



ESnet

ENERGY SCIENCES NETWORK

Evolution of ESnet - A Changing Landscape in Scientific Networking

Chin Guok

CTO

Energy Sciences Network

Lawrence Berkeley National Laboratory

CHEP 2023

Norfolk, Virginia

May 9, 2023



U.S. DEPARTMENT OF
ENERGY

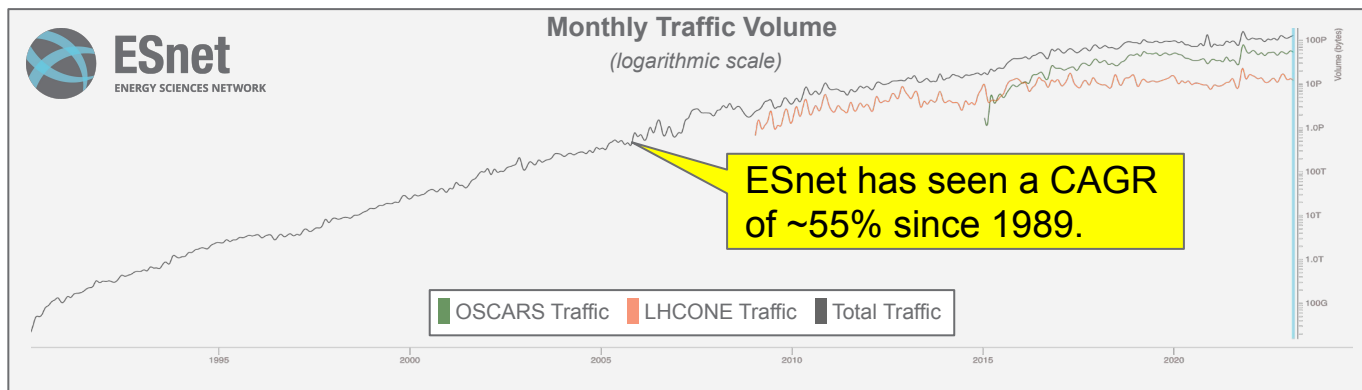
Office of Science



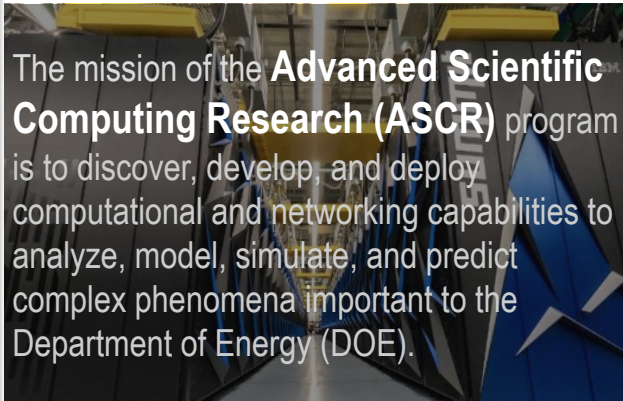
ESnet is a Science Mission Network

ESnet provides the high-bandwidth, reliable connections that **link scientists** at national laboratories, universities, and other research institutions, **enabling them to collaborate** on some of the world's most important scientific challenges including **energy, climate science, and the origins of the universe**. **Funded by the DOE Office of Science**, ESnet is managed and operated by the Scientific Networking Division at Lawrence Berkeley National Laboratory. As a nationwide infrastructure and DOE User Facility, ESnet provides scientists with **access to unique DOE research facilities and computing resources**.


ESnet's Mission is to enable and accelerate scientific discovery by delivering unparalleled network infrastructure, capabilities, and tools.



DOE Office of Science - Largest supporter of basic research in the physical sciences in the US




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).



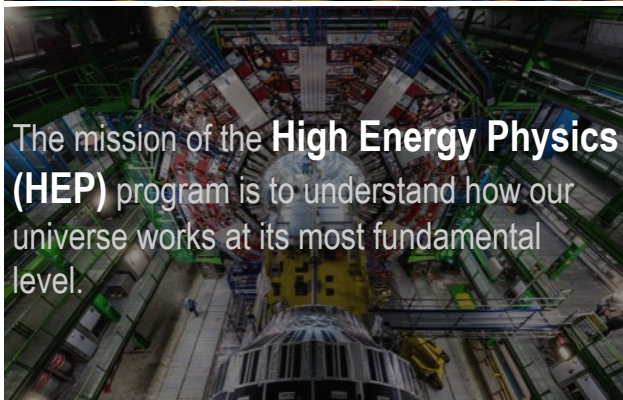
Basic Energy Sciences (BES) supports fundamental research to understand, predict, and ultimately control matter and energy at the electronic, atomic, and molecular levels in order to provide the foundations for new energy technologies and to support DOE missions in energy, environment, and national security.



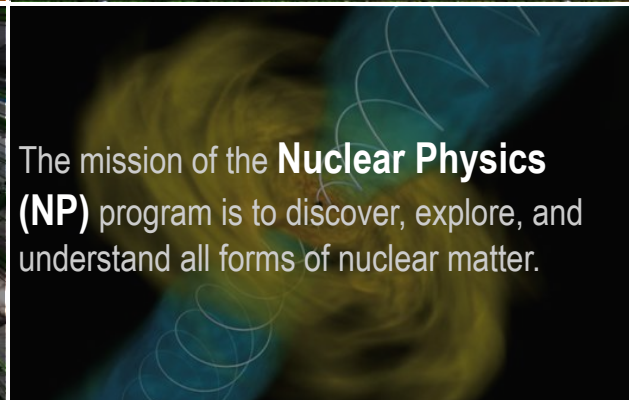
The mission of the **Biological and Environmental Research (BER)** program is to support transformative science and scientific user facilities to achieve a predictive understanding of complex biological, earth, and environmental systems for energy and infrastructure security, independence, and prosperity.



The **Fusion Energy Sciences (FES)** program mission is to expand the fundamental understanding of matter at very high temperatures and densities and to build the scientific foundation needed to develop a fusion energy source.



The mission of the **High Energy Physics (HEP)** program is to understand how our universe works at its most fundamental level.



The mission of the **Nuclear Physics (NP)** program is to discover, explore, and understand all forms of nuclear matter.

DOE Office of Science - Uniquely positioned for large scale collaborative science*

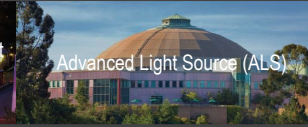
ASCR High End Computing (HEC)



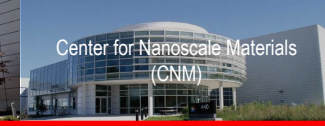
ASCR High Performance Scientific Network



BES X-Ray Light Sources



BES Nanoscale Science Research Centers (NSRCs)



BES Neutron Scattering Facilities



BER



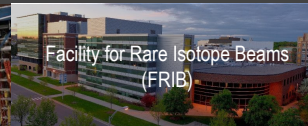
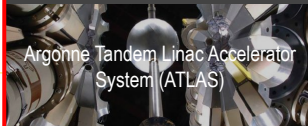
FES



HEP



NP



*DOE Office of Science facilities also support other collaborations, e.g., LHC, LSST, etc

DOE Office of Science - Uniquely positioned for large scale collaborative science*

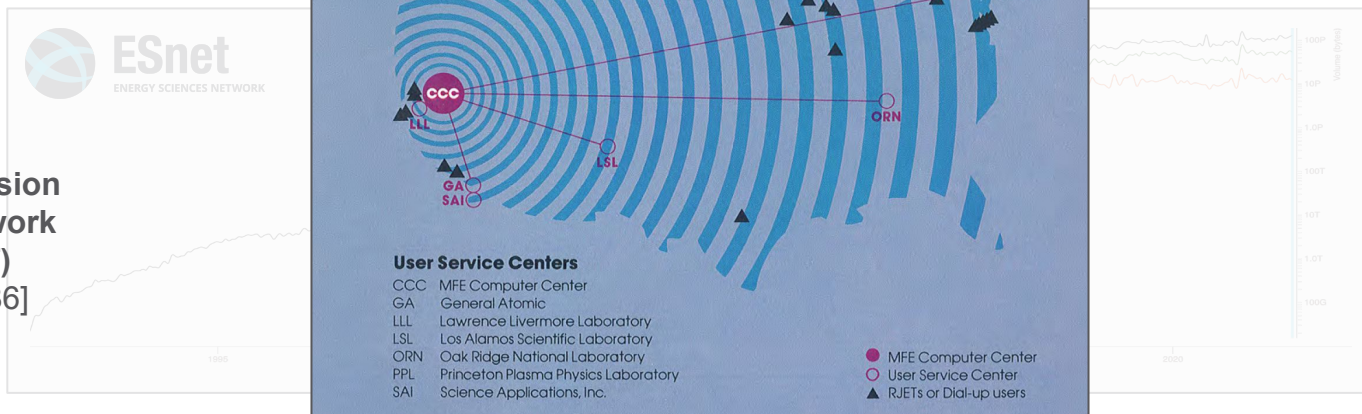
<p>ASCR High End Computing (HEC)</p> <p>Argonne Leadership Computing Facility (ALCF)</p>	<p>BES X-Ray Light Sources</p> <p>Advanced Photon Source (APS)</p> <p>Linac Coherent Light Source (LCLS)</p> <p>Stanford Synchrotron Radiation Light Source (SSRL)</p> <p>Advanced Light Source (ALS)</p> <p>National Synchrotron Light Source II (NSLS-II)</p>				
<p>Oak Ridge Leadership Computing Facility (OLCF)</p>	<p>BES Nanoscale Science Research Centers (NSRCs)</p> <p>Functional Nanomaterials (CEN)</p> <p>Center for Integrated Nanotechnologies (CINT)</p> <p>The Molecular Foundry (TMF)</p> <p>Center for Nanophase Materials Sciences (CNMS)</p> <p>Center for Nanoscale Materials (CNM)</p>				
<p>National Energy Research Scientific Computing Center (NERSC)</p>	<p>BES Neutron Scattering Facilities</p> <p>Spallation Neutron Source (SNS)</p> <p>High Flux Isotope Reactor (HFIR)</p> <p>Joint Genome Institute (JGI)</p> <p>Environmental Molecular Sciences Laboratory (EMSL)</p> <p>Atmospheric Radiation Measurement (ARM) user facility</p>				
<p>ASCR High Performance Scientific Network</p> <p>Energy Sciences Network (ESnet)</p>	<p>FES</p> <p>National Spherical Torus Experiment Upgrade (NSTX-U)</p> <p>DEMO National Fusion Facility (DEMO)</p> <p>Facility for Advanced Accelerator Experimental Tests (FACET)</p> <p>Fermilab Accelerator Complex</p> <p>Accelerator Test Facility (ATF)</p>				
<p>U.S. DEPARTMENT OF ENERGY</p> <p>Office of Science</p>	<p>NP</p> <p>Argonne Tandem Linac Accelerator System (ATLAS)</p> <p>Continuous Electron Beam Accelerator Facility (CEBAF)</p> <p>Facility for Rare Isotope Beams (FRIB)</p> <p>PHENIX Relativistic Heavy Ion Collider (RHIC)</p>				

This is ESnet

*DOE Office of Science facilities also support other collaborations, e.g., LHC, LSST, etc

Evolution of the Energy Sciences Network (ESnet)

**Magnetic Fusion
Energy Network
(MFENET)**
[1976 - 1986]



1970s

1980s

1990s

2000s

2010s

Present

Evolution of the Energy Sciences Network (ESnet)

I also agree that the Scientific Computing Staff should move forward to implement the Energy Sciences Network.

