been collected from multiple instruments of varying calibration standards mounted on many satellites. Pathfinder datasets include:

- Advanced Very High Resolution Radiometer (AVHRR) data sets held by NOAA,
- TIROS Operational Vertical Sounder (TOVS) data,
- GOES data (figure 3-6) held under NOAA contract by the University of Wisconsin,
- Special Sensor Microwave/Imagery (SSM/1) data acquired by NOAA from the Department of Defense,
- Scanning Multichannel Microwave Radiometer (SMMR) data recorded from the Nimbus-7 satellite, and
- Landsat data in the USGS archive at the EROS Data Center.[49]

NASA and Hughes have recently expanded the influence to be derived from Pathfinder data sets and Version O experience. NASA and Hughes appear to be moving toward using Version O as a testbed for further EOSDIS development, instead of replacing Version O with a different system for the EOSDIS EOS Core System ECS (box 3-4). NASA and Hughes plan to reuse the incremental development process, small development teams, "tirekickers,"[50] and other experience gained in Version O development and integration in subsequent versions of EOSDIS.

The work of NASA, NOAA, and the Department of Interior in developing the Pathfinder data sets is lauded by the scientific community. Congress may wish to encourage NASA to accelerate the Pathfinder activity to enhance the near-term benefits of EOSDIS. This action would also pro-

---

46 "Likewise, the nation's computer science community currently has very limited involvement in the Project, despite the fact that EOSDIS, to be successful, must implement the latest advances in scientific data management technology and, in some cases, stimulate the development of new technologies." *National Research Council Panel to Review EOSDIS Plans Interim Report,* April 1992, p. 16. The EOSDIS Advisory Panel also noted in October, 1993, "Experts in computing technology, with credentials comparable to those of the most prominent EOS investigators, have not had the opportunity to contribute to the architecture and design of EOSDIS."

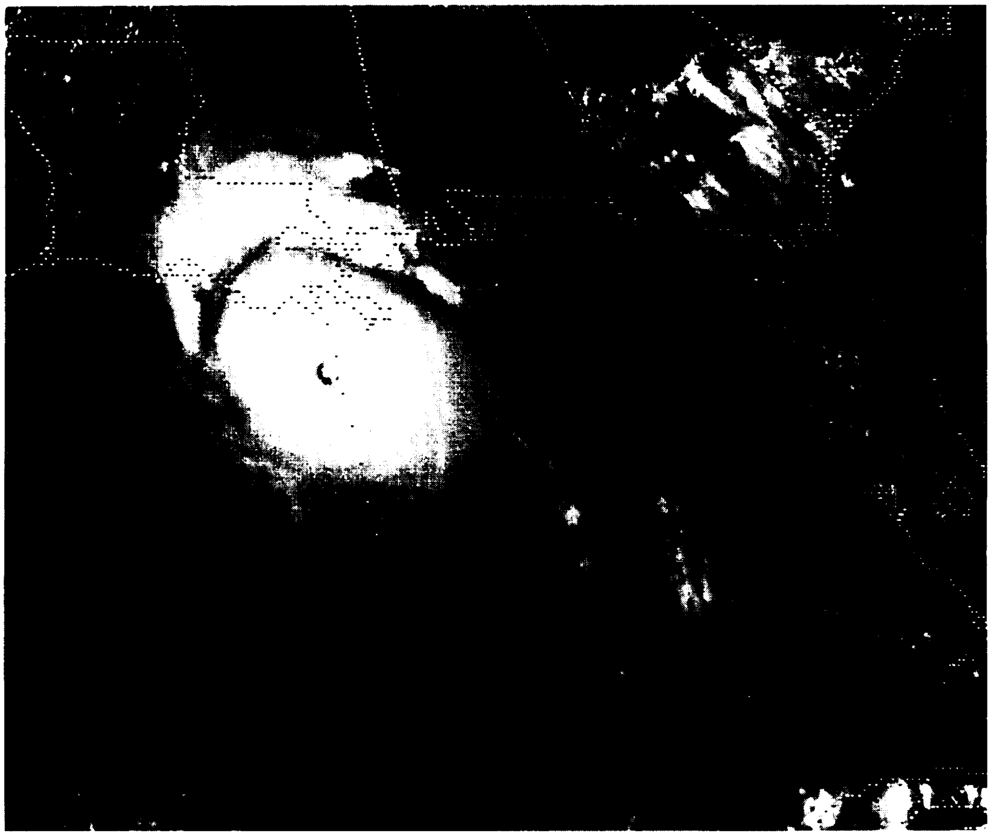
47 Researchers in academic computer science departments generally work with fairly small-scale interactive systems, and thus have little experience with large data flows—with some exceptions. Most experience with handling large data sets still resides in NASA projects, some science teams within NASA, and other agencies.

