United States General Accounting Office

GAO

Testimony

Before the Subcommittee on Space and Aeronautics,
Committee on Science, House of Representatives

For Release on Delivery
Expected at
9 a.m.
Thursday,
March 16, 1995

EARTH OBSERVING SYSTEM

Concentration on Near-term EOSDIS Development May Jeopardize Long-term Success

Statement of Jack L. Brock, Jr.
Director, Information Resource Management/
National Security and International Affairs
Accounting and Information Management Division

Mr. Chairman and Members of the Subcommittee:

We are pleased to be here to testify on the National Aeronautics and Space Administration's (NASA) development strategy for the Earth Observing System Data and Information System (EOSDIS). The Earth Observing System (EOS) is a comprehensive program to study global change by gathering and analyzing complete data about how the earth functions as a single, integrated system. Program costs through the year 2000 are estimated to run $7.25 billion. About a third of the cost for EOS will go to EOSDIS, whose success is imperative if the overall program is to succeed. Without the extensive processing, archiving, and distribution functions that EOSDIS will perform, the data collected by EOS' satellite-based instruments will be of little value.

As requested by this Committee, we recently reviewed the EOSDIS program to assess changes NASA has made since the National Research Council (NRC) issued a report on the EOSDIS program in January 1994. Today we would like to address issues and concerns as we see them at this point in our study.

On the basis of our review, we believe that NASA has taken positive steps to address the NRC's concerns. The NRC's primary recommendation was to redesign the system to be more open and distributed, with much more control and responsibility for data processing going to the scientists who will use the system. In response, NASA has revamped the technical architecture of EOSDIS to be more flexible and accommodating to the unique activities and

needs of individual data centers and researchers. Key members of the scientific community have endorsed these design changes.

Nevertheless, we have concerns about NASA's development strategy for EOSDIS and its approach to mitigating the significant risks inherent in this large, technically complex project. Today, I would like to discuss our concerns on two subjects related to EOSDIS development: (1) knowing who the expected users of EOSDIS will be and how they will use the system, and (2) finding the appropriate balance between near- and long-term system development needs. I would also like to discuss options that others have identified for improving EOSDIS and saving money. Those options deserve careful attention.

NASA's EOSDIS STRATEGY

Developing EOSDIS is a massive undertaking, both in terms of the system's scope as well as its cost. NASA currently estimates that it will spend over $2 billion to develop EOSDIS through fiscal year 2000. The sheer size of EOSDIS is staggering; its intended scope far exceeds that of any previous civilian data management system. Over its lifetime, the system could accumulate a mass of data equivalent to more than 1,000 times the amount of text stored in the Library of Congress.

NASA intends EOSDIS to be a comprehensive system that will bring together data from many sources to serve the needs of scientists performing integrated, interdisciplinary studies

of the earth. In addition to data from the EOS satellites, EOSDIS will include previously archived data; new measurements from other non-EOS spacecraft; various ground-, ocean-, and air-based measurements; and data-processing software developed by the scientific user community. As such, EOSDIS will be the one system responsible for archiving and distributing all NASA earth science data.

The goal of EOSDIS is to make this vast wealth of data easily accessible to as many as 10,000 earth scientists and 200,000 other users from government, academia, and public interest groups. Many of these users will bring little technical background to their interactions with EOSDIS. NASA's intent is to allow scientists and researchers from the many different disciplines to come together and collaborate on global change research in exciting new ways.

The bulk of EOSDIS development will be done under a single, comprehensive contract, known as the EOSDIS Core System contract. The Core System contract was awarded to Hughes Applied Information Systems in early 1993 and will cost NASA an estimated $930 million through 2003. Hughes will be responsible for building and integrating the major elements of EOSDIS, including hardware and software to be installed at eight data centers around the country. The system will be developed through a series of incremental hardware and software releases. A prototype called Version 0, not developed by Hughes, has been running since summer 1994. The Hughes-developed system will first become operational in

1997 to provide data archiving and distribution for the Tropical Rainfall Measurement Mission, and will begin processing and distributing data from the EOS satellites in 1998.

NASA HAS TAKEN POSITIVE STEPS TO ADDRESS SOME OF THE RECOMMENDATIONS OF THE NRC

The January 1994 NRC report on EOSDIS included a fairly broad range of findings and concerns about the planned development of EOSDIS. However, the Council's strongest criticism was leveled at the system's proposed architecture, which it thought was too rigid and centrally controlled. The NRC was concerned that researchers would not be able to go beyond receiving standard, predetermined data products, which would likely be inadequate for research topics that have not yet even been conceived. Furthermore, the Council found that although the system design included eight geographically distributed data centers, control was really centralized at Goddard Space Flight Center. For example, remote users had to go through a centralized office at Goddard in order to log onto the system. Also, only predetermined analyses were to be supported, with no flexibility to accommodate local development of supplemental products. As a result, scientists in the field would be severely constrained in trying to use EOSDIS to manage their own data processing at remote sites.

