Expanding the AG Community in a Closed Universe

Roger Edberg, Artic Region Supercomputing Center
Nickolas Jovanovic, University of Arkansas at Little Rock
Barbara Kucera, National Computational Science Alliance/EPSCoR
Don Morton, University of Montana

Introduction

There are many isolated institutions with limited resources in rural America and many of them are located in EPSCoR states. EPSCoR, the Experimental Program to Stimulate Competitive Research, is a multi-agency program that encourages collaborative efforts with State governments and research universities in 21 states and Puerto Rico. In the NSF component of this program, the main objective is to increase the research and development competitiveness of an eligible state through the development and utilization of the science and technology resources residing in its major research universities.

It is very difficult to improve research infrastructure in these states or within other organizations without massive investments and influx of expertise. The Access Grid serves as a mechanism for researchers to "plug in" at low cost and begin to meet, learn from, collaborate and participate with experts across the globe.

Most of the EPSCoR states are predominately rural with low population densities and few urban centers. They can be sorted into four groups: (1) Far West--ID, WY, NV, MT, AK, HI, NM; (2) Great Plains--ND, SD, NE, KS, OK, AR; (2) Southeast--KY, WV, LA, MS, AL, SC, PR; and (4) New England--VT, ME. Together they form about 17% of the country's population, but much less than that in terms of technological development. Currently, the EPSCoR states receive only about 10% of the total NSF funding. NSF realizes through its review criteria, that geographical balance in its funding patterns is desirable, and even though the top universities in the country may be in California and Massachusetts, it is not healthy for the scientific and economic enterprise of the country to concentrate NSF's resources in those states. Talent arises in all parts of the country, and most university students choose to attend an institution near their homes. Therefore, to optimally exploit our national talent for science and engineering, it is important to have excellent centers of research and learning in every state in the country and the AG can help resolve this situation.

Background

In an effort to bring Access Grid technology to remote areas, specifically EPSCoR states, Barbara Kucera, in her role as Alliance/EPSCoR liaison, sought participants for a proposal to NSF/EPSCoR. Six states demonstrated willingness to participate in a proof-of-concept activity that would provide the Alliance development team, lead by Rick Steven's group at the Argonne National Laboratory (ANL), with a research environment for the development of distributed data and visualization corridors, and for studying

issues relating to collaborative work in distributed environments. The participants would be in the thick of the action!

NSF/EPSCoR awarded a grant that provided equipment for six Access Grid Nodes (AGNs) to the University of Kansas; the University of Kentucky; Montana State University; North Dakota State University; University of South Carolina, and West Virginia University. The National Computational Science Alliance held Chautauquas in 1999 and 2000 at which the host sites--the Universities of New Mexico, Kentucky, and Kansas, Boston University and the Ohio Supercomputing Center--introduced Access Grid prototype technology to large audiences. These Beta-tests at the Chautauquas were enough to show the potential of the AG application, and revealed some defects that needed to be corrected. Continuing research and development is ongoing to optimize the functioning of the network protocols, the assurance of "quality of service," and desirable new features, including improved security. Usage has increased as institutions began to conduct remote meetings, site visits, training sessions and educational events using their AGNs.

As usage increased, the communication power of the AGN and its ability to bring geographically remote areas of academe into the mainstream of scientific research is being recognized. Other EPSCoR institutions sought inclusion at their own expense-Alaska, the Arctic Region Supercomputing Center (ARSC); Hawaii, two sites at Maui High Performance Computing Center and Manoa; Arkansas, U. of Arkansas at Little Rock; Maine, U. of Maine; and Montana, the U. of Montana at Missoula. The goal is for every EPSCoR state to have an AGN by the end of FY01--the eleven states that do not yet have the technology are AL, ID, LA, MS, NE, NV, OK, SD, VT, WY and the Commonwealth of Puerto Rico. We believe that if we can locate one AGN per state, the benefits will be demonstrated and other sites will recognize the potential for their institutions and seek to be included.