48 See U.S. Congress, Office Of Technology Assessment, *The Future of Remote Sensing from Space, OTA-* ISC-558 (Washington DC: U.S. Government Printing Office, July 1993) for a discussion of these programs.

49 Roughly 80 to 90 percent Of Version O data are from NOAA satellites. Pathfinder could be considered art exchange of technology for data between NOAA and NASA.

50 Interdepartmental engineering experts charged with testing overall system capability.

FIGURE 3-6: GOES–7 Image of Hurricane Hugo, Sept. 20, 1989



GOES images like this are effective in tracking severe storms in real time Historical data of the storms' changing form and track are useful in Improving scientists' understanding of the formation and evolution of severe weather patterns

SOURCE National Oceanic and Atmospheric Administration, 1994

vide more experience in providing Earth science data to the broad research community before NASA and Hughes implement later versions of EOSDIS.

Requirements-Driven Approach vs. R&D Experimental Approach

NASA is confident that an operational approach to the EOS Core System, integrating available commercial hardware and software to EOSDIS needs coupled with limited software development, will be sufficient to bring about dramatic improvements in the ability to use Earth science data. NASA and Hughes have designed the system to meet minimum standards of performance in all areas, an approach that decreases risk and is appropriate for the design and execution of an operational data system. However, this approach to EOSDIS will not push the state of the art in technology.

GAO and the NRC have criticized EOSDIS plans for insufficient attention to advanced technology development,[51] expressing concern that the contractor's near-term requirement to develop an operational system could detract from a thorough prototyping program to support the *long-term* needs of global change researchers.[52] Efficiently working with large, complex, and heterogeneous global change data sets may require special advanced technology. Much of this technology will not be available commercially, if scientific research is not considered a sufficient market.[53]

NASA does not usually sponsor the development of new technologies required for a flight program through the flight program budget itself, but rather uses other programs specifically established to sponsor flight and ground systems R&D. NASA previously intended to sponsor EOSDIS-related computer science research and technology through its computer/data systems R&D programs.[54] In response to external pressures, however, NASA has taken the unusual step of setting aside direct EOSDIS project funds to sponsor computer science research and advanced data systems technology development for EOSDIS. NASA is soliciting proposals, through a Headquarters Research Announcement, for technology development or adaptation for EOSDIS, and funding will be used for research and development at DAACS, Earth science organizations, and university computer science departments. Unfortunately, these steps may reduce the overall budget available for implementation and operations.

Congress has in the past had the opportunist y to direct NASA to strengthen the advanced technol-

---

51 As early as 1990 the NRC noted that: "According to NASA's development strategy, the EOSDIS Core System contractor will be responsible for initiating and conducting prototyping efforts after the contract is awarded and full-scale development begins. Prototyping is intended to be an ongoing aspect of the contractor's work. However, we believe that devolving responsibility for prototyping to the Core System contractor may make it difficult for NASA to ensure that the full range of critical technological risk areas are addressed in a timely fashion." The following technology areas were recommended for prototyping by the NRC:

1) data display & user interface,
2) browsing capability,
3) da(a formats & media,
4) accessibility of data and information,
5) cataloging,
6) search and query capabilities,
7) model and data interaction,
8) data structures,
9) data reduction algorithms, and
10) networking.

National Research Council, *The US Global Change Research Program: An Analysis of the FY1991 Plans* (Washington, DC: National Academy Press, 1990), p. 79.

52 The EOSDIS Advisory Panel noted in October 1993: "The system is being driven by detailed requirements, with little sense of the over-arching issues about information systems." The Panel also noted that Hughes' managers had "too little knowledge of the characteristics and computing styles of Earth scientists." GAO previously stated: "It is vital that NASA not allow the near-term operational requirements to prevent it from building a system that can ultimately provide a "next generation" of capabilities beyond what current Earth science data systems provide." U.S. Congress, General Accounting Office, *Earth Observing System: NASA's EOSDIS Development Approach 1s Risk>, 1* MTEC-92-24 (Gaithersburg, MD: General Accounting Office, February 1992), p. 33.

53 However several organizations outside NASA are pursuing technology development that would enhance EOSDIS capabilities. Sequoia 2000, a Digital Equipment Corp. project involving computer and Earth scientists at five campuses in the University of California, is pursuing a number of technology development efforts, including working with very large data sets using advanced query styles, searching for large objects, and techniques for working with diverse types of data. The Mitre Corp. is also exploring advanced query capabilities and object-oriented data management systems. Visualization techniques are being pursued at a number of research organizations, Including the IBM Watson Research Center, JPL, the Mitre Corp., and the Xerox Palo Alto Research Center.

