Strategic Plan FY2014-FY2023 ESnet

March 1, 2013

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

Table of Contents

1. Facility Overview

The Energy Sciences Network (ESnet) is the Department of Energy's high-performance networking facility, engineered and optimized for large-scale science.

Funded by the Office of Science (SC) and managed by Berkeley Lab, ESnet interconnects the entire national laboratory system, including its supercomputer centers and user facilities – enabling tens of thousands of scientists to transfer data, access remote resources, and collaborate productively.

Because network requirements for large-scale science are demanding, ESnet allocates significant effort to applied research, development, and innovation. Currently the fastest science network in the world, ESnet is regarded as an influential leader among its global peers. The foundation for this success is an unparalleled team of engineers and other contributors, each focused on the mission of enabling and accelerating scientific discovery. In the facility's most recent Operational Assessment Report, reviewers commented: "The entire staff conscientiously and continually lead their field."

ESnet was founded in 1986, soon after the creation of the global Internet. Since 1990, ESnet's traffic has increased by a factor of 10 every 48 months, roughly double the growth rate of the commercial Internet.[1]

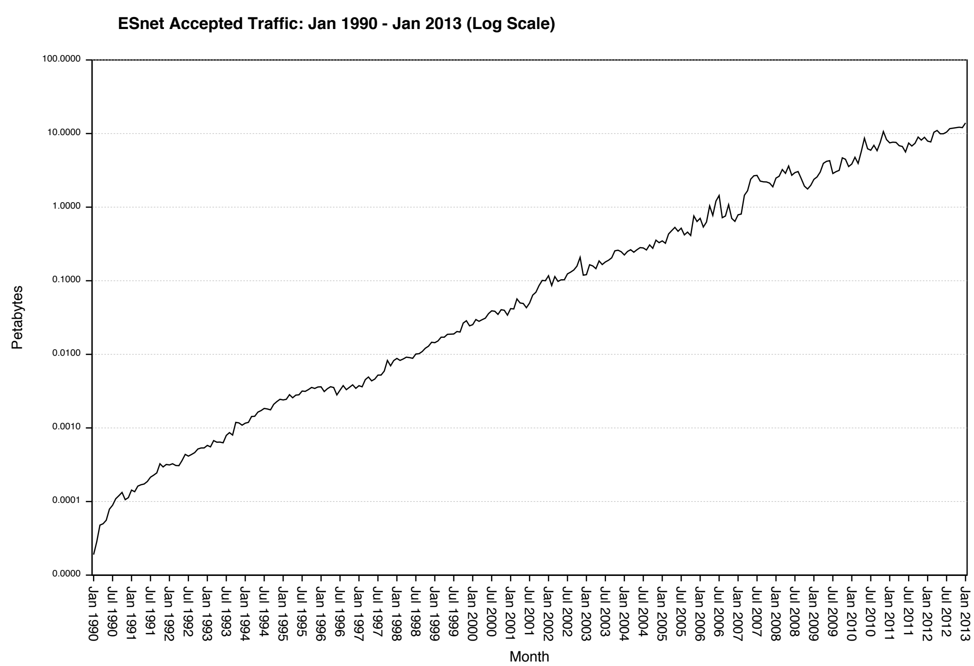


Figure 1: ESnet traffic accepted, PB/month, since 1990.

[1] IEEE 802.3™ Industry Connections Ethernet Bandwidth Assessment Ad Hoc Report, http://www.ieee802.org/3/ad_hoc/bwa/BWA_Report.pdf .

ESnet serves more than 40 DOE sites, connecting them to over 100 research and commercial networks worldwide. The richness of ESnet's global connectivity is motivated by the fact that 80% of its traffic originates or terminates outside the national laboratory complex. This pattern in turn reflects the collaborative, increasingly international nature of scientific research. ESnet serves other DOE Offices in addition to SC, including NNSA, EERE, IN, and EM.[2]

In late 2012, ESnet completed the most important upgrade in its history. With $62M of ARRA funding, the facility acquired rights to the underlying spectral capacity on a nationwide footprint of fiber-optic cable; built the world's first continental-scale 100 Gigabit-per-second (100Gbps) network in partnership with Internet2; and transitioned production traffic to the new infrastructure. As a result of this investment, ESnet can now support the data-mobility needs of distributed science more flexibly and cost-effectively than ever before.

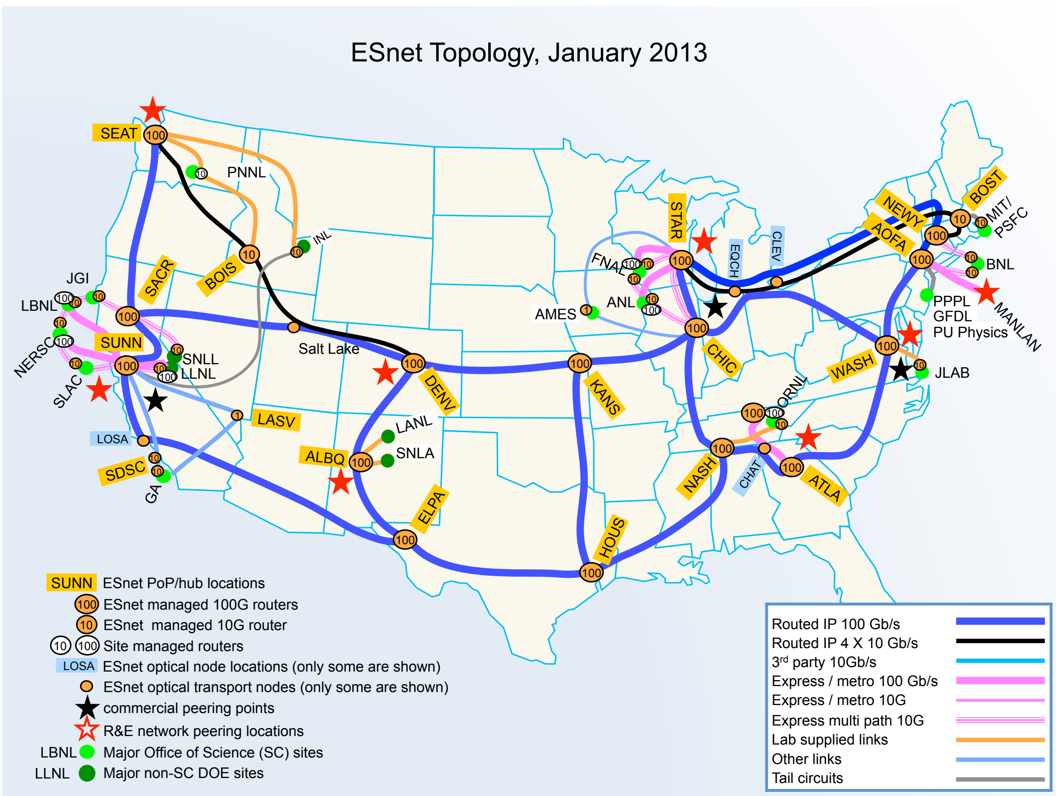


Figure 2: Current ESnet topology, after 100G upgrade in 2012.

ESnet's architecture and capabilities are tailored for its users, whose needs are derived through a formal requirements-gathering process. The facility sponsors periodic requirements reviews for each of the six SC program offices. These reviews focus on scientific projects and workflows, rather than technology. By first seeking to understand how disciplines and collaborations produce knowledge, ESnet obtains more useful and accurate requirements information. This approach has been highly successful, serving as a model for other facilities.

[2] The business model for these relationships is transitioning from *incremental cost-recovery* to *cost-sharing*.

To forge strong ties with its community and users, ESnet invests considerable energy in education, outreach, and advocacy. In a typical year, ESnet staff participate in more than 100 community and stakeholder meetings[3] and gives 60-90 technical presentations, including keynotes, tutorials, workshops, and informational sessions to help DOE scientists and IT professionals make better use of high-performance networks. The organization recently hired a Lead for Partnerships and Outreach, and will increase activity in this area significantly.

ESnet currently has 37 career staff, plus 2.5 part-time, term, contractor, or student employees; student numbers increase by 5-6 in summer months. Internet, systems, and software engineers make up nearly 80% of staff. ESnet's annual budget was $34.5 in FY12, with a request of $32M in FY13. In response to recent retirements, as well as chronic under-staffing, the organization plans to add 4-6 FTEs in the coming year.

It is important to understand that ESnet differs from a commercial provider of Internet services. Massive science data flows require different network capabilities than smaller flows generated by email, video, and web browsing. As a network carefully optimized for large-scale science, ESnet's capabilities include:

- an engineering emphasis on *lossless transport* (see **section 5.1** below)
- bandwidth guarantees, spanning multiple network domains
- a distributed performance monitoring platform
- a team of experts for rapid diagnosis of multi-domain data transport issues
- web-based visualization tools for monitoring the health of science data flows
- very high-speed (n x 100G) data transport, scaling cost-effectively
- global leadership in developing standards and applications to support data-intensive science flows
- global advocacy for network and security architectures to enable high-throughput data exchange
- extensive scientific outreach and requirements gathering activity
- a consistent record of architectural innovation

The structure of modern science now presumes the availability of reliable, high-bandwidth, feature-rich networks for interconnecting instruments and collaborators globally.[4] In this context, ESnet serves as a vital "circulatory system" for all DOE facilities, and for every mission space within in the Office of Science.

[3] Some of these occur using video conferencing services, which ESnet operates and provides for the entire SC community.

[4] The Large Hadron Collider at CERN may have been the first major experiment for which advanced networking was a design premise, but it certainly won't be the last. Dozens of facilities currently under construction worldwide make exactly the same assumption.

2. Mission, Goals, Drivers, and Strategic Connections

2.1 Vision and Mission Statements

ESnet's Mission

To enable and accelerate scientific discovery by delivering unparalleled network infrastructure, capabilities, and tools.

ESnet's Vision

1. Scientific progress will be completely unconstrained by the physical location of instruments, people, computational resources, or data.

2. Collaborations at every scale, in every domain, will have the information and tools they need to achieve maximum benefit from scientific facilities, global networks, and emerging network capabilities.

