

Computing and Networking Services

The primary mission of Computing and Networking Services (CaNS) is to provide the infrastructure and computing services within the W.R. Wiley Environmental Molecular Sciences Laboratory (EMSL) for an advanced computing environment that enables staff, visitors, and collaborators to effectively use computer and network resources for their scientific and business requirements. In supporting growing business and research needs of EMSL in the area of information sciences, CaNS secures global information access to our facilities by providing online remote access to both computing resources and scientific equipment. A large portion of the efforts undertaken by CaNS staff members involves providing customer support to EMSL researchers and offsite users. For offsite users, CaNS provides secure information access and dissemination among EMSL researchers and a global scientific user community.

Capabilities

One of the primary roles of CaNS is to provide computer support to all EMSL users and visitors, including scientists, technicians, and support staff. Computing support includes computer procurement assistance, setup, delivery, connection, upgrade installation, and 24-hour computer emergency support. System administration services include configuration management, software upgrades, security standards, account setup, and automated backup. CaNS is responsible for the design and implementation of EMSL's computing infrastructure, software, and Internet application development and support; conference room support; and management of computer maintenance contracts.

Expert Support and Services

- Security
- Desktop computing
- Scientific computing
- Infrastructure design and upgrade
- Web hosting and services
- Software application development and deployment
- Auditorium and conference room support

In 2004, CaNS staff members responded to approximately 6715 formal requests for assistance (a decrease of 4.1% over the 7000 requests made in 2003) and 1500 informal requests. Demonstrating consistency and responsiveness, CaNS staff members resolve 50% of support requests within a day and 90% within a week, giving priority to visiting scientists. Figure 1 shows a breakdown of computer support requests in 2004.

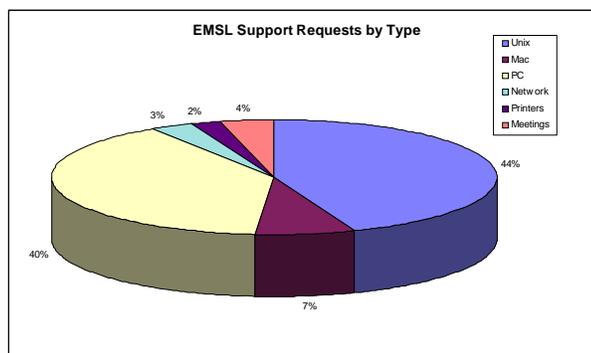


Figure 1. 2004 support requests by type.

EMSL has a history of growth in the number and variety of computer systems used by its staff members (Figure 2). In 2004, 377 systems were added to the support scope, and 406 systems were released, yielding a net decrease of 29 systems (1.5%) and a total of 1941 systems in use in EMSL.

