LBNL-62284

NERSC Cyber Security Challenges That Require DOE Development and Support

Brent Draney, Scott Campbell, and Howard Walter
NERSC Center Division
Lawrence Berkeley National Laboratory

January 16, 2007

This work was supported by the Director, Office of Science, Office of Advanced Scientific Computing Research of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Introduction

Traditional security approaches do not adequately address all the requirements of open, scientific computing facilities. Many of the methods used for more restricted environments, including almost all corporate/commercial systems, do not meet the needs of today's science. Use of only the available "state of the practice" commercial methods will have adverse impact on the ability of DOE to accomplish its science goals, and impacts the productivity of the DOE Science community. In particular, NERSC and other high performance computing (HPC) centers have special security challenges that are unlikely to be met unless DOE funds development and support of reliable and effective tools designed to meet the cyber security needs of High Performance Science. These needs include:

- in-depth auditing of users' authentication credentials and interactive session keystrokes
- intrusion detection/protection systems (IDS/IPS) and firewalls that can process 10–40 gigabit per second (Gb/s) single-stream traffic flows without impacting network and system performance
- firewalls designed for the dynamic port usage typical of Grid and other middleware (which will protect scientific work in a way similar to how commercial firewalls currently deal with FTP)
- role-based identity management that allows very large ad hoc collaborations to work together effectively without compromising security.

These special security needs of open HPC science are not likely to be met by commercial security products for one or more of the following reasons:

- Vendors lack interest in developing and supporting technology without a demonstrated, wide market in the commercial and/or commodity space.
- The higher network performance required for large data transfers is more stringent in HPC than in commercial sites, which are dominated by many small flows.
- Because of their use of advanced technology, high profile, and required openness, HPC centers need new security technology sooner than commercial market vendors could supply it.
- HPC sites have tens of thousands of systems — all working in unison. Rolling and synchronous updates for security are often not feasible without serious disruption of services, so the standard practice in the commercial space is not transferable.

To understand why commercial products will not, for the most part, be adequate to meet NERSC's security needs, it is informative to look at NERSC's resources, usage patterns, and data rates, and compare them to the larger enterprise market typically targeted by networking and security companies. While NERSC has some similarities to very large corporate networks (such as Google, Schwab, Amazon, etc.) in terms of number and

speed of processors, the distribution of computing resources, access patterns, and performance requirements are vastly different.

1. NERSC Today

The mission of the National Energy Research Scientific Computing Center (NERSC) is to accelerate the pace of scientific discovery by providing high performance computing, information, data, and communications services for DOE Office of Science sponsored research. NERSC annually supports over 2500 research scientists, more than 50% from universities, all of whom access NERSC resources remotely via a wide range of non-dedicated networks. NERSC resources, which consist of tens of thousands of high speed processors, petabytes of data storage, a 100 terabytes (TB) global file system which is accessible from all NERSC computational resources, and a 10 Gb/s connection to ESnet and the Internet, are a rich prize for a hacker. The challenge for NERSC and the Office of Science is to permit these researchers unencumbered access to NERSC resources while preventing the resources from being compromised.

NERSC resources are first and foremost open-science HPC resources, which differentiates them from both corporate and DOE classified resources. Open science means that one to several hundred researchers work on a "community" project from almost anywhere in the world. It also means that the researchers are developing and testing state-of-the-art computer codes which require execution privileges and unrestricted shell access. SSH is the primary means of interactive access, which eliminates the possibility of an adversary intercepting (sniffing) a user's password or credential as it travels through the network[1]. However, in a good example of the on-going balance the open security implies, because SSH encrypts all traffic, border intrusion detection/prevention systems are blind to the contents of a session.

The research codes developed by the NERSC users stress computer security since they can open and utilize a large number (64,512) of unprivileged ports and create new services such as permitting multiple researchers to connect to the running code to steer the computation and view the results.