54 Robert Price, Director, Mission to Planet Earth office, NASA/Goddard Space Flight Center, personal communication, January 1994.

ogy research component of the EOSDIS program. **Congress may yet wish to expand the higher risk technology development aspects of EOSDIS within NASA. This approach would have the potential to yield higher functionality of the system. Such a research effort would also have the potential to produce more generic technologies that might prove useful beyond meeting the operational requirements of EOSDIS.** Finally, an expanded technology development effort would enhance the oversight capability of NASA EOSDIS project staff.

On the other hand, successful and timely implementation of EOSDIS could be jeopardized if NASA and Hughes rely on custom-designed hardware and software, or new technologies without widespread commercial support or commitment. Most scientists currently desire basic online functionality with a small set of critical services presented in a way that matches how they work. Advanced graphics interfaces or similar '*extras" may be less important than simply having a system that is consistent throughout, works correctly every time, has a well-stocked archive of scientific data sets, performs quickly, and has a simple and inexpensive procedure to acquire data rapidly.

| Long-Term Archives

NASA has limited experience with operational Earth remote sensing data systems. Although the EOSDIS budget has fared no worse than other parts of the Mission to Planet Earth in recent program reductions, continuous operation and upgrading of an operational data system may prove a challenge for an agency historically oriented toward high-profile engineering hardware development and an emphasis on human spaceflight.

Since it is not known which data will prove useful in the future, and in order for scientists to understand the genesis of environmental changes they discover in the future, it is important to preserve all data.[55] Responsibility for long-term archiving of EOS data, however, has not been decided, and planning has barely begun for data maintenance after the 15-year life of EOS. NASA has promised to have all EOS data preserved for possible future use, and the pertinent Federal agencies are conducting negotiations concerning the means and mechanisms of long-term preservation of these data. The policy on archives will be an essential element in the long-term value of EOSDIS.

EOSDIS SCIENTIFIC INVOLVEMENT

NASA has sometimes conducted early mission planning and system development phases of space data systems without actively involving researchers and data consumers in the planning process. Insufficient scientific participation frequently has resulted in improperly implemented data systems and rejection of data systems by the scientific community. EOSDIS poses a special challenge, because its large scope could result in the domination of "system" concerns while science and service needs are overlooked.

Officials in the EOSDIS DAACS have already indicated that early and continuous involvement of the science community is the most important aspect of DAAC development. They also recognize that failure to involve scientists early in the planning can lead to a DAAC receiving little use by the scientific community.[56] To assist scientific input into EOSDIS development, NASA has

55 For example, when the AVHRR instrument was constructed, NOAA scientists were generally unaware of how important AVHRR data would be in following changes in vegetation. Now, NOAA distributes data on changes in vegetation throughout the year as a standard data product.

56 National Aeronautics and Space Administration, NASA Goddard Space Flight Center, Earth Science Data and information Systems Project, "EOSDIS Version O (V0): Lessons Learned," April 1993, p. 6. This document is tilled with references to the prime importance of a close working involvement between system development and the scientific community at the DAACs.

BOX 3-8: Science Advice in EOSDIS

The EOSDIS Advisory Panel of the NASA EOS Investigators" Working Group is the primary mechanism for obtaining user input in EOSDIS, with its 24 members drawn from the primarily academic community of 551 EOS Investigators Panel members have been on the procurement team for the ECS contract, and Panel members also communicate to Industry independently The EOSDIS Advisory Panel examines the "integrated picture' of EOSDIS, reviewing and assisting in planning, proposals, and system testing This group was Important in promoting the redirection of EOSDIS development toward a more open, evolutionary, and distributed system after the September 1993 Systems Requirements Review.

DAAC User Working Groups also provide essential guidance to EOSDIS. These groups provide "grass roots" input on science community requirements, data set needs and priorities, required functions and services, assistance in setting DAAC priorities, review and comments on DMC and EOSDIS system efforts, and assistance in the annual update of the EOSDIS science data plan Only half of the membership of these groups are EOS Investigators

The DMC User Services Working Group includes User Support Off Ice staff at the DMCS and NOM data centers, as well as members from the EOSDIS project at Goddard Space Flight Center This group is responsible for improving access to exiting data, developing common user services at all DMCS, and encouraging and gaining feedback from Version O use

Program scientists at NASA Headquarters take part in the MDIS Management Operations Working Group, which provides an overall review of EOSDIS program structure and performance, Insight into the larger outside Information systems world, and ties to the Earth Science and Applications Advvisory Subcommittee

Day-to-day scientific operational input and data product support IS provided by the DAAC project scientists on the staff of each DMC. At Goddard, the EOSDIS project has a project scientist on staff, as well as a scientist n the role of Science Data and External Interface Manager

EOSDIS at the system level and the DMCS attempt to be receptive to science advice through Individual comments and experiences from all users

