LBNL/PUB-5465, Vol. 1

Strategic Proposal

National Energy Research Scientific Computing Center

FY2002–FY2006

Ernest Orlando Lawrence
Berkeley National Laboratory

December 21, 2001

NERSC
NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER

BERKELEY LAB

LBNL/PUB-5465 Vol. 1 Rev. 1

Strategic Proposal for the National Energy Research Scientific Computing Center

FY2002 – FY2006

Lawrence Berkeley National Laboratory

Office of Science

U.S. Department of Energy

December 21, 2001

Executive Summary

NERSC, the National Energy Research Scientific Computing Center, is DOE's premier scientific computing facility for unclassified research. Over the last five years, NERSC — located at the Ernest Orlando Lawrence Berkeley National Laboratory — has built an outstanding reputation for providing both high-end computer systems and comprehensive scientific client services. At the same time, NERSC has successfully managed the transition for its users from a vector-parallel to a massively parallel computing environment. Building on a foundation of past successes, this strategic proposal presents NERSC's vision for its activities and new directions over the next five years. NERSC's continuing commitment to providing *high-end systems* and *comprehensive scientific support* for its users will be enhanced, and these activities will be augmented by two new strategic thrusts.

The first new component of NERSC will be the provision of intensive support for *Scientific Challenge Teams* funded by the new DOE Office of Science (SC) program called Scientific Discovery through Advanced Computing (SciDAC). DOE envisions that these large-scale teams will be formed to develop and deploy advanced modeling and simulation codes, as well as new mathematical models and computational methods that take full advantage of the new generation of terascale computers. These teams are representative of a shift from the single-principal-investigator model for high-end computing to a collaborative model aimed at producing "community codes" whose development is shared by entire scientific research communities.

A second new component of the NERSC strategy addresses another change in the practice of scientific computing. In recent years rapid increases in available networking bandwidth, combined with continuing increases in computer performance, and the increased use of large-scale, community data archives, are making possible an unprecedented *simultaneous* integration of computational simulation with theory and experiment. This change will have a fundamental impact on areas of science that have not yet made much use of high-end computing. By deploying critical parts of a *Unified Science Environment (USE),* NERSC anticipates playing a key role in the emergence of a new paradigm in computational science.

Herein NERSC presents a strategic proposal, together with a detailed implementation plan, for the years 2002–2006. NERSC proposes a strategy consisting of four components. The two ongoing components are:

- *High-End Systems* — NERSC will continue to focus on balanced introduction of the best new technologies for complete computational and storage systems, coupled with the advanced development activities necessary to wisely incorporate these new technologies.
- *Comprehensive Scientific Support* — NERSC will continue to provide the entire range of support activities, from high-quality operations and client services to direct collaborative scientific support, to enable a broad range of scientists to effectively use the NERSC systems in their research.

The new components are:

- *Support for Scientific Challenge Teams* — NERSC will concentrate its resources on supporting these teams, with the goal of bridging the software gap between currently achievable and peak performance on the new terascale platforms. This goal is explicitly stated in the SciDAC plan.
- *Unified Science Environment (USE)* — NERSC will enhance its architecture and systems as required to make NERSC the most powerful computational and data resource on DOE's Science Grid. Over the next five years, NERSC will use Grid technology to deploy a capability designed to meet the needs of an integrated science environment, combining experiment, simulation, and theory by

facilitating access to computing and data resources, as well as to large DOE experimental instruments.

Finally, NERSC will expand its collaborations with other institutions, especially with the other DOE SC laboratories, to systematically integrate into its offerings the products of their efforts in computational science. With this strategy NERSC will enhance its successful role as a center that bridges the gap between advanced development in computer science and mathematics on one hand, and scientific research in the physical, chemical, biological, and earth sciences on the other. Implementing this strategy will position NERSC to continue to enhance the scientific productivity of the DOE SC community, and to be an indispensable tool for scientific discovery.

Table of Contents

1. Introduction to the NERSC Strategic Proposal

NERSC, the National Energy Research Scientific Computing Center, is DOE's premier scientific computing facility for unclassified research. Located at the Ernest Orlando Lawrence Berkeley National Laboratory (LBNL), NERSC delivers high-end capability computing services and support to the entire DOE Office of Science (SC) research community. NERSC provides these services to the DOE community in conjunction with ESnet, the Energy Sciences Network, which provides NERSC with high-quality, high-bandwidth connectivity to the other DOE laboratories and major universities. This strategic proposal presents NERSC's vision for its activities and new directions over the next five years.

1.1 Building NERSC's Future Strategy on Its Record of Scientific Success

NERSC's strategy for its next five years is built on its record of success in enabling the scientific achievements of its clients, which are among DOE's most visible accomplishments. A small selection of recent successes is included in the Implementation Plan. These examples make it clear that the comprehensive scientific support provided by the Center is a key to the success of the scientific research of each user. NERSC's user orientation pervades the organization and is the unifying principle that underlies its activities.

NERSC's role as a general DOE facility for high-end scientific computing requires the Center to provide resources valuable to all SC programs, but NERSC must also respond to the specific needs of individual programs. Specific support for different programs, tailored to their varying needs, has been a key to the success of the Center. Indeed, as Figure 1-1 shows, the visible scientific accomplishments of NERSC clients span the range of DOE/SC programs.

The DOE/SC strategy of maintaining a flagship supercomputer facility that serves all its programs provides the best possible mechanism for technology transfer between the programs and projects. In addition, because the expertise developed in supporting one sector ultimately aids all sectors, NERSC's centralized operation provides maximum flexibility to DOE/SC: NERSC's resources can be quickly redirected if necessary to accommodate changing priorities in the DOE/SC community.

1.2 NERSCs Strategy Is Built on DOE Requirements for High-End Computing

NERSC constantly collects the requirements and needs of its clients and the DOE programs that fund them, through a number of formal and informal mechanisms. First, the NERSC User Group (NUG) meets twice a year with the NERSC staff to discuss both short-term and long-term needs. NERSC and NUG have monthly teleconferences to address short-term issues. NUG produces a new "Green Book" of user requirements independently every three or four years to provide the basis for major (and minor) system acquisitions, as well as the Center's services. NERSC conducts an annual user survey and publishes a self-evaluation based on its goals and metrics. In addition, NERSC staff interact frequently with the program officers in the principal offices — Biological and Environmental Research, Basic Energy Sciences, Fusion Energy Sciences, and High Energy and Nuclear Physics — as well as with the program officers of NERSC's program office, Advanced Scientific Computing Research. In this way NERSC learns about new directions in the programs, such as the more intimate coupling of experiment and high-end computing, in their early stages, allowing adequate time to chart the best implementation.

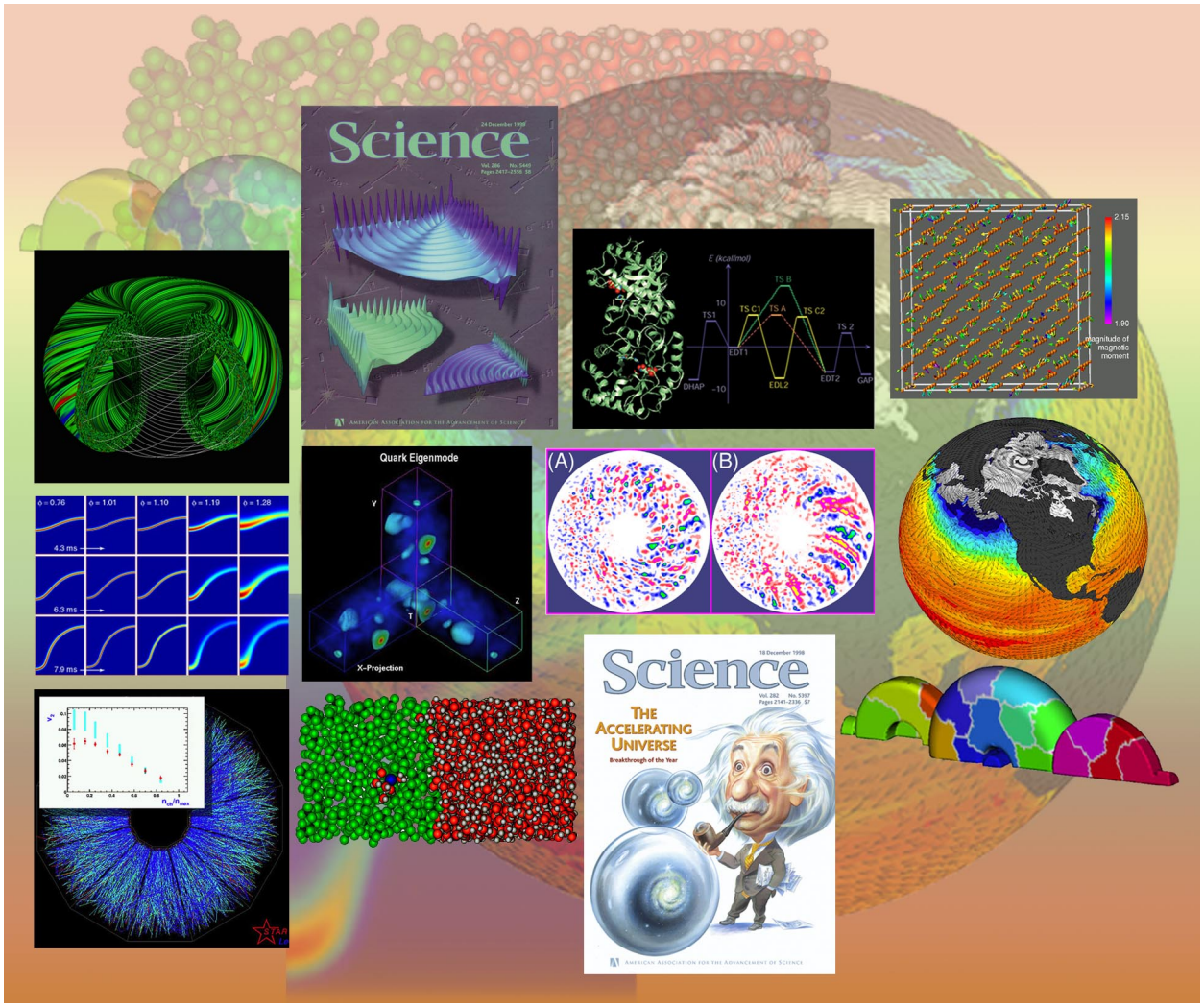

Figure 1-1. Well-recognized scientific accomplishments of NERSC's clients are representative of the Office of Science programs. (See the Implementation Plan for details on these accomplishments.)

One example of a changing DOE/SC priority requiring NERSC action is SC's recently funded initiative, Scientific Discovery Through Advanced Computing (SciDAC). The plan for the initiative, originally distributed in March 2000, envisions funding a small number of "Scientific Challenge Teams" (national consortia engaged in a community code-building effort for a particular discipline), as well as some "Integrated Software Infrastructure Centers" (to provide new computer science tools for the Scientific Challenge Teams). The success of this initiative will require that NERSC support a number of these teams with the most powerful computer resources available, as well as the intensive and flexible client support required to make high-end systems effective for these teams.

1.3 NERSC's Strategy Addresses Emerging Changes in DOE/SC Computational Science

In the 1980s physicist Ken Wilson and others described computing as the new "third component" of the scientific method, achieving parity with experiment and theory. There is abundant evidence that this vision of the critical role of computing has been realized in DOE/SC programs. At the present time, three additional trends are apparent: (1) the emergence of large, multidisciplinary teams for scientific research, (2) the convergence of computation, experiment, and theory on an interactive, real-time basis, and (3) the

importance of data in archives and repositories scattered around the world. These changes will form the basis of new directions and priorities for NERSC.

The first of these trends, typified by the SciDAC Scientific Challenge Teams, represents a break from the single-principal-investigator (PI) model for computational science that has dominated computing in many areas of the natural sciences for most of the last 50 years. The emergence of these teams is primarily motivated by the increased complexity of the science being investigated, but also by the need to close the growing gap between the sustained performance obtained by straightforward porting of single-processor codes and the peak performance capability of these systems. On vector supercomputers of the 1980s and early 1990s, many scientific codes realized 30% to 50% of peak performance. By contrast, scientific codes typically realize only 5% to 15% of peak performance on massively parallel supercomputers today. It is now clear that exploiting the full potential of massively parallel computers as scientific tools will require much greater effort and a wider range of skills than can be brought to bear by a single PI working with a few collaborators, trained in a single scientific discipline.