High performance implies both a large number of processors and high speed, parallel data flows. Today, NERSC computational systems have over 8700 processors which log system messages to a central syslog server. Later this year, Franklin, the newest NERSC Cray computer, will enter production with an additional 19,344 processors. Day-to-day security log analysis is quickly outstripping the capabilities of the human security analysts. Every day, about 2 TB of data enters NERSC from the Internet, and 0.4 TB leaves NERSC. Unlike corporate sites where the majority of traffic is relatively low speed, web server based transactions, the majority of NERSC border traffic is very large (>10 GB) bulk data transfers using tools such as FTP and Grid software. The data sizes mandate high speed, parallel flows, and traditional security hardware such as commercial

[1] Note it is still very possible to steal a user's credential or password — and credential theft is one of the most common forms of attack on NERSC.

firewalls are unable to keep pace. NERSC has seen the effect of using commodity security components when we receive problem reports of slow transfer rates between a scientist's local systems and NERSC. Many of the problems have to do with the limitations imposed by commercial security technology such as firewalls and Virtual Private Network (VPN) servers. Similar issues are seen for sites the institute "gateways" for users rather than direct access to individual systems.

Figure 1 shows the NERSC System Architecture for 2007.

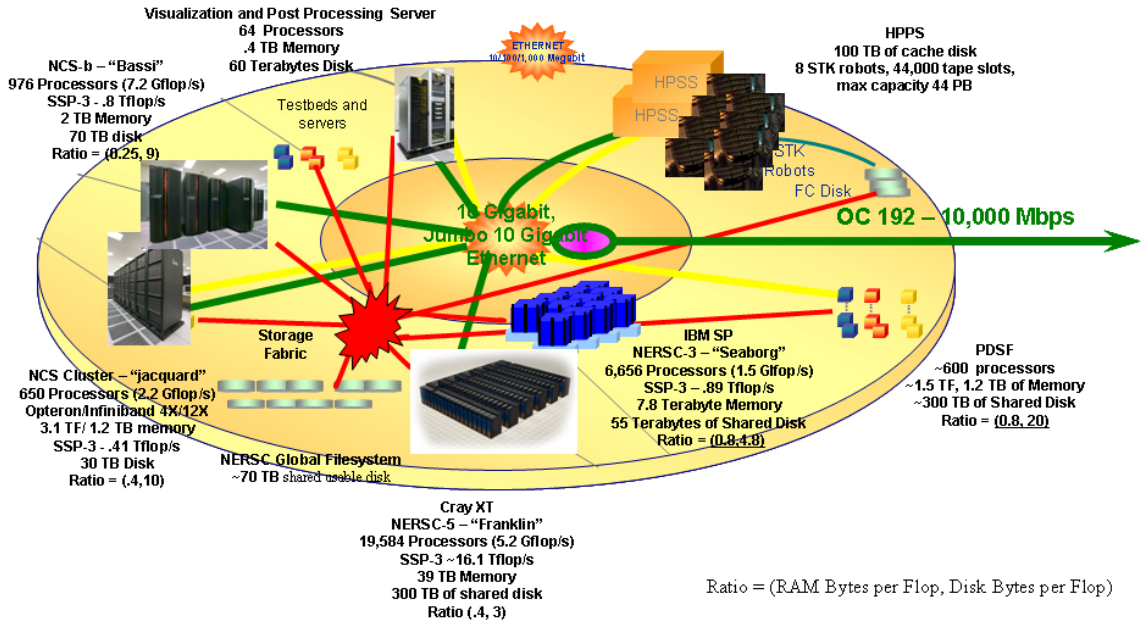

Figure 1. NERSC System Architecture for 2007.

2. NERSC in the Future

All HPC centers evolve in order to meet the national needs for competitiveness and to stay at the forefront of scientific productivity. A multitude of requirements, many that change the fundamental assumptions of the security models and practices, have to be met. While in the past, simulation and analysis of simulation results were essentially within a center, now there are communities of hundreds that share the work and share the results. Interdisciplinary teams of workers create applications that have 100,000–5,000,000 lines of code. The rapid expansion of some form of "virtual organization" is pushing the limits of what current software and system management can support. On the other hand, VOs are becoming a key way of doing large scale science — from climate to fusion to experimental data analysis.

In addition to creating new algorithms and applications that have differing profiles, each team requires new services. Hence, NERSC and sites like it are constantly expanding their service profile. Grid-based services, using new software protocols and layers, are quickly emerging with the existence of the Global Grid Forum, the Open Science Grid, and the NSF TeraGrid. Application teams, ranging from high energy physics to climate research, from astrophysics to quantum chromodynamics, from nanotechnology to life

science, all demand collaborative services and data access. Further, science teams will be moving more between sites with the DOE and other facilities, as allocations and functions change.