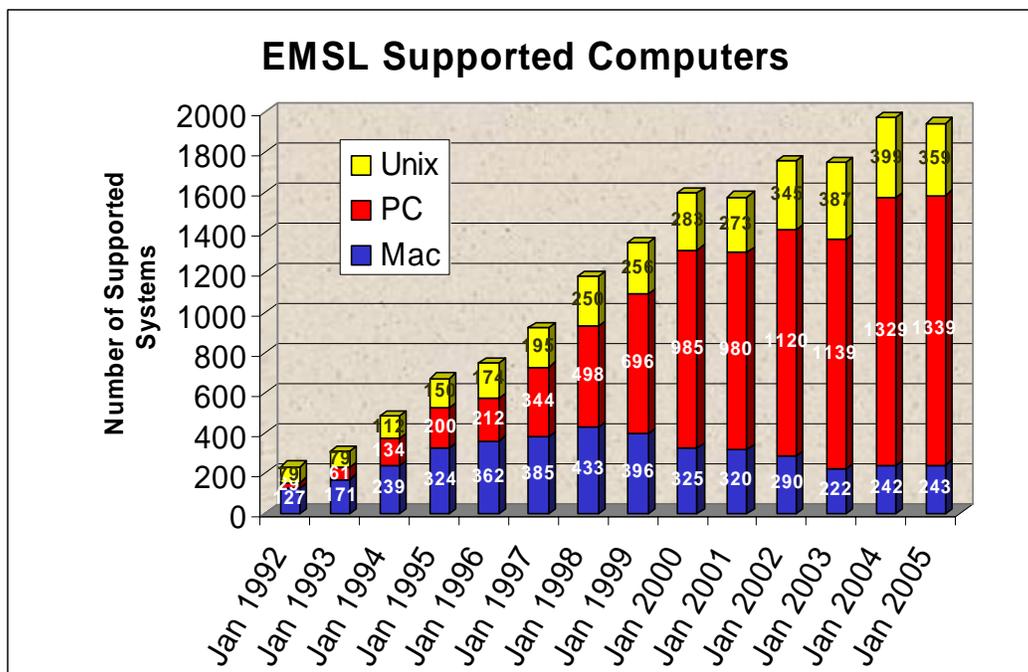


Figure 2. EMSL computer use trends.

In 2004, CaNS added 242 new scientific users, increasing its support scope to 1060 users. Of this total number of users, 498 use EMSL capabilities from offsite.

Support Queues. CaNS uses a software support request application to store and reply to submitted requests, and to record all dialogue surrounding problems. Users benefit by receiving a more thorough response, faster service, and an improved tracking system to ensure their problems are solved expeditiously. CaNS staff members benefit from the ability to better organize their work and identify chronic problems. CaNS also makes its support request system available to other projects in EMSL, such as NWChem and Ecce developers (see Molecular Science Computing Facility Overview section).

In 2004, a web-based product called Footprints was installed to replace the EMSL Support Help Queue (ESHQ) request system, which had been developed internally and had become expensive to maintain. All 24 projects that used the ESHQ system were migrated to Footprints, and the ESHQ system was discontinued.

Data from 2004 show that the 24 help queues were staffed by 99 administrators and experts, who together handled 14,500 support requests. Figure 3 shows historical use of the primary EMSL support queue, support@emsl.pnl.gov. The growth in support requests in 2004 was 10%.

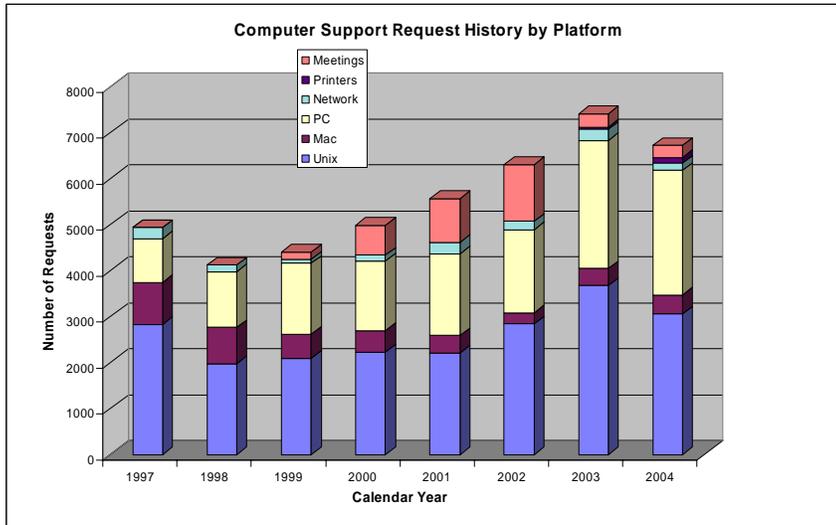


Figure 3. Computer support request history.

Security

A major focus in 2004 involved infrastructure and application upgrades designed to increase the security of EMSL's computing resources while continuing to provide open access for visitors and collaborators. This is an increasingly difficult task as new and more sophisticated security threats arise. Major project areas focused on the upgrade of SecureShell (ssh) access to include SecurID authentication and the creation of the EMSL enclave as a network separate from the Pacific Northwest National Laboratory (PNNL) network. Smaller efforts involved eradication of inherently insecure protocols from our network, including telnet, ftp, rlogin, and earlier versions of ssh.

EMSL Enclave. To encourage open access while protecting PNNL from computational vulnerabilities, EMSL and PNNL implemented the EMSL enclave. EMSL enclave is a series of firewalls that isolate the EMSL computer network from the PNNL network and also from the open Internet.

A hierarchy of access requirements was implemented to provide the EMSL and PNNL networks with appropriate levels of security while protecting them from compromise. Thus, staff and users access the PNNL network remotely from their offices using methods similar to those that would be used to access the PNNL domain from home. Many staff connect daily using Virtual Private Network (VPN) software, which affords security by directing the user through a virtual tunnel into the network, and a SecurID code, which provides user authentication. The VPN software is uniquely configured as a half-tunnel to allow users to connect to the PNNL network simultaneously with the EMSL network.