As indicated, there are 11 EPSCoR states with AGNs currently underway and 7 actively participating; the other four are nearly ready to go online. They have the equipment, space and staff, and are working out bugs. The doors are opening for these states, as discussed. If the remaining 11 EPSCoR states are not able to participate, the digital divide will widen even further for a large portion of our very bright students and scientists who happen to be located in geographically remote areas. Three examples of an unwillingness to accept this fate follow.

University of Alaska Fairbanks

Alaska could be described as the most "extreme" of the United States. It has North America's tallest mountains, including active volcanoes, wildest weather and coldest temperatures, largest concentrations of fish and wildlife, and arguably the hardiest and most individualistic residents. *[Note: Residents of other EPSCoR states may challenge this statement!]*

Alaska's population of 600,000 inhabits a total land area of 570,000 square miles (365 million acres), or one-fifth of the combined size of the lower 48 states. Half the state's

population is concentrated in the cities of Anchorage (255,000), Fairbanks (70,000 regional) and Juneau (30,000.) The remainder are widely dispersed in towns, villages and small bush communities.

Key industries in Alaska include oil and gas, mining, fisheries, agriculture, and tourism and recreation. The US military is also a major contributor to the state; it is home to more than 50,000 active duty personnel. Construction and operation of the Trans-Alaska Pipeline in the late 1970's brought a huge boost to the state's economy. Planning is underway for construction of a pipeline to deliver natural gas from Alaska's North Slope to the lower-48 states.

The Arctic Region Supercomputing Center (ARSC) at the University of Alaska Fairbanks (UAF) is the key provider of high-performance computing resources (HPC cycles, storage, visualization, networking) and expertise to the Alaska academic research community, which includes UAF, and the regional campuses at the University of Alaska Anchorage (UAA) and University of Alaska Southeast (UAS) in Juneau, Sitka, and Ketchikan. Each regional campus operates remote satellite campuses in neighboring communities as well. The central UAF campus is home to the main undergraduate teaching facilities and research institutes specializing in fields such as geophysics, marine science and arctic science.

ARSC's effort to join the Access Grid community is motivated by two factors: distance, and the technology imperative.

Distance, both within Alaska and to the "outside," is a hurdle all Alaskans (and visitors) have to deal with. Anchorage and Fairbanks are separated by a six-hour drive or one-hour flight. Air travel between Anchorage and Juneau, the state capital, is 1.5 hours. Flight time from Anchorage to Seattle is 3 hours, and to Chicago 6 hours. International flights from Anchorage to Asian and European hub cities such as Tokyo and Frankfurt are 9-10 hours. Senior ARSC staff routinely travel to Washington, DC, for management meetings and internationally to attend conferences. There is no such thing as a quick trip outside Alaska.

Productive, long-term interaction in science and technology requires more than phone calls and email. The Access Grid is a cost effective means of bridging the distance gap within Alaska, to the other 49 states, and internationally. Substantial savings in time, energy and travel and accommodation costs will be realized. Many UAF researchers already participate in traditional international research collaborations, exchanging ideas, data and results on a periodic basis and traveling on working visits when time and budgets allow. The Access Grid will allow more frequent and less costly contact and interaction, and accelerate the research process. ARSC is ideally poised to utilize the Access Grid to fight this "tyranny of distance," within Alaska and worldwide.

The Access Grid is also a vehicle for ARSC and UAF to increase their base of skills, knowledge and experience with new technologies in high-performance computing, networking and collaborative applications. We are interested in participating in efforts to

refine the AG technology and develop new tools and methods. There are many potential benefits, as outlined above.

ARSC has acquired and configured the AGN hardware/software suite, aided by a grant from the Next Generation Internet initiative. We plan to become fully operational in February, after completion of remodeling work on the AGN location: the Board of Regents Conference Room (109) in the Butrovich Building on the UAF campus. This large conference room, located adjacent to ARSC, can be conveniently accessed by ARSC and UAF faculty, staff and students, as well as UA administration, local and state legislators and external visitors. Indeed, the room is dedicated to regular meetings of the University of Alaska Board of Regents. It is the ideal location to showcase the Access Grid on the UAF campus.