The second change that is occurring in SC's portfolio is the simultaneous convergence of computing, experiment, and theory in scientific research. One example, familiar at NERSC because the Center has been supporting it since 1996, is the Supernova Cosmology Project. To find Type 1A supernovae, telescope images are analyzed using supercomputers, and the results are used to quickly direct telescopes around the world to observe and measure spectra of candidate objects. Simulations and predictions of the spectra based on the abundance of each elemental candidate are performed simultaneously to refine the identification. The entire effort requires the availability and scheduling of resources throughout the world.

The third change is a reflection of the first two. Large collaborations involve many institutions, each of which may specialize in the generation and management of parts of the data needed for the overall collaboration, and researchers today are more likely to use multiple data sources in their studies. This is the motivation, for example, for the National Virtual Observatory (NVO) project, which is developing the architecture for an Internet portal that will link all the major archives of astronomy data in the United States, along with the computational resources necessary to support comparison and cross-correlation among these resources. NERSC is storing data for the Supernova Cosmology Project that could advantageously be integrated into the NVO.

It was for such applications that the concept of Computational and Data Grids (discussed in detail in Section 5 of this proposal) was invented. For DOE/SC to be able to exploit this technology, NERSC, DOE's largest unclassified computing facility, must become an active node on the DOE Science Grid. Such a transformation is not accomplished by merely connecting to the Grid: it must be accomplished by the staged development of a new architecture for high-end computing that incorporates, schedules, and manages the computing and data resources at NERSC and elsewhere into a new and evolving national infrastructure.

1.4 Principal Components of the NERSC Strategic Proposal

NERSC's goal for the next five years is to address these changes in the context and requirements of the NERSC user spectrum, while continuing to strengthen NERSC's user-oriented approach. To accomplish this goal, this proposal focuses on four components of the NERSC strategy, two ongoing and two new (see Figure 1-2). The two ongoing components are:

- *High-End Systems* — NERSC will continue to focus on balanced introduction of the best new technologies for complete computational and storage systems, coupled with the advanced development activities necessary to wisely incorporate these new technologies.
- *Comprehensive Scientific Support* — NERSC will continue to provide the entire range of support activities, from high-quality operations and client services to direct collaborative scientific support, to enable a broad range of scientists to effectively use the NERSC systems in their research.

Figure 1-2. The four principal components of the next-generation NERSC are designed to serve the DOE science community.

The new components are:

- *Support for Scientific Challenge Teams* — NERSC will concentrate its resources on supporting these teams, with the goal of bridging the software gap between currently achievable and peak performance on the new terascale platforms. This goal is explicitly stated in the SciDAC plan.
- *Unified Science Environment (USE)* — NERSC will enhance its architecture and systems with the goal to make NERSC the most powerful computational and data resource on DOE's Science Grid. Over the next five years, NERSC will use Grid technology to deploy a capability designed to meet the needs of an integrated science environment, combining experiment, simulation, and theory by

facilitating access to computing and data resources, as well as to large DOE experimental instruments.

In order to better align the components of this proposal with DOE ASCR programmatic directions, we distinguish between three different levels for budgeting purposes:

- *NERSC Base Program* — The base program includes all systems and support for the base users, founded on the High-End Systems and Comprehensive Scientific Support components. It also provides specialized support for up to five special projects like the Scientific Challenge Teams. And it includes the deployment of basic Grid services, primarily for handling user data. The Grid-related effort in the base program will be restricted to integrating production-quality tools and middleware developed by the DOE Science Grid or by vendors which is ready for general introduction.

- *SciDAC Scientific Challenge Team Support* — This level contains all the special support for the Scientific Challenge Teams beyond the five special projects covered in the base program. All support for additional teams and the additional infrastructure support is included here.

- *USE Support* — This level contains all the additional activities required for an accelerated introduction of Grid services beyond what is in the base program.

The four components will be developed in the next four sections of this plan. Sections 2 and 3 describe the strategy for providing high-end computational services in the 21st century. Section 4 describes how NERSC will adapt the Center to support large Scientific Challenge Teams. Section 5 describes how NERSC proposes to become a resource for a new class of DOE science endeavors using Grid and collaborative technologies. A final section describes NERSC's relationship to other DOE projects and facilities, the proposed method of measuring the success of this strategic proposal, the milestone schedule, and the Base Center budget summary.

2. High-End Systems

Providing the most effective and most powerful high-end systems possible. This is the foundation upon which NERSC builds all other services in order to enable computational science for the DOE/SC community. High-end systems at NERSC mean more than highly parallel computing platforms — they also include a very large-scale archival storage system, auxiliary and developmental platforms, networking and infrastructure technology, system software, productivity tools for clients, and applications software. Our successful high-end system strategy includes advanced development work, evaluating new technologies, developing methodologies for benchmarking and performance evaluation, and acquisition of new systems.

NERSC plans to introduce the NERSC-4 system in 2003 and the NERSC-5 system in 2006, each with a three- to four-fold increase in capability over the baseline previous-generation system, bringing NERSC-5 to approximately 30 teraflop/s peak performance. Figure 2-1 shows the proposed peak computing power of NERSC. But computing power is only one measure of capability. NERSC will continue to increase the capacity of its storage system, reaching at least 15 Petabytes in 2006. A major effort will be made in developing and deploying a Global Unified Parallel File System (GUPFS). The development of GUPFS will take existing work as its point of departure, including the Global File System work initially developed at the University of Minnesota and being developed further at Sistina Corporation, developments within ASCI Pathforward in high-performance computer file system technology, and the IBM Global Parallel File System. This capability will not only increase scientific productivity by simplifying file access, but will also support Grid and archive storage improvements, which are described in Section 5.

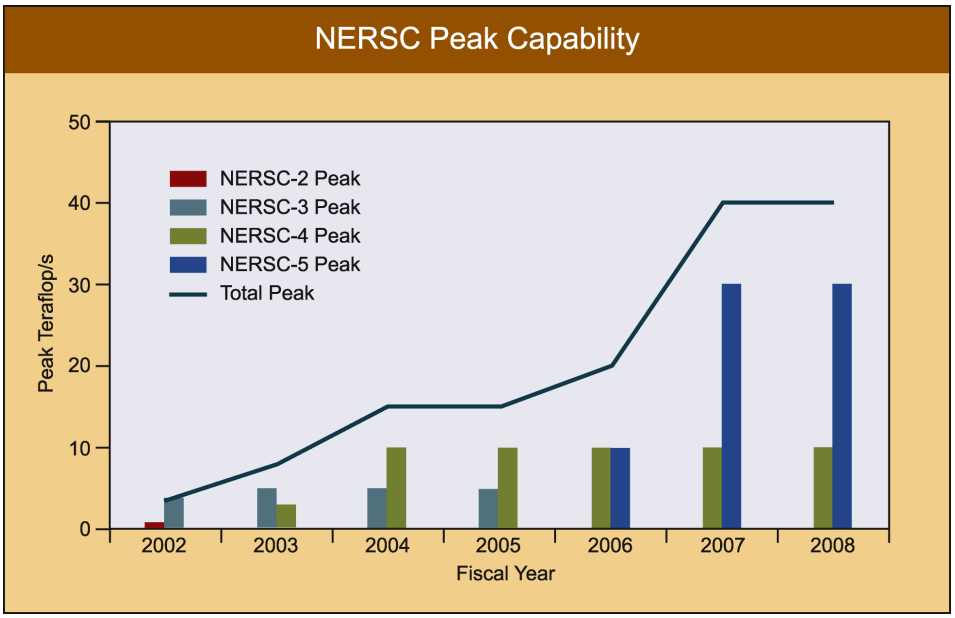

Figure 2-1. Computational capability growth of the NERSC Facility

Details of NERSC's strategy in high-end systems will be given in the following subsections. First is a description of the rapidly evolving technology base from which NERSC systems must be procured. This

is followed by a discussion of client requirements and system balance. With these established, NERSC then presents its strategy for acquiring and managing large-scale systems. This section then concludes with a discussion of the critical role advanced development activities take in defining NERSC's future.

2.1 Technology Changes

It is essential that NERSC continuously evaluate technology to determine what can best solve the problems of its clients. Key technology areas relevant to NERSC include supercomputer architectures, data communications technology, online disk storage, archival storage, and application development software.

Computing system architecture. From a systems architecture standpoint, there are two main choices for computational systems during the period of this proposal: clusters of symmetric multiprocessors (SMPs), and special architectures. Special architectures include Cray's MTA and SV2, IBM's Blue Gene and Blue Light, and academic research products such as UC Berkeley's iRAM. NERSC tracks the progress of special systems and evaluates their applicability to the DOE computational science challenges as part of our advanced development activity. NERSC staff members participate in design reviews and advisory boards of new special architecture projects, and in early evaluation efforts. The output from this technology assessment provides the basis for NERSC's decisions on new technology directions for its high-end systems.

From our current vantage point, the system architecture most likely to be chosen during the period of this proposal is a clustered SMP. This architecture class includes a broad set of systems:

- *Commercial, integrated SMP cluster systems*. Typified by IBM's SP and Compaq's Alpha servers, these systems employ commodity components for processor-memory nodes, but interconnect the nodes with proprietary, tightly integrated communications switches. They employ workstation-class CPUs and run proprietary UNIX-based operating systems.
- *Cluster systems made up of workstation-class components* (CPUs, memory, standard network interconnects). These systems may run vendor-proprietary UNIX or open-source Linux operating systems.
- *Completely commodity cluster systems*. These systems employ Intel or Alpha processors, standard networking such as Gigabit Ethernet for internode communications, and open-source software such as Linux. All the system software on these systems is either open-source or third-party software that runs across multiple platforms.

In all probability, the parallel message-passing workload, which constitutes a large fraction of the current workload on NERSC-3, can best be supported at the multi-Tflop/s scale by deploying a proprietary, commercial SMP cluster for NERSC-4. This may still be true when NERSC-5 is introduced.

Commodity clusters show great promise. They are not ready for full-scale NERSC usage yet, but they are getting closer. They are already in use at various sites (including NERSC) for applications such as experimental data processing, and they may soon be suitable to supplant vector systems for moderate-sized workloads. Significant advances in cluster software are expected over the next four to five years, in part due to the investments being made by the National Science Foundation in the Distributed Terascale Facility (DTF). Unfortunately, commodity cluster software does not appear to be targeted for use on high performance computing (HPC) systems. A commitment to HPC commodity clusters is lacking in the general open-source development community, with the notable exception of the DTF and possibly the Extreme Linux activity. There is greater commitment among certain system vendors, but their solutions

are still in development. Thus in order for NERSC to deliver a high-quality, capability-oriented production environment using commodity clusters, NERSC must participate actively in the development of cluster software for supercomputing, and must be prepared to integrate that software into the appropriate systems.

Data communications. Communication bandwidth is exploding, doubling in performance/price ratio every 9–12 months. This is due mainly to the incorporation of erbium-doped optical amplifiers, which permit multiple wavelengths to be carried on a single fiber. However, the end components, namely routers and switches, are advancing at only a Moore's Law rate, i.e., doubling every 18–24 months. This means the routers and switches will constitute an increasing fraction of costs when NERSC and other sites improve services and expand capacity. The border between the wide-area and the local network will remain problematic for some time. This is where transitions occur between media and technologies, as well as where monitoring and security protection are implemented, all of which put constraints on performance. Therefore, in the next several years NERSC will need to focus on the border between the wide-area network and the local network. ESnet provides NERSC's high-quality, high-bandwidth connectivity to the other DOE laboratories and major universities. NERSC needs ESnet to provide faster and more capable services to assure that the wide-area connectivity balances with the capability of NERSC's computational and data resources.

Disk storage. Storage capacity is increasing by a factor of 4 every three years, whereas storage bandwidth is increasing by a factor of 2 every ten years. This means that file systems will soon be capable of holding tens to hundreds of terabytes, but getting the data between the computational components and the media will be increasingly challenging. Storage area networks (SANs) are important new system components, but SANs have not yet shown the ability to integrate with heterogeneous systems in a high-speed manner. Several solutions are possible for integrating physical media into a parallel file system that spans multiple systems, but at the present time all of these solutions exhibit some performance limitations.