3. ESnet will foster the partnerships and pioneer the technologies necessary to ensure that these transformations occur.

2.2 Facility Strategic Goals

1. Ensure that ESnet continues to provide highly reliable data transport capabilities, optimized for the requirements of large-scale science, and scaling to accommodate the exponential traffic growth of DOE science missions.

2. In response to the challenges posed by the scientific data revolution, develop the ability to assess and rapidly improve how collaborations make use of high-performance networking.

3. Pioneer, develop, and sustain the architectures, protocols, and software necessary to achieve the facility vision.

4. Influence DOE facilities, labs, campuses, and networks around the world to optimize their infrastructures for end-to-end data movement.

2.3 Connections with DOE and ASCR Strategic Plans

The mission of the Department of Energy, as articulated in its May 2011 strategic plan, is to ensure America's security and prosperity by addressing its energy, environmental, and nuclear challenges through transformative science and technology solutions. Because ESnet serves every program office and nearly every major facility in DOE – including all SC user facilities, major SC collaborators, and NNSA labs and sites – *it broadly advances and* *sustains the entire spectrum of DOE missions*. For many DOE activities, including those in SC and the NNSA, ESnet offers vital enabling capabilities.[5]

ESnet's mission is tightly coupled with that of SC, which has the twin objectives of direct support for scientific research, and support for development, construction, and operation of open-access scientific user facilities. ESnet ensures that its strategies are aligned with those of SC by sponsoring ongoing network requirements reviews for each SC program office. These reviews are described in **section 4.3.1**, below.

The mission of the Advanced Scientific Computing Research (ASCR) program is to discover, develop, and deploy computational and networking capabilities to analyze, model, simulate, and predict complex phenomena important to the Department of Energy (DOE). As the scientific workflows initiated and supported by SC become more data-intensive, ESnet's capabilities become more strategic for SC. Any initiative to create a data-services facility for SC, either integrated across multiple sites or concentrated in one, will rely on ESnet for advanced networking capabilities to support activities such as data replication, data sharing, remote access to storage, and bulk data transfer.

2.4 Science Drivers

ESnet's method for assessing science drivers for SC is described in **section 4.3,** below. Several consistent themes have emerged from this process.

Exponential data growth. New instruments and computing systems continue to generate exponentially growing data sets. Moore's Law is transforming physics detectors and image-based sensors, and data growth in genomics exceeds Moore's Law by a considerable margin. The relentless growth in data associated with SC-funded science missions will require ESnet to scale its capacity aggressively, while maintaining the service quality required for high-throughput workflows.

Data mobility. Large data sets are moved to facilities where they can be analyzed; moved again for secondary analysis; and eventually shared with collaborators. This pattern applies broadly – whether the context is LHC computation, protein crystallography, or climate modeling. Data mobility – the placement of data where it can be productively analyzed – is a hallmark of data-intensive science, and the need for *efficient* data mobility motivates much of ESnet's strategic planning activity.

Centrality of computing and networking. Computing has become essential in virtually every area of science funded by DOE. From extreme-scale simulations to high-throughput analysis of instrument data, computation is ubiquitous. ASCR computational centers are therefore key components in the DOE science complex, and ESnet provides the high-performance services to make these facilities available and productive.

---

[5] ESnet is not regarded as a 'critical infrastructure' for NNSA, and is not managed, engineered, or funded for use as a life-safety or national security system.

Global collaboration. Science collaborations are growing in scope and scale. Collaborations in High Energy Physics, Nuclear Physics, and Earth Sciences can comprise hundreds or even thousands of collaborators. These collaborations are based on the assumption that the services provided by ESnet and the other networks of the world are consistent and reliable, and reach to the full length and breadth of the collaboration – including instruments, facilities, and participating institutions.

A few scientific workflow patterns merit special discussion.

Light sources. The evolution of beamline detectors at photon and neutron sources is following a *compounded* Moore's-law curve. The area of sensors is increasing exponentially (512 x 512, 1024 x 1024, 4096 x 4096), along with readout rate (0.5 Hz, 100 Hz, 1kHz, 1MHz). Each detector – and there may be dozens of these in a single facility – is connected to a data acquisition device inside a computer, where a device driver converts incoming data streams to files (for example, TIFF images) for scientific analysis. As a result, high-performance computing resources will soon be required for analyzing beamline data. Data rates from current-generation detectors are on the order of 1.5 Gbps, with rates of 8 Gbps expected in the next 18 months. As consequence of this growth, high-speed data transfer between light sources and neutron sources (ALS, APS, LCLS, NSLS-II, SNS, SSRL, and so forth) and HPC facilities (ALCF, NERSC, OLCF) will become critical to the scientific process of photon and neutron facilities.

Fusion science. Experimental fusion has a workflow that follows a regular duty cycle. The instrument (for example, a Tokamak) is run in *pulsed* mode: plasma is created for a short period of time, during which data is collected from thousands of sensors. For a period of roughly 15 minutes, data from the pulse is analyzed, and the configuration for the next pulse is determined based on that analysis. This duty cycle places significant demands on the network for consistent and reliable performance. This type of requirement stands in contrast to the more common need to move a 10 TB data set in a period of hours, for example. As more advanced fusion instruments have been developed (EAST in China, KSTAR in South Korea, and the future ITER in France), pulse times are growing longer, data volumes are increasing, and the distributed nature of the experimental teams continues to expand. Peak trans-Atlantic data transfer rates for ITER could reach 400 Gbps.

LHC science. Analysis workflows for ALICE, ATLAS, and CMS involve the processing of billions of individual events. Detectors in the LHC at CERN generate event data from particle collisions, and events are analyzed though a globally distributed collaboration involving thousands of scientists on multiple continents. The LHC experiments rely heavily on the automated distribution of large volumes of data; there is simply too much data (both in terms of number of events and data volume) to accomplish the required analysis on human timescales without automating data distribution. Many of the network capabilities ESnet has pioneered (for example, bandwidth reservations) were initially developed to serve the needs of LHC experiments.

JGI workflows. Most of the computing for the Joint Genome Institute (JGI) facility is now performed at NERSC. The use of NERSC resources by JGI has evolved over time, but a key component has been ESnet's dedicated bandwidth service, based on its OSCARS software. OSCARS makes JGI subnets available from within the NERSC data center, and supports direct access to NERSC computation and storage (for example, file system mounts) by JGI resources. OSCARS enables routine facility operations, large-scale transactions such as bulk data migration, and the "coupling" of geographically remote resources.

Earth System Grid. The Coupled Model Intercomparison Project / Phase 5 (CMIP5) uses the Earth System Grid (ESG) Federation for interoperable discovery, distribution, and analysis of large and complex data sets. There are two primary workflows: data replication and data distribution. The data replication workflow involves the transfer of many data sets between large storage systems at climate facilities. Data sets replicated are typically in high demand, and the replication process spreads load geographically across multiple data centers. The distribution workflow involves the discovery and download of specific data sets by scientists for analysis. Individual data flows are smaller, but there are many sets downloaded and the downloads can still be terabytes in size. *Both* workflows assume the ability to transfer data sets from tens to hundreds of terabytes in size – reliably – around the globe.

3. Facility S.W.O.T. and Dependencies

3.1, 3.2 Strengths, Weaknesses, Opportunities, Threats

Strengths	**Weaknesses**
• world-class staff • high capacity for innovation • culture of urgency • global reputation • new optical and dark-fiber assets • strong relationships (with Program Manager, Internet2, Labs)	• understaffed, oversubscribed • internal processes need modernization and attention • recruitment is slow, because facility is selective and competition is keen
Opportunities	**Threats**
• increase productivity of scientific user facilities through outreach • promote world-wide adoption of best-practice architectures for data mobility • transform network from black-box infrastructure to intelligent, programmable *instrument* • provide services and capabilities for non-DOE federal science partners	• talent-poaching from large, well-funded companies • budget cuts limit opportunities to advance mission • travel restrictions impair ability to lead and promote global change • operational activities crowd out innovation & strategic thinking

3.3 Risks, Dependencies, and Critical Gaps

Facility Risks

1. ESnet will **not be funded** to deliver the network capacity required to match SC's exponential growth in demand.

   As explained in **section 5.1** below, ESnet must be free of packet loss in order to avoid delaying large scientific data flows by a factor of 50 or more. Thanks to a $62M ARRA investment in new optical infrastructure, ESnet now operates a platform that can scale to meet exponential growth requirements of SC for the next 5-7 years, based on current traffic projections. However, accommodating this growth is not free – it requires purchasing optical and routing gear for every additional 100Gbps optical wave added to the network; these purchases in turn drive data center costs related to power, space and cooling, as well as additional effort costs. The **critical budgetary question for ESnet** is whether the declining cost-curve for 100Gbps components will offset the exponential rise in bandwidth required by SC. Because it would mean predicting the future of a market, there is no way to answer this question definitively. However, there are plausible scenarios in which ESnet's funding would need to increase significantly (by many millions of dollars) in order to accommodate growth. Now that the transition to 100Gbps technology is finished, ESnet will work actively with its optical vendors to build reasonable cost models for the next five years, and will brief the ASCR Facilities Director and ESnet Program Manager when these projections are complete.

2. ESnet will not continue to attract the world's **best Internet engineering talent**, or else lose critical team members to industry.

   The foundation for ESnet's success is an unparalleled team of Internet engineers and other contributors. Staff have been recruited from Google, Oracle, Sun, Wolfram Research, Level3 Networks, Internet2, UC Berkeley, NCSA, NERSC, ALS, JGI, BNL, LBNL, Ames Lab, LLNL, and many other prominent organizations. The facility is regarded as one of the best places to work in networking, and jobs at ESnet are coveted, in part because 'talent attracts talent.' The greatest threat to ESnet's talent pool is probably an extended period of venture-capital investment in networking start-ups; a period of substantial growth for highly-capitalized organizations that own big networks (Apple, Facebook, Amazon, Twitter); or a decline in ASCR funding that results in underfunding the growth of the facility. Since any of these scenarios is possible, ESnet leadership takes great care to motivate existing employees, hire wisely, refresh the talent pool, and perform succession planning. The general issue of retention is discussed in more detail in **section 7.3** below.