ESnet(1)

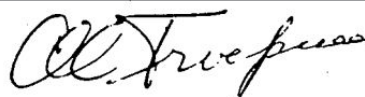
[1986 - 1994]

Building an open standards network

- IPv4
- BGP

Magnetic Fusion Energy Network (MFENET)
[1976 - 1986]

OCT 07 1986


Alvin W. Trivelpiece
Director, Office of Energy Research

United States Government
Department of Energy

memorandum

DATE: OCT 07 1986


TO: James F. Decker, ES-2

SUBJECT: ER Supercomputer Strategy

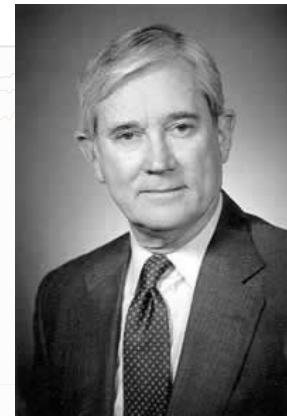
I concur with the recommendations which you present in your memorandum dated September 10, 1986. The acquisition of a Class VII Supercomputer in FY 1988 to support ER scientific has been given high programmatic priority. In view of the economic analysis presented, I agree that this supercomputer system should be installed at the National Magnetic Fusion Energy Computer Center (NMFECC) and that the NMFECC scope should be expanded to be a permanent ER supercomputer center. The establishment of additional ER supercomputer resources should be considered in the future as demand for supercomputer resources and budget factors indicate.

I also agree that the Scientific Computing Staff should move forward to implement the Energy Sciences Network. Funding for the supercomputer system will continue to be provided through the budget and reporting codes in the Office of Fusion Energy and Basic Energy Sciences which currently support this activity. However, the Scientific Computing Staff will assume budgetary responsibility and control of the activities under these budgets.

Please continue to keep me informed with regard to the progress of these important projects.


Alvin W. Trivelpiece
Director, Office of Energy Research

cc: R. Young, ES-60
I. Adler, ES-61
D. Nyquist, ES-63
W. Wallenmeyer, ES-20
C. Cavallini, ES-7
D. Stevens, ES-10
C. Delisi, ES-70