The Butrovich Building houses ARSC and its suite of HPC, visualization and networking systems. Close proximity to ARSC network hardware, including a connection to the Pacific Northwest GigaPop (and from there to the DREN), requires a minimum of additional infrastructure to accommodate the AGN.

Initially we plan to focus on collaborative teaching applications. Don Morton, a frequent visitor to ARSC from the University of Montana-Missoula, has expressed interest in working together with ARSC staff on courses in parallel programming and scientific visualization. Course lectures, demonstrations, tutorials and even exams can be run via the Access Grid. Knowledge and experience gained from this work will help us learn how to drive the AG equipment and work effectively. ARSC has several joint faculty in computer science and computational science that are keen to make use of the Access Grid for teaching.

Collaborative research and development and support of remote users are key areas where ARSC can use the AGN. Approximately 70% of ARSC's users are located in the states. The AG could prove to be a lifeline between ARSC and these remote users, at least those who have local AG access or some capability to send and receive content. Research and development projects could range from code debugging, optimization and parallelization to collaborative, distributed data visualization. The AGN could also facilitate and augment efforts to develop distributed visualization or virtual reality applications with users using ImmersaDesk or CAVE systems at remote sites, such as PMEL or AHPCC.

ARSC's users specialize in a variety of research fields, with an emphasis on high-latitudes and the arctic: space physics, geophysics, remote sensing, weather/climate, ocean circulation, hydrology, cold regions engineering, petroleum and mineral engineering, ecosystems, biology and chemistry. A number of researchers are working on cross-disciplinary modelling, for example, coupled models of ocean circulation, seawater chemistry and evolution of fish populations and bionutrients. These research areas offer lots of opportunities for collaboration.

There is huge potential for useful collaboration and discussion between ARSC and other supercomputing/HPC centers. Closer communication between centers would help us to

better organize, plan and support our respective user bases, and attract more users. ARSC and other centers will be able to interleave our expertise to provide a higher quality of user service. Extensions of the current AG technology may even allow users to mix and match the services and resources of various centers to better suit their HPC, storage, visualization and network needs. Likely partners for this type of collaboration are Albuquerque High Performance Computing Center, Maui High Performance Computing Center and the U.S. Army Engineer Research and Development Center Major Shared Resource Center.

In summary, ARSC sees the Access Grid as a tool for breaking the "tyranny of distance" within Alaska, from Alaska to the other 49 states, and internationally. Teaching, research and development, and management are all poised to benefit from this technology.

University of Arkansas at Little Rock (UALR)

Arkansas has a population of approximately 2.4 million, but it contains only one city with a population greater than 100,000, the capital city of Little Rock (pop. 190,000). However, the Little Rock metropolitan area has a population of approximately 500,000, and is located near the geographical center of the state. The southern and eastern half of the state contains the lowland valleys of the Arkansas and Mississippi Rivers, with agricultural land dominated by cotton, rice and soybeans. The northern and western half includes the Ouachita and Ozark Mountains. The state is about equally divided between highlands and lowlands and nearly half of the state is forested. Arkansas is the home state for Tyson Foods, Riceland Foods, and Wal-Mart. Little Rock is the home city of Alltel Communications (the sixth largest wireless telephone company in the U. S.), Acxiom Corporation (the largest demographics database company in the U. S.), and Stephens Incorporated (the largest off-Wall Street investment firm in the U. S.).

The University of Arkansas at Fayetteville, the only other Carnegie research university in Arkansas, is located in the extreme northwest corner of the state, approximately 200 miles away, or 3.5 hours by car. The University of Arkansas for Medical Sciences (UAMS), however, is located in Little Rock, just three miles from UALR.

