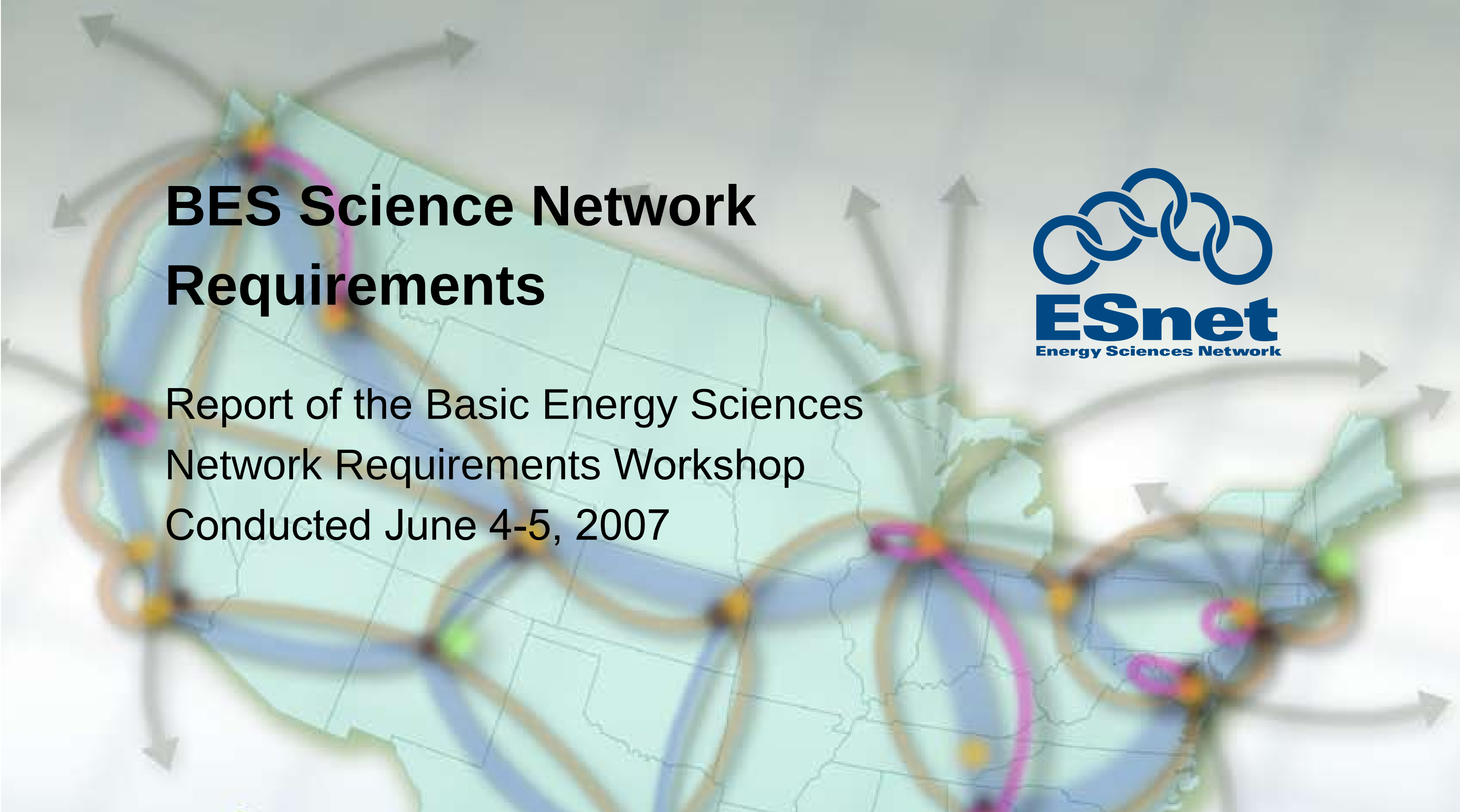
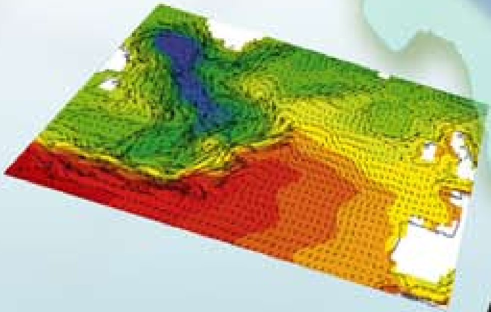
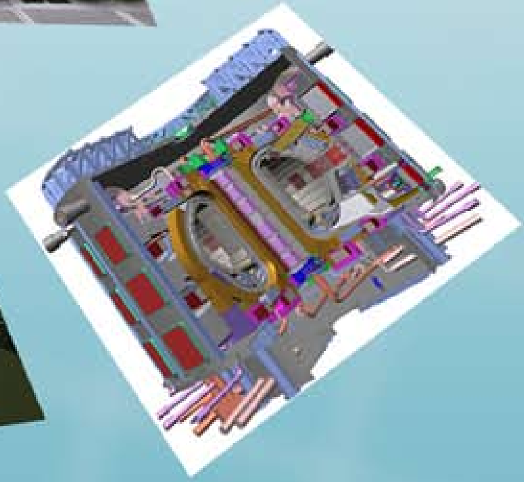


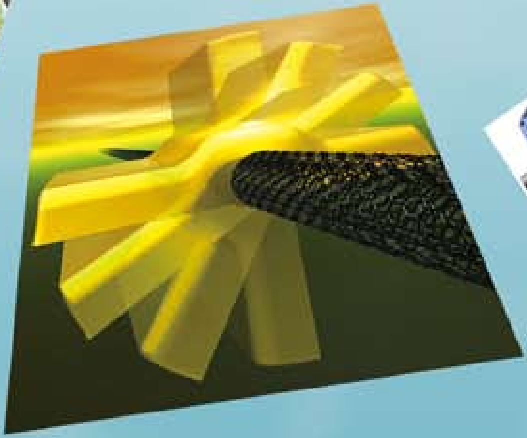
BES Science Network Requirements

ESnet
Energy Sciences Network

Report of the Basic Energy Sciences Network Requirements Workshop Conducted June 4-5, 2007







BERKELEY LAB

Office of Science
U.S. DEPARTMENT OF ENERGY

BES Science Network Requirements Workshop

Basic Energy Sciences Program Office, DOE Office of Science
Energy Sciences Network
Washington, DC – June 4 and 5, 2007

ESnet is funded by the US Dept. of Energy, Office of Science, Advanced Scientific Computing Research (ASCR) program. Dan Hitchcock is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory, which is operated by the University of California for the US Department of Energy under contract DE-AC02-05CH11231.

This work was supported by the Directors of the Office of Science, Office of Advanced Scientific Computing Research, Facilities Division, and the Office of Basic Energy Sciences.

This is LBNL report LBNL/PUB-981.

Participants and Contributors

Dohn Arms, ANL (Advanced Photon Source)
Anjuli Bamzai, DOE/SC/BER (BER Program Office)
Alan Biocca, LBNL (Advanced Light Source)
Jackie Chen, SNL (Combustion Research Facility)
Eli Dart, ESnet (Networking)
Bert DeJong, PNNL (Chemistry)
Paul Domagala, ANL (Computing and Information Systems)
Yiping Feng, SLAC (LCLS/LUSI)
David Goodwin, DOE/SC/ASCR (ASCR Program Office)
Bruce Harmon, Ames Lab (Materials Science)
Robert Harrison, UT/ORNL (Chemistry)
Richard Hilderbrandt, DOE/SC/BES (BES Program Office)
Daniel Hitchcock, DOE/SC/ASCR (ASCR Program Office)
William Johnston, ESnet (Networking)
Roger Klaffky, DOE/SC/BES (BES Program Office)
Michael McGuigan, BNL (Center for Functional Nanomaterials)
Stephen Miller, ORNL (Spallation Neutron Source)
Richard Mount, SLAC (Linac Coherent Light Source)
Jeff Neaton, LBNL (Molecular Foundry)
Larry Rahn, SNL/BES (Combustion)
Thomas Schulthess, ORNL (CNMS)
Ken Sidorowicz, ANL (Advanced Photon Source)
Ellen Stechel, SNL (ASCAC)
Brian Tierney, LBNL (Networking)
Linda Winkler, ANL (Networking)
Zhijian Yin, BNL (National Synchrotron Light Source)

Editors

Eli Dart, ESnet Engineering Group – dart@es.net
Brian Tierney, LBNL Distributed Systems Department – bltierney@lbl.gov

Contents

1 Executive Summary

The Energy Sciences Network (ESnet) is the primary provider of network connectivity for the US Department of Energy Office of Science, the single largest supporter of basic research in the physical sciences in the United States of America. In support of the Office of Science programs, ESnet regularly updates and refreshes its understanding of the networking requirements of the instruments, facilities, scientists, and science programs that it serves. This focus has helped ESnet to be a highly successful enabler of scientific discovery for over 20 years.

In June 2007, ESnet and the Basic Energy Sciences (BES) Program Office of the DOE Office of Science organized a workshop to characterize the networking requirements of the science programs funded by the BES Program Office. These included several large facilities, including the Advanced Photon Source at ANL, the Advanced Light Source at LBNL, the Combustion Research Facility at SNL, the Linac Coherent Light Source at SLAC, the Center for Functional Nanomaterials at BNL, the Spallation Neutron Source at ORNL, the Molecular Foundry at LBNL, the National Center for Electron Microscopy at LBNL, and the National Synchrotron Light Source at BNL. It also included input from the computational chemistry community.

Workshop participants were asked to codify their requirements in a “case study” format, which summarizes the instruments and facilities necessary for the science and the process by which the science is done, with emphasis on the network services needed and the way in which the network is used. Participants were asked to consider three time scales in their case studies – the near term (immediately and up to 12 months in the future), the medium term (3-5 years in the future), and the long term (greater than 5 years in the future).

In addition to achieving its goal of collecting and characterizing the network requirements of the science endeavors funded by the BES Program Office, the workshop emphasized some additional points. The most important of these is that there is widespread frustration involved in the transfer of large data sets between facilities, or from the data’s facility of origin back to the researcher’s home institution. The reasons for the difficulties vary, but the broad categories are lack of modern data transfer tools, lack of knowledge of host system tuning, and congestion or similar issues at “the other end” of the data transfer. To help address this, ESnet will enhance and expand its network performance tuning web pages to include information describing modern data transfer software, host setup and tuning, and network configuration.

In addition, it is clear from the case studies that many facilities funded by BES use similar technology for data detection and acquisition, and that these detectors are on a “Moore’s Law” growth curve in terms of their data production capabilities. This means that the data volume produced by the Light Sources, Nanoscience Centers, and the Neutron Science facilities will grow quickly, placing ever-increasing demands on storage and computational resources, as well as the wide area network.

2 DOE Basic Energy Sciences Facilities and Programs

All of the major DOE BES-funded facilities were represented at the workshop. The following sections described the data and networking requirements of each of these facilities.

2.1 Advanced Light Source at LBNL

Background

The Advanced Light Source (ALS) is a national user facility at Berkeley Lab that generates intense light for scientific and technological research. The ALS is one of the world's brightest sources of ultraviolet and soft x-ray beams, and the world's first third-generation synchrotron light source in its energy range. The ALS operates 40 beamlines and hosts more than 2,000 distinct users annually. ALS is a 24-hour operational facility and beams are available to users in excess of 4,000 hours per year.

