



ESnet

ENERGY SCIENCES NETWORK

Designing the next-generation facility: ESnet6 project

Inder Monga

Director, Energy Sciences Network
Division Director, Scientific Networking
Lawrence Berkeley National Lab

ASCAC

September 2017



U.S. DEPARTMENT OF
ENERGY
Office of Science



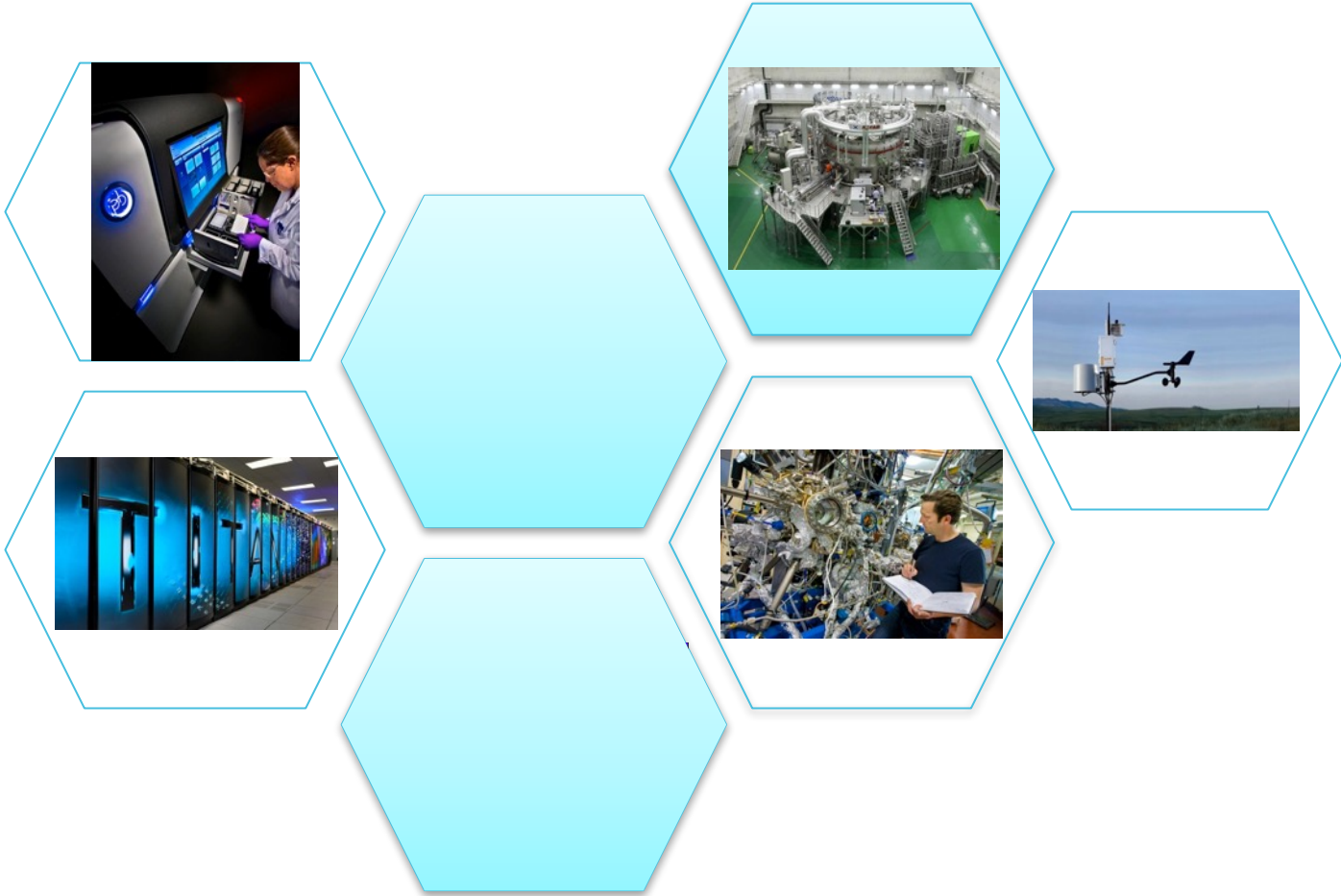
Networks are central to all 'smart' human life



Artificial Intelligence
Machine Learning



Additionally, Networks are central to science collaborations



Roadmap of this talk

ESnet today



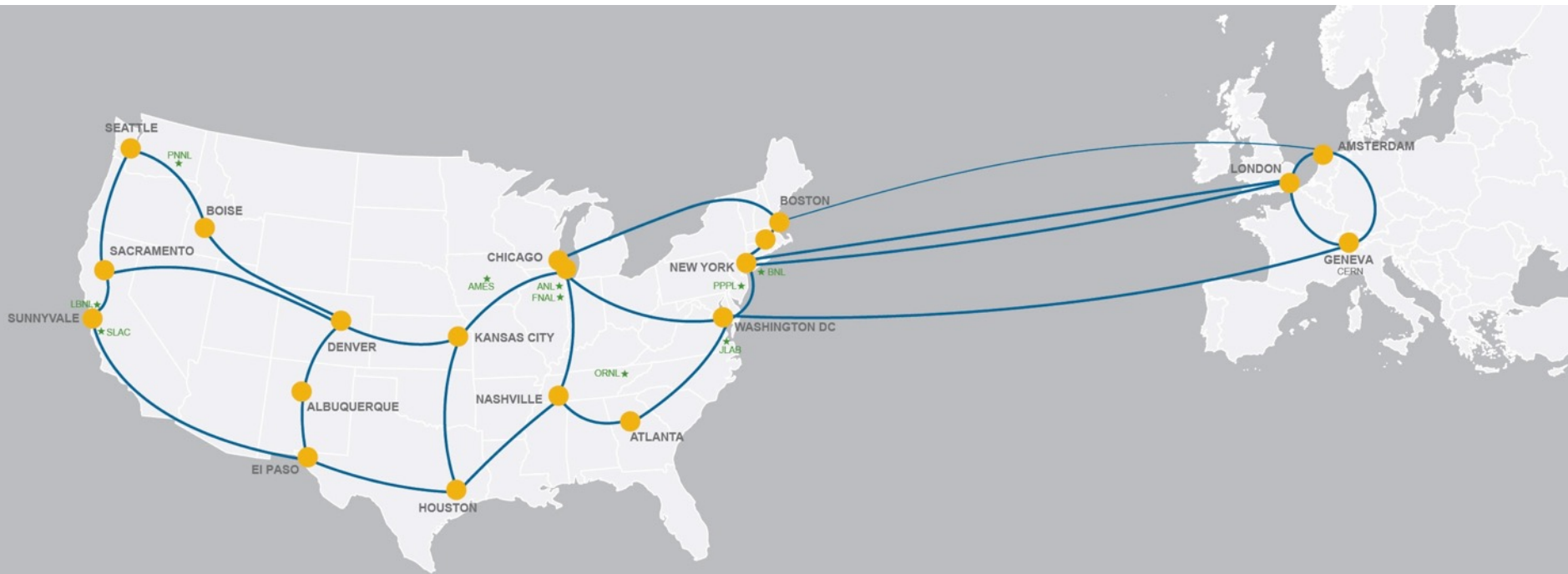
Next-generation requirements



ESnet6 progress



ESnet: DOE's international science network



ESnet is an **Office of Science facility** connecting most of the DOE labs, facilities, experiment sites and supercomputers