Archival storage. The market forces in archival storage are different, tending to push technology that has large jumps in performance but with less frequent cycles. This is due in part to the need to maintain data integrity and to convert media. While increases in disk capacity make it possible to expand online storage, it is unlikely that tape media will be replaced totally by disks in this time frame. In the petabyte range, it is not cost effective to have all the data on line all the time. Hence, NERSC will continue to operate robotic storage systems, able to use the form factor of the 3480 tape drive (even if inside is a set of small disk drives), while introducing more capacity to higher bandwidth. In any event, we will closely monitor this arena for developments that may suggest a change of strategy.

Application development software. With the increasing complexity of large-scale applications, software facilities such as parallel programming environments, debuggers, performance tools, and code validity checkers will become increasingly essential to our clients. NERSC already supports some commercial-grade application development software, such as the Etnus TotalView debugger. As other tools of this general type become available, NERSC will assess their utility and maturity, and then decide the appropriate time to offer them as supported software for our clients.

2.2 Client Requirements and System Balance

It is essential that NERSC understand the requirements, both near-term and long-term, of the computational scientists using its systems. These requirements come from DOE strategic and tactical programmatic goals, from leading-edge research teams, and, most importantly, from a formal process led by the NERSC User Group (NUG). This client-driven process produces the "Green Book," updated

every three to four years, which documents the requirements for the Center. Several disciplinary experts, who catalog future project needs, represent each DOE office. NUG then assembles and prioritizes these requirements. DOE strategic thrusts, such as SciDAC and the Grand Challenge program, are also carefully considered when planning future purchases, as they define in large measure the strategic applications areas for DOE. As such, NERSC will pay particular attention to the applications and methods used by the Scientific Challenge Teams working on Strategic Projects. Strategic Projects will be identified by the DOE and will emerge from SciDAC Teams, large-scale (Class A) projects in the base allocation, and other defined teams working at the very highest limits of scientific exploration and computational parallelism. A more detailed description of the latest Green Book requirements is provided in the Implementation Plan.

The latest Green Book argues that the DOE/SC community must increase computational resources capable of supporting very large-scale applications, since this remains the primary requirement for the scientific projects. Most projects anticipate substantial online storage requirements, with high bandwidth to the computational platforms. They also require higher network bandwidth to their sites. Some projects have very large data archive requirements, while others need large amounts of memory for their computational applications. Finally, all clients and projects want access to continued and expanded scientific support.

NERSC must support a diverse workload. NERSC's highest priority is to support *capability computing*, which we define here as the need to use more than one-fourth of an entire computing resource over an extended time period. NERSC also supports *large-scale computing*, which is defined as the use of more than one-eighth of the entire resource over an extended time period. Finally, NERSC supports a small amount of *related capacity computing*, which is comparable to running on a desktop system for a week. NERSC's systems and service architectures are currently designed to support a capability workload of up to five Strategic Team projects and about 20 large-scale, Class A projects. NERSC also supports 75 to 100 smaller, Class B projects, and 50 startup projects. We anticipate the Strategic Teams will use the majority of the system resources and also deploy some of the largest applications running on the systems.

2.3 Strategy for Systems and Major Milestones

There are three major areas of system design and implementation at NERSC: the computational systems, the storage system, and the network. The balance of the entire Center is determined by the requirements that evolve from the increased computational capability, plus independent requirements for other resources. New storage systems must be designed to support not just current work, but future workloads as well. Figure 2-2 shows the evolution of the NERSC system architecture between 2001 (left) and 2006 (right), with the introduction of the Global Unified Parallel File System and the Unified Science Environment integrating the discrete computational and storage systems. The following paragraphs provide detailed descriptions of NERSC's strategy for its computational systems, storage systems, and network.

2.3.1 Computational Systems Strategy

NERSC will acquire a new capability-focused computational system every three years. A three-year interval is based on the length of time it takes to introduce large systems, the length of time it takes for NERSC clients to become productive on new systems, and the types of funding and financial arrangements NERSC uses. At any one time, NERSC will have two generations of computational systems in service, so that each system will have a lifetime of five to six years. This overlap provides time for NERSC clients to move from one generation to the next, and provides NERSC with the ability to fully test, integrate, and evolve the latest generation while maintaining service on the earlier generation.

For NERSC-4 and NERSC-5, which are the two generations of new computational systems in this proposal's timeframe, we expect a factor of 4 increase in peak capability over the previous generation. These systems will be selected based on several criteria, including a good system balance, a good programming environment, effective system management tools, and — most importantly — useful sustained performance. We expect that these systems will be early-delivery, low-serial-number systems, since these provide NERSC clients with the most advanced functionality and performance at the earliest time, thus maximizing potential scientific impact. By targeting early-introduction production systems, NERSC rides the crest of Moore's Law.

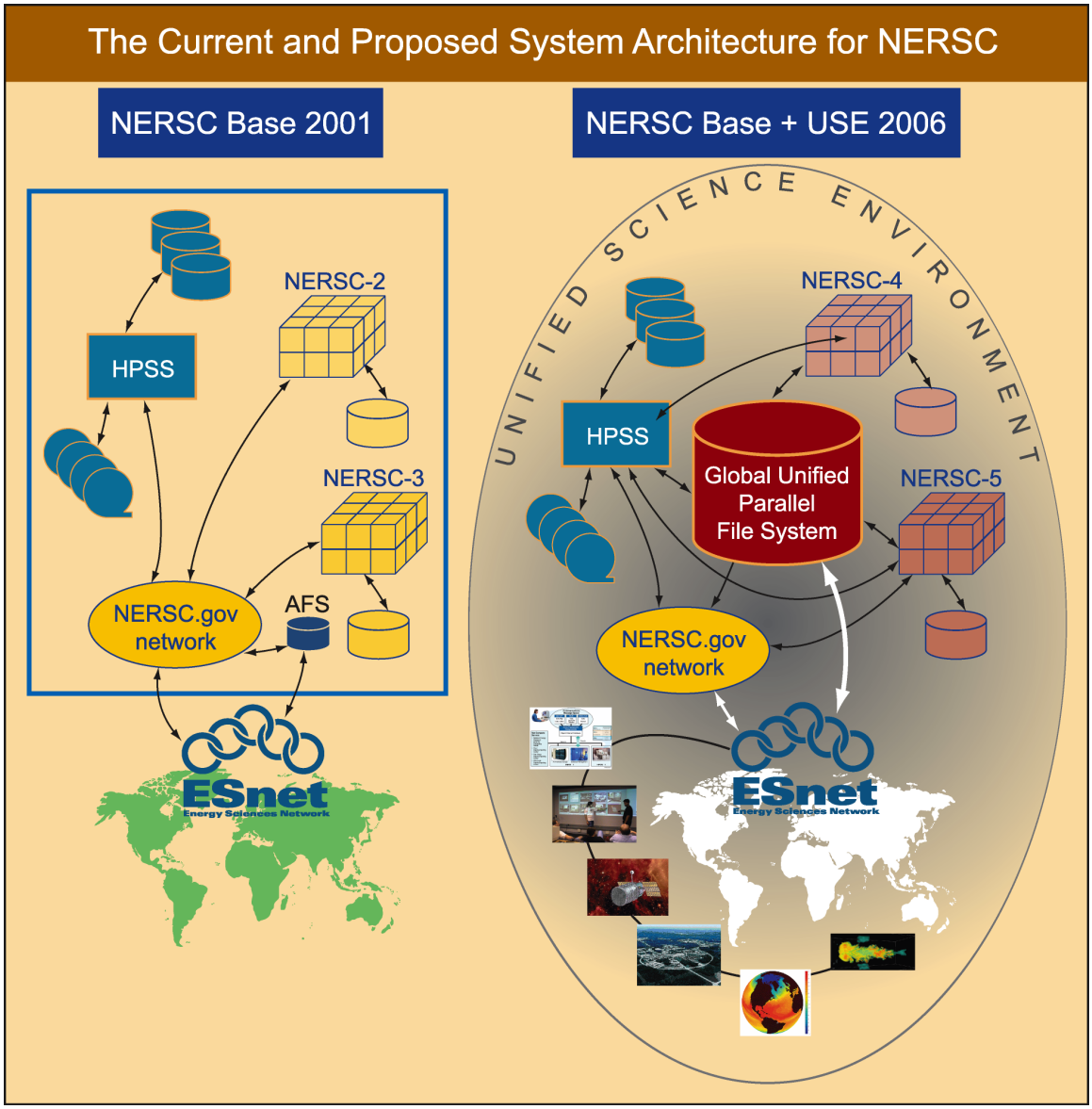

Figure 2-2. Evolution of the NERSC system architecture between 2001 (left) and 2006 (right).

The total annual investment in the computational supercomputer systems alone will be approximately one-quarter to one-third of the total NERSC annual funding. As in the past, lease-to-own payments will

be spread over three years, and it is possible that technology availability will dictate a phased introduction over one year to 18 months.

As mentioned above, we expect that NERSC-4 and NERSC-5 will very likely be commercial integrated SMP cluster systems. Special architectures will be considered, but it is not likely that these will be ready for high-quality production usage in the proposal's time frame. For example, projects such as Blue Gene and Blue Light will be available (if at all) only after 2005, since their first full-scale systems are slated for completion after 2004. Commodity cluster systems will also be considered, but based on our technology assessments, we do not believe it likely that these systems will be able to support the diverse and communication-intense applications at NERSC in this time frame. Cluster hardware will at best have a modest performance-per-dollar advantage, but cluster software in particular is significantly less mature than vendor-supplied software. This situation is expected to persist for the next 3 to 5 years. If NERSC were to procure a large cluster system, we would have to allocate significantly more of our personnel resources on system integration and support work than proposed.

NERSC will use the "best value" process for procuring its major systems. This 22-step process (described in the Implementation Plan) allows considerable flexibility for NERSC and also provides an opportunity for significant innovation by suppliers. The principal task of the acquisition team is to decide the best alternative among the available choices. One key metric we use is what we call the Sustained System Performance (SSP) metric, which is based on a benchmark performance integrated over three years. We will use the Effective System Performance (ESP) test to assess system-level efficiency, namely the ability of the large-scale system to deliver a large fraction of its potential resources to the users. In addition, NERSC plans to use the outcome of the SciDAC Performance and Evaluation Resource Center (PERC) and use the new *de facto* set of benchmark kernels being developed by that project.

2.3.2 Storage System Strategy

Over the time period covered by this proposal (2001 to 2006), NERSC plans to augment both the aggregate capacity and the transfer rate to and from the mass storage system. The aggregate capacity will increase more then tenfold, from two petabytes today to over 15 petabytes in 2006. Sustained data flow will also increase more than tenfold, from approximately 1.5 terabytes per day today to over 20 terabytes per day in 2006. Figure 2-3 shows NERSC's predicted growth in both capacity and transfer rate for the period covered by this proposal.

NERSC will continue collaborating in High Performance Storage System (HPSS) development, in order to improve archive technology. In particular, NERSC will help develop schemes to replicate data over long distances and to import and export data efficiently. The archive system, and all data resources at NERSC, will be the first focus of any Grid and USE activities within the base funding.

Shared storage will increase in importance for NERSC clients, as more and wider collaborations require transparent access to a common code base, data, and metadata. The inability to use shared storage in this environment results in higher costs, wasted resources, and reduced productivity for NERSC clients. In the current NERSC environment, sharing storage between the large systems cannot be done effectively, because

1. the large computational systems are not well suited to being storage servers

2. the current mechanisms for sharing storage (NFS, AFS, DFS) cannot deliver adequate high performance I/O bandwidth with low latency and low overhead.

NERSC will address this deficiency by implementing a Global Unified Parallel File System (GUPFS), which provides scalable high performance and high bandwidth to local systems as well as to the USE. The system will have a global, facility-wide file namespace. It will use storage area network (SAN) online data, integrated with HPSS. Using GUPFS in the NERSC environment will:

- Eliminate the wasteful replication of online files and inefficient transfer between systems.
- Allow effective use of disk storage by all connected systems through the aggregation of storage resources.
- Maintain the separation of storage from computational resources, which means simpler, independent administration and upgrades of the storage resources.
- Simplify and increase the usability of the programming environment by providing a uniform shared name space for user files.
- Simplify access to data and regulation of permissions.
- Provide improved ways to interface with archive storage.