To better enable data access across enclaves, a Windows-based distributed file system was implemented that provides users with access to shared data that is replicated on servers in two different buildings, thus providing reliability, backup, and disaster survivability. The new file share can be accessed from either the PNNL or EMSL enclaves without SecurID, thus solving a shared data access problem introduced by the EMSL enclave.

The enclave effort also included the development of separate services in the EMSL domain for non-staff, including EMSL visitors and students working in the facility. These groups were provided their own mail services and a terminal server (described below). They are able to access some PNNL services without the need for SecurID authentication.

The enclave implementation provided the foundation for easier access to EMSL by creating the framework for differentiating the EMSL and PNNL computer security requirements. The EMSL enclave implementation during 2004 was the result of a three-year effort involving multidisciplinary teams of staff from both EMSL and PNNL. Continued improvements to the architecture are planned in 2005.

SecureShell Remote Access. The ssh (ssh-gateway) access was completely re-engineered to support multiple enclaves and provide an enhanced level of security. Three new ssh-gateway servers that require SecurID authentication were deployed. SecurID provides a one time use authentication code that protects usernames and passwords from being compromised and then used to gain access to PNNL or EMSL systems. The three new ssh-gateway servers were deployed to both the PNNL and EMSL enclaves. Remote users that do not have a need to access PNNL resources can only use the EMSL gateway. Staff in the EMSL enclave, and remote-use staff and collaborators that require access to PNNL resources must use the PNNL ssh-gateway machine. As almost all EMSL staff now carry and use SecurID to perform their daily work, CaNS capabilities have grown to include administration and management of SecurID tokens and their use for authenticating to the network.

Instrument Networks. In 2004, a new System Update Service was implemented on the PNNL and EMSL networks. This service provides an automatic mechanism to apply security patches to Windows computers. The System Update Service, while ensuring security on Windows systems, is disruptive to scientific processes running on these systems as it can cause a system to reboot automatically, potentially causing loss of data on systems used to run experiments and collect data. Additionally, policy was changed to disallow pre-Windows 2000 operating systems on the network because of security vulnerabilities in older Windows systems. Some software applications used to manage instruments or collect data are no longer supported by the vendor, making an operating system upgrade impractical on these systems.

CaNS implemented several instrument networks within EMSL to solve these problems. An instrument network separates a group of computers from the main network, and provides gateway computers connected to both the PNNL network and the isolated instrument network. The gateway computer provides access and file transfer capabilities to the instrument network. The legacy operating system computers are not on the main network, so they are not subject to operating system upgrade requirements.

Terminal Services. Terminal Services allows staff members to access common applications, such as Microsoft Outlook and Office applications, from a server rather than from their desktop. These servers are frequently used by staff members who do not have a Windows-based system available. There are several advantages to using Terminal Services:

- eliminates the need for VPN running on the desktop

- allows access of email and other applications from home using VPN
- allows access of email and files from any computer.

EMSL Terminal Services was upgraded to better serve the EMSL community. CaNS currently operates four terminal servers that provide services to staff in the EMSL and PNNL enclaves, and to staff who connect remotely. Two of the terminal servers require SecurID authentication and are for users within the EMSL enclave. The third server provides regular access to users in the PNNL enclave. The fourth server is used by EMSL students and visitors and provides only the applications that these users can access.

Virtual VNMR Gateway. The virtual Varian Nuclear Magnetic Resonance gateway was completely re-engineered and upgraded to use SecurID authentication for remote instrumentation control and to allow communication through encrypted ssh communications. This effort resulted in increased security, replacing clear-text passwords with encrypted authentication.

XyLoc Pilot. XyLoc is a hardware device that automatically locks a computer when its user walks away. Upon return, it will automatically unlock the computer, thus eliminating the need to constantly lock the computer and re-enter passwords upon return. XyLoc can be set up for various office settings: one person accessing one computer, one person accessing several computers, or many people accessing a single computer (such as in a laboratory). In an EMSL pilot project designed to further secure the EMSL network and to ascertain the usability of the device, Xyloc was provided to all EMSL staff members who requested it.