Dependencies and Needs

1. ESnet depends on **external investments**.

   Because networking is an "end-to-end" problem, ESnet depends on many other organizations (networks, labs, universities, and SC facilities) to make infrastructure and engineering investments required to advance SC missions. To cite an obvious example, ASCR supercomputing facilities must invest in equipment for optimizing wide-area data transport in order for ASCR to take maximum advantage of its new 100Gbps national network.[6] Similar infrastructure investments need to happen worldwide in order for global networks to enable and accelerate discovery. Reduced investments in infrastructure by labs, campuses, and peer universities will prevent maximum utilization of wide-area capabilities carefully planned and engineered by ESnet. The outcome here is not simply a function of funding; it's related, perhaps more fundamentally, to how effectively ESnet can lead the world's networks towards a shared architectural vision. See strategic initiative #4 (in **section 6.1**, below) for further discussion.

2. ESnet depends on continued innovation from networking vendors.

   ESnet is growing roughly twice as fast as the commercial Internet. Special ARRA funding enabled the organization to deploy the first 100Gbps network at continental scale in the world, but this upgrade happened "just in time" from the perspective of DOE science. In the future, ESnet will continue to be among the earliest adopters of new networking technology, and will depend on continued innovation at every network layer. Disruptive innovations in optical and packet routing technology (for example, terabit waves or silicon-photonic integration) have the potential to make it cheaper to build high-performance networks at scale. ESnet depends on vendors investing in such innovation, in order to have a hope of scaling economically. It would be suboptimal for DOE, and certainly more expensive, if ESnet's traffic growth required it to diverge too far from commercially available architectures; this dynamic is familiar in the exascale program.

4. Science and Community Needs

4.1 User Community Description

As discussed in **section 5.1** below, it's more complicated to formulate a definition of "user" for a networking facility than a computing center, as a result of the Internet's different service model (*variable, best-effort, and ad-hoc*, rather than *fixed and reserved*). In some sense, every SC-funded scientist is an ESnet user, because the networking facility is fundamental, interconnecting all others, and providing access non-DOE collaborators as well. Unfortunately, this line of reasoning does not produce a satisfactory or well-bounded concept of "user" for comparison with other SC facilities. After consulting with ESnet's

[6] Fortunately, all three computing facilities are doing exactly that.

Program Manager, the facility has developed the user categories described below. These categories encompass just a small subset of DOE-funded scientists who benefit from ESnet: those who have network research objectives; and those with high data volume, network-intensive science workflow requirements who have requested dedicated bandwidth services. Proposed techniques for measuring ESnet's impact on scientists outside this small subset of users are described in **section 6.1**.

1) On-Site Users. These are network researchers granted physical access to ESnet's dark-fiber or network testbed resources, based on the nature of their research goals. These projects are peer reviewed, as described in **section 6.1**.

2) Remote Users. This category includes users who receive allocations of dedicated bandwidth capacity on ESnet itself, and users who are granted access to ESnet testbed resources.

3) Data Users. These are researchers who download or otherwise gain legitimate access to data sets collected by ESnet as part of an approved and peer-reviewed research project. (ESnet collects a great deal of monitoring data, related to utilization, energy consumption, bulk data transfer, routing, IPv6, flow dynamics, performance monitoring. Much of this data is of interest to network researchers, and all data is either publicly available, or available in anonymized form to approved researchers.)

4.2 User Community Engagement

Although a few large-scale collaborations such as ATLAS have the staff and resources necessary to make excellent use of advanced networking capabilities, most scientific collaborations are smaller and need assistance to achieve the same level of benefit. As an overarching goal, ESnet wants to ensure that its services are well understood, accessible, and easy to use for any scientist or collaboration.

Because this goal is so critical for addressing the "extreme data" challenges confronting the Office of Science, ESnet has defined two strategic initiatives in the area of user community engagement, partnership, and outreach (see **section 6.1**, which also describes ESnet's historical community engagement activities). The goal for *all* outreach activity – indeed the goal for everything ESnet does – is to achieve its vision of a world in which scientific progress is completely unconstrained by the physical location of instruments, people, computational resources, or data.

ESnet's new outreach initiatives will supplement engagement activities that are already vigorous. In a typical year, ESnet staff participate in more than 100 community and stakeholder meetings, giving almost as many technical talks, keynotes, tutorials, and webinars. ESnet also formally participates in multiple community events, including the Joint Techs workshop (the most important global gathering of research and education networkers, which it co-sponsors); the ESnet Site Coordinators' Committee (ESCC) meeting; the Large Hadron Collider Optical Private Network (LHCOPN) and Open Network Environment (LCHONE) meetings; the US ATLAS Throughput Working Group; the National User Facility Organization (NUFO); the GreenTouch consortium; TERENA (The Trans-

European Research and Education Networking Association) conferences; various Internet2 meetings; the Open Grid Form (OGF) standards body; the Global Lambda Integrated Facility (GLIF); and the Open Science Grid (OSG), among others.

4.3 Requirements Gathering

4.3.1 Data Gathering

ESnet's architecture and capabilities are tailored for its users, whose needs are derived through a formal requirements-gathering process. These twice-yearly reviews of network requirements focus on each of the six program offices funded by SC: Advanced Scientific Computing Research, Basic Energy Sciences, Biological and Environmental Research, Fusion Energy Sciences, High Energy Physics, and Nuclear Physics. The purpose of these reviews is to characterize near-term, medium-term, and long-term network requirements for the science supported by each program office. Reviews focus on *scientific case studies*, rather than information technology. By first understanding how disciplines and collaborations produce knowledge, ESnet can derive network requirements more accurately. This highly successful approach has served as a model for other research and education organizations, and for at least one other DOE facility (NERSC).

During each review, participants talk through written case studies prepared in advance, rather than presenting slides. A roundtable format is preferred, and anyone in the room can ask clarifying questions. Often the most useful conversation is opportunistic, when researchers in different collaborations hear and respond to each others' process-of-science explanations.

ESnet also sponsors (or participates in) requirements assessment outside the regular cycle of SC program office reviews. Recent examples include organizing a workshop to explore network requirements for the Belle-II experiment; and co-chairing the effort to gather 10-year network requirements for the Community Planning ("Snowmass") process within the High Energy Physics community.

4.3.2 Data Forecasting

ESnet's primary basis for modeling aggregate future data requirements is a consistent 23-year historical pattern: since 1990, total traffic accepted by the facility has increased tenfold every 48 months (see **figure 1**). ESnet has no reason to believe this rate of increase will change, because the underlying exponential technology drivers continue to operate. In addition, ESnet uses the programmatic reviews described above to discover, document, and monitor any new requirements.

While overall backbone utilization will likely continue on the current growth curve, trans-Atlantic, trans-Pacific, and commercial peerings may be on different paths, and could exhibit step-function changes. For example, the Belle-II experiment in Japan is projecting a 30 Gbps increase in trans-Pacific network use by 2017; the international ITER facility in France will cause a substantial increase in trans-Atlantic traffic; and future improvements of LHC luminosity will drive even higher trans-Atlantic data rates. ESnet also uses its requirements review process to identify capabilities needed by science experiments that are not obvious from aggregate traffic volume statistics. These include, for example, the ability to provide guaranteed bandwidth as a service through programmatic APIs for grid middleware components; or the ability to support international non-TCP, fixed data-rate protocols such as RDMA no later than 2013. By combining extrapolations based on historical data with capability requirements expressed by the science collaborations, ESnet can build and run a high-performance network capable of providing both the capacity and the capability required to support DOE's data-intensive science portfolio.

5. Center Operational Priorities

5.1 Center Priorities

Before discussing operational priorities for ESnet, it's helpful to note an important distinction between ASCR's computing and networking facilities.

The benefits of computing facilities have traditionally accrued to a limited number of PIs who received finite allocations of a scarce resource, through a process of competitive peer review. The service allocation has been **fixed and reserved**.[7] Other models are available, however. For example, users at NERSC who explore large data sets through web-based Science Gateways receive **variable, best-effort, ad-hoc** services. Although most services available at computing facilities still fall into the *fixed and reserved* category, the advent of rich web interfaces for exploring data may contribute to a parallel rise in demand for *variable, best-effort, ad-hoc* services. Neither service model is superior to the other, and both advance the ASCR mission. However, it's more complicated to formulate a definition of "user" for the second service category.

In the case of ESnet, the two models evolved in the **opposite order**. When it was founded 26 years ago, ESnet quite naturally inherited the service model for the newly launched global Internet, which was *variable, best-effort, and ad hoc*. In fact, this model was a key innovation of the Internet architecture, distinguishing it sharply from circuit-reservation systems (most notably public telephony) that came before. Over time, the science data flows carried by ESnet began to diverge from commodity Internet flows in their scale and service requirements. Starting ten years ago, ESnet pioneered, developed, and deployed a *fixed and reserved* networking model in response to the changing nature of its traffic. The software ESnet developed for this purpose is now used by 40 networks worldwide, and the *fixed and reserved* service model has been adopted by even more organizations. Today, roughly half of ESnet's traffic is handled through the *fixed and reserved* model, and it's reasonable to believe the ratio will grow over time.

With this context in mind, it should be easier to understand ESnet's operational priorities.

---

[7] Note the special definition of *reserved*. A fixed allocation is reserved for a named PI, but the allocation does not guarantee that any particular job will start or complete at a fixed time *t*; this is not, in other words, anything like a hotel or dinner reservation.

