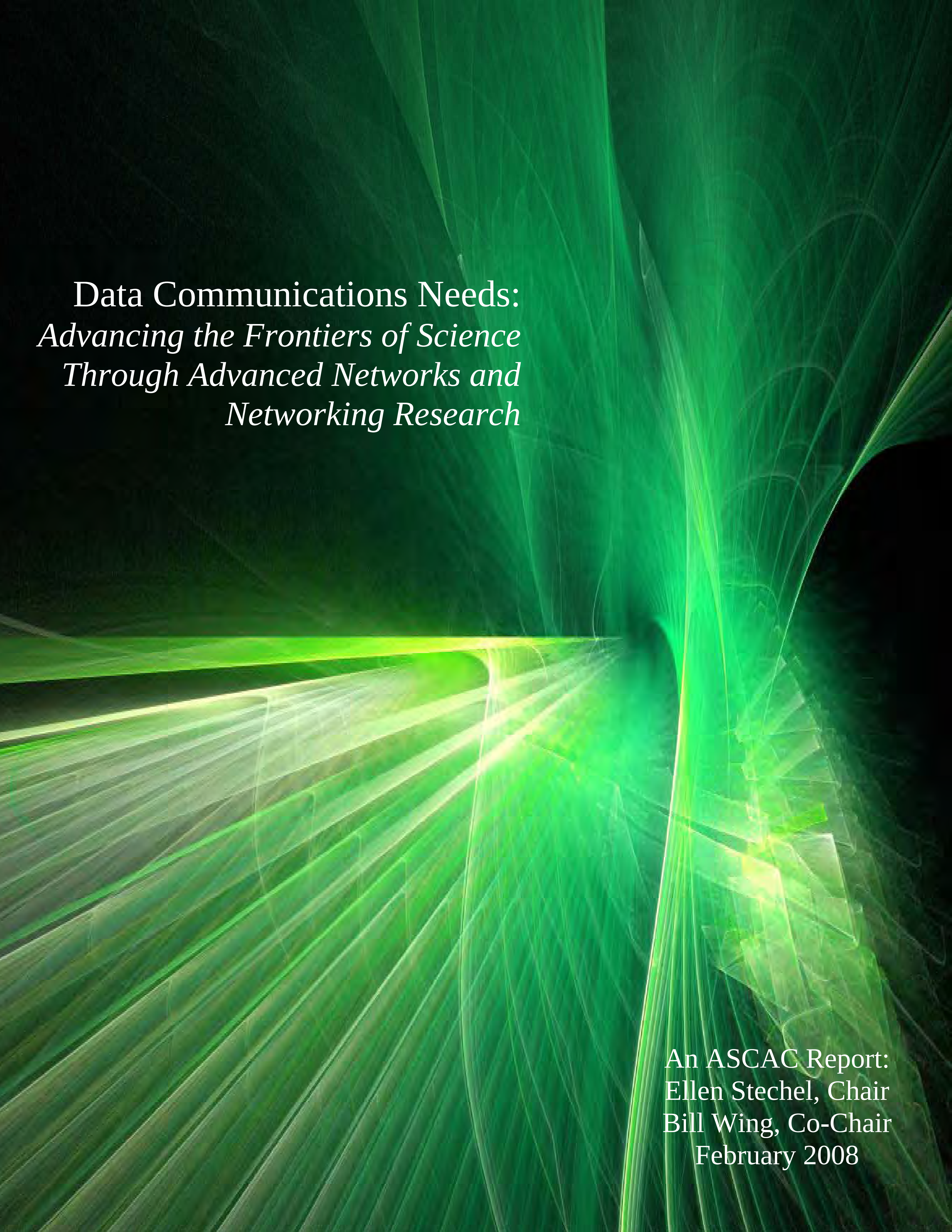


Data Communications Needs:
Advancing the Frontiers of Science Through Advanced Networks and Networking Research

An ASCAC Report:
Ellen Stechel, Chair
Bill Wing, Co-Chair
February 2008

FINAL DRAFT: 12 FEB 2008

Report of the Advanced Scientific Computing Advisory Committee Networking Subpanel

Data Communication Needs: *Advancing the Frontiers of Science* Through Advanced Networks and Networking Research

ASCAC Networking Panel Members

ELLEN B. STECHEL, Sandia National Laboratories (Chair)
WILLIAM R. WING, Oak Ridge National Laboratories (Co-Chair)
F. RONALD BAILEY, Advanced Management Technology, Inc.
JAMES CORONES, Krell Institute
IAN FOSTER, Argonne National Laboratory
HARVEY NEWMAN, California Institute of Technology
GWENDOLYN HUNTOON, Pittsburgh Supercomputing Center
LARRY RAHN, Sandia National Laboratories
HORST SIMON, Berkeley National Laboratory
JERRY SOBIESKI, Mid-Atlantic Crossroads

Submitted to the ASCAC on 12 February 2008 for consideration as a:
Report of the Advanced Scientific Computing Advisory Committee
March 2008

Table of Contents

APPENDICES

SIDEBARS

1 EXECUTIVE SUMMARY

This document reports the findings and recommendations of a subpanel convened to review the effectiveness of the Department of Energy's (DOE) existing network and networking research strategy in support of the department's scientific objectives. The subpanel interpreted its charge broadly in considering how the network will advance its utility for science and its anticipated role in enabling new and transformative modes of scientific inquiry as well as in enabling broad access to large data sets and unique analysis tools over the coming decade. The network technologies and services considered will require substantial and sustained research and development (R&D); therefore, the subpanel has based its findings and recommendations upon a ten-year horizon.

The first finding is that the Energy Sciences Network (ESnet) facility is doing an exemplary job in architecting, deploying, and operating a high-performance network infrastructure to serve DOE science needs. The successful basis for the operations of this facility has been a continuous requirements-driven process. The subpanel concurs that ESnet4 (the next-generation of DOE's science network) is critical in the relative near term and will play a continuing vital role in the long-term as new network architectures and services evolve and mature. Sustained and adequate funding for this facility remains essential. (§4.1[1])

A second finding arises from the following: High-performance networking is critical to programs in the Office of Science (SC,) including high-performance computing. The high-performance networks required will not automatically emerge from commercial R&D. Therefore, SC will need to fund strategic, high-performance networks and networking research. A corollary is that high-performance networks and networking research needs the same level of attention as given to high-performance computing. (§4.2)

The subpanel also observed a number of trends within DOE science:

- The network now underlies nearly every aspect of advanced e-science and, in fact, nearly all science activity.
- There is general acceptance that computational *simulation and modeling* has become the “third leg” of scientific inquiry. As such, it constitutes a major driver for cyber resources. Beyond the basic computational requirements, these applications depend upon tightly integrated access to storage facilities, visualization facilities, and advanced high-performance networks.
- Scientific inquiry is moving towards more quantitative understanding of ever more complex systems, whether that is from sophisticated experiments in SC's one-of-a-kind facilities or from large-scale modeling and simulation. Acceptance or refinement of the conclusions from such scientific inquiries depends on scrutiny from many eyes and from different research groups, approaches, and perspectives. This implies a need for data packaging, access, and mobility.
- There is an accelerating trend in the quantity and diversity of raw data captured from instruments, computations, and sensors and archived for reuse, again implying a strong driver for data handling, accessibility, and mobility.
- The continuing trend is toward *global collaborative e-science*, where science teams, science facilities, and science itself crosses a wide range of boundaries, including traditionally distinct disciplines,

---

[1] The section number, here and those below, point to further discussion in the body of the report.

funding agencies, and nations, once again indicating a major driver for data communication and data mobility.

- An increasing trend toward the development of new cyber systems by combining and integrating existing facilities is driving a need for a System of Systems approach to scientific infrastructure.

The magnitude and scope of these trends pose significant challenges and opportunities for DOE cyberinfrastructure development. Opportunities arising from advanced network technologies provide the avenues to overcome such challenges and to transform or redefine how one can conceive and construct large-scale e-science environments.

The following summary findings highlight the issues the subpanel considers most significant:

- In the coming decade, much of the network technology necessary to meet DOE (and other science-based agency) e-science requirements will *not* arise naturally from commercial R&D. The challenge then, given the central importance of high-performance networking to SC programs, is to construct a long-term and far-reaching network R&D program that sustains innovation from basic research to prototype to early deployment and culminates in a production network environment for use in advancing science. (§4.6)
- The pursuit of DOE science will continue to advance with an accelerating dependence on networks and related technologies to unite disparate teams, to allow efficient exchange of information and applications, and to enable new modes of scientific inquiry. The challenge is to develop network architectures and service models that support virtual organizations and provide reliable, predictable, and repeatable network performance, accessibility, and security. (§4.3)
- A key movement that will have profound impact on science, society, and the economy in the coming decade will be the development of national and global-scale, data-intensive (terascale to petascale and beyond) distributed “cyber environments.” The challenge will be to develop dynamic and intelligent (cyber) resource allocation architectures that will allow the DOE scientist to transparently and easily make use of resources at this scale. (§4.4)
- The network capacity and service capabilities anticipated for the next decade of science activities will likely be three to four orders of magnitude greater than current network architectures and technologies can effectively address today. The challenge will be to accelerate or develop a ten-year technology trajectory to achieve this projected need. (§4.6)
- There are many barriers to the development and adoption of promising new networking research ideas (across all agencies and industry.) These range from too narrow a focus (e.g., driven by already evident deficiencies,) to the classic “valley of death” between pure research and robust documented product. In the DOE environment, long-term delays in finding and leveraging useful new network technologies hinder the emergence of new scientific systems. The challenge is to establish mechanisms to identify promising networking research concepts and move them progressively through prototyping, experimental deployments, and ultimately into a production network service environment. (§4.5)
- Out of necessity, a few leading-edge science communities have pushed the state of the art in networks and networking services. The result has been vast discrepancies in the level of capabilities for data distribution and management available across the scientific enterprise. The challenge is twofold: first, to leverage and generalize the services developed in one context for the broader utility of the DOE science community and second, to develop advanced cyberinfrastructure service architectures

that allow future efforts to create, easily and effectively, the types of network (or other cyberinfrastructure) that they need. (§4.5)

Given these findings, the high-level recommendations from the subpanel are as follows:

1. Approach the development of advanced networking in a fashion similar to the manner in which SC develops its goals and objectives for petascale and exascale computing. The **network and networking services need to be an explicit and fundamental element of advanced petascale and exascale science,** of equal importance to leadership-class computing and large-scale scientific facilities in **enabling high-end science.** Hence, the subpanel encourages SC to create mechanisms whereby the DOE science community is encouraged to think broadly about how an unconstrained network resource might transform the conduct of science.
2. Establish mechanisms to manage the implications and issues that arise from the **System of Systems** aspects of DOE science, facilities, and programs. Such an approach will increasingly be required to enable interoperability needed for multi-disciplinary science. Advanced information technologies will enable this System of Systems; advanced networking architectures and capabilities will be particularly critical.
3. Reinvigorate an **aggressive, sustainable, long-term, and strategically focused networking research and development program** to create network-specific technologies that will allow DOE not just to increase the speed of existing systems but also to **transform the manner in which science is done**. The recommendation is that ASCR should convene an external committee to review this networking research program on a regular basis, in order to maintain a ten-year research horizon and integration across SC.
4. Formulate a deliberate **strategy to bridge the "valley of death"** for a networking research program that helps advance DOE science. Such a strategy would move concepts (whether from within ASCR's basic networking research program or from the networking research community more generally) through testbed deployment to production. The "valley of death" issue could be addressed with funding to support applied R&D and experimental or early adopter deployments. The strategy should involve increasing collaboration between the networking research scientists and operational/engineering facilities' personnel as concepts mature. Moreover, the application science community, as end-users, must be a continual and integral component of the collaboration.
5. Explore and develop methods for the automated monitoring, troubleshooting, diagnosis, and management of advanced network architectures and the applications that operate on those networks. The effective use of **next generation networks**, with unprecedented and rapidly expanding capacity and utility as well as complexity, **will require a new paradigm of operations and management**, including: end-to-end monitoring of the network (necessarily including the end systems;) automated methods for operating, managing, diagnosing, and alerting; and use of higher-level services to ensure optimal network resource utilization and workflow coordination and management. ASCR could convene a workshop to flesh out these and other priority research directions.
6. Integrate research in **data collection, archiving, curation, generation, pedigree, and access** into the ASCR networking research program. Effective integration of new networking technologies into the future data-management architecture is crucial to providing timely and secure access to and efficient migration or access to large datasets across an emerging cross-disciplinary globally, distributed science environment.

2 CONTEXT

Science, and DOE science in particular, is undergoing rapid changes in concert with the information technology revolution. Dramatic improvements in computer technology are enabling petascale simulations. This technology combined with new detection capabilities and magnetic storage technology is producing experimental datasets on the petascale. Optical communication technologies are enabling networks that can communicate this information among scientists and facilities across international collaborations. More broadly, unprecedented computation and communication capabilities resulting from the same information technology revolution are enabling scientists, and new approaches to science, in all areas. The roles of collection, storage, and communication of experimental data, international in scope, are becoming increasingly paramount to the scientific mission throughout the SC. In fact, this revolution is now affecting changes faster than the scientific community or its institutions is able to respond.[2] Within SC, ASCR is working to capitalize on this revolution for the benefit of DOE science and the nation: [3]

> The mission of the Advanced Scientific Computing Research (ASCR) program is to underpin DOE's world leadership in scientific computation by supporting research in applied mathematics, computer science and ***high-performance networks*** and providing the high-performance computational and ***networking resources*** that are required for world leadership in science. [Highlighting added.]

Over the next decade a number of large experimental instruments will come online that will significantly increase the demands on networking, communication, and connectivity. For example, SC will have a major role in operating experiments at the Large Hadron Collider (LHC) at CERN. In addition to these experimental programs, SC has begun operations of simulation facilities that will also generate enormous amounts of data. Taken together, the expectation is that the amount of data will soon exceed hundreds of petabytes. Today, within ASCR and the programs that involve networking research and implementation, including ESnet, considerable change is occurring to meet these important new challenges.

In anticipation of these and other significant transformations in the conduct of science, the Director of SC, Dr. Raymond L. Orbach, requested that the Advanced Scientific Computing Advisory Committee (ASCAC) convene a subpanel to examine the role and efficiency of networking and networking research within SC.

2.1 THE CHARGE

2.1.1 As Written

The subpanel should

- Weigh and review the organization, performance, expansion, and effectiveness of the current operations of ESnet.
- Consider the proposed evolution of ESnet, its appropriateness and comprehensiveness in addressing the data communication needs of SC **that will enable scientists nationwide to extend the frontiers of science**. [highlighting added for emphasis]

---

[2] "Preparing for the Revolution: Information Technology and the Future of the Research University," The National Academies Press, 2002, p45.

[3] ASCR website, http://www.science.doe.gov/Program_Offices/ASCR.htm.

Furthermore, the subpanel needs to

- Make suggestions and recommendations on the appropriateness and comprehensiveness of the networking research programs within ASCR with a view towards meeting the long-term networking needs of SC.

2.1.2 As Interpreted

The subpanel interpreted the words emphasized above "*that will enable scientists nationwide to extend the frontiers of science,*" as the organizing principle for their deliberations. The subpanel further parsed the charge into three elements:

(1) *Expansion and effectiveness of current operations*. In light of a recent and comprehensive Lehman Review[4] and a more recent Operational Review[5], the subpanel de-emphasized this element of the charge noting that the subpanel concurs with the findings and recommendations of these reviews.

(2) *Data communication needs as well as long-term networking needs*. The subpanel interpreted this element broadly and focused principally on higher-level needs, those that go far beyond what might come out in a requirements-driven process. The subpanel finds that the results of several requirements-driven workshops are already well documented[6] and well represent the foreseeable requirements over the next few years across all the relevant science domains in SC.

(3) *Appropriateness and comprehensiveness of the networking research*. The subpanel focused on the characteristics of advanced networks and their integration into enabling petascale science and on whether a vision and a strategy exist for the research portfolio. Resolution of these high-level issues with buy-in across SC will greatly enable development of detailed research strategies. Thus, this report deliberately avoids a prescriptive delineation of detailed research directions.

The charge was not explicit on the timeframe. The subpanel concluded that, in order to serve the charge best, a timeframe covering the next 10 years was appropriate.

In addressing the charge, the subpanel operated with the following assumptions:

- Others have already made the strong case for DOE science, a case that is uncontestable, including that for petascale computing, petascale datasets, and large-scale scientific facilities. Indeed, providing strategic major scientific facilities is one of the most important activities of the SC in DOE.
- SC leads all agencies in funding the physical sciences in the United States. The breadth of science supported by SC is as large as science itself, spanning high-energy and nuclear physics, condensed matter and interfacial science, fusion energy, chemical sciences, geologic sciences, biological and environmental sciences, applied mathematics, and advanced scientific computation.

---

[4] Lehman Review Reference [need to get appropriate reference]

[5] Operational Review Reference [need to get appropriate reference]

[6] Requirements documents: http://www.es.net/ESnet4/Case-Study-Requirements-Update-With-Exec-Sum-v5.doc; http://www.es.net/pub/esnet-doc/BES-Net-Req-Workshop-2007-Final-Report.pdf; http://www.es.net/pub/esnet-doc/BER-Net-Req-Workshop-2007-Final-Report.pdf.

Moreover, that support will grow with the increasing realization of the critical mission of SC, and through resulting legislative actions, such as the America COMPETES[7] Act.

- The DOE must meet critically urgent research needs to address its mission in confronting energy and national security issues. For example, this point was highlighted in the Basic Research Needs to Assure a Secure Energy Future report from the Basic Energy Sciences Advisory Committee (BESAC): - "*Considering the urgency of the energy problem, the magnitude of the needed scientific breakthroughs, and historic rate of scientific discovery, current efforts will likely be too little too late. Accordingly, BESAC believes that a new national energy research program is essential and must be initiated with the intensity and commitment of the Manhattan Project, and sustained until this problem is solved*."
- The high-end scientific and computational resources available to the DOE research community should be second-to-none. This point is called out explicitly in the American Competitiveness Initiative as one of the goals for its research: "*World-leading high-end computing capability (at the petascale) and capacity, **coupled with advanced networking**, to enable scientific advancement through modeling and simulation at unprecedented scale and complexity across a broad range of scientific disciplines* ..[8] [highlighting added]
- Petascale data, petascale computing, and experimental facilities for frontier DOE science create a triad forming the pillars of **petascale science**. It is petascale science (not just petascale computing) that is increasingly required as a critical element for addressing national and global priorities – such as understanding global climate change; innovating improvements in energy utilization; protecting our natural environment; applying genomics-proteomics to human health and new energy production; maintaining national and energy security; mastering the world of nanotechnology; and predicting, protecting against, and recovering from natural and human disasters –as well as addressing some of our most fundamental intellectual questions, such as the early formation of the universe and the fundamental character of matter.
- Cyber security will be an important element of advanced networks. However, its importance has risen to the level of task forces and dedicated workshops. Hence, extensive discussions in this report would most likely be duplicative of these parallel efforts.

