(~

DISTRIBUTED USER SERVICES FOR SUPERCOMPUTERS

Henry A. Sowizral

1 Apr 89

RIACS Technical Report TR-89.18

RIACS

Research Institute for Advanced Computer Science

DISTRIBUTED USER SERVICES FOR SUPERCOMPUTERS

Henry A. Sowizral

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report TR-89.18
1 Apr 89

This study examines user-service operations at supercomputer facilities. The question behind the study is whether a single, possibly distributed, user-services organization could be shared by NASA's supercomputer sites in support of a diverse, geographically dispersed, user community. It also identifies a possible structure for such an organization as well as some of the technologies needed in operating such an organization.

Table of Contents

This study is motivated by NASA's increased use of supercomputers and the associated need to support the researchers that use those facilities. The question is whether a single, possibly distributed, user-services organization could more effectively support a diverse, geographically dispersed, user community than the current structure, where each site provides its own user services support. If such an organization seems feasible, a possible structure for that organization and the technologies it would need to operate effectively should be identified.

1. INTRODUCTION AND MOTIVATION

The government has established supercomputer facilities at a number of universities and national research centers. No two of these supercomputer facilities are alike.

The supercomputer sites have been established as national resources. They operate as "facility class instruments," much as a facility class telescope or sub-atomic collider operates. But unlike telescopes or colliders, a supercomputer user need not be physically present at the instrument. The user can instead sit at his or her personal workstation, in the familiarity of his or her own office, and access the remote supercomputer via a communications network (NSFnet, ARPANET, CSNET, NASA Science Internet, etc.) This ability to work remotely means that supercomputer sites serve a substantial, non-local, user population.

A typical facility has many local procedures and policies that make it quite challenging for users to learn how to use that facility effectively. A user that interacts with multiple facilities has an even greater problem. To help users solve such problems, each supercomputer site dedicates a sizeable fraction of its operating budget to a user support facility.

The differences among the supercomputer sites and the large remote user populations might make it seem that site-specific user services provide the only realistic means for addressing use-related problems. But experience has shown that many user questions are fairly general in nature, and require little more than knowing where to locate specific information, such as how to determine the status of the supercomputer, or how to cope with subtleties in FORTRAN syntax. Many users call seeking help in locating documentation, in determining if the computer is operating, in accessing a particular supercomputer, or in learning about changes in policy.

Consider a new strategy: create a single point of contact with an 800 phone number and a single mail address. Using this simple concept, we can design a multi-level organization for all consultations involving NASA supercomputers. General questions would be handled immediately at the top level, while questions that require more detailed knowledge or site-specific information would be routed either to an expert in that specialty or to the appropriate person within a specific supercomputer facility.

An unified user services organization could provide a simpler, easier interface for supercomputer users. They would know immediately where to call or send mail when they have a problem or a question. An unified, jointly run, user services organization could allow the NASA supercomputer sites to more effectively use their operating budgets by freeing up staff that was assigned to manning "hot lines" at each site and reassigning those personnel to other tasks.

The telephone, computer mail, and surface mail seem essential to the first phase of such an organization. Other technologies also appear important in the medium term: software for capturing the state of a terminal's screen and multi-media mail for forwarding it to user services; software that consultants can use as an "institutional memory" to keep track of the state of the systems, answered and unanswered questions, and information about the users who asked questions; software that users can use for obtaining appropriate reference material, running the gamut from traditional documentation, to how-to tutorials (how do I compile a program, how do I restructure my disk files), through quite exotic computer-based animations or "movies" that could lead a user through each step in an activity. Other technologies, not yet mature, may prove to be as important in the future. These would require research funds to ensure their appropriate development with user support as one important focus. Included in this arena would be research in collaboration technologies and the use of multi-modal (audio and video) presentations in conjunction with digital technologies.

The sections that follow describe the study, provide an overview of the services provided by user services, summarize what we learned interviewing user services staff as well as users, list the issues to consider in establishing an unified user services organization, and provide recommendations for future action.

2. THE STUDY

We visited several supercomputer sites to discuss user-support needs with users and user-support staff. The purpose was to learn first hand what services were provided by user service facilities, how the organizations were structured, what they perceived user needs to be, how users perceived user services, what problems users had with user services, and what was seen as their successes as well as failures.

