U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

ESnet

ESnet Strategic Vision

Accelerating DOE scientific research over the next decade

Chin Guok <chin@es.net>

Chief Technology Officer

ESnet

5GRP

Osaka, Japan

Sep 17, 2024

Background and Motivation

"The mission of DOE's Office of Science is to deliver scientific discoveries and major scientific tools to transform our understanding of nature and advance the energy, economic, and national security of the United States."

ESnet

On May 29, 2024, ASCAC approved the facilities 2025-35 report responding to the charge issued in December 2023 by the SC Director


U.S. DEPARTMENT OF ENERGY

May 2024

2024 ADVANCED SCIENTIFIC COMPUTING ADVISORY COMMITTEE

FACILITIES SUBCOMMITTEE RECOMMENDATIONS

Ed Seidel
University of Wyoming, Chair

Amanda Randles
Duke University, Co-Chair

Rick Arthur
GE Aerospace

Keren Bergman
Columbia University

Bill Carlson
IDA Center for Computing Sciences

Ewa Deelman
University of Southern California

Ray Grout
National Renewable Energy Laboratory

Bruce Hendrickson
Lawrence Livermore National Laboratory

Dan Reed
University of Utah

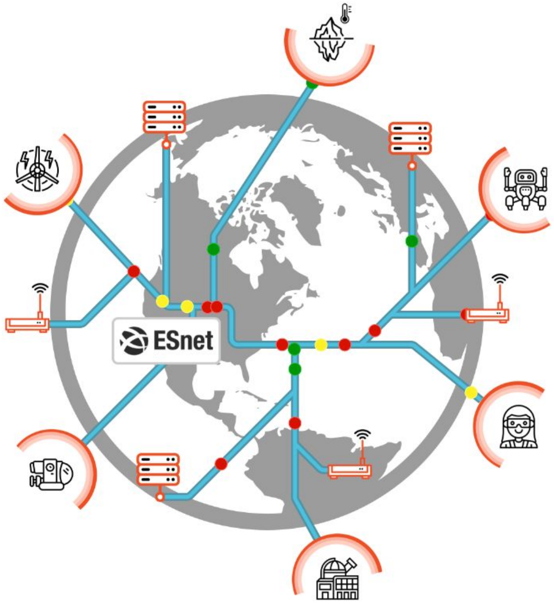
https://science.osti.gov/-/media/ascr/ascac/pdf/reports/2024/FinalReport_May_2024_2370379.pdf

This charge was only the third time the Office of Science (SC) conducted an SC-wide projects assessment. The previous two were extremely consequential.

- In 2003, Director Ray Orbach undertook the first projects assessment.
  - Delivered the *Facilities for the Future of Science* report.
  - ASCR Results:
    - Led to creation of the Leadership Computing Facilities;
    - Affirmed NERSC and ESnet upgrades.

- In 2012, OMB requested the second ever projects assessment.
  - Director Bill Brinkman elected to parallelize and democratize the process by issuing a charge to the six program Federal Advisory Committees (ASCAC, BESAC, BERAC, FESAC, HEPAP, NSAC). *Note: the 2023 charge used the same approach with the identical rating rubric.*
  - Acting Director Pat Dehmer succeeded Dr. Brinkman in 2013 and declined to issue a single rank ordering of all SC projects.
  - ASCR Results:
    - Led to the creation of the Virtual Data Facility, today the High Performance Data Facility.
    - Affirmed LCFs, NERSC, and ESnet upgrades.

Where are we today?



ESnet's Vision

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

ESnet's DNA includes integrating innovative ideas into operationally reliable infrastructure across the last 3 decades

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

ESnet(1)
[1986 - 1994]
Building an open standards network
- *IPv4*
- *BGP*

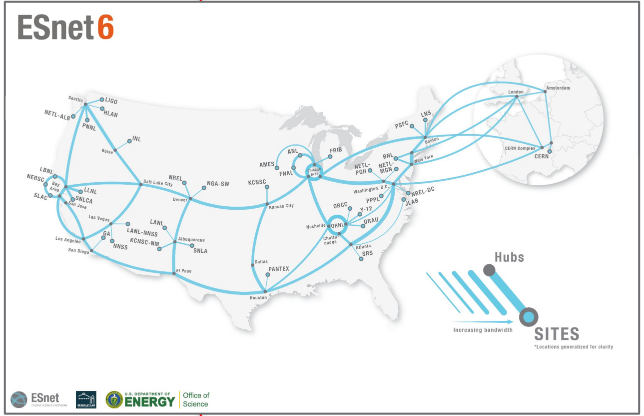
ESnet2
[1994 - 2000]
Adopting emerging technologies
- *ATM*
- *IPv6*
- *Video Conf Services*

ESnet3
[2000 - 2006]
Building predictable network services
- *QoS and TE (MPLS)*
- *OSCARS*

ESnet4
[2006 - 2011]
Deploying purpose build architectures
- *Science Data Network (SDN)*
- *ScienceDMZ*
- *perfSONAR*

ESnet5
[2011 - 2022]
Expanding the network breadth and depth
- *OLS operations*
- *Transatlantic connectivity*

ESnet6
[2022 - TBD]
Developing comprehensive network automation and visibility capabilities
- *ESnet6 Orchestration/Automation*
- ***ESnet6 High-Touch***