While this subpanel supports the bold statement from BESAC, the question arose whether the DOE science community could not also dramatically accelerate the historic rate of scientific discovery. Petascale science is an example of just such acceleration, enabled by advanced networks and especially networking (science) services derived from advanced networking research. Indeed, mission requirements, scientific leadership, and urgency (at least on the scale of a Manhattan project) require radical improvements in time to solutions and thus motivate newly enabled large-scale, high-throughput, and system-level approaches to science.

2.2 Background: Networks and Networking Research within SC

2.2.1 Networking Research

Within the purview of ASCR's networking research program, there have been significant investments in infrastructure, new concepts, and basic networking research in transport protocols, cyber security, and

---

[7] COMPETES: Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science

[8] ACI, 2006 http://www.ostp.gov/html/ACIBooklet.pdf

high-speed (>10 gigabits per second) large data transfers, as well as measurement and analysis of the ensuing traffic. For example, DOE installed and operated the first nation-wide, pure research, 20 gigabits per second (Gbps) network (UltraScience Net or USN[9]) as well as developed the switched-circuit basis for ESnet's Science Data Network (SDN.) Middleware research in ASCR has aimed to enable access to remote resources and to allow distributed teams to work together. Moreover, the collaboratory pilot projects provided early implementations of virtual research organizations for distributed, discipline-oriented applications, while the DOE Science Grid project explored the creation of multi-laboratory distributed computing infrastructure. Nevertheless, much of the department's networking research portfolio has been lost to other priorities, at least temporarily. ASCR remains in a unique position to enable a successful model of interdisciplinary, highly complex, system-level science through the development, verification, validation, and deployment of advanced concepts for networking services and technologies, should the organization reinvigorate an innovative networking research program.

2.2.2 Deployment of Advanced Networks for DOE Science

The primary mission of ESnet is to enable the large-scale science in SC and research efforts that depend on the following elements for exploring and solving scientific problems:

- Sharing of massive amounts of data
- Supporting thousands of collaborators world-wide
- Working in collaborative cyber environments
- Harnessing and sharing distributed computing, storage, and network resources
- Distributed data processing
- Distributed data management
- Distributed simulation, visualization, and computational steering
- Collaboration with the U.S. and international research and education community

In accordance with this mission, throughout its history ESnet has provided a reliable high-speed communications network infrastructure enabling thousands of DOE, university, and industry scientists and collaborators worldwide to use unique DOE research facilities and computing resources effectively, independent of time and geographic location. User demand to ESnet has grown by a factor of greater than 10^4 since ESnet began providing SC research-wide, networking services in 1988 – a 100% increase every year since 1990.

In August of 2006, ESnet partnered with Internet2[10] to deploy a highly reliable, high-capacity nationwide network that will greatly enhance the capabilities of researchers to participate in the DOE's scientific research efforts. The new network created through this partnership operates on two dedicated 10 Gbps wavelengths on Internet2's nationwide infrastructure and will seamlessly scale by one wavelength per year for the next four to five years in order to meet the needs of large-scale SC projects. This network will support new optical services like point-to-point dynamic circuits, which will serve as an advanced and dependable platform for scientists and researchers. ESnet also partners and collaborates with other national network infrastructure initiatives such as USN and National LambdaRail (NLR[11].) USN provides on-demand dedicated bandwidth channels at multi, single and sub lambda resolutions (SONET and Gigabit Ethernet) between its edges. Various types of protocol, middleware, and application research projects can make use of the dedicated channels provisioned by USN. ESnet provisions 10 Giga-

---

[9] http://www.icair.org/main_projects_optical.html#top9

[10] http://www.es.net/ESnet4/Internet2-ESNET-071607.html ; http://www.es.net/ESnet4/esnet.083106-1.html

[11] http://www.nlr.net/about/

bit Ethernet waves across NLR, a nationwide fiber optic infrastructure whose defining characteristic is its ability to support many distinct networks for the U.S. research and education community using the same core infrastructure.

As highlighted in the report on Facilities for the Future of Science: A Twenty Year Outlook[12], the ESnet upgrade will enhance the networking services available to support SC researchers and laboratories and to maintain their access to all major DOE research facilities and computing resources, as well as fast interconnections to more than 100 other networks.

The ESnet facility and networking research will be facing demanding capability and capacity needs arising from petascale data.

Sidebar 1: Networks 101

NETWORKS 101

Networks, in their simplest form, consist of a set of switching elements (essentially specialized high-speed computers) interconnected by high-speed telecommunications lines. Messages transit these switching elements one at a time – each element inspecting the message and forwarding it to its destination according to a common set of forwarding rules. This is fundamentally, how the Internet works: Users at home or at work send a message from their PC, through the Internet to a web server requesting a web page. The server sends a series of messages back to the user, these messages containing the contents of the website the user is surfing. The Internet Protocols, or IPs for short, collectively refer to the standards and common forwarding rules used within the Internet.

This process is essentially the same for high-end science and research networks. Researchers are able to move large datasets from repository to computing facility by breaking the dataset into a series of “packets” (messages), transmitting them through the cloud of switching elements (often times stretching across the globe), and reassembling the dataset at the destination. There exists well-known and universally adopted transport protocols (e.g., the Transport Control Protocol, or “TCP”) used to transfer data, and to ensure that it arrives completely and accurately.

The Internet was designed to be a shared medium. Messages from users can be aggregated from relatively slow access lines (DSL, cable modems, at 10^6 bits/sec, or 1 megabit/sec) into higher-speed backbone lines (10^{10} bits per second or 10 Gbps) to reach destinations across the country or around the world. Organizations can build networks in their campus and link them to similar commercial networks in their metropolitan region. In addition, those metro networks can link with national or international networks, creating an interoperable network – the “Internet” – spanning the globe. **ESnet** is an example of a national network that links to (or “peers with”) other national and international networks.

The shared infrastructure of the Internet allows for great cost efficiencies, particularly when many thousands of users are sharing the network. However, the same simplicity that made the Internet so attractive has exposed, over time, some serious problems. For example, if any user (or group) decides to flood the network with packets – such as might happen with a large file transfer –, the network switching elements can become overwhelmed and unable to keep up and consequently will ignore (“drop”) messages from the network. As this congestion builds, TCP will request a resend of those dropped messages, thereby making matters worse. The entire network risks a congestive collapse in which no user’s packets will be completely delivered and all communicating applications will grind to a halt.

To prevent this condition, end systems in the Internet use protocols that can detect congestion events and respond by sending *less* data –or, more accurately, sending the data at a more modest rate – until the congestion clears. This behavior is critical for the Internet to function, but it implies that *all* users will slow down when congestion is detected, whether one is surfing YouTube or calling an ambulance.

---

[12] http://www.science.doe.gov/Scientific_User_Facilities/History/20-Year-Outlook-screen.pdf

NETWORKS 101 (continued)

This same situation occurs even in high-performance networks when many different science teams all compete for the same, shared network resource and each has the ability to saturate the network with messages. This cross-flow traffic interference also occurs when different types of traffic interfere with each other– for example, file transfers saturating network links over which real-time traffic such as video is also flowing.

One refers to these networks as “best effort” because the network cannot guarantee that a packet or message will be delivered to the destination successfully, but the network will provide its best effort to do so.

Network friendly protocols such as TCP have other unexpected user consequences as well. For example, reliable transport protocols must retain messages at the source until acknowledgement of the successful receipt at the destination. When the Internet used relatively low speed communications lines and most file transfers were done between buildings on a single campus, this holding of such messages at the source required relatively little memory. However, in high-performance networks deployed since 1990, as the backbone capacity increases and spans the globe, the number of messages in transit between source and destination increases dramatically. The result is periods where the sender fills up memory with unacknowledged messages and then sits idle until the destination acknowledges receipt of those messages. For scientists attempting to move large datasets, this inefficient use of CPUs and high-performance networks results in significant delays and poor return on network investment. For network engineers, the bursty nature of such traffic can make it difficult to size properly the switching elements and communications lines in the network.

One means to address congestion and transport issues is to build networks with larger and larger backbone communications links. Greater capacity reduces the likelihood of congestion. Over the past decade, communications links have increased by a factor of four every few years, from 600 megabits/second, to 2.4 Gbps, to the current 10 Gbps: 40 Gbps technology is available but still cost prohibitive, and 100 Gbps is in the labs.

Solutions to all these examples of basic networking problems exist in today’s networked environments. Nevertheless, these demonstrate the range of issues that arise and need to be addressed in even simple network implementations in order to use the full potential of such networks.

New network concepts are emerging – often driven by the science and research sector. Two broad examples are photonics and network resorting. In the former, over the past 10 years, developments in the telecommunication optics field have introduced dense wave division multiplexing (DWDM.) DWDM technologies essentially allow as many as 160 colors of light – each carrying a separate optical communications link - to co-exist on a single fiber. This optical technology provides not only greatly enhanced fiber utilization, but also reduced power and size of the components, making such technology practical for even small organizations to deploy and operate high-capacity networks. The standards and common forwarding rules associated with these networks may not adhere to the standard Internet protocol or may allow the end user to determine the specific protocol used. Such networks are often referred to as hybrid networks since they use a mix of networking standards and technologies.

To date, networks have been designed to be as transparent as possible to the user application. That is, the user could send a packet from one PC to another without concern about network load, congestion, or length of time. However, new paradigms are emerging that view the network as a tangible and manageable application resource just as large computational clusters, data repositories, or sensors (e.g., LHC) are currently handled. In these emerging e-science environments, the applications are incorporating inexpensive DWDM-based communications links to connect the various locations and functional components directly involved in a distributed science project. Applications can now build their own network.

It is not clear how these new concepts will evolve or how they will ultimately influence network applications; but at least in these two examples, they appear to be taking the stage because of the advantage they bring the high-end science community and large, distributed applications in general. Hence, there is a critical need for a strong and sustained research mission to continue and guide this evolution.

Sidebar 2: Emerging Requirements for Future Networks

EMERGING REQUIREMENTS FOR FUTURE NETWORKS

Perhaps the most notable paradigm shift occurring within the science and high-performance networking community today is the notion of the network as a quantifiable application resource. This creates a fundamentally different approach to application development from that taken since IP networks began to receive wide acceptance. The application developer is now directly involved in defining network engineering and configuration requirements. The application designer (the scientists) is unfettered by limitations of the current available network or the local five-year network upgrade plans. The scientist is now thinking about how best to accomplish the science, rather than how to best accomplish the science within the networking technology limitations.

The radio astronomy community illustrates this point. For the past 30 years, the radio astronomy (RA) field has linked radio telescopes around the world to create a very long baseline interferometry (VLBI) instrument. All telescopes in a particular observation point at a single celestial object and listen to the RF noise emanating from the point. These observations are highly synchronized and recorded to tape or disc. The discs typically are shipped (FedEx!) to a correlation site where the individual streams are correlated with one another to create extremely high-resolution images of the cosmos. Only in the past five years has the data begun to be transferred over high-performance networks. The current model assumes approximately 25 telescopes (and often far fewer) are able to participate in a single observation at any time.

When asked how they would create such a large-scale science tool without consideration of conventional technical limitations, astronomers came up with a tool that spans a continent and contains 3000 antennae (telescopes) each generating with ten times the sensitivity of current radio telescope sensors. The resulting System of Systems for VLBI would generate well over 200 terabits/second – ~10^4 times more data than can be carried on a single high-capacity network link today. Yet when asked now how such a tool can be engineered, astronomers were surprised to discover that the key technology was actually tractable. Perhaps the most challenging issue will be the networking, for which total capacity requirements are projected at *four to six orders of magnitude greater* than existing processes. [See the “square kilometer array” project, http://www.skatelescope.org]

This example illustrates the issues that make many in networking community feel like deer in the headlights. Applications in the five- to ten-year time frame are well within the terabit/second range, namely, 100 times greater than our current maximum link technology provides. Given applications planning of the scale of the LHC or the VLBI (above), the technology trajectory of 100 gigabit/second links in five years becomes clearly inadequate to meet our projected applications requirements.

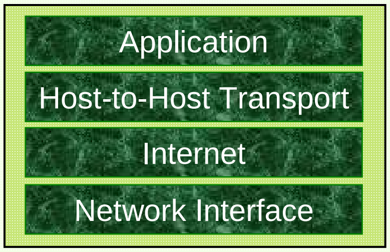
Another key aspect of networking that should be explicit is *scaling*, defined here as the ability to provide a similar overall network performance experience for every anticipated consumer. For example, a centralized database for routing information within the network will work fine for small network; however, as the network routing entries become numerous and updates increase because of the number of networks making changes, the centralized database becomes a choke point. Overwhelmed with the activity, it becomes a dangerous single point of failure for the network. Automated systems and protocols have been deployed to distribute the routing information throughout the network, thus improving both the performance and reliability.

Emerging e-science applications that individually pose challenges to existing network architectures will pose far greater challenges when scaling issues are considered. The prospect of scaling the network to support terabit applications is a tremendous challenge. We want to ensure that broad ubiquitous deployments of such applications are not hobbled by unanticipated scaling limitations.

Sidebar 3: Network Services

NETWORK SERVICES

A “network” is far more than the physical media (fibers, wires, etc.) that transport packets from sender(s) to receiver(s): it also includes the protocols, hardware, and software used to route packets, to map from domain names to network addresses, and to provide reliable end-to-end data transport. The Internet Engineering Task Force (IETF) defines four layers of network functionalities, starting from the top layer: application, host-to-host transport, Internet, and network interface. The layers have distinct responsibilities:



1. Placing and receiving packets on and off the network medium
2. Addressing, packaging, and routing
3. Providing the application layer with session and datagram communication services
4. Enabling applications to access the services of the other layers and defining the protocols that applications use to exchange data

High-speed transport can require innovations at every level in this protocol stack. However, the new research modalities discussed in this report require far more from the network than simple data transport. Indeed, the most challenging future requirements are likely to be for higher-level services –components categorized as *applications* – that provide important capabilities required to enable system-level, collaborative, and data-intensive science. These services often must be *network aware*, meaning that they can interact with the network to determine status and negotiate quality-of-service delivery. We give here four examples of where such services may be required.

- *Reliable, secure, and high-performance data distribution.* We frequently encounter the need to distribute large quantities of data to one or more participants. Meeting end-to-end requirements for performance, security, and reliability can require such methods as intermediate caching; end system reliable multicast; adaptive data reduction; and automated, policy-driven data placement and replication. [e.g., see the methods used in the LIGO Scientific Collaboratory, as described in A. Chervenak, R. Schuler, C. Kesselman, S. Koranda, and B. Moe, “Wide Area Data Replication for Scientific Collaborations,” *6th IEEE/ACM Int'l Workshop on Grid Computing*, 2005.] Such services may want to negotiate access to network bandwidth, storage, and computing capabilities embedded in the network or provided by participating end systems.
- *Service hosting and discovery*. Data, data analysis, and modeling and simulation capabilities are increasingly delivered to their user communities as network-accessible services. Enabling such “service-oriented science” approaches [See I. Foster, “Service-Oriented Science,” *Science*, 308, 814-817, 2005] can require new capabilities for publishing, hosting, and provisioning of such services.
- *Authentication, authorization, and accounting.* Both collaborative, system-level science, and the operation of the System of Systems that is modern science infrastructure require integrating many resources from many institutions, with different access control policies, and the participation of many people (and software systems) capable of acting with different roles and privileges. Resource consumption by both individuals and collaborative teams (“virtual organizations”) must often be tracked.
- *Management and troubleshooting.* The effective operation of the System of Systems that is the modern networked science environment will require increasingly automated operations and troubleshooting. New interfaces will be needed to access relevant data, and new services deployed and operated to meet these requirements.
- *End-to-end service with increasing degrees of intelligence.* These are needed to contain the manpower burden as the scale and complexity of the network and its workload expand.

As such systems evolve, their pervasiveness and level of knowledge of what is happening, on both the micro- and macroscales, increase. In addition, their level of self-awareness of which services exist, where, what is the state of each service, tasks in progress and queued, available diagnostic information increases likewise. Operators handle or mitigate only higher level or complex problems. It will be important for DOE to determine requirements for such services and to think carefully about how relevant new services should be provided and operated. Currently, responsibility for creating and hosting higher-level services rests with individual scientists or science communities. Over time, it may make more sense to shift the responsibility for certain services to specialist operations teams at individual sites or at other entities such as ESNet.

The Open Science Grid (OSG) [See www.opensciencegrid.org, and also R. Pordes et al., “The Open Science Grid,” *Scientific Discovery through Advanced Computing* (SciDAC) Conference, 2007] is exploring some of these issues. This DOE and NSF initiative is creating and deploying distributed computing infrastructure for data-intensive science. OSG sites agree to deploy on their physical resources services for user authentication, resource monitoring, job submission, and data access. The OSG consortium then runs certain centralized monitoring, troubleshooting, security incident response, and other services. ESNet provides authentication services.

3 FORCES and TRENDS

The science enterprise supported by SC is rapidly increasing in complexity. Experimental facilities now produce vast amounts of scientific data. Computational facilities provide simulations of unparallel complexity and fidelity. The scientific community increasingly self-organizes into teams that span both traditional disciplinary boundaries and geography, and these teams often change. Scientific discovery in many areas now depends on the ability to deal with very large sets of federated data resident on a variety of geographically distributed platforms. A science team's members can best explore experimental data and simulations results and collaborate with each other if all these elements network together in a transparent environment that includes the appropriate tools. From this perspective, the network environment is a critical, fully integrated component of the discovery process.

Three key technological drivers both drive this complexity and enable the scientific community to deal with it: the exponential increase in computing power and solid-state memory; the even more rapid increase in magnetic data storage capabilities; and the dramatic increase in global communication bandwidth. Dramatic advances in high-speed optical telecommunications technologies have made the last driver possible, namely fiber networks using optical amplifiers and wave-division multiplexing. As the cost to perform a fixed amount of computation, store, and/or transport a fixed amount of data dramatically decreases, scientists are now attempting calculations requiring orders of magnitude more computing and communication than was possible only a few years ago. Moreover, as noted in many present and future experiments, scientists are planning to generate several orders of magnitude more data than has been collected in the whole of human history.[13] In this section, we review in more detail the forces and trends evident in the conduct of DOE science that form the basis for the findings and recommendations presented in the remainder of this report.

3.1 Trends in Computation

We take for granted that computer speeds will continue to rise with each new hardware generation, that machines will have more memory than before, that disks will hold ever more information, and that the network will increase in data throughput. We further anticipate that advances will come from a combination of hardware, software, computer science, and mathematical or numerical algorithms, thereby providing increasing (but scalable) complexities as well as new features. Indeed, exascale (10^{18}) computing is not very far in the future. Moreover, the community is already exploring an exascale initiative.

It is also clear, however, that practical engineering and operational concerns increasingly limit the evolution of large-scale computational systems. Power consumption, heat dissipation, physical size, operational support, security concerns, and shear cost will drive large-scale e-science into an increasingly geographically distributed architecture from which large-scale science systems can emerge as needed from a global pool of available and pedigreed sub-systems and services.

