SCinet: 25 Years of Extreme Networking

Linda Winkler
Argonne National Laboratory
winkler@anl.gov

ABSTRACT

For 25 years, SCinet has provided SC conference attendees and the high-performance computing (HPC) community with a one-of-a-kind research network utilizing advanced technologies to internationally interconnect, interoperate, and showcase HPC applications in conjunction with scientific and network research at SC. SCinet is a focal point for the international high-performance networking community. It is the place for network engineers, network service providers, equipment vendors, and application developers in the diverse HPC communities to collaborate in demonstrating the latest research, innovative applications, and advanced technologies.

The event is made possible by the tireless efforts of over a hundred volunteers and the generous contributions of dozens of vendors and service providers. It requires a year to plan and two weeks to build, culminating in an operational network for four days, which then is dismantled, packed, and shipped in 24 hours.

SCinet makes a major contribution to the advanced networking field by bringing together unique resources in the form of hardware, software, and services along with expertise to architect, build and operate a state-of-the-art network to enable data-intensive science applications.

1. Overview

This paper is organized as follows. Section 2 introduces how SCinet was first integrated into the conference. Section 3 details the motivation behind SCinet. Section 4 provides insight into the team structure and their roles and responsibilities. Section 5 illuminates the scope of the SCinet efforts by presenting statistics captured from 2014. Section 6 presents the evolution of local area network protocols deployed by SCinet by highlighting the key technologies and significant inflection points. Section 7 paints the wide-area landscape as it has evolved over the 25 years. Section 8 describes the international network participation enabling SCinet to deliver network connectivity around the world. Section 9 provides details of challenges involving the network, which have been a part of the SC conference series. Section 10 captures the benefits of SCinet to the volunteers, vendors and the advanced networking community. Section 11 concludes with an overview of what SC15 has in store and offers suggestions for future SCinet directions.

2. In the Beginning

In 1991 the network was integrated into the SC conference for the first time. Networking was also added to the official title of SC in 1997, making it the International Conference for High Performance Computing, Networking, Storage and Analysis. SCinet was built to provide local-area networking (LAN) within the convention center, as well as wide-area connectivity to enable access to interesting remote resources.

From the beginning SCinet's mission has been to bring new technologies and capabilities to bear. Often this has meant relying on new products and untried interoperability. The 1991 network architecture was loosely based on the production network at the Pittsburgh Supercomputing Center (PSC) [1]. The first technologies to be deployed by SCinet were local-area connections at 10 Mbps Ethernet, 100 Mbps Fiber Distributed Data Interface (FDDI), and 800 Mbps High Performance Parallel Interface (HIPPI); wide-area network (WAN) connections to the National Science Foundation NSFNET [2] at 45 Mbps and Los Alamos and Sandia (Albuquerque) National Laboratories were provided at 45 Mbps and 155 Mbps, respectively, for a total wide-area capacity of 245 Mbps. The event demonstrated the interconnection of 12 supercomputers over a radius of 20 miles using HIPPI. This unprecedented access to multivendor resources on the exhibit floor as well as to remote resources allowed research exhibitors and vendors to display expanded capabilities.

A demonstration at SC91 was one of the first uses of wide-area networks to support a high-speed TCP/IP-based distributed application. The goal was to demonstrate real-time remote visualization of a large, complex scientific dataset. A Thinking Machines' CM-2 and Cray Y-MP at PSC were used to compute the visualization of a high-resolution MRI scan of a human brain. A user at a workstation in Albuquerque selected brain features of interest, and the values were transferred to the CM-2. To get the components of the distributed application to interoperate, engineers from PSC, Lawrence Berkeley National Laboratory and Cray worked feverishly up to the hour that the exhibits opened, in order to accomplish the first heterogeneous operation of the TCP large window option that made high-speed TCP transfer possible between Albuquerque and Pittsburgh. This demonstration was the beginning of the legacy established by SCinet in exploring emerging network protocols and technologies.

3. Motivation

The SC venue has become the high profile showcase for the latest standards and recommended best practices in high-performance computing and networking. SCinet is a purpose-built research network that explores the boundaries of networking technology through various workshops, demonstrations, and challenges, in addition to providing production services and Internet connectivity to conference operations and attendees. SCinet is the exclusive network service provider for the SC conference series. SCinet network engineers work closely with equipment manufacturers, service providers, application developers, and researchers to highlight new technologies at the conference. Factors distinguishing SCinet from other conference networks are the amount of extreme wide-area bandwidth provided and connectivity to national and international research and education networks.

 ACM 978-1-4503-4002-1/15/11…$15.00 DOI: http://dx.doi.org/10.1145/2830318.2830321

In the early days of the SC conference series, research exhibitors would drag their expensive supercomputing hardware around the country to each year's venue in order to demonstrate their applications. By 1999, network capacity at SC was considered sufficiently time-tested to allow exhibitors to leave the hardware at home and rely on SCinet wide-area network connectivity to deliver remote access to the same resources.

Lawrence Berkeley Lab's Image Based Rendering Assisted Volume Rendering made a splash at SC99. It included a number of components working together to enable direct volume rendering of large, time-varying datasets allowing on-the-fly visualization of data from several applications, from several data sources, using different computational data resources and viewing platforms [3].

4. Behind the Scenes

The planning cycle kicks off in November of the year preceding the next SC conference, with conversations between the network architect, routing and WAN leaders, vendors, service providers, and suppliers to evaluate which technology developments will become ready for prime time in the following year. With the desire to push the edge of what is possible, SCinet works with partners to carefully consider technologies available from multiple vendors to demonstrate capabilities and support services. Selecting architectural elements and products that are generally available as well as those that are experimental is fundamental to the need to balance the risk-reward tradeoff. Equipment evaluation and selection take place in the May–July timeframe, well in advance of the opening of the Connection Request System used to collect exhibitor booth connection details. The selected components are assembled together into a viable architecture before the build process begins in mid-October, roughly four weeks before the start of the conference.

SCinet is made possible each year at the SC conference through the tireless efforts of over a hundred volunteers, the generous contributions of dozens of vendors and service providers, and the invaluable support of the SC planning committee. It requires a year to plan and two weeks to build, culminating in an operational network for four days, which then is dismantled, packed, and shipped in 24 hours.