Fortunately, NASA and Hughes were already making changes to the EOSDIS design by the time NRC issued its report. Since the NRC report was released, NASA has adopted a new, distributed design. Although the heart of the system remains a network of computers at eight

data centers that will concentrate on storing data and generating standard products, the new design enables users to share their own research and data products. Users with whom we spoke were generally optimistic that the new EOSDIS design would better enable them to conduct their research.

NASA also deserves credit for strengthening the management of the program and filling vacant positions in the project office with highly qualified people. This is another area that the NRC had asked NASA to address.

Nevertheless, despite the improvements that have been made to the EOSDIS design, we have two major concerns about NASA's approach to developing the system:

- NASA still has not given adequate attention to EOSDIS users, their expectations, or their needs. NASA should take steps to better understand and incorporate user requirements before deciding on critical elements of the system.

- NASA's development strategy may be overly oriented towards large-scale system development work in the near term. Given the uncertainties regarding expected uses of the system and the level of functionality that can be provided, a more modest near-term effort may be warranted.

EXPECTED USERS AND THEIR NEEDS ARE NOT WELL UNDERSTOOD

The system's core users--the community of earth science researchers--are still not effectively involved in the development of EOSDIS. Furthermore, the needs of the large number of non earth science users that NASA expects are unknown. For maximum return on its investment, NASA needs to involve science users more directly in EOSDIS development and better define the broader user community.

For the near term, NASA will need to satisfy the 500 or so principal investigators that it is funding to carry out specific investigations using data from the EOS satellites. NASA's involvement of these investigators to date has been limited to the activities of advisory boards and committees.

The NRC recommended that NASA delegate greater authority, responsibility, and funding to the eight EOSDIS data centers to strengthen their position as representatives of the user community. The eight centers have historically had close relationships with earth science researchers, including the EOS principal investigators, who work in specific disciplines. Because these centers have been responsible for running software written by these researchers and providing analytical results and other automated support, they are closer than NASA project officials to understanding the diverse needs of the core group of researchers who will use EOSDIS in the near term. However, NASA has not yet acted on the NRC recommendation.

Much is still unknown about how earth science researchers will use EOS data. Two hundred and twenty-two standard data products have been designated to be processed and archived in EOSDIS. There has been considerable debate over which of these data products will actually be wanted. Several earth science researchers we contacted said they would prefer access to basic data sets, from which they would derive their own products. Another question is whether researchers will prefer to move EOS data to their own systems for analysis, or whether they will look to NASA to provide them with analytical tools within EOSDIS.

Beyond the near-term needs of the EOS principal investigators and other earth scientists, EOSDIS is intended to also serve a far wider audience. However, very little is known about these other users, which, according to NASA, could span such diverse groups as elementary school students, farmers, librarians, and policy makers. NASA has estimated that EOSDIS will serve as many as 200,000 non earth science users, many of whom will have no special knowledge of remote sensing data or access to sophisticated computers.

Over the years, the NRC has consistently recommended broad user involvement throughout development of earth science data systems to maximize the science return on space missions and improve the quality of data for users. Knowing who will use the systems and how they will use them must determine what systems actually get built. Since NASA does not know the true size of the expected user population of EOSDIS, which data sets will be most needed, or what types of analytical tools users will want, we believe the agency needs to be more cautious about establishing standard data products and other design specifics that could

become obsolete or counterproductive by the time the system is in widespread use. At the same time, NASA needs to make a better effort to define the system's expected users and involve them in the planning effort.

CONCENTRATION ON NEAR-TERM DEVELOPMENT MAY JEOPARDIZE LONG-TERM SUCCESS

NASA is currently investing heavily, through the Hughes contract, in near-term, full-scale development of EOSDIS. Hughes has been spending as much as $7.5 million each month on development work aimed at installing all new hardware and software at the eight EOSDIS data centers by 1997, in time to support the first of the EOS-related satellite missions. The existing Version 0 prototype of EOSDIS will temporarily operate in tandem with the new system, but will eventually be phased out.

Such emphasis on large-scale development in the near term may be unwise, for several reasons. As I already discussed, much is still unknown about who will use the system and how. This alone argues for a more conservative, incremental approach to system development. However, NASA faces an additional obstacle as well, namely that critical software technologies needed to ultimately make EOSDIS successful are still beyond the state of the art. Given this condition, greater emphasis on research and prototyping seems in order for the near term, with less emphasis on full-scale development.

The critical software technologies that will be needed are largely associated with handling the vast amounts of science data that EOSDIS is expected to contain. For example, current database search techniques, which were generally designed to meet the needs of transaction-based business applications, are inadequate to support the work of interdisciplinary earth scientists. A new kind of database management system is needed that will handle complex earth science images as nimbly as a business-oriented database system now handles payroll records. Advances are also needed in the way scientific data are organized and characterized within an information system. These technologies do not yet exist in operational form, although some research is underway. Much more work still needs to be done.

Aware of these obstacles, the scientific community advised NASA early in the program to adopt a "build a little, test a little" development strategy for EOSDIS rather than launching headlong into full-scale development. In 1992, we also recommended that NASA develop specific plans and identify resources for prototyping the full range of critical system elements. While NASA has officially subscribed to this "evolutionary" approach, its allocation of development resources nevertheless appears to be overly concentrated on near-term, full-scale development.