Sygate Pilot. CaNS staff members participated in the testing and roll-out of Sygate's Security Agent Software. This product is being implemented in EMSL to mitigate risk surrounding the use of the half-tunnel VPN connections from EMSL to PNNL and to allow the use of new remote access protocols, such as Grid computing. It is being evaluated for the eventual use on all PNNL desktop computers. The Sygate Security Agent provides the following key security features:

- an application firewall that can permit or deny network traffic based on the application initiating the connection
- intrusion prevention capabilities that detect malicious network traffic based on patterns observed in the traffic
- the application of different security policies based on the location of a system
- the ability to protect the system from revealing certain information about itself, such as browser type and operating system
- logging of security-related information.

Desktop Computing

Windows/Macintosh Desktop Support. The EMSL Office Computer Support team provided primary computer support services to more than 1300 Windows and 240 Macintosh computer systems. While the majority of these systems reside on user desktops, some systems are connected to specialized electronic instruments and devices that

support the EMSL research mission. Most of the older Windows operating systems have been replaced with Windows XP and the Office 2003 suite, which provide EMSL users with increased reliability and capabilities, such as the ability to connect remotely to a desktop computer located in EMSL. Macintosh systems in EMSL are primarily G3, G4, and G5 models running Mac 9.x or Mac OS X 10.3 and the Office 2004 suite.

Along with numerous support requests from PC and Macintosh users for personal digital assistant support, wireless configurations, and remote access, the team responded to increased requests related to the newly formed EMSL enclave, including administration of SecurID tokens and EMSL domain accounts for non-staff members.

Enhancements were made to the VisitorLAN wireless network and management of visitor accounts on that network. Badged visitors are now given automatic accounts that allow them to access the wireless network for the duration of their stay. Non-badged visitors who attend conferences and seminars in supported facilities also can gain access. CaNS installed and manages a publicly available computer in the EMSL lobby that allows visitors to access the Internet and their email accounts while visiting the EMSL facility.

Unix Desktop Support.

Linux continues to grow in popularity and is now the dominant Unix desktop operating system, increasing from a 35% share in 2003 to a 57% share in 2004 (Figure 4). We anticipate that Linux will continue to dominate the Unix desktop environment in the coming years.

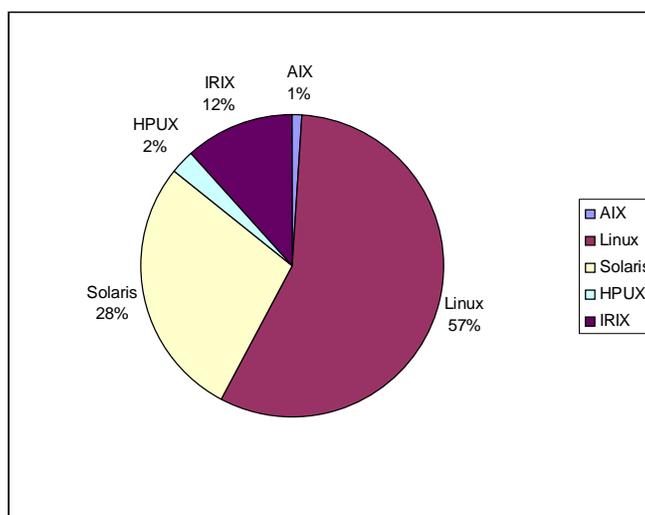


Figure 4. Distribution of Unix operating systems.

With the increase in the use of Linux, an effort was made to standardize a Linux version, and 400 licenses were purchased for Redhat Enterprise Linux version 3. CaNS played a major role in the testing and deployment of this version of Linux, with team members maintaining the central Linux installation server. This new server allows EMSL staff members to easily install Linux and subscribe their system to an update service that automatically updates their computer when new operating system patches are released. The Redhat Linux update service is analogous to the Microsoft Windows Update service.

In 2004, CaNS developed and deployed a new patch and configuration management tool for Unix desktop operating systems. Previous configuration management and patching of supported systems was carried out by a set of custom-written scripts that were replaced in 2004 by an open-source tool called cfengine, which runs on all supported Unix systems. The CaNS team ported the legacy scripts to this new architecture with an emphasis on Linux.

As a result, desktop machines are now automatically checked on an hourly basis for configuration file settings and security patches.

Scientific Computing

Windows Scientific Computing Support.