Numerous teams made up of talented engineers from around the world coordinate the various planning and engineering activities throughout the year to orchestrate the various aspects of SCinet services. Since 1998, the Commodity Network team has provided essential hard-wired connectivity throughout the convention center, to deliver commodity Internet service for workshops, tutorials, business operations, and attendees.

Wireless data services were first offered on the exhibit floor in 1999. By 2002, wireless services were available throughout the convention center venue for all attendees.

The backbone infrastructure required to support the SCinet architecture relies heavily on fiber optic cable to reach the distances required across the convention center facility. The fiber backbone is designed and installed by the Fiber team. In just a couple of days, each year the Fiber team deploys many miles of fiber optic cable on the exhibition floor to interconnect SCinet's equipment, external networks, and exhibitor booths to the SCinet Network Operations Center. Multimode fiber is run to exhibit booths to support Gigabit Ethernet connections. Higher-speed connections are supported over single-mode fiber.

The Routing team provides switching and routing services and works closely with exhibitors on configuration, optimization, and troubleshooting. Researchers interested in exploring network optimization, novel protocols, or tools are invited to collaborate with SCinet and to leverage the power and diversity of the SCinet infrastructure.

The WAN Transport team is responsible for engineering metro, national, and international connectivity to remote research and education resources to enable the data-intensive science applications that SC attendees want to demonstrate. In under a month, SCinet builds one of the world's fastest networks.

SCinet established an experimental networks (Xnet) component in 1999 to highlight bleeding-edge, prestandard network hardware and capabilities that typically do not exist outside the development laboratory. Beta and even alpha products that were common occurrences in the early years were later moved to Xnet to minimize risk to the general conference population.

Network monitoring and measurement instrumentation by SCinet became an integral part of the Network Bandwidth Challenge (detailed in Section 9) to determine the performance of each of the competing applications. SCinet utilizes advanced monitoring tools for this purpose. High-bandwidth applications provide a special case for demonstrating these tools. The Measurement team offers participants and attendees a view of performance in real time through a composite view of the network topology and telemetry across the infrastructure. During SC14, network utilization peaked at over 900 Gbps, supporting all high-performance research demonstrations, exhibitors, wireless traffic, and SC meeting room connections.

SCinet network engineers work closely with applications developers and researchers to ensure that the network performs as expected. Providing common network testing and debugging tools helps exhibitors analyze the performance of their application across the complex SCinet infrastructure. The Bandwidth Test Controller [4] is available to participants to run measurement tests to determine the achievable or available bandwidth, path, one-way latency or loss between hosts. The One-Way Active Measurement Protocol [5] application is available to determine one-way latency between end systems. The perfSONAR [6] test and measurement infrastructure is used for end-to-end monitoring and troubleshooting of multidomain network performance. Using host and network instrumentation, the team helps measure performance characteristics and traffic loads, presenting the results in real time.

The SCinet network security team must protect SCinet and attendee assets from external hackers and malicious activity. As a high-profile target for cyber attacks, SCinet must monitor network traffic for integrity without impeding network performance. Commercial firewalls cannot accommodate the data rates and functions SCinet requires. In the dynamic and open network environment created by SCinet, the Bro Network Security Monitor [7] developed by Vern Paxson at Lawrence Berkeley National Laboratory has been instrumental in providing network traffic analysis while minimizing the impact on performance. Bro has been a staple tool for the SCinet Network Security team since 2000.

Rather than requiring all network packets to flow through a firewall, Bro passively monitors IP packets, watching for suspicious activity. Bro is open-source software, licensed under the BSD license that includes a powerful scripting feature to analyze traffic. When Bro detects a suspicious traffic pattern, it can automatically take actions to limit or block the traffic, or it can inform security experts that anomalies should be examined in greater detail. Bro's traffic analysis components extend beyond the security domain, including performance measurements and troubleshooting. SCinet utilizes a Bro cluster to distribute the high-performance traffic across an appropriate number of backend PCs, all running dedicated Bro instances on individual traffic slices.

In protecting everything connected to the SCinet network, the data collected via high-speed fiber optic taps offers network security vendors and researchers a unique sample of extremely diverse network traffic. Various network security tools are exercised to analyze the traffic generated during the conference. The network traffic that SCinet monitors is considered proprietary and does not leave SCinet.

These teams and 10 others—Architecture, Communications, Helpdesk, Interconnect, IT Services, Logistics, Network Research Exhibition, Physical Security, Power, and Student Volunteers—play a pivotal role in guaranteeing that SCinet can successfully support network services for the activities of attendees, exhibitors, and committee operations during the conference.

5. By the Numbers

Delivering the advanced networking services to SC is no small undertaking. Below is a sample of the impressive numbers that quantify the scope of the SCinet efforts in 2014.

- 1.2 Tbps of WAN bandwidth delivered to the New Orleans convention center
- Network utilization peaked >900 Gbps
- 84 miles of fiber deployed throughout the New Orleans convention center
- 14 miles of fiber deployed on the exhibit floor to booths
- 41 fiber taps monitoring >300G of WAN traffic
- >$18M in loaned equipment and contributed services
- >125 volunteers
- 182 temporary wireless access points deployed
- >4800 simultaneous wireless clients supported
- 180 Ethernet switches installed throughout the convention center
- 28kW AC and DC power utilized to energize SCinet

SC15 promises to surpass these numbers. Watch the SCinet displays around the NOC for the latest information on this year's numbers.

6. LAN Technology Evolution

Over the 25 years that SCinet has been serving SC, many technologies have appeared and been vetted, with some amended, modified, enhanced, or replaced in order to facilitate the demonstration of high-performance computing and networking. Each technology area represents opportunities for researchers, developers, and vendors to get a unique glimpse into their area of interest. Figure 1 summarizes the lifespans of various LAN technologies SCinet has worked with through the years. Following is a brief discussion of some of the significant inflection points and key technologies employed by SCinet.