First, the facility must allocate sufficient resources to accommodate exponential growth in demand for both service models. If ESnet fails to stay ahead of the demand for both services, the result is undesirable *packet loss*. A small rate of packet loss (less than .01%) in a national-scale network causes hugely disproportionate effects (>50x reduction) on data transfer speeds for large flows, due to characteristics of TCP, the dominant Internet transport protocol.[8] For this reason, ESnet is designed to be *lossless,* to the extent possible.[9] Commercial networks such as Comcast or AT&T do not have the same design objectives, because the flows they support have much lower throughput and travel much shorter distances than DOE science flows. If ESnet failed to meet the exponential bandwidth demands of DOE science, or did so in a way that introduced packet loss, the user experience for DOE scientists would resemble the experience of watching YouTube in a highly congested wireless environment. Such an outcome would be detrimental for the productivity of every collaboration funded by DOE, both inside and outside the national lab system, and would significantly impair the SC mission.

Second, ESnet must dedicate enough of its underlying network capacity to meet the demand for *fixed and reserved* service. Massive science data flows are better served, generally, when migrated from the *variable* to the *fixed and reserved* model. (This generalization is less valid in a well-provisioned, lossless network such as ESnet, but the *fixed and reserved* model is still important for meeting resiliency and security requirements of many science collaborations, and for delivering network services that are not compatible with a best-effort routed architecture.) Network reservations are a finite resource, and it's possible to "oversubscribe" a network path. Up to this point, ESnet has added capacity rapidly enough to meet all reservation requests; further, utilization of network reservations has been high. But because these trends may not continue, ESnet will establish a **user group** to provide input and guidance on the question of allocation strategies for all Facility resources.

In addition to network capacity, ESnet has **other resources** that require allocation. The most important of these is access to a national-scale, multi-layer network testbed (in many respects unique in the world), designed to promote the research and innovation necessary for advancing the missions of ESnet and ASCR. The ESnet testbed is available to all DOE-funded network researchers, and accepts proposals every six months from researchers in industry, government, and academia. Proposals are peer-reviewed by a committee of lab, university, and industry networking experts. This review includes assessments of operational readiness in addition to merit, although proposals by projects funded by DOE ASCR are reviewed for operational readiness only. The architecture and service model for

---

[8] In a recent test, ESnet measured the effects of a 0.0046% loss (1 packet out of 22000) on data transfer rates at various distances. At DOE-relevant distances (for example, NERSC to ANL) and DOE-relevant transfer rates (10Gbps), this tiny rate of loss decreased data transfer times by 80x, an effect that renders the network virtually useless for the purposes of large-scale science.

[9] Creating a loss-free network requires careful engineering, and a combination of tactics: there must be adequate network capacity to begin with, deep and properly-configured packet buffers in core network components, and a test infrastructure that can automatically and continually verify the health (and implicitly the proper configuration) of all equipment. For this latter purpose, ESnet co-developed the perfSONAR software toolkit, which is now widely used in the research networking community.

the ESnet testbed are currently evolving to accommodate resources at other ASCR facilities and labs, and indeed outside the DOE complex (see **section 6.1**); this new model will be defined in a report due to ASCR at the end of April.

One final operational priority for ESnet is **high availability**. The importance of this requirement highlights another respect in which networking and computing facilities differ. ESnet normally achieves four-nines of availability to end sites (99.99%), a level of service which requires great engineering effort, operational discipline, and an appropriate level of funding to achieve. The requirement for high availability is a consequence the critical role Internet services now play in enabling and accelerating modern science. Even a 30-second outage of network connectivity to a national lab is conspicuous and disruptive. The recent $62M ARRA investment has greatly improved the underlying resiliency of ESnet and the national labs (especially at the optical layer), just as it has reduced the cost of adding network capacity.

5.2 Inter-Center and Inter-Facility Collaboration

Because networking is an "end-to-end problem," better collaboration with ASCR computing centers and user facilities is a major theme of this strategic plan. Approximately half of the initiatives described in **section 6** have increased collaboration as a primary objective or beneficial side effect.

One year ago, ESnet hired a Lead for Partnership and Outreach, the first staff member entirely dedicated to the function of outreach in the organization's history. The outreach team will grow by 1.5 FTE in the coming year, and will initially prioritize engagement with user facilities; for more detail, see **section 6.1**.

5.3 Connections with Berkeley Lab Plans

The broadest strategic goal for Berkeley Lab is **Carbon Cycle 2.0**, designed to accelerate the development of a carbon-neutral global energy system. This goal is supported by three strategic pillars: nanoscience, biology, and computing. The connection between ESnet and overall Berkeley Lab strategy is most obvious in the "computing pillar", which includes three components: simulation/modeling, extreme data, and ESnet itself. Each of these components is critical in the pursuit of broader Lab strategy, especially as DOE science enters a new era of "extreme-scale" data and computation. Facilities such as the ALS and (in the future) the NGLS require ever more substantial computing and networking resources for their discovery processes, and as that happens, the strategic importance of Berkeley Lab's computing pillar increases. Finally, ESnet plans to support and enable NERSC's new data initiative (see NERSC Strategic Objective #2) by providing the necessary network capacity, capabilities, and architectural consultation.[10]

---

[10] ESnet would, of course, eagerly provide the same capabilities to any SC Facility (ASCR or otherwise) requesting them, and will seek to do exactly that as part of its new outreach initiative described in **section 6.1** below.

6. Center Strategy for the Future

The **core strategic goal** for ESnet is delivery of highly-reliable data transport capabilities, optimized for the requirements of large-scale science, and continually scaling to accommodate the exponential traffic growth of DOE science missions. If ESnet fails to achieve this central goal, the facility cannot succeed in its mission. Having said that, the core goal is necessary but not sufficient. In addition, ESnet needs to develop the capacity to:

- **assess and rapidly improve** how collaborations use advanced networking; and
- **influence** DOE facilities, labs, campuses, and networks around the world to optimize their infrastructures for end-to-end data movement.

All of these activities are enabled by an **overarching strategy for innovation**: ESnet will pioneer, develop, and sustain the architectures, protocols, and software necessary to achieve its vision. Because ESnet's growth is so rapid, and the requirements of its science traffic so different from those of commercial Internet flows, an innovation strategy is key to the facility's success, and in turn to the success of SC science missions.

6.1 Major Initiatives

All initiatives discussed in this section are organized and grouped within the rubric of ESnet's four strategic goals, described in **sections 2.2** and **6**, and highlighted within blue boxes. The initiatives are selected (as the ASCR guidance directs) to focus on a time-scale of the next several years. Limitations of space require considerable compression in their description.

> Strategic Goal
>
> *Ensure that ESnet continues to provide highly reliable data transport capabilities, optimized for the requirements of large-scale science, and scaling to accommodate the exponential traffic growth of DOE science missions.*

On any given day, a large fraction of ESnet staff (from Internet engineers to deployment specialists; tools developers to systems administrators) is busy pursuing this central goal. Supporting activities vary widely, and include technology evaluation, vendor engagement, applied research, growth modeling, test and prototyping, financial analysis, procurement, and deployment.

Although this document does not describe an initiative in support of the above goal, **section 5.1** articulates the justification for maintaining a *lossless* network core; **section 3.3** describes the risks of failure; and **section 6.3** discusses the technical challenges, inflection points, and likely strategies associated with continuing to scale ESnet capacity to meet SC mission needs.

Strategic Goal

In response to the challenges posed by the scientific data revolution, develop the ability to assess and rapidly improve how collaborations use high-performance network capabilities.

Initiative 1: increase partnership, engagement, and outreach

Goal, Motivation, and Strategy for Producing Change

A world-class network is of little use, if collaborations and facilities don't have the information they need to benefit from its capabilities. The initiative described below aims to increase and track ESnet's activities in partnership, engagement, and outreach – with a goal of maximizing the ability of SC-funded scientists to use ESnet in accelerating their discovery workflows. This outreach initiative is particularly urgent, since ESnet's program requirements reviews and interactions with facility staff clearly indicate a looming data-intensity challenge (for analysis of "extreme data" as an ESnet driver, see **sections 2.3** and **4.2** above).

In the past, the ESnet outreach effort has been confined to three narrow activities:

- gathering requirements from SC Offices on a regular basis
- working with individual PIs on an ad-hoc basis
- engaging with lab networking staff, through regular meetings of the ESnet Site Coordinators Committee (ESCC) as well as more informally

Each of these activities is important, and each will continue. But as a coordinated strategy, they lack any method for making contact with a vital set of ESnet constituents: user facilities and scientific collaborations. The initiative described below provides a way of closing this gap, by developing outreach techniques tailored specifically for the needs of facilities and collaborations. The initiative represents an important – and for ESnet, unprecedented – expansion of activity for the purpose of direct engagement.

The strategy for producing this change has two forks:

- creation of rapid-response, data-mobility "tiger teams"
- implementation and tracking (using state-of-the-art customer relationship management tools) of partnership, engagement, and outreach activities targeted specifically for facilities and collaborations

Data-mobility tiger teams will be flexible in composition, but will always include ESnet engineering, outreach, and project management personnel, who will work in partnership with staff from relevant facilities, scientific collaborations, and labs. The objective of each tiger team will be clearly defined and narrowly scoped, with the aim of solving a significant data-mobility problem. The term "tiger team"can have negative connotations within DOE, but the intent here is to assemble:

> …undomesticated and uninhibited technical specialists, selected for their experience, energy, and imagination, and assigned to track down relentlessly every possible source of fault.[11]

Though the tiger team model is reactive, and will not scale broadly, it has very significant advantages in the short and medium term. First, it generates successful case studies, as well as productive relationships, which can be reused in broader outreach efforts. Second, it produces value for scientists whose problems are solved; there have been two recent examples of such activity involving ALS, NERSC and ESnet staff in the past year. Finally, this technique was discussed at the most recent ESnet Site Coordinators meeting, and was broadly endorsed by Lab representatives.