3.2 Trends in Collaborations and Virtual Communities

Large-scale science collaboration involving a dynamic community of many researchers, often international in scope, is increasingly common. The trend in the science enterprise is rapidly evolving to one of

---

[13] "e-Science and Its Implications," Tony Hey and Anne Trefethen, *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, Vol. 361, No.1809, Information, Knowledge and Technology. (Aug. 15, 2003,) pp. 1809-1825. Stable URL: http://links.jstor.org/sici?sici=1364-503X%2820030815%29361%3A1809%3C1809%3AEAII%3E2.0.CO%3B2-T

"virtual"[14] organizations using "virtual" facilities. Such communities are often self-organizing (sometimes rapidly reconfiguring) and depend on being able to establish trust among themselves – for example, personnel integrity, bona fide datasets, authorized access to shared resources. One can expect the nature of collaboration to change in unexpected ways as groups traverse geographic, cultural, and disciplinary boundaries in the pursuit of joint science objectives. Each (highly dynamic) collaborative group is trending toward using a set of allocated and owned resources carved out of a much larger distributed system of experimental, computational, data storage, software, visualization, and network resources. Many of these resources are not part of the directly funded DOE infrastructure.

SC created the SciDAC program to advance frontier scientific discovery by exploiting leadership class computing. Such advancement requires working across boundaries in SC. Moreover, SciDAC projects already consist of multidisciplinary teams of domain scientists, computer scientists, and applied mathematicians and increasingly are finding the need to include network scientists. SciDAC projects have also responded to the need for developing collaborative software environments and Grids where distributed resources and expertise combine to address complex questions that no single institution can manage alone.

3.3 Trends in Data

DOE science is growing ever more data-intensive. Many applications are beginning to depend on manipulating massive amounts of data, far more than can reside in memory, where such data arise from observational inputs (possibly from distributed sensor networks,) experimental measurements, large simulations, digital images, or videos. The ability to create, publish, and access these datasets – globally and securely – and a means of asserting their pedigree (authentic, verified, etc.) will be critical to developing integrated science systems in the coming decade.

Data management includes authorizing access to the data, certifying its veracity, publishing its structure, tracking its evolution and processes that produce new results, and linking new results into the system in order to establish the pedigree for those results. The management of data and workflow is becoming more complex as multiple researchers, using resources and services at multiple locations, use multiple scientific datasets for coupling simulation and experiment to address problems and potential solutions to ever more complex multidisciplinary global challenges. Experimental facilities are also relying more on working in concert with computation for data management, visualization, extraction of usable information, and interpretation of meaning. Observation may be rich in information content, but requires theory and modeling for "diagnosing" that rich information to turn it into knowledge and understanding.

Databases of a few terabytes are becoming common; only ones over 100 terabytes are now considered at the leading edge of what is practical. Disk capacities (measured as bits per square inch of magnetic material) are growing to accommodate this data, recently exceeding the historically growth of about 60% per year. Indeed, crucial data collections in the social, biological, and physical sciences are coming online and becoming remotely accessible. Modern genome research would be impossible without such databases. The high-energy physics community, for example, estimates that by about 2012 it will need a data archive capable of reaching the exabyte scale for data arising from four major LHC experiments. As a second example, the National Center for Atmospheric Research (NCAR) currently supports a petabyte of online data, and this data is growing at 10 terabytes per month.

---

[14] I. Foster, C. Kesselman, and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," International Journal of Supercomputer Applications, 15 (3.) 200–222. 2001.

3.4 Trends in Society

Under the influence of the evolving internet, significant worldwide changes are underway. While science-motivated networks serving major projects are pushing the leading edge, they do not dominate the network flows. An estimate of the flow of data on the Internet is greater than one terabit/second. This is in spite of a massive gap in technology, knowledge, and default system tuning and software. As applications grow (You Tube, streaming HDTV, eventually large scale binary data exchange for businesses and homes) the flow will grow to substantially greater than one petabit per second within the next ten years. At that point, all the "production of information" in all forms of humankind will be vastly Internet-dominated. The culmination of this trend is one in which the network will no longer be adjunct to computers. To some, possibly significant, degree embedded processing of information directly within the network becomes the norm. Google is driving this trend in Internet-based information.

3.5 Shifting Paradigm in Scientific Research

One consequence of crossing the petascale threshold is a widening gap between the needed capabilities and services of a network for a virtualized scientific enterprise and the needs and interests driven by commercial markets. Nevertheless, ESnet's SDN is already enabling the transport of large scientific datasets by early adoption of an emerging new paradigm –optical networks based on the materialization of an optical layer, operating entirely in the optical domain thereby beginning to enable high capacity end-to-end wavelength ("lambda") services that can provide, through wave division multiplexing (WDM), many "virtual" fibers on a single physical fiber.

In addition, we are now seeing the beginnings of a paradigm shift in the conduct of scientific research. The classic two modes of inquiry in scientific research, theoretical/analytical and experimental/observational, arguably have already grown to three modes of inquiry with modeling/simulation approaching equal footing with theory and experiment. With the emergence of this third mode of inquiry and petascale science, a fourth mode of inquiry is just beginning to emerge: collection sciences or *data intensive investigation*. With this shift, not only does the data need to be pedigreed, but the simulation and modeling codes that generate the next stage or derived datasets must also be pedigreed in order to build such datasets with confidence.

Simulations are beginning to match the complexity of the real world, with full spatial and temporal dimensions, incorporating realistic multi-physics models and thereby opening up a vast range of problems to quantitative investigations. Whereas one generally considers experimental observations as "perfect truth," experiments can only partially expose that truth, sometimes allowing imperfect deductions. In contrast, even fully exposed models are necessarily "imperfect truth." Moreover, if the model does not contain the key phenomena, no amount of simulation will provide reliable deductions either. However, as scientists are able to conduct experiments that are ever more sophisticated and to build, compute, and validate models that are more complex, the gap between experiment and computation is closing, revealing a detailed understanding of the underlying realities under investigation. Indeed, one can anticipate dramatic advancements in scientific understanding as increased capabilities allow scientists to better integrate the various modes of research.

Primary access to the latest findings in a growing number of fields is occurring through prepublications available on the Internet; secondary access is through preprints and conferences; only lastly is access through the traditional refereed archival papers. Consequently, there is higher risk for misinformation or misinterpreted information. Also, crucial data collections in the biological and physical sciences are now online and accessible broadly and remotely. However, many gaps remain in the practices, perceived value, and resources for generating metadata, storage, retrieval, and preservation and for repurposing, mining, analyzing, synthesizing, and integrating data resources. For example, the increasing access to distributed datasets created by other parties, coupled with the distribution of results without (or with only minimal) review, begs for a new means for curation, independent evaluation, and annotation. These and related functions will be carried out through virtual scientific communities enabled by new distributed data management capabilities and tools, advanced by current and future networking research, as well as updated policies among scientific research and sponsoring institutions.

The increased complexity and interdependence of the facilities, tools, and people in the continued formation of the science enterprise provide significant challenges for next-generation networking. To meet these challenges, it is necessary to view the network as one critical subsystem within the System of Systems that make up a dynamic whole.

3.6 Summary

Thirty years ago, the practice of science occurred at a single experimental or computational facility by single or small groups of researchers. The earliest role of the network, for example, at Magnetic Fusion Energy Computation Center (MFECC,) was to provide users access to computational facilities. Since then, networks have transformed the way scientists do research.

The DOE science infrastructure is evolving into a network-centric, multi-domain enterprise with international-scale collaborations. To date, the scientific community has seen only the beginnings of what is possible with advanced networks, especially with respect to tearing down long-standing disciplinary boundaries and accelerating the pace of scientific discoveries and the application of new knowledge. As the science process becomes less geographically focused (e.g., at a single laboratory or facility,) opportunities to incorporate resources from an increasing global pool become obvious and attractive. Flexible global, high-performance networks are fundamental to integrating these globally distributed resources into coherent, productive, and secure science environments. An advanced network will enable thousands of scientists from DOE laboratories, universities, and industry and their collaborators from around the world to work together effectively, with no limitations arising from spatial, temporal, or disciplinary separations. In addition, such advanced networks will enable the same researchers to maximize value from use of SC's world-class and unparalleled research facilities, both experimental and computational.

An advanced network has become an essential capability and infrastructural element for advancing DOE science and for accelerating the rate of discovery, the pace of scientific advances, and the speed of diffusion from recently created knowledge to the application of that knowledge in transformative ways. Moreover, networks and networking research are critical to enabling yet another emerging trend in science that promises enormous impacts on scientific discovery and application across the complete spectrum of DOE science. This trend has been described as a "systems perspective,"[15] 'System-level science,"[16] and "Systems Science,[17]" and is closely related to "integrated modeling environments."[18]

---

[15] Atkins Report, 2003; http://www.cise.nsf.gov/sci/reports/atkins.pdf, http://www.communitytechnology.org/nsf_ci_report/

[16] I. Foster, 2006

[17] Karen Schuchardt, et al, "Portal-based Knowledge Environment for Collaborative Science," *Concurrency and Computation: Practice and Experience*, Volume 19, Issue 12 , Pages 1703 – 1716 (25 August, 2007.)

[18] See DOE ASCR draft planning documents: http://www.sc.doe.gov/ascr/Misc/ASCRFacilitiesStrategicPlan.pdf and E3SGS: http://computing.ornl.gov/workshops/town_hall/energy_ecology.pdf

Sidebar 4: e-Science

e-SCIENCE

E-science processes are creating a network deluge in the capture, distribution, and processing of data. Providing the network resources to move large amounts of information around the world securely, fast, efficiently, reliably, and predictably are some of the most pressing issues for the networking research community. The current IP networks continue to perform well for many day-to-day functions and most conventional science applications and processes. Nevertheless, the unabated growth of the globally distributed e-science environment and the increasing number of these specialized science environments, combined with the explosive growth of the quantity of physical data, pose challenging issues for the network science and engineering community. Indeed, current networking technologies will not scale up to support such activities, if the current evolution continues (or more likely accelerates,) and especially when it crosses the threshold of revolution. Providing innovative network technologies and services not only will support but also will enable the revolution. Hence, these are just a sampling of the challenges networking researchers must address in order to meet the service requirements of future applications.

An overarching finding from the Blue Ribbon Advisory Panel on cyberinfrastructure [See Atkins Report, Jan 2003] was that a new age has dawned in scientific and engineering research, pushed by continuing progress in computing, information, and communication technology, and pulled by the expanding complexity, scope, and scale of today’s challenges. The capacity of this technology has crossed thresholds that now make possible a comprehensive cyberinfrastructure on which to build new types of scientific and engineering knowledge environments and organizations and to pursue research in new ways and with increased efficacy.

The Blue Ribbon Panel further envisions the use of cyberinfrastructure to build ubiquitous, comprehensive digital environments that become interactive and functionally complete in terms of people, data, information, tools, and instruments and that operate at unprecedented levels of computational, storage, and data transfer capacity. Increasingly, new types of scientific organizations and support environments for science are essential, not optional, to the aspirations of research communities and to broadening participation in those communities. These can serve individuals, teams, and organizations in ways that revolutionize what researchers can do, how they do it, and who participates.

These globally distributed e-science applications are sophisticated systems (System of Systems) that incorporate and rely on many other systems and sub-systems. They combine both hardware and software development, they incorporate shared instruments and facilities (e.g., computational clusters, LHC,) and they are often parts or larger initiatives and programs. The projects incorporate teams of researchers from many geographically distributed institutions, increasingly internationally so. These teams hail from private sector, universities, and even other governmental agencies. Developing the mechanisms and tools that can create a collaborative environment that allows and encourages these varied personnel and resources to work together is a multilevel challenge characterized by a System of Systems. Over the coming decade, the “network subsystem” will need to go far beyond simply providing the telecommunications links to carry packetized information. It will need a much closer integration of the application technical requirements, the data migration and storage facilities and middleware, security requirements, computational processing functions, and administrative and programmatic requirements.

4 Key Findings

Discussed in this section are the key findings, outlined in the Executive Summary. In each case, the subpanel makes recommendations for ASCR and SC. The intent of the findings and recommendations is to provide guidance and encouragement towards meeting the myriad of challenges arising from the trends discussed above.

4.1 ESnet's Exemplary Job

4.1.1 Findings

- The successful basis for the ESnet facility, to date, has been a continuous requirements-driven process. This facility has done an exemplary job in collecting and analyzing requirements, based on community input and workshops[19], especially the requirements update of February 2006 as presented to the subpanel by Eli Dart.[20]
- Furthermore, the subpanel finds that ESnet has done an excellent job of architecting, deploying, and operating a cost-effective, high-performance network infrastructure to serve DOE science needs, including deploying the SDN. In addition, ESnet has been highly effective in delivering bandwidth and staying ahead of the curve including rapidly and effectively productizing circuit switching, as prototyped by Ultra Science Net. ESnet has also been effective in understanding the need for network interfaces and software on the end-systems in order to ensure applications can adequately use the bandwidth and services associated with the backbone.
- The panel feels that the ESnet4 infrastructure is critical in the relative near term and will play a continuing vital role in the long-term as new network architectures and services evolve and mature. Sustained and adequate funding for these facilities remains essential.

4.1.2 Basis

The long-term traffic data shows a remarkably stable growth rate of tenfold every 46–48 months. LHC traffic peaks, during the summers of 2006 and 2007, indicate a growth factor of two in peak flows per year, which may exceed the historical trends. Nevertheless, the overall plan for the next three to four years appears solid and well justified, up to the 10–100 Gbps wavelength transition. The ESnet4 optical network backbone, together with metropolitan area networks (i.e., Bay area, Chicago, and New York plus Long Island, and other strategic locations throughout the U.S.,) will meet the needs effectively over the next three to four years. This will be accomplished by supporting a growing number of 10 Gbps links on a multiply connected national footprint, and by supporting a load of 30–50 Gbps on various legs by 2010–2011. Internet2, US LHCNet (DOE,) TransLight, Transpac (NSF,) and other international partners support the onward connections across the Atlantic and Pacific oceans. The corresponding links to the US LHC Tier-1 centers at the high-energy physics laboratories will increase from the present range of 20–60 Gbps (BNL-Fermi Lab) to 40–100 Gbps (BNL-Fermi Lab) over this time period, and U.S. LHCNet its current 30 Gbps to 80 Gbps across the Atlantic by 2010.

Hence, such networks are already, in aggregate, on the 100 Gbps scale and will reach the terabit/sec scale within the next five to seven years, driven by the convergence of three factors:

---

[19] See http://www.es.net/hypertext/requirements.html

[20] See Appendix D.

- Ability to use multiple 10 Gbps effectively over long distances, which has emerged over the past few years, driven in part by developments within the DOE-supported physics and computer science communities
- Increased affordability of 10 Gbps links, driven by the use of DWDM infrastructures
- Emergence of 40 Gbps links (beginning now) and 100 Gbps links (by 2010–2011.) For example, the expected completion for the 100 Gigabit Ethernet standards is by 2009, with first products expected by 2010.

4.1.3 Discussion

High-energy physics is clearly leading the way for other fields with its massive dataflows generated from *one-of-a-kind* experimental facilities. For example, the expectation is that LHC-related dataflows will expand from the current 100 Gbps to the 1000 Gbps (terabit/sec) range (roughly equal to 10% of the summed dataflow in the Internet today) sometime between 2010 and 2015.

Microscopy is leading the way in massive dataflows from (possibly remote-controlled) *distributed* experimental facilities. For example, the expectation is that dataflows from distributed electron microscope facilities providing multiscale images will reach 250 Gbps and beyond, sometime between 2010 and 2015, as well.

We can also expect a transition to occur in the 2010–2012 time-period on the ESnet backbone, as wavelengths on optical fibers modulated at 100 Gbps become available. Current projections are that at least three manufacturers of optical network transmission equipment will have these higher-capacity waves production-ready within 18–36 months from now. After the transition, the ESnet backbone capacity should rise to 500–600 Gbps, by approximately 2012. This is a remarkable expansion in a relatively short period. There is little doubt that technically, the advanced networks planned by ESnet, can and will be well used.

4.1.4 Recommendation

ESnet should continue to interact with the domain scientists supported by the offices in SC (BES[21], BER[22], FES[23], HEP[24], and NP[25]) to elicit impending changes in requirements and continue to focus on the modes of conduct in science. Nevertheless, DOE scientists need to have a plan, or a vision, for exploiting the performance capabilities of a network after the transition to wavelengths modulated at 100 Gbps become available.

4.1.5 Challenges to Implementing the Recommendations

Unlike petascale (and beyond) computing, it might be the case that only a few members of the community can contribute usefully to a vision beyond a five-year period. The reason is that the rate of technological advancement and change (including social change) is happening rapidly, enabling a change in the very conduct of science.

---

[21] http://www.sc.doe.gov/Program_Offices/BES.htm
[22] http://www.sc.doe.gov/Program_Offices/BER.htm
[23] http://www.sc.doe.gov/Program_Offices/FES.htm
[24] http://www.sc.doe.gov/Program_Offices/HEP.htm
[25] http://www.sc.doe.gov/Program_Offices/NP.htm

In addition, a large-enough increase in capacity often frees developers to speculate in new ways, leading to new concepts and new applications that were previously impractical or out of reach. Examples that can now be envisioned include a global real-time system, with either large streams for visualization and virtual collaboration with increasing resolution and "texture." Another might be terabyte (and then 10- and 100-terabyte data transactions, with the transaction time-constant adjusted downward over time, for increasingly agile data exchange and sharing among sites.

Sidebar 5: Networking in High-Energy Physics

NETWORKING IN HIGH-ENERGY PHYSICS

[See for example, "Networking for High-energy Physics at the LHC, H.B. Newman, on behalf of the International Committee Future Accelerators (ICFA) Standing Committee on International Connectivity (SCIC,) 2008.] LHC experiments will be carrying out full-scale network challenges in 2008, to help ensure that experiments can work as needed during LHC operations. A "Combined Computing and Readiness Challenge" (CCRC08) is planned for May 2008, shortly before the LHC is commissioned. The expectation is that these challenges will result in substantially greater network loads than previously experienced. There are several reasons for this: the target rates of each experiment will be larger, to test the computing model "at scale;" the experiments will work simultaneously; and the challenges involve far more than data transfers alone, so that the dataset production, management, and distribution tools are not yet all fully mature. By mid-2008, the efficiency of these tools, their scalability in terms of the number of successfully completed jobs per day, and their overall robustness should increase, resulting in a significantly greater data flow per experiment.

In considering the outlook in network requirements for the future, an additional factor is the rapid advance in some centers' basic ability to move data between storage systems over long-range networks. Some Tier2 centers (most notably some of the U.S. Tier2s) have shown that they can move data at rates of several hundred to as much as one Gbps. A pre-production demonstration, carried out last November at the Supercomputing 2007 conference, [See http://mr.caltech.edu/media/Press_Releases/PR13073.html] by a high-energy physics team showed that one can move data among relatively small storage systems over distances of 1,000 to 10,000 km at sustained high-speed rates using 10 Gbps links bi-directionally at high loading levels, limited mainly by the read- and write-speeds of the disks.