SOURCE National Aeronautics and Space AdministrationmvstratoP 1993

constructed an extensive system for providing science advice (box 3-8).[57]

In spite of this system of science advice and assertions about the importance of close involvement with the science community, observers have complained that the role of Earth scientists is limited to advisory committees, while the DAAC scientists have no direct input into basic design. development, and operations decisions. Also, some assert EOSDIS planning is conducted under almost exclusive advice from NASA-affiliated scientists, to the exclusion of other users .5X

57 "One of the first activities of EOSDIS was to &fine methods for increasing participation by the research community in the definition, testing, and rc-design of the system...The success of EOSDIS hinges on the users' being empowered to shape it to their needs—needs which will evolve with progress in Earth science research and with experience gained in manipulating the data systems." National Aeronautics and Space Administration, *EOSDIS: EOS Data and Information System* (Washington, DC, 1992), pp. 7, 23.

58 "The predominant users of EOSDIS are expected to be the thousands of Earth scientists who are not affiliated with the EOS program. However, NASA's planning for the system thus far has relied largely on input from the relatively small number of researchers funded directly by NASA. NASA's guidelines and mechanisms for obtaining further user input in the future do not provide assurance that c/// segments of the user community will he adequately represented" LT. S. Congress, General Accounting Office, *Earth Observing System: Broader Involvement of the EOSDIS User Community Is ,$'ceded,* IMTEC-92-40 (Gaithersburg, MD: General Accounting office, May 1992), p. 1.

Having individual investigators actually perform data processing, validation, and intercomparisons in EOSDIS development would provide important feedback on EOSDIS operations. Emphasizing this approach would be more expensive, but has proven to be crucial in past data systems such as the WETNET at MSFC.[59] At the same time, EOSDIS could be more effective if EOSDIS officials look beyond the advice of current users and successfully anticipate the likely modes of computer interaction of future users of EOSDIS. Hughes has sent teams of scientists and engineers to many science user facilities to gain better insight into the needs of scientific users of EOSDIS.

Data Pricing: User Fees?

Whether data centers will actually collect money from the research community for the use of EOSDIS remains an open question. Current pricing policy for EOSDIS. and all U.S. global change data, ensures that data will cost no *more* than the "marginal cost of reproduction." At present, several EOSDIS DAACS distribute their data free to the scientific community.[60]

User fees have the advantage of providing the recovery of a portion of the data system operations costs, without seriously impeding data use if prices are sufficiently low. User fees also provide accountability, serving as a constraint against users ordering vast amounts of data simply because they are free. They also encourage user involvement in the data system.

However, the costs of a billing system can sometimes outweigh the benefits. In some cases, especially for online data distribution, establishing a system to monitor payments, checks, purchase orders, etc., may cost ***more*** than giving data away (especially for smaller data sets), and may reduce overall use of the data. Many researchers (e.g., unfunded researchers doing exploratory research, graduate students, educators, some foreign researchers) cannot raise sufficient funds to purchase large quantities of data, even at incremental costs.

Would EOSDIS be "flooded" with requests if data were free? EOSDIS could depend on commercial data distribution networks to limit data demand, similar to current use of the telephone network. Users would be required to pay for the time they spend on the network accessing and transferring data, but no billing system would be required at the DAACS. Network access would still need to be relatively inexpensive, however, for EOSDIS data to have broad distribution.[61] Another alternative would be to institute the use of research vouchers, allocating a limited number of data credits per researcher. Time and storage limitations alone might serve to discourage an individual undisciplined acquisition of large data sets.

This issue needs to be resolved relatively soon to facilitate the appropriate design of EOSDIS. It is essential that whatever is decided, the data policy be consistent within the government, and supported by appropriate funding.

Equipment Requirements for EOSDIS Users

Use of EOSDIS will be determined in large part by the equipment required to access the network. NASA will provide Science Computing Facilities (SCFS) for use by the 551 EOS primary investigators, ranging from personal workstations to super-

s9 National Aeronautics and Space Administration, NASA Goddard Space Flight Center, Earth Science Data and Information Systems Project, "EOSDIS Version O (V0): Lessons Learned," April 1993.

60 As noted in chapter 2, the EROS Data Center, with by far the largest remote sensing data archive of the DAACs, has a user fee system in operation. The Alaska SAR Facility also charges for data. Goddard Space Flight Center expects to have a charging mechanism in place by 1994. The other DAACs do not currently charge for data, and do not plan to unless so directed by NASA Headquarters.

61 It is highly likely that future commercial data networks costs will be charged in terms of bandwidth, not bits, making EOSDIS data transfer economical compared to other uses.

computers, for algorithm development. Besides direct NASA support, many other sources of research funding, both within and external to NASA, provide funding for computing equipment and communications to support the needs of the broader Earth science community. For this larger community, NASA expects the minimum SCF required for full access to EOSDIS services would be commensurate with what is currently affordable with a "typical research grant" (about $10,000 fiscal year 1993 dollars).[62] EOSDIS plans to make available software toolkits for use on these computers.[63] Furthermore, NASA and Hughes have promised to design EOSDIS to provide user access and as many services as possible.