We chose three sites to visit, the Center for Theory and Simulation in Science and Engineering located at the Cornell National Supercomputer Facility, The National Center for Supercomputing Applications located at the University of Illinois Urbana-Champaign, and The Numerical Aerodynamic Simulation Facility located at NASA Ames. The first two sites, Cornell and Urbana-Champaign were chosen as representative of other NSF funded supercomputer facilities and because they had dissimilar hardware. Cornell runs an IBM 3090 supercomputer while Urbana-Champaign runs a Cray-2 facility. The third site at NASA Ames was chosen as representative of a NASA funded facility.

At the three sites we visited and talked with a broad range of individuals: consultants, directors of user services, documentation specialists, mangers of the outreach programs, trainers, managers of documentation services, research scientists, and users. We wanted to learn how different individuals viewed "user services" either by definition or through anecdote. We asked questions hoping to elicit such items as what they viewed as their user services strengths, what sorts of questions were asked by users, whether they perceived any gaps in services from both the user's and the provider's perspectives, and how they would structure an unified user services facility.

3. WHAT ARE USER SERVICES?

Though we seem to understand the term "user services" intuitively, different sites do provide different services. In general, "user services" tended to be the set of services provided by your particular facility. The disparity in services that results from this definition is a source of irritation, especially for users that use more than one facility or who developed their model of user service from their interactions with a facility they used previously.

Still, there is a set of core services and functions that most facilities provide. Users generally see "user services" as a place or phone number that they can either visit or call

with a problem and receive a more or less timely answer. This definition implies that user services provides little more than consultation services. But, consultation services constitute only one part of a user services facility; other services include training, documentation, and in-depth programming support.

3.1 The Role of User Services

User services must optimize a set of complex, interacting, constraints. It must provide the best service possible to users, given its staff. It must ensure that its staff remains interested and continues to grow. To ensure that users get the most out of their facility, the services staff must keep its users up-to-date on the latest development in programming techniques and on changes in the facility's hardware and software. And they need to do this in an atmosphere where the users are very goal driven, may not realize they need training, or may not wish to invest the up-front time to learn.

Because staff is limited, a user services consultant must walk a tightrope between aggressively pursuing a problem to its complete solution and understanding only enough of the problem to guide the user towards a possible solution. The consultant must do this in a manner that not only recognizes the needs of the user, but also those of the facility at large. The supercomputer facility is interested in making sure that users do not waste the resources of the machine--that, where possible, users have taken the time to maximize the efficiency of their programs, either by vectorizing their code or using the multi-processing capabilities of the underlying hardware. The users are interested in making sure that they do not waste their most precious resource, their own research time.

User services must also handle a broad range of users and user abilities. Some users like to solve their own problems whenever possible, and so only ask user services the most difficult of questions. Other users do not bother to look up anything at all and call user services to ask the simplest of questions. Not only must the user services consultant be a supercomputer jack-of-all-trades, but he or she must have a sufficient grasp of people to recognize and respond appropriately to their needs.

3.2 Services Performed

The range of user services includes many different items. We grouped them into four broad categories. Specifics varied according to the site. The first category, training,

consists of seminars, tutorials, and initial support. The second, publications, consists of newsletters, brochures, user guides, documentation, and interactive computer texts. The third, programming support, refers to the range of services provided users who are developing code and need specific programming support in various areas such as vectorization, visualization, multi-programming, or porting existing codes written for a different architecture. The fourth and last category, advocacy, includes consultant services for the user as well as the representation of the user's views with the administration.

Training

Training consists of any activity performed by user services that uses a prepared sequence of instruction. We divide training into four types based on duration and target audience: classes or summer institutes last longer than tutorials and cover the material in more depth, seminars provide an in-depth examination of one particular aspect of technique or science, while initial support or first time user training attempts to bring a new user up to speed quickly on the equipment, the software, and the interaction environment.

Classes/Summer Institutes. User services staff provide their users with fairly complete courses on important topics of interest. These include course such as UNIX and programming in C. Rather than spreading these course out over a ten or fifteen week period of time, a typical course lasts for five days and will meet from 8 am to 5 pm.

Tutorials. Some user services staff provide specialty courses that focus on a particular topic such as vectorization, multi-processing, or code porting. These courses tend to run for one to two days and can include hands-on participation. They usually explain the material in a general way and then quickly move to those concepts specific to a particular architecture. As an example, a tutorial on vectorization may spend two hours describing the general concept and then spend the remainder of the time talking about vectorization of code for use on a CRAY XMP, a CRAY-2, or an IBM 3090.