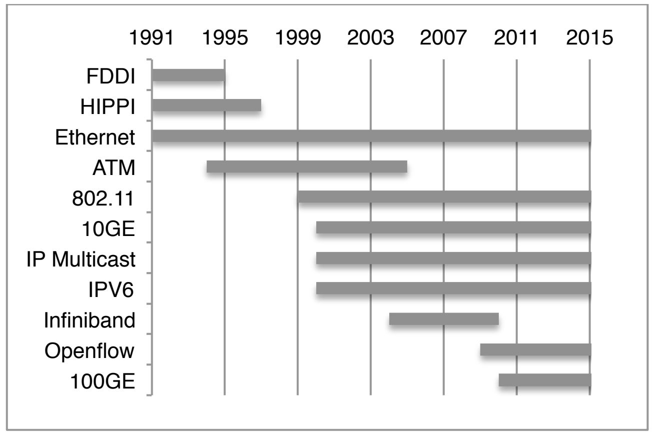

Figure 1. LAN Technology Lifespans

The FDDI American National Standards Institute (ANSI) standard defines an optical fiber-based token-ring network consisting of a primary ring carrying 100 Mbps and a secondary ring as a backup capable of carrying an additional 100 Mbps. When no backup ring was required, 200 Mbps could be achieved. FDDI supported a larger frame size (4,352 bytes) than the Ethernet standard (1,500 bytes), providing more effective data throughput rates. In the early- to mid-1990s, FDDI was an attractive campus backbone network technology because existing Ethernet networks offered only 10 Mbps data rates. By 1995 the introduction of Fast Ethernet with its greater speed, lower cost, and ubiquity made FDDI obsolete in local networks. SCinet supported FDDI from 1991 to 1994.

HIPPI is a standard point-to-point protocol for transmitting large amounts of data over relatively short distances, mainly on local-area networks between supercomputers and storage devices. HIPPI uses nonblocking crossbar switches to forward data with minimal processing. Switches provide error detection, but error correction is left to higher-level protocols. HIPPI copper cables were limited to 25 meters. HIPPI was popular in the late 1980s and into the late 1990s but has been replaced by faster standard interfaces. HIPPI was the first "near-gigabit" (0.8 Gbps) standard for network data transmission. It was specifically designed for supercomputers and was never intended for the mass market. Many of the features developed for HIPPI have been integrated into InfiniBand. The remarkable aspect about HIPPI is that it was developed at a time when Ethernet was limited to 10 Mbps and SONET at 155 Mbps was considered leading-edge technology. At SC, HIPPI first appeared in 1991 and was last used in 1996.

The Ethernet family of network technologies for local and metropolitan area networks was commercially introduced in the 1980s, and the first IEEE 802.3 working group standard emerged in 1983 [8]. The standard has been refined over time to support higher data rates over longer distance links. Ethernet is the only protocol that has endured the test of time for the 25 years SCinet has been in existence—albeit evolving to include higher bandwidth, improved media access control, and different physical media.

The following list captures the significant SCinet Ethernet milestones. The global market growth of Ethernet is unprecedented by any other networking technology.

- 1991 10 Mbps Ethernet supported
- 1995 100 Mbps Ethernet supported

- 1998 1 Gbps Ethernet supported
- 2000 802.3ad link aggregation utilized to increase bandwidth between devices
- 2000 preproduction 10G Ethernet demonstrated by Xnet
- 2002 10G Ethernet services offered to exhibitors
- 2004 first 10G Ethernet WAN circuit implemented
- 2010 100G LAN and WAN supported

Asynchronous Transfer Mode (ATM) is a telecommunication standard defined by ANSI and the International Telecommunications Union (ITU) to support a range of traffic, including data, and real-time, low-latency services such as voice and video. It was developed to meet the needs of the Broadband Integrated Services Digital Network, as defined in the late 1980s, and to unify telecommunication and computer networks [9]. ATM uses asynchronous time-division multiplexing and encodes data into small, fixed-size 53-byte cells, unlike IP and Ethernet that use variable-sized packets and frames. The connection-oriented model of ATM requires a virtual circuit to be established between two endpoints before data can be exchanged. Virtual circuits can be permanent, that is, dedicated connections preconfigured by the service provider, or switched, that is, set up (signaled) and torn down on a per-call basis.

The ATM Forum defined the LAN Emulation (LANE) standard in 1995 as a way to allow legacy networks such as Ethernet, Token Ring, and FDDI to use an ATM network for backbone connections. The integration of ATM with legacy protocols was not straightforward. As a connection-oriented technology, ATM required that virtual circuits exist between source and destination before any data could be sent. Data was transmitted in fixed-length cells requiring hardware segmentation and reassembly of frames. Legacy networks transmit data in variable-length frames over a shared connection-less network. LANE automated the setup of switched virtual circuits across ATM networks for clients. Alternatively, administrators had to manually configure permanent virtual circuits between end systems before traffic may be exchanged.

ATM was popular with carriers and many networking companies in the 1990s. By the end of the decade, however, the improved price/performance of Internet protocol products was replacing ATM technology for the integration of real-time and bursty traffic. For data rates of 155 Mbps and above, the cost of segmentation and reassembly hardware required for the fixed cell rendered ATM less competitive than Packet over SONET for IP. SCinet supported ATM in the local-area and wide-area networking environment from 1994 to 2004.

The IEEE 802.11 [10] standard is a set of media access control and physical layer specifications for implementing wireless LAN communications in the 2.4, 3.6, 5, and 60 GHz frequency bands. Established and maintained by the IEEE 802 LAN/MAN Standards Committee, the base version of the standard was released in 1997 and has undergone several amendments. SCinet first began supporting wireless LAN technologies on the exhibit floor in 1999. Ubiquitous wireless access quickly became the norm for the conference; and supporting the steadily increasing number of attendees and their ever-increasing number of simultaneous wireless devices has been an ongoing challenge. SCinet has supported 802.11 a, b, g, and n standards over the years to provide conference attendees the best possible wireless experience.

Over time various strategies have been employed to deploy wireless within convention center venues, including borrowing access points (problematic) from vendors and using existing wireless infrastructure in the convention center (unknown quality). In 2013, SC invested in its own wireless equipment to make the SCinet wireless installation suitable for future needs.