1970s

1980s

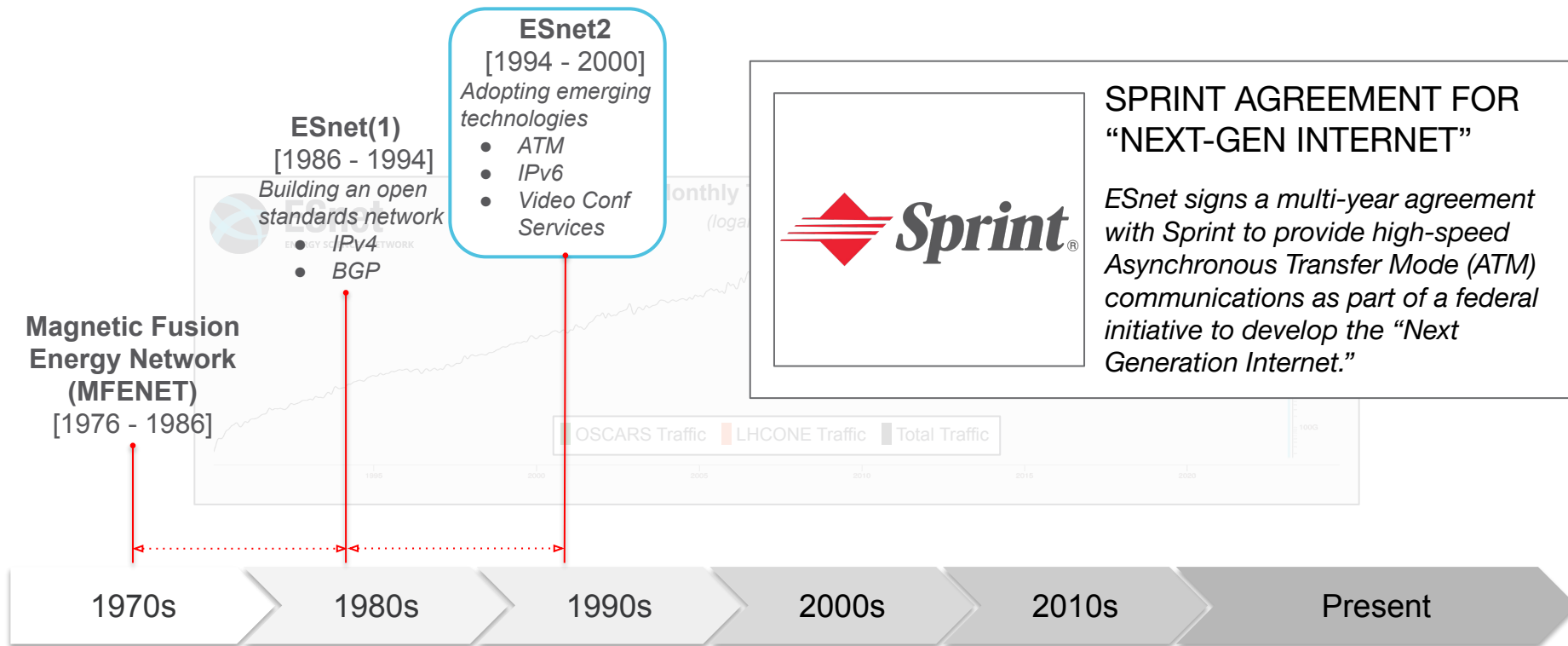
1990s

2000s

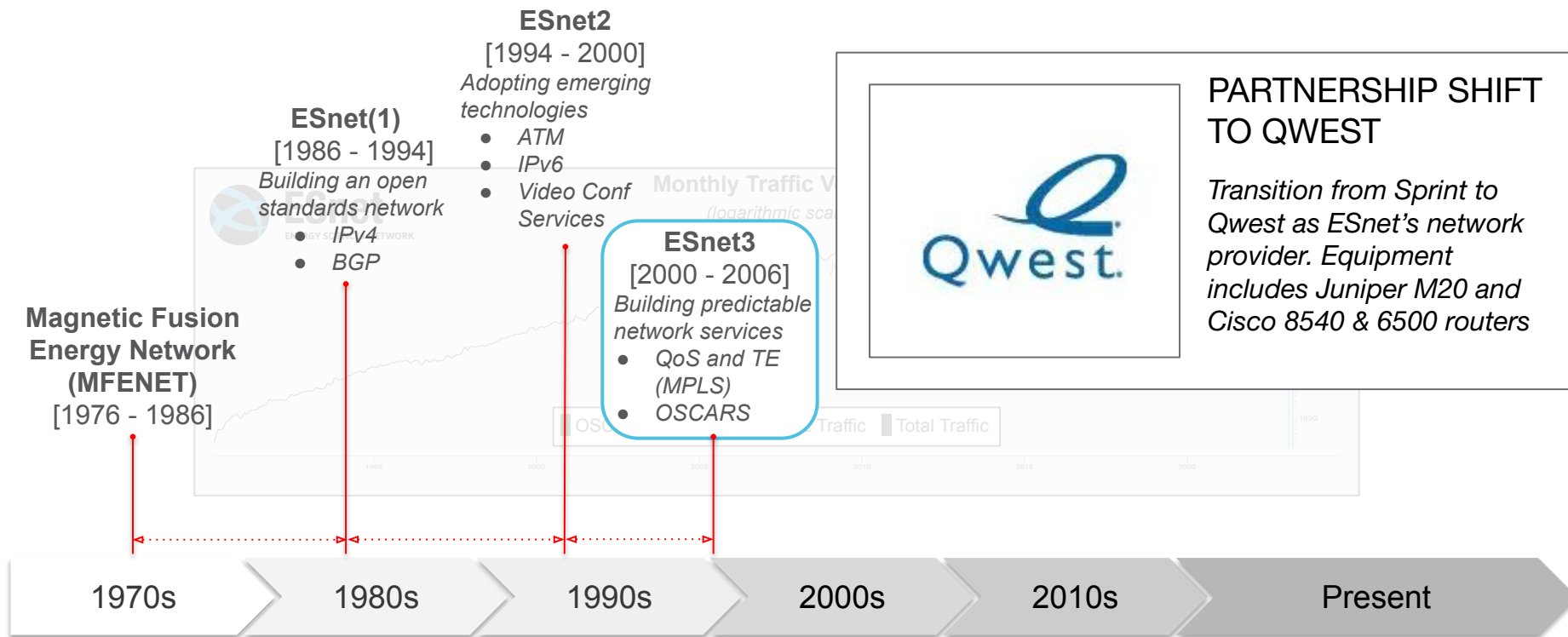
2010s

Present

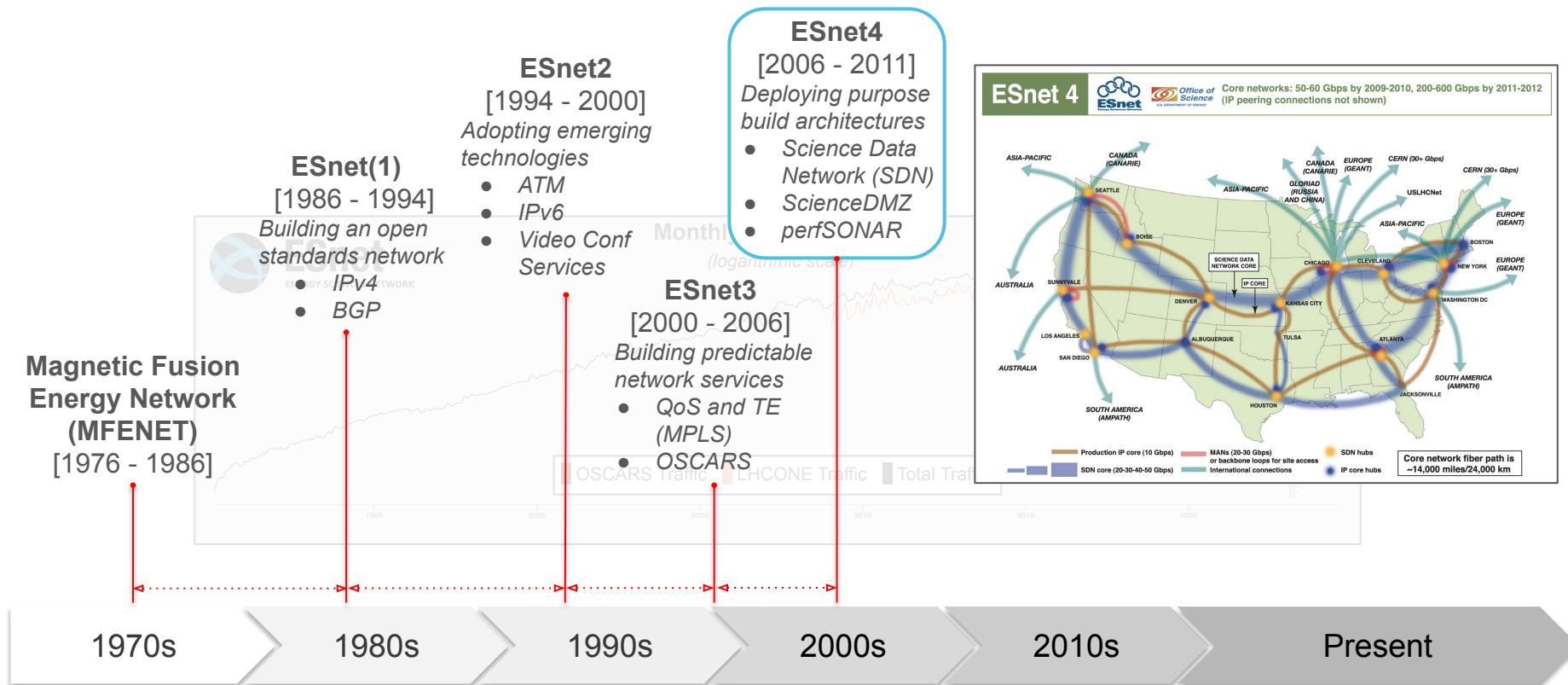
Evolution of the Energy Sciences Network (ESnet)



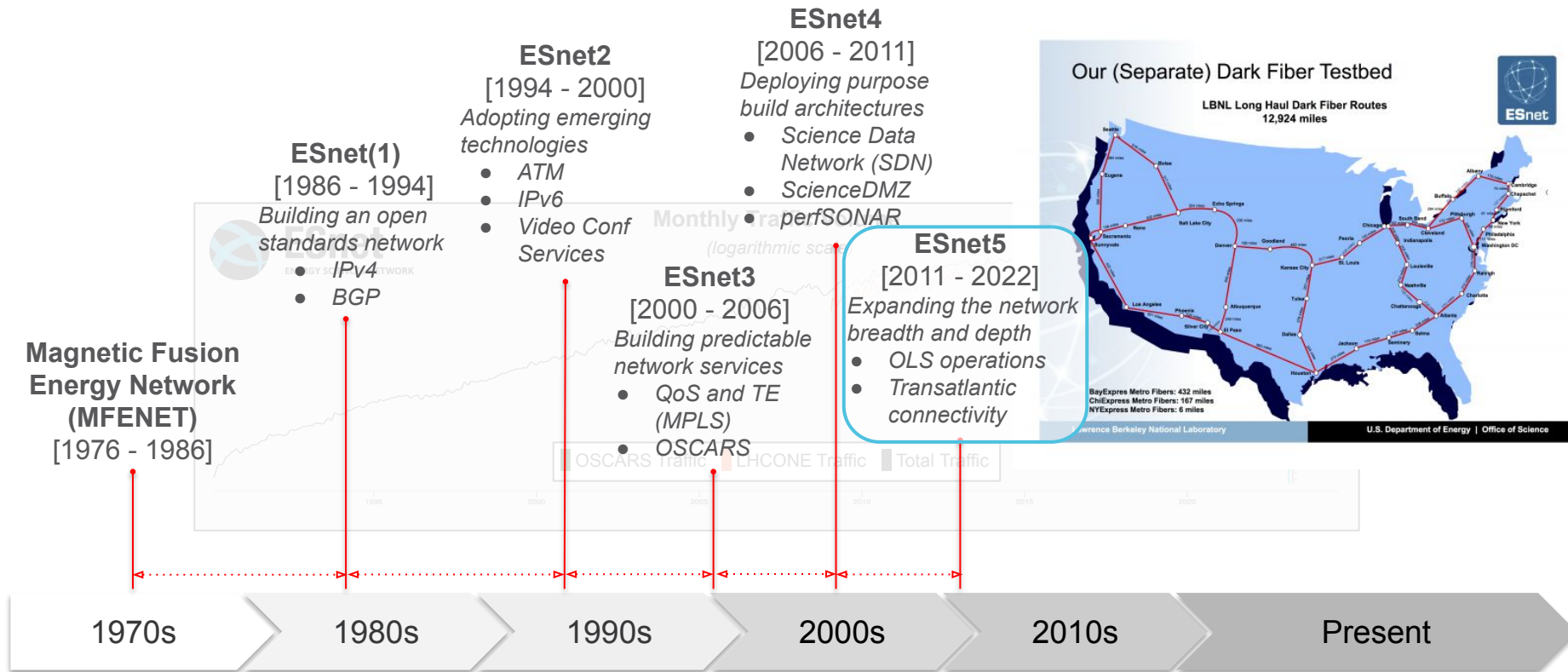
Evolution of the Energy Sciences Network (ESnet)



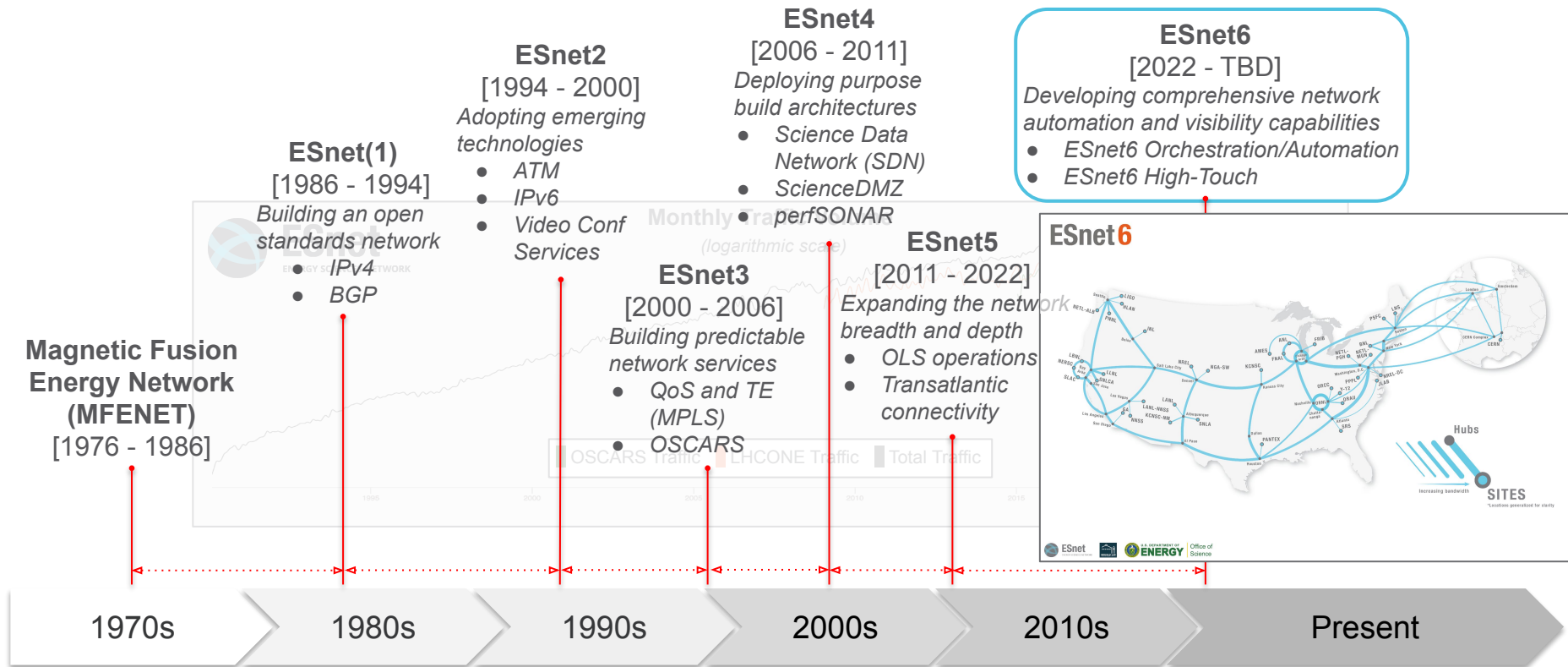
Evolution of the Energy Sciences Network (ESnet)