In 1999, UALR was authorized by the University of Arkansas System Board of Trustees to create a new College of Information Science and Systems Engineering. This new college was funded through a special act of the Arkansas legislature for the 1999-2000 biennium ($2 million in 1999 and $5 million in 2000 for operating expenses, as well as $5 million towards a new building). The college comprises the new Departments of Information Science and Systems Engineering, the previously existing Departments of Applied Science, Computer Science, Engineering Technology, and Construction Management, and a new Minor in Information Technology for liberal arts majors. The Department of Systems Engineering offers two concentrations: computer systems and telecommunications. The Ph.D.-granting Department of Applied Science offers five emphasis areas: Applied Computer Science, Engineering Science, Applied Physics, Applied Chemistry, and Applied Biosciences.

Nickolas Jovanovic, a professor in the new Department of Systems Engineering first learned about the Access Grid when he attended SC99 in Portland, Oregon. The Argonne National Laboratory (ANL) AG team demonstrated the Access Grid on the exhibition floor. Jovanovic met Barbara Kucera at the Alliance exhibit at the same conference. In addition, the Alliance sponsored an EPSCoR birds-of-a-feather discussion at SC99, at which Ken Bishop, University of Kansas, made a presentation about the Access Grid. When Jovanovic learned that the Alliance had the goal of establishing an Access Grid node in each EPSCoR state, he decided to find a way to locate the Arkansas node at UALR.

Jovanovic had loosely collaborated with scientists in the Mathematics and Computer Science Division at ANL for many years, and an initial attraction of the Access Grid was that it might improve his ability to collaborate with Argonne scientists. The other initial attraction of the Access Grid was to use it to launch the telecommunication concentration of UALR's new systems engineering program. Jovanovic had already received funding to build a parallel Linux supercluster to launch the computer systems concentration, and the Access Grid seemed to provide exactly the right example of a cutting-edge "telecommunication" system because it was based on Internet protocols instead of telephone protocols.

At about the same time, Jovanovic learned that the founding Dean of the new college, Mary Good, was serving on the External Advisory Council for the Alliance. Dean Good sent Jovanovic and two other UALR faculty members to visit NCSA. Barbara Kucera arranged for the visit, including a demonstration of the Access Grid, in March 2000. Shortly after the trip to NCSA, Jovanovic discussed the idea of building an Access Grid node at UALR with Dean Good (a chemist) and UALR Chancellor Charles Hathaway (a physicist) at a meeting to discuss ideas for the new college building. The Chancellor immediately authorized Jovanovic to build an Access Grid node with internal funding.

During the summer of 2000, Barbara Kucera submitted a proposal to NSF to obtain funding for five more EPSCoR Access Grid nodes, including the node at UALR. Jovanovic obtained letters of interest from both Alltel Communications and Acxiom Corporation in support of the proposal; however, the proposal was not funded.

Due to the commitment of internal funds, Jovanovic was able to build an Access Grid node at UALR without external funding. The UALR Access Grid node became operational in October 2000. Although much remains to be done, the UALR Access Grid node has already been used to conduct several research planning meetings between scientists at UALR and UAMS, here in Little Rock, and potential collaborators at NCSA, ANL, and the University of Utah. This series of planning meetings revolves around a new and unique scientific instrument that has been developed at UAMS for passively monitoring maternal and fetal biomagnetic signals during pregnancy. This instrument has the potential to improve the assessment of fetal development and the understanding of uterine activity during pregnancy if several information technology problems can be solved. The Access Grid meetings led to the submission of a three-year, three million

dollar group pre-proposal to the Information Technology Research program of the National Science Foundation.

A proposal from Arkansas has recently been submitted to the National Institutes of Health EPSCoR program. Among other things, this proposal seeks funding to build Access Grid nodes at UAMS and the University of Arkansas at Fayetteville in order to create an Arkansas collaboratory for biosciences. These are the only three campuses in Arkansas that are members of Internet2. Since other non-Internet2 campuses in Arkansas are involved in the biosciences collaboratory, it would be useful for scientists at these campuses to have some Access Grid capability, even without Internet2 membership.

