FINAL TECHNICAL REPORT

NASA GRANT # NAGW-2339

Mark R. Abbott
College of Oceanography
Oregon State University

MODIS Algorithm Development and Data Visualization Using ACTS

Abstract

The study of the Earth as a system will require the merger of scientific and data resources on a much larger scale than has been done in the past. New methods of scientific research, particularly in the development of geographically dispersed, interdisciplinary teams, are necessary if we are to understand the complexity of the Earth system. Even the planned satellite missions themselves, such as the Earth Observing System, will require much more interaction between researchers and engineers if they are to produce scientifically useful data products. A key component in these activities is the development of flexible, high bandwidth data networks that can be used to move large amounts of data as well as allow researchers to communicate in new ways, such as through video. The capabilities of the Advanced Communications Technology Satellite (ACTS) will allow the development of such networks. The Pathfinder global AVHRR data set and the upcoming SeaWiFS Earthprobe mission would serve as a testbed in which to develop the tools to share data and information among geographically distributed researchers. Our goal is to develop a "Distributed Research Environment" that can be used as a model for scientific collaboration in the EOS era. The challenge is to unite the advances in telecommunications with the parallel advances in computing and networking.

Rationale

The study of the Earth as a system will require enormous resources in terms of data, computer power, and scientific talent. The inherently interdisciplinary nature of the research, as well as its complexity, will require the development of new methods of scientific communications and collaboration to marshal these resources into an effective research program. It is no longer realistic to expect that an individual researcher, no matter how talented, can grasp the complexity of the Earth system and the associated data sets and tools necessary for its study. The proliferation of local research teams gives testimony to this need for group-based research.

In the past, such research groups typically have been based at a single institution with perhaps a few members located at other locations. With the advent of EOS, this "local mode" has been replaced by a more "distributed research mode" where collaborative scientific investigation requires the involvement of researchers at geographically dispersed institutions. However, collaboration using typical tools such as telephone and travel is difficult to maintain because of the need for flexible, "impulsive" contact between researchers, as well as the increased number of nodes that participate in a given investigation. As electronic networks have expanded, it has become somewhat easier to share data and information through file transfers and electronic mail so that distant team members can participate as effective members of the team. Networking has expanded the "base" of potential collaborators such that a researcher may belong to several groups that arise and disappear over time periods from a few months to several years.

The principal obstacle to such effective collaboration in the future will be inadequate communications bandwidth. The dilemma facing scientists is that as networking expands, performance often decreases because an increased number of users are demanding increased services. It is much the same dilemma that faces transportation planners; as soon as new highways are built, demand generally exceeds the capacity. While ground-based fiber systems will play a vital role in this improvement of networking capacity, it is not a complete solution. It will be several years before fiber reaches every researcher. Also, such fixed links do not allow much flexibility in network configuration, nor do they account for the degradation in performance that will occur as traffic leaves the high speed backbone for the local networks or as the networks become overloaded. It is analogous to estimating travel time simply by measuring distance between two cities without accounting for the delays that will occur as traffic moves through various low speed roads in towns, villages, etc. High speed fiber backbones are best suited for systems where there is a predictable, high volume load, such as between supercomputer centers.

The explosion of data from both space and ground/sea based sensors, and the increased level of collaboration that characterize EOS places strong demands on the communications network that will permit scientists to distribute data products and information in an interactive framework and in a timely fashion. Data from satellite and in situ sensors must be transmitted to processing and analysis centers, and the processed and/or blended products must be broadcast to a client community that is geographically dispersed. Interactive data analysis with scientists from remote institutions will require digital teleconferencing and integrated video data exchange. Current communications networks will be insufficient to meet the need.

The Advanced Communications Technology Satellite (ACTS) has several unique features which make it ideal for addressing these communications problems in the Earth Sciences and ACTS will provide a valuable adjunct to high speed ground networks. With its ability to deliver high bandwidth associated with the multibeam communications package, which includes both fixed and hopping beams, along with its flexibility in terms of network configuration, it will be possible to build up and tear down networks as the need arises. It is particularly well-suited to the "burst" or "bandwidth on demand" mode of traffic as occurs during occasional transmission of large data sets, video servers, or teleconferencing. The challenge is to integrate ACTS capabilities into existing network services so that the user is unaware of the exact nature of the physical link underlying the communications channel.