A number of network planning workshops and the roadmaps developed for the field by the ICFA SCIC (SCIC has been charged by ICFA to track such requirements), as well as ESnet and the major high-energy physics laboratories and some Tier1 sites, have foreseen the trends, and developed plans covering the next several years. U.S. LHCNet, for example, will move from three to four 10 Gbps transatlantic links across the Atlantic by early 2008, and has the goal of reaching eight 10 Gbps links by 2010. ESnet manages the connections to the U.S. high-energy physics labs, has begun to deploy ESnet4, where an SDN, whose expected capacity will reach approximately 50 Gbps by 2010-2011 complements the general-purpose 10 Gbps backbone. The longer term prospect, starting approximately in 2011 or 2012, is that standardized 40 Gbps or 100 Gbps links (generally on the present optical fiber infrastructure) will appear in production networks, including ESNet4. It is reasonable to expect that many of the major links supporting the program will reach several hundred Gbps by the middle of the next decade.

In spite of the bandwidth increases foreseen, the experience of the last few years has shown that the technical capability in the high-energy physics community to use the bandwidth effectively has progressed even faster. The outlook is thus that the available bandwidth will become a scarce resource as the perceived needs, as well as the ability of a growing number of sites to use the network efficiently increase. Depending on the rate of spread of knowledge and tools, the transition to a scarce-network-resource regime could even occur within the next two years.

Sidebar 6: Next-Generation Networks

NEXT-GENERATION NETWORKS: LIGHT PATHS AND DYNAMICALLY PROVISIONED CHANNELS

[See for example, "Networking for High-energy Physics at the LHC, H.B. Newman, on behalf of the International Committee Future Accelerators (ICFA) Standing Committee on International Connectivity (SCIC,) 2008.]

A transition began in 2005 by the major networks (GEANT2, Internet2, ESnet, SURFnet, CANARIE, SuperSINET, US LHCNet and many other NRENs) to the use of "hybrid" networks where the general purpose backbones of the major research and education networks are complemented by the use of point-to-point "light paths". Using a light path, a data transport request that has a sufficiently high priority can be given dedicated bandwidth, and service-level guarantees that often cannot be matched even in well-engineered shared networks. The use of light paths of 1 or 10 Gbps is becoming increasingly common in the high-energy physics community, as well as in astrophysics, for example, eVLBI projects. The use of dynamic provisioning of light paths in parallel with traditional network services that support general traffic has several advantages, including delivering high priority flows in time to meet deadlines, and isolating and protecting the many smaller flows from adverse impacts due to the larger flows.

Some of the major networks, notably Internet2 and US LHCNet, and SuperSINet3 in Japan, have turned to the use of optical multiplexers that use emerging standard protocols [See for example *Next-Generation Data Services Over SONET/SDH USING GFP, VCAT and LCAS* at http://www.cisco.com/en/US/netsol/ns580/ networking_solutions_white_paper0900aecd802c8630.shtml] to support dynamically switched network channels that can be dimensioned to have any bandwidth from 0.05 to 10 Gbps, in steps of 0.05 Gbps. The use of optical multiplexers has the added advantage that in case of a link failure, automatic switching to an alternate backup path can occur very rapidly and stably, so not to interrupt the data flows (especially those with priority.) The major networks are working together to ensure that compatible light paths can be provisioned, monitored and managed as required, across multiple administrative network domains in a global network environment. [See for example: http://tnc2007.terena.org/programme/presentations/show.php?pres_id=98 (on the GEANT2 JRA3 activity) and http://www.glif.is/meetings/2007/winter/controlplane/lehman-dynamic-services.pdf (from the GLIF Optical Control Plane working group).

4.2 Dependence of Petascale Facilities on High-Performance Networks

4.2.1 Finding

High-performance networks and their integration into petascale facilities (both experimental and computational) are critical to the successful realization of DOE's data-driven future. The high-performance networks required will not automatically emerge from commercial R&D. These networks will need the same level of attention as given to high-performance computing and will need to provide the following capabilities:

1. Quantified service
2. End-to-end service (application to application)
3. Workflow provisioning and management, and
4. Federated trust

4.2.2 Basis

Petascale science generates petascale data from petascale facilities (experimental, computational, distributed observations, or stored petascale databases,) which in turn requires petascale networks capable of delivering large datasets, fast-response real-time control, and interactive visualization. Indeed, a petascale experiment may require near-real-time guidance, whereby an application sends experimental data to a remote site for analysis in time for the results of that analysis to determine the configuration for the next experiment. This kind of need has already emerged within the fusion sciences community.

The development of the necessary advanced technologies, and associated products and services are quite unlikely to come from the commercial sector, at least not without government/private partnership and cost sharing. Furthermore, the marketplace will consistently under-invest in long-term research.

4.2.3 Discussion

Science makes advances through a continuous process of testing hypotheses, accepting those that pass rigorous examination and challenge, and then building on the accepted results in a continuing upward spiral of advancing scientific knowledge. The far-reaching implications of this scientific process go largely unnoticed. Nonetheless, those implications are unmistakable. As scientific experiments, both physical and computational, become increasingly data intensive with each new generation of facilities, this rigorous examination requires distributed access to ever-larger amounts of data on ever-shorter time scales.

However, with the results of a large simulation or experiment oftentimes represented as 20–30 terabytes of data (or more,) the process of data sharing and hypothesis testing takes on a qualitatively new character. One can no longer simply send a file as an e-mail attachment. As demonstrated today by the high-energy physics community, moving large amounts of data takes years of planning and development of new network techniques (e.g., hybrid connection-oriented services) and might well require installation of specialized network elements dedicated to and optimized for a particular network function. In short, the network becomes an integral part of doing high-energy physics. This statement does not imply that each DOE program office should develop and support its own network development program. It simply means that high-energy physics has envisioned a science environment that is so ambitious that it has had to pioneer new means of conducting that science. Because high-energy physics has a specific and proprietary model for sharing data among the participants, and because this model integrates fully into the science process itself, there is ample justification for internally funding the development and deployment of such a specialized high-energy network infrastructure.

The items enumerated above, however, represent much more generally applicable capabilities. Let is therefore consider them in more detail.

> ***Quantified service*** **–** The network, as a quantified and managed resource, is a complementary notion to the network as a transparent fabric, i.e., these are not mutually exclusive models. Future science applications will require both the simplicity of the transparent fabric and the control of the managed resource. As a specific example of quantified service, consider an experimental or computational dataset that occupies a known number of terabytes in a temp-store area on a supercomputer. Then, with quatified services, a user would be able to reserve S seconds of network bandwidth at B terabytes/sec, where $S*B$ equals the size of the dataset. Furthermore, the user would be able to have a realistic expectation that the data will clear the temp-store area when requested. In a more complex example, the user would be able to co-schedule both the bandwidth and the computing resources required for interactive visualization of intermediate results during a computational run. Hence, the scientist in these examples would no longer perceive the network as just a transparent fabric between a client and a server, but also as a quantifiable resource consciously designed into the application, similar to the computational processes, application data repositories, instruments and sensors.

End-to-end service – (application to application) provides for bridging the application-networking service gap. This issue exhibits itself in both the technologies and hardware architectures that link the end-system to the network (internal bus speeds, bus protocols, operating systems design, etc.) as well as the algorithms, protocols, and hardware technologies used to move the data across the wide area network. Supercomputers (and the applications running on them) often achieve poor wide area network performance, sometimes no better than that achieved by a PC. In light of high-performance I/O capacity (measurable in many hundreds of gigabits per second,) much better performance is possible. However, just as it requires detailed understanding of the inter-processor communications architecture in the design of algorithms to achieve efficient computational execution, it requires similar care in the design of I/O and networking services to achieve good performance over high-capacity networks across large geographic distances. The effectiveness of future network paradigms (such as dynamic circuits with bandwidth guarantees) will depend on engagement by ESnet and other involved network providers with the end-to-end system issues. Developing appropriate network services will require developing associated architectures that can monitor, manage, debug, and tune the network services and resources across administrative domains and through complex multi-protocol networks. Hence, DOE's plans for development must include the end-systems, not just as a value-added research topic, but also as a strategic requirement.

Workflow provisioning and management – In general, this refers to the co-scheduling of resources in such a way as to achieve the desired results following major run. This might mean not only reserving CPU node-hours but also reserving the necessary disk space and the network resources to interact with the computation (especially if the computation is being steered interactively) and handle the data as needed when the computational run is complete.

Federated trust – As DOE researchers represent a wider and wider base of workers from a more and more diverse set of organizations, the old models of granting access and control based on a user's physical presence and personal trust no longer works. Allowing faceless, possibly nameless entities to control DOE resources (e.g., circuit segments, file storage systems, and supercomputer node-hours) requires a formal scheme for building up hierarchical trust pyramids, coupled at the highest level, assigning reciprocal trust, reciprocal responsibility, and reciprocal control.

Determining the requirements and system characteristics to ensure scalability of the architecture is non-trivial. The long-term direction needs to enable an engineered real-time system that can reach a global (network and Grid-site) scale, perform non-stop, and be operated by few engineers (one of the requirements for scalability.) The real-time, all the time, and scalability requirements lead to the following special characteristics[26]:

- A services architecture that is fully distributed, with no single point of failure
- An underlying messaging fabric with high-performance
- Mutually auto-discovering registered services to make the system scalable and coherent
- Autonomous operation to provide a real-time response, reduce complexity from the user/operator point of view, and reduce the manpower load to a minimum, especially as the system scope expands

---

[26] A system with such characteristics has been in operation continuously for the past five years monitoring tens of thousands of compute nodes and >100 wide area network links, namely Caltech's MonALISA (http://monalisa.caltech.edu.) A MonALISA services infrastructure also underpins the globally scalable collaborative system EVO (http://evo.caltech.edu.)

- Ability to be enhanced in order to gather enough information to diagnose and thereafter to deal (semi-)autonomously with a growing list of complex situations
- End-host methods to profile system configuration and state in real-time, to deal with the end-to-end issues

The traditional mission of high-performance IP services will clearly not be sufficient. New technologies are evolving that can be expected to play important roles in meeting the objectives of the next decade. Theoretical research in DOE and the university research community may provide critical pieces of the puzzle. However, a disconnect exists between the theoretical research activities and the operations and engineering function that ultimately deliver new capabilities to the scientist – which is in the mission of ESnet and ASCR.

In summary, management of cyberinfrastructure resources involves integrating systems at every level and then presenting these systems as unified resources to the next higher level. The results can provide insights into how the scientific community can use the global cyber infrastructure in novel ways.

4.2.4 Recommendations

- ASCR and ESnet must go well beyond providing networking capacity and traditional networking services and enable an advanced network that provides quantified service, end-to-end service, workflow provisioning and management, and federated trust along with cyber-security.
- Effective use of these next-generation networks, with unprecedented and rapidly expanding capacity and utility, as well as complexity, will require a new paradigm of operations and management, including the following:
    - ✓ Monitoring the network from end-to-end and necessarily including the end systems. The monitoring must extend to following the progress (including problem trapping and mitigation) for individual flows, as well as for the overall state of the network and network performance.
    - ✓ Using autonomous software agents to operate, manage, diagnose, and alert by providing real-time information to network operators wherever and whenever needed.
    - ✓ Use of higher-level services to harness monitored information, to isolate and redirect problem flows as needed to ensure optimal network resource use and workflow coordination and management (as, for example, in Grid systems that coordinate the use of the network computing, experimental, and storage resources.)

Sidebar 7: Research Needs and the Future of Advanced Networks

RESEARCH NEEDS AND THE FUTURE OF ADVANCED NETWORKS

Petascale science and the growing trend to "virtualization" present many challenges to today's network technologies and services and hence present research needs for the future of advanced networks. Advanced networks must be able to support the management of the processes associated with managing data and services within the virtual environment. Workflow management tools will be necessary so that researchers are able to specify when and where an application will run, know the status of that application, and have data arrive at the right place and right time for analysis or for further computation on a specified resource. Hence, the network will necessarily become an **integral and an active** part of any petascale, workflow management system. The challenge will be to develop an **end-to-end** (application-to-application) agile, capable, affordable, and reliable network. Such a network will necessarily be much more than passive "pipes" connecting resources. It is reasonable to expect that by focusing on end to end, the network system will necessarily evolve to a service-oriented network architecture that will deliver an uninterrupted connected "path" from the application at one end to the application at the other.

Service-based application-to-application networking (or end-to-end) will need to provide network capabilities, such as error rate, bandwidth, delay, jitter, availability, and duration that are explicitly tailored (without manual intervention) to the needs of specific applications. For example, near real-time steering of experiments might use deadline scheduling with guaranteed bandwidth, so that experimental data can arrive and analyzed quickly enough that the results are useful in guiding the next experiment. As a second example, bandwidth guarantees, low latency, and reduced jitter will become necessary in order to meet the needs arising from requests for high quality video conferencing and real-time interactive data visualization.

In a scalable application-to-application network system, an end application must necessarily become "**network aware**" and capable of specifying its own identify, intentions, and network provisioning from the application I/O resource across multiple networks (WAN, MAN, LAN) and providers.

In addition, the petascale network must necessarily support an effective federated trust model that can authenticate users from institutions around the world that make up dynamic virtual collaborative communities. The incorporation of network security in the workflow architecture is necessary and critical from the outset. The network system must also include high-speed monitoring, intrusion detection, and auditing services.

4.3 Advanced Networks as Enablers of New Modes of Scientific Inquiry

4.3.1 Findings

- The co-evolution of computer and network technology and the practice of science enable valuable new modes of inquiry based on system-level non-reductionist perspectives, national- and global-scale multidisciplinary collaboration, and the collection and analysis of massive amounts of data. Research teams pursuing these new modes of inquiry rely in a fundamental way on continental and transoceanic networks with unprecedented capacity and on networking services with equally unprecedented capabilities.
- Both an advanced network and innovative networking research are critical to enabling such new modes of scientific inquiry that are of urgent importance to fulfilling DOE science and applied missions. Hence, the pursuit of DOE science will continue to advance with an accelerating dependence on networks and related technologies to unite disparate teams, to allow efficient exchange of information and applications, and to enable these new modes of scientific inquiry. The challenge is to de-

velop network architectures and service models that support virtual organizations and provide reliable, predictable, and repeatable network performance, accessibility, and security.

4.3.2 Basis

The new system-level, collaborative, and data-intensive approaches are permeating the core disciplinary science areas critical to DOE's energy and national security missions. The advancement and acceleration of these areas through the promotion of new modes of inquiry will significantly enhance the productivity of DOE scientists and impact of science.

As increasingly large collaborative teams engage in the distributed production and analysis of complex simulations and petascale data, demands grow for networks that are more powerful and more capable networking services. These demands include a need for raw end-to-end bandwidth to enable high-performance movement of large-scale experimental, observational, and simulation data, as well as derived data produced by analysis processes. Equally important is the need to address other aspects of the data life cycle, including capture, archival storage, annotation, sharing, and analysis. Thus, research on deployment of a wide range of supporting services must complement R&D on network transport.

Without such services, researchers in different fields or disciplines or in different organizations will adopt different formats, semantics, and representations of key information, or they will leave out critical quantities or descriptions. As a result, it will be difficult or impossible to combine, reconcile, or effectively use such data streams later or in a different context. Furthermore, lacking systematic archiving and curation of intermediate research results (as well as the polished and filtered publications,) data gathered at great expense is unlikely to be reused (at least not reliably) or repurposed, leading to unnecessary duplication and limited impact.

4.3.3 Discussion

Efforts aimed at enabling effective use of networks and associated networking services can, in addition to supporting individual disciplines, help break down artificial disciplinary boundaries. This latter achievement will not simply be nice to have: without it incompatible tools and structures are likely to perpetuate the isolation of scientific communities for many more years – a situation we cannot afford, given the global societal challenges that we face in the 21st century. Rising to the challenge of system-level science not only will force the breakdown of age-old disciplinary boundaries and force integration across the physical sciences, life sciences, and social sciences but also will require efforts to integrate the four modes of scientific inquiry: theoretical, experimental, computational, and data-intensive.

Network and Grid middleware will enable opportunistic and unanticipated forms of collaboration across disciplines, as well as encouraging the natural formation of new disciplines. Since we are just at the beginning of a massive paradigm shift, it is impossible to predict the outcome. Nonetheless, full globalization of the scientific enterprise will not occur until disciplinary and geographic boundaries are significantly more porous.

4.3.4 Recommendations

- The vision for networking research in SC needs to be broad enough to include the full range of distributed systems capabilities required to enable system-level, collaborative, and data-intensive science. The subpanel recommends that ASCR develop research programs that enable these emerging modes of inquiry.

- The advancement and acceleration of these areas through interdisciplinary networking research and cyberinfrastructure implementation needs to be high in the ASCR strategic plan and in its priorities for networking research and network deployments and operations.
- The vision for network and Grid services needs to include plans for developing or interfacing with capabilities to ensure the appropriate collection, curation, management, and archival for long-term access and reuse by scientists of the exponentially growing amounts of data. Such plans will be necessary for solving complex, coupled, system-level problems involving massive data collection, computation, and analysis from multiple perspectives.
- ESnet, networking, and Grid researchers should use requirements gathering workshops as well as other, more forward looking, venues to explore opportunities for cross-fertilization across SC offices, with the goal of identifying common cross-discipline requirements for enabling new but still immature modes of inquiry, such as system-level science, collaborative science, and data-intensive science. This recommendation is especially important because enabling infrastructure and applications often suffer from a chicken-and-egg problem: the infrastructure requires a diversity of successful applications for its sustained viability, while application scientists generally target fully deployed and hardened infrastructure.

4.3.5 Challenges to Implementing the Recommendations

While the challenge in achieving system-level science is often described in terms of enabling cyberinfrastructure[27], this transition is complicated by the requirement for culture shifts from reductionism, and needed modifications policies in recognition and rewards that affect scientists, policy, stakeholders (e.g., funding agencies, Congress,) and customers of scientific results. A change that might enable much enhanced data sharing and accelerated knowledge dissemination into applications might not be welcome by the scientists in all domains of science. Indeed, a simplistic “Build it, and they will come” strategy for infrastructure will likely meet limited success. Furthermore, in scientific communities where system-level or data-intensive approaches have not yet fully emerged, a requirements-driven process is not likely to produce transformative capabilities without invoking highly iterative and interactive approaches. It will be necessary to articulate a strong and compelling vision based on successes in other communities, as well as tools and components that implement fundamental elements enabling system-level approaches. The key question from scientists that will need addressing is: “What’s in it for me?”

Ensuring that all parties feel they gain more by collaborating and doing science in new ways is a delicate balancing act. It dares to envisage the creation of a powerful middleware infrastructure that supports new ways of doing collaborative science. If successful, it would enable different communities to come together and create robust, secure virtual organizations to attack new and increasingly complex problems, exploiting a wide variety of distributed resources.

---

[27] D.E. Atkins, K.K. Droegemeir, S.I. Feldman, H. Garcia-Molina, M.L. Klein, P. Messina, D.G. Messerschmitt, J.P. Ostriker, and M.H. Wright, “*Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Panel on Cyberinfrastructure,”* 2003.