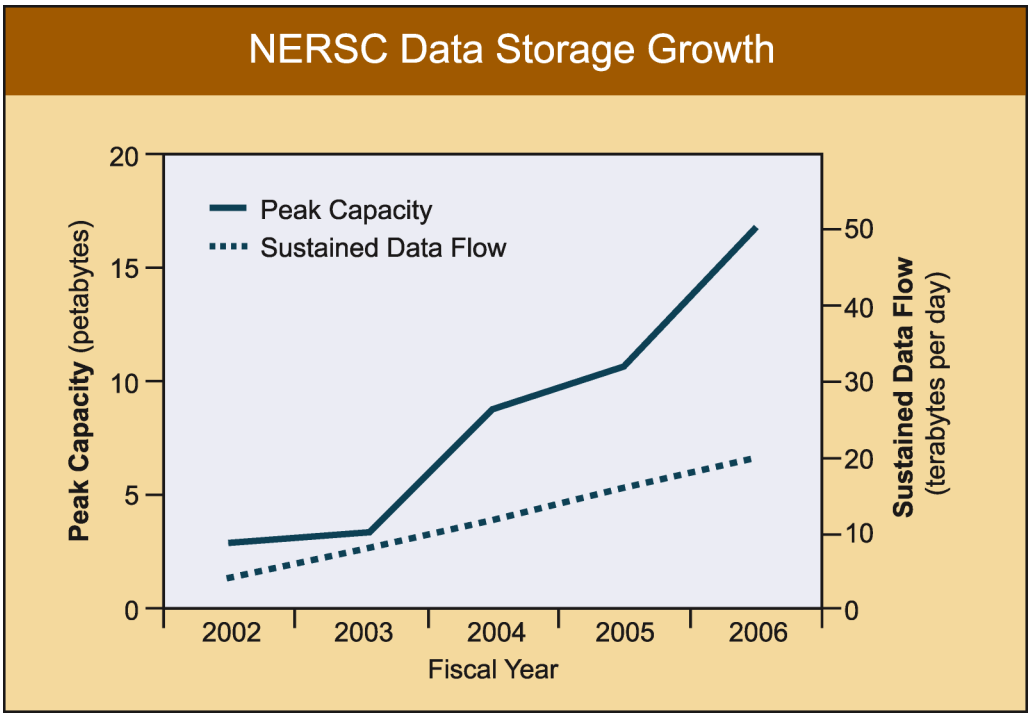

Figure 2-3. Predicted NERSC data storage growth, 2001–2006.

The USE environment, described in more detail in Section 5, will play an important role in the ability of researchers and workers to access and manipulate data in an efficient manner. NERSC has a long history of helping to develop and field advanced development software. In particular, NERSC is expert in testing early releases of software and providing important feedback to the development staff. NERSC is able to do this testing at a scale that is unique. Possibly more important is NERSC's expertise in applying

scientific application workloads to the software. It is in this manner that NERSC will help accelerate the deployment of Grid software.

2.3.3 Networking and Data Communications Strategy

NERSC must expand its networking and data communication capacity, as applications become more bandwidth intensive. The USE paradigm will change the common practice from bulk data transfers with some TCP/IP interactive traffic to many more bulk data transfers combined with inter-process communications. Since higher production bandwidth assumes bandwidth doubling at least every 18 months, and since clients indicate the need, NERSC will move to a new OC bandwidth level every three years in its production network.

Network tuning is increasingly critical to providing high performance network connectivity to NERSC clients. NERSC will continue to provide this service to NERSC clients, in particular the Scientific Challenge Teams. In collaboration with ESnet and other sites, NERSC will be active in deploying the latest enhancements in local and wide-area networking systems and protocols, such as 10 Gigabit Ethernet and Infiniband, to enable NERSC clients to move data and to provide system access for enhanced services such as remote steering, Grids, and visualization. NERSC will increase its emphasis on eliminating bottlenecks in the WAN-LAN boundary, since this is the key to success in many bandwidth-intensive applications.

As high-performance computing becomes more network-centric (the Grid, HPSS, cluster interconnects, etc.), the network will become the "glue" that holds everything together. NERSC must become a center of excellence in network engineering; this is the only way we will be able to deliver the full capability of our systems to our users.

2.4 Advanced Development

A vital advanced development activity has been a cornerstone of NERSC throughout its 25-year history. This is because NERSC has always attempted to deploy the most capable computing systems available, and these are often new and relatively immature systems. For example, when NERSC transitioned from the Cray C90 vector mainframe to the Cray T3E distributed shared memory system, system tools were very immature, and utilization hovered below 60%. NERSC responded by working with the vendor to develop checkpoint/restart capability and a new job-scheduling heuristic called "political scheduling." We were the first to deploy this new software, and it has largely been responsible for NERSC's current ability to sustain 95% utilization. We are now able to measure the impact of such enhancements by using the SSP and ESP metrics, which were developed in the course of the NERSC-3 procurement.

NERSC's development activities are directly motivated by NERSC's need to adequately support its DOE clients. The NERSC advanced development activity attempts to answer questions such as:

- What will NERSC's future capability systems be?
- Can we deploy a global parallel file system?
- What is the future of mass storage?
- What is the right cyber-security model?
- How do we integrate with the DOE Science Grid?
- What will future program development and performance-tuning software be like?
- How does NERSC support the emerging scientific software teams beyond bits, bytes, and flops?

Fortunately, we do not face such challenges alone. LBNL hosts ESnet and a significant portion of the DOE/SC Mathematical, Information, and Computational Sciences Division (MICS) mathematical and computing science research. These researchers can be called upon to help when needed. NERSC has also established partnerships with most of the other SC laboratories, in particular Argonne (ANL), Oak Ridge (ORNL), Brookhaven (BNL), and Pacific Northwest (PNNL) national laboratories. We coordinate our activities with our colleagues at Lawrence Livermore National Laboratory (LLNL), who have very similar computing systems, as well as with the National Science Foundation Partnerships for Advanced Computational Infrastructure (NSF PACI) centers, and the National Aeronautics and Space Administration (NASA) Information Power Grid project. Finally, NERSC looks forward to working with the new SciDAC Integrated Software Infrastructure Centers (ISICs). By leveraging their work where possible, NERSC can focus its own staff on problems that are unique to our mission. Some of these are briefly addressed below.

High-end system assessment. The most significant ongoing issue that NERSC faces is what its future capability computing systems should be. The NERSC advanced development staff evaluates potential systems to measure their range of applicability to DOE SC science, as well as their performance, reliability, and true cost of ownership. We do this by studying information provided by vendors, by evaluating early systems fielded by our colleagues, and by assessing testbed systems deployed at LBNL. In the course of this process, NERSC identifies deficiencies in the systems and communicates these deficiencies to vendors, so that they correct them if possible. In some cases we work with vendors or other sites to rectify deficiencies and increase functionality.

System software development. Looking to the immediate future, NERSC plans to examine the current work in checkpoint/restart for Linux PC clusters, and to port this to the NERSC environment. We will develop modular interfaces to next-generation interconnection networks such as Infiniband. In addition, NERSC will work with other DOE laboratories and the Scalable Systems Software ISIC on cluster management software. NERSC expects to work on GUPFS, based in part on the global file system initially developed at the University of Minnesota, to provide our clients with transparent access to their data from any NERSC system. NERSC is also collaborating with LLNL and Indiana University to integrate HPSS with GPFS.

Benchmarking and performance tuning. Understanding how well systems perform has been a hallmark of NERSC. In order to properly measure the performance of NERSC-3, we developed the SSP and ESP metrics. We are now developing some application benchmarks that characterize how systems perform on DOE/SC applications. In the long run, this will allow NERSC to develop a new suite of benchmarks that will provide a better basis for understanding the true capability of modern supercomputers. NERSC will also support the ISIC activity in high-end system performance by developing new benchmark and modeling methodologies, and by making the performance monitoring and tuning tools developed by this ISIC available to users.

Data archives. NERSC's development activities extend to its data archive as well. As one of the key development sites for HPSS, NERSC is working with the HPSS consortium to develop the capability to import and export tapes directly to and from HPSS. NERSC will deploy distributed archival technology and HPSS resource managers. In the longer term, NERSC will address issues such as whether or not large arrays of disks or perhaps some alternative technology will become a more attractive storage medium than today's implementation of HPSS at NERSC. NERSC will collaborate with the Scientific Data Management ISIC and integrate tools developed there.

Security. Internet security is and will likely remain a major area of ongoing effort at NERSC. As we discuss in Section 5, NERSC plans to integrate its capability systems into the emerging DOE Science

Grid. To adequately support our users in the future, NERSC will need to dramatically increase network bandwidth to and from NERSC. Unfortunately, today's firewalls are too restrictive to be compatible with such a goal. Instead, NERSC's perimeter defense is built around the reactive intrusion detection program called BRO, developed at LBNL. In order to continue using this approach, NERSC needs to increase BRO's ability to handle high-bandwidth networks. It will also need better traffic analysis algorithms, as well as the ability to react on a more autonomous basis. NERSC and BRO will have to learn how to integrate with the Grid's authentication and single sign-on protocols.

Grid software. In order to better support clients who will gain access to its services via the Grid, NERSC will investigate interfacing its queuing and job management systems to tools such as the Globus Resource Manager. It will also investigate portals to allow users real-time control of their jobs. Section 5 provides further details about the role of Grids and methods to accelerate their full deployment at NERSC.

Algorithmic tools. NERSC will continue to develop numerical algorithms, specialized code frameworks, visualization tools, and other scientific code critical to enabling high-priority DOE users to maximize their scientific productivity at NERSC. This latter work often involves long-term, sustained collaborations between NERSC staff members and clients, as described in Section 3. NERSC will also leverage software developed by the Common Component Architecture ISIC where appropriate.

Since long-term advanced development activities will be determined by future needs of NERSC clients, they cannot be fully predicted today. Instead, for the longer term, NERSC will use the following process for making investment decisions in the future. First, NERSC will ask if any of our colleagues are addressing the problem for us? Will a particular advanced development activity address the immediate needs of NERSC or its clients? Will the development activity reduce NERSC's long-term risk (e.g., early evaluation of potential future NERSC systems), or will it provide our DOE clients with a potential glimpse of their future computing platforms? Only when a project is shown to clearly address one of these challenges will NERSC invest its time and resources pursuing it.

In summary, advanced development has been, and will continue to be, critical to enabling NERSC to provide for its DOE users. This is because NERSC operates at the leading edge of systems and services where change will continue to be the norm.

3. Comprehensive Scientific Support

As described in Section 2, NERSC continues to provide early, large-scale production computing and storage capability to the DOE/SC computational science community. The NERSC systems will be of such a scale as to be unique or nearly unique in many aspects (e.g., computational abilities, storage capacity, etc.). The goal of NERSC's Comprehensive Scientific Support function is to make it easy and practical for DOE computational scientists to use the NERSC high-end systems. NERSC continues to reaffirm its commitment to excellent support for its base user community of about 200 projects and more than 2000 users. Activities described in this section are (unless noted otherwise) part of the NERSC base program.

NERSC is planning to provide Comprehensive Scientific Support through:

- providing consistent, high-quality service to the entire NERSC client community through the support of the early, production quality, large-scale capability systems
- aggressively incorporating new technology into the production NERSC facility by working with other organizations, vendors, and contractors to develop, test, install, document, and support new hardware and software
- ensuring that the production systems and services are the highest quality, stable, secure, and replaceable within the constraints of budget and technology
- participating in other work to understand and address the unique issues of using large-scale systems.

Comprehensive Scientific Support spans all the technology areas discussed in Section 2, and enables NERSC to provide exceptional support to its clients. It is the heart of the strategy that sets NERSC apart from other sites and greatly enhances the impact of NERSC's high-end systems. In this section we discuss three aspects of NERSC's Comprehensive Support program:

- robust services and service architecture (base program)
- focus on high-end teams (base program and SciDAC)
- new elements to our program (USE).

3.1 Robust Services and Service Architecture

In order to make the systems available and usable by NERSC clients, a basic set of essential services must be provided. The NERSC support model is expressed by its "service architecture," which defines the roles and interfaces for all the functions that make up Comprehensive Scientific Support. The functions involved are:

- essential services, including system monitoring and support, direct scientific support, and system management
- USE support
- Scientific Challenge Team support.

The essential services are discussed in detail below. The Scientific Challenge Teams and USE are discussed in Sections 4 and 5, respectively.

3.1.1 System Monitoring and Support

System monitoring and operations support all systems on a $24 \times 7 \times 365$ schedule. These tasks involve system monitoring, initial system troubleshooting, system backup, and management of the near-line and off-line storage media. Equally important, NERSC supports productivity-improving tools and techniques designed to automate or improve the operational tasks.