However, under the present product generation system, scientists not directly involved with the program cannot easily contribute to the development and distribution of new algorithms. Version O is criticized by some observers for requiring highly specific equipment and software to be usable. A need for special equipment, software, or formats could be a major liability, considering the many researchers expected to use the interagency GCDIS.

Breadth of EOSDIS services, based on special equipment requirements, could be one of the first casualties of any future program difficulties or budget shortfalls. **If Congress desires to maintain the advantages of broad use of EOSDIS, it will need to monitor this aspect of the program to ensure that highly specialized and advanced computer hardware and software does not become a requirement for use of EOSDIS.**

Cost Savings in EOSDIS?

Congress has provided EOSDIS with funding at the level of agency requests thus far, with additional funding added for CIESIN. However, the continued availability of such resources is likely to be a central issue in the program in the future, given the strained budget context in NASA, the deficit problem in the U.S. government, and the planned increase of over 50 percent in EOSDIS funding for fiscal year 1995—from $188 million to about $285 million. Also, cost overruns have been the rule, rather than the exception, in large data systems. Figure 3-7 illustrates changes in total EOSDIS funding between 1991 and 2000.

Some observers claim the sizing of EOSDIS is unrealistic, since it was not rescoped to reflect the new reality of the smaller EOS program with smaller data flows. The EOS satellites are now about the size of the current UARS, and the launches are staggered. A new, reduced set of requirements for EOSDIS, some argue, could cut the costs of the program significantly.[64]

The EOS Investigator Working Group sounded a general caution concerning EOSDIS staffing costs:

> "From experience with other data systems. we caution that costs are only moderately sensitive directly to storage volumes and processing operations; they are more sensitive to the work

---

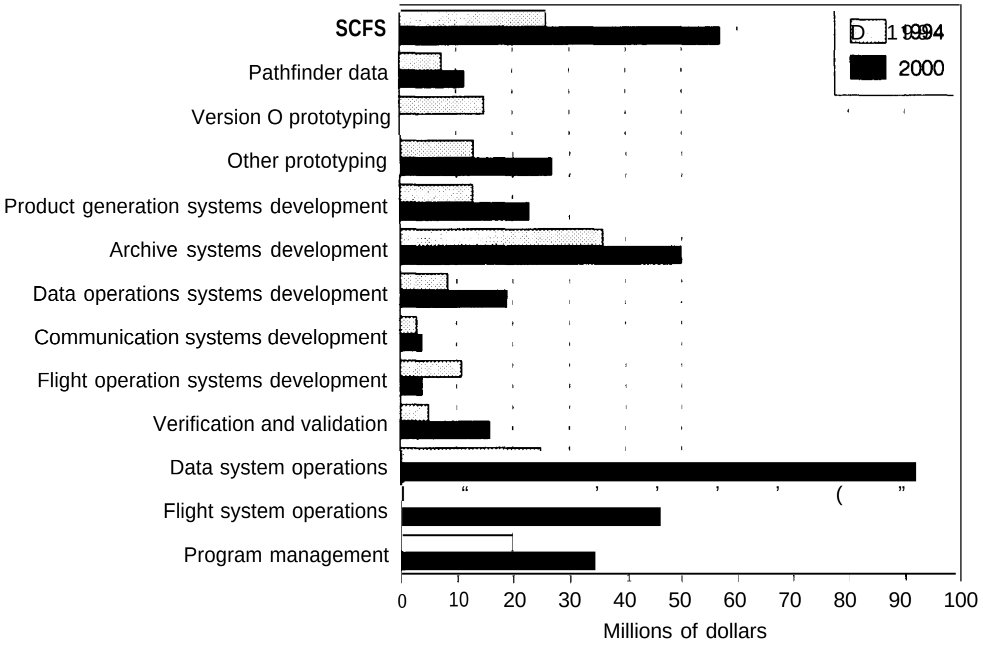
62 National Aeronautics and Space Administration, Modeling, Data, and Information Systems briefing to OTA, February 1993.

63 It is unclear what difficulties may be encountered from government restrictions on distribution of software. Hardware distribution difficulties were a major problem in the MSFC WETNET program. NASA ESDIS, "EOSDIS Version O: Lessons Learned," April 1993, p. 79.

64 An example is a recent ocean Wind scatterometer project at JPL, which was planned to cost $20 million. A new design review reduced the expected costs to $11 million, apparently with little negative impact.