In conclusion, the Access Grid node at UALR is providing opportunities for research collaboration that would probably not have come about otherwise. At the same time, the Access Grid provides a stimulating demonstration for new and potential undergraduate students in the new systems engineering program at UALR, as well as many possibilities for exciting research projects for graduate students at UALR.

University of Montana - Missoula

The State of Montana is the 4th largest in area, yet has a population of less than one million. Most of the territory is filled with mountains, forests, ranches, small towns, and several "cities," with the largest having a population of about 100,000. The University of Montana lies in Missoula, in the western portion of the state near the Idaho panhandle. The closest major cities to Missoula are Calgary (8 hours north), Seattle (10 hours west), Salt Lake City (6 hours south), and Minneapolis (24 hours east). Road travel to any of these cities requires negotiation of high mountain passes, often under less than desirable conditions. Air travel to/from Missoula is expensive, and often hampered by bad weather typical of the Rocky Mountains. In short, The University of Montana is located in a wonderful environment, but is quite isolated from expertise in major metropolitan areas.

Don Morton, a professor in the UM Department of Computer Science specialises in parallel, scientific computing. Geographic isolation has created difficulties in collaborating and learning from other experts in the field. Affiliation with the Arctic Region Supercomputing Center (ARSC), located in remote Interior Alaska, has provided Morton with collaborations that translate into helping Montana researchers with their scientific computing needs, and bringing such knowledge into the classroom.

During a visit to Alaska in June 2000, Morton learned that ARSC would likely be installing something called an Access Grid Node. This immediately prompted thoughts that, if only U. Montana had an AGN, collaborations with ARSC and other U. Alaska researchers might take place in a more workable fashion. There was immediate talk of both ARSC and Morton providing HPC training from their home bases, to personnel at other sites. After attending the Kansas Chautauqua in August 2000, it was obvious to Morton that an AGN at U. Montana was a necessity. It would add immensely to existing projects, and it would serve as necessary infrastructure to ultimately stimulate more, spatially distributed projects.

The new Associate Vice President for Information Technology shared Morton's enthusiasm, and saw the value in having an AGN. A meeting with the Vice President for Research, along with a presentation to the campus community, culminated in the UM Research Office committing approximately $60,000 for necessary room modifications and AGN equipment. As of mid-January 2001, the room has been modified and equipment is being assembled. Estimated debut of the UM AGN is late January 2001.

Initial use of the AGN will be to facilitate collaboration in a beginning NSF EPSCoR project, *Northern Rockies Center for Applied Computational Science*. The purpose of the project is to promote HPC among Montana researchers by proactively helping them move their research into the realm of parallel computers. The project enjoys collaborations with ARSC, and with researchers at Montana State University (210 miles east of UM). Travel funds were built into the budget to aid with these collaborations, but an AGN will now allow for more frequent face-to-face meetings. Morton and ARSC, in conjunction with another EPSCoR state supercomputing center, are exploring the possibility of jointly teaching a course in parallel computing, delivered over the Access Grid.

Other immediate uses will be to provide access to remote seminars for local students and faculty. Additionally, Morton has been engaged in interdisciplinary projects with University of Alaska researchers for several years, but communication has always been difficult, partly because of the different "languages" spoken among researchers from other disciplines. Even more problematic has been engaging computer science students in the work, trying to "relay" information to the students. Now, with an AGN, they will be able to hold small meetings consisting of one or two UM researchers, an ARSC specialist, Morton, and a student or two who would ultimately be doing the actual work. In fact, one U. Alaska researcher has agreed to pay an undergraduate student's travel, room and board for a summer in Fairbanks, and the Access Grid will now allow all to meet beforehand and coordinate the project.

Future plans for the UM AGN are to promote it in disciplines other than computational science, and play an active role (while producing interesting projects for students) in future growth and enhancements in the Access Grid infrastructure. Faculty and students from a diverse group of disciplines will be "introduced" to the UM AGN and provided a chance to collaborate and meet with groups at other facilities.