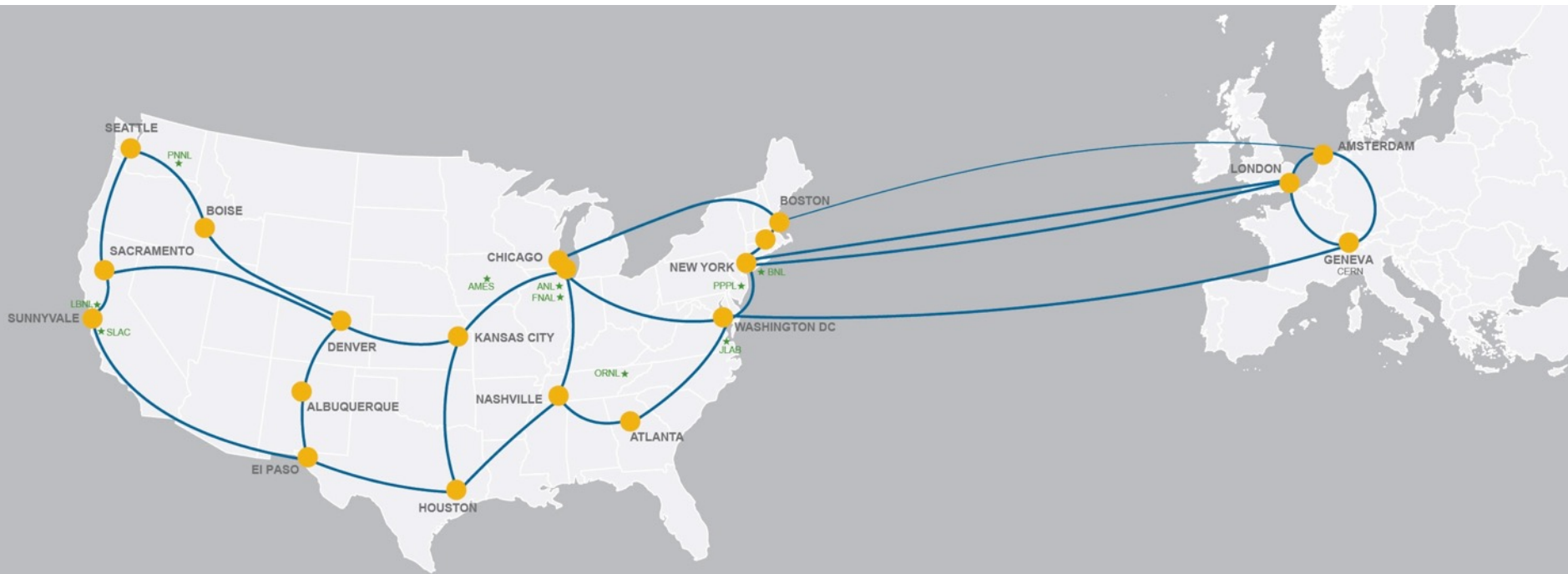
Connects to 100's of other science networks around the world and to the Internet



GLIF Map 2017: Global Lambda Integrated Facility Visualization by Robert Patterson, NCSA, University of Illinois at Urbana-Champaign Data Compilation by Maxine Brown, University of Illinois at Chicago Texture Retouch by Jeff Carpenter, NCSA Earth Texture, visibleearth.nasa.gov www.glif.is



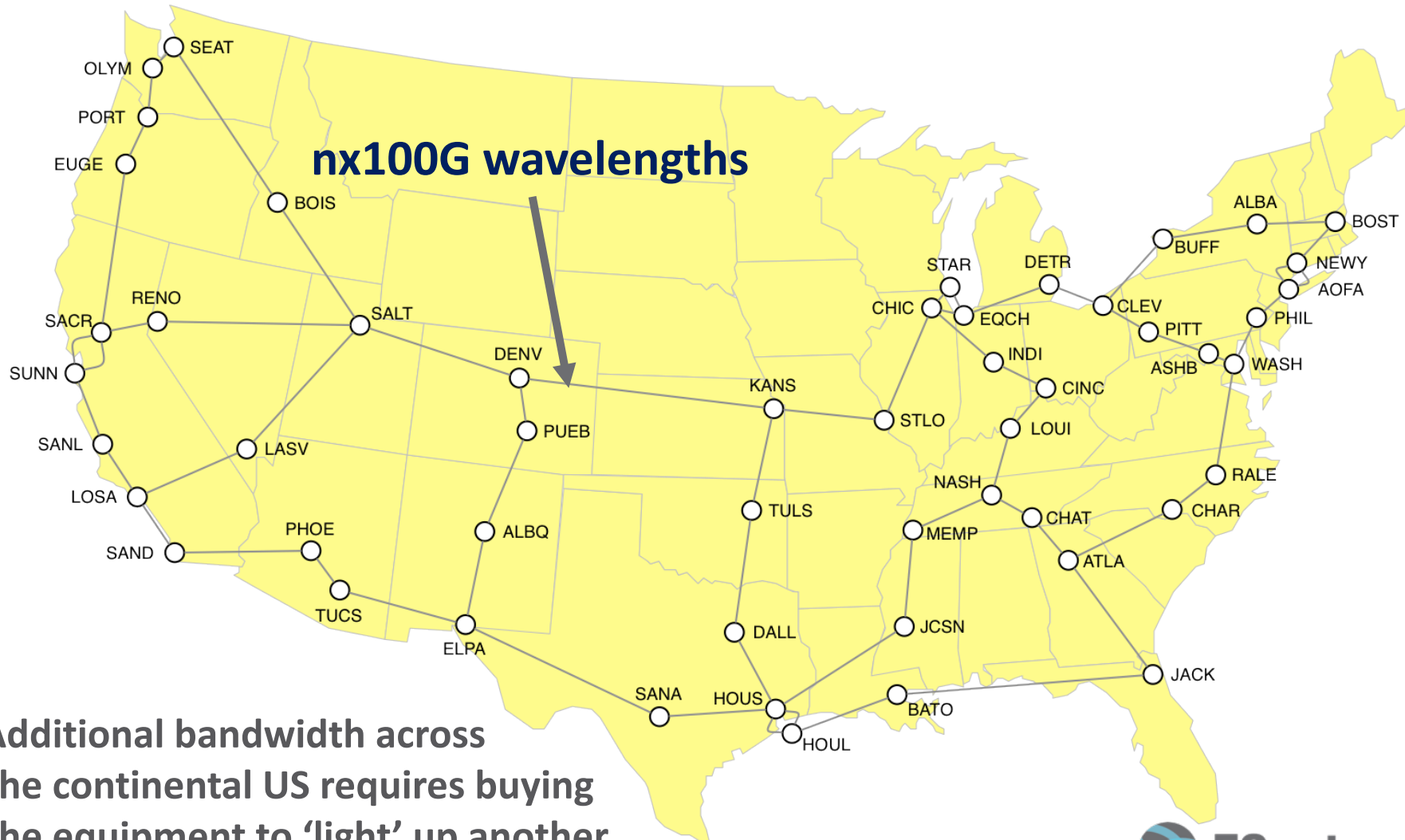
ESnet is engineered and optimized for DOE Big Data Science