The network requirements of ALS beamlines vary, and several beamlines at the ALS, facilitating multiple scientific programs, would benefit from enhanced network capabilities. With 40 beamlines collecting data simultaneously, data flow can be high. The ALS assists users both on-site and off, and datasets are shared over the network and via portable media. The data rates and size can vary widely depending on the experiment. Not all beamlines are data intensive, but there are a number of X-ray detectors struggling with high data accumulation rates. Examples of these include programs in x-ray tomography, protein crystallography, and microdiffraction. Future directives may also have significant network needs. Of the existing programs, the most data intensive at this stage are x-ray tomography and protein crystallography. Of the two, x-ray tomography will eventually be most strongly network-limited. X-ray tomography is a technique whereby the internal structure of an object can be visualized in three dimensions. With a synchrotron x-ray source, it is now possible to obtain tomographic images of micron-size systems spanning an unprecedented range of sizes and compositions, including ‘hard’ materials, such as samples of concrete; dense biological materials such as bones or teeth; and ‘soft matter’, such as an individual biological cells. Related work at Berkeley Lab with application to soft matter and biological systems is carried out at National Center for X-ray Tomography (NCXT).

Current Network Requirements and Science Process

At present, x-ray tomography is quite inefficient: current detectors collect flux for 5-10 ms, and readout for 1 s, resulting in a collector scheme with just 1% efficiency. The detector is a ~10M pixel camera. Tomography images are currently written to disk over standard GB Ethernet at speeds of ~ 100Mbs. About 1000 images make up a typical data set (10 GB). Users can produce up to 1 data set/hour, and over the course of 24 hours, 0.5 TB is routinely accumulated. Users currently take their data away on a portable hard drive, but they would greatly prefer to transfer their data via the network. As technology improves, via e.g. multiple parallel readout port schemes, readout times will diminish to less than 10 ms. Network demands will substantially increase concurrently, as images are generated every few milliseconds.

Techniques for tomographic reconstruction are very disk intensive, as ~ 1000 images need to be open simultaneously, and a high-performance workstation or cluster is required. The time required to complete the reconstruction depends on the dataset and the algorithm, and can range from 30 minutes to several hours. Subsequent visualization of 3D data is just possible with commercial software (currently, AMIRA) and a fast workstation. In the future, a centralized server for analysis of 3D data sets would require a fast network to move and manipulate TB image files.

Currently most users with large datasets take them away on portable hard disk or DVD. Though the present network could in principle handle some of these datasets, there are fundamental data transfer difficulties that most users have not overcome. Some of these issues might be addressed through increased awareness of existing tools and data transfer technology, which would better facilitate effective data transfer outside leading-edge communities, such as high-energy physics, climate modeling, and combustion. ESnet provides the network hardware, but reliable efficient “cargo-carrying” infrastructure is still needed for many scientists, including those at the ALS.

Network Requirements and Science Process – the next 5 years

In the next five years and beyond, dramatic improvements in detector technology are projected to enhance the tomography readout rate 100-fold, further improving efficiency and increasing the experimental data throughput. Near-term detector advances and higher x-ray fluxes will translate to faster data output of between 0.1-1GB/sec. Such advances will enable new modes of scientific discovery, such as “dynamic tomography”, where the dynamical evolution of micro or nanostructures can be observed through multiple tomographs taken in real time; this mode of measurement, and its associated analysis, would result in significant additional network demands. Faster local networks (InfiniBand, Myrinet) will be required to transfer data. Data analysis would be performed on clusters. A shared, distributed high-bandwidth network of clusters would be highly desirable for data analysis.

Beyond 5 years – Future Needs and Scientific Direction

In the long term, parallel detectors will increase readout speed, resulting in the 100-fold improvement mentioned above. The projected parallel CCD output rate (200 frames per second) exceeds currently available network bandwidth. Data set size is dictated by primarily by detector size and readout rate. These characteristics are driven by the same technology issues that drive computer performance – and are characterized by Moore's law (a doubling of performance every 18 months), and with the advent of parallel detectors, future networks will have to cope with Moore's law compounded with parallel readout. This will present significant challenges in data readout, storage, transportation and computations needed to extract the science from the data. Bandwidth requirements of 10s of Gbps are expected.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	Network Services
Near-term	• ALS has 40 beamlines, generating significant dataflow	• Typical x-ray tomography experiments generate 10 Gbytes datasets, at rates of 200MB/sec	• 3-5 Gbps	• Tools to facilitate moving large scientific datasets across the network
3-5 years	• ALS may have 50 beamlines, many with larger detectors that have significantly faster readout rates	• Data set size increase to 100 GBytes	• 10 Gbps, including need for remote visualization and analysis	• End-to-end quality of service
5+ years	• Remote control of experiments • “Dynamic tomography”, generating 4d (including time) data sets. • continued higher resolution detectors	• Real-time access to data and remote operation of the experiment • Analysis and visualization performed on a remote compute cluster or grid	• 10’s of Gbit/s • Quality of service for network latency and reliability	• Parallel network I/O between experiments, and visualization and analysis tools • Large scale datasets require cluster or grid computing for analysis

2.2 ANL BES Community External Network Needs

Background

Argonne National Laboratory is a multi-purpose lab with strong programs in all major applied and basic science disciplines. For the purposes of this summary, Basic Energy Science (BES) activities at Argonne can be broadly characterized as user facilities, computational sciences & facilities and basic physical science research. The following provides a summary of activities in these, insight into the scientific processes, resulting data and networking needs and the collaboration ecosystems surrounding the research.

User Facilities

As part of its science mission, Argonne designs, builds, operates and manages many scientific research facilities and makes them available to outside researchers from industry, academia and other government laboratories. Argonne user facilities provide world-class, unique capabilities and serve as discovery tools not only for many scientific programs at Argonne, but the nation's science R&D complex. Argonne operates six national user facilities for the U.S. Department of Energy's Office of Basic Energy Sciences:

- Advanced Photon Source
- Argonne Tandem-Linac Accelerator System
- Atmospheric Radiation Measurement Climate Research Facility

- Center for Nanoscale Materials
- Electron Microscopy Center
- Intense Pulsed Neutron Source

2.2.1 The Advanced Photon Source (APS)

The Advanced Photon Source (APS) at the U.S. Department of Energy's Argonne National Laboratory provides this nation's (in fact, this hemisphere's) most brilliant x-ray beams for research in almost all scientific disciplines. These x-rays allow scientists to pursue new knowledge about the structure and function of materials in the center of the Earth and in outer space, and all points in between. The knowledge gained from this research can impact the evolution of combustion engines and microcircuits, aid in the development of new pharmaceuticals, and pioneer nanotechnologies whose scale is measured in billionths of a meter, to name just a few examples. These studies promise to have far-reaching impact on our technology, economy, health, and our fundamental knowledge of the materials that make up our world.

Scientific Access

The APS is a open user facility that makes beam time available to the international scientific community through a peer-reviewed proposal process. Two access modes are available: General Users and Partner Users. General Users are those who require less than 10% of the beam time on a beamline in a given cycle. Partner Users are those whose work involves a greater scope and greater commitment by both the user and the APS. Specific requirements govern both modes of access.

APS Synchrotron Techniques

The unique properties of synchrotron radiation are its continuous spectrum, high flux and brightness, and high coherence, which make it an indispensable tool in the exploration of matter. The wavelengths of the emitted photons span a range of dimensions from the atomic level to biological cells, thereby providing incisive probes for advanced research in materials science, physical and chemical sciences, metrology, geosciences, environmental sciences, biosciences, medical sciences, and pharmaceutical sciences. The features of synchrotron radiation are especially well matched to the needs of nanoscience. This breadth of problems requires an extensive suite of probes. The basic components of a beamline, however, share general similarities as shown in the schematic diagram below.

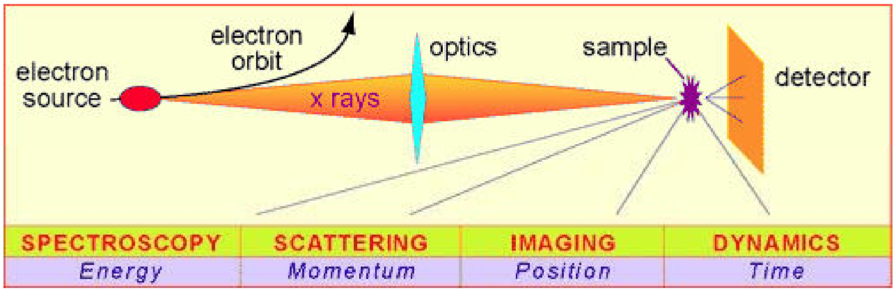


Figure 1 APS Synchrotron Techniques

The fundamental parameters that we use to perceive the physical world (energy, momentum, position, and time) correspond to three broad categories of synchrotron experimental measurement techniques: spectroscopy, scattering, and imaging. By exploiting the short pulse lengths of synchrotron radiation, each technique can be performed in a timing fashion.

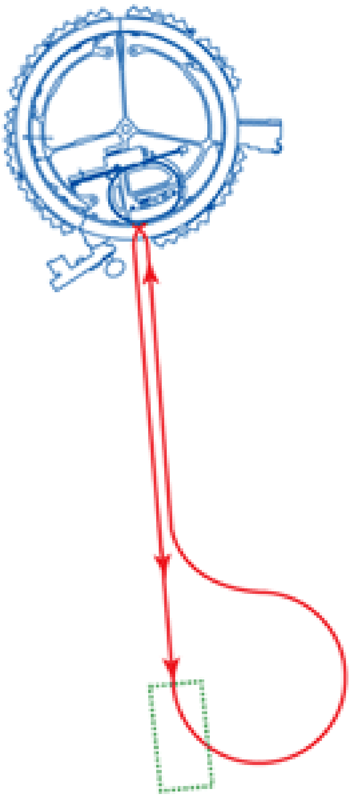
Spectroscopy is used to study the energies of particles that are emitted or absorbed by samples that are exposed to the light-source beam and is commonly used to determine the characteristics of chemical bonding and electron motion.

Scattering makes use of the patterns of light produced when x-rays are deflected by the closely spaced lattice of atoms in solids and is commonly used to determine the structures of crystals and large molecules such as proteins.

Imaging techniques use the light-source beam to obtain pictures with fine spatial resolution of the samples under study and are used in diverse research areas such as cell biology, lithography, infrared microscopy, radiology, and x-ray tomography.

The APS Upgrade

Advanced Photon Source (APS) has been operating for more than 11 years, and the capabilities and user impact of this facility grow continuously. Given the scope and productivity of the APS user community, it is essential that we plan for continuing improvements to our premier facility in order to deliver state-of the-art x-ray beams beyond the next decade. There is a wide spectrum of activities needed for progress, from beamline optics, detectors, and end-station equipment, to major accelerator upgrades. Innovative new machine concepts, especially so-called fourth-generation sources, are coming on the horizon.