Seminars. Seminars provide detailed information concerning work-in-progress or specific research results. In many cases, seminars are jointly hosted with the various research disciplines working at the supercomputer site.

Initial support. All supercomputer facilities provide some new user support. A new user typically receives documentation that introduces him or her to the operating system, the editor, and an application language. Some sites, when they establish a new account, include a typical "initial" environment configuration. The training portion of the new user orientation involves helping the user to log on to a workstation, and provides details on how to network into the supercomputer, how to transfer files to the supercomputer (BITNET, FTP, KERMIT), and how to get on-line help information. Some facilities provide free computer time (10 hours in cases we investigated) to "get acquainted" with the new environment.

Publications

Publications consist of any written documents generated by the user services staff. This includes paper documents and computer documents, whether static or interactive.

Newsletter. All supercomputer sites we interviewed regularly publish a newsletter. These include a listing of recent changes and planned changes, as well as articles featuring particular users, new commands, new applications, and quick tips. Dissemination can occur via the U.S. Mail, an on-line news service, or a herald that gets printed when a user logs into the system.

Documentation. No one can use a computer facility without having access to reference documentation. A typical user needs vendor supplied documentation as well as facility provided documentation. The facility must provide custom documentation because few facilities run unmodified versions of vendor systems; the support staff must document local features and modifications. Such documents might include tutorials, manuals, on-line versions of documents, and hierarchical help systems.

Interactive Documentation. Documentation can take on a descriptive form like a book or text file. It can also take on an interactive form, one that presents the steps needed to perform a particular task and then lets the user perform those steps as he or she reads the document. Some user services facilities are taking this idea and developing documents that coalesce information from multiple descriptive documents into such an interactive document. A typical document might describe the steps needed to perform a particular task such as compiling a program.

Programming Support

Programming supercomputers differs from programming other types of computers. Special skills are required to extract the best performance from programs written to execute on normal computers and now targeted to a supercomputer. Rather than waste CPU time, supercomputer facilities find it useful to help users port existing code to the supercomputer with an eye towards vectorization and multi-programming and, as their users and center staff become more sophisticated, to suggest additional improvements.

Some supercomputer user services facilities go further in supporting specialized activities. They have formed applications groups targeted at particular problems, whether it be computational chemistry, fluid dynamics, visualization, or physics. These special applications groups are available for in-depth consultation by those users who working in the same area.

Still another program targets visiting scientists. When the visiting scientist arrives he or she gets assigned a consultant from the user services staff. That consultant then has the task of caring for the scientist. The rationale behind the program: visiting scientists need more support, their time at the facility is limited, so some mechanism that shortens problem solution time is of great benefit.

Advocacy

Though most would view user services as flowing in only one direction from the consultant to the user, in fact the user provides much valuable information that the consultant should make available to appropriate people.

Consultation. A typical user-services facility receives a broad mix of questions running the gamut from the rudimentary:

> "How do I get an account?"; "How do I connect to the Cray?"; "What utility will let me do ...?"

through the seemingly capricious:

> "Why isn't my program working? I haven't changed a thing and the program is now computing the wrong value."

to the extremely complicated:

> "The compiler seems to act in a strange way. I used this construct but it didn't seem to get the right results when I turn on full optimization. Is there some known bug in the compiler that can explain this phenomenon?"

or

> "I've just gotten an illegal op-code after recompiling my code with the new compiler. It got the error inside a library routine. Is there an unannounced incompatibility?"

The bulk of the user services requests, typically about 25%, are information requests, such as "Is the system up," "Who is the right person to ask about x," "Where can I find documentation on x," and so on. Other categories of questions include: new user questions, bug reports, and questions from people affected by changes to the system. As an example, when one center was having particular problems with network connectivity, a large fraction of the questions/complaints were focused on networks. In another situation, when only vectorized code could be queued for execution, many users began asking questions concerning vectorization.

Acting on User Information. Many times users will find bugs in vendor software or hardware. The consultant must take an active role in ensuring that the local staff as well as the vendor are made aware of the problem. Similarly, users may point out flaws in stated policy or planned operation. It is the role of the consultant to ensure that information gets to the appropriate responsible party, to track the responses, and to inform the user of the status.

A tremendous amount of user dissatisfaction occurs from a lack of feedback. Even though a user's request or statement may be generating considerable change, because that user is not kept informed he or she may feel ignored or worse still inconsequential.