We do this by offering super high bandwidth, unique capabilities and protocols for big data data management

Current ESnet optical footprint

shared infrastructure, 50% of fiber capacity



Additional bandwidth across the continental US requires buying the equipment to 'light' up another wavelength across the fiber

ESnet5 network constructed from equipment purchased in 2009-2011



Roadmap of this talk

ESnet today



Next-generation requirements



ESnet6 progress



Next-generation network (ESnet6)

Project Mission Need

- Exponential growth in network **CAPACITY** needs
 - 72% year-on-year traffic growth since 1990
 - Cost effective solution to increase capacity as needed
- Network Life Cycle: Improve **RELIABILITY**
 - Replace aging infrastructure
 - Increase the cyber-resiliency of the network
- Network **FLEXIBILITY**
 - Support increasingly complex workflow models
 - Flexibility at all layers of the network is needed to support wide spectrum of science requirements

1. Capacity

HEP data projects predict 10x-20x growth in data, including new compute models (2021-2027+)

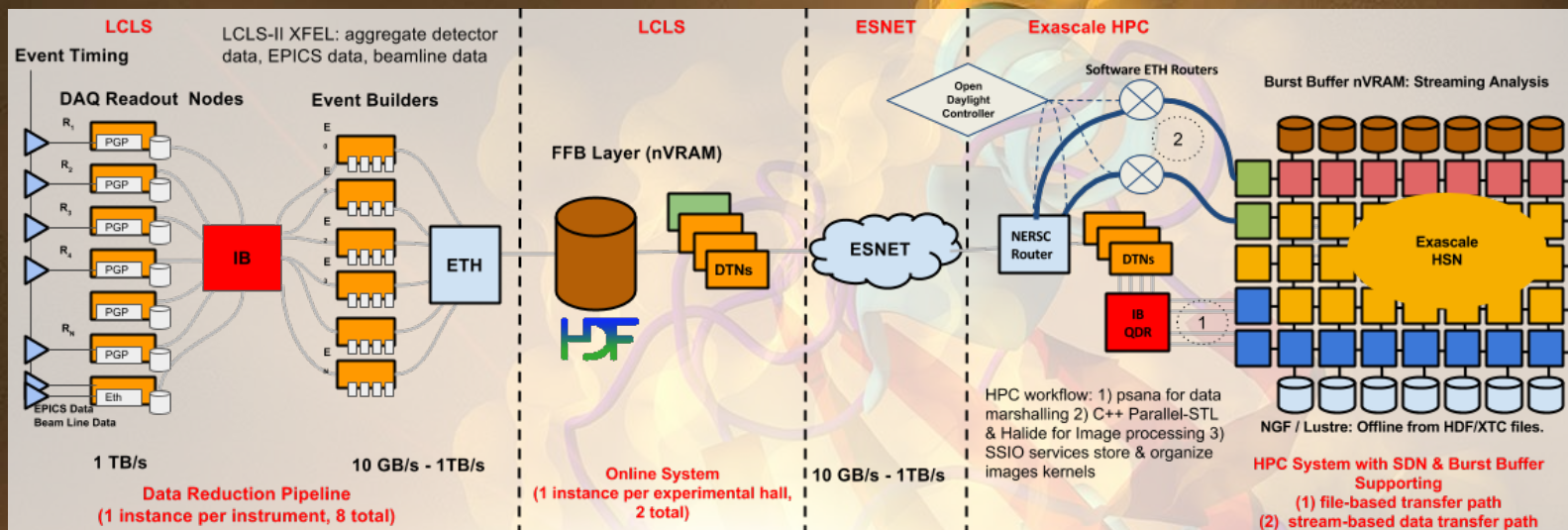


Very rough estimate of new RAW data per year of running using a simple extrapolation of current data volume scaled by the output rates (in PB, follow-up of 2015).

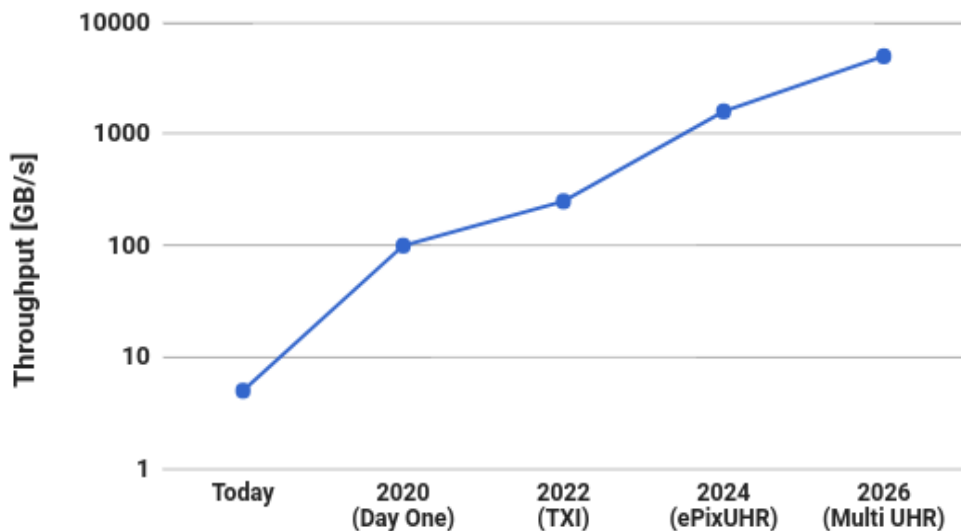
To be added: derived data (ESD, AOD), simulation, user data...