Some observers argue that enough new information has been learned about the production of data products to warrant re-evaluation of the EOSDIS product generation system by DAACs and PIs to determine whether changes could bring more quality data products while saving resources. The present strategy is to plan each product in great detail. Top levels of NASA prepare and update detailed plans (to avoid delay of data products to "fine tune" the algorithms, or a focus on only a small subset of the data). As an alternative, however, more control could be passed back to the investigators and project groups. An agreed set of goals and delivery schedules for primary products would be required, but secondary products could be more creatively developed by investigators. This might give NASA many more good products, although there would be more failures as well. This approach would promise delivering more and better data products, lowering costs, and increasing productivity and satisfaction of the scientists. NASA and Hughes moved toward this conception at the EOSDIS Progress Review, December 1993.

FIGURE 3-7: Estimated EOSDIS Program Funding



SOURCE National Aeronautics and Space Adminstration Office of Technology Assessment, 1993

> force needed for system engineering, algorithm integration, etc."[65]

The EOSDIS Advisory Panel noted in October 1 993:

> "By far the greatest expense in EOSDIS is the sum of the salaries for maintenance and operations, and the currently proposed staffing levels seem high."

EOSDIS plans require excellent data delivery at all times of the day; some costs could be saved if this capability were reduced.[66] If researchers can wait up to a week for most data products, the dif-

65 National *Aeronautics* and Space *Administration, Adapting the Earth Observing System to the Projected $8 Billion Budget: Recommendations from (he EOS investigators,* October 1992, p. 39. The EOSDIS Advisory Panel noted in Oct., 1993, that"...the putative costs seem too sensitive to the floating-point operations needed to create the EOS standard products, when the constraints are more likely to be the population of users who can be served and the rate at which the system can deliver products to users,.."

66 If 8-hour shifts are substituted for 24-hour shifts, DAAC operations costs are estimated to be reduced by 8.2% to 17.9% (depending on "non prime time" operations levels, and assuming processing and electronic access/distribution of data during *'non prime time"). This is an estimated savings of $15.7 to $34.4 million through October 2002. "Cost/Performance," Joe Guzek, EOSDIS Progress Review, Dec. 13-14, 1993, Landover, MD.

ference in costs could also be expected to be substantial.

However, underestimation of required data system personnel would also be a serious problem. User services (establishing communications for users, training researchers and computer experts in the use of the system, solving communication and network problems for users, providing information on data product generation and delivery, special requests for data, etc.) usually have required *more* effort than initial estimates. In EOS, with a projected 100,000 to 200,000 users, the number of people dedicated to user services can be expected to be large.[67]

While many climate problems require global data for many years, there may be ways to acquire samples from the data stream rather than store the entire data set in order to reduce the volume sharply. These opportunities could make it easier to pursue the scientific search while reducing costs.

EOSDIS has ambitious plans for providing data online. A few EOSDIS data streams will be needed by NOAA for real-time weather forecasting, and NASA plans to deliver these data to NOAA rapidly. Some observers claim the rest of the EOSDIS data products do not need to be available in real-time, as currently planned. Short delays in data transmission offline might result in significant cost savings.

Observers have noted that to achieve good service at reduced cost, some competition is usually necessary. While this is difficult to achieve without some duplication, furthering efforts to provide choices to researchers might result in overall higher efficiency, better service, and lower cost. Placing the entire responsibility for meeting diverse user needs through a single, pre-planned data and information system could be very difficult. EOSDIS could become a very "brittle" system if EOSDIS were "monolithic," and the only means of communication between researchers and NASA officials. Examples of resilience - enhancing alternatives would include the use of the CIESIN network, direct broadcast of data, commercial high-volume/high-speed lines, and NREN.

As technology and economics change, the system must adapt to enable functions to migrate to where they can most economically be performed (e.g., the shift from centralized mainframe computers in the 1970s to today's distributed workstations). NASA and Hughes plan to isolate functions where technology change is most likely to occur, so these functions can easily be changed or replaced as technology matures.[68]

Some observers point out that many of the functions in EOSDIS might well be provided by the private sector. This view posits that it is inappropriate and inefficient for government to plan and build operational science networks in an era of rapidly expanding technical capability. For example, special funding for networks of very high bandwidth would be redundant if sufficient bandwidth becomes widely available and inexpensive commercially. As noted in chapter 2, computer processing speed, and storage capacity and access, largely funded by the commercial sector, have been increasing markedly in recent years.

As an alternative or supplement to EOSDIS, NASA could rely on direct-broadcast of data from satellites to ground stations at scientists' research institutions. Proponents of this approach claim costs increase dramatically when government performs computing tasks, noting many researchers already receive data directly over communications links. Reliance on this strategy would, however, hamper the fundamental goal of fostering scientific interdisciplinary research. Such a plan also might increase costs, not decrease them, since each user would require the ability to process raw data signals to final products, a costly process for many types of data. Using a few well-controlled facilities (DAACS) is advocated as a less expensive and more effective system.