InfiniBand is a communications standard that defines an input/output architecture used in high-performance computing featuring high-throughput and very low latency. It is used to interconnect servers, switching equipment, storage, and embedded systems. Introduced in 1991, single data rate operates at 2.5 Gbps. InfiniBand is a fabric architecture that leverages switched, point-to-point connections with data rates up to 120 Gbps. InfiniBand links can be aggregated: most systems use the quad data rate (QDR) 4X aggregate (40 Gbps) introduced in 2008; typically, 12X links (120 Gbps) are used for clusters, supercomputers, and inter-switch interconnections. New InfiniBand data link encoding schemes have been introduced roughly every three years to maintain pace with CPU/GPU developments and the increasing performance demands of HPC applications and users.

In 2004, we witnessed the adoption of InfiniBand as a clustering interconnect solution, outperforming Ethernet in latency and price. The OpenIB Alliance (later evolved to the Open Fabrics Alliance) [11] was founded to focus on developing a vendor-independent, Linux-based InfiniBand software stack. In 2005, InfiniBand was implemented to interconnect storage devices. Linux distributions first included InfiniBand support, and the first OpenIB showcase appeared at SC. In 2004 and 2005 the Storcloud Challenge (described in Section 9) made extensive use of InfiniBand, tying together unprecedented quantities of multivendor storage devices on the exhibit floor. By 2009, InfiniBand started to take hold in the market over Gigabit Ethernet as the internal interconnect technology for supercomputers. In 2010, SCinet offered exhibitors QDR 4X and 12X connections. Distance limitations of InfiniBand cabling made it impractical to support across the exhibit floor. InfiniBand connections were last offered across the SC exhibit floor in 2010.

7. The Wide-Area Landscape

In the early years, individual exhibitors arranged special WAN circuits into the venue. This approach clearly did not scale. SCinet stepped up and collected requirements to fulfill the needs for all exhibitors. Figure 2 summarizes the lifespans of the WAN technologies SCinet has worked with. Following is an introduction to some of the noteworthy inflection points and significant WAN technologies employed by SCinet.

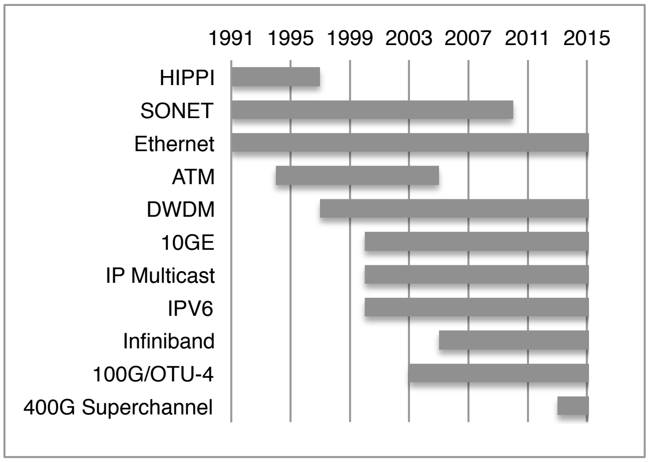

Figure 2. WAN Technology Lifespans

The wide-area capacity brought to bear each year is dominated by bandwidth to national and international research and education networks. In 2014, only a small fraction of the capacity, roughly 8%, was dedicated to the commodity Internet. The lion's share of the wide-area capacity requirements is dominated by applications highlighted on the exhibit floor.

The extreme network bandwidth and rich connectivity that SCinet brings to the SC conference are made possible through collaborations with various network service providers and carriers. SCinet leverages regional research and education (R&E) network service providers listed in Table 1 to extend national WAN services into the convention center. Negotiations often entail locating and procuring dark fiber into convention centers or working with carriers to extend their reach the last mile(s) into the conference venue. Most convention centers lack the necessary metro fiber infrastructure to support SC requirements. Building resilient metro fiber paths is often challenging and is not delivered until the 23rd hour.

Table 1. Research and Education Network Collaborations

Regional Research and Education Network	**SC Venue (year)**
Three Rivers Optical Exchange (3ROX)	Pittsburgh (04)
Corporation for Education Network Initiatives in California (CENIC)	Phoenix (03), Reno (07)
Florida Lambda Rail (FLR)	Orlando (98), Tampa (06)
Front Range GigaPop (FRGP)	Denver (13)
Lonestar Education and Research Network (LEARN)	Austin (08, 15)
Louisiana Optical Network Initiative (LONI)	New Orleans (10, 14)
Mid-Atlantic Crossroads (MAX)	Baltimore (02)
Network for Education and Research in Oregon (NERO)	Portland (09)
Pacific Northwest GigaPop (PNWGP)	Seattle (99,05,11)
Pacific Wave	Seattle (05,11)
Utah Education Network (UEN)	Salt Lake City (12)

Synchronous Optical Networking (SONET) is an ANSI standard that defines optical signal rates and formats and a synchronous frame structure for multiplexed digital traffic on optical networks [12]. The SONET transport containers allow for the delivery of various protocols including traditional telephony, ATM, Ethernet, and TCP/IP. Typically deployed by carriers using Linear Automatic Protection Switching, SONET long-haul circuits were traditionally expensive. Based on a basic frame format and speed structure, SONET uses Synchronous Transport Signal (STS), with STS-1 as the base-level signal at 51.84 Mbps. SONET offers a hierarchy of transmission speeds from 51 Mbps to 39 Gbps (OC-1 at 51 Mbps, OC-3 at 155 Mbps, OC-12 at 622 Mbps, OC-48 at 2.4 Gbps and OC-768 at 39 Gbps). Multiple lower-level signals can be multiplexed to form higher-level signals. The transmission of IP packets directly over SONET (POS) became popular by IP service providers seeking cost-effective, efficient solutions for interconnecting routers at high-speed while circumventing the 25% ATM cell tax overhead: SONET offered roughly a 2% overhead tax by comparison [13]. SCinet supported POS WAN circuits from 1991 to 2009; it last supported POS connections to exhibit booths in 2002. In 2004, SCinet built the fastest SONET-rate WAN connection at OC-768. In 2010, all WAN connections utilized optical transport network (OTN) framing, replacing POS.

Dense wavelength division multiplexing (DWDM) technology uses multiple wavelengths to transmit a number of bidirectional optical carrier signals onto a single optical fiber pair. The first commercially viable wavelength-division multiplexing systems appeared around 1995. Interoperation with legacy routing and switching equipment is achieved through optical-electrical-optical translation at the edge of the optical transport.