Through extensive discussions with our community, we have come up with some serious options that have been reviewed for technical feasibility by an external machine advisory committee. At this point, we favor the concept of the energy recovery linac (ERL), which offers revolutionary properties, especially for x-ray imaging and ultrafast science

An APS upgrade and the associated next generation instrumentation would likely result in close to an order of magnitude increase in data production and therefore data transport needs. Much of the increase would translate to external bandwidth requirements.

2.2.2 Argonne Tandem-Linac Accelerator System (ATLAS)

ATLAS is a National User Facility and as such hosts 200 - 300 Users each year. It is supported by the Office of Nuclear Physics of the Department of Energy. The Users come from U.S. universities and national laboratories as well as from foreign institutions. The facility is also accessible to industrial Users.

The ATLAS facility is a leading facility for nuclear structure research in the United States. It provides a wide range of beams for nuclear reaction and structure research to a large community of users from the US and abroad. These beams are used mostly to study nuclear reactions of astrophysical interest and for nuclear structure investigations. ATLAS provides a wide variety of experimental instruments for a large community of nuclear scientists. In 2006, there were 436 active users, including 75 graduate students.

2.2.3 Atmospheric Radiation Measurement Climate Research Facility

Argonne manages the Atmospheric Radiation Measurement (ARM) Program for the U.S. Department of Energy (DOE), operating outdoor research stations on the Southern Great Plains , the North Slope of Alaska and the Tropical Western Pacific.

ARM is the DOE's largest global change research program. It was created to help resolve scientific questions related to global climate change, with a specific focus on the crucial role of clouds and their influence on radiative feedback processes in the atmosphere. ARM's primary goal is to improve the treatment of cloud and radiation physics in global climate models so they can more accurately simulate and predict future climate conditions.

Figure 2 ARM Sites

Each site has been heavily instrumented to gather massive amounts of climate data. Using these data, scientists are studying the effects and interactions of sunlight, radiant energy, and clouds to understand their impact on temperatures, weather, and climate.

In addition to our permanent facilities, the ARM Mobile Facility will provide the Program with the capability of performing atmospheric measurements similar to those at the other ARM sites for periods up to a year at a time anywhere in the world.

2.2.4 Center for Nanoscale Materials

The recently constructed Center for Nanoscale Materials (CNM) is a 83,000 gross-square-foot structure that includes advanced research facilities for interdisciplinary nanoscience and nanotechnology research.

The key feature of the CNM is its ability to support all stages of research on nanoscale materials, from synthesis and patterning through metrology, compositional and structural determination, and physical phenomena characterization.

The CNM provides cleanroom laboratories, specialized research laboratories, offices and conference space, and space for technical and office support personnel.

The facility adjoins the west side of the Advanced Photon Source (APS) and is designed to be compatible with existing APS structures. Through its physical attachment to the APS, CNM researchers will benefit from the unique capabilities of the APS X-ray source, which will facilitate the study of materials and phenomena at nanoscale resolutions. Capabilities offered by at the CNM will include synthesis and fabrication of nanoscale materials, instruments to characterize the fundamental properties of nanoscale materials, a dedicated beamline at the APS (hard X-ray nanoprobe), advanced data-collection systems and remote-presence capabilities, training and collaborative outreach facilities simulation, modeling and visualization utilizing dedicated high performance computing resources. These operations are inherently data and bandwidth intensive. The CNM will heavily leverage remote collaboration capabilities. Aggregate raw data production from CNM is expected to be in the TB/week range. Add to this routine traffic and traffic from collaboration services and internal bandwidth requirements could reach peaks of 40Gbps. How much of this will be external to CNM and how much will traverse ESnet is unknown at this time.

The CNM will be fully operational by Fall 2007.

2.2.5 Electron Microscopy Center (EMC)

The EMC Conducts materials research using advanced microstructural characterization methods, maintains unique resources and facilities for scientific research for the both the Argonne and the national scientific community and develop and expand the frontiers of microanalysis by fostering the evolution of synergistic state-of-the-art resources in instrumentation, techniques and scientific expertise. The EMC maintains extensive capabilities for remote tele-presence microscopy, the capabilities of which are expected to be expanded upon on an ongoing basis.

Currently, EMC is building a new state-of-the-art laboratory space for advanced electron microscopy. The Sub-Angstrom Microscopy and Microanalysis Facility (SAMM) will house four new instrument rooms all of which will be remotely operated and available to researchers throughout the world.

2.2.6 The Intense Pulsed Neutron Source (IPNS)

The Intense Pulsed Neutron Source at Argonne National Laboratory is a National User Facility for performing neutron scattering experiments to determine the properties of materials by studying atomic arrangements and motions in liquids and solids. IPNS provides qualified users with reliable optimized neutron scattering instruments; provides users the assistance of experienced scientific and technical staff; and ensures the safe and timely completion of users' experiments. The IPNS accelerator system consists of an H-ion source, a Cockcroft-Walton preaccelerator, a 50 MeV Alvarez linac, a 450 MeV Rapid Cycling Synchrotron (RCS), transport lines and ancillary subsystems (controls, diagnostics, services).

IPNS operates 13 major investigational instruments that are utilized by an international user community.

2.2.7 Argonne Leadership Computing Facility

At Argonne, efforts have focused on the application of high-performance computing systems to the study of phenomena that cannot easily be approached via theory and/or experiment alone. Also of interest is the study of the methods and tools by which high-performance computing applications are constructed and the fundamental limitations of numerical simulation. The data intensive nature of computational science has raised the importance of the development and application of methods and tools required to make sense of massive quantities of multidimensional data, including the efficient collection and organization of that data, and the reduction of dimensionality via data mining, visualization, and other techniques.

Projects making use of ANL Laboratory Computing Resource Center (LCRC) resources represent a wide cross-section of Laboratory expertise, including work in biosciences, chemistry, climate, computer science, engineering applications, environmental science, geoscience, information science, materials science, mathematics, nanoscience, nuclear engineering, and physics.

High-performance computing systems planned for deployment in 2009 and later have the potential to achieve a petaflop/s of computing power. If these systems fulfill their promise, they will definitely change the nature of scientific questions that can be pursued via simulation. To this end, Argonne has established the Argonne Leadership Computing Facility (ALCF).

Bioinformatics

Bioinformatics has become essential in order for scientists to analyze and manage the unprecedented growth of genomic and molecular data. At Argonne, our goal as computational biologists is to develop new methods and technologies for acquiring, organizing, analyzing, and visualizing such data. We apply these methods and technologies in two principal areas: sequence analysis, comparative analysis, phylogenetic and evolutionary studies; the systems-centered biology and statistical genomics.

Today, scientists have accumulated enormous volumes of genomic and enzymatic data. Analysis of such genomic data is complicated and computationally intensive.

Modeling of the Earth System

The grand challenge for climate modeling is to predict future climates based on scenarios of anthropogenic emissions and other changes resulting from options in energy policies. The challenge for this project is to transform an existing, state-of-the-science, third-generation global climate model, the Community Climate System Model, to a first-generation Earth system model that fully simulates the coupling between the physical, chemical, and biogeochemical processes in the climate system. This project will improve the representation of carbon and chemical processes, particularly for treatment of greenhouse gas emissions and aerosol feedbacks in collaboration with the DOE Atmospheric Science Program, DOE Atmospheric Radiation Measurement Program, and DOE Terrestrial Carbon Programs.

Methods to improve scalability to thousands of processors are being introduced, maximizing the length and number of simulations that can be performed, and facilitating the aggressive schedule of simulations required for scheduled national and international climate change assessments. Computations tend to be CPU and memory latency bound. Post-processing requires a great deal of memory and disk I/O. Collaborators include LANL, ORNL, NCAR, LBNL, University of Chicago, UW-Madison. Datasets generated by simulation codes are compared to coordinated measurement programs. Analysis is performed on maps from computer simulation data compared with maps of gridded observations. Software requirements include: simulation codes, MPI, custom/specialized viz and analysis tools for meteorological statistics and map projections. Data management is ad-hoc and individually controlled for the most part. Some simulations are stored at central sites. No central meta data service exists for all simulations and observational data. Some shipping of data occurs by postal system because scientists want *all* of the data locally for analysis. These can be 10s-100s of TB. In the future it would be advantageous if data could be left where it was generated and visualizations and analysis tools could know how/if to move it as needed for calculations.

Molecular Dynamics

Scientists use theoretical and computational methods to advance understanding of the structure, dynamics and function of biological macromolecular systems at the atomic level. The computational approach called molecular dynamics (MD) consists of constructing detailed atomic models of the macromolecular system and, having described the microscopic forces with a potential function, using Newton's classical equation, to literally "simulate" the dynamical motions of all the atoms as a function of time. The calculated trajectory, though an approximation to the real world, provides detailed information about the time course of the atomic motions, which is nearly impossible to access experimentally. All-atom MD simulations are used to rigorously compute conformational free energies, and binding free energies.

Scientific discovery is carried out through the development of effective simulation strategies and analysis of simulation data. Data is unstructured and managed by individuals. Myproxy authentication is utilized. They are developing workflow with swift (http://www.ci.uchicago.edu/swift/) for the rapid and reliable specification, execution, and management of large-scale science workflows. Moving toward gridftp for data movement.

Physical Science Research

Basic Science seeks solutions to a wide variety of scientific challenges. This includes experimental and theoretical work in materials science, physics, chemistry, biology and high-energy physics. Of these, biology and high-energy physics have the most intensive computing and networking needs.

Biology and in particularly the computational biology disciplines as mentioned above result in very large data sets that are made available to collaborators worldwide.

High-energy physics have similar needs, typically transporting data sets in the 10’s of Gb range on a regular bases between accelerator facilities such as CERN, Fermi and Brookhaven as well as university collaborators. The high-energy physics community at Argonne also makes extensive use of grid computing facilities.

Current Network Requirements and Science Process at ANL

Argonne maintains 2 main network connections – a recently upgraded 10Gbps ESnet connection and a 10Gbps Metropolitan Research and Education Network (MREN) connection. The balance of traffic over these links is approximately 75% and 25% respectively. Baseline aggregate utilization for the laboratory is approximately 300Mbps with spikes reaching into the 2Gbps range.

Currently, the most demanding networking needs can be found in the user facilities and computational areas which account for approximately 75% of current external bandwidth requirements. Occasional needs exist in these areas for high-bandwidth, or dedicated bandwidth connections to other research institutions. Basic physical science programs on the other hand can generally be characterized by commodity IP traffic around the range of 2% of available bandwidth with steady incremental growth.

Network Requirements and Science Process – the next 5 years

Near-term growth in network needs will be driven by expansion of user facilities, higher instrument resolution and resulting bandwidth demands. Higher resolution and real-time science in turn, is driving simulation and visualization in a similar direction, the net result being the potential for higher growth rates over the next 5 year.

Another important factor is a significant paradigm shift in the way science is carried out. Faster, cheaper and more capable internet-connected instruments, a richer suite of collaboration tools and a research community that knows how to use them and demands high levels of connectivity and availability have already begun to transform high-availability, high-performance networks from an internal necessity to a global necessity for science. To address these challenges, Argonne has embarked upon the Digital Laboratory Initiative (DLI) which is aimed at enabling science through the aggressive and strategic use of IT resources. It is anticipated that DLI will have a significant impact on Argonne's external connectivity requirements.