By utilizing existing data streams and terrestrial communications paths, one could establish a "Distributed Research Environment" among three institutions which will then serve as a model testbed for future development of the ACTS network after launch in 1993. We see the communications network as one of the key weaknesses to achieving success with EOS and therefore believe that the ACTS-based communications concept in the should be advocated in the EOS context.

Results

There are four general areas where ACTS could make a unique contribution to the EOS communications network.

First, remote data collection platforms at sea and in the terrestrial environment will be increasingly be used both for direct capture of satellite observations (i.e. HRPT transmissions) and to provide observational data from ground or sea based sensors. The requirement will be for transmission of these data to remote analysis and processing centers. Current remote platforms are severely bandwidth limited; the capability of vastly increased multipoint to point data communications paths with high bandwidth will immediately open many new observational opportunities.

Second, ACTS could provide a point to multipoint distribution network for derived products resulting from EOS sensors. For example, data from distributed earth station reception could be merged at a central location to form, say, a sea-surface temperature map of the U.S. coastal waters, and distributed to a wide range of researchers interested in the using the product for input into numerical models of for further researcher-specific analysis.

Third, geographically dispersed researchers could use the network capabilities for joint interactive data analysis. The combined video, data and voice requirements place high bandwidth demands, even with digital compression. Many researchers are not and likely will not be accessible by fixed land lines. Interactive applications and teleconferencing are very sensitive to high latency. Satellites have often not been considered a viable solution due to the delay incurred by the two round trips to geosynchronous orbit that the connection must take. With its on board routing, ACTS eliminates one of those hops and can compete favorably with the land based Internet. Its high bandwidth and flexibility also make it a viable solution for video and data analysis teleconferencing.

Finally, ACTS could act as an advanced TDRS for point to point transmission of satellite data to ground stations. This would improve the current limitations in the amount of data stored and transmitted during the time the satellite is in the field of view of the receiving station and opens opportunities for distributed data reception of specific data streams to a distributed data analysis network.

Future Capabilities

Publishing results and corresponding with peers is the backbone of scientific research. It allows researchers to make their findings known to a wide audience. Unfortunately, as science and technology progress, this body of knowledge will grow at an ever increasing rate. It is difficult for a scientists and engineers to follow all relevant journals in their particular field, let alone explore other fields for relevant information. Present methods of publishing are too slow, not in a digital format, and offer only a tiny fraction of the actual data and work that went into their making.

It is often difficult to retain the knowledge base that went into the development of a particular scientific idea, paper, or model. Present solutions generally have inadequate bandwidth and are ephemeral. That is, there is little corporate memory or any methods to retrieve the various paths, dead ends, intermediate data sets, comments, brainstorming sessions, etc. that are essential to scientific understanding. The result is an incomplete record; a complete record would prove invaluable not only in allowing other researchers to recreate the process used to develop the theory, but also for teaching.

With the increasing amount of data and processing that goes into a research paper, printed media will become increasingly inadequate at conveying the information. More and more researchers will request access to the raw or partially processed data to verify or extend the research. The "paper" will also incorporate a wide variety of data types, ranging from model code and visualizations, to records of audio and video conversations. CD-ROMs appear to be a useful form to maintain these "papers" and the associated data sets. Researchers may wish to access these CD-ROMs over the network, much as they can browse through library listings. This again requires a high bandwidth network connect in order to transfer the data in a reasonable time.

Similar statements can be made about video servers. As numerical models produce increasing volumes of data, we expect that video animations will become an increasingly common tool for display and analysis of model output. The difficulty with this medium is that it is expensive to duplicate and distribute to the science community. One might store such animations on a video server, much as there are news and mail servers on Internet at present. This would allow researchers to access video segments of interest for playback on their local workstations.

Shared-X from Hewlett-Packard is an example of a software system that allows users to view identical windows on distributed workstations. Applications running under the X windowing system can be displayed simultaneously on multiple machines. All participants may interact with the shared application and have their actions distributed to the other participants. This system, coupled with an audio link, would allow scientists to examine common data sets in a manner that is quite similar to what can only be done locally now. The North Carolina Supercomputer Center is developing a similar system for model manipulation and visualization.