All in all, the presence of an Access Grid Node removes, in a significant way, the isolation experienced by Montana researchers and educators. Now, for the first time, personnel will be able to participate, on a large scale, in national and international meetings. Researchers will be able to enjoy healthier collaborations across vast distances. And, what makes this all so incredible is that these same people will be able to keep living in one of the most beautiful areas on Earth!

Problems and Future Directions

The current AG technology has certain limitations, but that is part of the purpose of this retreat--to identify areas of concern to enable the development team to decide where to focus their energies next. Many of those in the "resource poor" regions lack the expertise to easily connect to the Access Grid. Taking the time necessary to install and operate an AGN is prohibitive at numerous institutions, and this difficulty is only compounded when local expertise is lacking. Key areas of concern are:

- Purchasing and installing AGN components requires a significant amount of patience, experience, or willingness to learn a lot of technical material. Jumping in to create an AGN still requires a huge time commitment, prohibitive at some organizations. There appears to be work in the creation of Plug 'n Play AGN's, and this is absolutely necessary if we are to move AGN's out of "geekdom" and into the realm of more "normal" environments.

- As with numerous beginning, distributed, projects, there is a fair amount of technical information available on the Web, but it is disjoint, not critically reviewed for changes and errors, and sometimes hard to find. Even if a centralized index of all these sites was created (and there are initiatives to do that), it would still be disjoint and very difficult to use, and its accuracy would always be in question. A comprehensive "Dummies Guide" is clearly needed. Although a "cookbook" approach might be helpful to some users, telling operators what to click when, there should be a source that spends a lot of time on background and theory, trying to foster an understanding of the entire AG infrastructure and the role an AGN plays in this environment.

- Cost is sometimes the determining factor for a collaboration choice other than an Access Grid Node.

- The focus of the AG is group-to-group communications and great strides have occurred in the past year in this effort. However, discussion of the current AG technology limitations is one focus of this retreat--to identify areas of concern to enable the development team to decide where to direct their energies next. One very important factor to enable the AGN to compete with (if not accommodate) the H.323 protocol, will be the availability of a secure "AG in a Box," able to be run by a single operator who is not an expert in audio-visual technology.

We are all familiar with the "can you hear me" syndrome at the beginning of each AG session when the audio levels are checked, etc., by the node operators. The analogy of a telephone comes to mind. Nearly anyone can make a point-to-point telephone call; however, if you want to talk with seven people in various locations, you can't call them all at once (unless your telephone is much more sophisticated than mine) without the assistance of an operator. With an AGN, you can reach one person at one site or at many geographically dispersed sites, but each site currently needs an "operator." Perhaps we should consider the merits of having an AGN virtual operator--a system whereby users

(with minimal training) could indicate who would be included in the session, the start time, and the room reservation. At the appointed time, the participants would go into the AGN (which is always left running) and participate. The only thing the user would have to do would be to indicate the end of the session. This is not a simple solution, but is offered to stimulate discussion.

Conclusion

Access Grid Nodes offer, for the first time, a means for geographically isolated sites to collaborate in a natural manner. Organizations may now share their local pools of expertise with others and, in turn, may now access expertise to make up for local shortcomings. In effect, the World Wide Web evolves towards collaborative webs of expertise. Significantly, it is now possible to imagine small organizations in remote areas contributing to these collaborative webs, and benefiting from them. Everyone gains from this increased access to information and collaboration.

It is important for EPSCoR states to recognize and adopt this emerging opportunity, and it is important for them to have a relatively painless approach for connecting to the Access Grid. Funding issues still exist and may require innovative solutions. The pioneering efforts of numerous groups have clearly opened the doors for others to follow, and it is believed that an evolution towards turnkey AGNs will make it possible for more organizations to participate, and for more disciplines within each organization to participate.

01-25-01