The other consequence of a highly-mobile, inherently digital diverse scientific community is the increase in geographic reach. No longer are users and collaborators confined to high performance research networks. They reside on just about every network around the world and expect access to their data and experiments from anywhere. Science has become globalized. It’s important therefore that ESnet maintain strong peerings with and sufficient bandwidth to other research, educational and commercial networks.

Major Collaboration Sites for ANL

The primary endpoints for high-speed data transfers into out ANL for BES facilities are Fermi National Lab, Brookhaven National Lab, UCLA, and CERN.

Summary Table for ANL facilities

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	**Network Services**
Near-term	• APS, CNM, SAMM, ARM	• 24x7 operation, • multi-gigabyte data sets • remote operation, • analysis / visualization capabilities	• Current 10 Gbps • Occasional need for dedicated circuits	• Grid / PKI infrastructure • Video conference bridging services
3-5 years	• Addition of Digital Laboratory Initiative	• increases in scientific collaboration services	• 10-20 Gbps • Occasional need for dedicated circuits • Good connectivity to commercial providers	• Grid / PKI infrastructure • Video conference bridging services • Additional collaboration services
5+ years	• APS/ERL upgrade		• 100 Gbps • Occasional need for dedicated circuits • Better connectivity to commercial providers	• Grid / PKI infrastructure • Video conference bridging services

2.3 Center for Functional Nanomaterials at BNL

The Center for Functional Nanomaterials is a DOE funded facility whose mission is to bring the technological promise of nanoscience to reality. The center consists of seven separate centers which concentrate on different areas of nanoscience. These are centers for nanopatterning, ultrafast optical sources, electron microscopy, radiation chemistry, material synthesis, proximal probes, and theory and computation. All these smaller centers perform data collection, data storage, data processing and visualization. In the case of the theory and computation center the data is created by a simulation. However even in that case a model cannot be considered realistic unless it can be compared with actual data measured by one of the other centers. A medium size cluster of computers will do the data processing and analysis. The visualization at the nanocenters will be done by special workstations equipped with large memory graphics cards to handle the large image and volume data, which are the output of the data processing. In the case of molecular imaging and electron microscopy this visualization will be in stereo and output to a special 3D theatre located in the theory and computation center. The theory and computation center will use a leadership class computation facility consisting of a 120 teraflop BlueGene computer. The output from these simulations will also be visualized on the graphics workstations and in the 3D stereo theatre.

The largest data generated in the nanocenter is expected to be from molecular dynamics and material simulations at the theory and computation facility. Depending on the size of the molecule, number of atoms of the material and the duration of the simulation these can be several gigabytes. Data sets of this size are too big too handle on a workstation as they exceed the graphics memory and are processed on a central computer, which outputs an animation. The network challenge in this case is to process the data generated from the central computational facility and play the animation remotely on the scientist's desktop. This requires a fast network connect probably optical fiber if the visualization is to be interactive and played back at 30 frames per second.

Summary Table.

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	**Network Services**
Near-term	• Center for Functional Nanomaterials	• Local Data collection, storage, and processing	• 1 Gbps	
5 years		• Multiple compute centers	• 10 Gbps	
5+ years	• Medical and Energy applications	• Remote data collection, processing and compute	• 20 Gbps	• Remote user management • Open Grid computing infrastructure

2.4 Combustion Research Facility at SNL

Background

The goal of combustion science is to gain fundamental and predictive understanding of the complex multi-scale physico-chemical processes involved in combustion energy systems. This research spans atomistic scale chemistry and molecular properties, interfacial science of catalytic reactions and multiphase chemistry, through multiple scales of continuum scale chemical and fluid dynamics. The multi-scale nature of combustion science benefits from the construction of predictive models, such as those describing the interactions of reaction chemistry and microscale mixing, that can be assembled into full predictive models of device scale combustion processes. As such, the community relies on interdisciplinary communication and integration of research results for model development and validation by experiment and direct computational simulation. The community is taking advantage of advanced laser diagnostics and imaging detectors, along with leadership scale computational resources that are producing information at unprecedented rates, and are challenging many assumptions about how the collaborative scientific process can efficiently operate. These changing paradigms often require advanced networking capabilities and new software tools to use them.

The combustion research community is diverse and geographically distributed, involving many research institutions around the world. One of the focal points for the combustion science community is Sandia's Combustion Research Facility (CRF), a special-purpose DOE BES collaborative user facility. It is distinguished by its multidisciplinary integration of research activities that range from exemplary fundamental research in combustion chemistry, reacting flows, and laser-based diagnostics, to applied research focused on high-impact combustion systems such as internal combustion engines, coal and biomass combustion, industrial burners for process heat, high-temperature materials processing and manufacturing, and related environmental and defense applications.

The CRF as a whole averages about 100 users (those engaged in a research project with one or more staff) per year and about 900 visitors (those wanting to see the labs, attending technical workshops, or wishing to discuss future collaborations) per year. The core basic research typically involves the broadest collaborative interactions. While most of these activities are experimental, research in theory and computation is an important and growing aspect of CRF collaborative research. Tera- and peta-scale direct numerical simulations (DNS) of combustion currently offer the most challenging requirements to ESnet.

In an effort to anticipate the networking requirements of the combustion science community, the relevant characteristics of examples of CRF experimental and simulation research are summarized below.

Current Network Requirements and Science Process

2.4.1.1.1 Combustion Experiments

The advent of groundbreaking advanced laser diagnostics, imaging detectors, and computer data acquisition at the Combustion Research Facility has revolutionized experimental investigation of the structure and dynamics of turbulent combustion in the past twenty years. The CRF, and in particular the reacting flow program is configured to make multi-scalar point and line measurements in addition to planar laser-induced fluorescence imaging of select flame marker species (e.g. OH and CH) to obtain the instantaneous spatial flame structure and mixing field. In addition particle-based velocity measurements are used to characterize the flow field. The resulting unique, detailed experimental benchmark data serve as the basis for evaluation and development of turbulent combustion models throughout the community. In particular, these efforts plus those of other researchers throughout the world are coordinated and leveraged through the CRF led International Workshop on Measurement and Computation of Turbulent Non-Premixed Flames. Within this construct, researchers meet to exchange data, validate and develop statistical models against the benchmark data, and plan new experiments on a biennual basis. Through constant web and email collaborations this consortium of researchers has made progress well beyond that possible by individual researchers. Current and projected data rates are of the order of several GB of raw data from a given experiment.

2.4.1.1.2 Combustion Simulations

The goal of the simulations is to advance the state of the art in the understanding and predictive modeling of reacting flows through the use of detailed computational studies. The work in this area involves investigations of fundamental turbulence-chemistry interactions using complimentary high-fidelity numerical simulation approaches, direct numerical simulation (DNS) and large-eddy simulation (LES), and the analysis and validation of chemical models in the context of low Mach number laminar flame-flow interaction.

Direct Numerical Simulation (DNS) is a first-principles description of chemically-reacting flows, i.e. a description based on continuum-mechanics statements for conservation of mass, momentum and energy. DNS provides unique high-fidelity descriptions of turbulent convective transport, molecular diffusion transport and chemical kinetics, ideally suited to studying chemistry-turbulence interactions in flames, in which all relevant physical and chemical scales are resolved, both in space and time. Because of its stringent spatial and temporal resolution requirements, DNS is a computationally intensive approach that requires massively parallel computing power and its domain of application is limited to fundamental studies in canonical configurations. With the continual and fast-paced growth in scientific computing hardware and software technologies, DNS is the natural companion of detailed experimental laboratory-scale studies of ignition and combustion at moderate Reynolds numbers. These simulations are being used to address issues ranging from reactive scalar mixing, extinction and reignition, flame propagation and structure, flame stabilization in autoignitive flows, autoignition under homogeneous charge compression ignition (HCCI) environments, and differential transport of soot in turbulent jet flames. In addition to the new understanding provided by these simulations, the resultant data are being used in apriori validation and improvement of engineering subgrid models.

The LES effort is designed to complement the DNS. The primary objectives are to establish a set of high-fidelity, three-dimensional, computational benchmarks that identically match the geometry (i.e., experimental test section and burner) and operating conditions of selected experimental target flames and to establish a scientific foundation for advanced model development. The goal is to provide direct one-to-one correspondence between measured and modeled results at conditions (high Reynolds number) unattainable using DNS by performing a series of detailed simulations that progressively incorporate the fully-coupled dynamic behavior of reacting flows with detailed chemistry and realistic levels of turbulence. The LES resolves the large-scale structures and rely on subgrid models for turbulent mixing and reaction, and therefore are complementary to the DNS efforts which focus on small-scale mixing and reaction. The focal point is the series of flames that have been studied as part of the experimental reacting flow research program at Sandia’s Combustion Research Facility.

Information from these simulations combined with detailed laser-based experiments of carefully designed benchmark flames and ignition problems present new opportunities for understanding of turbulence-chemistry interactions and for the development of predictive models for turbulent combustion in practical devices. State-of-the-art terascale DNS and LES solvers are routinely running at DOE Leadership-Class facilities at ORNL and NERSC.

Scientific Process

The scientific process for peta-scale simulation involves several stages categorized as:

Production run preparation and software readiness

- Determine and implement optimum programming model for multi-core processors through performance monitoring and use of OpenMP and MPI.
- implement and test collective I/O and standardized I/O formats (pnetcdf or hdf5) scalable to peta and exascale architectures.
- perform preparatory runs – i.e. coarse mesh runs to determine spatial resolution requirements and refine selection of numerical and physical parameters for production runs.

Perform production run

- submit 2-3 million cpu-hr job on Crays at ORNL on the 119 Tflop machine that will run for 7-10 days. This may happen 3-4 times in a year (involves ~5 people).
- write restart files out each hour (1GByte). Over the course of the run an aggregate of about 20 TB data is written to scratch.
- monitor the health of the run daily by using sftp to send diagnostics (x-y plots, isocontour plots back to SNL Livermore, CA). Develop automated workflow and Dashboard to monitor data and manage data movement.

Postprocessing

- morph data to N processor-domains for analysis and viz.
- archive restart files and morphed data to HPSS at ORNL or NERSC. The data needs to be archived for 5-10 years since it will be revisited multiple times by the modeling community.
- move morphed data from scratch disk to analysis machine (Beowulf cluster or SGI Origin at ORNL).
- move morphed data from scratch disk to cluster at SNL, Livermore for analysis and parallel volume rendering. (protocols used include parallel streams using bbcp or rsync with 4 streams.)
- perform parallel analysis and viz. on data. This is a highly iterative process with portions of the analysis requiring new analysis and other portions relying on existing analysis approaches. The first stage of analysis (typically up to 1 year after data is generated) is performed by the data creators. Subsequent analysis of the data is by collaborators at universities and labs in the U.S. and abroad. Current collaborators are at University of California at Davis, Iowa State University, Lawrence Livermore National Labs, University of Illinois, University of Utah, and Cambridge University.