1970s | 1980s | 1990s | 2000s | 2010s | Present

By the numbers*

17 National Laboratories

28 DOE User Facilities

271 R&E, Commercial, and Other Networks

126 Staff and contractors

23 States Where Staff Work

15,000 MILES of dedicated fiber-optic cable

3 100 Gbps trans-Atlantic circuits

3 400 Gbps trans-Atlantic circuits

99.999% Site Uptime

1.7 EXABYTES Transited ESnet in 2023

46 TBPS Aggregate Capacity

400 GBPS to 1 TBPS Network Backbone

4.86 User Satisfaction Rating (out of 5)

*NB: As of Mar 31, 2024

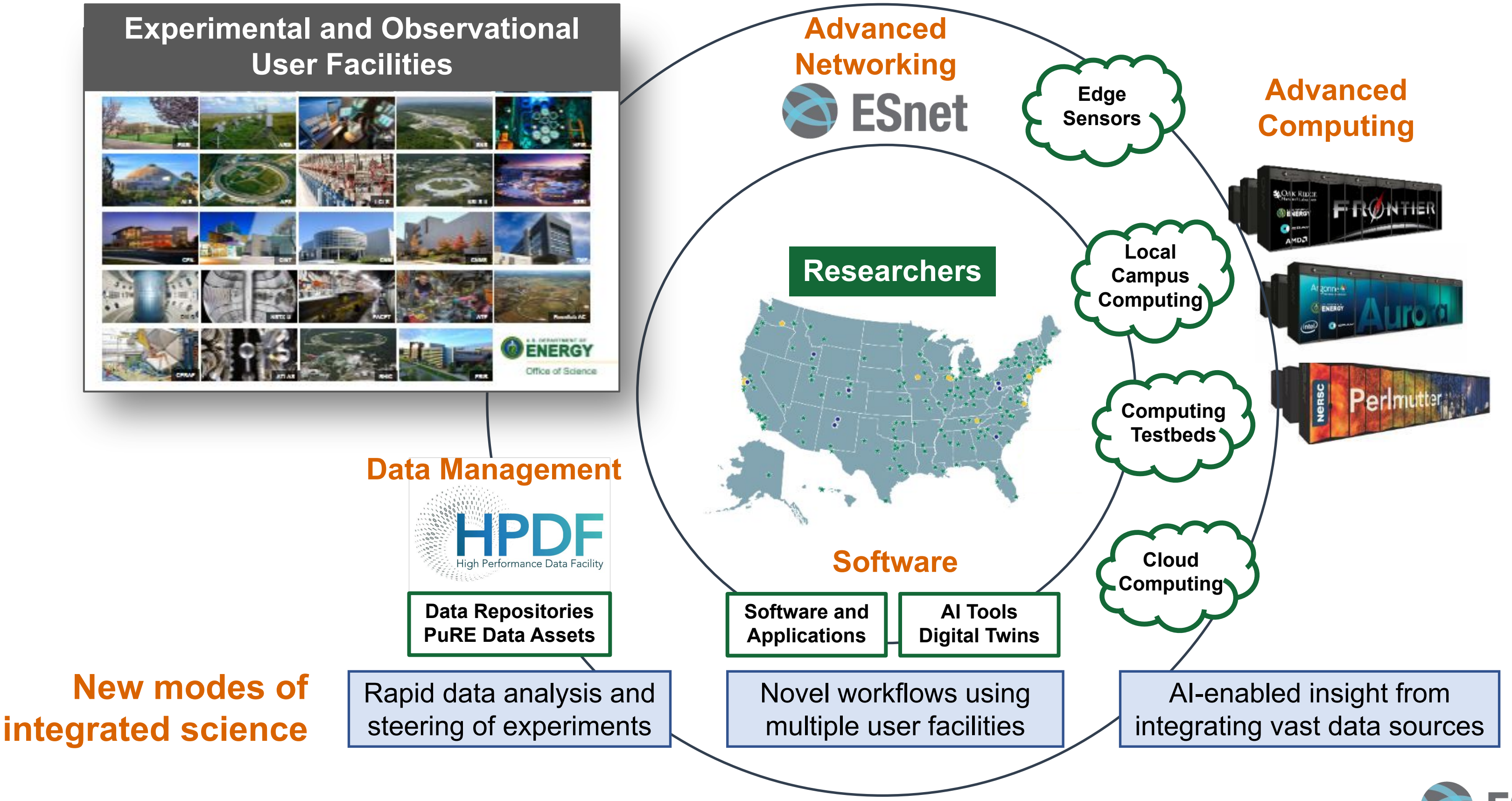
ESnet

What is driving our evolution?

- Integrated Research Infrastructure (IRI)
- High Performance Data Facility (HPDF)
- Artificial Intelligence (AI) for Science

ESnet

ESnet's central role in Integrated Research Infrastructure (IRI) - Connecting researchers and user facilities to HPC, HPDF, Cloud, testbeds, and the edge



DOE SC Integrated Research Infrastructure Architecture Blueprint Activity (IRI-ABA) - Understanding the ecosystem

Convened over **150 DOE national laboratory experts** from **all 28 SC user facilities** across **13 national laboratories** to consider the **technological, policy, and sociological challenges** to implementing IRI.

IRI Science Patterns (3)

Time-sensitive pattern has *urgency*, requiring real-time or end-to-end performance with high reliability, e.g., for timely decision-making, experiment steering, and virtual proximity.

Data integration-intensive pattern requires combining and analyzing data from multiple sources, e.g., sites, experiments, and/or computational runs.

Long-term campaign pattern requires sustained access to resources over a long period to accomplish a well-defined objective.

IRI Practice Areas (6)

User experience practice will ensure relentless attention to user perspectives and needs through requirements gathering, user- centric (co)-design, continuous feedback, and other means.

Resource co-operations practice is focused on creating new modes of cooperation, collaboration, co-scheduling, and joint planning across facilities and DOE programs.

Cybersecurity and federated access practice is focused on creating novel solutions that enable seamless scientific collaboration within a secure and trusted IRI ecosystem.

Workflows, interfaces, and automation practice is focused on creating novel solutions that facilitate the dynamic assembly of components across facilities into end-to-end IRI pipelines.

Scientific data life cycle practice is focused on ensuring that users can manage their data and metadata across facilities from inception to curation, archiving, dissemination, and publication.

Portable/scalable solutions practice is focused on ensuring that transitions can be made across heterogeneous facilities (portability) and from smaller to larger resources (scalability).

ESnet

IRI-ABA implications for ESnet(7)

Pattern Blueprints

- **T**ime **S**ensitive
- **D**ata **I**ntegration **I**ntensive
- **L**ong-**T**erm **C**ampaign

Practice Areas

- **R**esource **C**o-**O**perations
- **W**orkflow **I**nterfaces **A**utomation
- **P**ortable **S**calable **S**olutions
- **C**ybersecurity **F**ederated **A**ccess
- **S**cientific **D**ata **L**ifecycle
- **U**ser **E**xperience

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 metics

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

Building capabilities to support IRI touches many ESnet areas

Predictable (end-to-end) network services

- OSCARS guaranteed b/w dynamic provisioning
- Operational measurement & performance monitoring

High bandwidth and rich connectivity

- ESnet6 capacity deployment
- Transatlantic spectrum
- ESnet Cloud Connect for Virtual Private Clouds (VPC)

Application/network interaction

- SENSE multi-domain resource orchestration
- OpenAPI & OpenTelemetry

"Friction-free" data movement

- ScienceDMZ data transfer optimization architecture
- Petascale DTN for HPC oriented data transfers

Network computational storage

- DTNaaS in-network caching
- EJFAT FPGA based real-time DSP processing for edge compute

Multi-modal network connectivity

- CBRS pilot field deployment (EESA) Mt. Crested Butte, Co

Programming constructs

- SURF's Workflow Orchestrator
- SENSE and NSI network interdomain APIs
- JANUS container management

Common access framework

- Federated Identity and Access Control Management

Resource allocation policies

- ESnet does not have an allocation policy, but this may change to support IRI (e.g., time sensitive workflows)

Development and testing environments

- ESnet Testbed Next-Gen
- FABRIC (NSF) Testbed
- ROVER network orchestration testing environment

Co-design services

- GRETA network integration
- EJFAT ESnet JLab FPGA Accelerated Transport

Engagement and partnerships

- DOE program requirements reviews
- ESnet ConFAB

Indicators show ESnet's current experience or implementation status in the area.