CaNS supports computers used to control scientific instruments during experiments or to collect data from experiments. A variety of Windows, Macintosh, and other operating systems are installed on these computers. The estimated distribution is detailed in Figure 5.

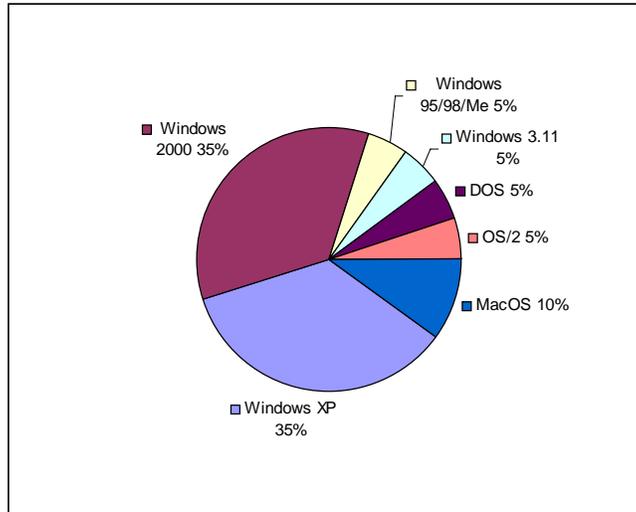


Figure 5. Distribution of Windows/Macintosh operating systems installed on computers used to control research instruments or collect data (excluding office systems).

Data Storage Capability.

A 2-terabyte disk array was installed on the 11-T magnet to accommodate a data storage requirement in the High-Field Magnetic Resonance Facility.

CaNS staff researched, procured, and installed an eight-bay hard disk enclosure. The enclosure has a redundant disk configuration in case of drive failure. The project was a success, and we anticipate procuring and installing more of these arrays as data needs increase.

Clustering. A number of CaNS-supported Linux clusters experienced growth in both CPU power and disk storage this year, including the Ransac and Aeolus clusters described below. In total, more than 12 terabytes of disk space and 100 CPUs were added to six Linux clusters. The number of supported Linux clusters is now 13, and they range in size from 8 to 180 CPUs. Customers made extensive use of our test cluster, which has both Gigabit Ethernet and Myrinet networking. Having both network topologies allows a customer to determine if the added expense of Myrinet is worth the cost.

The System Assessment Capability (SAC) program is an integrated system of computer models and databases used to assess the impact of waste remaining on the Hanford Site. SAC has been developed for DOE and will help decision makers and the public evaluate the cumulative effects of contamination from Hanford. SAC can be run in a stochastic mode and has been used recently on the *Hanford Site Solid Waste Program Environmental Impact Statement* and was also part of the DOE-Battelle award fee for FY 2004. The Ransac cluster of 180 CPUs and 4 terabytes of disk storage is the main compute engine for this project.

The beowulf cluster Aeolus has been used for research on understanding the impacts of climate change at the regional scale. The PNNL-developed regional climate model is the

main software being run on the machine. The regional climate model is a numerical code written in Fortran that simulates the dynamical and physical processes of the climate system at spatial resolution typically about 10 to 60 km. Several long-term simulations are being performed to examine the current and future climate conditions in the United States and assess how climate change may potentially affect water resources, air quality, and ecosystems. The cluster has 32 CPUs and total disk storage of about 18 terabytes, currently almost fully utilized, to archive model outputs for analysis and use in driving impact assessment models.

CaNS staff members are becoming more proficient in Windows server clusters and provided build support for a 26-node (52-processor) cluster that was made for Idaho National Laboratory (INL). We have been asked to participate in two Windows cluster projects for FY 2005 and expect the number to grow.

Infrastructure Design and Upgrades

Collaboration Enclave. EMSL funded a capability enhancement project to implement a collaborative network enclave (CNE) to provide high-performance connectivity that bypasses the PNNL central firewall complex. The CNE project procured infrastructure servers, a network switch, and a prototype data-transfer system. Project staff developed a host-based security model in which the central firewall complex component was replaced by a centrally managed firewall on each CNE system. PNNL will be providing funding in 2005 to integrate the CNE into the PNNL Enterprise Network Security Plan and to add in a hardware-based intrusion detection system. The CNE methodology will be the primary high-performance network for collaboration with external networks and systems when its implementation is complete, allowing EMSL to take advantage of higher Internet bandwidth capabilities anticipated in 2005.