Network Requirements and Science Process – the next 5 years

Production runs, in-situ analysis and visualization, and data sharing

- By 2008, DNS simulations on 250TF production runs will produce 100 Tbytes of data per year at NCCS/ORNL, and by 2010/2011 20 PF runs will produce 10 TB of data per checkpoint file (i.e. 90 G grid * 14 variables *8 bytes/variable) = 10 TB. If 200 checkpoint files are written out at approximately 1 file per hour in a 7-10 day run, the data generation rate ~ 20 Gbits/s. If we move data elsewhere at the rate it is generated then we need a network to move data at ~20 Gbits/s. The network will need an even higher bandwidth to account for overheads due to protocol, metadata, contention, etc.
- The goal is to stream data to an analysis and rendering machine as it is produced rather than waiting until the run is complete. In this manner, known analysis tools can be applied to the data as it is generated to get a first glimpse understanding of the underlying physics, and subsequent iterative analyis can be performed off-line. Automated workflow scripts using Kepler will facilitate the data streaming, morphing, archival, and analysis. Data will need to be moved to platforms and archival storage within a supercomputing center as well as to SNL Livermore, CA to two open network clusters for analysis and rendering.
- DNS simulations will be instrumented with in-situ feature detection, segmentation and tracking to enable data reduction and querying on-the-fly, thereby reducing the amount of data for further analysis and enabling steering of adaptive I/O.
- Web-based portal developed for sharing simulated benchmark data with modeling community of ~50-100 international collaborators at universities, national labs, and industry. A scalable, extensible framework will be developed for analyzing large data and comparing data with experiments. The framework will include capability for standardized formats, translators, graphics, parallel library of combustion analysis software, parallel feature detection/tracking library, inference software for automatic model generation, and query tools that can operate on portions of the data at the supercomputing facilities where the data resides. Reduced data and remote viz. will be sent back to institutions via ESnet.

Beyond 5 years – future needs and scientific direction

We envisage as machines progress towards exa-scale, it will be possible in combustion to consider hybrid multi-scale simulations of turbulent combustion in the gas phase interacting with catalysis at a surface in a fully coupled simulation. Similarly, it will become possible to consider multi-phase turbulent combustion (dense and dilute sprays) through atomistic simulations of the condensed phase and to handle their coupling with the continuum gas-phase turbulent combustion. A third example would be to consider the deflagration to detonation transition in simulations that handle both flame propagation as well as embedded shocks – again, a hybrid simulation involving solution of a large system of PDE’s using method of lines together with Monte Carlo or Molecular Dynamics approach for treating internal shock structure. Therefore, the analysis and visualization software framework will need to handle heterogeneous data types to an even greater extent than before. We also anticipate there will be more emphasis placed on in-situ segmentation, tracking and query-driven analysis and model inference to glean insight from such large data. Hence, there may be a shift in paradigm from moving large amounts of data over the network to greater emphasis on remote intelligent and versatile data reduction and visualization.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	**Network Services**
Near-term	• 250 Tflop DNS simulations running for 1 week on Cray at ORNL, 'Science Day 1' in 2008. • INCITE simulations run for 1-2 weeks at a time, 3-4 heroic runs per year at ORNL.	• Data generated on terascale machine at ORNL or NERSC • Some of analysis is performed at ORNL, other analysis and data reduction is performed at SNL or NERSC • Data sharing with geographically distributed collaborators	• key sites: ANL, NERSC, PNL, ORNL, LLNL, UC Davis • 5-10 Gbps	• data management and workflow management tools
3-5 years	• 20 Petaflop DNS simulations running for 10-12 days per run	• 50-100 collaborators from around the world	• 20 Gps • connections to universities, labs) both in the U.S. and abroad (U. K, Germany, France, Japan, S. Korea).	• portal-based tools for data sharing
5+ years	• 100 Petaflop DNS simulations performed with embedded multi-scale models	• Extensive community-wide collaboration and data sharing.	• real-time remote analysis and rendering • 100 Gbps	• collaborative visualization support

2.5 Computational Chemistry

Background

Computational chemistry is used to address scientific questions in many areas, including catalysis, solar energy, hydrogen economy, environmental remediation, and biology. To enable researchers to make major advances in areas of interest to DOE it is crucial to incorporate the kinetics and dynamics of chemical processes that drive many (radiation or photo induced) chemical reactions and transformations. Obtaining accurate kinetic and thermodynamic results requires the capability to model an accurate potential energy landscape, to determine and discover reaction pathways, and to extract statistical information through sufficient sampling of the configuration space. Molecular dynamics simulations on large biological systems are similarly driven by the need for extensive sampling of the configuration space.

Current Network Requirements and Science Process

Statistical sampling of large numbers of configurations (hundreds of thousands or even tens of millions) at multiple temperature regimes can be distributed to many groups of processors, which can be distributed geographically (i.e. grid based computing), with a minimum amount of communication overhead, and are particularly well suited for petascale computing resources. Large quantities of data will be generated that need to be analyzed, visualized, and mined to discover reaction pathways.

Molecular dynamics simulations on biological systems of about 50,000 atoms in the time range up to tens to hundreds of nanoseconds can be routinely carried out. These generate data sets in the range of 1-5 TB in size. However, in order to study biologically relevant phenomena, and to obtain information regarding the dynamics during transitions between the different conformational states and the reactions that they induce, it is necessary to simulate for significantly longer simulation times. In addition, biological systems of interest consist of millions of atoms.

Network Requirements and Science Process – the next 5 years

Long time scale molecular dynamics simulations of hundreds of nanoseconds combined with large biomolecular systems of possibly a million atoms will produce trajectory data that needs to be stored, retrieved, and subsequently analyzed or visualized. For example, a 100 nanosecond simulation of a biomolecular system of 10 million atoms could potentially generate 4 petabytes of data. It is anticipated that most simulations will be much smaller, and that the number of these large scale runs is limited and will require possibly weeks of computer time. With such huge quantities of data, it is clear that tools and methodologies need to be developed to analyze data “on the fly”. With on-the-fly analysis, data sets might be reduced to 10-100TB in size before they are moved across the wide area network. Long term data storage will be important, especially for large “hero” style simulations of which the results could be used for analysis by various researchers. At PNNL, the EMSL supercomputer center (MSCF) provides its users with archive storage on disk for relatively quick and easy access. To avoid single point of failure, backup of essential data at other locations will be crucial.

With the emergence of computing platforms that are a hundred teraflops, and with petascale computing platforms on the horizon, computational chemistry is on the verge of entering a new era of modeling. These huge computing resources will enable researchers to tackle scientific problems that are larger and more realistic than ever before, to include more of the complex dynamical behavior of nature, and to start to answer new and different scientific questions. To take full advantage of petascale architectures, and achieve high computational performance on such computer hardware, it will be necessary to significantly improve the efficiency of current high performance scientific software. The effort will have to include scientists from various disciplines, including computational chemists, mathematicians, and computer scientists. Expertise needs be drawn from groups in various disciplines that are working on the development of new advanced numerical methods and scalable computational chemistry algorithms, multilevel parallelism, and are working on the development of new ways to effectively utilize large numbers of processors, network topologies, and hierarchical memory structures. A collaborative infrastructure is required to enable interactive collaborative development and shared access. To counter cyber attacks and terrorism, many computing facilities have resorted to the use of security measures one-time passwords. Standardization of secure access to computing resources will be necessary to ensure an effective collaborative environment.

Beyond 5 years – future needs and scientific direction

By now large petascale systems will be generally available for computational chemistry research. Simulation sizes will increase and become more expensive, as higher levels of theory are applied. These large computational resources will be capable of generating multiple petabytes of data that need to be stored, archived, analyzed (locally by moving the data from the computing center, remotely through remote visualization, or on-the-fly), and made available to other computational chemistry researchers. Sharing data will be pertinent to enable effective use of the results from the large-scale simulations that are too expensive to reproduce. It is also expected that secure access to computing resources be standardized across the DOE campus.

Summary Table

Time Frame	**Science Instruments and Facilities**	**Process of Science**	**Anticipated Requirements**	
			Network	**Network Services**
Near-term	• Large scale kinetics and dynamics simulations that generate huge data sets.	• Monte Carlo simulation • Data sharing and retrieval • 1-5 TB/week of data	• 1+ Gbps	• Reliable data transfer services
5 years		• Remote data mining and visualization • “On the fly” data analysis • 1-4 PB/week of data • 10-100TB data sets moved over network	• 10 Gbps	
5+ years		• Distributed computing at various supercomputing sites with tightly coupled simulations • 10PB/week of data	• 20+ Gbps	• Standardized secure access to computing resources

2.6 *Linac Coherent Light Source at SLAC*

Executive Summary

The LCLS facility is under construction at the Stanford Linear Accelerator Center (SLAC). LCLS will produce coherent X-ray pulses with an intensity 10^9 times greater than synchrotron radiation facilities. The pulses will be capable of femtosecond time resolution. The LCLS experimental program will start in 2009. A typical experimental setup at the LCLS will initially involve a megapixel X-ray detector, read out once per pulse. As the LCLS approaches its design performance, experimental data rates will approach hundreds of megabytes per second. Expected advances in detector technology will increase this rate by several powers of two during the first 10 years of operation. Network data rates may be similar to that at which data will be written to mass storage – around 0.5 Tbytes/day initially, rising to over 6 Tbytes/day in 2012 and over 30 Tbytes/day by 2015.

LCLS Network Requirements

The LCLS will be a flagship resource for the DOE Office of Science program. LCLS construction builds on SLAC's more than four decades of accelerator research, construction and operation and will re-use a major part of SLAC's original linear accelerator. LCLS also builds on SLAC's photon science program that began two decades ago with the first synchrotron radiation facility.

LCLS breaks new ground by using a high-energy "wiggled" electron beam to create coherent synchrotron radiation pulses. The instantaneous intensity will be 10^9 times greater than SLAC's synchrotron radiation facility, and pulse compression techniques will allow the LCLS to achieve femtosecond-scale time resolution. The LCLS will open up new frontiers in science, particularly those of Ultrafast Science, exploiting the short LCLS pulses to study the time dimension at the femtosecond scale, and of single-molecule Coherent X-ray Imaging, allowing molecular structure to be studied without the changes and limitations introduced by crystallization.

The raw data rates that will be produced by LCLS experiments will be determined by the machine performance and by detector performance. LCLS is designed to produce 120 Hz of pulses and, eventually, to run with high up time and high efficiency. The first round of detectors now under construction are typically centered on megapixel X-ray detectors. At 120 Hz, such detectors will produce up to 250 megabytes/s. Much of LCLS science is eagerly awaiting the development of higher resolution detectors. These will push raw data rates up towards 1 gigabyte/s by the middle of the next decade. Much of this raw data will not be amenable to online data reduction and will be recorded and distributed for later analysis.

LCLS is spurring the creation of a new community of scientists. At present, this community is just beginning to recognize its own existence. The community is still far from having a well-developed view about how LCLS science will be done. The traditional approach to exploiting synchrotron-radiation facilities has been to turn up with samples to be exposed, take data for a few days, and then take away a DLT or DVD containing the relatively modest amount of data collected. Much of LCLS science will be data-intensive and compute intensive. For example, the coherent imaging of single molecules will require that each sample of tens of terabytes is subjected to intense computational analysis. This analysis is itself, a research topic, and it is very likely that the same date will be both analyzed on site at SLAC and transmitted to other computer centers where other analysis approaches will be tried.