SCinet's first encounter with wavelength division multiplexing involved the National Transparent Optical Network (NTON) network in 1997. NTON, cofunded by the Defense Advanced Research Projects Agency (DARPA) and the National Transparent Optical Network Consortium, provided an open research platform for Next Generation Internet trials [14]. Members of the consortium included Nortel Networks, GST Telecommunications, Lawrence Livermore National Laboratory, Sprint Communications, and the Bay Area Rapid Transit District. NTON was a 2,500 km 16- wavelength division multiplexed network deployed using in-place commercial fiber from GST Telecommunications and provided OC-48 and OC-192 SONET services. NTON spanned from San Diego to Seattle, with add/drop locations in Los Angeles, Oakland, and Portland linking government, university, and private sector labs; and it was able to interface with many of the broadband research networks in the United States. NTON supported an ATM virtual topology over the optical layer and demonstrated the transparency feature of WDM by concurrently transmitting several data formats over different wavelengths.

NTON first provided services to SC 97 by provisioning two OC-12 connections into San Jose and one OC-3 and one OC-12 into the NASA National Research and Education Network (NREN) at NASA Ames Research Center [15]. In 1999, NTON carried traffic from PNWGP in Seattle, the University of Washington, and agency networks such as the DREN (DOD), ESnet (DOE), NREN (NASA), and vBNS (NSF) [16] to the Portland Convention Center over OC-48 and OC-192 waves on a 4-node 8-wave Nortel DWDM transport system. Of the 8-wave capacity, one wave was utilized for SCinet production services and the remainder of the capacity was part of Xnet. In 2000, NTON provided OC-48 connectivity between Los Angeles and the San

Francisco Bay area, as well as peering with Internet2 in Los Angeles in support of SC.

In 2001–2003, SCinet used DWDM across the exhibit floor modeled after a metropolitan area network. SCinet has built a metro DWDM system every year since 2002 to deliver the extreme quantities of bandwidth required into the convention center while utilizing a minimum number of scarce metro fiber strands.

The Optical Transport Network (OTN) defined in ITU-T G.709, also commonly called "digital wrapper," is the industry-standard protocol providing an efficient way to multiplex services onto optical channels. One of the key benefits of OTN technology is that multiple networks and services, such as legacy SONET, can be combined seamlessly onto a common infrastructure with Ethernet, storage and video applications. Key elements have been added to improve performance and reduce cost. These include: management of channels in the optical domain and forward error correction to improve error performance and enable longer optical spans. OTN also provides a standard method for end-to-end optical wavelength management without the need to convert an optical signal into the electrical domain. The enhanced multiplexing capability of OTN allows different traffic types—including Ethernet, storage, digital video, as well as SONET—to be carried over a single Optical Transport Unit (OTU) frame, either OTU-1 at 2.7 Gbps, OTU-2 at 10.7 Gbps, OTU-3 at 43 Gbps, or OTU-4 at 112 Gbps. WAN circuits implemented by SCinet have utilized OTN standards since 2006.

Around 2010, coherent optical transmission at 40 Gbps and 100 Gbps began to be deployed, enabling higher data rates to be transmitted over long-haul (typically >2,000 km) optical transmission networks. DWDM superchannels were deployed in field trials conducted in 2010–2011 offering a solution to the problems of increasing optical transport capacity beyond 100 Gbps and providing the flexibility to maximize the combination of optical capacity and reach. Implementing a superchannel with multiple optical carriers reduces the requirement for exotic electronics, allowing the higher bandwidth to be delivered much more quickly than other options would allow. SCinet implemented a 400 Gbps superchannel across the Denver metro area in 2013.

In 1995, the Information Wide Area Year (I-WAY) [17,18] was an experimental environment created for SC for building distributed virtual reality applications and for exploring issues of distributed wide area resource management and scheduling. The goal of the I-WAY project was to enable researchers to use multiple networked supercomputers and advanced visualization systems to conduct very large-scale computations. By inter-connecting 11 wide-area ATM testbeds and agency networks (ACTS, ATDnet, BAGnet, CalREN, CANARIE, CASA, DREN, ESnet, MAGIC, NREN, vBNS), 17 supercomputer centers, 5 virtual reality research sites, and over 60 applications groups, the I-WAY project created an extremely diverse wide-area environment for exploring advanced applications. The I-WAY involved coordinating connectivity between multiple carriers (Ameritech, AT&T, Bell Atlantic, MCI, Pacbell, Sprint) to interconnect resources and extend networks to San Diego where the applications were demonstrated. This was the first time that some of the early testbeds were interconnected; and some of the carriers were uncomfortable with the notion of interoperability at this juncture. Nevertheless, the environment provided a glimpse of the future for advanced scientific computing and identified a number of areas where research was needed to achieve the vision.

Dr. Michael Norman from the National Center for Supercomputing Applications used computing resources at the National Science Foundation supercomputing centers as well as the Maui Supercomputer Center across the I-WAY to simulate a >100 Gflops distributed supercomputer to assess the scalability of several algorithms developed by the Grand Challenge Cosmology Consortium. Results of the simulations were visualized by using the CAVE Automatic Virtual Environment to interactively navigate the enormous datasets of virtual universe [19].

Through the years, SCinet has interconnected with numerous experimental network testbeds to facilitate access to remote computing and scientific resources. Many of the early testbed networks (ACTS, ATDnet, BAGNet, CASA, HSCC, MAGIC, NTON) have evolved or been replaced by stable production infrastructure (CalREN, Internet2, National LambdaRail, vBNS) and agency networks (DREN, ESnet, NOAA, and NREN).

The following list captures many of the significant SCinet WAN milestones.

- 1991 first WAN connection into SC
- 1997 first demonstration of DWDM
- 2002 the last ATM WAN circuit supported
- 2004 the last time ATM was exercised
- 2004 OC-768 WAN circuit implemented
- 2004 first 10G/OTN WAN circuit appeared
- 2004 National LambdaRail introduces Layer2 circuits to SCinet WAN portfolio
- 2005 first long-haul InfiniBand demonstrated
- 2010 SONET/POS disappears; replaced by 10G/OTN
- 2010 first 100G/OTN WAN circuit implemented (Chicago to New Orleans)
- 2013 first trans-Atlantic 100G link is available; ANA-100G NYC to Amsterdam [20]
- 2013 first 400G superchannel employed
- 2014 ANA trans-Atlantic link upgraded to 200G
- 2014 first trans-Pacific 100G link; Seattle to Singapore
- 2014 1.2 Tbps WAN capacity delivered to New Orleans

From 2005 to 2009 InfiniBand protocols were utilized over long-haul links to efficiently move large datasets between data centers over the wide area. InfiniBand reappeared in 2014 between Singapore and the SC New Orleans venue, spanning in excess of 20,000 km.