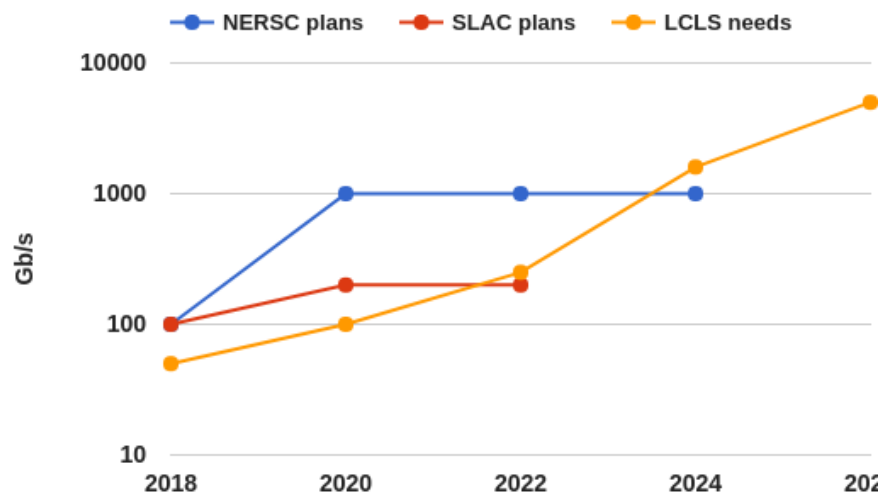
LCLS Science Data (2020 – 2026+)



Peak Throughput (prior to data reduction)



Border Network



This assumes 10x data reduction is achieved

Potential use of ‘Digital Twins’– applications for big data on exascale machines



Industrial
manufacturing

Computational
biomedicine

Experimental
science

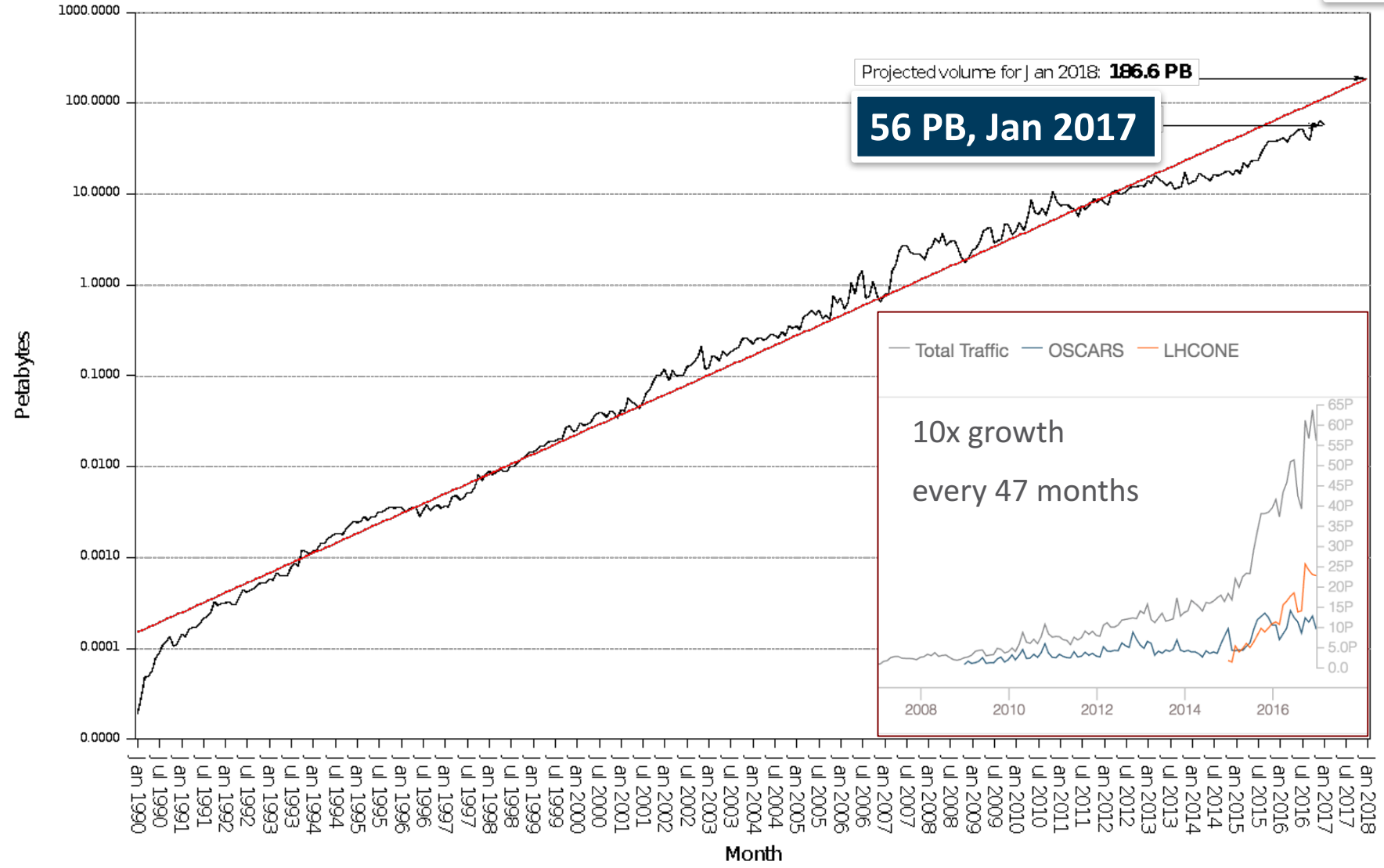
“Ability to run digital twin experiments routinely, including with comparative analytical capabilities during the experiment. This capability will include multisource streaming analysis in distributed computing environments”: *Section 3.6.2.2.3, BES Exascale Requirements Report, <http://exascale.org/bes/>*