4. WHAT WAS LEARNED?

Interaction with the people at the various supercomputer sites provided a wealth of valuable information. Though some of the information was anecdotal in nature, it was clear that similar problems plagued the facilities and that similar solutions were being discovered independently. The themes that recurred most frequently were that:

- There is sizeable growth in supercomputer user populations,
- Staffing and retaining user-services staff is quite difficult,
- Structuring the consulting staff requires great thought,
- Keeping user-services consultants up to date is a never ending process,

- Time spent in user training greatly reduces the frequency and duration of user-consultant interactions,

- Off-site users need special consideration but resources to support their special needs are lacking,

- Other supercomputer sites are a good source of information but not enough effort is spent on site-to-site coordination.

4.1 Growth

Supercomputer facilities are changing rapidly as scientists learn about their capabilities and limitations. Pioneers in the scientific use of supercomputers now think of scientific computation as just one more tool that they can use in pursuing their research goals. Their less venturesome colleagues are also beginning to see computational science as an important technique for them and they too wish to use supercomputers.

This rapid growth in supercomputer use is placing considerable strain on user services. The user population is growing rapidly, more rapidly than staffing budgets. Additional strain arises from the very broad range of programming abilities found among the users. To support this disparate group of users, user services must provide an extremely broad range of support on a wide variety of topics. Not only is user services expected to provide the "normal" set of services, but they are also expected to provide more unconventional services, such as tutorials on state-of-the-art programming techniques including vectorization, multi-processing, and visualization. They must sometimes also provide in-depth consultant services, and even develop specialized software. Many user services groups are finding it difficult to "keep up."

4.2 Staffing

A user services facility needs to maintain a fairly eclectic staff. Its consultant must understand several computer architectures, the supercomputer itself as well as the various workstations used in accessing the supercomputer. He or she must have some understanding of the programming languages, the supported applications, and also the language of science, if for no other reason than to talk intelligently with users.

Keeping consultants interested

Finding good consultants is hard. Keeping them interested in providing user service is harder. A typical consultant works in overload mode when first hired. A consultant may

know quite a lot about the hardware and software involved with the supercomputer, but he or she will not know all the answers to all the questions that users will asked. As consultants fill in their gaps in knowledge, they tend to feel less and less challenged by their work, and begin to feel more dissatisfied with their job. One method for alleviating this problem arises naturally from a consultant's individual interests, namely specialization in an area of interest. Another technique involves rotation of duties between front-line consultation and more in-depth support. Flexibility appears to be quite important especially in matching a consultant's personality with their assigned task. Employee satisfaction is very important. Without it a facility will have very high turnover rates.

Good Characteristics

Finding "people" people, that is, consultants that like to help people, is extremely important. Such consultants will derive additional personal rewards from helping others. A certain level of maturity is extremely important. Though age provides an important indicator in this arena--it was noted that science graduate students make better consultants than undergraduates--the issue really speaks to professionalism. How does the person view the job: is that person a user consultant because he or she is providing a service or is that person a consultant because they know more about the machine than other people. Though a potential consultant's expertise may be crucial to the supercomputer facility, assigning them a job as a consultant may not be appropriate.

Appropriate technical expertise

Computer science students were not good hires for front-line user-service support. They tended to focus more on problems that interested them technically. They identified more with the computer science issues presented by the problem than the user's need to solve the problem rapidly. On the other hand, people whose primary background was either in the sciences or engineering and who also had a background in the use of computers tended to be good supercomputer, user-service consultants. They identified more closely with the users and their needs. And, they tended to limit their search to solutions that were "good enough;" solutions that let the user get on with the research rather than solution that were elegant.

Staff Composition

Many supercomputer facilities are located on university campuses. Their proximity to educational institutions is an important ingredient in transferring the information learned to the next generation of scientific and engineering professionals. That same proximity permits a supercomputer facility to augment their professional staff by hiring students. Different facilities hold different viewpoints on hiring student user-support staff. Some facilities neither hire students nor part-time professional because they believe that part-time help cannot stay current on the status of a supercomputer's operating environment and at the same time provide support services. Others believe that part-time support personnel can be effective in providing first level support, but only if they work a minimum of 20 hours per week.

4.3 Structuring the Consulting Staff

Many of the questions asked of consultants do not require detailed knowledge of the supercomputer. As a result, supercomputer facilities try to optimize the use of senior staff. Some facilities use a different staff to handle "walk-in" questions. The front-line staff answers the less involved questions and transfers the more complex questions to the appropriate senior consultant. Different facilities structure such services differently.