The data volumes being handled at NERSC are 20 to 10,000 times more than just ten years ago. This will be expanded as new services such as geographically distributed, parallel file systems, are deployed.

3. Current Security Posture

Figure 2 shows the generic network design of the NERSC network, which is segmented into functional areas.

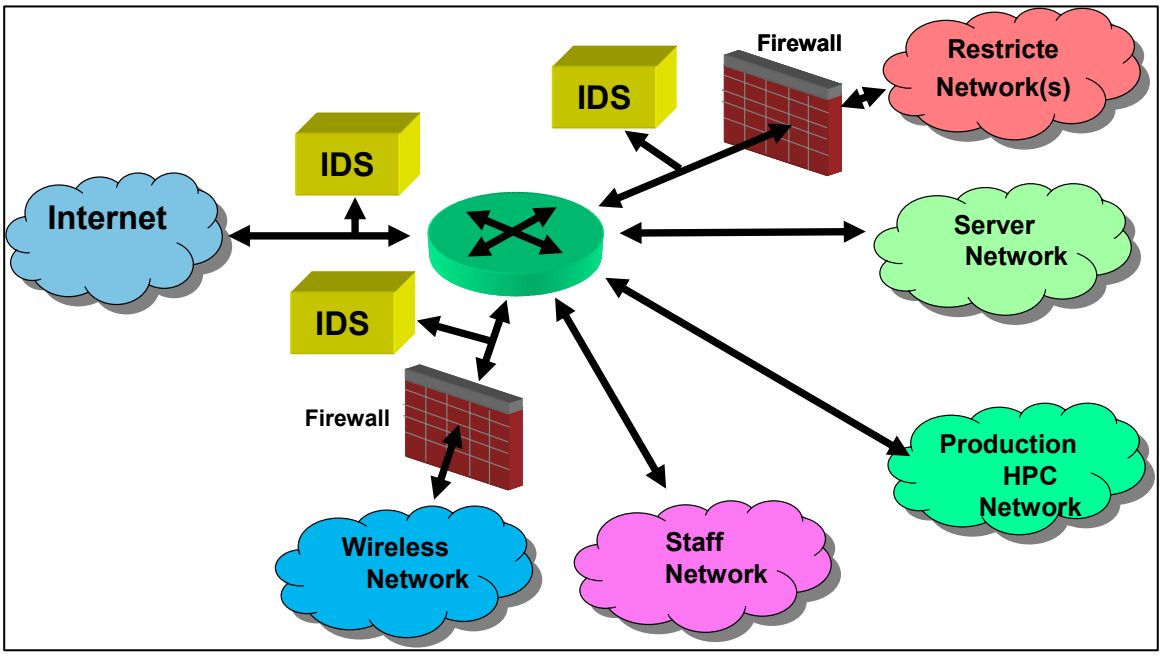

Figure 2. A general view of a segmented network in an HPC facility.

This arrangement fits into the general security posture at NERSC, which stresses the need to do security within an open science environment by extensive monitoring and proactive administrative practice. Examples of this include the following operational production practices:

- *Internal and external network monitoring*. This is used for scan and attack detection, as well as providing a detailed protocol analysis useful for forensic analysis. Systems used for this include Bro, Snort, and a network traffic bulk recorder which are tied together with locally developed scripts.
- *Centralized syslog services*. Allows for a single analysis point as well as getting logging information off of a host if it is compromised. This also provides a convenient way to do detailed login analysis of user activity such as SSH logins across the entire center.
- *Selective use of firewalls and port blocking*. Firewalls are used internal to the NERSC network for critical non-user functions where the need for heightened security

outweighs the performance penalty. In addition, a small number of services are blocked both on the overall border and between network segments when these services are determined to never be needed by the NERSC community.