Exponential growth, a 27 year trend

1 EB
Jan
2021*

ESnet Accepted Traffic: Jan 1990 - Jan 2017 (Log Scale)

— Actual
— Exponential regression with 12 month projection



Exponential growth is deceptive, and then explosive

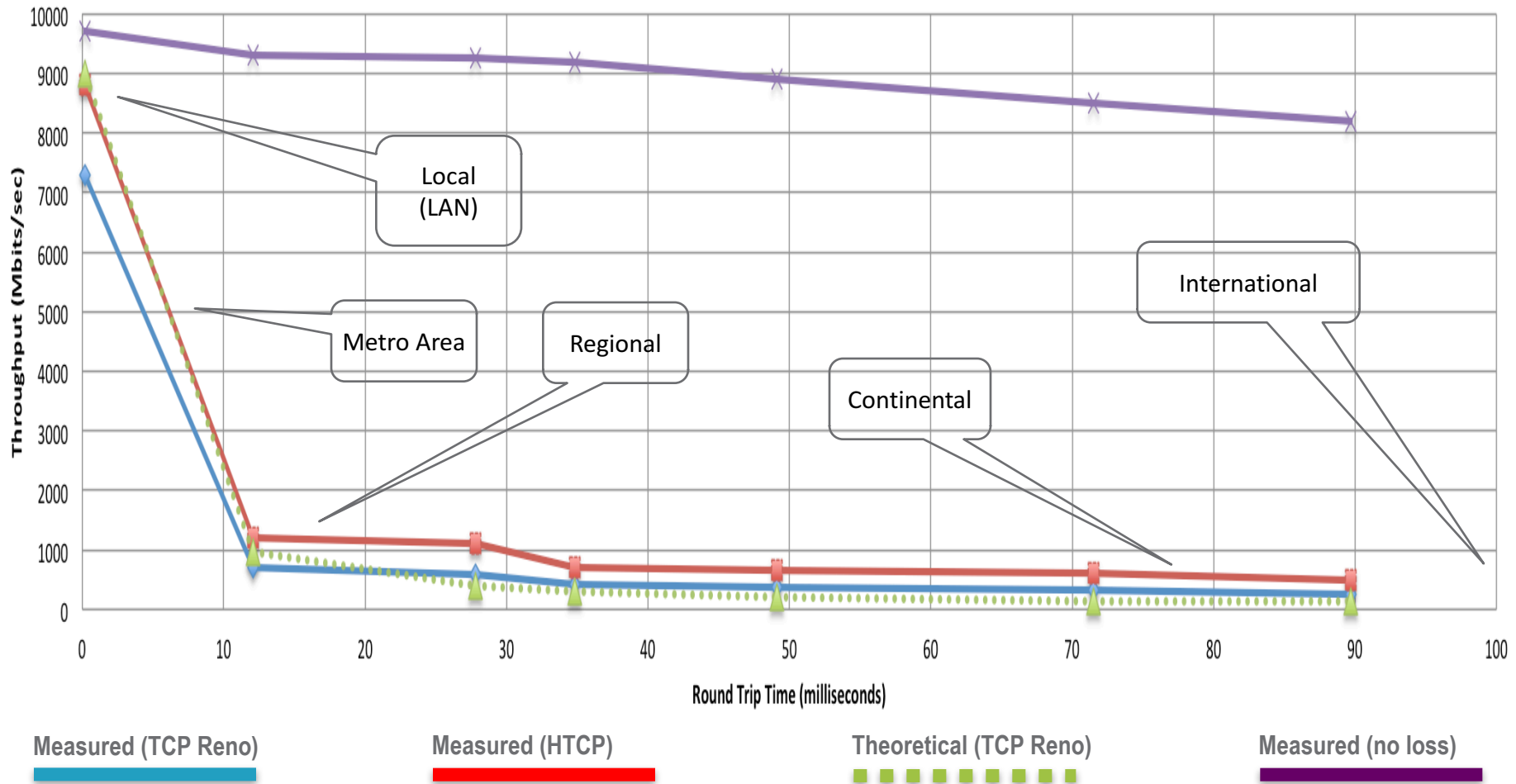


2. Resiliency

Big data flows require almost *lossless* networks

Throughput is equally important as bandwidth

Throughput vs. Increasing Latency with .0046% Packet Loss

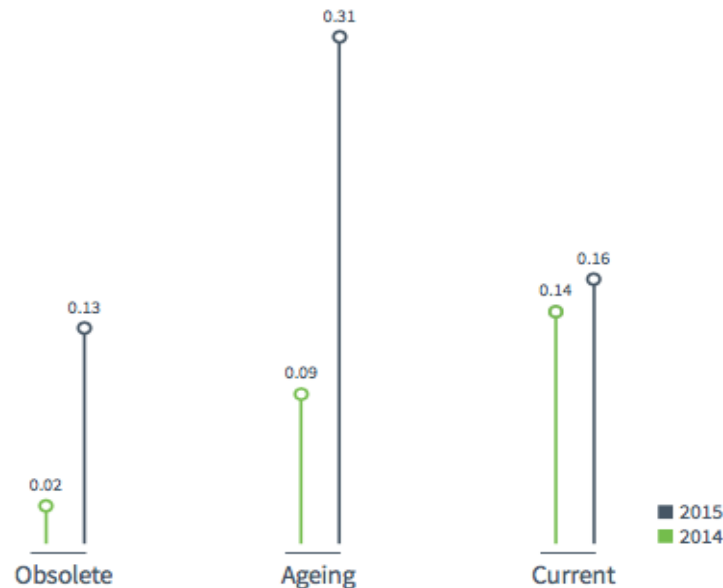


See Eli Dart, Lauren Rotman, Brian Tierney, Mary Hester, and Jason Zurawski. The Science DMZ: A Network Design Pattern for Data-Intensive Science. In *Proceedings of the IEEE/ACM Annual SuperComputing Conference (SC13)*, Denver CO, 2013.