Single Point of Contact

One structure that failed involved the use of a single point of contact. In this case a secretary received all incoming calls and, after asking some questions to help localize the problem, transferred the call to a particular consultant. Two difficulties arose from this approach. First, the secretary did not always understand the nature of the problem correctly and so transferred many of the questions to an inappropriate consultant. Second, the user asking the question did not diagnose the source of the problem correctly and so got transferred to an inappropriate consultant. That consulting group changed its structure to one where the consultants themselves answered incoming calls. The first consultant to talk to a user owned that user's problem or at least its initial diagnosis.

The single point of contact failed for at least two reasons. One, it introduced another person into the chain of people a user needed to talk to before he could get an answer. Though the secretary could answer some questions and thus off-loaded some work from the consultants, each phone call required the user to explain the problem to at least two

people. The secretary was perceived as shielding the user services staff, not augmenting them. Second, many time the secretary chose the wrong consultant adding yet another level of indirection.

You Hear It, You Own It

A second structure, where the consultant that "hears a problem, owns that problem," works better from both the perspective of the consultants as well as that of the users. The consultants get variety. The users get informed answers with a minimum of intermediaries. But the "you hear it you own it" method suffers from an unusual problem: inconsiderate users. Some users will call and ask the same question more than once. Each time talking to a different consultant, effectively shutting down user services as two or more consultants work to solve the same problem.

This scheme can also create some "bookkeeping" problems. Take the example of a user who asks a complex question of one consultant and then later has another, unrelated question to ask. The user needs to keep track of which consultant is solving which problem. Otherwise the user could not ask follow up questions nor could he or she inform the consultant of any new information that might affect the problem's solution. Similarly, consultants need to cross check to ensure that a user does not call in multiple times asking the same question thereby monopolizing all the resources in user services.

Two Level Structures

More complex, two level structures are also possible. First level consultants hear all queries, if they know the answer or can find one rapidly, they own the query. Otherwise, they pass the question along to a second level consultant. They choose which second level consultant to assign either by availability or on the basis of specialization.

4.4 Consultant Awareness

A major problem facing a user services facility, and one that will pose a serious problem for a distributed user services facility, is keeping the user services consultants aware of changes at the supercomputer facility. Consultants need to know the current status of the computer systems, whether the computer is up or down, whether interactive logins are allowed, whether communications equipment have been overloaded, and so on. User services consultants also need to know about recent changes to the system;

whether the FORTRAN compiler has been upgraded, what new applications have been installed, what applications have been removed or replaced by different codes, and what policy changes have occurred. Lastly, user services consultants need to know about proposed changes.

User services consultant need a close working relationship with the operations staff so they know what is currently happening; what hardware is down, the status of the network, what software changes have occurred. The consultants also need a close working relationship with the system programmers so they are aware of planned changes or perceived problems.

It was noted that housing consultants near the operations and the systems programming staff had significant benefits, by encouraging informal encounters. User services consultants were more likely to meet the operations or systems staff in the hallways. Such chance encounters would allow the consultants to ask "trivial" questions that they might not wish to ask more formally. By asking the question in the hallway the consultants were not "interrupting" the other person and so more likely to ask a less pressing question.

A single user-services facility that serviced multiple sites would need to carefully implement explicit procedures for communicating such operational information.

4.5 User Training

User services have found that time spent on training users reduces the time they spend with users solving problems. However, they have also found that users normally learn about a facility and its use by reading newsletters and documentation rather than by attending classes or tutorials. Users find it difficult to take the time out of their schedule to learn a particular topic well.

In situations where users do attend a course, the course has high attrition rates. Users seem to constantly evaluate a cost-benefit function. If they do not perceive themselves receiving large immediate value from their attendance, they cut their losses and stop attending the course.

A large disparity in backgrounds of the users attending a course causes a large part of the attrition. The people attending tend to be bright, competent programmers who know their field but had little formal training in computer science. Another factor is course structure. Because a course instructor must bring everyone up to a common level before he or she can introduce the more complex concept, the first day or two may not be of benefit to some attendees. In fact, it is likely that this will be the case for at least half the class.

Other problems with developing training courses stem from a lack of uniformity in the particular sub-discipline. Many users wish to learn about graphics so they can present their data in a more useful manner. Defining graphics or visualization is easy; however, describing how to do it in general is extremely difficult. It may require too much investment from a user's perspective to be worthwhile.