The second fork of this initiative (*implementation and tracking of partnership, engagement, and outreach activities*) is more strategic in nature, because it promises to deliver a scalable and comprehensive outreach solution. ESnet took the step of hiring a Lead for Partnerships and Outreach over a year ago, and the Partnerships Lead has presented her outreach strategy to ESnet's Program Manager and Berkeley Lab's ALD for Computing Sciences. The primary goals of this effort are to:

- build awareness of ESnet in communities that have not made optimal use of advanced networking (especially small to mid-sized collaborations)
- target facilities used by these collaborations; partner with user-support and outreach staff to create new communications channels for information and possibilities for education
- develop a set of early adopters (through the tiger team process, for example) to act as advocates for these communities
- partner with complementary tool providers such as Globus Online, to expand the reach and impact of ESnet outreach
- add content on www.es.net dedicated to the needs of scientific users
- monitor the impact of outreach *at the network layer* by benchmarking and measuring traffic in and out of facilities over time
- implement customer relationship management (CRM) technology to capture user data, gauge outreach effectiveness, measure satisfaction, and target specific communities with focused and customized information
- in the longer term, ensure that every DOE user facility has the tools, knowledge and resources it needs to inform scientists about ESnet capabilities and direct users to ESnet when appropriate for problem solving and support

Required Resources

The most important resources required for this initiative are human. ESnet has hired a Lead for Partnerships and Outreach; is recruiting for a half-time Outreach Coordinator; and will recruit for a technically oriented Outreach Engineer in the next two months, to support tiger team and other activities. In addition, approximately .5 FTE is allocated from the Network Engineering Group to Outreach on a steady-state basis. It's likely that the

---

[11] J. R. Dempsey, W. A. Davis, A. S. Crossfield, and Walter C. Williams, "Program Management in Design and Development."

Partnerships and Outreach team will grow over time, and may double again in the next few years, assuming outreach efforts are demonstrably valuable. The notion of partnership also implies commitment from other organizations; one example is the recently announced partnership between ESnet and Globus Online. Although ESnet does not have the power to commit funds or resources of other organizations, the facility will aim to make outreach partnerships (such as tiger teams for data mobility) so useful and attractive that outside organizations have a compelling motivation for making these commitments with ESnet.

Risks

The primary risk of this initiative is failure to advance ESnet's strategic goal of improving the way scientific collaborations use advanced network capabilities. The customer relationship management tools described above will reduce this risk.

> Strategic Goal
>
> *Pioneer, develop, and sustain the architectures, protocols, and software necessary to achieve our vision.*

Initiative 2: real-time discovery, correlation, and visualization tools

Goal, Motivation, and Strategy for Producing Change

A perennial frustration (and deep technical challenge) in networking is the difficulty of obtaining useful information about "bits on the wire". It is trivial to count those bits, but much harder to extract meaning – for example, to assess the real-time impact of high-performance networks on a particular global research collaboration. In the initiative described below, ESnet aims to lead the world in developing discovery and visualization tools to identify significant network flows in real time, correlate them with scientific activity, make accurate deductions, and perform useful interventions.

For the past two years, ESnet has built the capacity to develop world-class tools. This initiative began with the creation of a dedicated ESnet Tools Team, currently a group of two: an experienced network and software engineer, plus an experienced web designer and developer. This team has been focused on the functionality of ESnet's customer portal, MyESnet (http://my.es.net). This portal provides feature-rich tools customized for networking staff (and network-savvy scientists) at each ESnet site, and also designed for use by program management within ASCR. MyESnet is currently the most polished and useful operational interface for any research network in the world. However, this functionality is only the beginning of what the Tools Team can accomplish.

Now that ESnet has developed the necessary skills (data analytics and visualization for network-based information), it can build tools of greater impact for scientific collaborations. The foundation for this new activity is creation of a global database with community-curated metadata documenting cyberinfrastructure resources relevant to science: supercomputers, clusters, data repositories, mass storage systems, archival storage, data transfer notes, science gateways, performance monitoring nodes, and network service interfaces. In the words of one member of the Tools Team, the database will be a "who's who of IP addresses." Building on this project, the Tools Team would use nearly real-time network flow information to create correlation and visualization engines of relevance to SC collaborations, or of operational relevance to ESnet. The opportunities for extracting real-time intelligence from network include:

- diagnosing performance problem with accuracy and precision
- recommending corrective action, again with accuracy and precision
- dropping large flow into optical circuits for the purpose of maximizing utilization of expensive infrastructure
- applying appropriate security policy when necessary
- analyzing traffic with fine-grained discrimination of science impact
- identifying opportunities for specific facilities, or specific collaborations, to make better use of ESnet resources

<u>Required Resources</u>
The size of the ESnet Tools Team will need to grow from two FTE to four at minimum in order to pursue the opportunities described above, but this increase is more than justified by the prospect of obtaining the functionality described. Developing these tools will also require consultation and collaboration with the network research community; ASCR-funded researchers have, for example, already studied statistical methods for identifying probable big data flows based on historical data transfer records.

<u>Risks</u>
The primary risk associated with this initiative is under-estimating the effort required to achieve the goals described, and in particular underestimating the difficulty of establishing a community-curated repository for information about global cyberinfrastructure resources. Careful scoping and watchful management will mitigate both risks.

Initiative 3: next-generation network testbed

<u>Goal, Motivation, and Strategy for Producing Change</u>
For the foreseeable future, ESnet will continue to be among the earliest adopters of network technology in the world – the result of unrelenting exponential growth, combined with demanding service requirements for science data flows. For precisely the same reasons, the organization's focus on applied research, development, and innovation (and collaboration with ASCR-funded network researchers) will continue and increase.

During most of its history, ESnet has not had access to a realistic, national-scale platform for research and innovation. However, the facility's $62M ARRA-funded Advanced Networking Initiative (ANI) included dedicated funds to support a national network testbed, and that testbed proved to be extremely useful for ESnet and the broader network research community.[12] The goal of the testbed was to provide a realistic environment for researchers and the private sector to experiment with new technologies, protocols, and applications at 100Gbps. The topics explored included non-TCP transport; multi-layer and

[12] A closeout report for the ANI Testbed was submitted to DOE ASCR in December 2012.

multi-domain architectures; protocol scaling; component testing for high performance; protection and recovery; automated flow classification; and high-throughput middleware.

Based on the success and utility of the original ANI testbed, ESnet has proposed and ASCR has agreed to continue funding it, although the architecture and service model will evolve considerably. In fact, ASCR has requested (and ESnet is preparing) a report describing its future vision for a network testbed, due in April 2013. Table 1 briefly summarizes major differences between the original and future testbed architectures and service models.

	ANI Testbed	**Future Testbed**
Who can use it?	DOE, .edu, industry.	DOE, .edu, industry.
Possible topologies?	Fixed and constrained by availability of physical circuits.	Programmable for each collaboration, through a combination of dedicated 100Gbps waves and OSCARS circuits on the spare capacity of the new 100Gbps ESnet footprint.
Attached resources?	Only those provided by ESnet.	Any resources provided on a temporary or permanent basis by ESnet, DOE supercomputing centers, labs, user facilities, universities, exchange points, or other network testbeds.
Bare-metal access to network components and systems?	Yes.	Yes (on ESnet-managed resources).
OpenFlow capability	Yes, with 1Gbps NEC switches.	Yes, with multi-vendor deployment on national footprint; optical OpenFlow possible.
Connectivity to other national-scale testbeds such as GENI, and those in Europe?	No.	Yes, with peering at the Starlight exchange point in Chicago.
Access to 13,000-mile dark fiber footprint for disruptive optical research?	Yes, but researchers must supply optical components and cover associated costs.	Yes, but researchers must supply optical components and cover associated costs.

Table 1. Major differences between ANI Testbed and Future Network Testbed.

Required Resources

The ESnet staff requirement for the testbed is only 1 FTE, which includes .4 FTE for systems support/upgrades, .4 FTE to provide support to experimenters, .1 FTE for network engineering support, and .1 FTE to manage the proposal process and other management tasks. The testbed itself requires a dedicated 100Gbps wave, and a minimum of 40Gbps spare capacity to several major DOE labs to perform relevant experiments. OpenFlow components require 10Gbps of spare capacity between a number of ESnet hubs. The purpose of the testbed is to perform research on next-generation architectures, so testbed hosts will need to be upgraded/replaced every 2-3 years. Finally, a 100Gbps peering at an open exchange point in Chicago is needed to federate the testbed with other networks.

Risks

Many of the operational risks of managing a testbed for the network research community have already been mitigated in the new plan; in particular, the risk of inexperienced researchers requiring disproportionate assistance will be addressed by an operational readiness review for all proposals.

Initiative 4: intelligent network services

Goal, Motivation, and Strategy for Producing Change

The power of the Internet is largely attributable to the success of well-documented and widely deployed network protocols. Some of these, like TCP/IP, are indispensable for any Internet communication. Others, like HTTP, are vital for particular applications or services.

Currently, there is a widely publicized effort to build new protocols for interacting with network components directly – through simple, open interfaces. This model is called "software-defined networking" (SDN), and OpenFlow is the most well-known SDN protocol. If the promise of OpenFlow can be realized, it will be possible to program complex, multi-vendor networks such as ESnet, optimizing them for specific uses, such as science data transport. For example, OpenFlow could be used to detect science flows that require special handling, and move them into dedicated optical paths – which is to say, OpenFlow could play an enabling role in **Initiative 2** for real-time tools. As explained in **section 5.1** above, science applications designed for moving large data sets can be poorly served by the current Internet architecture (specifically, by the combination of statistically-multiplexed, best-effort IP packet delivery and TCP, which is highly sensitive to packet loss).

Although any fresh wave of technology generates hype and hyperbole that must be discounted, ESnet firmly believes that the Internet is passing through an inflection point of programmability, and that OpenFlow and related protocols hold great promise in advancing the facility's vision of scientific discovery unconstrained by geography. Certain scientific communities (notably High-Energy Physics, but there may be others) seem willing and able to embrace intelligent network services soon, for the purpose of ensuring more consistent and predictable data transfer rates.[13]

The initiative proposed here is threefold. First, ESnet will continue to lead protocol standardization efforts which advance the cause of DOE science, and particularly in the area of building interoperable interfaces for managing end-to-end bandwidth reservations

---

13 ESnet recently participated in a workshop on this topic, which brought together technologists from national research networks with application and middleware developers from the LHC experiments.

(so-called dedicated circuits), which are described in **section 5.1** above.[14] In pursuit of this goal, ESnet's Chief Technologist has been co-chairing the effort to standardize the Network Services Interface (NSI), an SDN protocol through which applications and middleware components can request dedicated circuits across multiple administrative domains – and eventually, get performance information, configure overlays, and obtain other services.