Given the novelty of the science, there is no detailed network use model. A good working assumption, that reflects experience in experimental high energy physics at SLAC, is that the data volume transmitted over the network will not be less than that written to mass storage at the experimental site. The requirements detailed below are a first guess, made on this basis.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Networking	Middleware
Near term	(Facility comes on-line in 2009)			
3-5 years	• The 4 initial instruments at LCLS/LUSI will generate an aggregate of perhaps 150 Gbytes/day (20 Tbytes/day if all systems work perfectly) • Some data analysis will be accomplished on computing systems that are remote from the LCLS	• Some real-time data reduction followed by straightforward analysis • recording of full data stream for offline analysis. • near real-time analysis and visualization to ensure data quality.	• 60 Mbits/sec sustained • 2 Gbits/sec peak	• Workflow management • Reliable data transfer
2012	• Instrument pixel counts double. • Massive local data analysis facilities progressively coupled with remote supercomputer centers	• Add organized data distribution and remote visualization	• 750 Mbits/sec sustained • 4 Gbits/s peak	• Security (authentication and access control) to permit direct interaction with the instrument remotely.
2015	• Instrument pixel counts double again • Integrated local/remote computing environment	• Iterative analysis of the data with the use of models running on supercomputing systems.	• 4 Gbits/sec sustained • 8 Gbits/s peak	

2.7 *Molecular Foundry at LBNL*

Background

The Molecular Foundry at Lawrence Berkeley National Laboratory (LBNL) is a user facility charged with providing support to research in Nanoscience underway in academic, government and industrial laboratories around the world. The Foundry provides users with instruments, techniques and collaborators to enhance their studies of the synthesis, characterization and theory of nanoscale materials. Its focus is the multidisciplinary development and understanding of both “soft” (biological and polymeric) and "hard" (inorganic and micro-fabricated) nanostructured building blocks and the integration of those building blocks into complex functional assemblies.

The Molecular Foundry is one of five DOE Nanoscale Science Research Centers devoted to enabling scientists from a wide range of disciplines to engage in multidisciplinary nanoscale research. User proposals consist of multidisciplinary projects in nanoscience submitted by scientists throughout the nation (and internationally) seeking to broaden the boundaries of their research with the help of Foundry scientists and facilities. Proposals are reviewed on a rolling basis throughout the year and research begins immediately on approval; the review process is highly selective and only those with the highest level of scientific merit are accepted.

Current Network Requirements and Science Process

The Molecular Foundry consists of six different scientific sub-disciplines, or facilities, each with different core capabilities to assist users. The two dominant data-intensive programs within the Molecular Foundry are currently the Theory of Nanostructured Materials Facility, and the Imaging and Characterization Facility. As many of our users are off-site, network bandwidth is a critical determining factor with respect to our capabilities. Currently Molecular Foundry users routinely generate 10's of GB of data with either experimental imaging equipment, or theoretical simulations. Due to network bandwidth limitations or lack of efficient parallel transfer tools installed at the endpoints, these datasets are usually analyzed piecemeal, and if necessary, large files are transferred via DVD or portable hard disk.

Network Requirements and Science Process – the next 5 years

Over the next 5 years, simulation data set sizes are expected to increase to the size of terabytes, primarily due to expanded computational resources that will permit the simulation of larger nanostructured systems and phenomena on longer time scales. In addition, enhanced demand for remote access, visualization and control, and analysis of simulation and experimental data will result in greater network requirements. The creation of data archives will also increase the load on the network as the data sets are collected and then analyzed. Upon completion of their projects, data transfer for off-site storage will be extremely time consuming for users. Remote simulation and visualization will also require larger bandwidth, even with use of better compression algorithms.

Beyond 5 years – Future needs and scientific direction

Providing the capability of remote operation of experimental equipment, real-time visualization of simulations, and analysis of terabyte datasets, will open up new routes to scientific discovery, but with greater data rates and storage criteria than are currently possible, roughly 10 gigabits per second and beyond.

We envision several ways in which a significant expansion of network bandwidth and latency can strongly impact scientific research and discovery at the Foundry, for both experiment and theory. Examples of these are:

- Remote control of simulation. The ability to visualize and adjust simulations and experiments in real time from remote locations will open up new doors to scientific discovery. High-resolution on-the-fly visualization of the results of nanoscale molecular dynamics simulation, for example, isosurface plots on a large real-space grid, are planned in which remote users can view and readjust aspects of the simulation. At low-resolution or with some compression, this is currently possible over LAN speeds of 1 Gb/s, and would require similar guaranteed bandwidth over the WAN for multiple simultaneous users across ESnet.

- Advanced visualization and analysis. Enhanced bandwidth can drive scientific discovery by enabling rapid, high-throughput analysis of large datasets. For example, analysis of terabyte-scale datasets currently generated by nanoscale electronic structure and molecular dynamics simulations requires a large-memory graphics workstation. At present, reading such a dataset directly into memory over Gigabit ethernet takes ~3 hours. A single iteration on the dataset would cause another 3-hour delay while using the entire network during transfer for a single user. The ~3 hour iteration would drop to 15 minutes by moving to 10 Gigabit ethernet. Future datasets are expected to be significantly larger – from 10's and 100's of TB, to petabytes – network demands for analysis of a fraction of this information, i.e. TB datasets, will increase accordingly.
- Remote control of imaging experiments. Beyond five years, the ability for users to control remotely imaging and fabrication experiments is an ultimate goal at the Molecular Foundry. Viewing high-resolution digital images at frame rate in real time from a remote location will require significant bandwidth.
- Remote analysis of imaging experiments. After finishing experiments returning to their home institutions, some users many need remote access to commercial analysis tools available at the Foundry to finish their projects. Additional bandwidth may be necessary to support real-time analysis of large datasets.
- File transfer rates. The Theory Facility has a 296 processor compute cluster that currently generates large data files that can be hundreds of GB or TB. In the future, quick and efficient transfer of these files (on the order of minutes instead of hours) will significantly enhance user collaborations.
- Grid-based computing. Looking into the future, sharing cluster resources with other labs over a grid would have many benefits for large-scale nanoscale simulation. A more robust ESnet would need to provide high-speed, low-latency interconnects required for efficient parallel processing to enable computationally intensive parallel jobs.
- Advanced collaboration tools. Finally, high-resolution video conferencing will be necessary for collaborative projects in which large data sets are shared with users. A network with bandwidth and jitter guarantees will be required to maximize the utility of these tools.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	Network Services
Near-term	• The Molecular Foundry is a national user facility, and is thus a highly collaborative experiment and analysis environment	• Typical imaging experiments generates GB of data • Simulations can generate 1 TB datasets • Simulations are primarily carried out by remote users, who must transfer 100 GB datasets back and forth to analyze results	• 1-2 Gbps for data transfer	
5 years	• Increase in computational power and file storage • Higher-resolution detectors • Multichannel TB datasets • 10 TB or beyond simulation datasets	• Remote data analysis by off-site users • multi-terabyte datasets • Remote operation of high-resolution imaging experiments and simulations • Ten's of remote users may be active at any given time	• 1-10's of Gbits/sec for remote visualization, data mining, and analysis • 1-10 Gbyte movies generated by each instrument every sec for remote viewing across United States and internationally	• Remote I/O • Secure remote interface • Collaborative use of common, shared data sets – version control on the fly • International interoperability for collaboratory infrastructure, repositories, search, and notification
5+ years	• Factor of 10 increase in resolution of experiments and computational power of simulations • Petabyte archives of simulation data	• Streaming visualization of 5-10 terabyte data sets • Real-time remote operation of experiments and simulation	• Beyond 10Gbit/sec • Quality of service for network latency and reliability	• Remote collaborative simulation steering, mining, visualization, and analysis

2.8 National Center for Electron Microscopy at LBNL

Background

The National Center for Electron Microscopy (NCEM) at Lawrence Berkeley National Laboratory (LBNL) is one of the world's foremost centers for electron microscopy, and micro- and nanoscale characterization. NCEM features several unique instruments, complemented by strong expertise in image simulation and analysis. NCEM also maintains one-of-a-kind instruments for imaging of magnetic materials, and develops techniques and instrumentation for dynamic in-situ experimentation.

The newest and most exciting programs at NCEM is the TEAM (Transmission Electron Aberration-corrected Microscope) project. The TEAM project brings together 5 leading microscopy groups supported by the US Department of Energy's Office of Science to jointly design and construct a new generation microscope with extraordinary capabilities. These leading groups include those at Argonne, Oak Ridge, Brookhaven, and the University of Illinois. The TEAM project will construct a new generation electron microscope designed to incorporate aberration-correcting electron optics, to develop a common platform for a powerful new nanocharacterization instrument, and to make this instrument widely available to the materials and nanoscience community. With the TEAM microscope it will become possible to study how atoms combine to form materials, how materials grow and how they respond to a variety of external factors. This work will improve designs for everything from better, lighter, more efficient automobiles, to stronger buildings and new ways of harvesting energy. The TEAM project is part of DOE's 20-year roadmap of Facilities for the Future of Science. With the addition of several new high-resolution microscopes as part of this project, network needs are expected to grow accordingly.

Current Network Requirements and Science Process

NCEM is continually augmenting its capabilities, which currently include several electron microscopes with 4k x 4k images, generating 0.5 GB per day. Due to network bandwidth limitations, these datasets are usually analyzed piecemeal, and relevant files are transferred via DVD or over the network.

Network Requirements and Science Process – the next 5 years

The ability to control experiments remotely will greatly enable scientific discovery. Halting an experiment and changing course in real time, steering the measurement toward the most interesting phenomena, would allow more efficient pathways to future breakthroughs. Further, the ability to analyze data in remotely will allow NCEM to fulfill its mission and better serve users. Offsite users could have high-bandwidth access to their data and any commercial analysis tools long after their visit. However, it is clear that viewing high-resolution digital images in real time from a remote location will require significant bandwidth. For example, within two years typical high-resolution datasets will have the size 4096x4096 pixels, each 16 bit, yielding 30 MB per image. Viewing these images at 30 frames/second would require the transfer of 1 GB of information per second. Currently, the camera described above does not exist, but detector development is underway, and an increase in data transfer needs would need to be addressed in the next 5 years. Compression will also reduce the load on the network, and compression algorithms will also be explored to relieve network strain. In the near future, multichannel datasets will become possible, as diffraction patterns or other additional data are accumulated at each pixel, greatly increasing the dataset size to 10's of GB and beyond.