Sidebar 8: System-Level Science

SYSTEM-LEVEL SCIENCE

Foster and Kesselman [See I. Foster and C. Kesselman, "Scaling Systems-Level Science: Scientific Exploration and IT Implications," *IEEE Computer*, November, 32-39. 2006] define system-level science as the "integration of diverse sources of knowledge about the constituent parts of a complex system with the goal of obtaining an understanding of the system's properties." Providing for the required integration mentioned in this definition raises considerable challenges. Meeting these challenges requires implementation of many new elements to an infrastructure system that would enable scientists to overcome barriers of distance, interdisciplinary discovery and collaboration, community and resource management, data and model publication, continuous infrastructure adaptability, metadata and data provenance, and translation and standardization of data formats. Many of these elements are strong drivers of networking science that can deliver the concepts and research implementations that inspire participation in the implementation of system-level science.

The powerful transformative impact that system-level science is expected to produce makes it an imperative consideration for a mission agency such as the DOE and for the ASCR vision of enabling DOE science through computational and networking research and development. An excellent example is the recent successes in climate research enabled by unprecedented international collaborations of governments and scientists, which has now led to this year's Nobel peace prize. Indeed, the Earth System Grid [See http://www.earthsystemgrid.org/, and D. Bernholdt et al., "The Earth System Grid: Supporting the Next Generation of Climate Modeling Research," in *Proceedings of the IEEE*, 93 (3), 485-495. 2005] made the climate-model simulation data, which underpinned the IPCC analysis, available to the international community. More generally, there is broad recognition that the emergence of system-level science "can enhance the discovery of knowledge gaps and assumed knowledge that "just isn't so," clarify research priorities, and potentially accelerate scientific impact on industrial development, economic competitiveness, and societal needs." [See Karen Schuchardt et al., "Portal-based Knowledge Environment for Collaborative Science," *Concurrency and Computation: Practice and Experience*, 19(12), 1703-1716, 2007.]

For these types of problems, the time-honored reductionist, or subsystems, approach, in which key phenomena are isolated and analyzed in depth, is approaching a status of diminishing returns. The approach for the future must be systems based, in which simulations are developed in the context of encoding all known relevant physical laws with engineering practices, production, use, distribution, and environmental factors. Even with petascale scientific facilities, given the constraint of the current conduct of science, the added remarkable capability will enable us to do "old science" very well, but will not enable the new science that we need. [*See Networking and Information Technology Research and Development: Scientific and Technical Aspects*, Dave Nelson, Director National Coordination Office for Information Technology Research and Development, October 2003.]

Despite the exponentially growing knowledge base, we would be wise to recognize that the following quote remains true today:

"The wiser you are, the more you believe in equality, because the difference between what the most and the least learned people know is inexpressibly trivial in relation to all that is unknown."

Albert Einstein

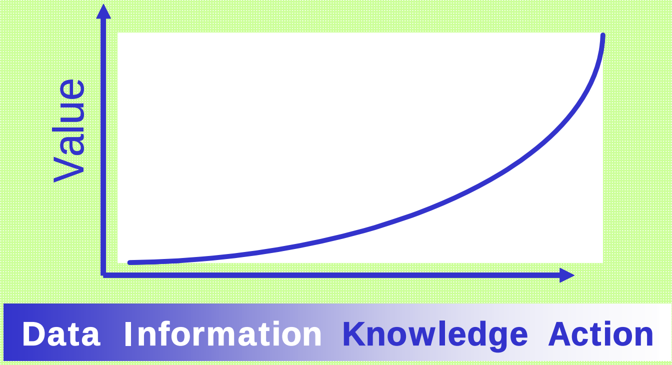
Sidebar 9: Data Challenges and Issues

DATA CHALLENGES AND ISSUES

Large instruments such as advanced light-sources and neutron facilities are producing huge streams of data. In addition, large streams of data are arising from the growing numbers of ubiquitous arrays of small sensors and from capacity, capability, and Grid computing. Currently, preservation of much of this scientific data occurs in ad hoc and fragmented ways. All too often, the data ends up in "data mortuaries" rather than archives.

Metadata will be vital for storing and preserving scientific data. Such metadata not only must contain information that annotates the data with semantic tags, but also must provide information about its provenance and its associated user access controls. In order to construct intelligent search engines to facilitate automated discovery, each separate community and discipline will need to come together to define generally accepted metadata standards for their community Data Grids. It will be advantageous, possibly critical, that such standards are interoperable across disciplines.

With the imminent data deluge, the issue of how to handle the vast outpouring of scientific data becomes of paramount importance. Up to now, it has been feasible manually to examine the experimental or computational data, in order to identify potentially interesting features and discover significant relationships between them. Such a mode of inquiry is not scalable. In the future, in view of the massive amounts of data created by simulations, experiments, and sensors, perpetuating such a mode of inquiry will no longer be practical. Hence, it will be necessary to automate the discovery process, which proceeds from data to information to knowledge to action, to the greatest degree feasible. At the lowest level, this requires automation of data management with the storage and organization of digital entities. At the next level is the need to move to automatic information management, which will require automatic annotation of scientific data with metadata describing interesting features both of the data and of its storage and organization. At the next higher level is the need to progress beyond structure information toward automated knowledge management of our scientific data, which will include the expression of relationships between information tags as well as information about the storage and organization of such relationships.



The scientific digital libraries, created by either highly collaborative experiments or simulations, will come to require the same sort of facilities as conventional electronic libraries: including a set of services for manipulation, management, discovery, and presentation. In addition, these scientific digital libraries will require new types of tools for data transformation, visualization, and discovery (such as data mining.) The community will also need to solve the problem of the long-term curation of such data and derived information and knowledge. Generating the data is one thing, moving the data is another, but preserving the data in a form usable by scientists (or nonscientists) other than those who were involved in its creation is fundamentally a different issue for the scientific community.

4.4 Petascale Facilities as a System of Systems

4.4.1 Findings

- The integration of networking services into SC petascale facilities (computing and experimental) to enable an e-science environment needed to address 21st-century challenges results in a System of Systems.
- The complex interdependency of ESnet, networking research, DOE science programs, and facilities poses some unique management and development challenges, because that interdependency represents such a System of Systems.
- The data access, integration, workflow, and federation capabilities of the next generation of network and Grid middleware will necessarily play a key role in enabling both the virtualization and the realization of petascale science with novel, emergent, and transformational characteristics.
- A key movement that will have profound impact on science, society, and the economy in the coming decade will be the development of national and global-scale, data-intensive (terascale to petascale and beyond) distributed cyber environments. The challenge will be to develop dynamic and intelligent (cyber) resource allocation architectures that will allow the DOE scientist to transparently and easily make use of resources at this scale.

4.4.2 Basis

The high-end computing community is growing; usage is also growing, and an increasing number of users prefer not to be constrained to a specific machine. Most users want to interact with a set of autonomous service providers (leadership class and capacity computing centers as well as large one-of-a-kind experimental facilities) as a virtual facility that allows ease of navigation and enables resources to be treated as building blocks. The data generated and stored must be accessed by or distributed among collaborating centers worldwide. This means that distributed software infrastructure must become easier to deploy and manage in such a System of Systems. The heterogeneity and autonomy in this System of Systems must be leveraged. This process requires agreement about services, interoperability, local control, and central coordination.

While it has not been stated this way before, the mission of the network and networking research could reasonably be considered as one of ***enabling SC's extraordinary collection of facilities and petascale science as an even more extraordinary System of Systems***. In recognizing the DOE science as such, one then invokes more "systems think" in which one focuses on the whole, not on the parts, of a complex system. One concentrates on the interfaces and boundaries of components, on their connections and topology of arrangement. However, then one then exposes all the interfaces for which explicit responsibility tends to stop short of the boundary, thereby creating a gap. As a subsystem, the network tends to handle "bulk transport" very effectively, but in the system, it does not necessarily handle the boundary layer effects very well (cyber security is a good example.) Generically interface-like issues are not unique to SC and are symptomatic of a System of Systems, in general.

Hence, ESnet has traditionally provided the high-performance networking "plumbing" among the DOE laboratories and increasingly to the collaborative research community in the universities and associate research institutions. Moreover, facilities such as Ultra Science Net are developing advanced hybrid network technologies. However, the notion of the overall integration of cyber resources (e.g., network, computation, data storage, instruments) into a set of interoperable and extensible building blocks, as well as the upper layer services and agents to manage these resources effectively is critical to moving the science programs forward. Turning raw data into knowledge that can be accessed ubiquitously, can be replicated for performance or security, and can then be integrated with other distributed systems will drive many of the software and hardware technologies required in the network. This is a System of Systems and any approach needs active collaboration among networking, computer, computational, and domain scientists.

4.4.3 Discussion:

Many science disciplines, sub-disciplines, programs, and projects are relevant to DOE's missions. The science processes have evolved to the point where the results of one discipline or sub-discipline are the basis for models and experiments in another discipline (e.g., the relationship between ocean modeling and atmospheric or climate modeling.) No longer are paper publications the only product of science. The datasets themselves constitute a product, and the model that generated the results becomes another tool for other researchers. This integration process of melding one model into the science process of another project occurs in a largely ad hoc fashion today, where every additional inclusion requires substantial manpower to integrate and to validate in a way that allows the scientist to produce reliable and rigorous results.

From a more network-centric aspect, the emerging notion that the network is itself a cyber resource, which calls for consideration in a manner much like other computational, storage, and instrument resources, explicitly highlights the necessity to understand how the network allows these other resources to interweave more intimately. The network is not just a transparent substrate on which these other resources and processes rely – it is itself an essential resource to design into the science processes. In one sense, this idea breaks the established concept of computational models developed independently from the data repository architectures developed independently of the instruments. Like the science processes mentioned above, these cyber resources must interoperate seamlessly with each other and the science processes that wish to use them.

To break the ad hoc and time-consuming nature of mating interdisciplinary systems and heterogeneous resources, it is necessary to approach the issue from a System of Systems perspective. The goal is new science – and accelerating the pace of new science - made possible by an integrated system: the (previously) separate systems are the components. An analysis of the interfaces between these entities, together with a formalized means for standardizing such interfaces, would provide a number of important benefits. Such benefits include allowing science teams to publish data and tools in a fashion that would make it possible (if not simple) for any scientist to stand on the proverbial shoulders of the scientists who have paved ground ahead of them – in ways that have never before been practical.

4.4.4 Recommendations

- ASCR should develop a Systems of Systems view of petascale science, focusing on the whole, not just the parts, of this complex system, thereby creating a concentration on the interfaces and boundaries of components and on their connections and arrangement.
- The networking research program should seek cross-fertilization with the mathematics program, especially with respect to the science of networks and science of System of Systems, to advance the understanding of the complexities of emergent behavior in the science Grids and the System of Systems.

4.5 Research and the "Valley of Death"

4.5.1 Findings

- Theoretical and fundamental networking research is often published and then forgotten. There is no sustained process to nurture promising concepts into experimental pilots and then mature them into deployment and operational services.
- The dead zone, where there is a notable lack of funding between early networking research results and when those results are rediscovered or reinvented and developed into new operational network capabilities is called the "valley of death". This valley of death can delay introduction of important new technology into the cyberinfrastructure and hence could retard science progress by many years.
- There are many barriers to the development and adoption of promising new networking research ideas (across all agencies and industry.) These range from too narrow a focus (e.g., driven by already evident deficiencies,) to the classic valley of death. The challenge is to establish mechanisms to identify promising networking research concepts and move them progressively through prototyping, experimental deployments, standards, and ultimately into a production network service environment.
- Out of necessity, a few leading-edge science communities have pushed the state of the art in networks and networking services. The result has been vast discrepancies in the level of capabilities for data distribution and management available across the scientific enterprise. The challenge is twofold: first, to leverage and generalize the services developed in one context for the broader utility of the DOE science community and second, to develop advanced cyberinfrastructure service architectures that allow future efforts to create, easily and effectively, the types of network (or other cyberinfrastructure) that they need.

4.5.2 Basis

All too often exploration of advanced research in network protocols or technologies occurs with broad interest and fanfare but the impact of the research languishes on one side of the valley of death while the research funds dry up or interest wanes. The reasoning is often that the research is complete with published papers to cite and that adoption into standards or new products and services should be the responsibility of either the private sector or the operations and engineering functions of production networks. Unfortunately, the needs of the private sector are generally different from (or several years behind) that of the advanced science and research community. Hence, the utility and importance of the research results go unrecognized or is not sufficiently mature to be releasable into the open-production network services environment. Therefore, the research sits, often for years, perhaps until an emerging market requirement results in rediscovery or reinvention.

In another dimension, research may progress on behalf of one segment of the community, but the results do not develop in a way that brings the benefits to a larger community of potential users. This may be an unintentional result of specific groups simply focused on specific and narrower objectives. Nevertheless, the result often is chasm-like differences in technical capabilities between well-funded high-priority science programs and smaller programs or projects.

4.5.3 Discussion:

The so-called valley of death is, in reality, an issue in the high-level mission (and funding) process. Basic research sponsoring organizations typically have neither the mandate nor the budget to continue to fund a promising research concept beyond the laboratory. This shortcoming means that much good research gets written up and then sits on a shelf for five to ten years until some emerging market drives an industrious vendor or developer to dust off the paper (or worse, to recreate the idea) in an attempt to solve a problem that has appeared in the commercial market. The result is that advanced research and science users do not benefit from the relevant research until years later when the commercial market has finally matured and vendors are shipping production-level software and hardware that ESnet can then purchase and deploy as a production-quality service for the DOE science community.

The subpanel believes that the research community can provide enough market influence and sufficient insight into future information processing and dissemination that with sustained life-cycle funding[28] the DOE science community would benefit from advances in networking technologies likely five to ten years sooner than with traditional technology commercialization processes. Ignoring “valley of death” issues will slow the pace of advancement of science; overcoming such issues will accelerate the rate of scientific discoveries and the rate at which those discoveries influence solutions to global societal challenges.

4.5.4 Recommendation

ASCR should develop strategies to overcome the various chasm-like and valley of death issues. In particular, ASCR needs to establish processes to review networking research results, as well as to select and fund promising capabilities for further development, with the express intent to accelerate the availability of new capabilities for the science community. Research products should include, where appropriate, participation in the applicable standardization organization or process. Acceptable end products of network research should include standards documentation, such as an IETF RFC.

Sidebar 10: The Grid

THE GRID

The vision for a layer of Grid middleware that provides a set of core services to enable such new types of science and engineering is due to Ian Foster, Carl Kesselman, and Stephen Tuecke [See I. Foster, et al., 2001.] Within the Globus project, they have developed parts of a prototype open source Grid Toolkit [See I. Foster and C. Kesselman, 1997.] Their choice of the name “Grid” to describe this middleware infrastructure resonates with the idea of a future in which computing resources, compute cycles, and storage, as well as expensive scientific facilities and software, is accessible on demand, like the electric power utilities today. Grid technologies are now extensively available on a wide range of campus, regional, national, and international sites that provide infrastructure and services to numerous scientific projects.

Science Grids (which are themselves an instantiation of a System of Systems) are enabled by cyber processes (network and Grid services) that incorporate many facilities, typically distributed around the world – and of many different types – computational facilities, storage facilities, instruments that generate unprecedented quantities of raw data, and visualization facilities.

In addition, the intent in the use of the terminology “Grids” was to invoke the vision of “utility grids” in which the service is readily available with no prior expertise – just plug into the outlet. However, the future electrical grid, is likely to take on more characteristics of e-science Grids or a System of Systems, including much more distributed control and distributed decision-making.

---

[28] “Life-cycle” here implies from basic research through applied research and development through to experimental pilot deployments and, finally, production deployment.

Sidebar 11: System of Systems

SYSTEM OF SYSTEMS

The health and vitality of U.S. science and technology depend on the availability of the most advanced research facilities. SC builds and operates the world's finest collection of scientific facilities, which each year host more than 20,000 researchers and students from universities, private industry, and other government agencies. These large and exceedingly complex one-of-a-kind facilities have been critical enablers in many of the most important scientific discoveries over the past six decades. The distributed facilities are open to a wide range of researchers (on a peer-reviewed basis) and shared with the science community worldwide. Without question, investment in these facilities yields extraordinary scientific breakthroughs and provides vital societal and economic benefits.

These user facilities provide resources … that speed up experiments by orders of magnitude and open up otherwise inaccessible facets of nature to scientific inquiry. Many of the important discoveries made in the physical sciences in the second half of the twentieth century were made at – or were made possible by – user facilities.

Dr. Hermann A. Grunder, former director of Argonne National Laboratory, in Congressional testimony, July 2003

It is more than an academic exercise to recognize that the system of facilities operates with the characteristics of a System of Systems. *[See for example: Carlock and Fenton, "System-of-Systems (SoS) Enterprise Systems for Information-Intensive Organizations," Systems Engineering, Vol 4, No. 4 (2001) pp. 242-261]* Indeed, the emergence of a new paradigm of scientific application, i.e., the e-science or Grid application – is an instantiation of the abstract concept of a System of Systems. Such a system typically exhibits the behaviors of complex adaptive systems, including operational independence of elements, managerial independence of elements, evolutionary development, geographical distribution, interdisciplinary action, heterogeneity of systems, a system of networks, and typically emergent behavior.

System of Systems science is an emerging discipline not yet claimed by any organization or agency, although, it is a science that can provide a critical organizing principle. The elements of SC's System of Systems include its computing systems and network; its experimental facilities, instruments, and user centers; observational networks; data, information, and visualization resources; virtual organizations and distributed collaborations; ASCR's research portfolio; relevant research portfolios in the other program offices; and researchers, facility staff, center staff, and program managers. This System of Systems produces large amounts of heterogeneous, geographically dispersed data, with complex relationships and couplings between experiment, computation, data, and people. The sought after emergent behavior, so hard to predict and often surprising in nature, is the scientific breakthroughs and revolutionary advancements that change the way people think and do things.

The network is a sub-system that forms one element of the System of Systems. The network knits elements together so that the whole can function as a system, despite the distributed control. It is the whole (rather than any single subsystem) that "owns" the goal of making frontier science possible, as well as DOE mission-oriented basic and applied research.

4.6 Need for a Coherent Networking Research Program

4.6.1 Findings

- In the coming decade, much of the network technology necessary to meet DOE (and other science-based agency) e-science requirements will not arise naturally from commercial R&D. The challenge then, given the central importance of high-performance networking to SC programs, is to construct a long-term and far-reaching network R&D program that sustains innovation from basic research to prototype to early deployment and culminates in a production network environment for use in advancing science.
- The network capacity and service capabilities anticipated for the next decade of science activities will likely be three to four orders of magnitude greater than current network architectures and technologies can effectively address today. The challenge will be to accelerate or develop a ten-year technology trajectory to achieve this projected need.
- The advancement of networks and their services couple intimately to the capabilities they support, such as data management, and to capabilities on which they depend, such as a secure open science environment. Thus, research programs in all distributed, multi-domain science must include networks and network services. Currently, however, ASCR has neither a coherent networking research program nor a shared vision for the R&D required for achieving successful petascale connectivity and the network services necessary to achieve system-level science.