In addition to leading standardization work to support its science mission, ESnet will prototype and deploy software interfaces in partnership with at least two highly distributed collaborations, very likely ATLAS and KBase, though others are possible. Initially using the next-generation testbed described in **Initiative 3**, ESnet will expose web-services based SDN interfaces for each collaboration's middleware stack, with a focus on dedicated circuits.

Finally, ESnet will expand the intelligent network services (including NSI and OpenFlow) supported by its OSCARS software – the leading network middleware for building and managing dedicated network circuits.

<u>Required Resources</u>
Pursuit of this initiative will require one or two additional software developer FTEs, over a period of two years. In addition, ESnet will work closely with other networks and projects attempting similar projects, especially the NSF-funded ANSE initiative (PI Harvey Newman, Caltech).

<u>Risks</u>
Principal risks with this initiative: effort could be wasted; networks might not be evolving into programmable systems; or if so, innovations might add minimal value within common scientific data workflows. Such risks are endemic to all activities that involve research and innovation. The technical consensus of numerous ESnet engineers is that SDN paradigms are transformative, helpful to large-scale science, and here to stay.

<u>Strategic Goal</u>

Influence DOE facilities, labs, campuses, and networks around the world to optimize their infrastructures for end-to-end data movement.

[14] This is not unexpected, since ESnet was a pioneer in this area, and uses dedicated circuits to carry about half of its data flows.

Initiative 5: increase worldwide deployment of best-practice architectures

Goal, Motivation, and Strategy for Producing Change

Consider that:

1. 80% of ESnet's traffic leaves the DOE complex.
2. The remaining 20% crosses multiple administrative boundaries (for example: ALS -> LBLnet -> ESnet -> ANL -> ALCF), because networking is an end-to-end service.

These two facts have wide-ranging consequences, and they suggest another way in which ASCR's networking and computing facilities differ. ESnet's success is tied to the health of its peer organizations (global and campus networks). This means that ESnet could attract the best staff, develop the best services, and foster the most visionary innovation – but all that would be minimally useful to DOE if the organization did not also lead the global networking community *strongly and effectively.* ESnet's obligation for global leadership can be derived from DOE's global traffic flows.

The initiative described below is an effort to increase worldwide deployment of best-practice end-to-end architectures, including ScienceDMZ,[15] performance monitoring, and guaranteed bandwidth services. Wide deployment of these architectures would benefit many scientific communities, but would especially help **climate, genomic, and photon** science – where ESnet's requirements workshops have revealed that an inability to transfer data hampers productivity.

ESnet has spent significant effort creating a knowledge base for the networking community (http://fasterdata.es.net); this site receives more hits each week than ESnet's homepage. Recently, the facility also helped initiate what may be a tipping point in deployment of science-friendly campus architectures in the United States: NSF published an important call for proposals ("CC-NIE") referencing ESnet's ScienceDMZ architecture and its knowledge-base, and is funding dozens of campuses to upgrade their infrastructures accordingly.

This is a very welcome development, but ESnet needs to continue its effort to promote best practices, especially in Europe and Asia – and also in the United States, among campuses and regional optical networks that have not yet been funded or inspired to make changes.

This initiative dovetails with **Initiative 1** ("increase partnership, engagement, and outreach") and will involve several of the same staff, but it's treated separately in this plan because the scope is largely technical, and the natural audience is network operators and funding organizations, not science facilities and users.

In order to increase worldwide deployment of best-practice architectures for data mobility, ESnet will pursue a four-part strategy:

1. Collaborate with funding agencies, promoting European and Asian equivalents of the recent NSF CC-NIE proposals.

[15] ScienceDMZ is a cyberinfrastructure design pattern pioneered by ESnet that optimizes campus infrastructure for high-speed data transfer.

2. Enhance the [fasterdata]() knowledge base with videos, additional how-to documentation, and more detailed discussion of security models for data-intensive science.
3. Participate in at least one tiger-team activity as described in **Initiative 1**, to document the clear performance advantages that result in ScienceDMZ implementation.
4. Continue to serve as global advocates and thought-leaders by keynoting conferences, organizing panels, sponsoring webinars, and publishing content on ScienceDMZ and related topics.

Required Resources

In order to step up activity in this area, it will probably be necessary to transition an existing network engineer into the role almost full-time, and backfill for his or her position in the Network Engineering Group.

Risks

The primary risk is failure to impact a substantial number of DOE-relevant networks and facilities, which in turn would perpetuate the status quo for smaller collaborations and non-HEP science. The risks of *not* attempting this effort are greater than the risks of attempting it.

6.2 Systems Software and Libraries

This section was probably intended primarily for computing centers. It's important to note, however, that the ability of global research networks to advance the SC mission depends on software tools, middleware, and applications – for which there are few reliable funding or sustainability models. The most critical examples for ESnet are OSCARS (to enable network reservations) and perfSONAR (for network diagnostics), both open-source projects heavily supported by the facility, and each requiring about 1 FTE for ongoing support. In general, support for useful but inherently non-commercial software tools is hard to secure outside of large, well-funded collaborations. Software will continue to play an important role in the facility's future, as intelligent network services drive integration of science applications with network middleware.

ESnet is keenly interested in the new funding model[16] being pursued for Globus Online, another important community tool for data mobility. But even if the "freemium" model sustains this particular project, that model is more appropriate for client-facing software than network middleware, which has a much smaller direct constituency.

6.3 Technology Strategy

ESnet's technology strategy is driven by the goals of scaling to accommodate exponential growth; maintaining a *lossless* network infrastructure; and making intelligent use of available network resources, as explained in **section 5.1** and **6.1**

[16] https://www.globusonline.org/news/globus-online-announces-big-data-sharing-service/

The $62M ARRA investment enabled ESnet to obtain, for the first time, access to terabits of optical capacity in a nationwide fiber transport system, which is shared with Internet2 (the research and education system interconnecting US universities). Without this investment, exponential traffic growth would certainly have lead to unsustainable yearly costs for ESnet. The upgrade from ESnet4 to ESnet5 was the most important technology transition in the facility's history, and it's unlikely a transition of similar magnitude will occur again before 2019.

Smooth Scaling

Between now and 2019, ESnet will deploy optical transponders, routers, and other network components to make incremental use of available optical spectrum, as the science drivers described in **section 2.4** continue to push an exponential demand curve. In the near term, ESnet will leverage its new optical platform to add capacity just in time, and at a cost that is proportional to the raw cost of optical transponders. These costs will follow a downward price curve; whether the shape of that curve completely offsets the exponential growth in traffic is a question of great importance for ESnet's budget, and it should be noted that additional optical waves create other costs as well (see **section 3.3** on both counts). For a rough chronology of upgrades that can be anticipated now, with current knowledge about growth patterns and roadmaps of optical vendors, see **section 6.5** below.

Disruptive Change

Disruptive improvements in technology are possible – and would be welcome, if affordable. One class of breakthroughs would be the advent of ultra-high-speed (400Gbps or 1Tbps), long-reach optical components. Even though such speeds have been demonstrated in the lab, deployment of these technologies will be expensive because of operational limitations regarding distance, non-ITU grid spectral boundaries, and complex modulation formats. Leading optical researchers predict a decline in the rate of growth of network speeds due to these limitations, unless new architectures that involve optical parallelism (for example, multi-core fibers) are adopted. Much more likely, if not certain, is the release of 200Gbps components, which ESnet will adopt as soon as they are cost-effective.

ESnet partners with vendors, other networks, and researchers to test, influence and support product innovations that promise to drive down costs, deliver needed capabilities for science, or produce disruptive change. In that spirit, ESnet is monitoring several major areas of research in high-performance networking. These include:

1. Gridless Optical Infrastructure

To support 400Gbps+ capacity per wavelength, optical infrastructure will require flexible allocation of spectral capacity, rather than being limiting it to the ITU-specified, 50Ghz spacing supported by all equipment today. Spectral blocking through fragmentation is one major concern that will need to be addressed before such technologies are adopted.

2. Packet-optical integration

Many vendors are pursing the integration of optical and routing components, a process that is almost certain to result in products that can be purchased in time to activate ESnet's dark fiber footprint in 2019. The promise of this approach is elimination of an expensive layer of hardware, by collapsing two separate network transport systems (optical and routing) into one.

3. Programmable network components
ESnet's science requirements reviews have documented the need for applications and science workflows to receive *predictable* network performance. This motivates the goal of building protocols and interfaces to support intelligent network services, as described in **section 6.1** above. ESnet will continue to work closely with leading vendors to develop and prototype architectures for network programmability.

<u>Software</u>
Most of ESnet's lasting innovations have come in the form of network architecture and network middleware. As explained in **section 6.2**, the role of software development will play an increasing role in the facility's future, as science requirements motivate the integration of applications and programmable network components.

<u>Partnerships for Innovation</u>
ESnet has participated in numerous successful technology partnerships and first-in-world demonstrations with networking and telecommunications firms (including two OpenFlow achievements in the past six months). Partners among vendors include Ciena, Infinera, Juniper, NEC and Brocade; among networks: Internet2, SURFnet, REANNZ; among research institutions: UC Berkeley, University of Virginia, Indiana University, University of Chicago, Stanford, and UC San Diego. In addition, ESnet is collaborating with NSF and will shortly engage with international funding agencies to promote best-practice architectures for data mobility in university campuses.

6.4 Power, Space, Cooling

The issues of power, space and cooling are less urgent for networking than computing facilities – in part due to the distributed nature of networks, and in part due to lower absolute power consumption.

However, many network collocation facilities were designed decades ago at modest power density. As networks carry more bits per second, with fixed energy costs per bit transferred, network providers are starting to take steps such as charging for metered power, increasing collocation rates, and limiting rack density. It's too early to project how these developments will affect ESnet's operating costs in the medium and long term, but it's certain that operating costs associated with power, space, and cooling will increase.