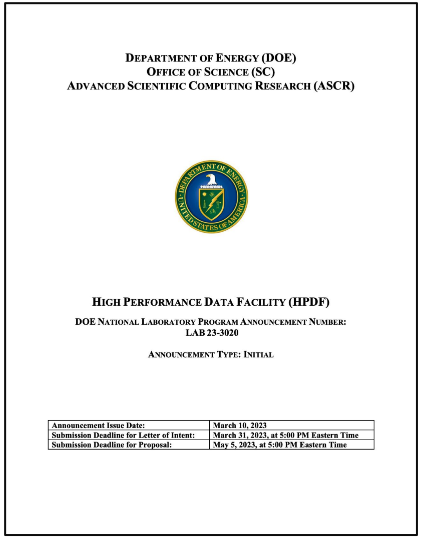
ESnet

What is driving our evolution?

- Integrated Research Infrastructure (IRI)
- High Performance Data Facility (HPDF)
- Artificial Intelligence (AI) for Science

ESnet

ESnet services are integral to the High Performance Data Facility (HPDF) distributed facility



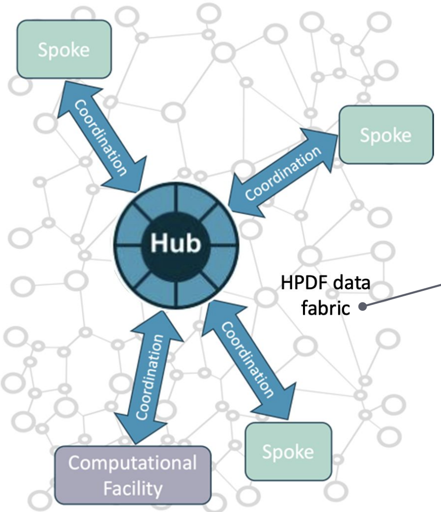
DEPARTMENT OF ENERGY (DOE)
OFFICE OF SCIENCE (SC)
ADVANCED SCIENTIFIC COMPUTING RESEARCH (ASCR)

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



Integration with ESnet will be key for

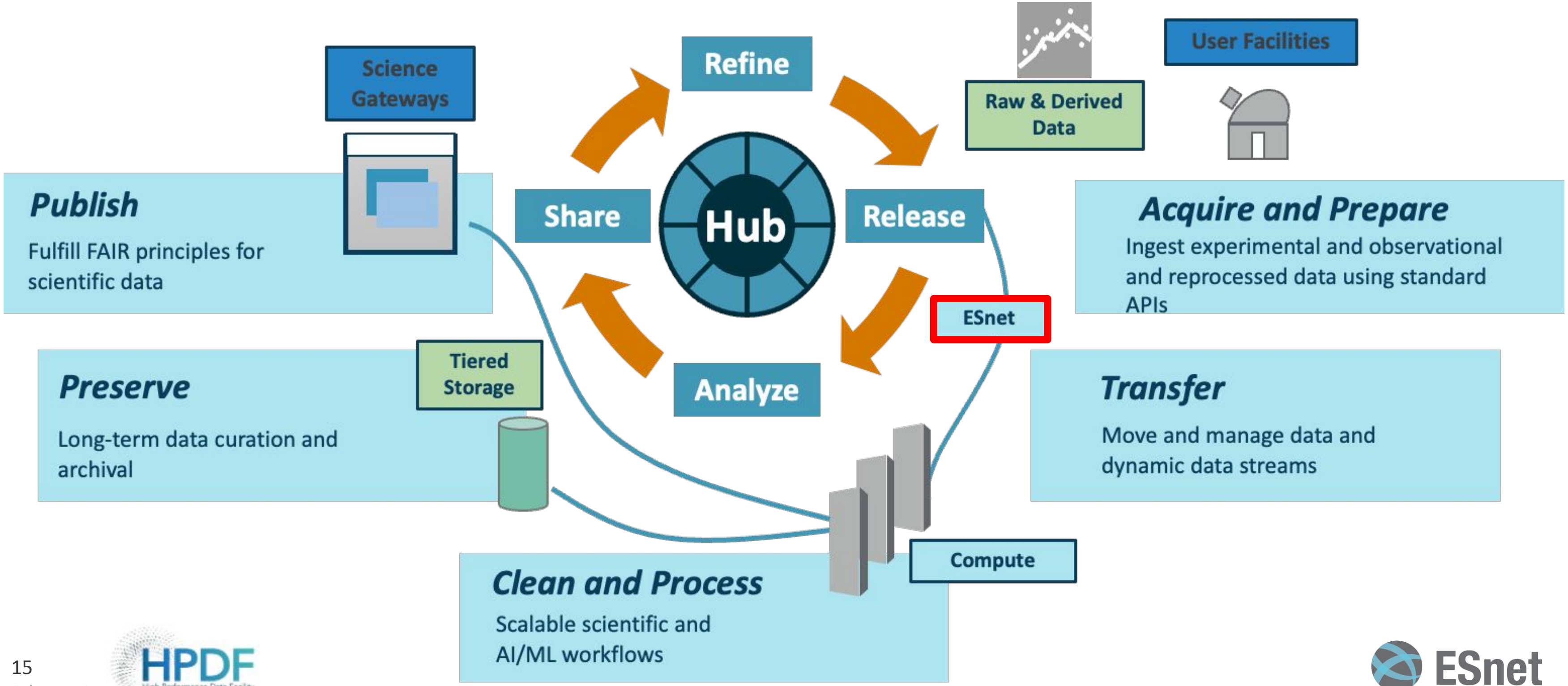
- Advanced data services to handle science workflows
- Geographically and operationally resilient active-active failover
- Deploying distributed computing or storage resources between Hub(s) and Spokes

“HPDF will provide crucial resources to Office of Science programs to attack fundamental problems in science and engineering that require nimble **shared access to large data sets**, increasingly aggregated from multiple sources.”

“The facility will be designed to dynamically configure computation, network resources and storage to **access data at rest or in motion**, supporting the use of **well-curated datasets** as well as near real-time analysis on **streamed data** directly from experiments or instruments.”