The help desk provides direct client assistance, as well as managing and resolving client problem reports. It is important that the client community be able to ask for assistance in the way most effective for them — not just what is most efficient for NERSC. Thus, NERSC supports telephone, e-mail, and Web interactions with timely acknowledgement and response resolution. The help desk is the first level of the direct scientific support described below, but it plays a key role: once a client reports a problem, NERSC will manage it until it is resolved, not just send the client to another group or have the client manage the problem. The combination of help desk, direct consulting support, account support, and extended support constitutes the basic user services.

NERSC system and storage staff provide basic system administration and remedial maintenance 24 hours per day. Each system has an assigned point of contact who responds to system issues and problems. A system manager is also assigned to each system; this manager is responsible for the overall operation and support, as well as being an expert in the particular hardware and software. Vendor personnel, possibly on site, are available to ensure that the systems operate well and provide high reliability.

3.1.2 Direct Scientific Support

Direct scientific support focuses on helping the DOE scientific community become more productive in its computational and data management work. The key components of this activity are described below.

Advanced Consulting and Support

Consulting staff solve and manage client problem reports and requests for assistance, particularly with regard to programming and application development. NERSC provides live phone coverage from 8 a.m. until 5 p.m. (local time) Monday through Friday, with basic help desk support around the clock. The staff introduce new techniques, systems, and technologies; help analyze and debug problems with user codes, as well as with systems and applications software; report problems to vendors; and track problems so they will be corrected in a timely manner. Consulting staff provide software support for a complex set of tools, libraries, and environments that exist on some or all of the NERSC systems, such as MPI, TotalView, and performance analysis tools. Over 200 different software packages, including visualization and client interface software, are supported.

Since consulting means working closely with NERSC clients, it is important that the staff assess and understand client needs and requests and advocate for them to the NERSC organization as well as the vendor community. Furthermore, the consulting staff are intimately involved in testing and evaluating the changes brought on by new system functions. They assist in developing and maintaining benchmark suites in a manner that allows them to be used in system regression testing, and they test new application software such as compilers, libraries, and tools for advanced programming.

Direct Collaborative Support

NERSC works directly with scientists on major projects that require extensive scientific computing. In these collaborations, the NERSC staff member is frequently a scientist experienced in the field of study

who is also knowledgeable in the computing needs of the project. Recent examples of these collaborations include:

- The modeling of metallic magnet atoms (a collaboration with ORNL, the Pittsburgh Supercomputer Center, and the University of Bristol, U.K.). This collaboration won the 1998 Gordon Bell prize for the best achievement in high-performance computing.
- Climate modeling, a collaboration with the Geophysical Fluid Dynamics Laboratory (GFDL), which developed a massively parallel version of GFDL's Modular Ocean Model (MOM) code, used by researchers worldwide for climate and ocean modeling.
- The Supernova Cosmology Project, a collaboration with LBNL's Physics Division. NERSC's supercomputer analysis confirmed the existence of the oldest, most distant Type Ia supernova ever discovered.

In all these collaborations, NERSC staff provide high-quality intellectual involvement, not merely consulting support, in the project.

Training and Documentation

NERSC provides advanced training and client instruction in the use of the latest technology. The NERSC staff develop skills in new areas and share them with clients by creating, updating, and presenting all the relevant external and internal training information related to using NERSC systems. These activities include monthly training, multiple days of intensive classes, lectures, seminars, and symposia presented in collaboration with other groups. Additionally, NERSC staff organize documentation provided by vendors and make it available as conveniently as possible (on the Web) to our users.

NERSC uses videoconferencing and teleconferencing to provide the timeliest information to the clients, rather than using infrequent face-to-face meetings. But until recently, each method has had shortcomings. In part because of the development and deployment of the Unified Science Environment and collaborative tools, it will be possible to cost-effectively deliver training and information largely independent of where the provider and consumer are.

In the coming years, NERSC will capture its training content — classes, seminars and lectures — in a form that allows digital distribution. Real-time video broadcast (not studio production quality, but nonetheless competent and effective) of training and other events will make it far easier for clients to have the most current information on large-scale systems, programming, and algorithm development. Once captured, the content can be played later (although without interactive discussion and questions), so a library of training information will be available.

Account Management and Allocations Support

The accounts and allocation support staff maintain and extend several major tools that provide NERSC clients with the ability to manage their project resources. The core part of the effort is in the use, support, and extension of the NERSC Information Management system (NIM). This system manages all accounts and projects for NERSC systems, automatically installs accounts and accumulates usage data of clients and projects, summarizes it, and implements resource restriction if a project or client exceeds its allocation. All accounts and allocations are periodically reviewed and validated, with unused accounts being disabled.

Accurate and timely accounting reports are important to NERSC's client community and DOE. Usage reports are important to judge and improve the service and effectiveness of the NERSC systems. Much of

the effort needs to be done by NERSC, since new systems generally have weak accounting software, and the need to maintain a vendor-independent system precludes the use of most vendor-specific software. The responsibilities include maintaining, adapting, and porting the NIM software for current and new systems as necessary, and producing weekly, monthly, and yearly accounting summaries for clients.

3.1.3 System Management

System management is the term used to describe system administration (mentioned above) and the advanced tasks of resource management, system tuning, system improvement, and developing new functionality. By aggressively using advanced scheduler functions, NERSC has been able to double the amount of computational capability delivered by some systems, compared to that delivered by the standard vendor software. Systems have complex interactions of memory, CPU time, I/O, and networking that make it complex to balance high utilization and fast turnaround for a diverse set of clients and disciplines. NERSC constantly performs tuning and balancing to assure that the resources are well used but also have the best possible response.

NERSC will respond quickly to special requests from the NERSC community for processing and services. NERSC is able to provide highly effective processing — be it high priority, very long runs, massive data, or other special needs — because the system managers can configure the systems to respond to multiple needs for resources.

Proper system management includes cyber security. NERSC uses a best-practices approach, applied with the philosophy of ensuring that known security problems are fixed, systems and communications are monitored for inappropriate activity, and security incidents are responded to swiftly. NERSC uses and improves advanced monitoring and reactive tools that limit inappropriate access but provide the best level of security with minimal impact on performance and function. The cyber security system has components that are in the network as well as in every computational and storage system.

3.1.4 System Improvements

NERSC system staff members are highly skilled at testing and integrating new system hardware and software with little or no service disruption. Technology comes from vendors, the open-source community, the academic community, national laboratories, NERSC staff, and other sources. NERSC's job is to deploy the best and most appropriate technology in a cost-effective manner. Technology — whether hardware, software, or a combination thereof — enters the process and goes through different phases based on its source, maturity, and function. The phases are:

- experimentation and development
- evaluation, observation, and external testing
- prototyping
- testbed
- early use
- general or special use
- final full service.

At each phase, the technology is evaluated for:

- readiness to progress to the next phase

- potential impact on NERSC clients (both immediate and long term)
- overlap with existing functions
- costs (both initial and ongoing)
- benefits and risks.

NERSC schedules explicit reviews to allow the technology to progress to the next phase. Throughout the phases of the process, NERSC staff provide feedback and analysis of the technology to the supplier. NERSC may assist in the development of new requirements for vendors and other groups, and, when appropriate, develop key technology that is important to the success of NERSC clients. NERSC staff interface with vendors and with other sites and visitors in this process.

3.2 Focus on High-End Teams

NERSC has supported large teams of scientists doing high-level computation since its earliest days. Most recently, NERSC enhanced the DOE Grand Challenge projects by deploying the "Red Carpet Plan" of enhanced support. This plan featured a single point of contact for each project and a close relationship with the scientific teams. In the coming years, NERSC will have several levels of teams to support. These teams will be both computational and experimental, and they will come from existing NERSC clients, the new SciDAC teams, and evolving new major experiments. NERSC will be able to continue to support up to five teams as part of its continuing base program. Additional teams will be supported as part of the SciDAC Scientific Challenge Team support funding. Section 4 discusses NERSC's proposed support approach in more detail.

NERSC takes a multifaceted approach to its support of high-end teams. Computational and/or disciplinary experts will be assigned to each team as a point of contact (POC). These focal points not only directly collaborate with the Scientific Challenge Teams, but also serve as an advocate for the team within NERSC. Part of the POC's job is to coordinate with other NERSC staff to deal with issues and needs the teams have; in essence, the POC is just the tip of the support iceberg. NERSC will also expand its expertise in supporting new needs the teams have: collaborative tools, software development, or new areas not yet identified.

3.3 USE Support

The Unified Science Environment (USE) will lead to a new set of software and functions that will expand the services provided by NERSC. The USE functions and implementation are presented in detail in Section 5, but we will mention here some aspects of how USE will influence NERSC's Comprehensive Science Support program. The tools developed and deployed by the USE activity will support Scientific Challenge Teams. Uniform access to computing and data resources across NERSC and Science Grid sites, as envisioned by USE, will facilitate the collaboration among geographically distributed teams. Grid middleware services to support problem-solving environments and workflow frameworks are critical for large integrated teams. USE will deploy tools for integrating human collaboration with computing tools and tools for remote access (e.g., visualization and data). The support for virtual organizations (for example, code and data libraries and team authorization), which is an important element of the USE, will be required in particular for the Scientific Challenge Teams.

NERSC will support USE software that will allow clients to come to NERSC for assistance in accessing and implementing these technologies. Our staff will improve the functions of USE in the same way they have supported and improved the functions of the major systems NERSC provides. USE will be

integrated into the NERSC production systems and services in a seamless manner, so that, at the end of this period, USE will be a common and effective way of working for the staff and client base.

USE support activities are part of the USE level funding.

4. Support for Scientific Challenge Teams

The arrival of large, highly parallel supercomputers in the early 1990s fundamentally changed the mode of operation for successful computational scientists. In order to take full advantage of the new capabilities of these parallel platforms, scientists organized themselves into national teams. Called "Grand Challenge Teams," they were a precursor to the "Scientific Challenge Teams" that NERSC anticipates as its leading clients in the next decade. These multidisciplinary and multi-institutional teams engage in research, development, and deployment of scientific codes, mathematical models, and computational methods to maximize the capabilities of terascale computers. NERSC responded by creating the "Red Carpet" plan, described in Sections 3.2 and 4.2.1.

In March 2000 DOE launched a new initiative called "Scientific Discovery through Advanced Computing" (SciDAC). SciDAC defines and explicitly calls for the establishment of Scientific Challenge Teams. These Teams are characterized by large collaborations, the development of community codes, and the involvement of computer scientists and applied mathematicians. In addition to high-end computing, teams will also have to deal increasingly with issues in data management, data analysis, and data visualization. The expected close coupling to scientific experiments supported by the USE environment (described in the Section 5) will be an essential requirement for success for some teams. Scientific Challenge Teams represent the only approach that will succeed in solving many of the critical scientific problems in SC's research programs. These teams are the culmination of the process of users moving to ever-higher computing capability, and NERSC's new structure enables that entire process (see Figure 4-1).

In addition to working with Grand Challenge Teams, NERSC has already worked successfully with other similar teams. DOE's scientific portfolio abounds with examples, a few of which appear in the introduction to the Implementation Plan of this proposal. In the field of climate prediction, for example, the complexity of the coupled system of atmosphere, oceans, and a host of energy transfer mechanisms has resulted in the aggregation of an extensive and well-organized team of researchers in fields ranging from fluid dynamics and chemistry to applied mathematics, who are building the Community Climate Model. It is perhaps the best-known example of a community code, and it demonstrates the advantages of such codes. First, they unite the field. Also, researchers needing further code can write to the existing code rather than create new work from scratch, thus increasing their productivity significantly. DOE/SC also funded the development of NWChem, a community code infrastructure for quantum chemistry and molecular dynamics that is changing the practice of the discipline of theoretical chemistry.

Similar efforts are beginning in other areas as well, with the DOE materials science network, the Numerical Tokamak Turbulence Project, and the comprehensive terascale accelerator simulation environment for the U.S. particle accelerator community.

NERSC plans to continue its support for up to five special projects or teams as part of the NERSC base program. Services for the strategic teams will include Web page hosting, CVS services, specialized consulting and computational algorithmic support, and special system services and request processing. Additional teams beyond that will be supported through SciDAC Scientific Challenge Team support funding. In Sections 4.2.2 and 4.2.3, additional tasks will be described which fall under the SciDAC Scientific Challenge Team support funding category. Several of the USE tools (USE level funding) described in Section 3.3 will also facilitate the large team collaborations.