4.6.2 Basis

Relying primarily on vendor R&D is appropriate with respect to directed research on fundamental optical science or components. However, commercial network development will not produce the networking services and technology needed to achieve DOE's petascale science objectives. The reason is that commercial networks and network elements (switches and routers) are optimized for the most cost-effective aggregation possible of ever-larger numbers of proportionately ever-smaller flows (web views, Google searches, e-mail, WEB 2.0 applications, music downloads, etc.) In contrast, DOE's mission (science and application) requires being able to handle a wide range of data types and flows. For example, proportionately ever-smaller numbers of ever-larger flows culminating in the need to handle single point-to-point flows at the petascale. DOE's mission may also require the handling of an exceedingly large number of very small data flows, such as that coming from sensor data. These two examples demonstrate the wide range of data types, which must be supported long term within the DOE network envirnomnet.

The situation is analogous to that in high-performance computing. Commercial applications (from Google to commercial banking) rely on running hundreds or thousands of parallel independent applications on massively parallel systems. With the exception of embarrassingly parallel applications, using commercial cluster systems would scale miserably if they attempted to run, for example, one of DOE's large combustion, materials science, or climate codes. Recognizing this divergence from commercial drivers, DOE has invested significant resources to ensure that the leadership-class architectures will be useful. DOE has also invested significant resources to develop the algorithms and software needed to take full advantage of these unique capabilities.

Petascale science implies petabyte file transport, and simple arithmetic shows that moving petabytes requires multi-100-gigabit circuits as well as the network interfaces and software to move the data from the end-system onto the network. Guiding the development of all of these components so that they will be useful to DOE will require close collaboration with network vendors.

The revolution in optical technology that led to the development of the first DWDM (dense wave-division multiplexing) optical transport systems 10 years ago is continuing to accelerate. Explosive development is occurring in such areas as the following:

- IP over DWDM (in other words, building DWDM technology directly into the router or switch ports and then allowing those same ports to tune themselves to a particular color as determined by the needs of the moment)
- Multi-way, electronically reconfigurable add/drop systems that allow arbitrary selection and steering of circuits through a network mesh
- Application of multi-level modulation schemes to optical signals, thus increasing the data rate without increasing clock speeds
- Integration of control of the DWDM elements of a network into the same framework used today to control the peering of network domains

Collectively, these developments will allow protocols like GMPLS (Generalized Multi-protocol Label Switching,) BGP (Border Gateway Protocol,) or OSPF (Open Shortest Path First) to be aware of the optical state of the network (rising applications of forward error correction on a link for example, or approaching amplifier saturation.) A systematic overall knowledge of the state of the network, combined with knowledge of the physics of the network elements, will enable algorithmic network management – proactive hitless switching of circuits from a path that is becoming marginal to one with excess capacity, for example.

In addition to advances in network technologies, a second driver is the need to support the emerging paradigm for conducting petascale science. End-to-end services – including the ability to schedule, reserve, and provision network resources as determined by a particular application – is necessary to support researcher collaboration and system interoperability in multi-disciplinary petascale science. Because science applications communicate with end-to-end network services, these services depend on advancements at all network layers. The services must be reliable, robust, scalable, and readily invoked by the science application.

Realizing the benefits of these developments will require significant DOE to investment – on par with its applied mathematics and computer science efforts – in developing middleware and integrative network technologies that can take advantage of these advances. Again, it will not be sufficient to assume industry will provide appropriate solutions: in fact, industry predictably will not. Moreover, solutions are not likely to emerge from NSF, which generally focuses on academic-scale (single professor plus small number of graduate students and perhaps a postdoctoral researcher) efforts in protocol, security, and similar software development. Only DOE and DOE's science facilities have the complete end-to-end control of network infrastructure necessary to leverage fully these advances. An optical status-aware GMPLS is only marginally useful if its control does not extend all the way to the end devices and applications.

Networks are becoming more complex, not simpler. The underlying transport medium, at one time no more than a passive piece of copper wire, is now an elegant, complex, highly technical part of the system. Future data-management architectures for a cross-disciplinary, distributed science environment will require specially tailored advanced network services. However, since such a network must support a secure, open science environment, networking research may be required to ensure cyber security.

4.6.3 Discussion

Given the new portfolio of hybrid networking services (conventional packet switched services melded with dynamic circuit provisioning) being deployed in the global science community, one must try to divine how these capabilities will affect the way DOE e-science environments are realized in 2018. The notion that traditionally high-end and parsimoniously rationed resources could become plentiful and easily manipulated by individuals or small groups leads to a view of the future where applications are no longer simply particular codes that run on a particular supercomputer at a particular data center. Rather, construction of these applications will consist of combining large-scale IT components, active science cyber agents distributed across the globe, and other resources in ways that previously took months or years to configure into a workable and productive e-science workflow environment. Scientists will be able to envision, construct, test, and refine massive (by today's standards) applications quickly and easily. The global resource pool, coupled with sophisticated integrative software, will define the computer rather than a disparate (albeit high-end) set of IT facilities that require extensive network engineering to stitch them together into a whole cloth. The previously mentioned sophisticated integrative software will transform the network from a transparent fabric into an explicitly managed and dedicated network resource. This is a key area of research and development: How can we provide the science user with complete control to acquire and construct the e-science environment necessary to make science breakthroughs and do so with ease and security?

Providing scalable distributed-system architecture scalable applications is vital. This is not something easily done by network provider organizations, networking researchers, or discipline researchers on their own. It will require a synergy and likely a new (possibly SciDAC-like) form of development and deployment effort. Developing these network technologies in an isolated networking lab will no longer be a viable path forward. They will have to evolve out of cross-disciplinary applied research, where development of the networking component arises from close interaction with the applications as a whole.

Besides the high-level application perspective, one must consider the potential for e-science environments that dwarf present or near-future telecommunications technologies. In radio astronomy, for example (see the Networks 101 sidebar,) a scientist might envision a continental scale instrument or set of instruments that generates in excess of 200 Terabits/second (2×10^{14} bits/second) on a 24×7×365 basis. No present-day network, technology, or architecture can scale to support an application of this type or scope. This circumstance is not unique to radio astronomy; abating climate change, defending against unconventional warfare abroad and at home, and ensuring global and ecological sustainability are all likely to present similar challenges. An analysis of expected evolution of current network technologies (i.e., the prospects for 100-gigabit transmission link, line rate switching and forwarding, etc.) indicates that our present telecommunications and networking technology trajectory is perhaps as much as three orders of magnitude too timid to address this type of science need. Optical and photonic telecommunications technologies will undoubtedly play a critical role in the quest for ever-greater network performance, but these must be pursued explicitly so as to ensure that the interests of high-end e-science are served and addressed in a timely manner.

For transmission over these highly distributed environments, cyber security research is an overarching and critical need for DOE. The report of the Cyber Security Research Needs for Open Science Workshop, July 23-24, 2007, presents many key research topics, including development of an open science security architecture that provides secure software, end-to-end data security, secure information management, resiliency, monitoring, detection and responses, situational awareness, federated trust and user-friendly implementation and use. In addition, security of the network itself is a concern, especially for new hybrid networks presenting potential control system vulnerabilities.

DOE will not be able to rely on market forces alone to create the types of technologies it needs in the timeframe e-science demands.

4.6.4 Recommendations

- The networking research agenda must include not only basic research to formulate rigorous foundations for new network concepts but also applied research where promising concepts can be deployed into experimental, pilot, or testbed-like environments. Such experimental deployments will allow the science community to provide critical feedback. After refinement, new capabilities can move into the production environment to serve the broadest possible community. This iterative process will allow the DOE science community to drive the network evolution to their needs. Hence, the networking research agenda must embrace a strategic goal to create solutions to the above-mentioned “valley of death” network system issues. Addressing such networking research objectives will require a long-term vision for fulfilling the potential of good ideas as well as a long-term commitment to a thorough life-cycle perspective, including involvement in the development of standards.
- Formulation of the research agenda must have science applications in mind while considering the overarching enabling technologies, including cyber security.
- ASCR needs to develop a strategic vision for the network and for networking research (including Grid research) and to involve a wide range of stakeholders in order to gain acceptance or “buy-in.” ASCR, ESnet, the science application communities, and the networking research community should work together to develop the vision and ensure that it complements work by other agencies (such as NSF,) and commercial investments.
- ASCR’s networking objectives should go well beyond the most easily envisioned needs of the science community. If the science community can envision a challenging networked environment, then the challenge to the networking community should be to design the network architecture and technology to support 10, 100, or even 1,000 such environments simultaneously, without undue labor or dollar cost. The network architecture and services developed should explicitly allow the scientist to focus on the science and the mission needs rather than on the cyberinfrastructure, and in a way that accelerates time to solution and time to societal impact.
- ASCR should convene an external committee to review this networking research program on a regular basis, in order to maintain a ten-year research horizon and the required complementary attributes.

4.6.5 Challenges to Implementing the Recommendation

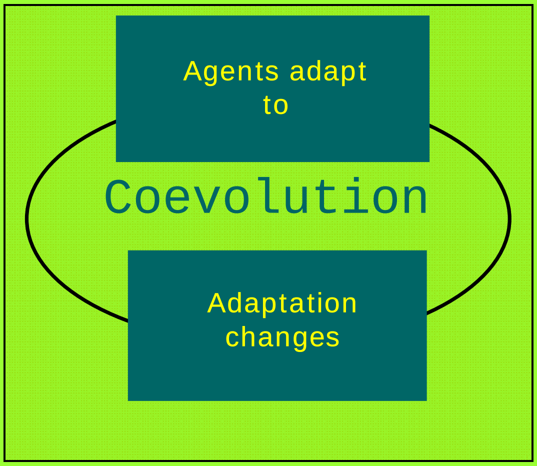
- If the organization of the research portfolio tilts too heavily toward the domain scientists’ immediate needs, the focus overemphasizes procurement of existing technologies, and computer scientists become viewed merely as consultants and implementers. If the weight shifts too heavily toward computer science, the focus is likely to miss important future needs of users, or to shift efforts too heavily toward creating new technologies with insufficient attention to stability and user support.
- Producing and maintaining widely usable, reliable software is at least one – possibly several –orders of magnitude more difficult than generating an initial high-quality prototype.
- A deliberate strategy is essential to developing the right teams so that the results ultimately serve many disciplines with broadly useful tools and methods, and potentially with an underlying services and communications fabric that would have an extensive, transformational impact on research. Such

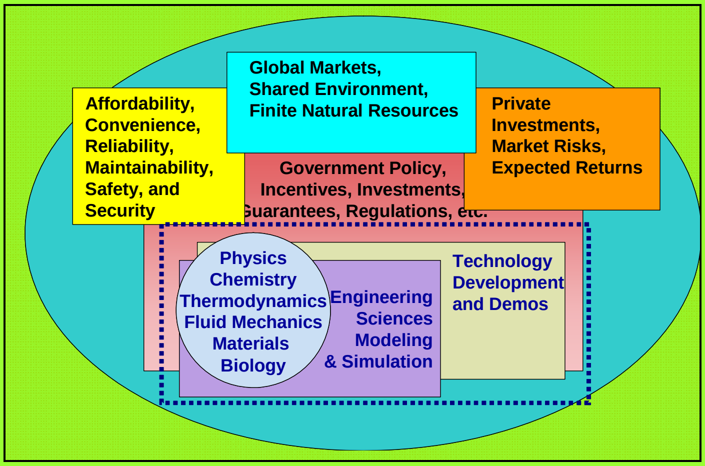
a strategy might require technically adept liaison sub-communities formed within each discipline, to ensure the overall effectiveness and impact.

Sidebar 12: Complex and Complex Adaptive Systems

COMPLEX AND COMPLEX ADAPTIVE SYSTEMS

Addressing 21st–century global societal issues will depend on how well and how fast researchers can move beyond reductionist approaches – which view science, engineering, technology, economics, business, markets, and policy in isolation – to embrace entire systems, which may be complex and adaptive. (i.e., in which the individual components change their rules based on experience.) Complex systems generally have the following characteristics: multiple interactions between many components; nonlinear relationships; experimental domain is large; underlying model is typically unknown; no analytic formula for the response surface; oftentimes patterns are what matter; and the whole is not the sum of the parts. Increasingly, a system-level perspective is relevant to DOE’s mission regarding global economic, energy, and environmental challenges, increasing interconnectedness, increasing social and economic inequities, the spread of capitalism and democracy, and ethnic and religious conflict. As we move into system-level science, the traditional expectations of achieving detailed understanding that can be validated and lead to predictive science must yield to explanatory and exploratory science, since a hallmark of complex systems is emergent behavior and emergent behavior is typically surprising. Prediction most likely becomes impossible, except perhaps in a statistical sense. The goal becomes *improved* decision-making, which does not require full predictability.





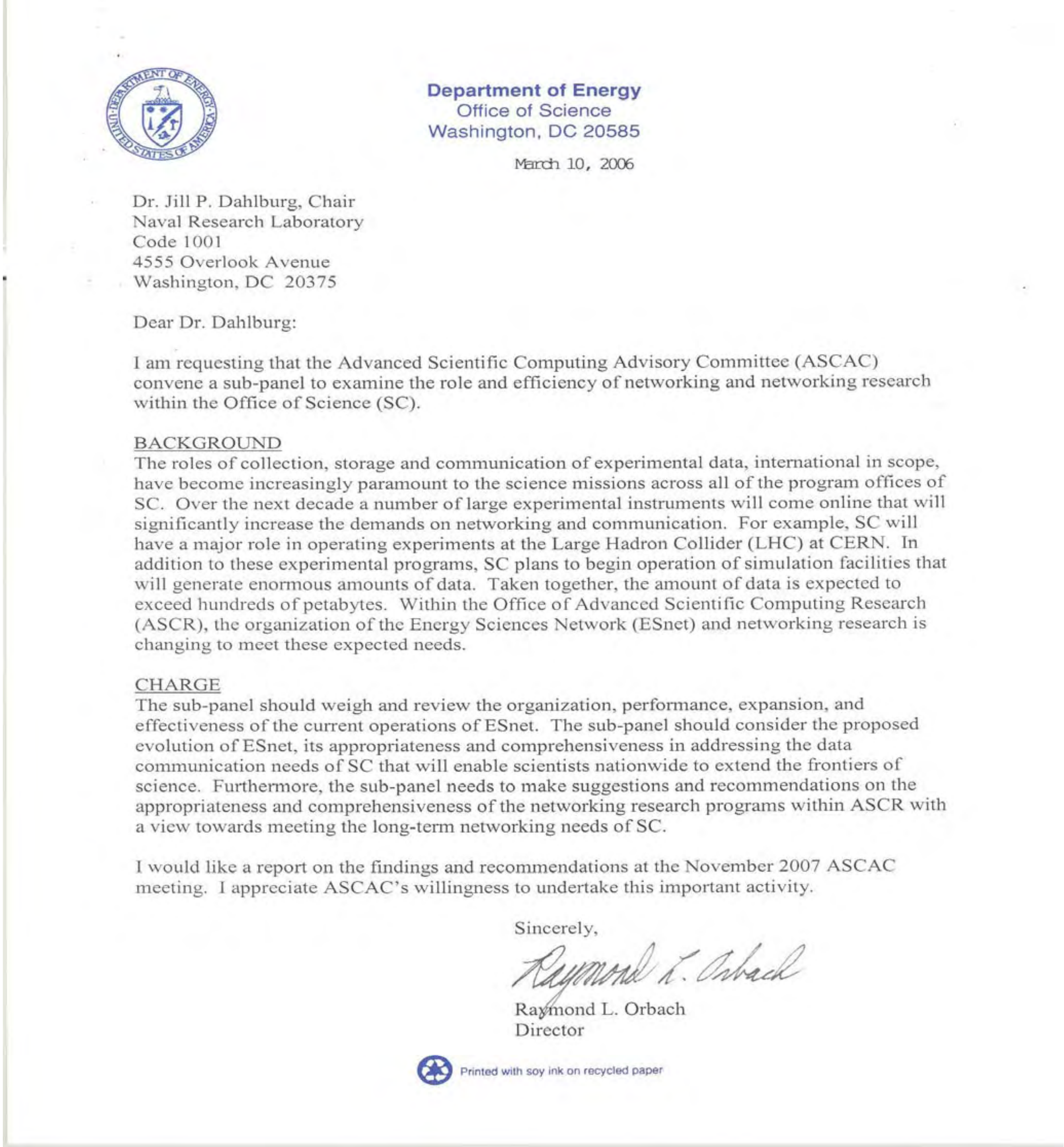
5 CLOSING REMARKS

This document reports the findings and recommendations of a subpanel convened to review the effectiveness of the DOE's existing network and networking research strategy in support of extending the frontiers of science. The network technologies and services considered will require substantial and sustained research and development. Indeed, the subpanel saw a need to reinvigorate an aggressive, sustainable, long-term, and strategically focused networking research program to create network-specific technologies that will allow DOE not just to increase the speed of existing systems but also to transform the manner in which science is done, including enabling system-level science and data-intensive science. Moreover, this will further require that ASCR formulate a deliberate strategy to bridge the "valley of death" for a networking research program to provide critical enabling technologies to advance the frontiers of DOE science. Such a strategy would move concepts (whether from within ASCR's basic networking research program or from the networking research community more generally) through testbed deployment to production. Any such strategy should involve increasing collaboration between the networking research scientists and operational/engineering facilities' personnel as concepts mature. Moreover, the application science community, as end-users, must be a continual and integral component of those collaborations.

The report discussed issues arising from the System of Systems aspects of DOE science, facilities, and programs and the intimate relationship between data mobility through the network and data management. Indeed the subpanel recommended that ASCR integrate research in data collection, archiving, curation, generation, pedigree, and access into the networking research program. Effective integration of new networking technologies into the future data-management architecture will be crucial to providing timely and secure access to and efficient migration or access to large datasets across an emerging cross-disciplinary globally, distributed science environment.

APPENDICES

A. The Charge Letter



Department of Energy
Office of Science
Washington, DC 20585

March 10, 2006

Dr. Jill P. Dahlburg, Chair
Naval Research Laboratory
Code 1001
4555 Overlook Avenue
Washington, DC 20375

Dear Dr. Dahlburg:

I am requesting that the Advanced Scientific Computing Advisory Committee (ASCAC) convene a sub-panel to examine the role and efficiency of networking and networking research within the Office of Science (SC).

BACKGROUND
The roles of collection, storage and communication of experimental data, international in scope, have become increasingly paramount to the science missions across all of the program offices of SC. Over the next decade a number of large experimental instruments will come online that will significantly increase the demands on networking and communication. For example, SC will have a major role in operating experiments at the Large Hadron Collider (LHC) at CERN. In addition to these experimental programs, SC plans to begin operation of simulation facilities that will generate enormous amounts of data. Taken together, the amount of data is expected to exceed hundreds of petabytes. Within the Office of Advanced Scientific Computing Research (ASCR), the organization of the Energy Sciences Network (ESnet) and networking research is changing to meet these expected needs.