HPDF Will Set the Standard for Data Life Cycle Management

Data science requires curated and annotated data that adheres to FAIR principles, and data reuse will be a metric for HPDF. Office of Scientific and Technical Information services will complement HPDF to provide full life cycle coverage.

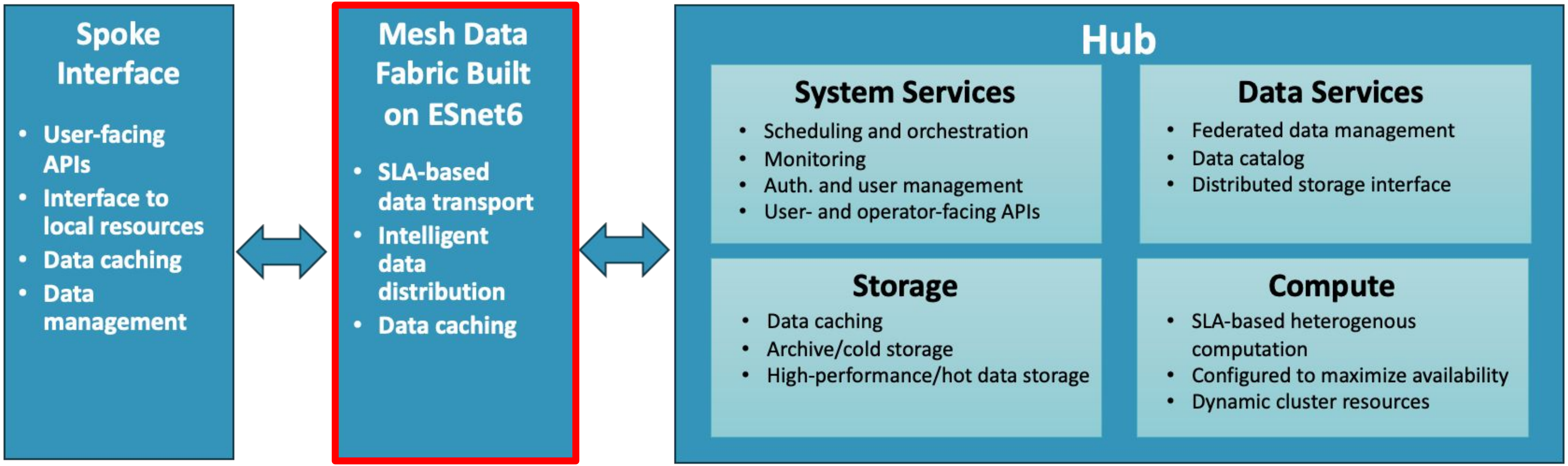


Boehnlein et al

HPDF Technical Design Core Capabilities

Hub Computing and Data Infrastructure	Distributed Spoke Infrastructure	Data-centric Orchestration of Hardware, Software, and Services
• High uptime • Experiment-friendly availability • Data-driven agility • Support for new technologies • Data storage, management, and interoperability • Data preservation	• User support • Scientific application tailoring • Hardware resources that mirror, supplement, or complement hub resources • Low-latency or high-bandwidth coupling of HPDF services to edge compute	• High availability • High-performance mesh data transport fabric • Secure data paths • Monitoring • Orchestration

High-Level HPDF Technical Concept



Design methodology, qualification, and approach:

- Pilot and phased delivery, enable early development, fine tune design
- Use of proven technologies to ensure a reliable, robust platform
- Hardware distributed and replicated at both sites to improve reliability and geographic diversity
- Modular heterogeneous approach to support a broad range of analysis

Approach to delivery and modularity allows composition adjustment during the design phase

What is driving our evolution?

- Integrated Research Infrastructure (IRI)
- High Performance Data Facility (HPDF)
- Artificial Intelligence (AI) for Science

ESnet

Distributed, large data infrastructure and movement needed for foundational and large-parameter AI models

Chapter 19: Data Infrastructure

- In pursuit of the active collective memory concept introduced above, we may imagine a **malleable, tiered set of AI foundation models with high bandwidth connections**.
- These varying foundation models would also connect and coalesce, as **relationships are discovered between the different data sources**, either by the growing understanding of domain scientists or through connections made by computational analysis at scale.
- Data science applications require <u>new capabilities</u>, such as **fast, smart response to new data (e.g., from a new experiment); rapid, random access reads (e.g., when training a foundation model); edge or in-transit processing capabilities (e.g., to filter out interesting events from a high-rate experimental data stream); and continual update of data and knowledge bases as new data appears**

https://www.anl.gov/sites/www/files/2024-05/AI4SESReport-2023-v7.pdf

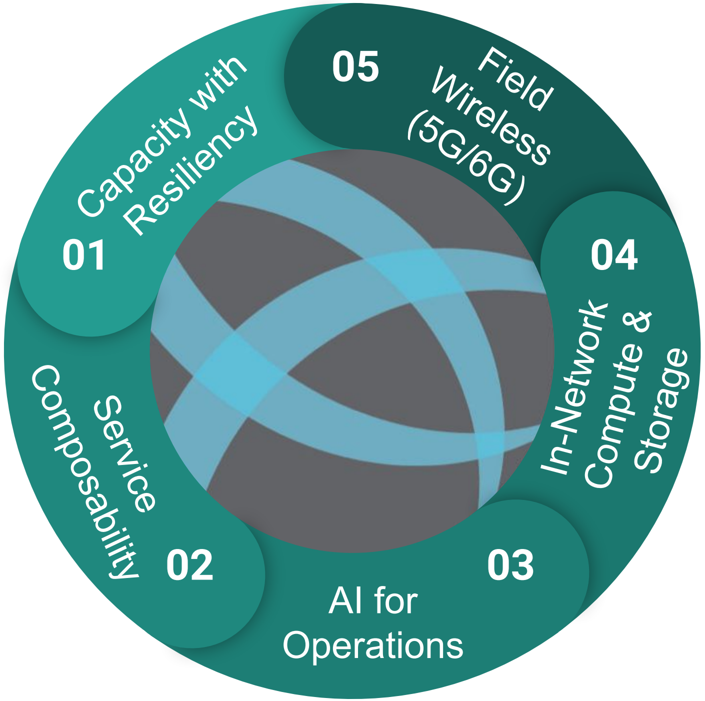
Distributed, large data infrastructure and movement needed for foundational and large-parameter AI models

"The scientific community believes AI can have a foundational impact on a broad range of DOE missions, including science, energy, and national security. Further, DOE has unique capabilities that enable the community to drive progress in scientific use of AI, building on long-standing DOE strengths and investments in computation, data, and communications infrastructure, spanning the Energy Sciences Network (ESnet), the Exascale Computing Project (ECP), and integrative programs such as the NNSA Office of Defense Programs Advanced Simulation and Computing (ASC) and the SC Scientific Discovery through Advanced Computing (SciDAC) programs."

https://www.anl.gov/sites/www/files/2024-05/AI4SESReport-2023-v7.pdf

ESnet

What is in our future?