Evolution of the Energy Sciences Network (ESnet)

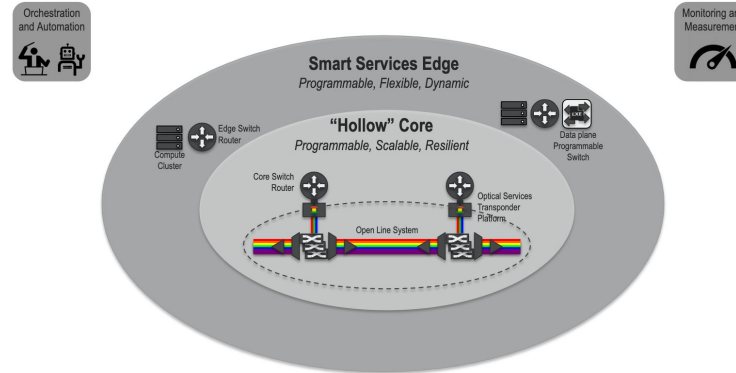


Evolution of the Energy Sciences Network (ESnet)



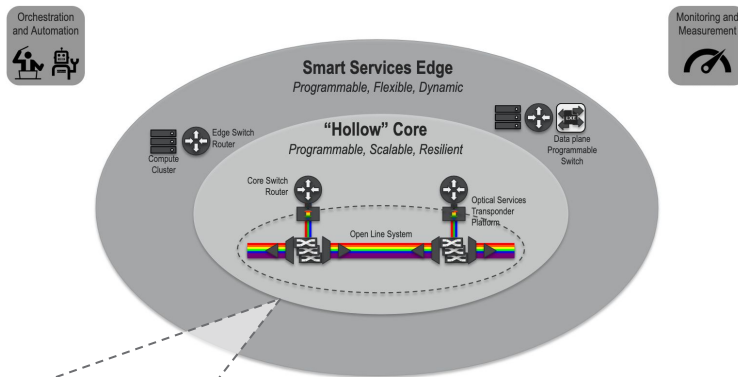
ESnet6 Design and Build (in a nutshell)

ESnet6 “Hollow” Core Architecture

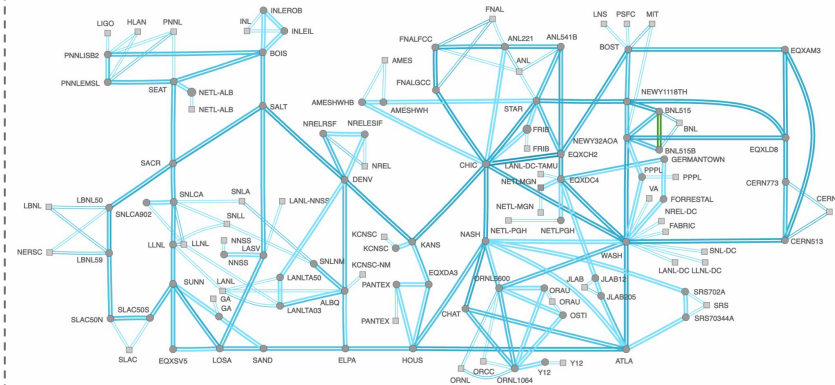


ESnet6 Design and Build (in a nutshell)

ESnet6 “Hollow” Core Architecture



ESnet6 Network Topology



ESnet6 Design and Build (in a nutshell)

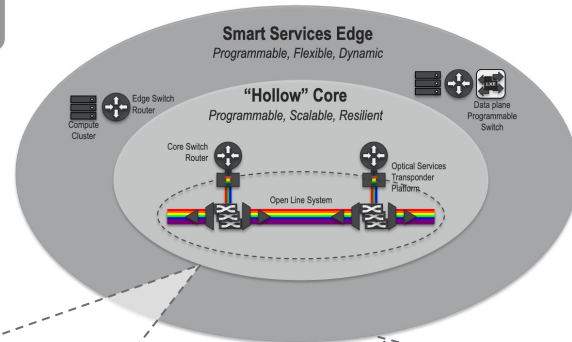
ESnet6 “Hollow” Core Architecture



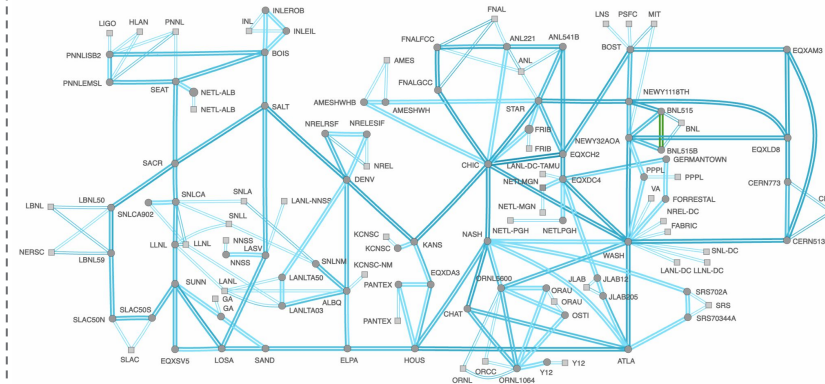
Orchestration
and Automation



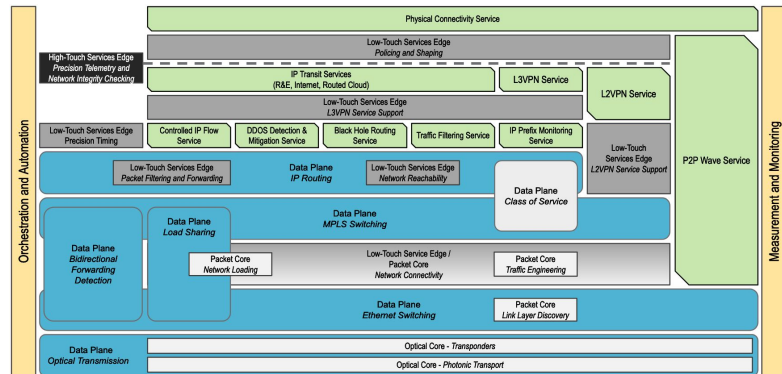
Monitoring and
Measurement



ESnet6 Network Topology

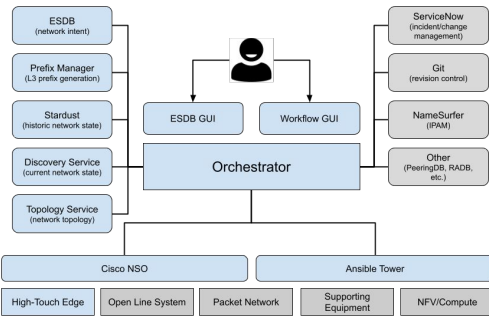


ESnet6 Services and Capabilities Structure

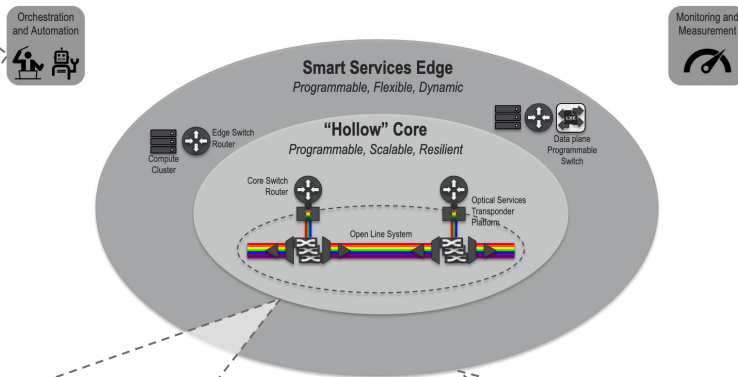


ESnet6 Design and Build (in a nutshell)

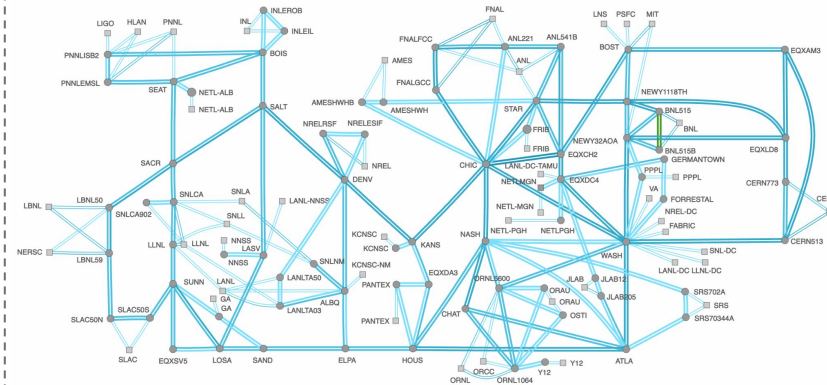
ESnet6 Orchestration & Automation Framework