During the past ten years we have observed an explosion of interdomain Layer2 circuit provisioning as a key mechanism for delivering service to researchers. SCinet routinely stress tests new features and services, including Generalized Multi-Protocol Label Switching (GMPLS) control of routers to create circuits on demand and the dynamic creation of interdomain Layer2 circuits. The complex wide-area portfolio features the Layer2 R&E networks from ESnet and Internet2 along with key U.S. R&E exchange points including AMPATH, AtlanticWave, MAN LAN, Pacific Wave, and StarLight.

The WAN bandwidth delivered at SC has grown exponentially in order to support the insatiable demand from exhibitors to meet the needs of their high-bandwidth applications. Figure 3 illustrates the SCinet wide-area capacity plotted on a logarithmic scale.

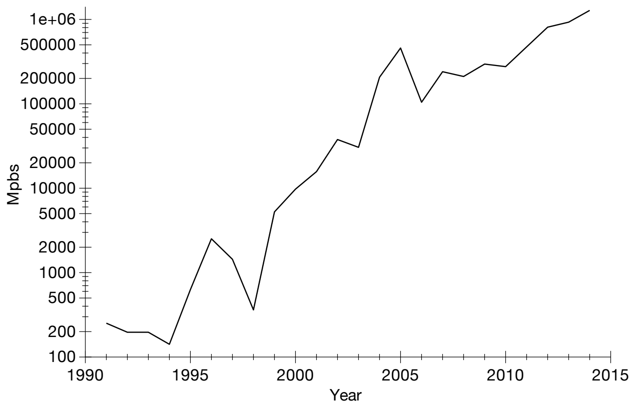

Figure 3. SC WAN Capacity

As the demand for bandwidth by applications has increased, WAN circuit sharing has been problematic. In many cases the requirements of the demonstrations demand careful scheduling of the WAN circuits, as many of the demos cannot co-exist. The need to timeshare/schedule the use of WAN circuits continues to be challenging.

8. Global Reach

SC is an international event with global impact. Many of the demonstrations require multidomain coordination. Working in collaboration with exchange point operators in the United States as well as other countries (i.e. Netherlands, Japan, and Korea), SCinet is able to provision end-to-end lightpaths spanning the globe. To shorten the turn-around time for these setups, the Japan R&E Network Operations Center (JP-NOC) has set up shop on the SC exhibit floor since 2006.

The persistent Global Lambda Integrated Facility (GLIF) infrastructure has provided a global foundation for the support of demonstrations at SC since 2002. GLIF is an international consortium that promotes the paradigm of lambda networking [21]. GLIF participants are national R&E networks, consortia, and institutions supporting lambdas. The GLIF community shares a common vision of building a framework, in which the central architectural element is optical networks to support the most demanding E-science applications. GLIF was formally established in Reykjavik, Iceland, in August 2003. The expansive reach and generosity of the GLIF community have enabled researchers to demonstrate network capabilities around the globe at SC. GLIF participants include Australia, Brazil, Canada, CERN (the European Organization for Nuclear Research), China, the Czech Republic, Denmark, Finland, Iceland, Japan, Korea, Netherlands, New Zealand, Norway, Poland, Russia, Singapore, Sweden, Taiwan, the United Kingdom, and the United States.

9. Network Challenges

Between 1999 and 2009 Xnet showcased various bleeding-edge technologies. By establishing a venue for demonstrations that were allowed to fail, Xnet provided a crash-and-burn environment to highlight potentially disruptive technologies while not interfering with the ability of SCinet to deliver production-quality services to the remainder of the conference attendees. In 1999 (Portland), Xnet featured WAN connectivity over NTON, as described in Section 7. In 2000 (Dallas), 10G Ethernet was first demonstrated between computer clusters utilizing multiple vendors on the exhibit floor. In 2001 (Denver), DWDM was highlighted across the exhibit floor. Software-defined networking (SDN) made an early appearance in 2004 (Pittsburgh) with Dynamic Resource Allocation via GMPLS optical networks (DRAGON) [22] from Mid-Atlantic Crossroads GigaPOP and USC/ISI and Dynamic Resource Allocation Controller (DRAC) [23] from Nortel. In 2009, ESnet first used OSCARS [24] to dynamically provision lightpaths across the ESnet backbone to support SC demonstrations [25]. In the years that followed, research on control plane management has continued to be a hot topic. In 2005 (Seattle), high-speed encryption over WAN circuits and Ethernet super jumbo frame (16k–64k) experiments were included. In 2009 (Portland), Xnet included the exhibition of Openflow controllers. Many of the technologies first demonstrated by Xnet were later included as standard SCinet service offerings once the technologies reached maturity.

By 2010, SCinet found it increasingly difficult to attract suitable breakable demonstrations to Xnet. In some ways a victim of its own success, as SC grew, efforts to attract vendor technical pre-GA product demonstration have been over ridden by corporate marking departments. Xnet and the Research Sandbox activities tried to define compelling rationale to encourage participation but were insufficient to overcome the marketing veto.

In 2000, the SCinet team created an extensive wide-area network and worked with the research community to find applications that could take advantage of the unique high-powered network. SCinet solicited proposals for inventive bandwidth-intensive application demonstrations as a way to challenge the community to show that, by removing bandwidth as an obstacle, researchers can unleash advanced applications capable of carrying out scientific and collaborative missions leading to exciting new discoveries. The event became the SC Network Bandwidth Challenge (BWC) [26]. Applications were selected based on their ability to both fully utilize the SCinet network infrastructure and deliver innovative application value. The challenge became somewhat of a network stress test as researchers developed applications to push the boundaries of the network infrastructure.