Wireless Upgrade. CaNS staff members, in collaboration with the Wireless Infrastructure Team, brought 802.11a and g wireless networking coverage to EMSL offices and conference rooms. This upgrade significantly increased available wireless bandwidth in EMSL, where the 11-Mbit 802.11b protocol was installed three years ago. This initial deployment and testing of the 802.11a wireless protocol is a precursor to a broader installation throughout PNNL. A wireless device must have 802.11a or g support to take advantage of the new service. The current implementation seamlessly switches among a, b, and g protocols.

Disk-Based Backup. The EMSL disk-based workstation backup system has been performing well. While this system has typically been used to back up Unix systems, the EMSL Windows servers are now also using this same disk-based backup. In 2004, CaNS was able to eliminate three older Windows backup servers and transfer storage capacity to the centralized backup solution.

Distributed File Services. The CaNS strategy for high availability in its distributed file services (Andrew File System or AFS) configuration paid off several times in 2004 as hardware component failures were repaired without disruption of service to users. Besides high availability, the EMSL distributed file system provides uniform access for users to their personal and project areas from any Unix, Windows, or Macintosh system; enhanced and secure file-sharing capability; consistent file backups; access to shared software; access to

web site content; and consistent tools across platforms. Of the 2.4 terabytes of available space, 1.4 terabytes are in use, which leaves plenty of space for both transient files and long-term storage. Figure 6 shows the historical usage versus capacity.

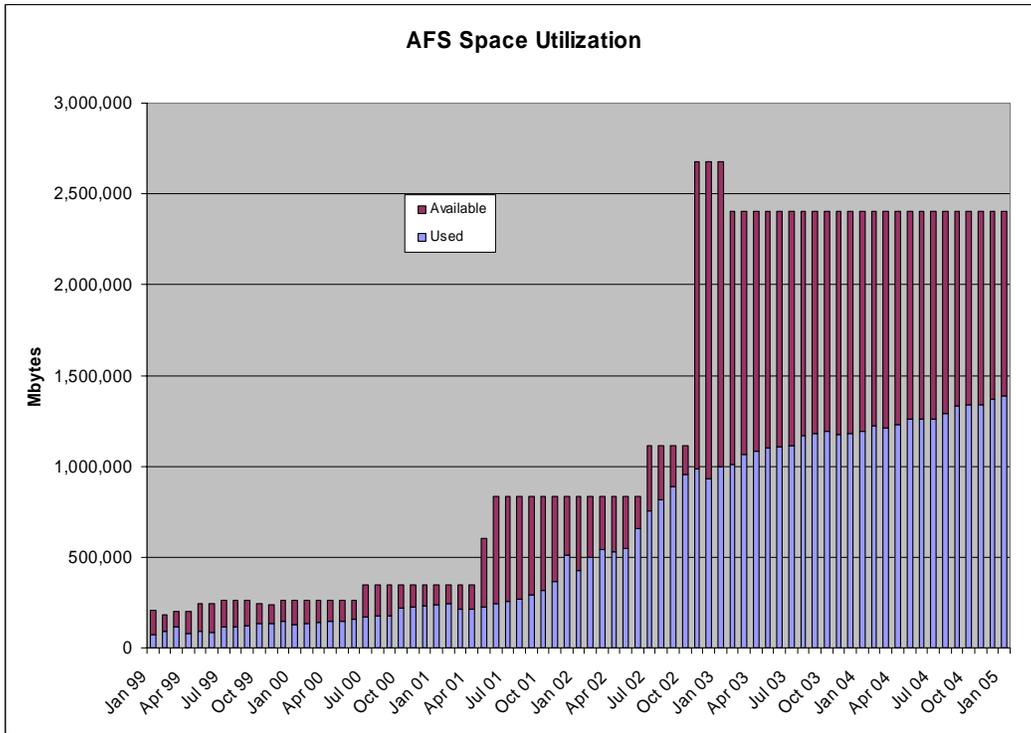


Figure 6. AFS capacity utilization.

Web Hosting and Services

EMSL Web Services. Increased security concerns required EMSL to split its web server and move the server with external access to a more secure configuration. To accomplish this upgrade, the layout of the server configurations was redesigned, and the redesigned layout was implemented. A new server was purchased and deployed to house the external server, and the internal web server was moved to its own system. The external server will be put behind proxy servers in 2005 to manage secure internal and external access.