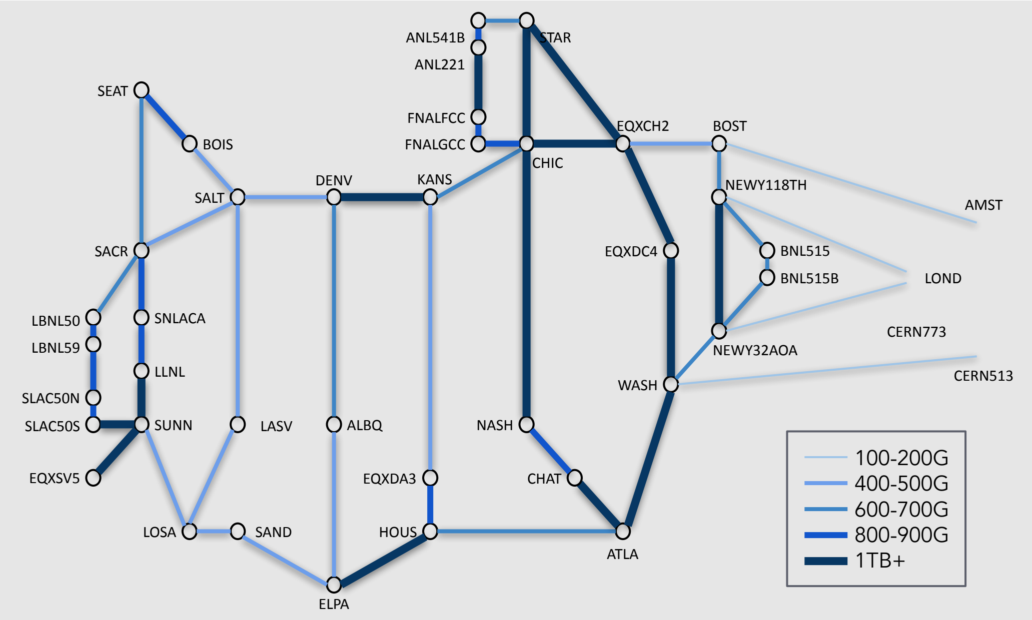


ESnet

ESnet6 laid the foundation for future of data-intensive DOE science

Enough base capacity and ability to cost-effectively add more provides unconstrained access to data, no matter how big or distributed

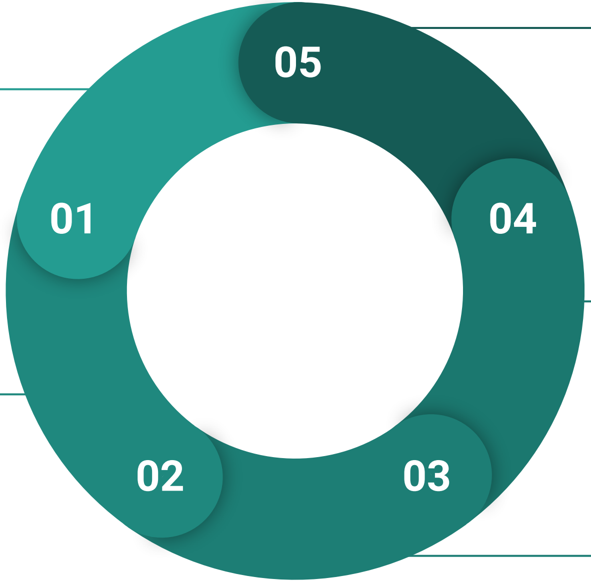
- A **software architecture and network workflow orchestration framework** that hastens our ability to **create new, custom services** and allows to **scale the network** rapidly
- **An innovative packet microscope capturing every packet *header*** across the network: an enviable unlabeled data-set to explore building a **foundational AI model for network**
- **Well-defined service architecture** that reduces snowflakes and brings consistency across sites



ESnet7: Build new capabilities and services on ESnet6 foundation to deliver on science needs, including ASCR strategy on IRI, HPDF, and AI for Science

Capacity with Resiliency
Scale the capacity of the network cost-effectively while increasing resiliency

Network-Application Service Composability
Automation and orchestration-related capabilities for intelligent decision-making and custom workflow services



Advanced Wireless
Transform field and small sensor-driven science by deploying private 5G and LEO connectivity for near real-time data collection and processing.

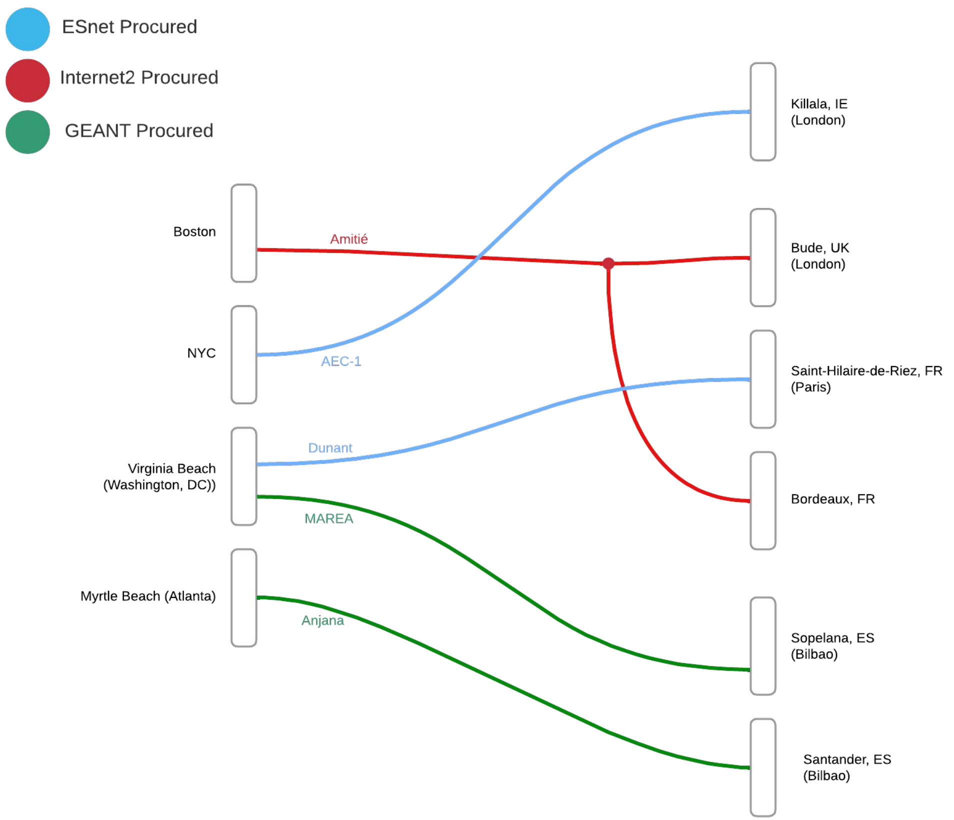
In-Network Storage and Compute
In-network pre-processing, caching, and data transfer assists for science workflows

Artificial Intelligence for IT Operations (AIOps)
Data and analytics-driven systems operations leveraging AI/ML models

Preliminary work in all areas is informing potential ESnet7 components and reducing risk.