CHARGE
The sub-panel should weigh and review the organization, performance, expansion, and effectiveness of the current operations of ESnet. The sub-panel should consider the proposed evolution of ESnet, its appropriateness and comprehensiveness in addressing the data communication needs of SC that will enable scientists nationwide to extend the frontiers of science. Furthermore, the sub-panel needs to make suggestions and recommendations on the appropriateness and comprehensiveness of the networking research programs within ASCR with a view towards meeting the long-term networking needs of SC.

I would like a report on the findings and recommendations at the November 2007 ASCAC meeting. I appreciate ASCAC's willingness to undertake this important activity.

Sincerely,

Raymond L. Orbach

Raymond L. Orbach
Director

Printed with soy ink on recycled paper

B. Acknowledgments

The subpanel is grateful for the active participation and contributions by Bill Johnston, Eli Dart, and the ESnet staff, Charlie Catlett, Richard Mount, Bruce Gibbard, Ray McCord, Michael Strayer, Walt Polanski, Daniel Hitchcock, Robert Lindsay, Nancy White, and Christine Chalk, during the subpanel's deliberations. The subpanel is also indebted to Gail Pieper and Andrea Heacock-Reyes for their invaluable assistance in preparing the final report and to Nicole Blaisdell for the artwork on the back cover.

C. Subpanel Biographies

Dr. F. Ronald Bailey is a senior consultant at Advanced Management Technology, Inc. Dr. Bailey also holds the position of principal at HPC Consulting. Prior to these positions, he was president and CEO, AppsPoint Corp.; program director, Raytheon Corp.; and consulting professor, Stanford University. He also held several management positions with NASA Ames Research Center. Dr. Bailey has served on DOE SciDAC Cross-cut Review Panel, 2006; DOE Office of Science Advanced Scientific Computing Advisory Committee ,2005-present; Pacific Northwest National Laboratory Advisory Committee, 2005-present; chair, PNNL Computational and Information Sciences Directorate Review Committee, 2005-present; chair, DOE Office of Science, Cray X1 Review Panel, 2004; NSF PACI Review Panel, 1998; NSF PACI Proposal Review Panel, 1996; chair, DOE High-performance Computing Research Centers Advisory Committee, 1992-1995; Program Committee, International Symposium on Computational Fluid Dynamics, 1994-1996; International Committee, Office National d'Etudes et Recherches Aerospatiales publication: “La Recherche Aerospatiale,” 1990-1994; DARPA Submarine Hydrodynamics and Hydroacoustics Technology Center Advisory Board, 1990-1991; NSF Supercomputer Center Renew Committee, 1989; Federal Coordinating Committee on Science Engineering and Technology, Subcommittee on Computer Networking, Infrastructure and Digital Communications, 1988-1989; IEEE/ACM Supercomputing Conference Steering Committee, 1988-1994, chair 1990; general chair, IEEE/ACM Supercomputing 1989 Conference and External Advisory Board, Supercomputer Computations Research Institute, Florida State University, 1987-1989. Dr. Bailey earned his doctorate at Iowa State University in aerospace engineering and engineering management. He has published 41 papers in the areas of computational fluid dynamics and high-performance computing. Dr. Bailey's has been involved in computer networking in support of high-performance computing for over 22 years, beginning with the founding of NASA's first IP network (NASnet) and the Bay Area Regional Research Network. Later, as founder of AppsPoint Corp. he was engaged in design and development of advanced networks. He is currently involved in development of methods and technology for network security at NASA.

Dr. James Corones is presently president of the Krell Institute, a 501c(3) corporation that he founded in 1997. He has had an extensive career in academia, research institutions, and government laboratories. He holds an ScB in physics from Brown University, and a Ph.D. in physics from Boston University with an emphasis on mathematical physics. He began his academic career as an assistant professor at Iowa State University in 1973, being promoted to professor in 1982. From 1978 through 1997, he held various research and administrative posts at the Ames Laboratory – a Department of Energy laboratory managed by Iowa State University. His administrative posts included Program Director – Applied Mathematical Sciences, Program Director – Environmental Technology Development, Deputy Director and then Acting Director. His research interests have been in linear and nonlinear wave propagation, including extensive work in acoustic and electromagnetic inverse scattering problems. He conducted this research under the auspices of the Department of Energy, the Office of Naval Research, the Air Force Office of Scientific Research, and the National Science Foundation. He was a Fullbright fellow and twice a National Academy of Sciences exchange scientist. He has served on various editorial and advisory boards. His primary interest on this subpanel is in helping to assure that ESNET continues to effectively serve the needs of the science and engineering research communities.

Dr. Ian Foster is director of the Computation Institute at the University of Chicago and Argonne National Laboratory; senior scientist and associate division director in the Mathematics and Computer Science Division at Argonne National Laboratory, where he leads the Distributed Systems Laboratory; and the Arthur Holly Compton Distinguished Service Professor of Computer Science at the University of Chicago. An earlier project, Strand, received the British Computer Society Award for technical innovation. Foster led research and development of software for the I-WAY wide-area distributed computing experiment, which connected supercomputers, databases, and other high-end resources at 17 sites across North America in 1995. The Distributed Systems Laboratory, which he leads, is the nexus of the multi-institute Globus Alliance, a research and development effort that encourages collaborative computing by providing advances necessary for engineering, business, and other fields. Furthermore, the Computation Institute addresses many of the most challenging computational and communications problems facing Grid implementations today. Dr. Foster’s research has resulted in the development of techniques, tools and algorithms for high-performance distributed computing and parallel computing, as a result he has been denoted as “the father of the Grid”. Dr. Foster’s honors include the Brittish Computer Society Award for technical innovation, the Lovelace Medal of the British Computer Society, and the Gordon Bell Prize for high-performance computing (2001.) He was elected Fellow of the American Association for the Advancement of Science in 2003.

Prof. Harvey Newman received a Sc. D. degree from MIT in 1974. He has been on the faculty at the California Institute of Technology since 1982 and is currently profession of physics there. Dr. Newman co-led the MARK J Collaboration that discovered the gluon, the carrier of the strong force, at the DESY laboratory in Hamburg in 1979. His current activities in physics include the use of precision photon measurements to search for the Higgs particles thought to be related to the generation of particle masses in the universe, with the Compact Muon Solenoid (CMS) at CERN's Large Hadron Collider (LHC) that will begin operation starting in 2007. He also leads the Caltech MINOS group studying neutrino flavor oscillations. Since 1998 he has served as the chair of the Collaboration Board of the U.S. contingent of CMS, which includes more than 600 physicists and engineers from 48 universities and Fermi and Lawrence Livermore National Laboratories. Dr. Newman has had a leading role in the strategic planning, development, operation, and management of international networks and collaborative systems serving the high-energy and nuclear physics communities since 1982, and he served on the Technical Advisory Group for the NSFNet in 1986. He originated the Data Grid Hierarchy concept adopted by the four LHC high-energy physics collaborations. He is the PI of the LHCNet project, linking the United States and CERN in support of the LHC physics program, and is the PI of the NSF-funded UltraLight project developing the next generation of network-integrated Grid systems. He chairs the ICFA Standing Committee on Inter-regional Connectivity. He and his team have established 11 Internet2 Land Speed Records since 2002, and he led the international consortium that won the SC2003, SC2004, and SC2005 Bandwidth Challenge Awards, with a 2005 record of 151 Gbps.

Ms. Gwendolyn Huntoon is the director of networking at the Pittsburgh Supercomputing Center (PSC). In this role, she directs the networking group at PSC, which supports supercomputing users, carries out advanced networking research, and provides consulting and training to universities and research centers nationwide. She also directs the Three Rivers Optical Exchange (3ROX), a high-speed network aggregation point that provides advanced network services to over 20 research, higher education and

K12 groups in Western Pennsylvania and West Virginia. Ms. Huntoon is also the director of operations for National LambdaRail, a major initiative of U.S. research universities and private sector technology companies to provide a national scale infrastructure for research and experimentation in networking technologies and applications. From 2001 – 2004 she served as the first executive director for the Quilt, a coalition of advanced regional network organizations. Throughout her tenure at PSC Ms. Huntoon has focused on developing and then supporting an environment that fosters networking research within a high-performance network environment. To that extent, she has been principal investigator, co-principal investigator, or senior personnel on a number of federally funded research grants. Examples includ: lead PI on the NSF funded NLANR Engineering Services project that provided engineering support for the advanced network community and co-PI on the NSF funded Web100 project, that focused on developing standards and software to support better end-to-end network performance.

Dr. Larry Rahn is a senior scientist in the Center for Biological and Energy Sciences at Sandia National Laboratories, Livermore, CA. After earning his Ph.D. in physics at Kansas State University and performing postdoctoral studies at Michigan State University, Dr. Rahn joined Sandia, where he contributed to early efforts that led to the founding of the Combustion Research Facility (CRF,) a DOE Basic Energy Sciences User Facility. He became a principal investigator in the BES Chemical Science Program at the CRF, earning a promotion to Distinguished Member of Technical Staff and election to Fellow of the Optical Society of America. After seventeen years of research in laser-based combustion diagnostics, he managed the Reacting Flow Research Department at the CRF for almost a decade, when he was promoted to senior scientist. While a manager and in his Senior Scientist role, he led research in collaborative data sharing environments. The most recent of these resulted in the DOE Collaboratory for Multi-Scale Chemical Science (CMCS,) the infrastructure of which is being developed and used for follow-on projects. His research has resulted in more than 60 journal publications in Raman and nonlinear optical spectroscopy, combustion diagnostics, molecular physics, solid-state physics, and collaborative data sharing environments. Dr. Rahn is currently on assignment to DOE BES to assist in program management.

Dr. Horst Simon was named associate laboratory director (ALD) for Computing Sciences at Berkeley National Laboratory in 2004. In this role, Dr. Simon represents the interests of the lab’s scientific computing divisions, NERSC, and computational research in the formulation of Laboratory policy, and he leads the overall direction of the two divisions. He also coordinates constructive interactions within the computing sciences divisions to seek coupling with other scientific programs. Dr. Simon joined LBNL in early 1996 as director of the newly formed NERSC Division and was one of the key architects in establishing NERSC at its new location in Berkeley. The NERSC Center is DOE’s flagship supercomputing facility for unclassified research funded by DOE’s Office of Science and currently supports 2,677 users at more than 300 institutions. Under his leadership, NERSC has enabled important discoveries in fields ranging from global climate modeling to combustion to astrophysics. Dr. Simon is also the founding director of Berkeley’s Computational Research Division, which conducts applied research and development in computer science, computational science, and applied mathematics. His research interests are in the development of sparse matrix algorithms, algorithms for large-scale eigenvalue problems, and domain decomposition algorithms for unstructured domains for parallel processing. His recursive spectral bisection algorithm is regarded as a breakthrough in parallel algorithms for unstructured computations, and his algorithm research efforts were honored with the 1988 Gordon Bell Prize for parallel processing research. Dr. Simon was a member of the NASA team that developed the NAS Parallel Benchmarks, a widely used standard for evaluating the performance of massively parallel systems. He is also one of four editors of the twice-yearly “TOP500” list of the world’s most powerful computing systems.

Mr. Jerry Sobieski is the director of Research Initiatives for the Mid-Atlantic Crossroads (MAX,) a consortium of almost 50 research and higher education institutions in the Washington, DC, region. He is responsible for developing strategic and multi-institutional networking research programs that address the needs of the next generation of globally distributed e-science applications. Mr. Sobieski is PI on the NSF DRAGON Project, an experimental optical network testbed in the Washington, DC, metro area developing GMPLS-based dynamic hybrid service architectures. Besides his work on DRAGON, Mr. Sobieski and his team are part of the Testbed Support Center for the Internet2 Hybrid Optical/Packet Infrastructure (HOPI) and they support the Global Information Grid Experimental Facility – a DoD-funded advanced technology testbed in Washington, DC. Mr. Sobieski has served on the Technical Advisory Committee for the Internet2 Abilene network, the HOPI Design Team, the Atlantic Wave Engineering and Governance committees, and NetWorkMaryland Engineering Advisory Board. Before joining MAX, Mr. Sobieski was the director of advanced networking for Highway1, a non-profit organization in Washington, DC, focused on presenting the technical and policy issues of emerging global advanced internetworking to members of Congress and industry associations. From 1997 until 1999, Mr. Sobieski worked for the Internet2 organization as part of the Abilene design and implementation team. Mr. Sobieski's career has focused on high-performance computing, from systems development in vector supercomputers in the 1980s, to heading the Laboratory for Parallel Computation at the University of Maryland Institute for Advanced Computer Studies in the 1990s. He has worked closely with industry and academia in developing and deploying advanced computational technologies in the areas of seismic processing, image processing of satellite imagery for climate and land cover dynamics, and radio astronomy, as well as advanced networking architectures and technologies. Mr. Sobieski holds a B.S. in computer science from the University of Houston. Ongoing interests and activities include design and modeling of application specific network topologies for resilience, security, and deterministic performance; ultra-high capacity photonic packet switching and transport systems for distributed computing and Grid architectures; and the design and engineering of global network infrastructure. His interest in the network arises from his high-performance computing background. Scientific computing has infused his entire career, driving him into advanced networking in order to feed the applications. He is interested in emerging models of applications and networking integration, in particular in large parallel and globally distributed computational applications.

Dr. Ellen Stechel (Chair,) received her Ph.D. in chemical physics from the University of Chicago, 1978. Her principal areas of expertise are transportation and energy technologies and computational chemical and materials sciences. She currently manages a department in the Energy Futures Group at Sandia National Laboratories, where she builds research programs and capabilities in energy technologies to simultaneously reduce the nation’s dependence on fossil energy and reduce greenhouse gas emissions. Her department also is responsible for managing the Department of Energy’s Atmospheric Radiation Measurement Facility on the North Slope of Alaska, which plays a critical role in reducing uncertainties in climate science models. Dr. Stechel joined Sandia, in 1981 as a technical staff member in the Condensed Matter Physics Department. From 1994 to 1998, she was the manager of the Advanced Materials and Device Sciences Department, which worked on experimental, theoretical, and computational projects in novel materials with a number of applications and devices in mind. Prior to rejoining Sandia, she worked at Ford Motor Company from 1998 to 2005, where she covered a range of energy and environmental programs at Ford and in Universities (e.g., Ford/MIT Alliance and BP/Ford/Princeton Carbon Mitigation Initiative,) including building a sustainability science program, overseeing Ford Research

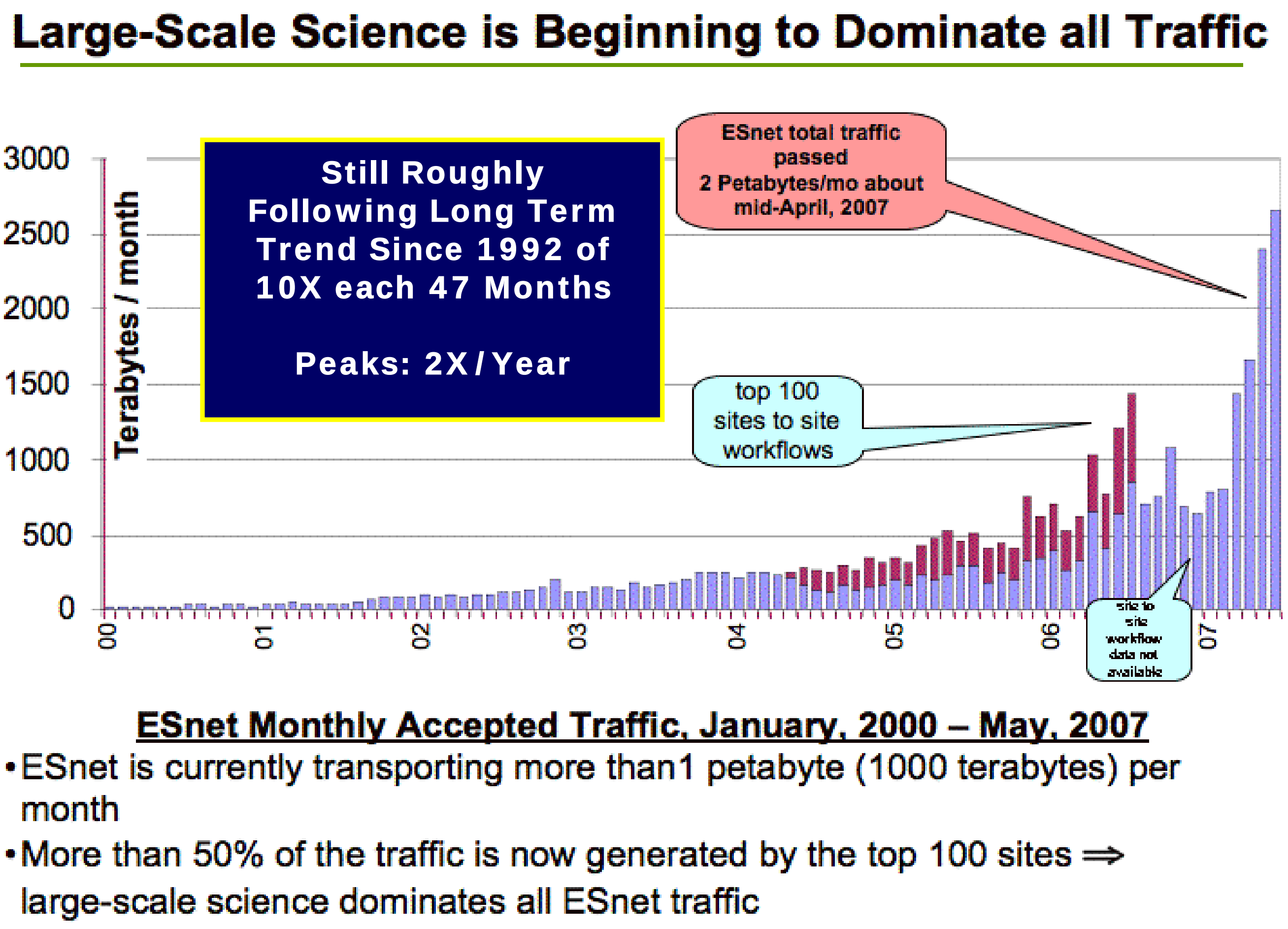
Lab’s atmospheric chemistry and climate change programs, as well as proving and deploying new low emission technologies on many of Ford’s North American vehicles. In 2005, she worked on contract for the Department of Homeland Security, Science & Technology Division, Office of Research and Development. Upon completion of that temporary assignment, she returned to Sandia, Albuquerque, to form her current department in the summer of 2006. As a result of her varied career, her experience base has touched multiple aspects of the science, engineering, technology, business, and policy enterprise in a number of research fields, including chemistry, physics, materials, surfaces, the environment, and computational science, in which she has published more than 65 peer-reviewed articles. Her professional activities include senior editor of the Journal of Physical Chemistry 1998-2000; chair of the Division of Physical Chemistry, American Chemical Society, 1998; and co-founder, Strategic Oversight, and Scientific Oversight for the Computational Materials Science Network (CMSN,) DOE/Basic Energy Sciences. Dr. Stechel's interest in the network includes its role in enabling a transition in the practice of science to foster computational discovery, multi-disciplinary research, system-level science, data-intensive research, collaborative research, multi-organizational research, research from a distance, and the acceleration of discovery to application.