4.1 Leveraging SciDAC and Other Efforts

One primary motivation for the establishment of the SciDAC Scientific Challenge Teams is to bridge the software gap between currently achievable and peak performance on terascale platforms. With the previous generation of vector supercomputers, many scientific codes realized 30% to 50% of the peak performance of the supercomputer. By contrast, with the current generation of microprocessor-based parallel supercomputers, scientific computing codes often realize only 5% to 15% of the potential peak performance of the computer. This gap will continue to increase with increasing use of parallelism. Closing this potential usage gap represents a major challenge to the SC scientific computing community.

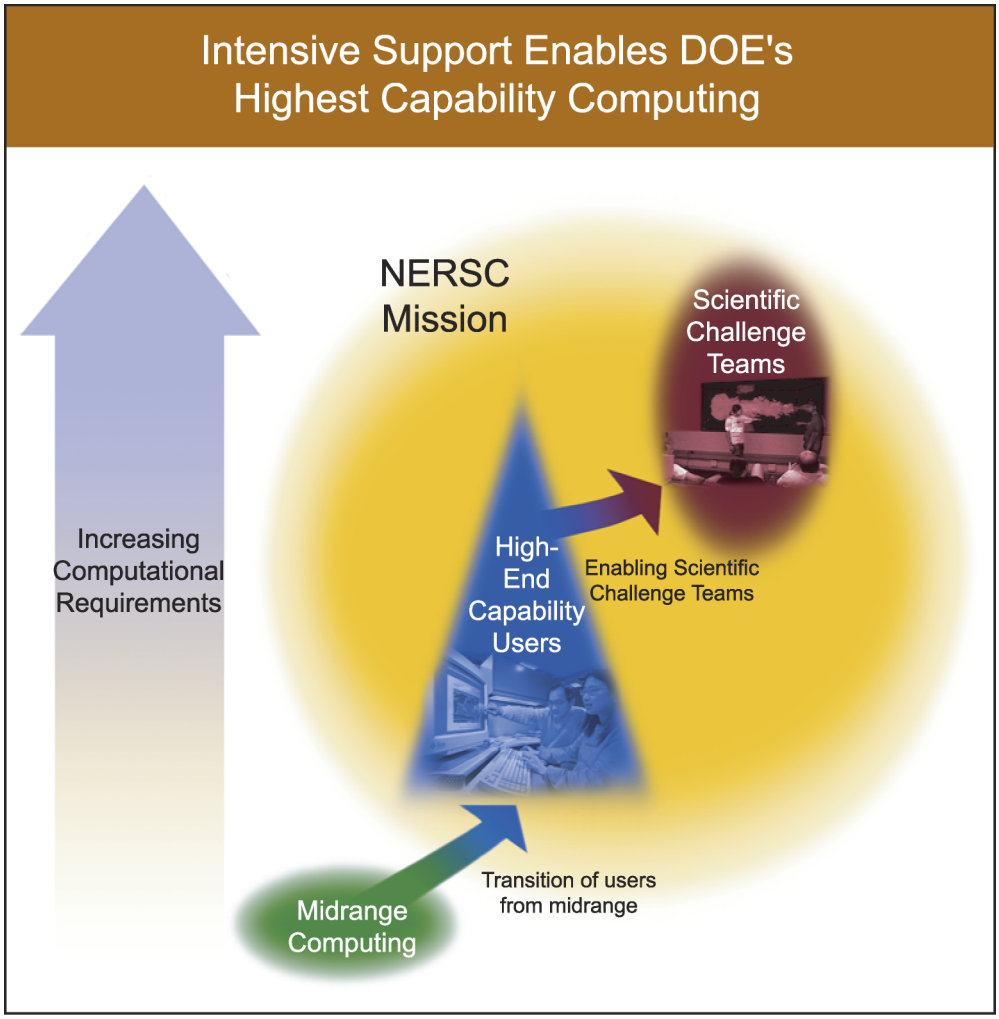

Figure 4-1. NERSC facilitates the transition to high-end capability computing, and enables Scientific Challenge Teams through intensive support.

NERSC's proposed focus on Scientific Challenge Teams anticipates these new teams and their special requirements. In addition, we expect that many other scientists will organize themselves into similar collaborations in the coming years. SciDAC will be the catalyst for a fundamental shift in computational science from the principal-investigator model to the collaborative-team model. SciDAC also proposes to establish a set of new centers in computer science and applied mathematics, called Integrated Software

Infrastructure Centers (ISICs), with the goal of supporting research, development, and deployment of software in order to

- accelerate the development of and protect long-term investments in scientific codes
- achieve maximum efficiency on terascale computers
- enable a broad range of scientists to use simulation in their research.

Since the success of these ISICs will be measured by their impact on applications and Scientific Challenge Teams, NERSC will seek a close collaboration with these centers. In particular, NERSC plans to collaborate closely with ISIC staff in order to facilitate the tasks of deployment, installation, and support for the ISIC tools on the NERSC platforms. NERSC also plans to provide the advanced support in the use of the tools for the Scientific Challenge Teams using the NERSC platforms, all under SciDAC Scientific Challenge Team support funding.

4.2 Strategy and Major Milestones

NERSC's strategy for the next five years is to build a focused-support infrastructure for the Scientific Challenge Teams consisting of four components:

- integrated support and collaboration from the NERSC staff
- deployment of tools developed by ISICs
- deployment of Grid and collaboration technologies (USE)
- building the software engineering infrastructure.

NERSC will implement this strategy by pooling staff resources that are already planned elsewhere. Of these four components, the additional support beyond the base program, the deployment of ISIC tools, and the software engineering support function are supported through SciDAC Scientific Challenge Team support funding. The deployment of Grids and collaboration technology is supported through USE level funding. By leveraging the base program level comprehensive scientific infrastructure, NERSC will integrate all four components into a comprehensive support structure for the Scientific Challenge Teams. NERSC's plan for each of these components is described in the following paragraphs. The USE support is described in Section 3.3.

4.2.1 Integrated Support and Collaboration from the NERSC Staff

During the second round of the Grand Challenge Program, from 1997 to 2000, NERSC served as the computing facility for eight Grand Challenge projects. NERSC provided focused support by developing the "Red Carpet" plan for the Grand Challenge teams. This plan revolved around building individual relationships with the users, as well as providing a NERSC staff member as a point of contact (POC) to expedite any problems or concerns. In order to extend the same level of support and collaboration for the Scientific Challenge Teams, NERSC will continue to use this Red Carpet plan, and add some new elements, described below. As with the Red Carpet plan, each Team will be provided with a single POC named from the NERSC staff. The POC will handle all requests by the Challenge Team, and in particular will facilitate special requests, e.g., special queues, early access to new systems, etc.

The Scientific Challenge Teams will also have direct access to some of the special staff resources at NERSC, e.g., the Scientific Computing Group and Visualization Group. NERSC will encourage

collaboration in areas such as algorithms, visualization, data management, and software engineering tools. Such collaborations range from extended consultation to scientific collaboration for the duration of a project and can span the entire range of comprehensive scientific services. We expect to build individual long-term working relationships between NERSC staff and Challenge Teams, which will increase the productivity of the Team and contribute to the professional growth of the NERSC staff involved.

NERSC provides direct collaborative support for scientists on major projects that require extensive scientific computing. In these collaborations, the NERSC staff member is frequently a scientist experienced in the field of study and is knowledgeable in the computing needs of the program. As the examples in Section 3.1.2 show, this effort has produced significant accomplishments. We propose to continue this high-quality intellectual involvement in these collaborations.

4.2.2 Deployment of Tools Developed by ISICs

The seven recently selected ISICs will further develop and deploy software and tools for increasing the productivity of the Scientific Challenge Teams. ISICs have been formed in the areas of numerical algorithms and libraries, system software tools, performance tools, and data management. Since the focus of the ISICs will be deployment of tools in direct support of the Challenge Teams, it is obvious that NERSC must be a focal point for this deployment activity. NERSC will facilitate tool deployment through a number of activities; however, the ultimate responsibility for deployment rests with the ISIC staff.

NERSC will provide each ISIC with a single point of contact, and also give ISICs special access to NERSC, in collaboration with the Scientific Challenge Teams. NERSC will develop a software testing and integration methodology for ISIC software, and also provide a beta-test environment for ISICs. NERSC is already working to support the ACTS Toolkit, a set of DOE-developed tools that make it easier to write parallel scientific programs. ISIC support can be seen as a direct extension of the ACTS Toolkit support model, which provides documentation, assistance, and training, as well as second- and third-tier support.

4.2.3 Building the Software Engineering Infrastructure

This component represents a new area of support for NERSC. NERSC's main role here will be to advocate for software engineering, and to provide the tools and training. It is important for NERSC to evaluate the needs of the Challenge Teams first and not push "solutions." NERSC will engage the scientists and motivate them to consider software engineering practices.

The Scientific Challenge Teams will present a wide range of software experience. This range of experience will provide a challenge for NERSC. On one hand, we expect to work with some well-established communities, such as the high energy and nuclear physics data analysis communities, who have worked with code teams of hundreds of developers and tens of millions of lines of codes. On the other hand, we might find newly constituted groups, who are united only by the desire to work together, but who have not yet thought at all about the tools required. NERSC will therefore propose a variety of tools and services, and will modify them as the requirements of the Challenge Teams evolve. There is a wide range of opportunities for software engineering and other support of the teams; these will be discussed in detail in the implementation plan. NERSC will also gain experience by using software engineering infrastructure in its own projects.

In order to be the focus of the Scientific Challenge Teams, NERSC must support the establishment of code repositories for the Challenge Teams. We can do this by providing tools for managing these

repositories. NERSC will enable nightly builds, automatic notification of errors, developers' builds, etc. Currently CVS, gmake, and custom tools are the most widespread tools. However, if required, NERSC will also provide more expensive tools. An advantage of a centralized facility such as NERSC is that several teams can share these tools.

NERSC will consider developing new tools, e.g., a standard (but customizable) layer above CVS, gmake, etc., for large-scale software distribution, installation, and release management. NERSC might also integrate automatic notification (e.g., HyperNews forums) and generation of documentation (e.g., Doxygen).

Another area where NERSC can increase the productivity of its clients is by providing tools for workflow management. One set of such tools is currently being developed under ASCI funding at Sandia National Laboratories. These tools will be become critical when USE is fully operational.

5. Unified Science Environment

A new science paradigm is emerging in which multiple national assets — computational, data, and experimental — are simultaneously employed in the process of scientific discovery. The underlying systems are geographically distributed and managed by different organizations. Projects such as the DOE Science Grid are constructing prototype infrastructure of software and services that will automate the task of using these systems, in concert, for large-scale scientific problem solving. When integrated with NERSC's high-end computational and data storage resources, NERSC calls this the "Unified Science Environment" (USE).

5.1 Scientific Motivation for the Unified Science Environment

Numerous examples of the convergence of computing, experiment, and theory in science are emerging in DOE/SC programs. The fact that the Office of Science builds and operates major experimental facilities, including the light sources and neutron sources used by the nation's chemists, materials scientists, and biologists, focuses this trend in DOE's mission. Already being discussed is the possibility of theoretical simulations being used to steer experiments at the Advanced Photon and Advanced Light Sources. At the same time, remote telepresence of researchers is being developed as a routine mode for use of those facilities. Computational and Data Grids, such as the GriPhyN and Particle Physics Data Grid projects, are being built in high energy and nuclear physics. The National Earthquake Engineering Simulation Grid aims at tying together the nation's major civil engineering instruments and simulation capabilities. The international, satellite-based earth sciences community is planning for a Grid to do the large-scale data management and simulation needed to effectively use huge volumes of earth-sensing data. Individual disciplines are developing experimental efforts in the area of Grids, and the computer science tools for such structures are rapidly maturing.

Grid technology is a collection of tools and services that facilitate the building and managing of "virtual" systems that integrate distributed, heterogeneous, multi-organizational resources on demand. These resources might include the different computing and data systems operated by a supercomputer center like NERSC, as well as a diverse collection of user-controlled computing and data systems and scientific instruments.