One of the limitations to Shared-X and similar projects will be available bandwidth to remote sites. Bitmaps can be many hundreds of kilobytes in size. The ability to display them quickly requires high bandwidth between the data source and the remote display. In a local environment where all machines are likely to be connected with a 10 Mb/s Ethernet, the network performance is usually acceptable. The Internet backbone is only presently being upgraded to 45 Mb/s (DS-3). With hundreds of users nationwide, only a tiny fraction of this bandwidth will be available to a single user. This will cause a severe drop in performance when an application is shared between remote sites. A communications satellite would be an ideal solution as it has the necessary bandwidth, can be easily reconfigured, and even with the round-trip delay will still compare favorably with the interactive response time of the Internet.

For researchers separated by long distances, teleconferencing offers a viable alternative to travel. Travel cannot be justified for short meetings and informal contacts that are an essential part of a working team. Existing teleconferencing capabilities via telephone are awkward, expensive, and not suited for the transmission of visual information. Video teleconferencing is limited to only a few locations and is inflexible in terms of scheduling and cost. One of the obvious applications of increased network capability is video teleconferencing using the networking/display capabilities of desktop workstations. Various computer hardware and software vendors are beginning to demonstrate desktop solutions running over standard networks (at least between local area workgroups). Apple Computers, Digital Equipment Corp., IBM, and Microsoft are some of the companies that have demonstrated capabilities at various computer shows.

While T-3 bandwidth is still required for two-way, high resolution video, several commercial units can function over T-1 lines using compression techniques. Long haul networks, such as Internet, have not been used as yet, because of inadequate bandwidth. The recent emergence of ATM/SONET promises the ability to handle multiple data types with sufficient bandwidth. Many vendors are actively working on incorporating ATM hardware in their workstations.

Most, if not all, of the hardware components necessary for multimedia and teleconferencing exist today. Video equipment manufacturers (such as Tektronix and the Grass Valley Group) provide the necessary video equipment such as switching equipment, multiplexers, etc. Telecommunications vendors provide the necessary routing, data conversion, and transmission equipment through protocols such as ATM. The remaining challenge is in the area of operating systems. For example, UNIX is not well-suited to the requirements of teleconferencing. Although many vendors are working on new technologies, most are in early stages of development.

The computer vendors provide the ability to control the network and store the data. Unfortunately, these areas are very poorly integrated currently. Differences in interfaces, data formats, language and paradigms, and an extreme lack of software hamper efforts to integrate all information transfer into a single package. Presently, "integrated multimedia" all too often actually means completely independent subsystems that happen to be in the same room.

Many vendors are claiming "multimedia" solutions, such as Sun Microsystems, Silicon Graphics, NeXT, and Apple, but these are all in early stages of development. Challenges appears to be in the software side, where it has been difficult to provide an application environment to provide the links between various data and media types. NeXT is perhaps the furthest along, but other companies have significant research and development activities.

One positive example is QuickTime from Apple, which is a system that is designed to handle video as another data type. Various tools are provided so that video records can be incorporated into other Apple applications, much as existing graphs from one application can be cut and pasted into other applications. If, for example, one recorded a video teleconference, one could use QuickTime to cut and paste the entire record (or appropriate subsets) into a meeting document. Another researcher could read the document, and then double-click on the video icon to play the video of the teleconferencing session. This provides an efficient means of transferring information as well as a method to maintain a database of these video records. QuickTime has recently been adopted by Silicon Graphics for its future products.

Often voice is inadequate to express an idea, which is why we point to things and use hand gestures. Often the topic of discussion is either text or graphics residing on a researcher's workstation. The ability to have multiple workstations sharing the same view of data, with multiple pointing devices would allow researchers to identify and manipulate the data directly. Several companies have been conducting research in this area and created prototype systems. US West has developed a prototype system for distributed medical collaboration. Called the Telemedicine project, it integrates voice, video, and shared computer applications. All functions are controlled by the user using a UNIX workstation. Connections are created using a telephone-like model. After the connection is established, users may speak to each other, view each other via remote camera, and share applications and computer graphics. Users have control of the remote camera's pan and zoom so they may "look around". Also, users may place "pointers" on the computer screen to indicate an item of interest to other viewers.

Summary

As the capabilities of hardware and software improve dramatically in the early 1990s, it will become possible to achieve "media integration" to support researchers located within academic departments and research units. With the coupling of these technologies to high bandwidth land and satellite links, it will be possible to expand this integration to include researchers distributed worldwide. Such global-scale "enterprises" are essential if we are to make significant progress in understanding the Earth as a dynamic system. ACTS capabilities can play an important role in these activities.