Network equipment approaching 10 years in service by 2021, platforms 2+ generations old

Ageing devices are more likely to fail

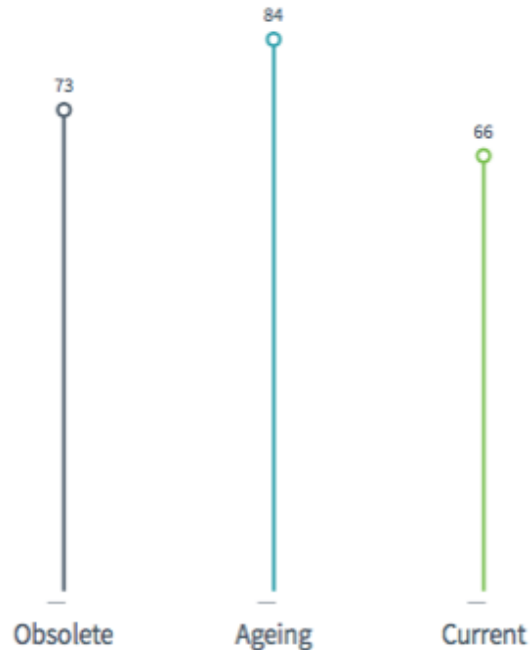
Figure 7: Percentage of failure rate by lifecycle stage



This year we've seen an **increase in the number of incidents across the board**, and a pronounced **increase in the number of incidents on ageing devices**.

Cyber-security vulnerability needs to be managed closely with ageing devices

Figure 16: Devices with at least one vulnerability, by lifecycle stage



Current devices have the lowest level of vulnerabilities, at 66%. What's surprising is that ageing devices have the highest, at 84%.

How can ESnet assist increase cyber-resiliency?

ESnet Security-Assist

Automatic

Automating security processes to reduce resource drain and threat response times

Subscription/Opt-in

Assistance where a site chooses to permit on an ongoing basis until explicitly terminated

One-off Requests

Assistance provided upon site request that happens only once and is not repeated

Self-Serve

Assistance activated directly through the my.es.net portal, and doesn't require any interaction with ESnet staff

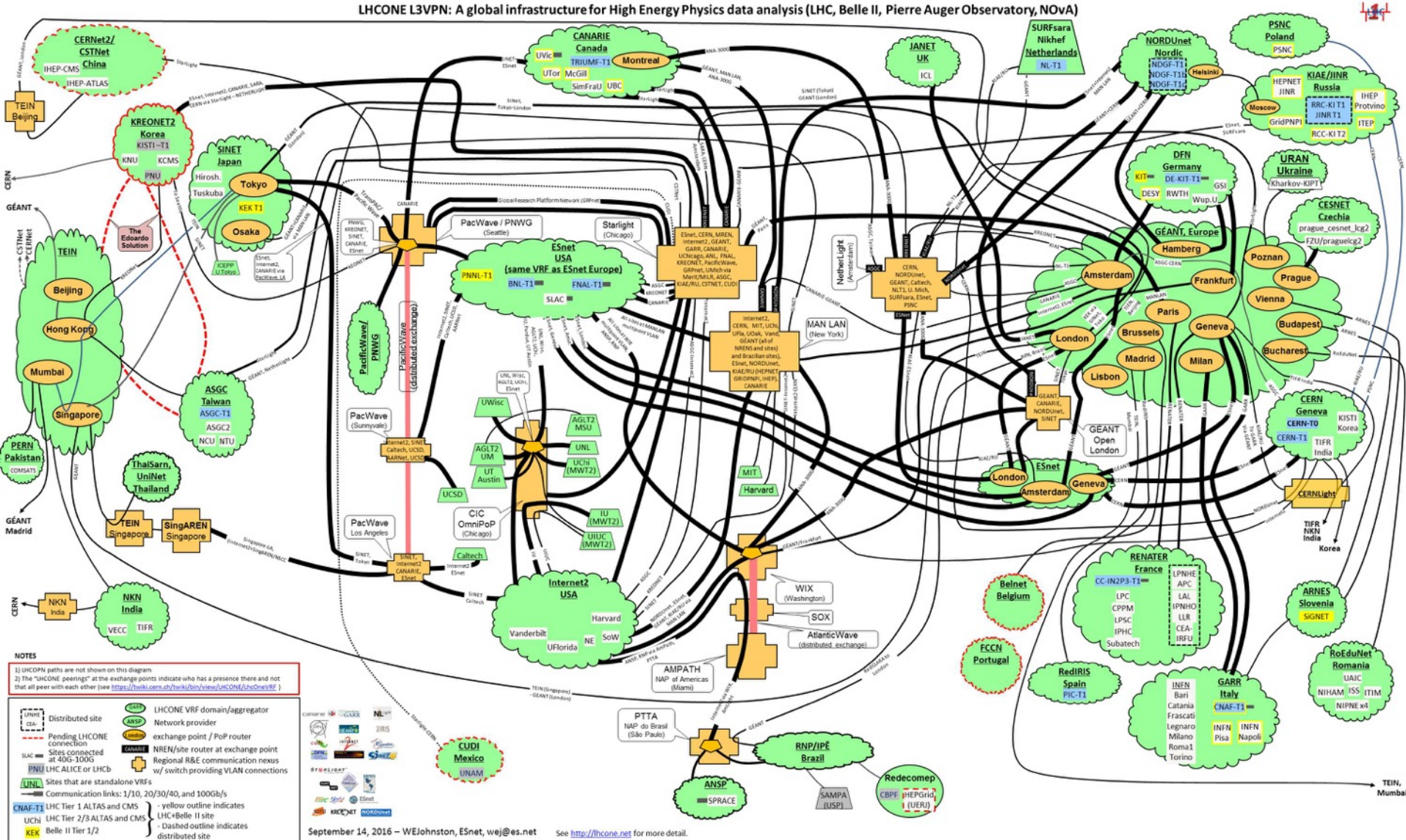
We plan to provide a full suite of self-serve assistance



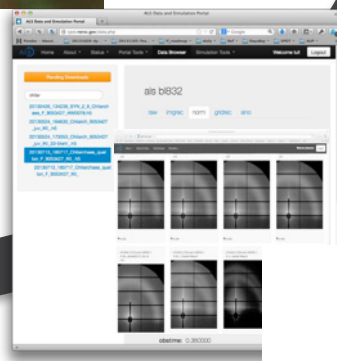
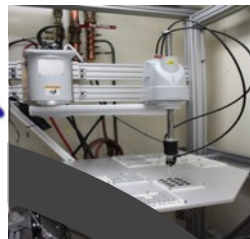
3. Flexibility