The integration of high-end computing, storage, and data management capabilities with distributed scientific application environments will enable the routine use of large-scale, multi-institutional, distributed science problem-solving environments. A compelling reason for the integration of simulation with experiment is that each can drive and modify the other in a time-constrained interaction. Data-driven projects like Supernova Cosmology that use observation to refine simulation, and simulation to drive observation within a narrow time window, and use large-scale data archives to provide persistence for both, cannot be done without this sort of integration. Grids provide the software foundation required to make the building and use of these distributed applications easy and routine. The Unified Science Environment (USE) will homogenize the computational simulation and data management environments and facilitate building multi-component applications that execute codes, catalogue, store, and access data, and integrate collaborators at different locations. It will also foster the integration of supercomputing and large-scale data management into real-time observational and experimental science.

The model for USE derives from changes that are, of necessity, occurring in how science is done: computing, data, and collaboration must be tightly coupled with theory and experiment across many organizations in order to enable the next generation of science. Examples of the potential of — and the necessity for — a unified approach to computing and science may be found in many of DOE's large-scale

science projects, such as accelerator-based science, climate analysis, collaboration on very large simulation problems, and observational cosmology. These activities occur in widely distributed environments and under circumstances that are constrained by the timing of the experiments or collaborations, and are essential to advancing those areas of science. The NERSC USE will provide high-end computing, storage, and data management in this integrated environment, and thereby facilitate DOE's large-scale science.

5.2 The Role of Grids at NERSC and NERSC's Role in the Unified Science Environment

Grids will play an important role in NERSC, and NERSC will play an important role in Grids. Though Grids provide the middleware for managing and accessing widely distributed resources, NERSC must add the extreme high-end computing and data storage for Grids to have a significant impact on large-scale DOE/SC programs.

Grid middleware provides the user with a uniform view of the job- and data-management environment across heterogeneous systems. This environment has a single, consistent security model and strong security services that are not obstructive. Tools are available in this environment to manage complex sequences of tasks. The DOE Science Grid will initially incorporate limited, non-production computing and storage resources at LBNL, ANL, PNNL, and ORNL. In addition, it will provide support for the development and deployment of applications in the Science Grid. Inclusion of NERSC in the DOE Science Grid will make high-end services available to NERSC computational scientists through the uniform Grid environment (see Figure 5-1). The resulting combination of Grid access to desktop, midrange, and high-end services creates the USE.

5.2.1 Integration of Large-Scale Science, Supercomputing, and Large-Scale Data Management

The role of supercomputing and data management in large-scale scientific research is changing in several significant ways. A growing number of research efforts have developed computing models based on continuous and reliable access to levels of computing and data storage services found at only a handful of supercomputing facilities such as NERSC. While some of these research programs have computing models based solely on computational simulation, many now also address analysis of large data sets acquired from detectors at remote accelerators, earth-based observatories, or satellite-based instruments. For example, the LBNL Supernova Cosmology research program will, in both its present project and its satellite-based successor, SNAP, base its daily operations and observational schedules on comparisons between recently analyzed data and recently computed model simulations, both of which require substantial supercomputing and data storage resources. This time-constrained use of supercomputing in the heart of the observational cycle represents the new aspect of computing integrated with science. Implicit in these computing models is the assumption that coordination and operation of both these large central facilities and extended environments, within which instruments and researchers are distributed across the globe, are straightforward and easily managed. They are not easily managed now, and improving their ease of management is the sort of capability that integrating NERSC and Grids will address.

To a great extent, this automation is required because of the sheer complexity of the operations involved. For example, in Supernova Cosmology, input data and calibration files need to be located in tertiary storage, staged to disk before analysis programs can begin. Resulting data files will need to be archived, cataloged, and published to other collaborating programs. And, in the presence of software, hardware, or communication errors, these activities need to be rescheduled in a carefully controlled sequence to ensure their proper completion. Some experiments, notably the Supernova Factory, will require that these

operations occur with a minimum of human intervention around the clock, whenever observational data from earth and space instruments have been transferred to the HPSS data storage system at NERSC.

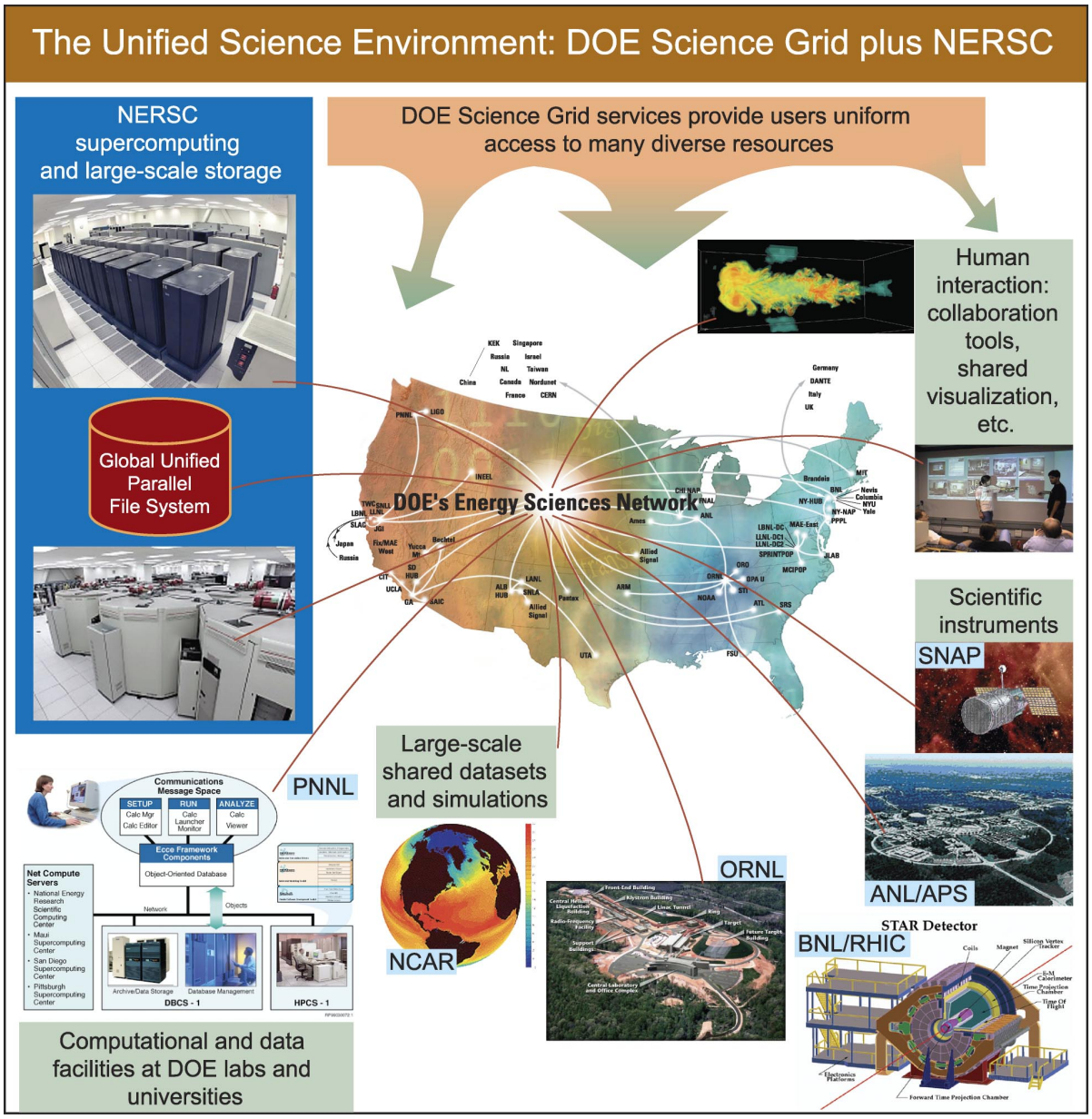

Figure 5-1. The role of NERSC as the largest computational resource in the DOE Science Grid

These computing and data environments are complex collections of large systems involving high-speed networks, very large-scale data storage, and world-class computing resources that need to act in a coordinated, reliable, predictable, and — most importantly — transparent way. The level of integration between supercomputing resources and data analysis and simulation environments anticipated by these research programs is substantial and nontrivial to achieve. DOE is increasingly involved in such applications. As the focus of a number of such large-scale computing efforts, NERSC has the unique

opportunity to provide an integrated set of software tools that will facilitate the development and operation of systems of this scale. In particular, the success of these scientific research efforts will depend on the development of software tools that provide users with transparent, high-level mechanisms to control and monitor the collection of computing and storage activities that constitute an individual work process. Scientific workflow management tools, for example, and the underlying Grid services for managing distributed resources, are one of several key technologies that will ultimately change the way in which supercomputing and large data storage resources are accessed and integrated into large-scale scientific research. Data management tools that provide for cataloguing and publishing information about datasets, and services that manage the movement, replication, and caching of large datasets, will be crucial to large-scale, data-driven, distributed scientific collaborations. These tools are being developed in various Grid R&D projects, and deployment of these tools within the NERSC environment will allow users of the facility to focus more fully on their individual research efforts and less on software tool development.

5.3 Grid Technology and Deployment: Leveraging the DOE Science Grid and Other Efforts

Current Grids are mostly based on services from the Globus and Condor software packages, and emerging data Grid and Web based portals. Globus services provide a standard way to define and submit jobs, manage the code and data associated with those jobs, and locate and monitor the available resources across geographically and organizationally dispersed sites. They accomplish this by working with, rather than replacing, the existing site mechanisms. They also provide a consistent set of security services based on X.509, PKI, or Kerberos authentication, proxy certificates to carry the user authentication to remote resources, and a set of secure communication primitives based on the IETF GSS-API. Secure telnet, remote shell, and secure ftp are provided by using these services. Condor-G provides job management on top of the Globus services, ensuring that one or more associated jobs that might run on remote resources execute once and only once. Both Globus and Condor services provide for communicating with remote jobs, etc.

A sophisticated set of Grid data services is being developed by the NSF GriPhyN and EU DataGrid projects for managing massive data sets in support of the global high energy physics community. Over the next few years these will provide for cataloging, querying, accessing, and managing replication, location, and movement of very large data sets from a worldwide collection of data sources.

Grid portal work at half a dozen institutions is defining and building the Web services and primitives that will provide all Grid services though the user's Web browser. Advanced services that will provide for brokering, co-scheduling, advance reservation of CPU capacity, network bandwidth, and tertiary stored data availability are currently being developed. Collaboration services are also being developed that provide for secure distributed collaboration group management, messaging, versioning and authoring, and the definition and management of "virtual organizations." These services are being integrated with the basic Grid services, frequently through Web Grid services.

Also under development is the infrastructure to provide comprehensive system performance and status monitoring; examples include the status of user codes, data transfers, number of processors available, and other tasks. This infrastructure will provide a necessary basic debugging layer for development teams and also provide the basis for tools to aid the job of the NERSC Comprehensive Scientific Support staff.

All of these services and tools are being coordinated through work in the Global Grid Forum, which is working on standardizing the protocols and APIs of the Grid services and tools. This IETF-like organization involves hundreds of researchers and developers from the U.S., Asia, and Europe.

These technologies are being deployed and tested in the DOE Science Grid, in which NERSC is a partner. The Science Grid, together with diverse development activities throughout the global Grids community, will provide a steady flow of technology that will form the basis of USE at NERSC. NERSC, in turn, must adapt these development efforts and forge them into a robust and large-scale data management and supercomputing environment that supports a new and integrated approach to DOE's large-scale science.

6. Collaborations, Metrics, Milestones, and Budget

6.1 Collaborative Efforts in Technology Development and Deployment

NERSC, as the largest Office of Science computing resource, must continue to develop its collaborations with the other programs of the Mathematical, Informational and Computational Sciences (MICS) Division. In the area of the Unified Science Environment, NERSC must leverage the work in Grids and collaborative technologies performed elsewhere in the MICS program, but also lead in placing large-scale production computing facilities in the DOE Science Grid at the rate commensurate with the maturity of the software. In other areas of technology development, NERSC must, in order to meet its responsibility to the Office of Science program that funds it, maintain close connections and strategic collaborations with programs funded by MICS in the other DOE laboratories, as well as universities. Recently NERSC has begun to participate in a series of in-depth discussions with the MICS-funded activities at the ACRF sites at Argonne National Laboratory and at Oak Ridge National Laboratory. The purpose of the discussions is to coordinate the activities of NERSC and the ACRFs, and to explore areas of further collaboration and opportunities for technology transfer. In addition, selected collaborations with other facilities and agencies will be pursued. Some collaborative efforts that NERSC will continue to pursue and expand are listed below.