ESnet

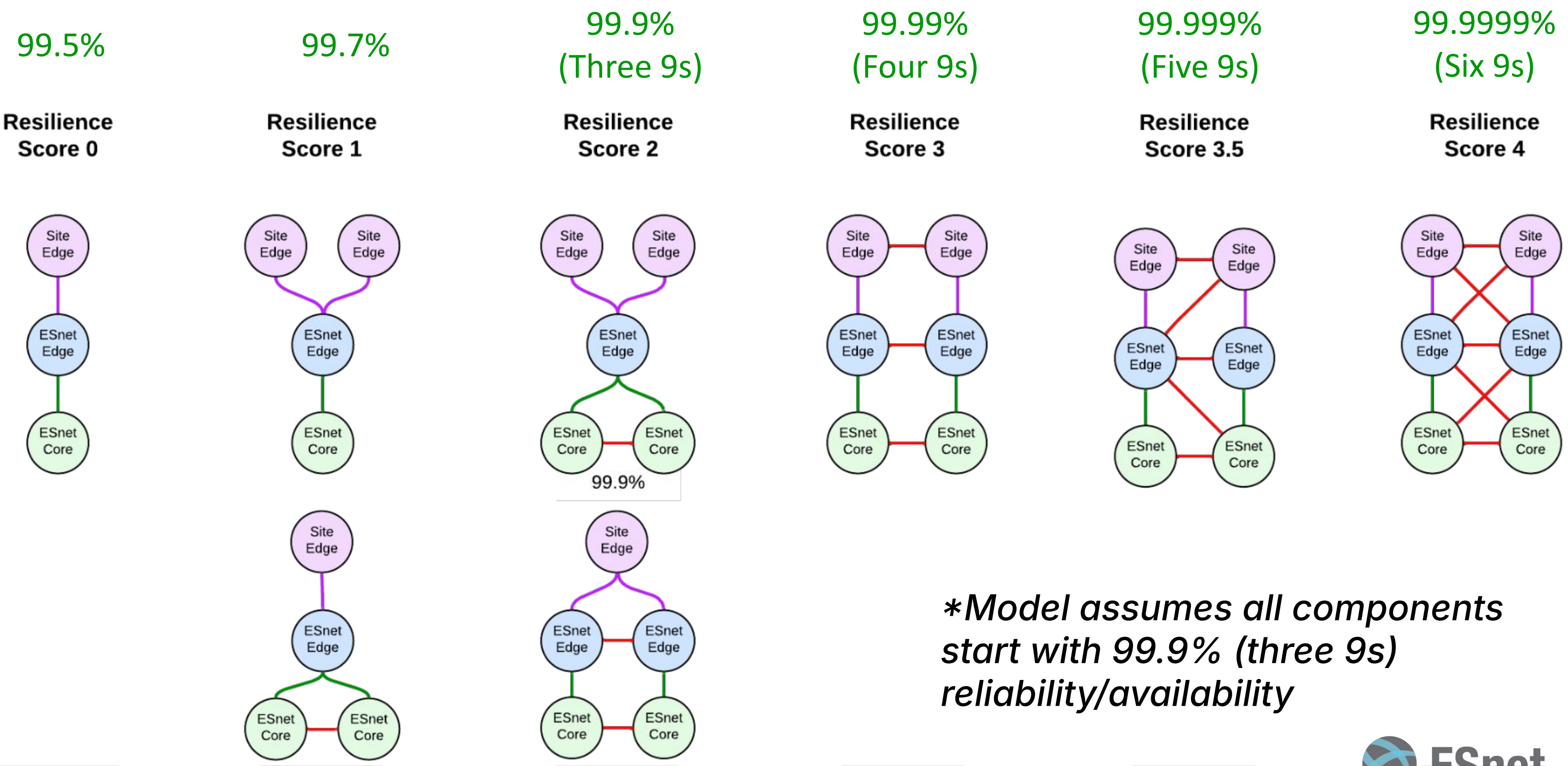
Capacity with Resiliency: Continue enhancing fiber footprint terrestrially and internationally



- 2027 commitment to LHC community: 3.2Tbps of ESnet transatlantic bandwidth
- Plan: collectively acquire optical spectrum across 4+ diverse cables
- Subsea spectrum IRUs are 15+ year contracts - low operating costs
- Subsea spectrum should provide ~10Tbps aggregate to ESnet (to meet HL-LHC commitments plus growth)

ESnet

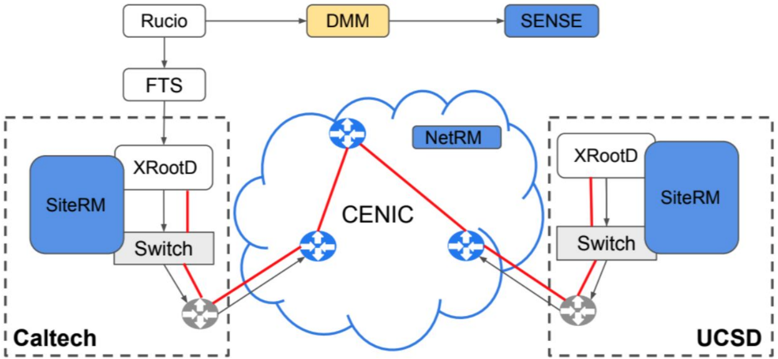
Site Resiliency: Driving understanding across the DOE complex on resiliency and infrastructure risk, crucial for meeting the needs of IRI and HPDF workflows



∗Model assumes all components start with 99.9% (three 9s) reliability/availability