ESnet has been interested in network energy consumption for some time. It was the first national network to monitor and visualize energy consumption in real time, and it also sponsored an award-winning student research project on the topic of path computation and 'green' power sources.

Finally, ESnet is taking steps to reduce its own energy footprint by investing in efficient blade servers, implementing virtualization technology, and replacing numerous inefficient UPS components with an enterprise-grade system. Within two years, ESnet will move (along with NERSC) to the Computational Research and Theory building, which will be the world's most energy-efficient HPC data center.

6.5 Summary Roadmap and Timeline

2013/2014

- ESnet5 backbone requires new 100Gbps wavelengths as early as late 2013, with incremental capacity upgrades until 2019 or slightly beyond.
- Testbed roadmap established. Programmable testbed deployed.
- Intelligent network services integration with applications.

2014/2015

- Outreach and tiger teams for data mobility have a significant impact on productivity of science collaborations. Customer relationship management tools demonstrate utility of segmenting scientific communities and carefully tracking impacts.
- Intelligent network services provide more predictable use of the network for some important middleware components and applications.

2016

- Current router platform does not have capacity for expansion; significant purchase of new routers throughout network.

2017

- 1 Tbps capacity required (10x100Gbps) on all backbone routes.

2018

- Exponential data growth on ESnet continues (increasing 10x every 48 months), resulting in deployment of additional waves and >1.5 Tbps backbone. Preparation to expand onto unused, national fiber optic footprint commences.

2020

- ESnet5 optical fiber capacity is fully utilized; ESnet6 is deployed on nation-wide dark fiber footprint.

7. Human Resources

7.1 Current Workforce

ESnet staff members are inspired to work for the organization, and therefore loyal, which has resulted in low turnover. Even in the height of the dot.com boom, when Internet engineers were in short supply, the organization lost just two members of its core engineering team. That loyalty did contribute, however, to an aging workforce. In the past two years, six key staff members retired – with a combined ESnet work history of nearly 100 years. They were replaced by early and mid-career staff for the most part, all of whom gained valuable experience by participating in ESnet5 upgrade.

Culturally, ESnet has a strong work ethic. This leads to heroic performance in times of crisis, but it also exacerbates a chronic (and long-standing) problem of understaffing compared to peer organizations. The resulting overload will eventually lead to frustration and poor morale, if not addressed. Recently, and despite the demands of the ESnet5 project, there has been good progress in hiring, but the gap is not entirely closed (especially in the area of outreach, business operations, Internet engineering, and software engineering).

80% of ESnet staff members are engineers. The core skill required is Internet (or wide-area network) engineering, for which the supply of outstanding candidates is limited, especially at the optical layer, due to worldwide demand. Almost all of ESnet's Internet engineers have systems and security experience as well, and many are also skilled software developers. ESnet staff are also expected to have outstanding communication and collaboration skills.

For the most part, ESnet has been extremely successful in recruiting world-class talent. Staff have previously worked at Google, Oracle, Sun, Wolfram Research, Level3 Networks, Internet2, UC Berkeley, NCSA, NERSC, LBNL, JGI, BNL, Ames Lab, LLNL, and many other prominent organizations.

7.2 Future Workforce

Like many of the world's research networks, ESnet is devoting more effort to developing advanced services and software capabilities, which requires hiring proportionally more staff. Drivers for the change are explained in **section 6.1,** above. For an illustration of this large-scale trend in the context of several dozen European networking organizations, see **Figure 3**.

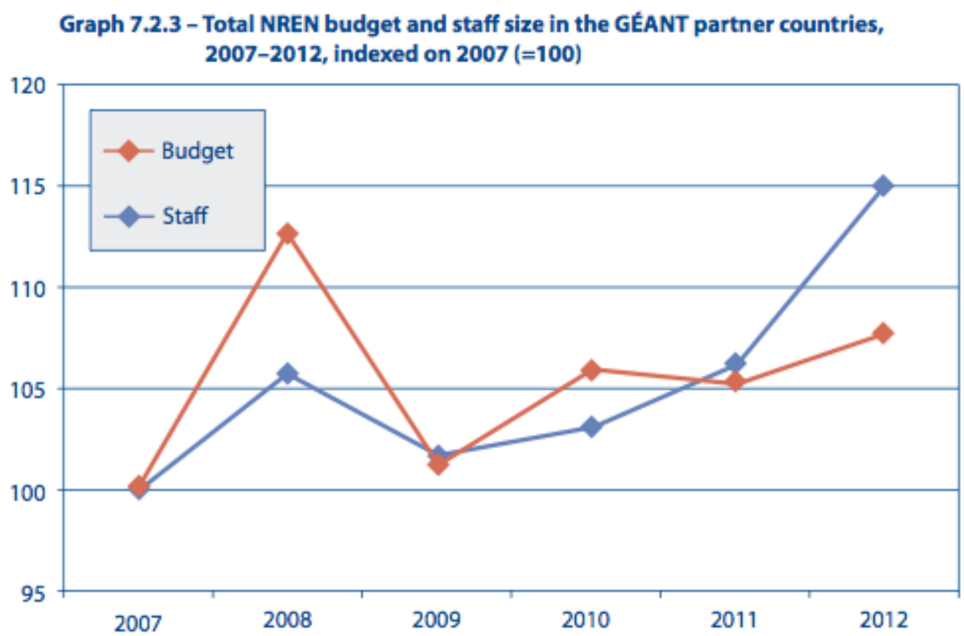


Figure 3. Source: TERENA Compendium 2012, www.terena.org/compendium.

In addition to this overarching trend, staff increases will be necessary to support exponential traffic growth, trans-Atlantic network extension, and possible trans-Pacific. The facility is in the process of modeling staff requirements driven by network scale and complexity, as part of the larger project of projecting 5-7 year costs (see **section 3.3.1**).

7.3 Obstacles and Strategies

<u>Major Obstacle</u>
Persistent shortage of the most skilled and qualified engineers, especially:

- experienced Internet engineers with a background in science, research networking, or optical technology
- engineers with superb communication and collaboration skills
- software developers and web app developers familiar with networking
- engineers who are women

<u>Strategies</u>

- sponsor 5-6 student internships per summer
- ensure that facility Director is personally involved in strategic recruitments, making calls to promising candidates
- ensure that facility Director is "always recruiting" in public presentations, by stressing the satisfactions of ESnet's mission and talent pool
- maximize use of social networking in recruitment
- build from within, giving promising early-career staff plenty of responsibility and interesting projects
- fight back when Silicon Valley tries to recruit key staff (by listening carefully, identifying the source of temptation, and trying to construct a local remedy)
- actively participate in high-profile conferences such as SC and NANOG, to identify future ESnet staff
- seek assistance at the lab level to increase representation of women on engineering teams
    - most recent hires at the leadership layer have been women, but the pipeline of *Internet engineers* who are women is disappointing, industry-wide
    - hiring students and targeted recruitment appear to be successful strategies

8. Conclusion

This document is a roadmap for achieving ESnet's vision of the future. In that future, scientific progress will be completely unconstrained by the physical location of instruments, people, computational resources, or data. In addition, collaborations at every scale and in every scientific domain will have the information and tools they need to achieve maximum benefit from global networks and emerging network capabilities.

In the next decade, ESnet is very well-positioned to achieve this vision. Its key driver is data intensity across the spectrum of SC science, and its key asset is superb engineering staff. As one reviewer commented in the most recent Operational Assessment Report for ESnet, "the entire staff conscientiously and continually lead their field."

Because science data flows have exacting requirements and continue to grow exponentially, ESnet will increase activities in applied research, development, and innovation. In order to avoid a technology roadmap that diverges substantially from the industry norm, ESnet will engage and collaborate with vendors. In order to assure that its activities enable and accelerate discovery, ESnet will build a strong outreach program.

Most network traffic on ESnet originates or terminates *outside* the DOE complex. For this reason, ESnet cannot succeed unless it fosters partnerships and pioneers technologies necessary to ensure that architectural changes occur nationally and internationally. ESnet's mission requires it to innovate, influence, and lead on a global scale.

The Office of Science invests $5B annually in supporting transformational research and operating world-class facilities throughout the United States. ESnet is the essential "circulatory system" for this investment. When ESnet's data-transport capabilities are interrupted or even slightly impaired (see **section 5.1**), the productivity of every collaboration funded by SC suffers. Although the Internet is often taken for granted, the health and success of DOE's networking facility should not be. SC cannot thrive without a strong and constantly-innovating Energy Sciences Network.

Appendix 1: Process for Developing this Strategic Plan

This strategic plan has been a collective effort. The facility solicited broad input from its staff in several ways: a staff-wide retreat; focused meetings; an online 'virtual white board' for ideas and brainstorming; and circulation of drafts for comment.