The numerous applications that contributed to the BWC from 2000 to 2009 not only used SCinet bandwidth and advanced features efficiently but also illustrated that applications could step up to the challenge, given increased network speed to unleash important applications with the potential to carry out scientific missions in future years. Applications developers were willing to commit significant time and energy to the BWC for the opportunity to test next-generation network capabilities in advance of production services. Often, doing so required refactoring the application design in order to achieve the next level of performance. SCinet brought together a dedicated group of experts willing and able to dissect, inspect, evaluate and tune systems to optimize applications network performance.

A team from Lawrence Berkeley National Laboratory demonstrating Visapult, a prototype application and framework for remote visualization of terascale datasets, won the 2000 BWC "Fastest and Fattest" category for overall best performance. Utilizing 2.4 Gbps POS connectivity across NTON between Los Angeles and Oakland, and the High Speed Connectivity Consortium (HSCC) backbone between Los Angeles and Dallas, the Visapult team recorded a peak performance level of 1.48 Gbps over five-second sample periods. Bandwidth restrictions over the HSCC backbone prevented higher performance levels [3, 26].

SC Global played a part in the conference from 2001 to 2006 by using Grid technologies to link diverse communities from around the world and enable international remote researchers to participate in and contribute to the conference. SC Global used the Access Grid [27]; an ensemble of resources including multimedia large-format displays, presentation and interactive environments, and interfaces to Grid middleware and to visualization environments.

Remote sites had the opportunity to participate in activities on the SC exhibit floor, and many offered content to the conference—including workshops, presentations, panels, and birds-of-a-feather sessions in areas ranging from science, computing, and networking technologies to interactive art exhibits and cultural exchanges. In 2005, the addition of SC Desktop created the opportunity for virtual attendees to remotely participate in parts of the technical program. The challenge for SCinet was to support wide-area IP multicast protocols on an international scale.

The Storcloud Challenge [28], introduced to the conference in 2004 and continued in 2005, was an initiative to provide high-performance storage to HPC applications. Storcloud had two goals: (1) to provide high-performance storage in support of the conference presenters' applications and (2) to serve as a showcase for vendors of high-performance storage systems to demonstrate their capabilities and ability to scale and to illustrate InfiniBand interoperability with other vendors' storage and network solutions. The networking infrastructure to support the Storcloud Challenge was implemented by SCinet.

SCinet replaced Xnet and the BWC with the SCinet Research Sandbox (2010–2012) as a way to continue the focus on network research activities within the conference. Those efforts have evolved into the SCinet Network Research Exhibition (2013–2015) and the Innovating the Network for Data Intensive Science workshop (2014–2015). Through these efforts, SCinet continues to work with the SC Planning Committee to raise the visibility of network research within the technical program.

As part of the Network Research Exhibition at SC14, the Caltech Network Team presented a demonstration titled Intelligent Software Driven Dynamic Hybrid Networks with Terabit/sec Science Data Flows. The demonstration showcased efficiently moving large scientific datasets between Large Hadron Collider data centers by leveraging intelligence through different layers of software and hardware to dynamically reconfigured network components to realize optimal performance. To achieve the goal of 1 Tbps between storage and memory, the configuration of the Caltech booth along with that of their collaborators (iCAIR, Michigan, and Vanderbilt) provided more than 1 Tbps of LAN capacity and 400 Gbps over the WAN. SCinet linked the SC conference in New Orleans to Caltech, CERN, Victoria, B.C., the University of Michigan, and San Paulo for the demonstration. The servers and networks deployed were a modest representation of the global operations and cooperation inherent in global science programs handling hundreds of petabytes of data per year [29]. Look for the next progression of this application on the exhibit floor at SC15.

10. Benefits

Over 100 volunteers from national laboratories, educational institutions, HPC sites, national and international research networks, government agencies, equipment vendors, and telecommunications carriers from around the world come together to design, build, and deliver SCinet. This is an excellent training opportunity for network engineers to gain practical hands-on experience designing and building a state-of-the-art multivendor network to support HPC applications. Participating in SCinet gives engineers a chance to interact with leading-edge technology before making recommendations to management. The "try before you buy" approach is an invaluable experience that is difficult find.

Each year vendor partners loan or donate a large percentage of the equipment and services used to build the impressive network infrastructure. The SC high profile, multivendor showcase attracts vendors and service providers to participate in a unique proving ground in which to exercise new products. SCinet participation gives vendors and developers valuable feedback on their products and tools, whether beta-level deployments, open-source projects, or commercial systems.

Vendors and equipment manufacturers also target SC as a venue in which to announce new products and features. SC is considered the premier event for reaching the HPC decision makers. Being able to reference deployment in SCinet is a powerful endorsement because of the high-caliber team and the stress-testing nature of the event. Vendors have used SCinet to harden their products, and it allows them to deploy products in ways that are difficult to simulate in their development labs but are clearly relevant to the HPC environment that the decision-makers who attend SC represent.

Both vendors and engineers find the hands-on experience and "learning by doing" training extremely valuable. Participating vendors are able to set aside their commercial competitive nature and collaborate to achieve SCinet. The unique relationship and the collegiate nature of SCinet often surprise the first-timers but continue to draw interest from new participants. In the spirit of SCinet, standard vendor competitiveness is replaced with the concept of "doing your very best and offering your very best technology." While missteps have occurred from time to time, everyone involved in SCinet learns from the experience and strives to work through obstacles and ensure that issues get corrected. If vendors were punished for every slip-up or subpar performance of new cutting-edge products deployed in SCinet, SCinet would see only mundane, generally available products being offered.

SCinet is recognized by applications developers, hardware and software manufacturers throughout the networking industry, agency sponsors, and decision makers in industry, academia, and government research organizations for its long-standing reputation as a high-caliber team with a proven track record of organizing a high-profile event to showcase scientific applications. SCinet contributes to the advanced networking field and the HPC community by orchestrating advanced hardware, software, and services from industry, academia, and research labs into a state-of-the-art environment loosely modeled after the IETF creed of rough consensus and running code. After a year of planning and two weeks of construction, SCinet spends four intense days working closely with applications developers, and network researchers to harness the capabilities needed to enable advances in data-intensive scientific discovery.