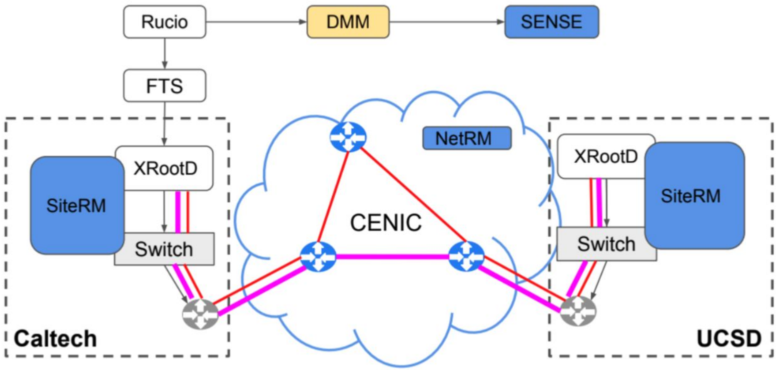
ESnet

Network Application Service Composability: Integration of data management software with network APIs enabling applications to request network outcomes



For **non-priority** Rucio request, Rucio will contact the Data Movement Manager (DMM) and receive endpoints that use the (red) path for **best-effort data movement**.

SENSE provides end-to-end (network) orchestration functions for **IRI Time-Sensitive Pattern** workflows

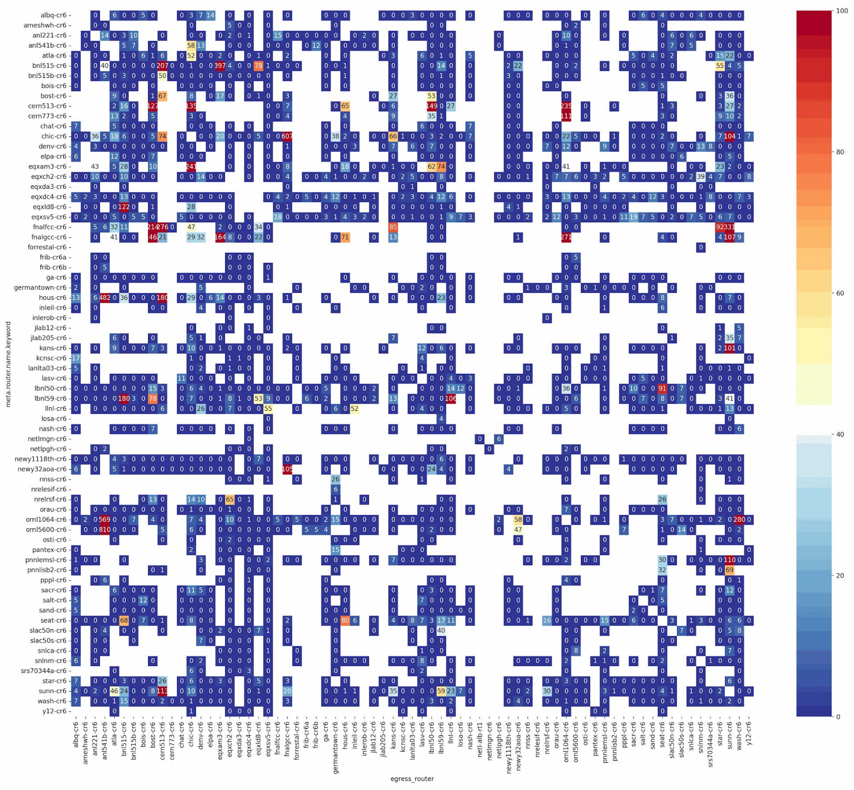


For **priority** Rucio request, Rucio will contact the DMM for endpoints associated with the (pink) guaranteed bandwidth path. The DMM would concurrently request a bandwidth allocation from SENSE to set up the guaranteed bandwidth path. SENSE will instruct both the SiteRm and NetRM to implement specific routing and QoS, facilitating an **end-to-end guaranteed bandwidth data movement**.

ESnet

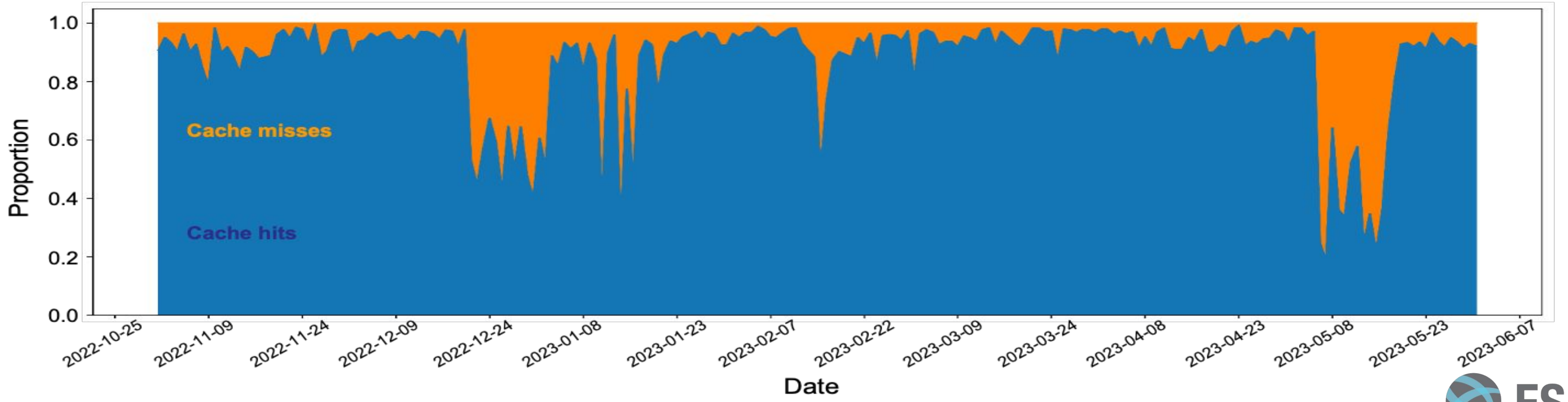
AIOps: Long-term capacity prediction

- Use flow data and topology information to understand capacity requirements
- Conduct what-if analysis and simulations for traffic matrix and network scenarios
- Predict effects of outages on traffic loads



In-network Storage and Computing: Science data caching improves data accessibility

- Pilot installations in 5 locations: currently supporting LHC CMS, DUNE, and LIGO datasets
- Leveraging Open Science Grid caching solution with ESnet's DTN-as-a-service virtualization software stack
- Early studies show lower latency of access for scientists and the reduced traffic on network backbone — a win-win result