Lastly, it became hard to decide where training ends and education begins. Should training courses provide basic computer science training? How much? When should it refer the user to course curricula at a university?

4.6 Off-site Support

A large fraction of a supercomputer's users are located off-site. In fact, facilities run user-outreach programs to increase remote use. Those programs inform potential off-site users about the supercomputer facility, its availability, and its use. Some programs include a modest amount of free supercomputer time. The free time is usually enough for a potential user to log on, get acquainted with the operating system, and possibly even compile and run a small program.

The individual user-outreach programs vary in style and structure among supercomputer sites, but they usually require a remote institution to designate an on-site user contact.

The success or failure of the remote site's program rests heavily on that site's local contact. If that person is well trained and comfortable with using the supercomputer, then users at the remote site will perceive the supercomputer as a friendly and useful system.

Not all remote sites can allocate a full-time person to serve as a point-of-contact for that institution's users. If not, the lack of support can seriously impede the use of the remote facility. One way of solving this problem, adopted at some remote institutions, is the designation of a single point-of-contact for multiple supercomputers--a first step towards out concept of a single point-of-contact for all supercomputer users.

This local "user services" consultant acts as a remote front-line service. He or she can help with local questions that the remote staff may know nothing about such as setup information for locally used terminal and networking equipment. That person also acts to answer the simpler questions that do not require the detailed knowledge available at the supercomputer site. Lastly, that support person provides a modicum of face-to-face support for those users who need to "get to know" the person who is helping them and for those situations where the consultant must observe the symptoms directly.

4.7 Center-to-center Communication

Most of the supercomputer sites have experienced the same problems. Yet, as a group they do not communicate with one another. Some sites have communicated with one another especially when they have similar hardware and software and a need arose. And as systems changed the inter-site communications changed. Yet, many structural problems are quite similar across all facilities. Our discussions elicited an interest within the supercomputer facilities in increasing their communications among one another.

5. ISSUES

Establishing a single supercomputer user-services facility raises a number of concerns. How does the new facility handle questions, the referral of questions to other consultants, the tracking of questions to ensure their proper handling, and, variations in policy among the various sites?

5.1 Handling Questions

As discussed earlier, some users will call and ask for answers to the most basic of questions while others call for help only after they have exhausted every source of information at their disposal. How should the new organization handle questions? Should it answer simple requests directly and provide the answer? Or should it refer the user to a specific document and page number? The first solution encourages user-services abuse,

but it also acts like a service organization. The latter help user help themselves, but could engender a high degree of antagonism. In general the major questions is, what is the correct role of the new user services organization?

5.2 Triage

The first thing that a user services consultant must perform is, for want of a better term, *triage*. A consultant must decide who best can answer the question. If it is one that this consultant can answer, that is, the requestor needs to know if the system is up, the phone number for accessing the system, or the question falls into the consultant's area of specialty, he or she should provide the answer directly. More complex questions or those outside the consultant's specialty will require referring the requestor to another consultant. Referring questions to some other consultant may be the only recourse since no one person can be an expert in everything.

Referring a user's problem to some other consultants may sound like a viable and reasonable technique for solving a problem. However, users do not like to be handed off repeatedly. Informally, user services staff believe that users will graciously handle no more than two transfers. No one likes to repeat their problem over and over again.

Much thought must go into an appropriate structure for question referral. In fact, referral may be wrong. Rather, an appropriate structure may have the consultant that receives the original question providing the answer even if that consultant might need to obtain help from someone else. If referrals are to be used then care must be taken in choosing who should receive the referral. The front-line consultant must diagnose the user's problem well. This includes assessing the capabilities of the user. Many time users may pre-diagnose their problem incorrectly.

5.3 Tracking

Users get quite irritated with user services when they feel that their problem is being ignored. Consultants must make sure to handle questions that cannot be answered directly with care. If a question will take more than a day or two to answer, then the consultant should contact the user and explain this fact. It is also important that the consultant provide a way for the user to contact the consultant directly. A consultant cannot make himself unavailable during the time he must research an answer to a

question. Sometime the user has found the answer to the problem on his or her own and wishes to contact the consultant to save them time and trouble. In other cases, the user wants to perform a status check. Unavailability in the latter situation can cause extreme duress.