The EMSL website continues to be moderately busy with about 510,000 visits in 2004. The website absorbed over nine million hits during the year. Figure 7 shows the history of user visits.

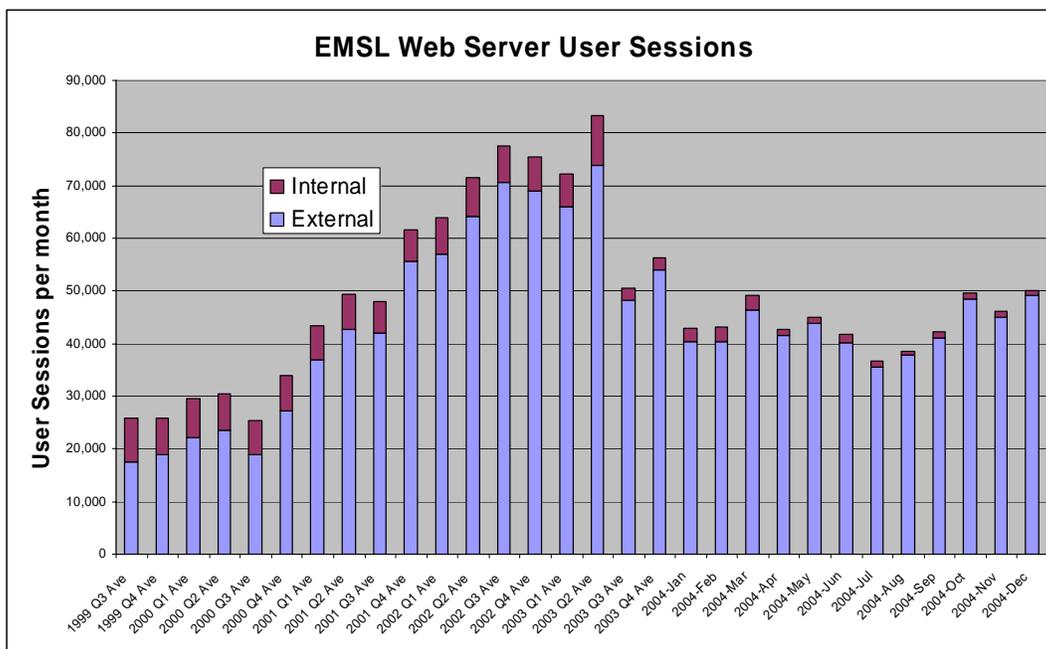


Figure 7. History of user visits.

Software Application Development and Deployment

EMSL User System (EUS). The EUS web-based, work-flow tracking system facilitates the use of EMSL resources, starting from the initial request for use of resources to the tracking of publications resulting from each use. EUS was implemented on October 1, 2002. During 2004, several enhancements were made to the system. EUS assists the EMSL Business Office by automatically sending a request for user agreements and providing standardized text for subsequent emails. Additional restrictions were put into place to limit access to proposal information, especially proprietary proposals. Support for proposal submissions and approvals was provided for Collaborative Access Team and Grand Challenge proposals. User information managed by the EUS has been enhanced to help track EMSL users' publications through an in-house system and to store foreign national visitor assignment and badging information. The EUS now provides over 50 reports that allow EMSL to track statistics on users, such as their professions and institutions, areas of science, and funding sources.

EMSL Resource System (ERS). The ERS web-based tool was developed to track usage of major EMSL scientific instruments and facilities. In 2004, several enhancements were made to the ERS. Several new reports were developed, including one that lists resource usage associated with proprietary proposals and another that lists instrument usage by proposal and publications resulting from a proposal. A recurrence feature was added to simplify scheduling a resource with regularly occurring usage. Scheduling of instruments has been simplified by allowing schedules to be modified by dragging a cursor across the schedule. The ERS project funded an enhancement implemented in the EUS that includes tracking instrument life-cycle information (e.g., cost and life expectancy).

Unix Software Repositories. Software repositories provide shared access to frequently used software including freeware and floating license managers for commercial scientific applications. Applications in the repository include computer languages and interpreters; editors and debuggers; configuration management tools; documentation tools and pre-viewers; communication and collaborative tools; visualization, plotting, and graphing tools; numerical tools and libraries; data acquisition tools; and data management, structure, and format conversion. More than 150 applications available via this repository have been built and compiled to operate over the various operating systems that CaNS supports. The repository is made available via the AFS distributed file system and provides the benefit of consistent applications that users do not have to build themselves.