- *Active vulnerability scanning*. All systems are scanned for general security vulnerabilities. In addition, specific application scanning is done for specialty systems such as database servers.
- *Aggressive administrative controls*. Systems are carefully monitored and maintained. This includes aggressive system patching and centralized configuration management on major systems.
- *System process accounting*. On all main systems, some form of process accounting records user activities for forensic analysis in the event of an attack.
- *Central account management*. NERSC proactively manages all accounts on all systems. Accounts are restricted and then removed after a relatively short idle period. This decreases the vulnerability from password sniffing. Further, account information on all NERSC systems is frequently validated again a known correct database in order to find unexpected privileges and accounts. Advanced password formation is enforced and checked periodically.

A number of these practices, particularly those relating to system monitoring and the need for flexible user services, differentiate the needs of NERSC from those found in the standard implementations for commercial situation described in Appendix 1. As expected, the presence of large numbers of direct execution shell accounts on high performance computing systems poses one of the most significant differentiating factors. In light of this, there are a number of improvements that need to be included in the protection of an open, scientific facility.

The first category of improvement is user activity monitoring. This includes SSH keystroke monitoring, as well as accessing the more ad hoc Grid-related protocols. Integrating cross-system process accounting data into a more complete view of the activities of a given user's activity across an entire facility would also go a long way toward a improved protection

At a strategic level, the overall network design for both NERSC and most typical large corporations is shown in Figure 3, which illustrates NIST guidelines for placing filtering routers, intrusion detection systems (IDS), and firewalls within a site's overall network architecture. What is different is the placement of the actual systems that serve the user community. Essentially, the major systems in an open facility need to be placed on layers closer to the external network.

This network diagram describes three main segments: an external network providing services to the public and external users, a main network for services to staff, and a special internal network for protecting sensitive systems. Deciding which segment a system ends up in is determined by its function (i.e., who is supported and what services are provided) rather than by the system's size or cost. Along the same rationale, the size

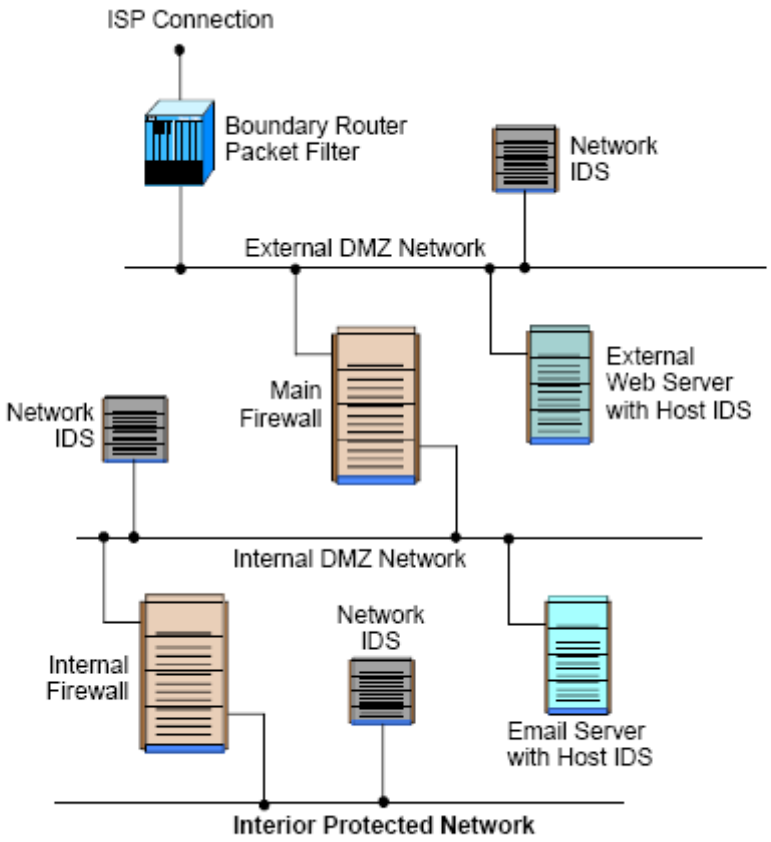

Figure 3. Recommended firewall design for most large organizations
(from *Guidelines on Firewalls and Firewall Policy,* NIST publication 800-41).