Flexible, virtual, Global Science Networks

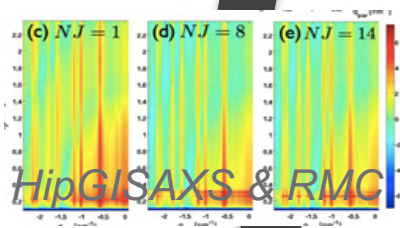
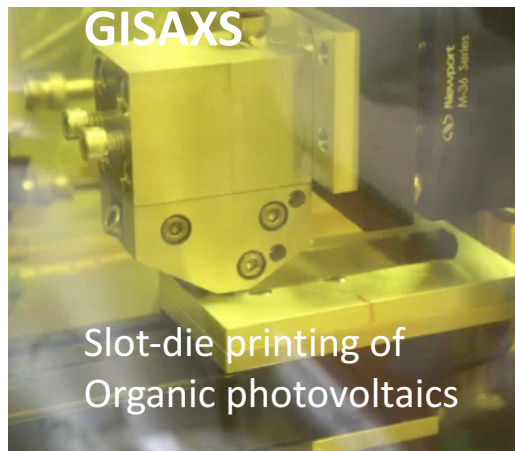
LHCONE L3VPN: A global infrastructure for High Energy Physics data analysis (LHC, Belle II, Pierre Auger Observatory, NOvA)



Computing, experiments, networking and expertise in a “Superfacility” for Science



Liu et al, “Fast printing and in situ morphology ...”. Adv Mater. 2015



Orchestration of end-to-end data workflows

data source to data sink

ASCR Research funded project looks at orchestrating data between Science DMZ resources across various labs, universities and HPC centers

SENSE SDN Control Plane Architecture for End-to-End Orchestration

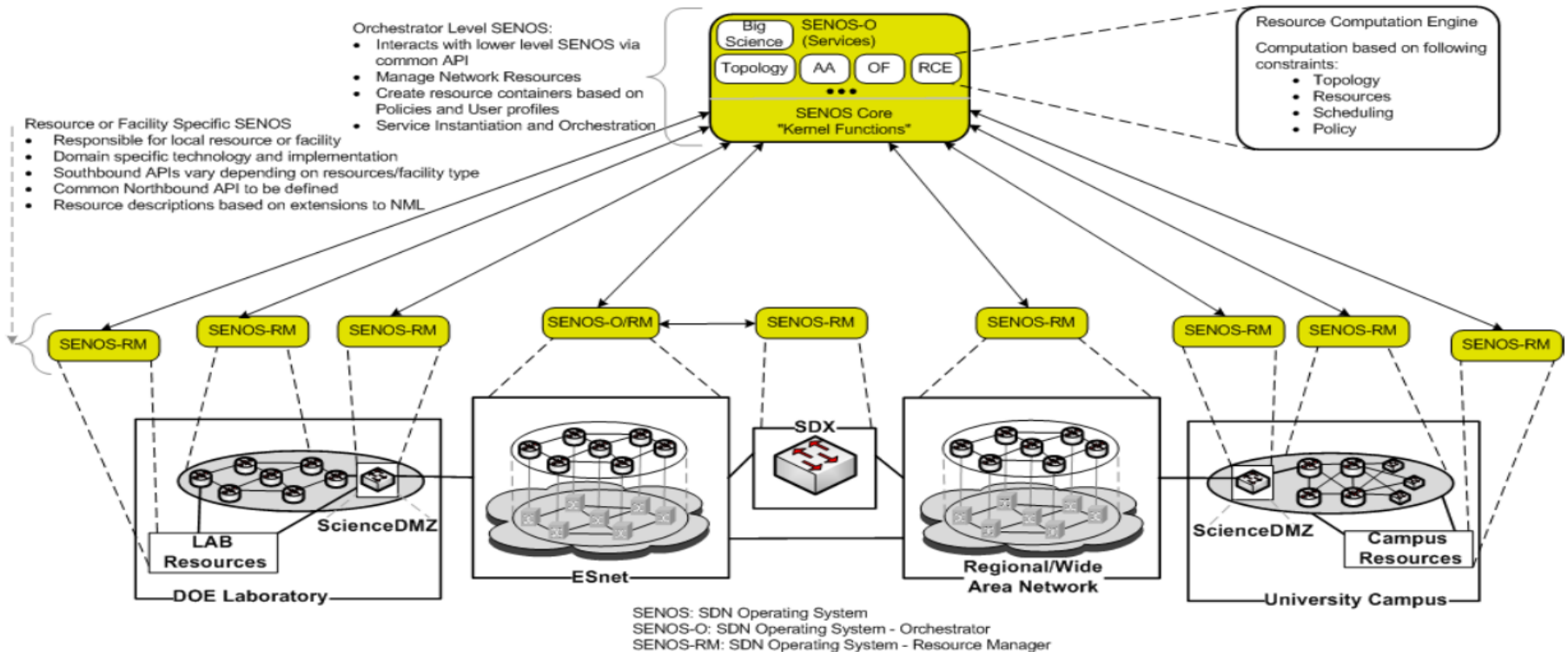
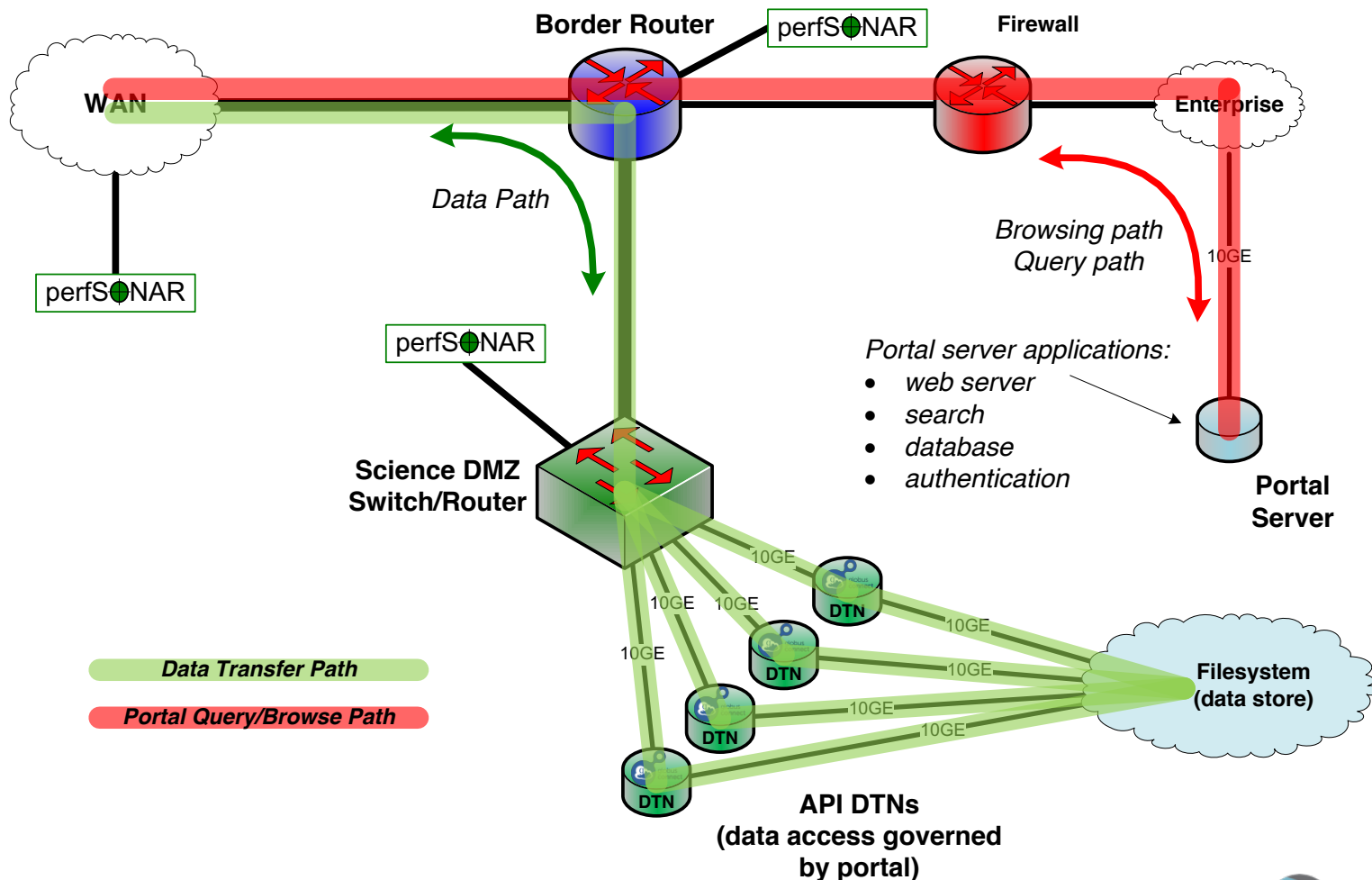