- Argonne National Laboratory: Grids, PC clusters, visualization.
- Oak Ridge National Laboratory: Probe and other storage and data technology, supercomputer system administration tools.
- Pacific Northwest National Laboratory: Support for NWChem at various levels, and active development of new capabilities in the NWChem suite of programs.
- ASCI and Lawrence Livermore National Laboratory: Computational platform technology exchange regarding the fielding of terascale systems.
- ESnet: NERSC shares operations monitoring with ESnet, and the coupling of ESnet with the USE effort will provide critical underpinnings for deploying USE and the DOE Science Grid.
- NASA: Coordination of NERSC and LBNL Grid supercomputing efforts with NASA's Information Power Grid program will be accomplished by Bill Johnston at LBNL, who is also the NASA project manager for the Information Power Grid.
- National Science Foundation supercomputing centers: Evaluation of the Tera multithreaded architecture, benchmarking, and data management. The primary connection of NERSC with NSF Centers is with NPACI (National Partnership for Advanced Computing Infrastructure) and the San Diego Supercomputer Center (SDSC). Given the high likelihood of SDSC's involvement in the new NSF Distributed Terascale Facility, we expect close cooperation between SDSC and NERSC on Grid supercomputer technology.
- National Center for Atmospheric Research: Collaboration with DOE's Community Climate System Model.

We also expect that the DOE Science Grid will spark much closer collaboration between NERSC and many of the midrange facilities at the other SC laboratories. The Science Grid will provide a much closer coupling of these facilities with NERSC than has existed before.

6.2 Relationship to SciDAC and Other DOE Computing Facilities

In this proposal NERSC configures itself explicitly to support the SciDAC and other strategic Scientific Challenge Teams. Upgrades of the NERSC terascale hardware are directed toward meeting the high-end needs of those teams, and the NERSC scientific support architecture focuses on them. SciDAC will establish a number of Integrated Software Infrastructure Centers (ISICs) to which NERSC will build new connections. The transfer of much of the technology developed in the ISICs to production use at NERSC is an important mechanism for the success of the ISICs. NERSC will work with the ISICs to determine which of their products are appropriate for the NERSC user community, and to deploy those technologies on NERSC systems.

In the coming years the DOE laboratories will continue to develop ever more robust local computing facilities. NERSC's commitment to helping researchers move their codes from midrange computing resources to the most powerful high-end computing that DOE can offer will require NERSC to continue to work with those facilities, as it has done in the past. The work with ORNL to deploy common system management tools for supercomputing systems is an example of such collaborations. Smoothing the path for researchers to move between these facilities, and between these facilities and NERSC, will require more extensive joint efforts.

6.3 Metrics for Success of the NERSC Strategic Proposal

At the highest level, the ultimate measure of success is the quality of science enabled by NERSC. Certainly there will be cover stories in *Science* and *Nature* in the future, maybe even another "Breakthrough of the Year." It is difficult for NERSC to produce a comprehensive metric that will accurately capture its scientific impact. That measurement will largely be accomplished through periodic outside peer review. At the technical and operational level, NERSC establishes goals for each year and measures itself against these goals in a report titled "How Are We Doing?" While this goal-setting process is important in measuring operational effectiveness on a year-to-year basis, it is not the single metric for evaluating a strategic plan.

In general, NERSC will formulate a set of general questions about progress made compared to the strategic proposal. We propose to evaluate ourselves and invite evaluation of NERSC's progress by grading our performance according to these questions. Each question is possibly a subjective qualitative assessment, but we propose to compute a quantitative average in order to measure our performance against the objectives of the strategic proposal. Here are some suggested questions:

1. High-End Systems
    - Did NERSC increase the computational capability by at least a factor of 3 in each of the two major new systems, NERSC-4 and NERSC-5?
    - Did NERSC increase the archival storage by a factor of 8 over the five-year period?
    - Did NERSC succeed in having production resources on the Grid?
    - Did NERSC succeed in providing a NERSC-wide fast file system for production resources?
    - How well did NERSC accomplish all the above systems goals, while maintaining an overall balanced system?
    - Did the advanced development activity yield on average one new technology transfer per year into the production environment?

2. Comprehensive Scientific Support
    - To what extent did NERSC maintain a high-quality level of essential services, as measured by the user survey?
    - How well did NERSC innovate its support and introduce new concepts and ideas?
    - How well did NERSC succeed in building a support infrastructure for Scientific Challenge Teams?
3. Scientific Challenge Teams (if funded)
    - Did NERSC support more than the five strategic teams in an integrated way, enabling the teams to produce new scientific results?
    - How well did NERSC provide extended comprehensive services and thus increase the Scientific Challenge Teams' productivity?
4. Unified Science Environment (if funded)
    - Did NERSC meet the implementation plan for USE as proposed above?
    - Did NERSC support at least five teams that use USE in a day-to-day production manner, enabling the teams to produce new scientific results?
5. Outside Evaluations
    - How well did NERSC perform in independent outside reviews (e.g., the LBNL Director's Review of Computing Sciences)?
    - How well was NERSC's work recognized by its peers in the high performance computing community (measured by papers published, awards received, etc.)?

6.4 Schedule of Principal Milestones

To successfully implement this proposal, all of the components of this proposal must meet important milestones. The staff must be deployed appropriately, and the modest growth in staff proposed over the five-year scope of the proposal must be staged correctly. The deployment of the NERSC-4 and NERSC-5 computers, which will have peak capabilities of approximately 10 teraflop/s and 30 teraflop/s, respectively, must be staged together with increases in storage capacity. The milestones for deployment of Grid capabilities depend on those for networking and hardware resources. Thus the necessary milestones for the components of this proposal are frequently interdependent. The major milestones for the Base Program are summarized below.

FY02:

- Initiate Strategic Project Team support
- Archive capacity at least 2.75 petabytes
- Production network at OC-12
- NERSC-4 RFP release

FY03:

- Initial NERSC-4 Phase 1 system installed

- Initial Grid deployment — basic services installed on non-production storage system, e.g., PROBE; GridFTP made available on the Probe/HPSS system
- Archive capacity at least 3.5 petabytes
- Production network connection upgrade to OC-48
- Staffing for FY03 increase in FTE levels complete

FY04:

- NERSC-4 Phase 1 in full production, NERSC-4 Phase 2 installed
- Grid preproduction: portal-based Grid job submission and management installed, and high-speed Grid access to production HPSS established
- Archive capacity at least 8 petabytes
- Introduce collaborative consulting support
- Web portal for HPSS files

FY05:

- NERSC 5 RFP release
- Archive capacity at least 8 petabytes
- Linux checkpoint/restart initial release
- Grid production, Phase I: NERSC supporting large-scale science collaborations; collaborative services integrated with the NERSC Grid
- GUPFS prototype deployed

FY06:

- NERSC 5 Phase 1 system in production, Phase 2 system installed
- Archive storage capacity greater than 16.5 petabytes
- Production network connection upgrade to OC-192

6.5 Budget Summary

Base Center Budget Summary (Dollars in Millions)	**FY02**	**FY03**	**FY04**	**FY05**	**FY06**	
Personnel (in Full Time Equivalents, FTE)						
High-End Systems	31	35	35	35	35	
Comprehensive Scientific Support	33.5	34	34	34	34	
Management	1	1	1	1	1	
Total FTEs	65.5	70	70	70	70	**TOTALS in $M**
Total Staff Costs	*$13.01*	*$14.65*	*$15.38*	*$16.15*	*$18.22*	$77.41
Hardware						
Computational Investment	$7.85	$11.00	$9.93	$9.93	$9.35	$48.06
Storage	$1.26	$1.58	$1.80	$2.19	$2.30	$ 9.13
Networking & Security	$0.82	$1.00	$1.20	$1.40	$1.47	$ 5.89
Maintenance & Facilities	$5.30	$4.27	$5.98	$6.50	$6.83	$28.87
Total Systems Costs	*$15.23*	*$17.85*	*$18.91*	*$20.02*	*$19.94*	$91.95
Grand Total	$28.24	$32.50	$34.29	$36.17	$38.16	$169.36

Scientific Challenge Team Support Summary (Dollars in Millions)	**FY02**	**FY03**	**FY04**	**FY05**	**FY06**	
Personnel (in Full Time Equivalents, FTE)						
Scientific Challenge Teams	0	1	2	3	3	
Total FTEs	0	1	2	3	3	**TOTALS in $M**
Total Staff Costs	*$0.0*	*$0.22*	*$0.46*	*$0.72*	*$0.76*	$2.16
Hardware						
Support Systems	$0.0	$0.13	$0.25	$0.25	$0.25	$0.88
Total Systems Costs	*$0.0*	*$0.13*	*$0.25*	*$0.25*	*$0.25*	$0.88
Grand Total	$0.0	$0.35	$0.71	$0.97	$1.01	$3.04

Unified Scientific Support Environment Budget Summary (Dollars in Millions)	**FY02**	**FY03**	**FY04**	**FY05**	**FY06**	
Personnel (in Full Time Equivalents, FTE)						
Unified Scientific Environment	0	3	4	6	8	
Total FTEs	0	3	4	6	8	**TOTALS in $M**
Total Staff Costs	*$0.0*	*$0.65*	*$0.96*	*$1.44*	*$2.02*	$5.07
Hardware						
Support Systems	$0.0	$0.25	$0.50	$0.75	$0.75	$2.25
Total Systems Costs	*$0.0*	*$0.25*	*$0.50*	*$0.75*	*$0.75*	$2.25
Grand Total	$0.0	$0.90	$1.46	$2.19	$2.77	$7.32

Composite Budget Summary (Dollars in Millions)	**FY02**	**FY03**	**FY04**	**FY05**	**FY06**	
Personnel (in Full Time Equivalents, FTE)						
Base Center	65.5	70	70	70	70	
Scientific Challenge Teams	0	1	2	3	3	
Unified Scientific Environment	0	3	4	6	8	
Total FTEs	65.5	74	76	79	81	**TOTALS in $M**
Total Staff Costs	*$13.01*	*$15.52*	*$16.80*	*$18.31*	*$21.00*	$84.64
Hardware						
Base Program Systems	*$15.23*	*$17.85*	*$18.91*	*$20.02*	*$19.94*	$91.50
SCT Support Systems	$0.00	$0.13	$0.25	$0.25	$0.25	$0.88
USE Support Systems	$0.00	$0.25	$0.50	$0.75	$0.75	$2.25
Total Systems Costs	*$15.23*	*$18.23*	*$19.66*	*$21.02*	*$20.94*	$95.08
Grand Total	$28.24	$33.75	$36.46	$39.33	$41.94	$179.72

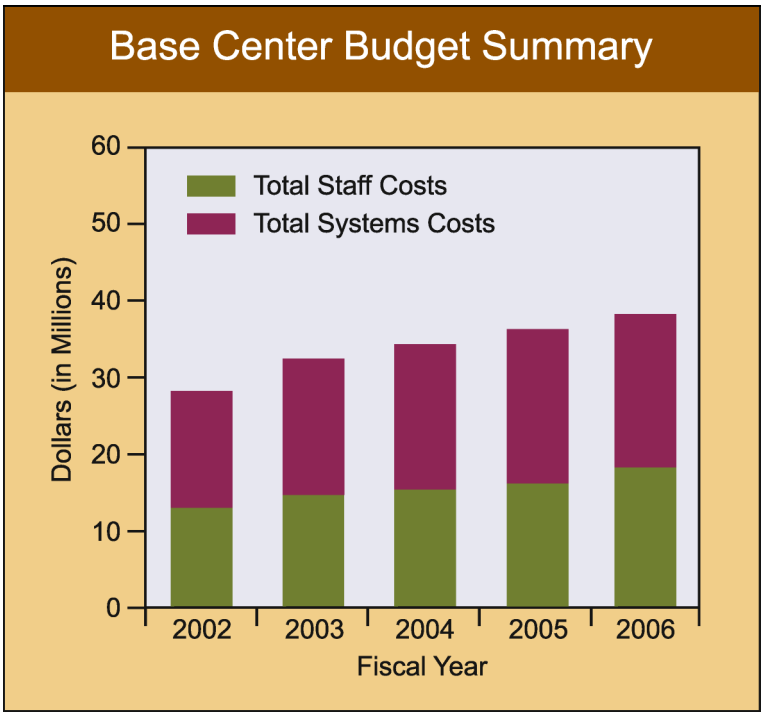

Figure 6-1. Executive budget summary for the Base Center.