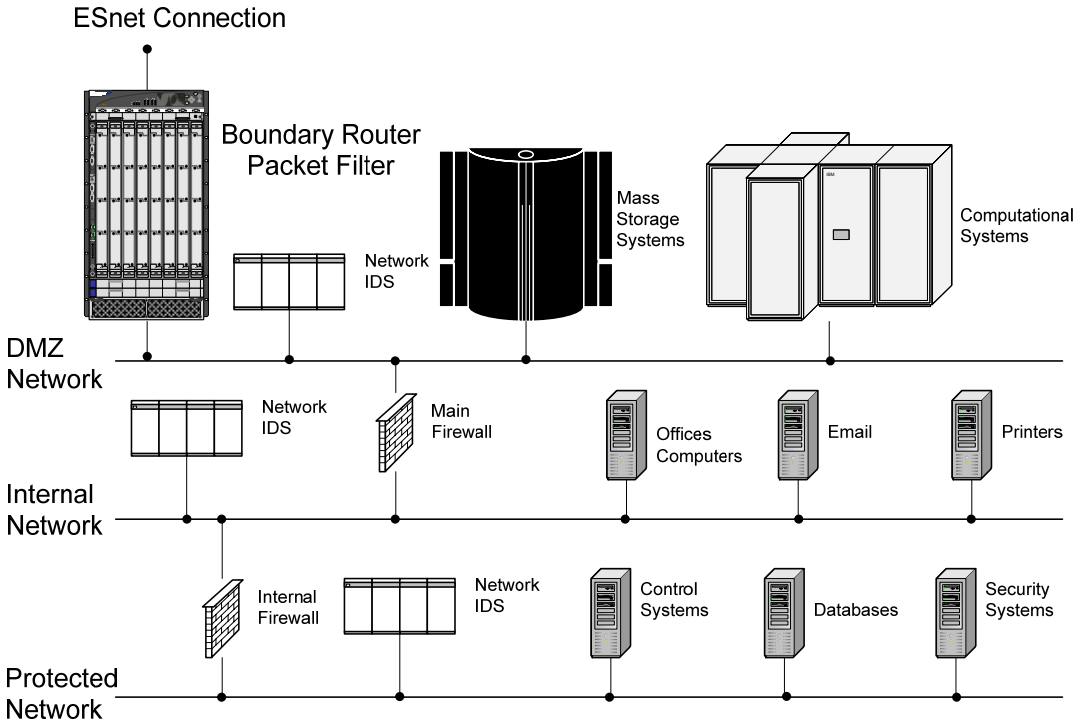

Figure 4. NERSC's network security configuration.

and performance for each of the different network segments are driven by the systems and services provided along that segment.

For an example of this network design, we can look at a generic web service provider. In their design, web servers are placed on the external network, where the public can have access. However, almost all of the work and information is held on servers and data base systems that are on other networks that are protected from both the Internet and its own web servers in the event that one is compromised. Well defined transactions pass from the external servers to the internal, often protected by application-specific firewalls.

NERSC places its HPC systems on the external network, where NERSC users can have high performance access, while simultaneously protecting its main network from both the Internet and its own HPC systems in the event of system compromise (Figure 4).

4. Near-Term Goals

The security challenges facing NERSC, as described above, can be collected into three basic problem sets: network performance and dynamics, application complexity and diversity, and a complex user community that can have transient affiliations with actual institutions. To address these problems, NERSC proposes the following four general solutions:

1. *Auditing user and system activity across sites.* This describes a consistent set of secure tools used to gather end-to-end audit information, as well as the ability to detect and respond to actionable events. Currently, IDS logs, session connections, batch job files, process accounting (pact), system status monitoring (nagios), application performance monitoring (IPM, PAPI) and much other data is collected, at least in real time. Each collection is separate and the parameters for analysis, with the exception of IDS logs, is non-realtime. This is in part due to the size of the data, and in part due to not having defined relationships to look for. The ability to join all this data together in real time would provide an enormous improvement in protection. This vision should be that of the commercial credit card companies — be able to proactively detect possible changes in behavior at an individual user basis and then take measures to quickly verify and possibly limit the damage.
2. *Firewall port configuration in real time.* This provides the ability for an authenticated and appropriately privileged user to modify firewall rules in order to provide for a greater level of flexibility in supporting new protocols until firewalls can be taught to monitor control traffic and respond dynamically.
3. *Cross-site/virtual organization identity management and access control.* As shown in Figure 2, even an individual site has multiple protections. The output and logs of each device are only correlated today by very smart security and system managers. The goal would be to have automated coordination of the observations of these devices, not only within a site, but across sites. Furthermore, today's expansion of virtual organizations means that there is less direct management of the work of groups. Understanding how to identify participants in VO is challenging, since operating systems know only of users and groups. Hence, there is a need for new models and approaches to granting/revoking responsibility, authority and privilege, recording use

and assigning roles. The goal is to provide the ability to uniquely identify individuals and systems for access control so that, in the event of a hostile act, the ability to respond is improved.