Auditorium and Conference Room Support

CaNS supported approximately 450 symposia, conferences, and meetings throughout the year. With the help of an Internet media streaming server installed on the EMSL network for electronic streaming of audiovisual presentations, EMSL streamed about a dozen live events, allowing offsite users to view these meetings in realtime via the Internet. CaNS worked with other staff in a collaborative effort to evaluate new streaming software. They also supported CD and DVD creations. Major events supported by CaNS in 2004 include:

- “Russian Delegation for the Nuclear Cities Initiative” was a collaboration between Russian delegates and PNNL. It involved presentations running simultaneously on two screens in both Russian and English, and a public address system and five microphones used for a panel discussion with both Russian and English interpreters.
- PNNL’s Computational Information Sciences directorate sponsored “Onramp to the Future,” which involved many external speakers from organizations such as Microsoft and IBM. CaNS support included videotaping the events for subsequent presentation on the World Wide Web, and producing and creating website pages that contain all the viewgraphs and video.
- “Partnership Discussion Day” was a collaboration among PNNL, Washington State University, Oregon State University, and others to promote communication or science needs between universities and colleges. CaNS staff gathered PowerPoint files from the meeting and created CD handouts of the presentations.
- The “Explorations in Nanoscale Science and Engineering Workshop” workshop put on by Johns Hopkins University Center was attended by almost 90 high school students and their parents. CaNS supported the event, which held activities in the Battelle Auditorium at PNNL, the EMSL auditorium, two conference rooms, and two EMSL laboratories.

Staff

Linda M. Connell, Technical Lead
(509) 376-2375, linda.connell@pnl.gov

Cheryl E. Pomeroy, Administrator
(509) 376-2499, cheryl.hall@pnl.gov

Scott W. Campbell, IT Specialist
(509) 376-2952, scott.campbell@pnl.gov

Timothy S. Carlson, Senior Research Scientist
(509) 376-3423, timothy.carlson@pnl.gov

Boyd L. Champion, IT Specialist
(509) 376-2287, boyd.champion@pnl.gov

Patricia M. (Shelly) Fangman, Research Scientist
(509) 376-5382, shelly@pnl.gov

Phillip P. Gackle, Senior Research Scientist
(509) 375-2130, philip.gackle@pnl.gov

Dan N. Hagedorn, Senior Research Scientist
(509) 376-8206, dan.hagedorn@pnl.gov

Cheryl A. Wilmoth, IT Specialist
(509) 376-2327, cheryl.hartwell@pnl.gov

Zoë C. Johns, Research Scientist
(509) 376-2038, zoe.johns@pnl.gov

Bryce P. Kaspar, IT Specialist
(509) 376-3949, bryce.kaspar@pnl.gov

Jeffrey Larsen, IT Specialist
(509) 376-1729, jeff.larsen@pnl.gov

Tim J. Lawson, Research Scientist
(509) 376-2137, timothy.lawson@pnl.gov

Thomas J. Mathieu, User Support Coordinator, Senior Research Scientist
(509) 376-9711, thomas.mathieu@pnl.gov

Nicholas A. Nanni, Technician
(509) 376-1714, nick.nanni@pnl.gov

Douglas J. Nordwall, Senior Research Scientist
(509) 376-4308, doug@pnl.gov

Shaun T. O'Leary, Senior Research Scientist
(509) 376-3505, shaun@pnl.gov

Andrew E. Rakowski, Senior Research Scientist
(509) 375-2238, landrew@pnl.gov

Michael S. Russcher, IT Specialist
(509) 376-0300, mike.russcher@pnl.gov

Aaron I. Thronas, Research Scientist
(509) 376-1170, aaron.thronas@pnl.gov

Ralph S. Wescott, Senior Research Scientist
(509) 372-6901, ralph.wescott@pnl.gov

Matrixed Staff

Gerald A. Funnell, Electrician
(509) 372-6276, gerald.funnell@pnl.gov

Bruce J. Herrin, Electrician
(509) 376-0512, bruce.herrin@pnl.gov

Students

Tyson A. Tucker, Student
(509) 528-8169, tyson@pnl.gov