Beyond 5 years – future needs and scientific direction

Beyond five years, remote operation of the next generation of higher-resolution detectors now on the drawing board, with resolution on the order of 20k x 20k pixels, would require even more bandwidth; up to 25 GBytes/sec . Finally, we expect a growing need for real-time interactions among NCEM users and staff, as well as the interactive visualization and processing of very large datasets, drive additional network requirements beyond bandwidth, including low latency for voice and video over the network.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	Network Services
Near-term	• NCEM is a national user facility, and is thus a highly collaborative experiment and analysis environment • 8 TEM's by 2008	• Typical TEM experiments currently generate 500MB/day, though multichannel data sets will exceed several GB's • Each TEM experiment only gets a few days per month - high productivity is critical	• 1-5 Gbps	
5 years	• Up to 1 additional instrument every 2/yrs • Higher resolution detectors	• Remote operation of high-resolution TEM experiments	• 5-10 Gbits/sec for remote experimentation, visualization, and analysis	• Secure remote interface
5+ years	• Factor of 10 increase in resolution of TEM experiments will scale network needs accordingly	• Real-time remote operation of experiments • Comprehensive integrated simulation	• 10's of Gbit/s	• Quality of service for network latency and reliability

2.9 National Synchrotron Light Source at BNL

Background

NSLS is a synchrotron light source that consists of two storage rings (UV and X-Ray), providing intense and tunable X-Rays, UV and Infrared (IR) light sources for about 2300 experimenters per year. NSLS currently has a total of 60 active IR/UV/X-Ray beamlines or end stations, serving wide range of experimental techniques, from diffraction, protein crystallography, imaging, to all ranges of spectroscopy. NSLS is a very productive and reliable facility, operating in a 24x7 mode for average 44 weeks per year. About 70% of users come from academia, others from national labs, industry or international. Data generated at the facility depend on the Internet, together with portable media, to be transferred back to user's home institutions for further and detailed analysis. Over the years, with enhancements in beam intensity and instrumentation, and especially with wide adaptation of modern area detectors by many beamlines, we have been witnessing a steady increase in generated data. Currently we are generating about 1 TB/day, 250-300 TB/year.

A new state of the art third generation synchrotron light source, NSLS-II, is planned to be commissioned and operational in 2014. About 20 beamlines at NSLS will be continuously improved, and will be moved to NSLS-II at the onset of the operation of NSLS-II. Thus the data rate at BNL light sources will constantly increase, due to improvements in instrumentation, optics, and detector upgrades. In 2014 and beyond, network demands will increase significantly as NSLS-II is set to operate.

Current Network Requirements and Science Process

The basic network needs for the NSLS users is to transfer raw and preliminarily analyzed data from instruments at NSLS to their home institution for detailed analysis. Typically, users do preliminary analysis at the beamline's workstations, or transfer data to their home institutions for processing, to guide their experiments.

The NSLS network requirements are based on data intensive instruments and techniques, which mostly involve modern detectors. They are typically associated with science programs such as Macromolecular/Protein crystallography (PX), Quick EXAFS (Qexafs), imaging, spectroscopy using multi-element detectors. At the moment, PX programs generate a majority of data volume.

The NSLS currently has 10 PX beamlines, including 8 bending magnet (BM) beamline employing 4kx4k CCD, or 2kx2k CCD, and the two undulator beamlines using ADSC Q315 (6kx6k CCD). Depending on sample being studied and detector resolution users choose to use, a typical BM beamline has data rate estimated to be in the range of ~20-50GB/day, 8 beamlines will generate about 160-400 GB/day. For the two undulator beamlines with Q315 CCD detectors, readout time is ~1sec for a full 6kx6k image, data rate is estimated to be 260 GB/day/beamline, for 520 GB/day. Subtotal for PX programs in NSLS is about 700-900GB/day.

Other Programs: (a) We developed a quick EXAFS program, where the photon energy is scanned continuously; fluorescence signals are converted digitally through fast ADCs. In the early stage, the system can typically generate 6MB in 30s, or routinely 3-6GB/day. With improvement in the technique, data rates could increase to 10-20GB/day. (b) We have two CCD detectors in the detector pool to be used upon users request (the usage rate is very high). The data rate depends on experiments, but can reach about 10-30GB/day/detector. (c) We have three 13-element spectroscopy detectors, recording full spectrum for each data point. Data rate is about 2-3 GB/day/detector. (d) Microfocusing beamline using CCD detector: 50-80GB/day. Contribution from non-PX programs is estimated to be ~100-150 GB/day.

With 60 beamlines, we estimate that the tesotal data rate for NSLS is about 1TB/day, or 250-300 TB/year. We generate almost as much data as RHIC at BNL.

Network Requirements and Science Process – the next 5 years

- A new undulator beamline for SAXS (small angle x-ray scattering) employing CCD detector will be in operation starting 2008. This could add ~20-50 GB/day of data.
- Multi-element Si detector array (400 element) for microprobing with microbeam scanning applications, data rates can reach 1MB/s, or 80GB/day.
- PX program detector upgrades, with faster readouts also would push data rate higher. At the moment, many of the CCD detectors in use have readout time in 10sec range for full 4kx4k image, many users/stations choose to use the 2x2 bin readout (1sec) for 8MB rather than 32MB image. Newer CCD detectors mostly have full resolution readout within 1sec.

- In about 3-4 years, Million-Pixel fast readout 2D area detector which is being developed at BNL for LCLS, is expected to be available for use. With 1*ms* readout (improved version for NSLS), data rates could reach 2GB/s.

Beyond 5 years – future needs and scientific direction

NSLS-II will be coming up for operation in 2014. About 20 beamlines will be transferred to NSLS-II from NSLS for immediate operation during the commissioning of NSLS-II. The brighter beam will shorten the exposure times for existing experiments, new beamlines will be built, new techniques will be explored, generating much higher volumes of data of up to 20 TB per day.

Summary Table

Time Frame	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
			Network	**Network Services**
Near-term	• ~60 active NSLS Synchrotron Radiation beamlines operating	• continuous operation (24x7, 44 weeks/year) • data at rate of about 1TB/day, 300TB/year. • Remote operation and collaboration between on-site staff and outside experimenters	• Bulk transfer, For the whole facility, average 1 TB/day • 1 Gbps network	• remote control of experiments and visualization. • Reliable data transfer services.
5 years	• New and updated beamline instruments and optics	• Million Pixel 2D detectors push data rate to 2GB/s	• 3TB/day data transfer • QoS to support remote steering, remote visualization • 5-10 Gbps network	• Remote collaboration with on-site staff and experimenters • Remote steering and visualization • Reliable data transfer
5+ years	• Improved detector readout and new type detectors significantly increase data rate • NSLS-II coming into operation in 2014,	• data volume reaching 5TB/day • Increased reliance on remote visualization and steering	• 10-20TB/day data transfer • 30-50 Gbps network	• Live video/audio for environmental monitoring of experiments • QoS capabilities for remote collaborations

2.10 Spallation Neutron Source at ORNL

Executive Summary

The Neutron Scattering Science program for DOE is currently located at four facilities: 1. the Spallation Neutron Source at Oak Ridge National Laboratory, 2. The High Flux Isotope Reactor at Oak Ridge National Laboratory, 3. The Lujan Center at Los Alamos National Laboratory, and 4. The Intense Pulsed Neutron Source at Argonne National Laboratory. The facilities fall into one of two categories: accelerator based pulsed sources in which neutron time of flight can be measured, and reactor based facilities in which neutrons are produced continuously with a wide spectrum of energies. Each facility will have multiple instruments each designed to exploit a particular measurement strategy for sculpting the neutron beam in the primary flight path and then a secondary flight path in which neutrons are scattered from a sample producing patterns (such as Bragg diffraction patterns) and then detected by highly specialized neutron detectors. Measurements can range in time from a second to a day or more depending on what is being measured. Data can be collected in either histogram mode or event mode. Histograms can range in size from a few Megabytes up to 10GB. Event files can vary widely in size and are directly proportional to the amount of neutron flux reaching the detector and the amount of measurement time.

These are user facilities in which users compete for neutron beam time via a peer reviewed proposal system. Users awarded time prepare and bring samples to the facility and place their samples in the beam sometimes within special sample environment chambers, which can control variables such as temperature, pressure, or magnetic field. Experiment teams are typically comprised of a small number of people (less than 10) and an experiment team may typically be awarded 3 days to perform their experiment. It is expected that SNS users will typically produce many GB of data during their stay. The raw data are collected by the Data Acquisition System (DAS), which has a local capacity of 1-2TB of storage space. Once a measurement is complete, these data are cataloged and stored in a centralized repository at SNS. In close physical proximity to the DAS is an instrument analysis computer, which reduces the data from instrument, coordinates to scientific coordinates thus transforming data from pixel space to “Q” space. When a facility is operational, data are collected 24 hours a day.

Once data are located within the centralized repository at SNS, they are available for further analysis. Researchers try to “fit” their data to atomic models in order to understand structure and atomic dynamics. Fitting is a computationally intensive operation which is well suited to cluster or grid computing. Users may also take the step to employ instrument and materials simulation Monte Carlo tools to produce simulated results. These Monte Carlo tools are also well suited to cluster or grid computing. The science areas supported by neutron scattering are wide ranging including: chemistry, complex fluids, crystalline materials, disordered materials, engineering materials (stress/strain), magnetism, polymers, and structural biology.

Neutron Scattering Science Future

The typical experiment model for these user facilities has been to assist the user to produce data, but once their data were produced and reduced, the user was on their own to extract the science from these data. This model has tended to produce “islands” of communities each adopting their own data formats and analysis software. This model has not been ideal for collaboration and makes scientific discovery an arduous task requiring each user to have some level of computer science savvy in order to work with their data – in addition to having expertise within their own science area.

With the building of SNS, there has been the opportunity to update this user-operating model. A centralized data repository has been produced and raw data are stored in an HDF5 based format called NeXus, defined by the neutron scattering community at large. Data are “owned” by the experiment team and user IDs are provided via the ORNL XCAMS username clear-text password mechanism. As HFIR is in close proximity to

SNS, providing a centralized data repository with authentication/authorization has been straightforward.

Other US neutron scattering facilities have not provided a user authentication/ authorization system for managing data, so the ORNL XCAMS system is available and extensible to these facilities as well. At this point, it has been demonstrated that data from the Lujan Center can be stored in the centralized repository maintained by SNS. Users are like the idea of having all of their data located in one easy to locate place. As a result, the other US facilities also wish to utilize the “data portal” concept to enable their users to find and work with data. In addition to the DOE facilities, there are two other US facilities which may partner with this DOE lead effort: 1. the NIST Center for Neutron Research (NCNR) funded by DOC, and the Low Energy Neutron Scattering (LENS) facility funded by NSF. Currently NCNR operates the largest neutron scattering user program in the US. Eventually there will be a desire to interact and share data with neutron scattering facilities internationally. To do so will require a more robust user authentication system.

The long term goal of neutron scattering science is to more closely couple computation with experimentation. Users wish to come to a facility to not only “take measurements”, but to “perform an experiment”. Users don’t want to go home following their facility time with just data, but a first look at the science their data holds. This will require a more fully integrated system that brings high performance computing directly to the experiment. To be useful, specialized computing tools must be made available to these users to automatically provide functionality that they must now do themselves. These tools should help with experiment steering and control, but also with the process of fitting data to structure. Tools including a variety of modeling and simulation software must be encapsulated for these users. In addition, it will be necessary to characterize and utilize the instrument resolution response functions in order to properly utilize models. On top of all this is the ability to visualize and interact with data in order to explore it. The reality for user facilities is that there is a throughput limitation on the number of users it can support, thus to accelerate scientific discovery it is essential to facilitate helping users extract science from their data.

Network Interconnectivity

Near term network connectivity supporting data management for neutron scattering would resemble the spoke and hub model with facilities flowing data to SNS. This includes the following facilities producing data: SNS and HFIR at ORNL, Lujan Center at Los Alamos, IPNS at Argonne, NIST Center for Neutron Research (NCNR) in Gaithersburg, MD, and the Low Energy Neutron Scattering (LENS) facility in Bloomington IN. Eventually, perhaps within the next 5 years, interoperability with international sites such as ISIS in the UK, ANSTO in Australia, and J-PARC in Japan are envisioned.