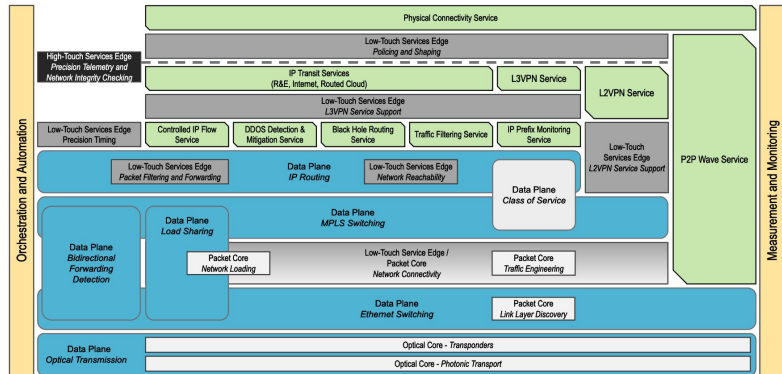
ESnet6 “Hollow” Core Architecture



ESnet6 Network Topology

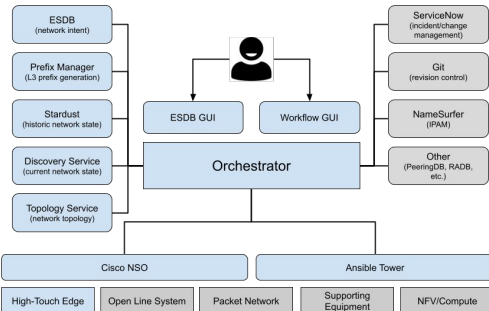


ESnet6 Services and Capabilities Structure

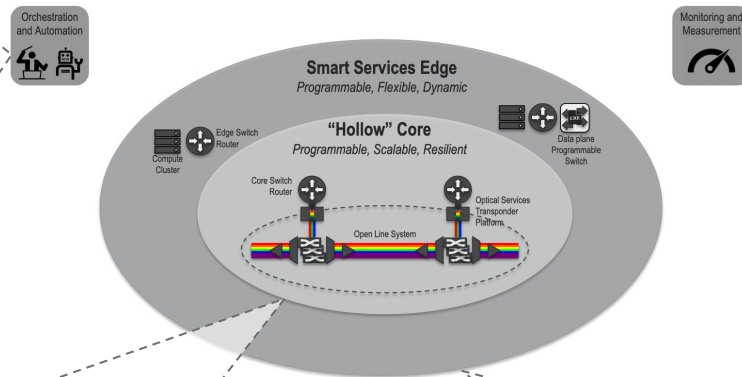


ESnet6 Design and Build (in a nutshell)

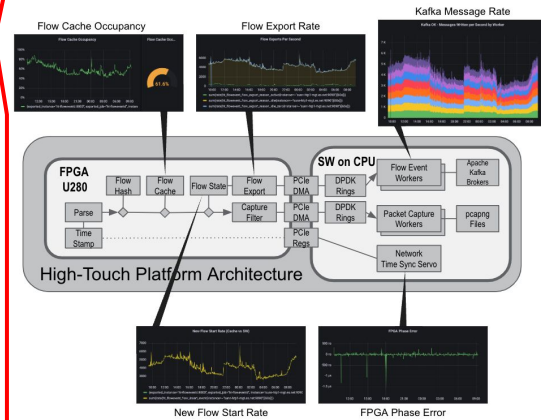
ESnet6 Orchestration & Automation Framework



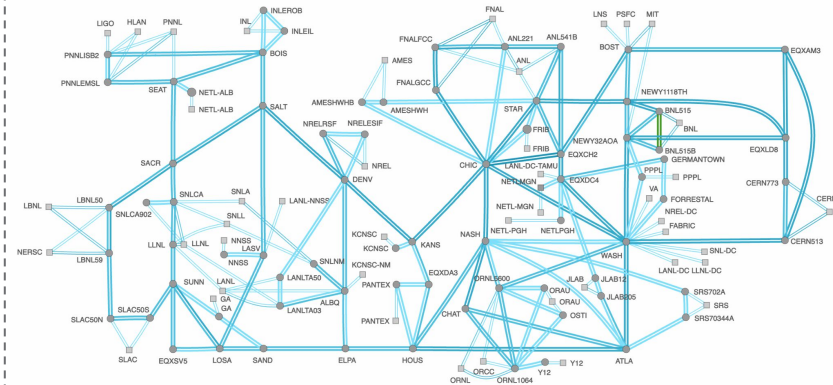
ESnet6 “Hollow” Core Architecture



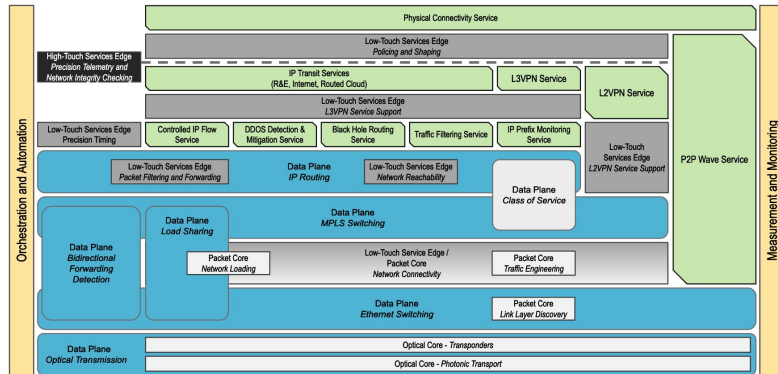
ESnet6 High-Touch Precision Network Telemetry Platform




ESnet6 Network Topology



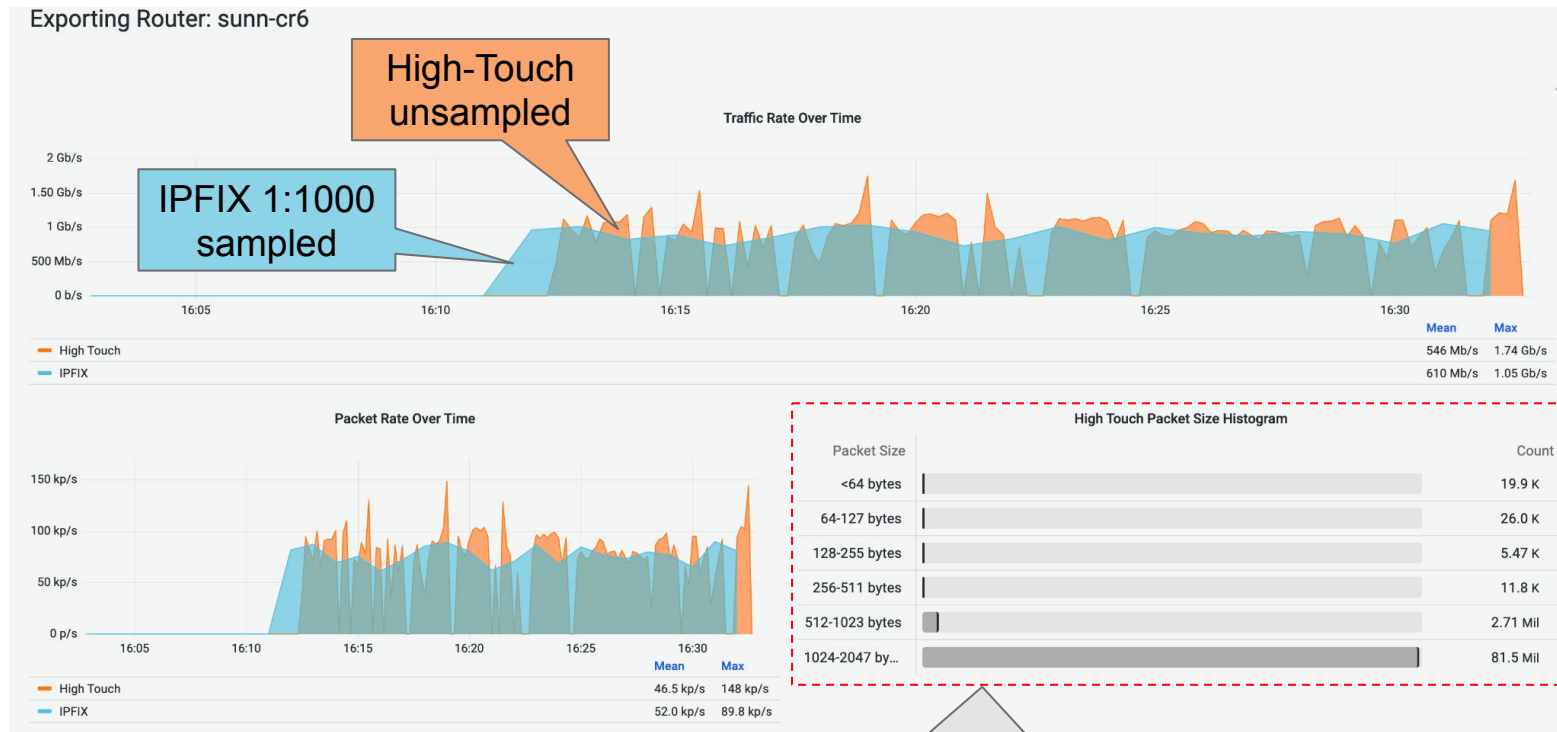
ESnet6 Services and Capabilities Structure



18



Visibility into Network Flow Performance



Packet distribution histogram
generated by High-Touch
FPGA component



Easy Integration with Data Science and ML Libraries

Host fingerprinting to detect anomalous connections

Background radiation monitoring for security breaches

Data import lag analysis for understanding ingest issues

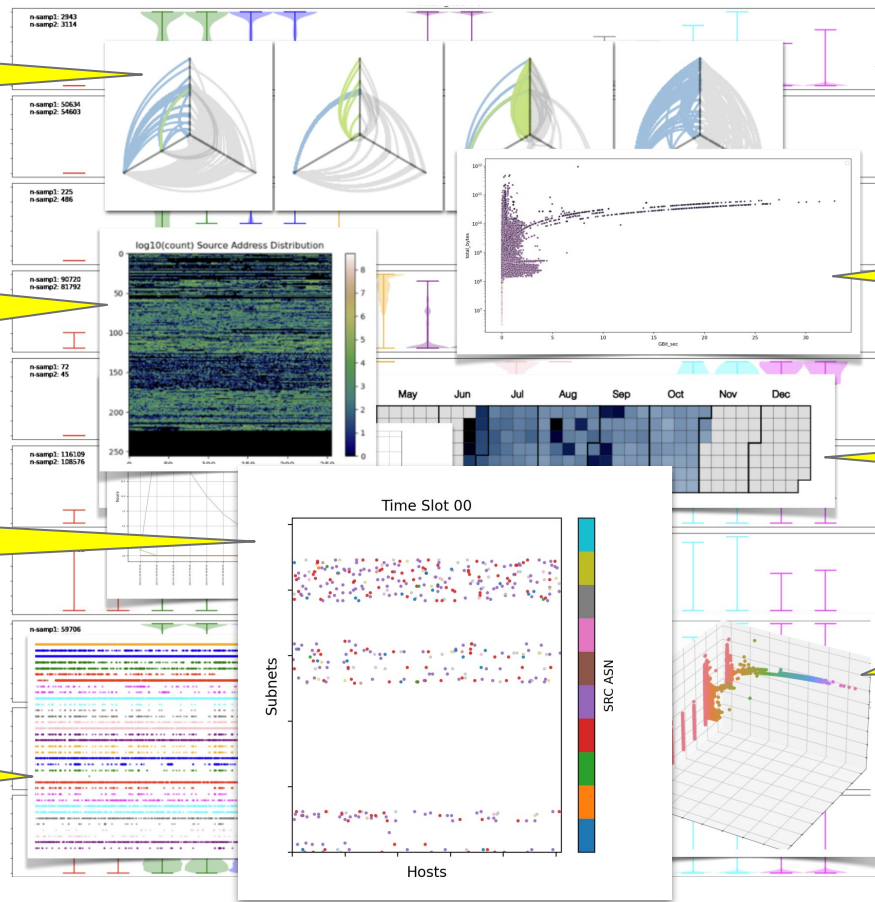
Port and subnet scanning to identify malicious activities