---

67 "Members of most user communities will continue to want to talk to knowledgeable user service personnel via telephone-specially as the number of data products and their complexity increases...training is not a trivial matter, especially for a large number of data products with frequent changes." Pitt Theme, "Demographics", EOSDIS Progress Review, Dec. 13- I 4, 1993, Landover, MD.

68 "EOSDIS Progress Review: Introduction," D. Butler, J. Dalton, EOSDIS Progress Review, Dec. 13-141993, Landover, MD.

There are other alternatives. As noted in chapter 2, many government data distribution programs, including remote sensing data systems, have derived and distributed products with fairly modest costs. While these programs do not achieve the broader goals of EOSDIS, they do provide less expensive models of data distribution.

Commercial Relevance

NASA has planned and developed EOS as an operational scientific data system, relying on USGCRP goals and scientific and technical considerations for program planning and execution. NASA has not designed EOS to stimulate the Earth remote sensing market, nor as a "test-bed" for advanced Earth remote sensing technologies, nor to contribute to the national goal of "global competitiveness." The two original EOS instruments with the most potential commercial relevance, HIRIS and SAR, were deleted from the program in the 1992 restructuring. EOS data are generally low-resolution, and land observations have not been emphasized in the program, limiting the commercial value of EOS data.

Nonetheless, a strong potential may exist for commercial value of some EOS data. While current EOSDIS plans to make all data available almost immediately could destroy the commercial value of similar data from other sources, such as from the Sea Star Satellite (ch. 4), easy access to EOS data by the commercial sector could result in valuable enhancements ("value-added" products) that could satisfy various needs the government cannot meet.[69] NASA and Hughes are conducting studies of the potential commercial relevance of EOSDIS data.

DATA FORMATS/STANDARDS

The issue of data formats for remote sensing has been debated for several years, especially in NASA. Until recently, NASA and Hughes had supported only high density format (HDF) for EOSDIS data storage and retrieval. However, this implied anew way to ingest data that very few scientists have used. Many scientists might desire to continue to obtain small datasets in a simple binary or character format, such as ASCII.

Data formats should be easy to use. Formats also should not substantially increase data volume or slow down the processing of large datasets. Finally, formats must be capable of allowing data processing on primary workstations and PCs. The appropriateness of HDF by these criteria is debated in the science and computing communities. Designating a present format system as the standard for future EOS data would doubtless cause problems with using the data. Instead, NASA and Hughes plan to provide translators within EOSDIS so users can easily access data in different formats.

Is EOSDIS "Distributed" and "Evolutionary"?

At the first EOSDIS ECS system requirements review in September 1993, Earth scientists expressed concern that EOSDIS architecture appeared too centralized and inflexible to evolutionary change. The EOSDIS Advisory Panel of the EOS Investigators' Working Group, in October 1993 concluded the system was not a distributed system:

> Instead it is a system of geographically dispersed elements with tightly centralized management [a central architecture forced to reside at several geographic locations] . . . Essential elements of a distributed system-competition among elements, and different approaches, distributed responsibility, power, and resources—are missing .70

The Panel also stated EOSDIS was not an evolutionary system:

---

69 For example, Orbital Science Corp. is currently attempting to develop a commercial market for value-added enhancements to data that will be collected by its SeaWiFS ocean color sensor. See U.S. Congress, Office of Technology Assessment, *The Future of Remote Sensing from Space*, op. cit., footnote 2, ch. 7.

70 EOSDIS Advisory Panel, EOS Investigators' Working Group Payload Panel meeting, Oct. 4-6, 1993.

> Instead its developers focus tightly on the near future, use tools and standards that are already obsolete, view "technology insertion" as synonymous with evolution, and have little vision of the computing environment of this century and the early part of the next.[71]

The Panel noted the EOSDIS design had changed little since 1990, despite important technological achievements in the architecture of distributed information systems since then.

In response, Hughes is conducting studies of alternative ECS architectures, with study teams selected from top universities in computer and Earth science.[72] NASA is also funding the development of prototypes and discipline-specific functions, and encouraging increased involvement in ECS development and funding for added functions and services to meet the needs of specific science disciplines.[73] The EOSDIS ECS system is now being designed to accept alternative implementations at all levels, including new developments not created by NASA or Hughes, as well as test marketing new ideas, products, and methods.[74]

EOSDIS needs to maintain the flexibility to deal with different methods of data management among the DAACS, since different science communities will have different data management needs. The report of the NRC Panel to Review EOSDIS Plans Interim of September 1992 stated DAAC managers did not have well-defined authority or accountability in building EOSDIS, that DAACS were not sufficiently involved in EOSDIS implementation, and their primary role appeared to be simply to operate hardware and software at their sites after delivery.[75] According to the NRC Panel, 'The centralized management of the design and implementation of EOSDIS functions at each DAAC is not conducive to active user involvement and responsiveness to changing technology."[76]