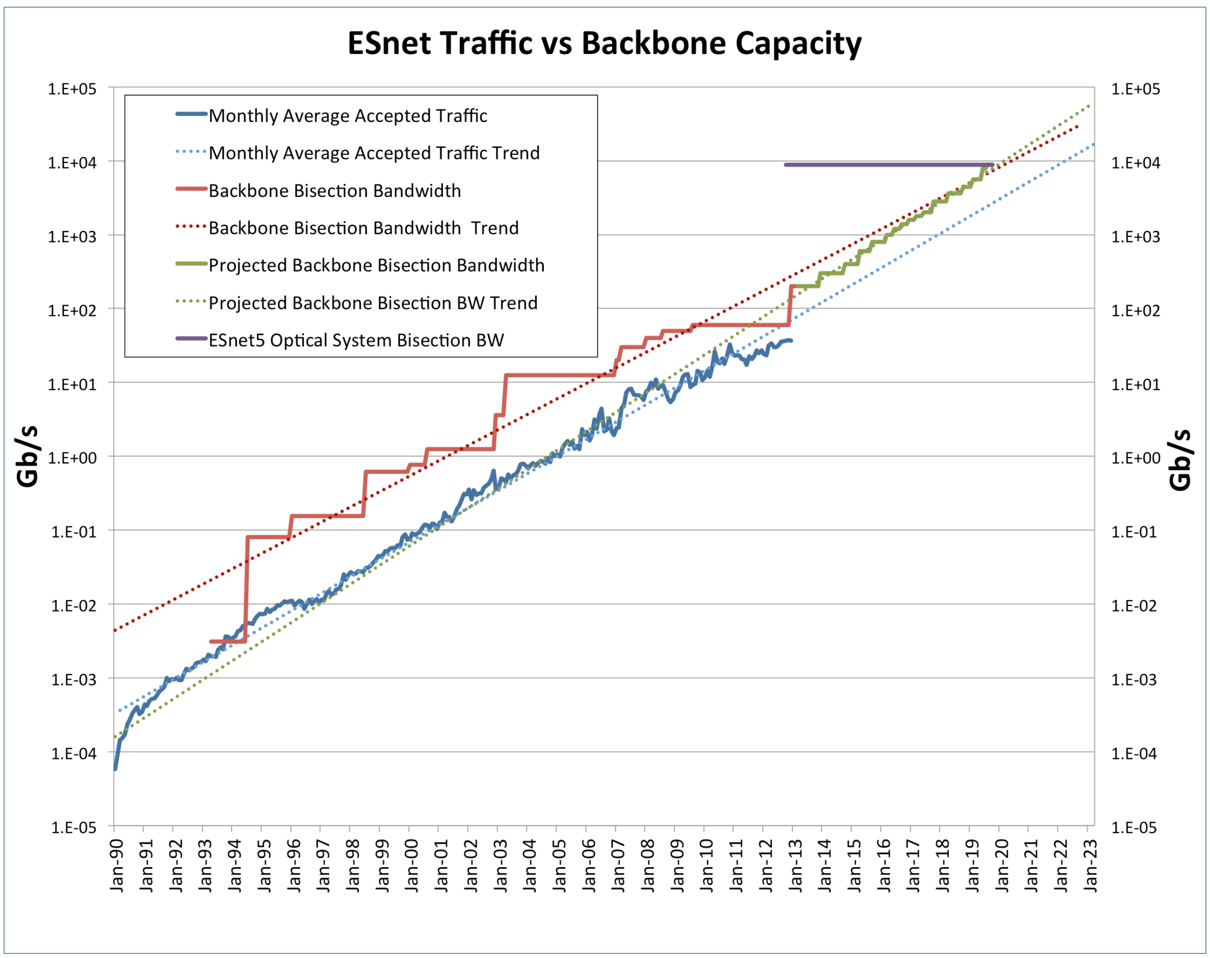
In addition, the facility Director presented core ideas from the plan to a meeting of national laboratory networking coordinators; during a plenary talk at an international networking conference; and to a sub-committee of ASCAC.

A two-page summary was circulated to ASCR computing center directors. ESnet also consulted with the NERSC director and staff, the Berkeley Lab ALD for Computing Science, and the ESnet ASCR Program Manager. Many suggestions for improvement were incorporated into the final draft.

Appendix 2: ESnet Traffic Growth Projections

The following chart shows the historical and projected future relationship between ESnet's monthly accepted traffic volume (in Gbps), and its backbone bisectional bandwidth (also in Gbps). In simple terms, the red line must 'stay above the blue' in order for the network to be resilient against a major fiber outage and also remain non-blocking. The green line shows projected growth rate in backbone capacity.

ESnet anticipates adding new 100G waves to the national footprint staring as early as Fall 2013. By 2015, new optical and routing chassis will be required. In 2016, the complexity of ESnet5 will exceed that of ESnet4 (in terms of devices under management). In 2017, ESnet will be a Terabit network (10x100G). In 2018 and 2019, ESnet will prepare to light its dark fiber, turning up the first 200G wave on ESnet6 in 2020 (at which time the facility will go back and convert ESnet5 from 100G to 200G technology).



Appendix 3: ESnet Budget Scenarios and Analysis

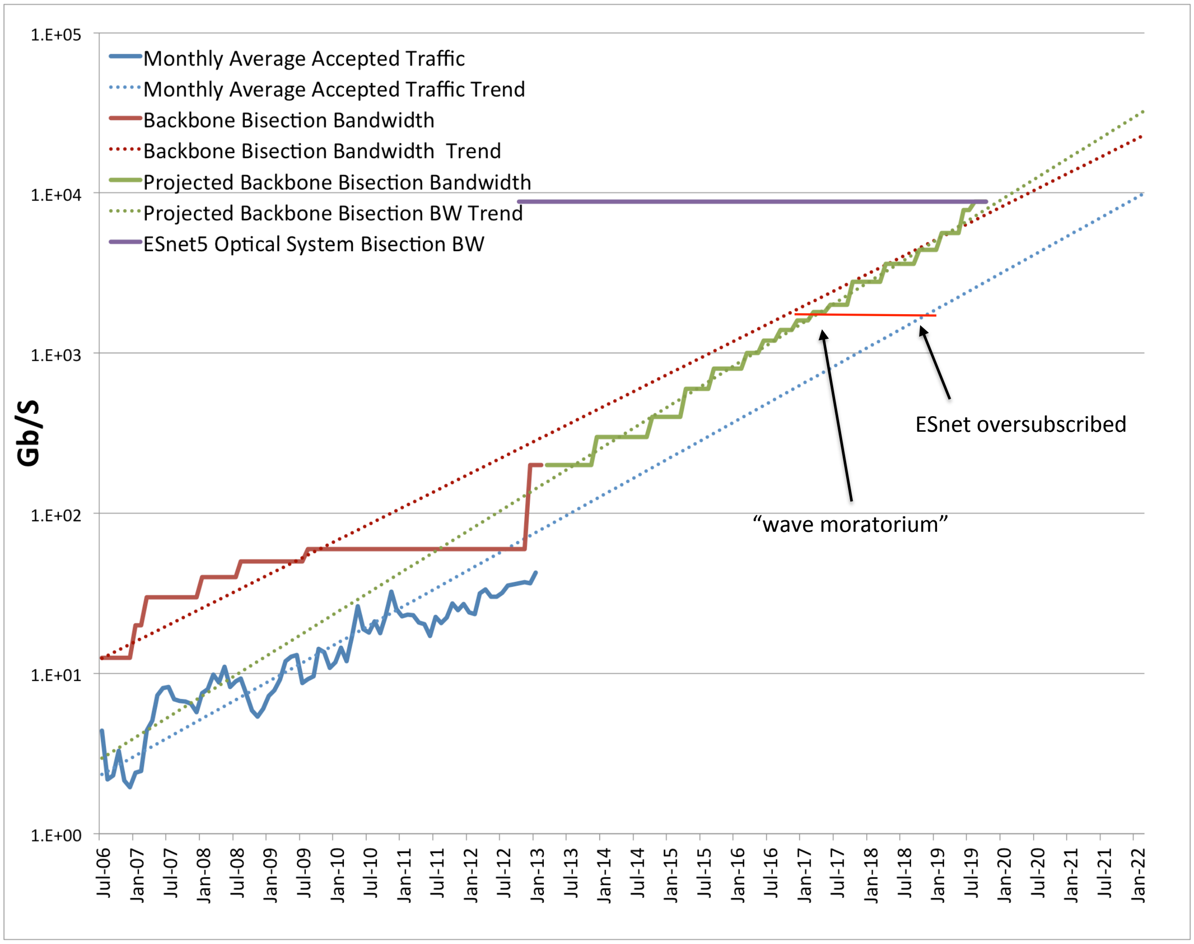
The Office of Science has requested that this strategic plan include two long-term budgetary scenarios: one in which programmatic needs are fully funded, and another in which funding is constrained to grow by 1% annually.

Over the next decade, ESnet's costs will be dominated by optical and routing equipment, plus the labor required to maintain a growing number of optical waves. Both the technology roadmap and downward price-curve for hardware components are well understood for the next 2-3 years, but more poorly understood in the 5-10 year range. Consequently, any long-range budget forecast will have a degree of uncertainty. ESnet is making several assumptions in presenting these budget scenarios.

- First, we assume that no technological breakthrough will radically change the facility's cost structure in the next 10 years. Incremental technology improvements and steady declines in component costs are factored in, but step-function disruptions are not.
- We assume that 200Gbps technology will not achieve rough cost-parity with 100Gbps technology until 2019, and that 400Gbps technology will not be attractive in long-haul operations, due to its shorter reach. Again a breakthrough is possible, but we do not factor one into these budgetary models.
- Finally, we assume that traffic demand will continue at historical rates. With current information, ESnet has no reason to believe otherwise.

Starting in FY2016, the cost to light additional waves on the ESnet5 infrastructure will increase significantly, as a result of exponential traffic growth. As the 'fully funded' budget scenario indicates, ESnet's budget authority will need to rise in order to meet ongoing demand. In approximately 2020, ESnet will have exhausted spectral capacity on the current fiber, and will need to deploy optical gear to light its dark-fiber footprint (using 200G wave technology); after that, ESnet5 will be converted from 100G to 200G technology.

In the 1% budget scenario, ESnet will encounter **severe problems** beginning in FY17. The facility would curtail certain programmatic activities (including trans-Atlantic extension), but more importantly it would stop adding additional waves, and begin allocating existing capacity through peer review. Assuming ESnet stops adding waves in 2017, then by 2021, network demand would be 10x greater than ESnet capacity. Long before that, the lack of adequate data transport within the DOE would severely constrain the productivity of every program, facility, and national lab. Smaller scientific projects might try to adapt by shipping disks, but most larger projects (including ATLAS and CMS) would not have that option, and their science workflows would be devastated. The following diagram (which 'zooms in' on a portion of the diagram shown in Appendix 2) illustrates the timing and effect of the 1% funding scenario.



For those readers more familiar with thinking about funding in connection with supercomputing facilities, an analogy may be helpful. Imagine an HPC center that is constrained from procuring any new systems. Over time, center capacity would become negligible compared to the growing demand for computation; eventually, the center would be effectively useless. A ‘wave moratorium’ for ESnet would have the same effect.

Because ESnet functions as a vital circulatory system for the entire DOE complex, capping its growth would fundamentally compromise the $5B investment made each year by the Office of Science.