Cluster attribute analysis for identifying common traffic characteristics

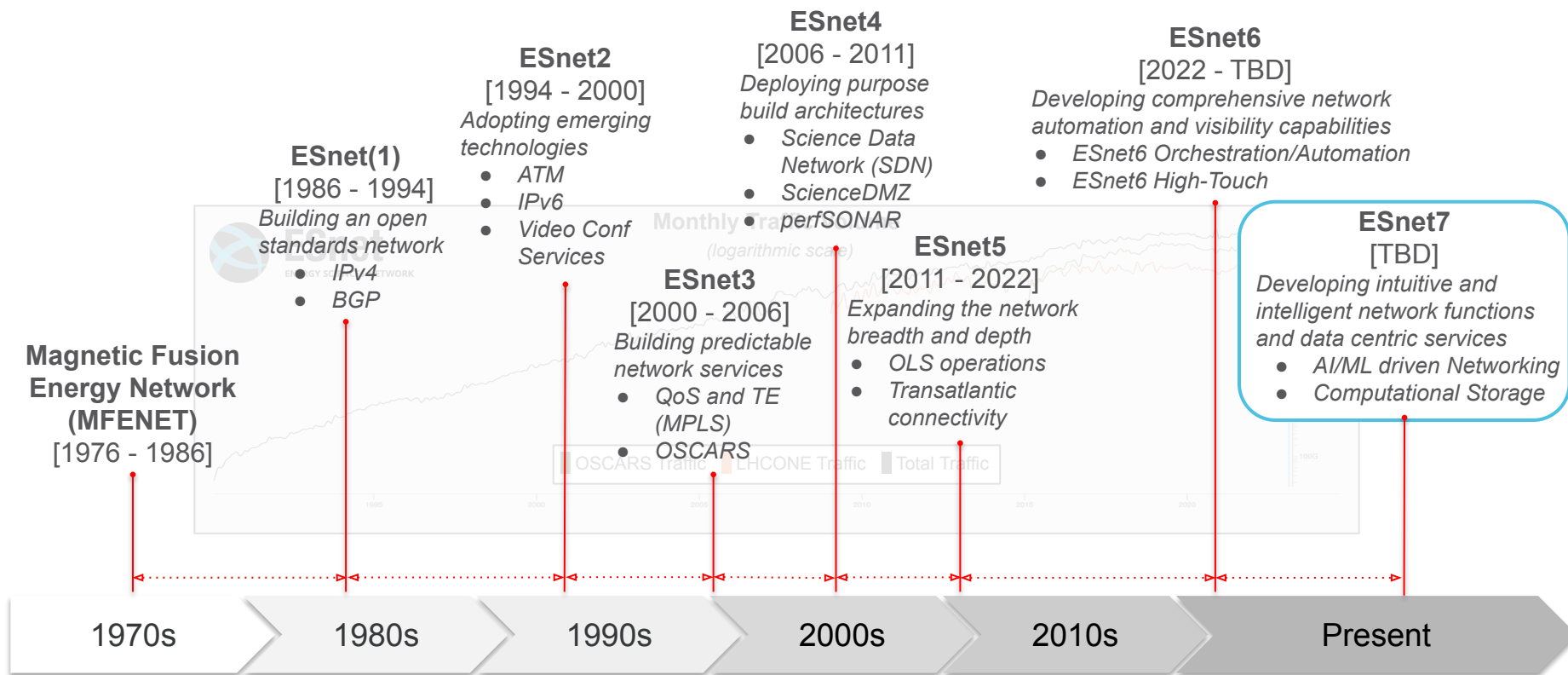
Throughput analysis to determine data movement performance

Calendar view of activity for network planning

Packet size analysis to understand flow characteristics and behaviors



Evolution of the Energy Sciences Network (ESnet)



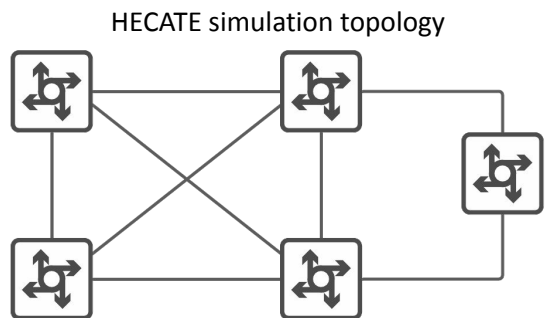
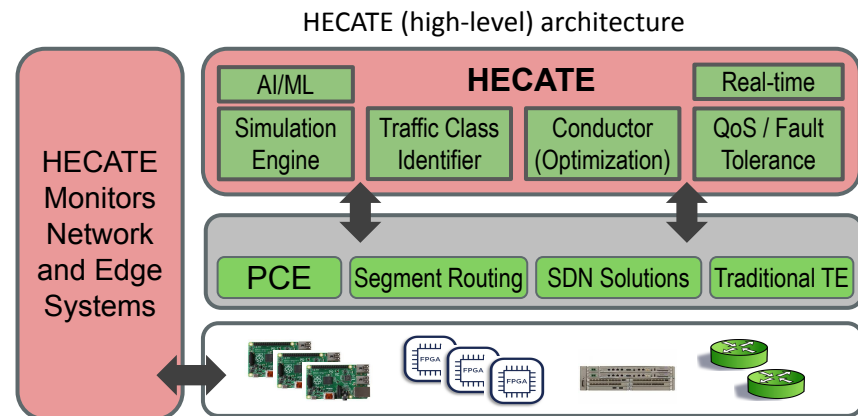
Optimizing Traffic Flows in Real-time : HECATE*

Scientific Achievement

Multi-objective path optimizer driven by historical endpoint behaviors, line and predicted health, and topology.

Significance and Impact

Every network is different and often need constant human attention. We develop a reinforcement learning approach coupled with unsupervised learning to help HECATE learn optimal patterns and then optimize the network when HECATE is turned on.



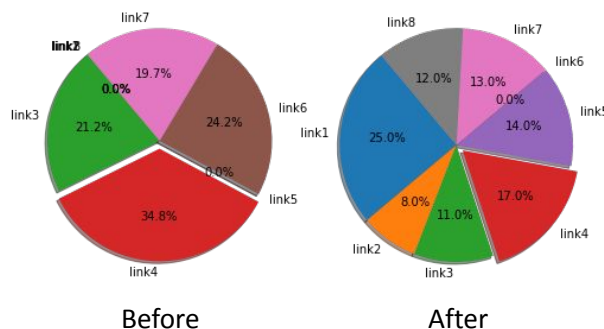
Site characteristics
(slow time refresh)

Link characteristics
(fast time refresh)

Traffic classes:
Jumbo, interactive, default

Parameters:
Loss, delay, jitter

Improvement of load sharing across all links and overall average utilization



Research Details

- No-compromise on performance: Hecate monitors network “health” and actively reroutes traffic
- Caters to many applications: Hecate self-learns traffic classes to guarantee service
- Seamlessly integrate multiple network solutions
- Deployable as hardware solution



*Patent filed – Deep Learning informed Traffic Engineering

In-Network Data Caching

Scientific Achievement

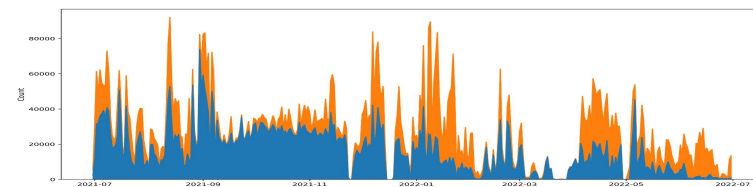
We experimented and demonstrated the capability of a network-based temporary data cache; how in-network caching mechanism helps network traffic performance, how much data can be shared within the network, and how much network traffic volume can be reduced consequently.

Significance and Impact

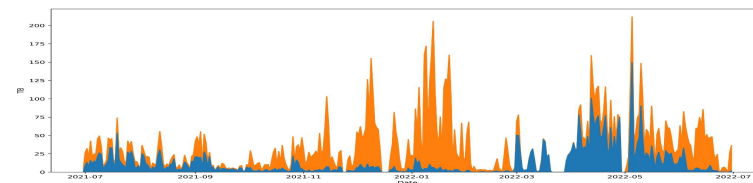
- In-network services such as temporary data caching could potentially have a big impact on traffic engineering and how the remote data is being accessed.
- Data caching mechanism in a region is expected to reduce the redundant data transfers, saved network traffic, and lower data access latency improving overall application performance.
- It also provides the unique capability for a network provider to design data hotspots into the network topology. The appropriate bandwidth resources and traffic engineering techniques can manage traffic movement and congestion.