As importance grows to jointly use neutron scattering and X-Ray data (in order to perform co-refinement), it will be important to be to quickly move data between SNS and the DOE light source facilities. Initially the APS at Argonne will be the most likely facility for this interaction. To solidify this need Argonne National Laboratory is currently working on a proposal for the Argonne “Scattering and Imaging Institute” to be established at Argonne with the charter to support scientific research involving both SNS and APS.

As the Center for Nanophase Material Science (CNMS) is connected to the SNS, it is anticipated that individual users will use both facilities. To support this, in the near future CNMS instruments will be able to place data in the data repository at SNS . As CNMS users are likely to use the other nano-centers, good connectivity between these nano-centers and SNS will also required.

Users accessing data maintained in the repository maintained by SNS are envisioned to be from virtually anywhere ranging from the facilities listed above, to numerous academic institutions, user's home internet connections, and public access points such as hotels and coffee shops. The following are institutes where the bandwidth requirements are envisioned to be highest: University of Tennessee, Caltech, Michigan State University, Iowa St. University, University of Maryland, University of Binghamton NY, and Middle Tennessee State University. This list of institutions is sure to grow.

Coupling with computing facilities is another source of traffic. Currently SNS is interacting with the Neutron Science TeraGrid (NSTG) Gateway co-located at ORNL. NSTG is currently a data source for SNS as simulated experiment data are currently produced on the TeraGrid and moved back to the SNS data repository. In the future, there may be interest to flow streaming data from SNS for processing on the TeraGrid. In addition, universities in close contact with SNS may also desire to interconnect with the TeraGrid having a high bandwidth interconnect to support data movement between TeraGrid and their institution. Interoperability with the Open Science Grid may also be desired.

Future connections with computing centers includes closer coupling with the DOE supercomputing facilities located at ORNL, ANL, and LBL. Working with these high-end computers will enable petascale computing to be coupled with neutron scattering science. Bi-directional data flow will be required – flowing experiment data to the computing facilities, and flowing computational results and simulated data back to the SNS data repository. There will be a desire to schedule computer time in order to provide results during experiments – ideally while researchers are sitting at their instruments. Such functionality will require network connectivity reserved or dedicated in order to facilitate timely data flow.

Volume of Users

Typical experiments last on the order of 3 days and experiment teams may typically be comprised of 1 to 5 experimenters (nominally 3). The SNS operating schedule includes 200 days of operation. Currently there are 3 instruments in operation with 5 to 7 more to come on-line before the end of calendar year 2008. A full complement of instruments for target station #1 would be 23 instruments. Thus in round numbers, SNS can expect a few thousand facility users per year when fully functional. NCNR which currently has the most active user program in the US, has between one and two thousand users a year. IPNS and Lujan each typically have less than 500 users per year. Note that these numbers are just rough estimates. Also note that not all facility users will be computing users.

Data Flow

Instruments produce data that are initially stored locally at the instrument. At SNS and HFIR, a messaging system is in place that is triggered upon completion of a “run” and this initiates data flow from the instrument. (A run is comprised of the user pushing the acquisition start button, collecting data, then pushing the stop button.) The raw data are “packaged” into NeXus files that utilize HDF5. The data management system is then notified of the new data and pulls this data and catalogs the data from the run. Experiment data are kept locally within SNS. In the near future, data will also be archived in the ORNL HPSS tape archive system. Data archives will be integrated with the data catalog so that users can locate their data independent of the physical location. The desire would be to keep all data collected on-line at SNS (in addition to the tape archive), but the practicality of this desire is still being evaluated. The “push” architecture described above where a facility indicates when it has data is naturally scalable to support additional facilities.

Summary Table (SNS)

Time Frame	**Science Instruments and Facilities**	**Process of Science**	**Anticipated Requirements**	
			Network	**Network Services**
Near-term	• 9 instruments	• 1000 users / year • Processing • Simulation • remote data access	• 1 Gbps needed • key endpoints: LANL NIST	• ORNL XCAMS system to support user authentication
3-5 years	• 16-18 operating instruments	• Same as above, but number of users will be continuing to increase.	• 10 Gbps needed • key endpoints: LANL NIST, ANL, Univ Indiana.	• Envision utilizing DOE grid certificates for user authentication.
5+ years	• additional 22 instruments.	• increased need for simulation.	• 30 Gbps needed • key endpoints: same, which increased traffic to ANL.	• May enable authentication via non-DOE grid certificate authorities such as international partner organizations.

Summary Table (Theory and simulations* requirement for CNMS)

Feature	Science Instruments and Facilities	Process of Science	Anticipated Requirements	
Time Frame			Network	Network Services
Near-term	• Electronic structure calculations, quantum many-body simulations, and statistical simulations on large clusters, high-end capacity computing machines, and capability computing machines. • Electron transport in nanostructures. • Emergent behavior in strongly correlated electron systems. • Statistical physics in non-equilibrium systems.	• Users generate output that can be a significant fraction of the available memory of a machine. • Users will want to transfer data files back to their home institutions for future reference, analysis, etc. (several times per day). • Memory ranges from hundreds of GB on clusters to tens of TB on capability machines.	• Network bandwidth needed from computer centers (NCCS) and CNMS to user home institutions. • Data transfers between computer centers and visualization capabilities at CNMS. • Number of large users will be of order 10-30.	
5 years	• Same as above • Simulations used in real time for data analysis at the neutron scattering facilities (SNS and HFIR).	• Same as above but with magnitude scaled by factor 100. • Users move experimental data and simulations results between SNS, computer center, and CNMS.	• Same as above • Computing and storage systems need to be staged to allow seamless transfer of data and simulational capabilities between SNS/HFIR, CNMS, and computer centers (NCCS, NERSC, ANL).	
5+ years		• By 2015 we are talking about ~100PB of memory on exa-scale machine.		

(*) From discussions with experimental groups in the CNMS we concluded that they would follow in the shadow of solutions provided for theory and simulation needs as well as solutions provided for SNS.

3 Issues

The following issues were reported at the workshop.

SNS Connectivity Issues

There is a need for better connectivity between the SNS at ORNL and the NIST Center for Neutron Research (NCNR) in Gaithersburg, Md. There was a discussion of possible options during the workshop, and a decision was made to approach NIST regarding a connection between NIST and the MAX Gigapop. ESnet and many other research and education networks have a high-bandwidth presence at MAX, and a connection between the NIST Gaithersburg facility and MAX would provide NIST with the ability to connect to ESnet as well as other networks at high bandwidth (ESnet's connection to MAX is currently 10Gbps).

Data Transfer Issues

Several attendees indicated that one of the barriers to using the network for data transfer was a lack of expertise in the setup and maintenance of high-speed data transfer systems. In addition, in many cases (e.g. the light sources) the facility users do not have the knowledge to troubleshoot data transfers at their home institution.

For example, NSLS representatives at BNL provided the following:

> *Like in most of other light sources, data intensive users tend to find that portable media (hard disk, DVD) a convenient option to move data to their home institutions. This can create problems – for example, sometimes a team member might have to stay late just to finish burning DVDs. Currently, the bandwidth provided by ESnet4 backbone is more than enough for network transfer of user's data. We should make every effort to take advantage of the bandwidth to promote the use of network for data transfers. Currently users are not using the network for the following reasons: (1) The facilities are not providing a high-speed data transfer service, many times because of cyber security concerns. (2) Our users have a lack of bandwidth at their home institutions. As user's home institutions improve their network, and in fact many are already affiliated with Internet2, this situation could be improved over time. (3) There is a lack of technical expertise and dedicated computing/networking staff to tackle technical difficulties at the users home institutions. For example, in high energy physics, dedicated staff for networking and data storage are available to solve technical issues, whereas in the case for light source users, group tends to be small, and dedicated computer/networking staff at user's home institutions are mostly non-existent.*

In addition, ORNL provided the following case study from February 2007:

> *It took 2 weeks and email traffic between network specialists at NERSC and ORNL, sys-admin David Skinner at NERSC, Norbert Podhorski of SDM, and combustion staff at ORNL (Ramanan Sankaran) and SNL to move 10 TB of INCITE data from NERSC to ORNL. After iterating on the protocol to use (scp,*

> *rsync, bbcp) we converged on 4 bbcp transfers at once (16 channels each) from NERSC/davinci to ORNL/ewok. Under the best scenario we could achieve data transfer rates of 24MB/sec. However, practically most often we could only see ~12 MB/s during the day, more than one from ten transfers failed, creating additional delays in the workflow, and usually we had 3 transfers at a time instead of 4.*

Clearly there is a need for education and documentation on setting up and configuring data transfer systems and tools, and for installing and using network performance analysis and troubleshooting tools. ESnet has begun to address this by building a new web site: http://fasterdata.es.net, which contains information on bulk data transfer tools and techniques.

4 Overall Summary and Conclusions

It is clear from the case studies that many facilities funded by BES use similar technology for data detection and acquisition, and that these detectors are on a “Moore’s Law” growth curve in terms of their data production capabilities. This means that the data volume produced by the Light Sources, Nanoscience Centers, and the Neutron Science facilities will grow quickly, placing ever-increasing demands on storage and computational resources, as well as the wide area network.

An important finding at the workshop was the discovery of widespread frustration involved in the transfer of large data sets between facilities, or from the data’s facility of origin back to the researcher’s home institution. The reasons for the difficulties vary, but include the lack of modern data transfer tools, lack of knowledge of host system tuning, and congestion or similar issues at “the other end” of the data transfer. To help address this, ESnet will enhance and expand its network performance tuning web pages to include information describing modern data transfer software, host setup and tuning, and network configuration. ESnet will also continue to work with users and site to help with end-to-end performance issues.

The combination of instruments, facilities and science process represented by the BES research portfolio will consume significant networking resources over the next 5 years. The BES instruments and facilities tend to be data producers – the data must be sent to their users’ remote sites. As detectors and data acquisition systems are upgraded, the BES sites will easily consume 100 Gbps of network resources in aggregate in the next 5 years. Since many of the users of these instruments and facilities are from academia, there is a clear requirement for flexible, expandable, robust peering with university networks (i.e. Internet2) in addition to the need for high-bandwidth connectivity to the laboratories that house the instruments.

5 Acknowledgements

This work would not have been possible without the contributions and participation of those who provided information and attended the workshop. ESnet would also like to thank the BES and ASCR program offices for their help in organizing the workshop and providing insight into the science supported by the BES program. In addition, the LBNL conference support and logistics staff were very helpful.

ESnet is funded by the US Dept. of Energy, Office of Science, Advanced Scientific Computing Research (ASCR) program. Dan Hitchcock is the ESnet Program Manager.

ESnet is operated by Lawrence Berkeley National Laboratory, which is operated by the University of California for the US Department of Energy under contract DE-AC02-05CH11231.

This work was supported by the Directors of the Office of Science, Office of Advanced Scientific Computing Research, Facilities Division, and the Office of Basic Energy Sciences.

This is LBNL report LBNL/PUB-981.