Decentralization also has its risks, however. To build an integrated, interoperable system requires sufficient central authority to ensure interoperable system architecture and interfaces. As a project serving multiple agencies, EOSDIS requires smooth and efficient interpersonal communications, as well as computer communications, in a highly complex environment. Parochial interests need to be controlled to some degree. Completely autonomous DAACS, each with its distinctive system architecture, data formats, and so on, was one of the primary reasons for the development of EOSDIS. While insufficient input from DAAC management would endanger system responsiveness to scientists, excessive DAAC autonomy might endanger integration and interoperability of EOSDIS.

Instead of increasing the authority of DAACs as a means of dealing with centralization, NASA might arrange a system having a manager for each cluster of major products for related disciplines. This manager would make agreements on how to develop products that would be stored in the DAAC. Data distribution could be separated from product generation, with the DAACS and advisors having most control over distribution while science experts have control over product generation. This is similar to older NASA project management philosophy in which a single manager has control over the priorities and the level of ef-

---

7. Ibid.

72 D. Butler, J. Dalton, "EOSDIS progress Review: Intrxduction," EOSDIS Pn)gress Review, Dec. 13-141993, Larrdo\cr, MD.

73 Ibid.

74 Gal] McConaughY, ESDIS Project, "The Evolving Context of EOSDIS (Focus: Science Support)" EOSDIS Progress Review, Dec. 13-14, 1993, Landover, MD.

75 NASA has been criticized for allowing the DAAC managers little influence over the operation and maintenance of ECS as a whole, no financial control over the long-term strategy of the DA AC, and no responsibility to reallocate resources to maximize the provision of services.

76 A reliance on standard data products alone could be too rigid. For example, users would have difficulty in automatically combining data from different sensors, altering products to meet new scientific needs, or revising algorithms to meet various purposes. The DAACs currently have little control over the forms in which they receive their data, the management and evolution of the ECS, or budgets.

fort of each task, with less influence from NASA headquarters and more authority from those in the "field," along with freedom to cut across organizational boundaries to accomplish tasks.

The level of autonomy at each DAAC could have a significant impact on the success of EOSDOS. Congress may wish to consult with NASA management, DAAC management, and informed members of the global change research community, to monitor the appropriate level of centralization in EOSDIS management.

I The New Future of EOSDIS: "UserDIS" and the "Earth Science Web"

The UserDIS is a vision of the future information infrastructure in which there will be a multiplicity of data sources and information integrators available to scientists and other users of Earth science and global change data. EOSDIS would be one of the key providers of data services in this "Earth Science Web" of easily accessible pooled computing and data resources.[77]

A Hughes study of this issue found: '*There are many things which ECS could provide without leaving its mission envelope for GCDIS/UserDIS."7[8] In response to ideas from the EOSDIS Advisory Group and the NRC Panel to Review EOSDIS plans, NASA and Hughes have recently promised to design EOSDIS ECS as part of a larger environment from which users can freely find, invoke, and selectively combine services.[79] While focusing on Earth science data and its users, other uses would not be excluded by the new EOS - DIS design architecture. The distinction between user and provider would be elmininated, effective y using the computer resources and expertise in the distributed user community beyond EOSDIS. Responsibility, power, and resources would be dispersed throughout the Earth Science Web, with any provider having the ability to add a new idea to the Web. No restrictions would be placed on the number of providers, their locations, and the services and data they offer.[80] Beyond the DAACS, the UserDIS would accommodate autonomous provider sites dealing with researchers and research groups as individuals rather than relying on sponsoring "institutions''.[81]

If EOSDIS is to evolve toward UserDIS, as advocated by the NRC Panel to Review EOSDIS Plans, specific EOSDIS goals should be limited, relying instead on the entrepreneurial spirit of DAACS and other organizations. The Panel expects the cost of communication and switching to drop dramatically in the 1990s, meaning a variety of approaches to computing not previously envisioned would be made available by entrepreneurial companies and other organizations. The role of EOSDIS would be to remain open, not excluding the use of new developments or other users and uses of the system.

77 Gail McConaughy, "The Evolving Context of EOSDIS (Focus: Science Support)," op cit.

78 Mark Elkington, "GCDIS/UserDIS—Background and Issues," EOSDIS Progress Review, Dec. 13-14, 1993, Landover, MD.

79 Gail McConaughy, "The Evolving Context of EOSDIS (Focus: Science, Support)," op Cit.

80 Mark Elkington, "GCDIS/UserDIS—Background and Issues," op Cit.

81 Ibid.