11. 2015 and Beyond

As this paper was being prepared, the architectural details for this year's event were still being finalized. Three aspects of the SC 15 architecture are noteworthy. Openflow has been adopted as a tool to provision Ethernet drops to exhibit booths. Whether the technology will provide the necessary degree of flexibility and resiliency remains to be seen. The hope is that we can build upon this experience to explore more sophisticated tools to control the network. An unprecedented 1.6 Tbps of WAN capacity will be delivered from a single rack. Firewall technology will be engineered to help protect assets when exhibitors are not able to provide the necessary level of system integrity SCinet requires in order to keep these end systems from inflicting further damage to other attendees. Next-generation firewall products and tools will be put to the test performing analytics on the SC network traffic at data rates of 400 Gbps.

Looking to the future, SCinet must build upon its strengths and find ways to eliminate or improve upon its weaknesses. The focus should be on a small number of goals with the highest return. Continuing to build upon the strong community involvement is a key to future SCinet potential for success. Opportunities to engage the network research community should be evaluated. SCinet should seek new opportunities to explore networking research in the telecommunications industry beyond the product development-marketing horizon. SCinet should establish an advisory committee to solicit feedback. SCinet must continue to respond to a multifaceted challenge. SCinet will push the boundaries of what is possible through demonstration of multi-vendor interoperability of new technologies while bridging the gap between theory and practice. SCinet will provide advanced capabilities to customers and at the same time will learn how to manage new infrastructure and services. SCinet must strike a balance between innovation and stability requiring mindful risk assessment and management. SCinet must keep things rewarding and relevant for both the highly capable engineers who build it, the vendors and service providers who support it, the applications developers who use it and the attendees who pay to observe the event. Careful planning and assessment will go a long way toward preserving the strong commitment and track record SCinet has in making SC a success.

Acknowledgments

I thank the following individuals for their assistance in preparing this paper: Maxine Brown, Eli Dart, Paul Daspit, Cees de Laat, Jon Dugan, Marc Lyonnais, Pat Teller, Brian Tierney, Kevin Walsh, and Rod Wilson. My work is supported by the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357.

References

[1] Alex Nussbaum. SCinet '91 – Supercomputing '91s gigabit network. SIGUCS Newsletter, Vol. 22, no. 1, April 1992, 24–35.

[2] About NSFNET. http://nsfnet-legacy.org/about.php

[3] Bethel, W.; Tierney, B.; Lee, J.; Gunter, D.; Using High-Speed WANs and Network Data Caches to Enable Remote and Distributed Visualization. IEEE. ACM/IEEE Supercomputing Conference, Nov 2000.

[4] Bandwidth Test Controller. http://software.internet2.edu/bwctl/index.html

[5] One-Way Active Measurement Protocol. http://software.internet2.edu/owamp/index.html

[6] perfSONAR. http://www.perfsonar.net/

[7] V. Paxson. Bro: A System for Detecting Network Intruders in Real-Time. Computer Networks, Vol. 31, nos. 23–24. 1999, 2435-2463.

[8] IEEE 802.3 'Standard for Ethernet' Marks 30 Years of Innovation and Global Market Growth (press release). IEEE. June 24, 2013.

[9] ATM Forum, The User Network Interface (UNI), v. 3.1, ISBN 0-13-393828-X, Prentice Hall PTR, 1995.

[10] IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. 2012 revision. IEEE-SA, April 2002.

[11] OFA Overview. https://www.openfabrics.org/index.php/organization.html

[12] Telcordia GR-253-CORE, Synchronous Optical Network (SONET) Transport Systems: Common Generic Criteria, GR-253, issue 5, October 2009.

[13] IP Over Direct Links (SONET). http://www.cs.rutgers.edu/~rmartin/teaching/fall08/cs552/position-papers/021-01.pdf.

[14] Nortel Networks. National Transparent Optical Network Consortium (NTONC). Final Technical Report AFRL-IF-RS-TR-2004-303, October 2004.

[15] Jude A. George. Network Quality-of-Service Demonstration at Supercomputing 97. NASA Technical Report NAS-98-006. June 1998.

[16] John Jamison, Randy Nicklas, Greg Miller, Kevin Thompson, Rick Wilder, Laura Cunningham and Chuck Song. vBNS: Not your father's Internet. IEEE Spectrum, Vol. 35, no. 7. July 1998, 38–46.

[17] T. DeFanti, I. Foster, M. Papka, R. Stevens, T. Kuhfuss. Overview of the I-WAY: Wide Area Visual Supercomputing. International Journal of Supercomputing Applications, Vol. 10, no. 2. 123-130.

[18] I. Foster, J. Geisler, B. Nickless, W. Smith, S. Tuecke. IEEE. Software infrastructure for the I-WAY high-performance distributed computing experiment. Proc. 5th IEEE International Symposium on Distributed Computing, 1997, 562-571.

[19] Maxine Brown. Supercomputing '95 information Architecture: GII Testbed and HPC Challenge Applications on the I-WAY.

[20] ANA-100G. https://www.nordu.net/content/ana-100g

[21] Global Lambda Integrated Facility. http://www.glif.is/

[22] Lehman T; Sobieski, J.; Jabbaru, B; DRAGON: a framework for service provisioning in heterogeneous grid networks. IEEE. Communications Magazine, Vol. 44, issue 3. March 2006, 84-90.

[23] Travostino F.; Keates R.; Lavian T.; Monga I.; Schofield B. Project DRAC: Creating an applications-aware network. Nortel Technical Journal, February 2005, 23-26.

[24] Guok, C.; Robertson, D.; Thompson, M; Lee, J.; Tierney, B.; Johnston, W.; Intra and Interdomain Circuit Provisioning Using the OSCARS Reservation System. IEEE. 3rd International Conference on Broadband Communications, Networks and Systems, 2006. October 2006, 1-8.

[25] StarGate Demo at SC09 Shows How to Keep Astrophysics Data Out of Archival Black Holes. http://www.sdsc.edu/News%20Items/PR112009_stargate.html

[26] William T.C. Kramer. SCinet: Testbed for high-performance networked applications. IEEE. Computer, Vol. 35. no. 6. June 2002, 47–55.

[27] Access Grid. http://accessgrid.org

[28] Storcloud. http://www.sc-conference.org/sc2004/storcloud.html

[29] Intelligent Software Driven Dynamic Hybrid Networks With Terabit/sec Science Data Flow (press release). Caltech. November 23, 2014.