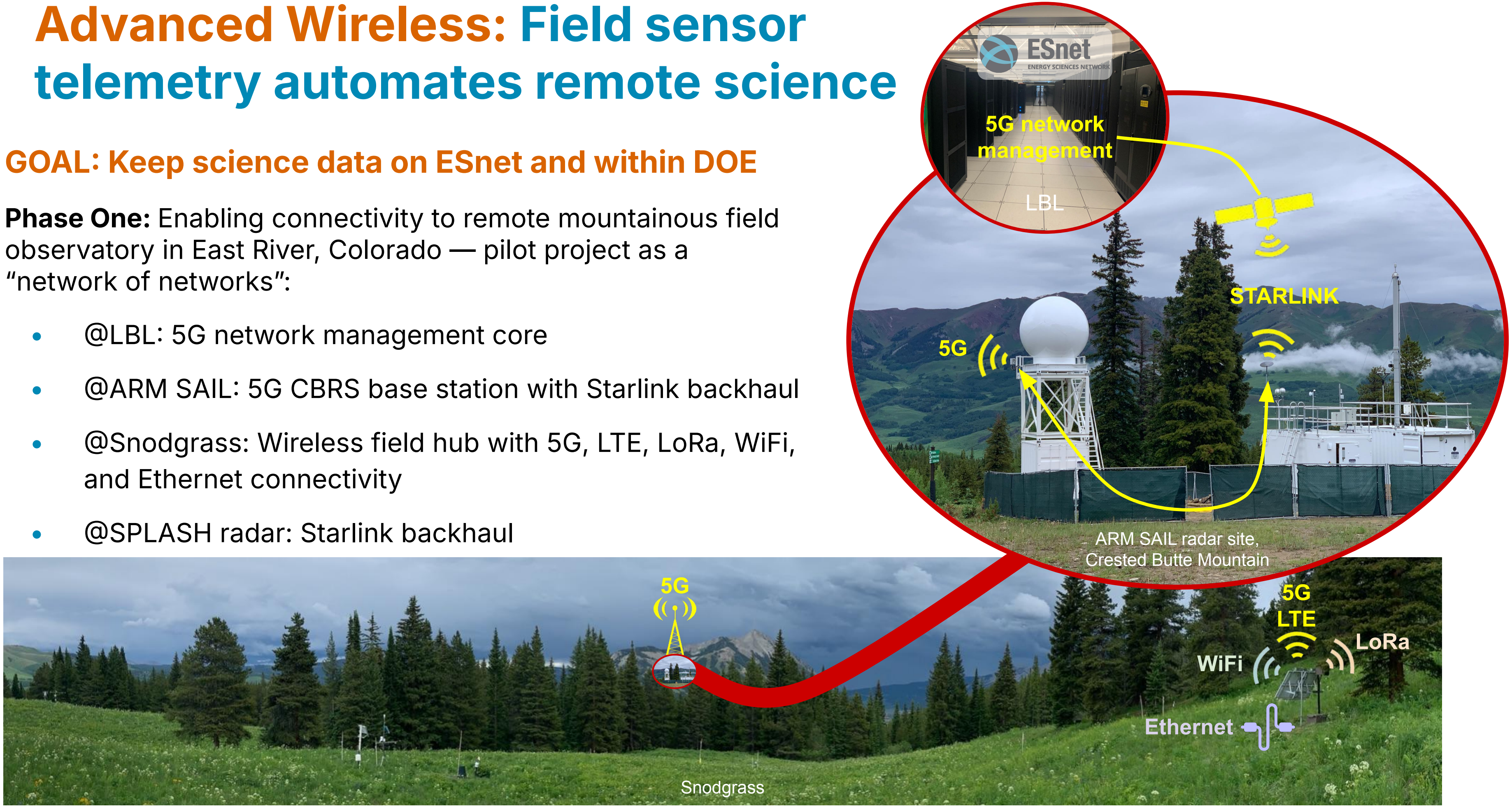
ESnet

Advanced Wireless: Field sensor telemetry automates remote science

GOAL: Keep science data on ESnet and within DOE

Phase One: Enabling connectivity to remote mountainous field observatory in East River, Colorado — pilot project as a "network of networks":

- @LBL: 5G network management core
- @ARM SAIL: 5G CBRS base station with Starlink backhaul
- @Snodgrass: Wireless field hub with 5G, LTE, LoRa, WiFi, and Ethernet connectivity
- @SPLASH radar: Starlink backhaul



2024 ASCAC Facilities Subcommittee Recommendations

ASCAC Facilities charge report summary: Overarching recommendations (direct quotations)

Recommendation 1: Ensure the continued support and development of all five ASCR computational facilities reviewed—ALCF, OLCF, NERSC, HPDF, and ESnet—as they are central and essential to all SC science programs and broader national science and engineering research programs.

Recommendation 2: Science demands integration. We advocate viewing ASCR facilities not as isolated entities, but as integral components of a single, larger integrated computational *ecosystem* (henceforth referred to as *Ecosystem*), with a single governance model. ... *Further, this integrated ecosystem is required for programs of other agencies, and industry. Its critical role in bolstering national scientific and technological capabilities, as well as its status as a model internationally, cannot be overstated.*

ASCR views #3 as related also to the Research and Advanced Computing Technology Divisions

Recommendation 3: A comprehensive, coordinated R&D program delivering multiple prototype computing systems over a five-year timescale must be mounted to inform pathways for this integrated ecosystem, operational by 2034, due to (a) rapidly evolving economic and technical landscapes of the semiconductor and computing industries and (b) changing research practices.

“Recommendation 1 is necessary but not sufficient for success."

ESnet

ASCAC Facilities charge report summary: Overarching recommendations (direct quotations)

ASCR views #3 as related also to the Research and Advanced Computing Technology Divisions

Recommendation 1: Ensure the continued support and development of all five ASCR computational facilities reviewed—ALCF, OLCF, NERSC, HPDF, and ESnet—as they are central and essential to all SC science programs and broader national science and engineering research programs.

Rec ... ving
ASC ... f a
sing ... ed to
as E ... rated
eco ... Its
criti ... lities,
as w ... l.

Individual Facility Assessment

"Without Esnet, the entire vision collapses; none of the facilities, nor the integrated Ecosystem, nor IRI could function."

Rec ...
delivering multiple prototype computing systems over a five-year timescale must be mounted to inform pathways for this integrated ecosystem, operational by 2034, due to (a) rapidly evolving economic and technical landscapes of the semiconductor and computing industries and (b) changing research practices.

"Recommendation 1 is necessary but not sufficient for success."

“Before you marry a person, you should first make them use a computer with slow Internet to see who they really are.”

- Will Ferrell

Chin Guok <chin@es.net>

ESnet