Research Details

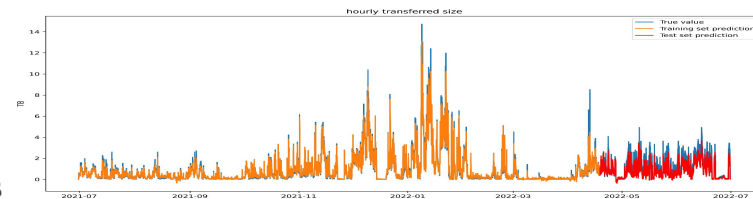
- ESnet cache node as a part of SoCal Petabyte scale cache in collaboration with Caltech, UCSD, and US CMS.
- Studied 1-year's operational from SoCal Repo from Jul 2021 to Jun 2022.
- On average 67.6% of file requests were satisfied by the cache, which translated to 4.5PB (35.4%) of requested bytes (12.7PB) served by the cache.
- Network traffic was reduced by up to 29TB per day due to cached data.
- Sim et al. "Effectiveness and predictability of in-network storage cache for Scientific Workflows", IEEE ICNC, 2023



Daily file requests (count) - cache misses (in orange) and cache hits (in blue)



Daily traffic volume - cache misses (in orange) and cache hits (in blue)



Hourly volume (in bytes) of cache misses from the LSTM model output



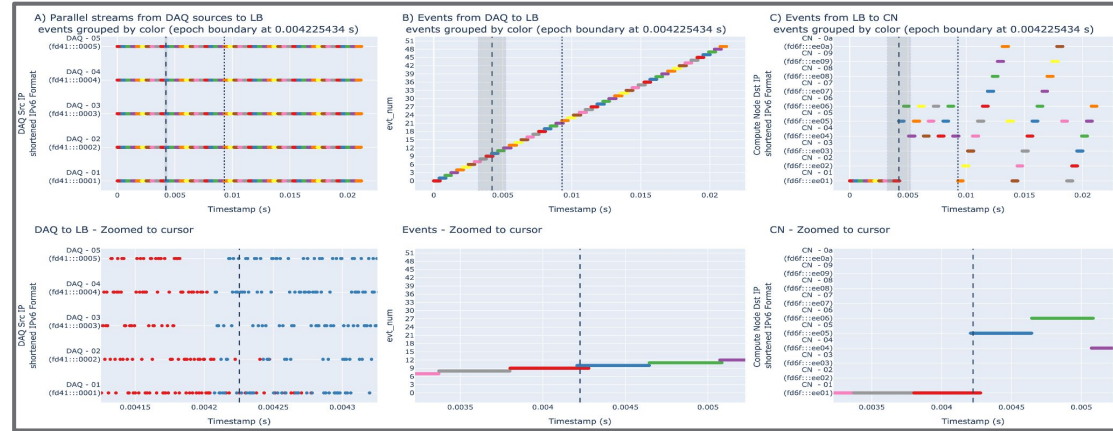
EJ-FAT FPGA Accelerated Transport Load Balancer

Scientific Achievement

The real-time load balancer is designed to support WAN latencies for geographically distributed accelerator facilities and high performance computing centers, and has been successfully integrated with JLab's ERSAP processing pipeline for end-to-end event processing.

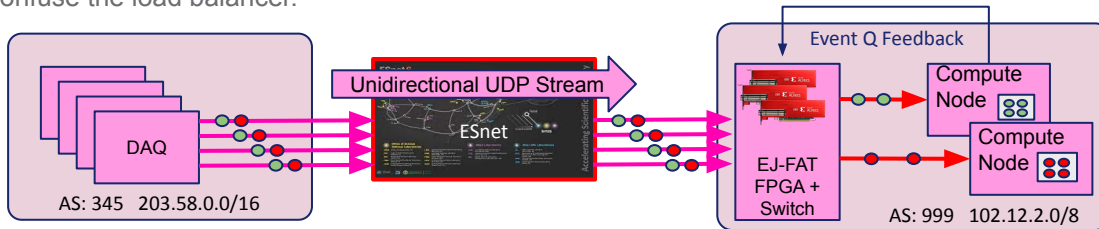
Significance and Impact

- Horizontal scale. Keep adding parallel FPGAs and switches to achieve Terabits of throughput. All the elements work together to get related pieces of data to each compute node.
- Multi-Domain. DAQ source only need to know 1 dst IP for the load balancer. Compute nodes can register their IP address with the LB and receive work.
- Work is broken into 1uS or shorter packet bursts. Zero packet loss or accidental reshuffling in the load balancer, even in dynamic environments with compute nodes changing on the fly. Overlapping event times do not confuse the load balancer.



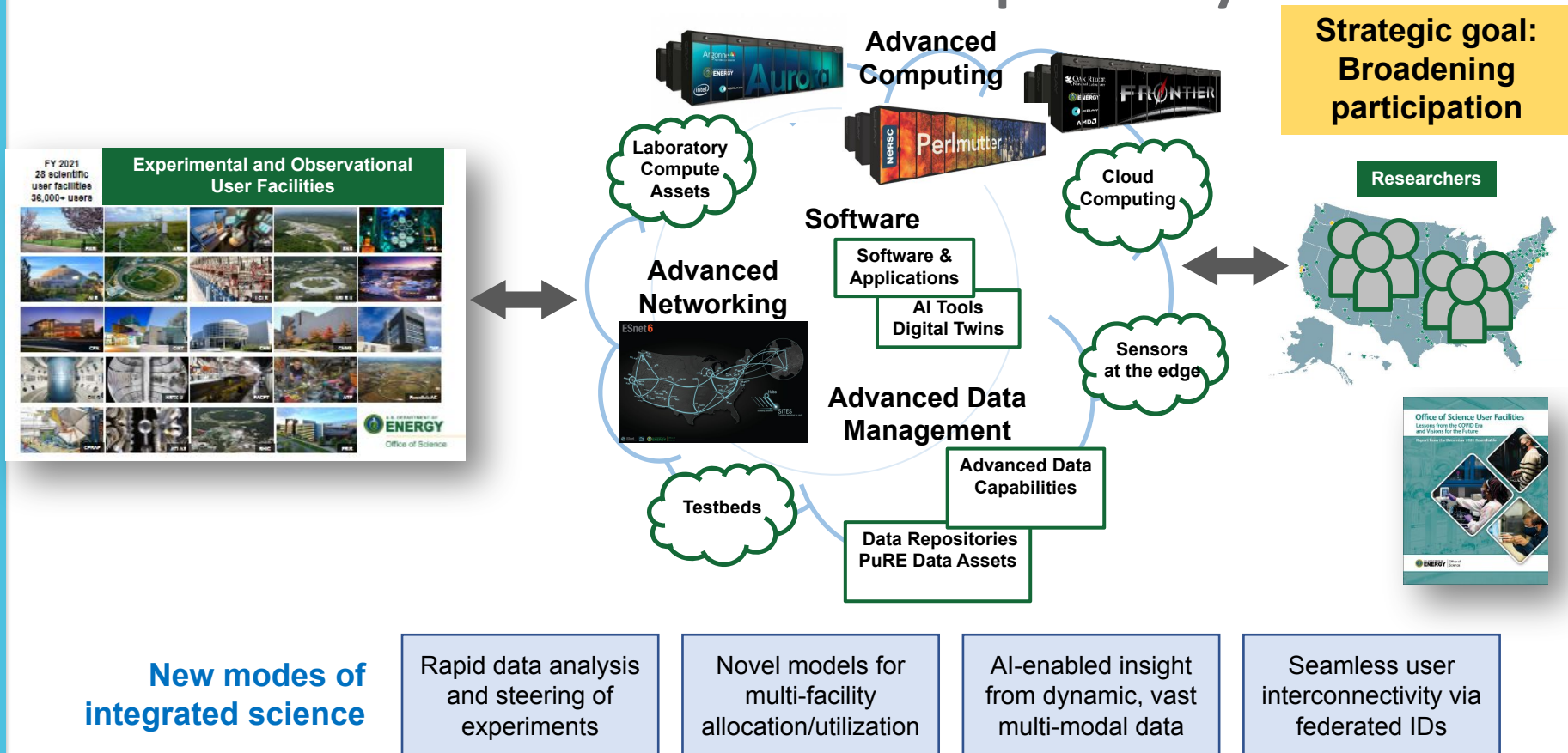
Research Details

- Separation of IP Addresses between labs
- In network sorting of Event Data
- Stateless load balancing
- Compute node feedback for dynamic LB
- Hit-less reconfiguration of LB table
- Unidirectional UDP streaming



What is the big picture objective?

The vision: A DOE/SC **integrated research ecosystem** that transforms science via seamless interoperability



DOE ASCR Integrated Research Infrastructure Vision Statement

“To empower researchers to seamlessly and securely meld DOE’s world-class research tools, infrastructure, and user facilities in novel ways to radically accelerate discovery and innovation.”

Addressing national and societal grand challenges and unlocking new opportunities around energy, science, and technology for US competitiveness will require **highly coordinated, collaborative research and integrating capabilities across our world-leading facilities**, which currently operate largely independently. We can achieve this vision if the facilities, projects, and science communities have the right **incentives, governance, and operating structure** to enable them to **deliver an integrated research platform** – accelerating time-to-discovery and time-to-innovation.

DOE ASCR IRI Task Force contemplated operational models and guiding principles [CY2021]

ASCR Integrated Research Infrastructure Task Force

March 8, 2021

Toward a Seamless Integration of Computing, Experimental, and Observational Science Facilities: A Blueprint to Accelerate Discovery

About the ASCR Integrated Research Infrastructure Task Force

There is growing, broad recognition that integration of computational, data management, and experimental research infrastructure holds enormous potential to facilitate research and accelerate discovery.¹ The complexity of data-intensive scientific research—whether modeling/simulation or experimental/observational—poses scientific opportunities and resource challenges to the research community writ large.

Within the Department of Energy's Office of Science (SC), the Office of Advanced Scientific Computing Research (ASCR) will play a major role in defining the SC vision and strategy for integrated computational and data research infrastructure. The ASCR Facilities provide essential high end computing, high performance networking, and data management capabilities to advance the SC mission and broader Departmental and national research objectives. Today the ASCR Facilities are already working with other SC stakeholders to explore novel approaches to complex, data-intensive research workflows, leveraging ASCR-supported research and other investments. In February 2020, ASCR established the Integrated Research Infrastructure Task Force² as a forum for discussion and exploration, with specific focus on the operational opportunities, risks, and challenges that integration poses. In light of the global COVID-19 pandemic, the Task Force conducted its work asynchronously from April through December 2020, meeting via televideo for one hour every other week. The Director of the ASCR Facilities Division facilitated the Task Force, in coordination with the ASCR Facility Directors.