Dr. Bill Wing (Co-Chair,) received his Ph.D. in physics from the University of Iowa in 1972. His research interests are in network monitoring and instrumentation, network simulation, and protocol development. Dr. Wing is a senior research staff member in the Networking Research Group, Computer Science and Mathematics Division, Oak Ridge National Laboratory (ORNL); co-PI on the DOE UltraScience Net Project; and project manager of the ORNL FutureNet Infrastructure Project. Dr. Wing joined the Fusion Energy Division immediately after completing his PhD. He developed and applied a wide variety of diagnostic instruments for characterizing the fusion plasmas in experimental devices there. He received a patent for one of these, a Gigacycle Correlator. While there, he started using early laboratory-scale computers (PDP-8, PDP-12, and a PDP-10) for data acquisition and analysis. He led the in-house programming group responsible for writing data acquisition software and developed an integrated data acquisition system that spread throughout the fusion community. His interest in computerized analysis and modeling led to an interest in networking (ORNL’s Fusion Energy Division was one of the first backbone nodes on the Magnetic Fusion Energy Network, which linked Fusion sites to the MFE computer center at Livermore.) In 1991, he moved from the Fusion Energy Division to the Office of Laboratory Computing to help improve ORNL’s position in the high-performance computing and networking community. He has chaired the ESnet Site Coordinating Committee, as well as serving as ORNL’s representative on it. He was chair of the SCinet committee for SC’99 in Portland and again for SC 2001 in Denver. He served as ORNL’s representative on the National Information Infrastructure Testbed, fills ORNL’s chair on the National Lambda Rail board of directors, and is co-PI on DOE's UltraScience Net Testbed. In addition, he is network architect and project manager for the ORNL FutureNet DWDM infrastructure being built to link ORNL and selected research universities in the Tennessee Valley to Chicago and Atlanta at multiple 10Gbps speed.

D. ASCAC Networking Subcommittee Meeting April 13, 2007, (summarized by Eli Dart)

- **For all the science cases the following were identified by examining the science environment:**
  - Instruments and facilities
    - Location and use of facilities, instruments, computational resources, etc.
    - Data movement and storage requirements

  - Process of science
    - Collaborations
    - Networking services requirements
    - Noteworthy patterns of use (e.g., duty cycle of instruments)
  - Near-term needs (now to 12 months)
  - 5-year needs (relatively concrete)
  - 5- to 10-year needs (more uncertainty)
- **The requirements collected from the case studies form the foundation for the current ESnet4 architecture**
  - Bandwidth, Connectivity Scope / Footprint, Services
  - We do not ask that our users become network experts in order to communicate their requirements to us
  - We ask what tools the researchers need to conduct their science, synthesize the necessary networking capabilities, and pass that back to our constituents for evaluation
- **We have collected requirements from diverse science programs, program offices, and network analysis – the following summarizes the requirements:**
  - **Reliability**
    - 99.95% to 99.999% reliability
    - Redundancy is the only way to meet the reliability requirements
      - Redundancy within ESnet
      - Redundant peerings
      - Redundant site connections where needed
  - **Connectivity**
    - Geographic reach equivalent to that of scientific collaboration
    - Multiple peerings to add reliability and bandwidth to interdomain connectivity
    - Critical both within the US and internationally
  - **Bandwidth**
    - 10 Gbps site to site connectivity today
    - 100 Gbps backbone by 2010
    - Multiple 10+ Gbps R&E peerings
    - Ability to easily deploy additional lambdas and peerings
  - **Service guarantees**
    - All R&E networks must interoperate as one seamless fabric to enable end2end service deployment
    - Guaranteed bandwidth, traffic isolation, quality of service
    - Flexible-rate bandwidth guarantees
  - **Collaboration support**
    - Federated trust, PKI (Grid, middleware)
    - Audio and video conferencing
  - **Production ISP service**

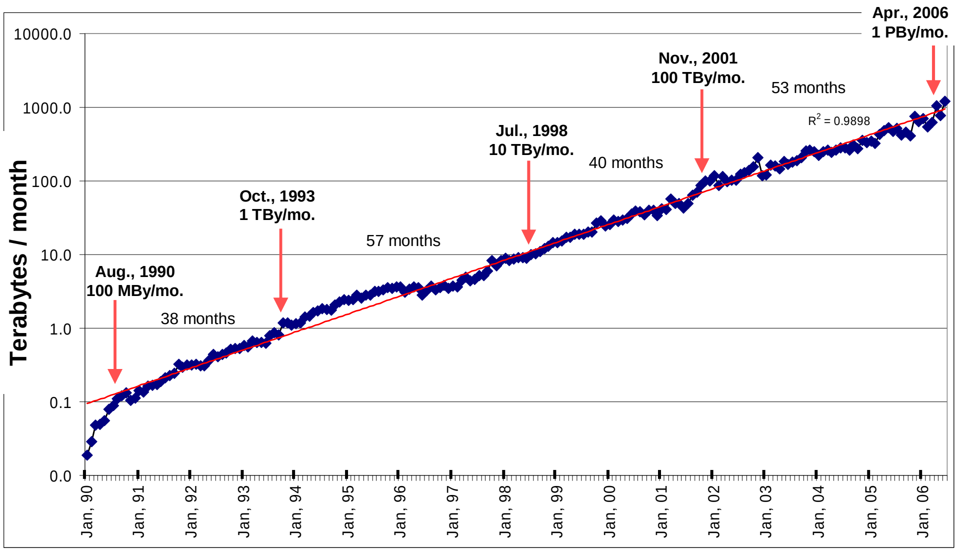


Large-Scale Science is Beginning to Dominate all Traffic

ESnet Monthly Accepted Traffic, January, 2000 – May, 2007

- ESnet is currently transporting more than1 petabyte (1000 terabytes) per month
- More than 50% of the traffic is now generated by the top 100 sites ⇒ large-scale science dominates all ESnet traffic

ESnet Traffic has Increased by
<u>10X Every 47 Months</u>, on Average, Since 1990

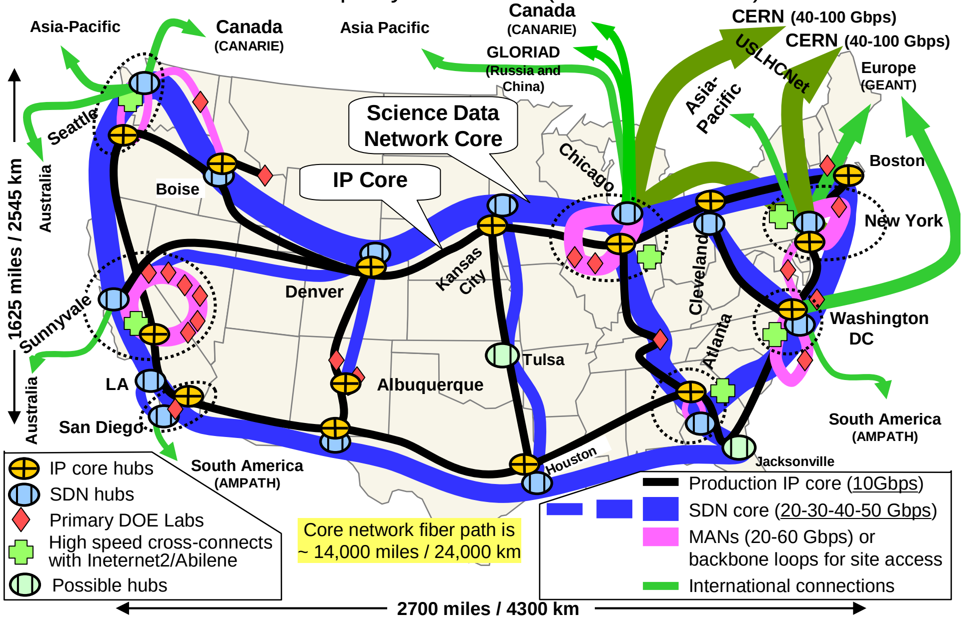


Log Plot of ESnet Monthly Accepted Traffic, January, 1990 – June, 2006

1

ESnet4

Core networks 50-60 Gbps by 2009-2010 (10Gb/s circuits),
500-600 Gbps by 2011-2012 (100 Gb/s circuits)



E. Selected References

ASCR Website, http://www.er.doe.gov/ascr/About/about.html

Collaborative Science, "Portal-based Knowledge Environment for Collaborative Science," *Concurrency and Computation: Practice and Experience*, Karen Schuchardt, et al, Volume 19, Issue 12 , Pages 1703 – 1716 (25 August, 2007.)

Cyber Infrastructure – "Revolutionizing Science and Engineering Through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure," January 2003, http://www.nsf.gov/od/oci/reports/atkins.pdf

"Cyber Security Research Needs for Open Science" Workshop Report, July 23–24, 2007, http://www.er.doe.gov/ascr/Misc/CSWorkshopFinalReport.pdf

"Data-Management Challenge," Report from the DOE Office of Science Data-Management Workshops March–May 2004 http://www.er.doe.gov/ascr/ProgramDocuments/Final-report-v26.pdf

Fast Data Transfer, http://monalisa.cern.ch/FDT/

Grid: I. Foster, C. Kesselman, and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," International Journal of Supercomputer Applications, 15 (3.) 200–222, 2001.

The Grid: Blueprint for a New Computing Infrastructure, Editors-Ian Foster and Carl Kesselman, Morgan Kaufmann Publishers, 2006

"Scaling System-Level Science: Scientific Exploration and IT Implications, Ian Foster and Carl Kesselma, Published by the IEEE Computer Society (November,) 32–39, 2006.

"e-Science and Its Implications," Tony Hey and Anne Trefethen, *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, Vol. 361, No.1809, Information, Knowledge and Technology. (Aug. 15, 2003,) pp. 1809-1825. Stable URL: http://links.jstor.org/sici?sici=1364-503X%2820030815%29361%3A1809%3C1809%3AEAII%3E2.0.CO%3B2-T

ESnet Website, http://www.er.doe.gov/ascr/Facilities/ESnet.html

Exascale Report: "Modeling and Simulation at the Exascale for Energy and Environment," Report on three Town Hall Meetings in 2007, http://www.er.doe.gov/ascr/ProgramDocuments/TownHall.pdf

"Facilities for the Future of Science: A 20 Year Outlook" http://www.science.doe.gov/Scientific_User_Facilities/History/20-Year-Outlook-screen.pdf and Four Years Later, An Interim Report, August 2007, http://www.science.doe.gov/Scientific_User_Facilities/History/20-Year-Outlook-Interim%20Report(August).pdf

"MonALISA: An Agent based, Dynamic Service System to Monitor, Control and Optimize Grid based Applications," I.C.Legrand, H.B.Newman, R.Voicu, C.Cirstoiu, C.Grigoras, M.Toarta, C. Dobre, CHEP 2004, Interlaken, Switzerland, September 2004.

Optical Networking Testbed 3 Workshop, September 7–8, 2006, Tokyo, Japan, http://www.nren.nasa.gov/workshops/pdfs9/ONT3_FINAL_Report_June2007_V15.pdf

"Service-Oriented Science," I. Foster, Science, 308, 814-817, 2005.

System of Systems: Carlock and Fenton, "System-of-Systems (SoS) Enterprise Systems for Information-Intensive Organizations," Systems Engineering, Vol 4, No. 4 (2001) pp. 242-261.

"A Vision for DOE Scientific Networking Driven by High Impact Science,"W. E. Johnston, W. T. C. Kramer, J. F. Leighton, and Charlie Catlett, March 15, 2002.

F. Abbreviations/Acronyms

ALCF – **A**rgonne **L**eadership **C**omputing **F**acility http://www.alcf.anl.gov/

ASCR – **A**dvanced **S**cientific **C**omputing **R**esearch

ATDnet – DARPA's Advanced Technology Demonstration Network

BESAC – **B**asic **E**nergy **S**ciences **A**dvisory **C**ommittee

BGP – The **B**order **G**ateway **P**rotocol is the core routing protocol of the Internet. It works by maintaining a table of IP networks or "prefixes," which designate network reachability among autonomous systems. It is described as a path vector protocol. BGP does not use traditional IGP metrics but makes routing decisions based on path, network policies and/or rule sets.

CERN – European Organization for Nuclear Research

CHEETAH – NSF's **C**ircuit-switched **H**igh-speed **E**nd-to-**E**nd **T**ransport **A**rc**H**itecture project http://cheetah.cs.virginia.edu/

COMPETES – **C**reating **O**pportunities to **M**eaningfully **P**romote **E**xcellence in **T**echnology, **E**ducation, and **S**cience

CPU – **C**entral **P**rocessing **U**nit

DOE – **D**epartment **o**f **E**nergy

DRAGON – NSF's **D**ynamic **R**esource **A**llocation via **G**MPLS **O**ptical **N**etworks project http://dragon.east.isi.edu/twiki/bin/view/Main

DSL – **D**igital **S**ubscriber **L**ine

DWDM – **D**ense **W**ave **D**ivision **M**ultiplexing

ESnet – DOE's Energy Sciences network http://www.es.net/

eVLBI – **E**lectronic **V**ery **L**ong **B**aseline **I**nterferometry

GÉANT2 – Pan-European research and education network http://www.geant2.net/. The GÉANT2 network connects 34 countries through 30 national research and education networks (NRENs,) using multiple 10 Gbps wavelengths

Gbps – Gigabits per second

GENI – NSF's **G**lobal **E**nvironment for **N**etworking **I**nvestigations

GigaPoP – Gigabit per second Point of Presence

GLIF – **G**lobal **L**ambda **I**ntegrated **F**acility

GMPLS – **G**eneralized **M**ulti**p**rotocol **L**abel **S**witching, also known as Multi-protocol Lambda Switching, is a technology that provides enhancements to Multi-protocol Label Switching (MPLS) to support network switching for time, wavelength, and space switching as well as for packet switching. In particular, GMPLS provides support for photonic networking, also known as optical communications.

HOPI – The **H**ybrid **O**ptical and **P**acket **I**nfrastructure Project http://networks.internet2.edu/hopi/

ICFA – **I**nternational **C**ommittee **F**uture **A**ccelerators

IETF – **I**nternet **E**ngineering **T**ask **F**orce is a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet.

I/O – **I**nput/**O**utput

IP – **I**nternet **P**rotocol

ISP – **I**nternet **S**ervice **P**rovider

IT – **I**nformation **T**echnology

LAN – **L**ocal **A**rea **N**etwork

LCF – **L**eadership **C**omputing **F**acility

LHC – **L**arge **H**adron **C**ollider

LSN – **L**arge **S**cale **N**etworking

LUSTRE – An open source distributed file system

MAN – **M**etropolitan **A**rea **N**etwork

MAN LAN – Manhattan Landing exchange point

MAX – **M**id-**A**tlantic **E**xchange

Mbps – **M**ega**b**its **p**er **s**econd

MFECC – **M**agnetic **F**usion **E**nergy **C**omputation **C**enter

MPLS – **M**ulti-**P**rotocol **L**abel **S**witching, a standards-approved technology for speeding network traffic flow and making it easier to manage. MPLS involves setting up a specific path for a given sequence of packets, identified by a label put in each packet, thus saving the time needed for a router to look up the address to the next node to forward the packet to. MPLS is called *multiprotocol* because it works with the Internet Protocol, Asynchronous Transport Mode, and frame relay network protocols.

NCAR – **N**ational **C**enter for **A**tmospheric **R**esearch

NCO/NITRD – National Coordination Office for the NITRD Program

NITRD – The Federal **N**etworking and **I**nformation **T**echnology **R**esearch and **D**evelopment Program

NLR – **N**ational **L**ambda**R**ail http://www.nlr.net/

NMS – **N**etwork **M**anagement **S**ystem

NREN – **N**ational **R**esearch and **E**ducation **N**etworks

NSF – **N**ational **S**cience **F**oundation

O&M – **O**perations and **M**anagement

OEP – **O**pen **E**xchange **P**oint

OMB – **O**ffice of **M**anagement and **B**udget

ONT – **O**ptical **N**etwork **T**estbed

OSCARS – DOE/SC's **O**n-demand **S**ecure **C**ircuits and **A**dvanced **R**eservation **S**ystem project http://www.es.net/oscars/

OSG – **O**pen **S**cience **G**rid

OSPF – The **O**pen **S**hortest **P**ath **F**irst**,** a hierarchical interior gateway protocol (IGP) for routing in Internet Protocol, using a link-state in the individual *areas* that make up the hierarchy. A computation based on Dijkstra's algorithm is used to calculate the shortest path tree inside each area.

PCAST – **P**resident's **C**ouncil of **A**dvisors on **S**cience and **T**echnology

PITAC – **P**resident's **I**nformation **T**echnology **A**dvisory **C**ommittee

PKI – **P**ublic **K**ey **I**nfrastructure

PoP – **P**oint **o**f **P**resence

PPP – **P**oint-to-**P**oint **P**rotocol

PNWGP - **P**acific **N**orth**W**est **G**iga**P**oP

PSC - **P**ittsburgh **S**upercomputing **C**enter, http://www.psc.edu/

QoS – **Q**uality **o**f **S**ervice. On the Internet and in other networks, QoS is the idea that transmission rates, error rates, and other characteristics can be measured, improved, and, to some extent, guaranteed in advance. QoS is of particular concern for the continuous transmission of high-bandwidth video and multimedia information. Transmitting this kind of content dependably is difficult in public networks using ordinary "best effort" protocols.

R&D – **R**esearch and **D**evelopment

R&E – **R**esearch and **E**ducation

RA – **R**adio **A**stronomy

RF – **R**adio **F**requency

RFC – **R**equests **f**or **C**omments is an internet standards related document.

SC – Office of Science

SCIC – **S**tanding **C**ommittee on **I**nternational **C**onnectivity

SciDAC – **Sci**entific **D**iscovery through **A**dvanced **C**omputing

SDN – **S**cience **D**ata **N**etwork

SOA – **S**ervice-**O**riented **A**rchitecture

SONET – **S**ynchronous **O**ptical **NET**work

StarLight – NSF-funded international optical exchange point

Tbps – Terabits per second

TCP/IP – Transmission Control Protocol/Internet Protocol

TeraGrid – NSF's grid computing initiative

USN – DOE/SC's **U**ltra**S**cience **N**etwork, an experimental research testbed funded by DOE Office of Science. The next generation DOE large-scale science applications have requirements that will drive extreme networking, in terms of sheer high throughput and dynamically stable connections. The UltraScience Net goal is to enable the development of hybrid optical networking and associated technologies to meet the unprecedented demands of large-scale science applications. http://www.csm.ornl.gov/ultranet/

VLAN – **V**irtual **L**ocal **A**rea **N**etwork

VLBI – **V**ery **L**ong **B**aseline **I**nterferometry

WAN – **W**ide **A**rea **N**etwork

WDM – **W**avelength **D**ivision **M**ultiplexing

XFP – (10 Gigabit Small Form Factor Pluggable,) a hot-swappable, protocol-independent optical transceiver, typically operates at 850 nm, 1310 nm, or 1550nm, for 10 Gbps SONET/SDH, Fibre Channel.