Almost as annoying as having a question dropped on the floor is not being able to reach user services. Consultant must always answer their phones and mail as soon as possible and they should add the harder problems to their list of things-to-do. A consultant may be away researching a problem in detailed documentation found in the machine room, known only to a system developer, or known by the hardware/software vendor, but other consultants should know this and provided nervous users with status.

Tracking is as important in avoiding duplicate work as ensuring that all questions get answered. Some users have been know to call multiple times to ask the same question. Without a good tracking system two or three members of the consulting staff may end up researching the same problem for the same person.

5.4 Policies

Policy issues are some of the hardest problems to handle even in single facility situations. Computational projects are now being rejected at some facilities because users still have not vectorized their code. Few mechanisms exist for modifying policies or even of informing administrations of the impact of their policies on users. Users are notoriously good at circumventing policies to minimize their own discomfort. Because user services is the user's contact point, they will act as the liaison between user and supercomputer facility administrators. Attention must be payed to this role to minimize feelings of being ignored.

6. CONCLUSIONS

We believe that it is possible for a single, unified, user-services organization to provide service for users of supercomputer facilities in a more effective manner while using fewer resources. Such an unified facility could improve site-to-site communication while simultaneously lobbying the supercomputer site to improve uniformity where appropriate. Increased uniformity among the sites would benefit users substantially. A user would

need to learn fewer site specific idiosyncrasies. Supercomputer facilities, too, would benefit. They would increase their potential base of users.

User-services staff at the various supercomputer facilities have very similar functions. Though the particular services may vary from site to site, core services are available at each site. The most important core service, answering user questions, has a component that is not site specific. Many questions do not require site specific information for the consultants to provide useful answers. The various outreach programs in place at some supercomputer sites already have introduced a layered user-services organization, at least for remote-users. Aspects of such programs can help in structuring a single-point-of-contact user-services program to serve users uniformly.

7. RECOMMENDATIONS

We believe a pilot program should be started immediately. This pilot program should involve a minimum of two supercomputer sites, each with substantial remote user populations. We envision the new user-services organization as supported by grant funds, augmenting the user-services staff at the chosen supercomputer sites. The new organization would serve as an user-services front-line. Its staff would receive and process all incoming telephone calls and mail, answer those questions within their capabilities, and transfer the remainder to an appropriate person within the various supercomputer user-service organizations.

We believe such a pilot program, working with a representative user community, will provide insight into the feasibility of such an organization and will also illuminate problem areas that must be addressed before expanding to include other supercomputer centers.

The pilot program should:

- Develop a mechanism for improving inter-site communication,

- Establish a working group to develops a set of standards for user-service support,

- Establish appropriate communications procedures for use by the target user community in contacting the user-services organization,

- Identify the role of part-time consultants and incorporate them into the program,

- Evaluate the role of PSCN and the evolving National Research and Education Network as the supporting element in a distributed user-services facility.

- Install appropriate hardware and software at each user site to, at a minimum, provide phone service, electronic mail, and possibly screen capture facilities,

- Establish an evaluation team with expertise in human factors, computer technologies, and organization effectiveness and efficiency,

- Develop a means for inserting new concepts, testing those concepts, and reviewing their effectiveness.

Initially, the pilot facility should operate with currently available technologies: 800 phone numbers, computer mail, and screen capture capabilities. The program should also evaluate, develop, and insert enhancing technologies into the user-service function when they become sufficiently useful. At this time, the most promising technologies include collaboration and multi-media technologies for expanding the range of user-consultant interactions and hypermedia technologies for enhancing user access to documentation.

Government and industry are already funding work in these three relevant areas: collaboration technologies, hypermedia, and multi-media communication. NSF has a major new initiative, the National Collaboratory, to explore the use of collaboration technology in support of scientific research. Industry, too, has shown considerable interest in collaboration technologies. In attempting to overcome the lack of spontaneous interactions among remote colleagues, the Xerox corporation is experimenting with an "electronic hallway." This system allows its user to take a "stroll" through the "hallway," look into those offices with "open doors" and possibly start conversations. DARPA is funding ongoing work in hypermedia and multi-media technologies. The commercial sector has developed hypermedia products and continuing to develop such products.

This research is laying the foundation for developing applications in many different areas. However, few funding sources are targeting these technologies towards solving the problems related to user-services.

We see a need for a systematic approach to the application of these technological advances to user services and the development of an appropriate organizational structure for their implementation within the US.