The work of the Task Force began with these questions: Can the group arrive at a shared vision for integrated research infrastructure? If so, what are the core principles that would maximize scientific productivity and optimize infrastructure operations? This paper represents the Task Force's initial answers to these questions and their thoughts on a strategy for world-leading integration capabilities that accelerate discovery across a wide range of science use cases.

B. Brown, C. Adams, K. Antypas, D. Bard, S. Canon, E. Dart, C. Guok, E. Kissel, E. Lancon, B. Messer, S. Oral, J. Ramprakash, A. Shankar, T. Uram, <<https://doi.org/10.2172/1863562>>

Areas

- AL Allocations
- AC Accounts
- DA Data
- AP Applications
- SC Scheduling
- WF Workflows
- PB Publication
- AR Archiving

Principles

Flexibility. Assembly of resource workflows is facile; complexity is concealed

Performance. Default behavior is performant, without arcane requirements

Scalability. Data capabilities without excessive customizations

Transparency. Security, authentication, authorization should support automation

Interoperability. Services should extend outside the DOE environment

Resiliency. Workloads are sustained across planned and unplanned events

Extensibility. Designed to adapt and grow to meet unknown future needs

Engagement. Promotes co-design, cooperation, partnership

Cybersecurity. Security for facilities and users is essential.

What does this mean for networks*?

Promoting networks as “first class” resources, similar to instruments, compute and storage, e.g.,

- Accessible
 - Security frameworks for accessing (selected) services
 - APIs to interact with services
- Controllable
 - Resource/service selection/negotiation
 - Service scheduling
- Transparent
 - Resource (general) availability
 - Service (specific) status
- Adaptable
 - Ability to integrating compute and/or storage into the network
 - Rapid prototyping of new services

****Networking is an end-to-end service, inter-domain interoperability and service consistency is critical!***



SC Integrated Research Infrastructure Architecture Blueprint Activity (IRI-ABA) [CY2022]

- **Aim:** Produce the **reference conceptual foundations** to inform a coordinated “whole-of-SC” strategy for an integrative research ecosystem.
- **Approach:**
 - Invite DOE experts across the SC User Facilities, SC National Laboratories, and key SC enterprise stakeholders to participate in a series of activities and events:
 - Gather and analyze **integrative use cases** that inclusively span SC programs and user facilities
 - Develop overarching design principles and one or more “architecture blueprints” that will address the chief IRI **design patterns** in an efficient way.
 - Identify **urgent program and lab priorities** and early win opportunities.
- **Intended outcomes:**
 - Produce a shared understanding across SC and DOE of IRI requirements, operational and technical gaps and needed investments, and a common lexicon to describe these.
 - Position SC Programs to contemplate future investment decisions.
 - Explore leveraging existing SC and ASCR resources and services as well as identifying new needs for research and capability gaps for new resources that do not yet exist.
- **Timeline:** February through September 2022.

Common recurring sentiments across the user interviews

Data Management

- Users are overwhelmed with large and growing amounts of data to manage, reduce, analyze
 - Users need to move data across facilities and use different systems at different steps of data processing chain
 - Users need bespoke data movement and workflow solutions, and long-duration support for data/metadata.

Automation/AI

- Users need reliable automation & seamless access and try to compensate via human effort.
 - Users need automation, and anticipate AI, but struggle with skills and application of these novel technologies

Heterogeneity

- Users face mismatches between resources, tools and needs
 - Users need heterogeneity in scale and type of resources but have platform fatigue learning many different platforms
 - Users need workflows to be at the center but need software APIs and standardization/uniformity
 - Users have a spectrum of computing needs from elastic computing (matching need to available resources) to urgent computing (near real-time/just-in-time, on-demand)

Ease of Use

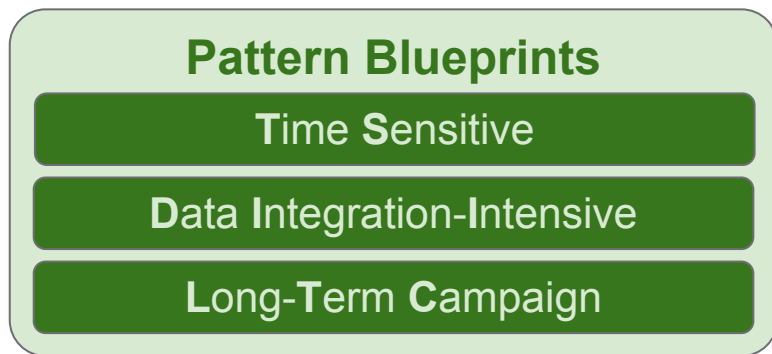
- Users find infrastructure hard to use
 - Users encounter a lack of transparency about workflow tools and resources, and many different use policies and cybersecurity barriers
 - Users need infrastructure to be easier to use and be more coordinated across resources and facilities

Workforce Skills Gap

- Users and teams struggle with workforce and training needs
 - Users (and their organizations) struggle with lack of skills, oversubscribed staff, recruiting, and retention
 - Users experience gaps between their working knowledge and skills and those of infrastructure experts.
 - Users need support and expertise in data science.



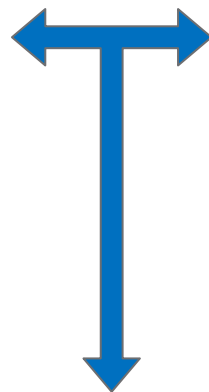
IRI ABA Design Phase



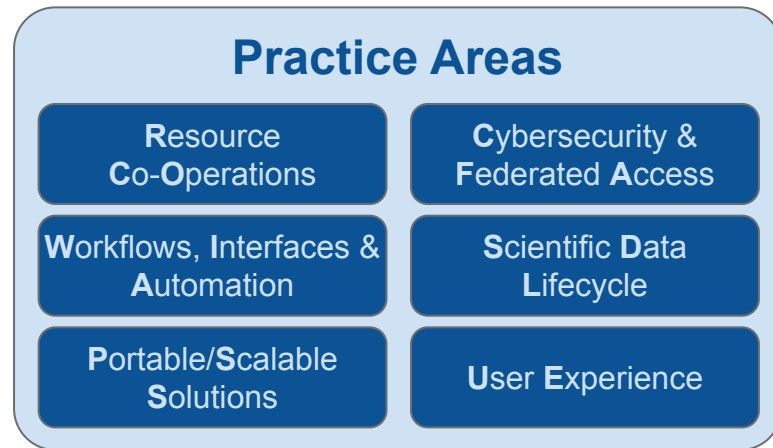
Overarching IRI Principles

Blueprint Compare & Contrast

Governance/Steering



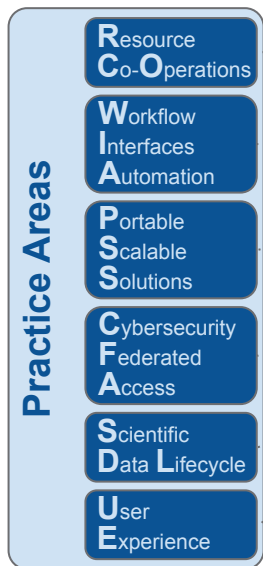
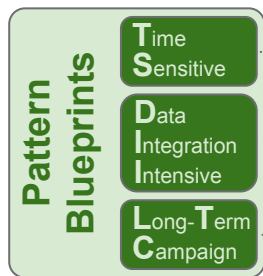
Focus Topics



Coming soon!



IRI ABA Implications for ESnet



Enable **predictable (end-to-end) network services**, e.g., guaranteed bandwidth/latency/jitter, load-balancing, network resiliency

Provide **high bandwidth and rich connectivity**, e.g., capacity planning, Cloud-connect/peering strategies

Support **application/network interaction**, e.g., availability, provisioning, verification, monitoring

Facilitate **“friction-free” data movement**, e.g., low-impedance architectures, data movement tools

Provide/support **network computational storage** capabilities, e.g., workflow integrated edge compute, in-network compute, in-network data caching

Support **multi-modal network connectivity**, e.g., wireless sensor nets

Advocate for supported **programming constructs**, e.g., orchestration/automation, inter-facility APIs, common (portable) programming and runtime environments, software lifecycle, “standardization”

Collaborate on **common access framework**, e.g., cybersecurity, federated access, resource allocations

Support **resource allocation policies**, e.g., (guaranteed/transferrable) resource allocations, facility metrics

Encourage **development and testing environments**, e.g., (federated) testbeds, prototyping collaborations

Facilitate **co-design services**, e.g., design patterns, standard practices

Empower **engagement and partnerships**, e.g., outreach, practice groups, forums



High Performance Data Facility (HPDF)

DEPARTMENT OF ENERGY (DOE)
OFFICE OF SCIENCE (SC)
ADVANCED SCIENTIFIC COMPUTING RESEARCH (ASCR)



HIGH PERFORMANCE DATA FACILITY (HPDF)
DOE NATIONAL LABORATORY PROGRAM ANNOUNCEMENT NUMBER:
LAB 23-3020

ANNOUNCEMENT TYPE: INITIAL

Announcement Issue Date:	March 10, 2023
Submission Deadline for Letter of Intent:	March 31, 2023, at 5:00 PM Eastern Time
Submission Deadline for Proposal:	May 5, 2023, at 5:00 PM Eastern Time

“To meet these challenges, SC is advancing the Integrated Research Infrastructure (IRI) vision: DOE will empower researchers to seamlessly and securely meld DOE’s world-class research tools, infrastructure, and user facilities in novel ways to radically accelerate discovery and innovation.”

“The High Performance Data Facility (HPDF) will serve as a foundational element in enabling the DOE Integrated Research Infrastructure.”

“DOE requires a dynamic and scalable data management infrastructure that is network-integrated with the DOE computing ecosystem, with diverse capabilities”

“DOE requires a diversified computing ecosystem that can provide researchers with access to an appropriate computing resource at the appropriate time”



Questions...

Chin Guok <chin@es.net>