4. *Detecting security issues in application middleware.* This includes code analysis to identify problems in application codes before compromises can occur, as well as the ability to identify configuration errors in running applications.

With these tools, it will be possible to address the specific issues differentiating HPC sites such as NERSC from large corporations such as Google.

5. Long-Term Goals

There are three general issues seen for the long term goals: *data volume, application complexity,* and *information integration*. *Data volume* not only represents problems related to increased network speeds, but also of storage, file systems and distributed applications. *Application complexity* deals with distributed applications (such as Grid services) and their related virtual organizations. *Information integration* is related to user activity auditing, but will take the individual problems solved in the short term goals and look at them from a collective perspective.

As stressed above, there is no complete package which can be purchased for the emerging technologies being described here. Production quality tools need to be developed and assembled in such as way as to maximize flexibility and interoperability.

To address these long term goals, the following general solutions are proposed:

1. *Network capacity* — A fundamental issue needing to be addressed is doing production-quality network intrusion detection and analysis at speeds up to 100 Gb/s. Given current limitations in hardware design, some sort of hardware acceleration will be required.

2. *Process accounting abstraction* — A tool that will allow system and security managers to gather process accounting information for a given user across a number of systems via the same interface.

3. *Cross-site file system issues* — In addition to the issues created by inter-system file system use, more complicated problems arise when these are extended across site boundaries.

4. *Project accounts and VOs* — As projects become larger and more collaborative, creating a means for assigning and tracking user identities as well as associated permissions and accounting data in a way that is consistent across arbitrary locations will be necessary.

5. *Tools for identifying correlations between different security events* — In the short term goals, individual tools will be developed for security analysis and anomaly detection for a set of given functions. What we are looking for is to tie together

different data types to get a significantly clearer picture of a given security event. For example, correlating an outbound IRC connection and process accounting info, or web server error messages and network traffic.

Appendix A

Table 1 describes functional requirements for each of the three network segments.

Table 1. Network Comparison: NERSC vs. Large Corporation

		NERSC	Large Corporation
External Network	Traffic patterns	Thousands of very large connections (100 MB–100 GB) through a single border.	Billions of very small connections (10 KB) geographically distributed.
	Network size	Very large.	Medium.
	Network rates	10–40 Gb/s linked with 10 gigabit routers. Supporting data rates >1 Gb/s. Systems with 10 gigabit interfaces.	1–10 Gb/s linked with 10 gigabit routers. Supporting connection rates >100,000 connections/second. Systems with 1 gigabit interfaces.
	Supported systems	HPC systems, archival mass storage systems, remote visualization, web servers, email gateways, DNS servers.	Arrays of redundant web servers, email gateways, DNS servers.
	Protocols	Access is primarily authenticated SSH logins, Grid protocols, and FTP large file transfers.	Access is primarily unauthenticated web queries with a smaller number of authenticated transactions.
	User base	Thousands of users working collaboratively.	Millions of individual users.
Main Networks	Traffic patterns	Interactive sessions, transactional data.	Interactive sessions, transactional data, large database synchronization.
	Network size	Small campus office.	Very large campus office.
	Network rates	100–1000 Mb/s linked with 1 gigabit routers.	100–1000 Mb/s linked with 10 gigabit routers.
	Supported systems	Small number of offices, internal email servers, file servers, printers, internal web servers, non-sensitive databases.	Very large number of offices, internal email servers, file servers, printers, internal web servers, development and research systems, non-sensitive databases.
	Protocols	Web, email, print, SSH.	Web, email, print, SSH.
	User base	~60 staff and visitors.	Thousands of staff and visitors.
Sensitive Networks	Traffic patterns	Interactive sessions and transactional data.	Interactive sessions, transactional data, large database synchronization.
	Network size	Very small.	Large.
	Network rates	100 Mb/s protected with 100 Mbit firewalls.	1–10 Gb/s protected with arrays of 1 Gbit firewalls.
	Supported systems	Sensitive databases, security systems, configuration management/control systems.	Sensitive databases, security systems, configuration management/control systems, proprietary data systems, financial systems.
	Protocols	SSH, SNMP, syslog, database communications.	SSH, SNMP, syslog, database communications.
	User base	~25 system/security analysts and DBAs.	Thousands of programmers, researchers, DBAs, system/security analysts.