NASA is devoting relatively few resources towards research and prototyping activities. NASA-sponsored research projects focused on advanced software technologies for EOSDIS only began last year and are currently funded at a modest $4 million per year. Although

more money has gone to prototyping, NASA recently reduced these activities severely when faced with a significant budget cut.

Prototyping is an engineering technique in which partial, experimental versions of computer-based systems are rapidly and inexpensively built to validate requirements and test the feasibility of key functions before production. Taking an evolutionary approach means relying on a substantial ongoing level of prototyping activities to evaluate new concepts and implementations as they are invented.

However, following the rebaselining of the EOSDIS project in July 1994, NASA applied steep cuts to prototyping. Funding at the project level was cut from $69 million to $49.5 million, a 28 percent reduction. Funding for Hughes to conduct future prototyping work was also cut 64 percent, from $72 million to $26 million. In comparison, overall spending for EOSDIS was reduced by only 10 percent. Furthermore, project officials told us that future reductions in the EOSDIS budget would again impact directly on funding for developmental prototyping. This de-emphasis on prototyping clearly jeopardizes the "build a little, test a little" strategy.

By choosing this strategy, NASA has attempted to preserve the substantially larger and more costly effort by Hughes to develop a complete, new data system for the near term. Although Hughes plans to use commercial-off-the shelf hardware and software wherever it can, extensive software development is still needed. Standards for large, distributed database

systems are not yet fully developed, and there is no commercially available software to implement many of the required functionalities. As a result, Hughes estimates that over 1 million lines of code will have to be written in-house. Approximately 22 percent of the overall contract costs is allocated to system development. Actual costs may well prove to be substantially higher.

We recognize that NASA needs a system in the near term to provide data handling for the satellites that will be launched beginning in 1997. However, it appears reasonable and prudent to limit this near-term development to a smaller-scale effort. While the existing Version 0 system does not have the capacity to support the upcoming satellite missions, it could be enhanced to provide such support at least on an interim basis. Deferring full-scale development would allow a better, more appropriate system to be designed and built later, when technology and standards have further advanced and user needs are better known.

OPTIONS FOR RESTRUCTURING EOSDIS DEVELOPMENT DESERVE CAREFUL CONSIDERATION

The scientific community has also been concerned about the approach NASA is taking with EOSDIS. In December 1993, NASA's advisory panel of scientific users suggested that NASA commission independent studies of alternative architectures for EOSDIS that could result in a better, cheaper, and more easily evolvable design. NASA responded by commissioning three teams of independent university researchers to examine alternatives to the planned

architecture, or overall design, of EOSDIS, at a cost of over $1 million. These teams reported their results to NASA in September 1994.

The teams offered a range of recommendations for improving the effectiveness of EOSDIS and saving money. Some recommendations had to do with specific implementation details while others made significant changes to the system's overall architecture. NASA appears to be making a good effort to address the recommendations having to do with implementation details, but it has not considered making any high-level changes to the system's architecture.

The proposed architectural changes are largely aimed at trying to streamline near-term development and be more responsive to users. For example, one team suggested that greater effectiveness, as well as substantial cost savings, could be gained by reallocating the roles and responsibilities of the data centers. Under this proposal, large-scale data processing and archiving would be confined to two "super" data centers, which presumably could do the job more efficiently and cheaply than eight, as currently planned. The remaining data centers would be freed up to concentrate on their primary role of interacting with scientists and responding to data requests. In fact, since a smaller investment in new hardware, software, and system support would be required at each of these centers, additional centers could even be added without increasing the overall cost.

The two other teams, while not recommending the same approach, also suggested modifications in the number of data centers and their roles and responsibilities in an effort to be more responsive at a lower cost.

We have not had an opportunity to examine in detail the expected costs and benefits of the alternative architectures and thus cannot make a final determination on which, if any, NASA should adopt. However, we believe that NASA needs to look for ways to scale back its near-term systems development effort and take a more cautious approach with greater user involvement. These studies show that alternatives to the current EOSDIS architecture are possible that may help achieve those ends.

In conclusion, Mr. Chairman, I would like to stress that for EOS to succeed, EOSDIS must work well. EOSDIS will be the critical link for turning the data that are gathered by the EOS satellites into useful knowledge about global change. NASA has made good progress in defining a technical architecture for EOSDIS and, in that regard, has responded well to the findings of the NRC.

However, further adjustments could both save money and enhance the system's effectiveness. NASA needs to look at ways to bring users into closer collaboration on system development. The agency also needs to reexamine the pace of system development to see if too much development is being done too early, and too little is being invested in needed research and prototyping. In making that assessment, NASA needs to be realistic about what enhanced

functionality can be provided in the near term, given technical as well as resource limitations. The alternative architecture studies that have been completed may offer valuable options for scaling down near-term development and enhancing user involvement.

Mr. Chairman, this concludes our testimony. We will be happy to answer any questions you or other members of the Subcommittee may have at this time.