Applications	Development	Developed using formal methods by a small group of professional developers	Developed as research codes by small to very large groups of scientists.
	Interfaces	Well defined interfaces and protocols	Rapidly changing interfaces
	Time Period	Once developed and tested – applications tend to not change for a substantial period	Change on the order of weeks to months
	Users	Large numbers of users who use predefined functions but are not developing or changing code	Small to moderate number of users who use the codes for production science runs while also modify the codes to provide improvements
	Number of Applications	10-20 of applications	500-1,000 applications

Looking at the primary differences between the two requirement sets, a number of items stand out as unnecessary for large corporations, yet essential for the NERSC model. This constitutes a set of requirements for which there is a lesser financial incentive for networking and security vendors. These basic differences include:

1. **Effective security for high speed data connections.**
   While large corporations have Internet links at 1–10 gigabit speeds, few if any are trying to sustain single data connections above a few Mb/s. A firewall or IDS designed to function well with millions of connections per second will not be appropriate for use in a network that has a few data connections at greater than gigabit speeds. Typical methods for large corporate IDS and firewalls involve load-balancing network traffic across arrays of gigabit systems. These arrays cannot support individual connection speeds above 1 gigabit but can aggregate to 20 gigabits given enough flows. For security systems to function correctly, the load balancing must be consistent for any given host pair, guaranteeing that no more than 1 gigabit per client can be reached.

2. **Most common services and protocols.**
   For most large corporations, the main interface for their business is a public web server. This is in contrast with the NERSC environment, which provides dozens of unique services along with ad hoc services developed by the user community for their own projects.

   A byproduct of this is that while there is a single general protocol (HTTP) which is used by large corporations, there are dozens of unique protocols used by NERSC. Each of these services represents a different cross section of security threats. In several cases (such as with the Grid services), there are a number of vectors that do not have a significant parallel within the business community. This represents a significant security problem which lives outside the scope of a typical enterprise environment, but which needs to be addressed by NERSC.

3. **Heterogeneous systems.**
   In the same way that the number and variety of applications and protocols is greater in the NERSC environment, the number and variety of operating systems on the DMZ is also larger. For most large corporate web services, there is a well defined operating system and architecture which can be linearly scaled as required for performance reasons. NERSC acquires major systems through an open procurement process and fields multiple systems — often from different vendors — at the same time. This results in a mix of OS types which greatly complicates the problem of whole site security.

4. **Shell access for a geographically diverse user community.**
   A significant difference between the typical web based service providers and NERSC is interactive shell access. Corporate security methods typically center on preventing and detecting shell access. Because a major part of NERSC's mission requires providing this access, a greater level of system auditing and keystroke analysis is required.

   In addition, the user community tends to be spread across multiple DOE HPC facilities as well as dozens of universities. Complications arise with user identification and communication in the event of a compromise. This is in direct contrast to the millions of (relatively) anonymous transactions processed daily by web based service providers.

5. **Interfaces, functions and decomposition.**
   Most commercial software is formally developed, has well defined interface functions, and is decomposed into different functions. This allows the use of application-specific firewalls, restricted virtual machines, and other technology to limit risk. In the open science activities supported by NERSC and others, the scientists are both the developers and the users of the codes. Rather than a few well-defined applications, HPC facilities run hundreds, even thousands of different codes, and the facilities have no control and for the most part little understanding of the application behavior.

These points represent differences between the environments found in most general web service providers and that found at NERSC (as representative of the HPC community). For NERSC, they also represent specific real-world security problems requiring workable solutions that can be extended across multiple sites in a flexible enough manner to address ad hoc virtual organizations which traverse traditional boundaries and relationships.