Figure 1. SENOS End-to-End Orchestration



Modern data portals disaggregate data movement from data websites/portals



Roadmap of this talk

ESnet today



Next-generation requirements

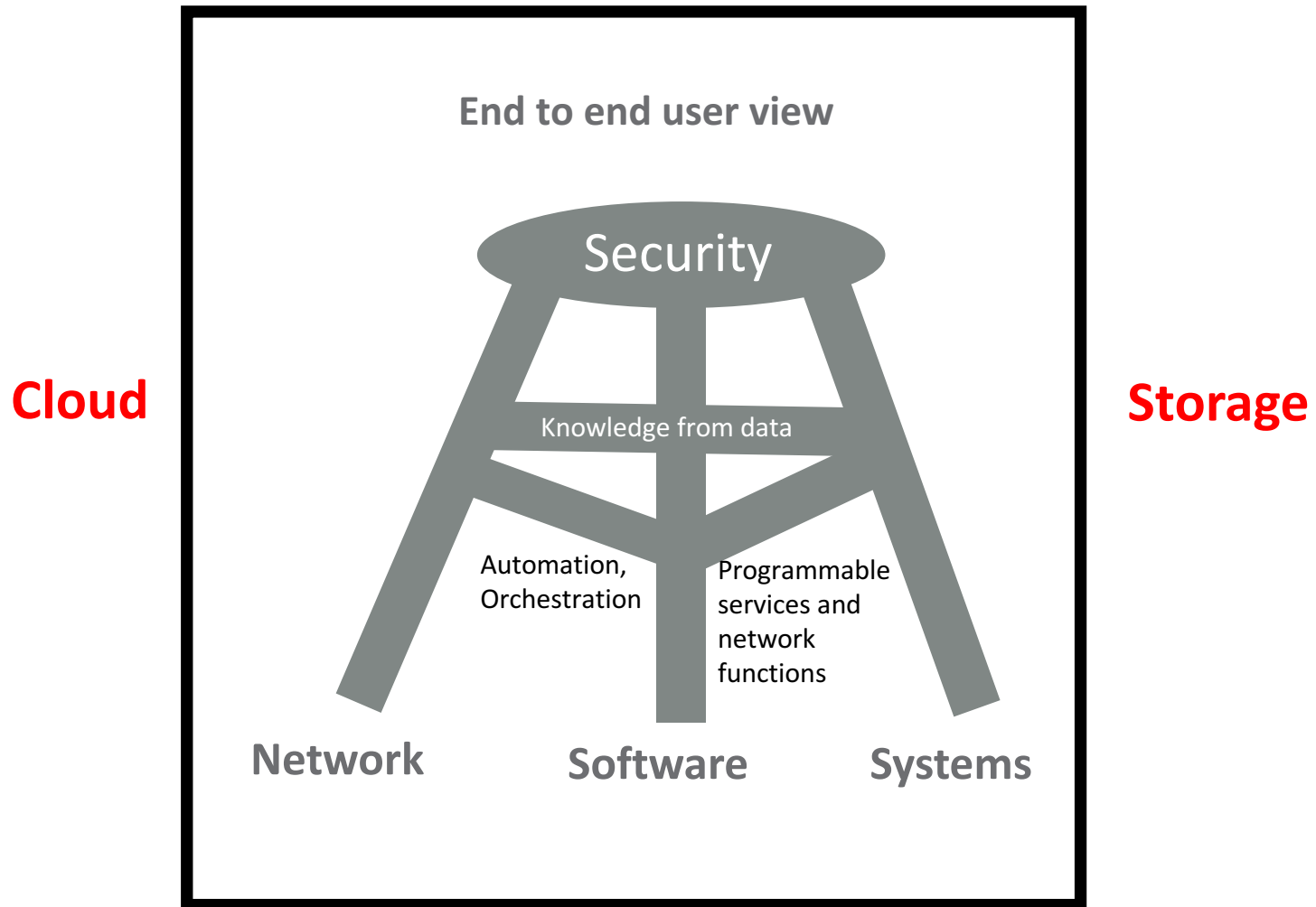


ESnet6 progress



ESnet6: 'Integrative thinking' for future architecture

Compute



Science Instruments

Technical requirements: input to ESnet6 Design

1. Capacity Plan

- Predict usage
- Determine overheads (e.g. burst multiplier, resiliency requirements, short-term growth trends)

2. Services

- Document workflows
- Develop service portfolio*

**NB: Service Portfolio in conjunction with architecture design drives the technical requirements*

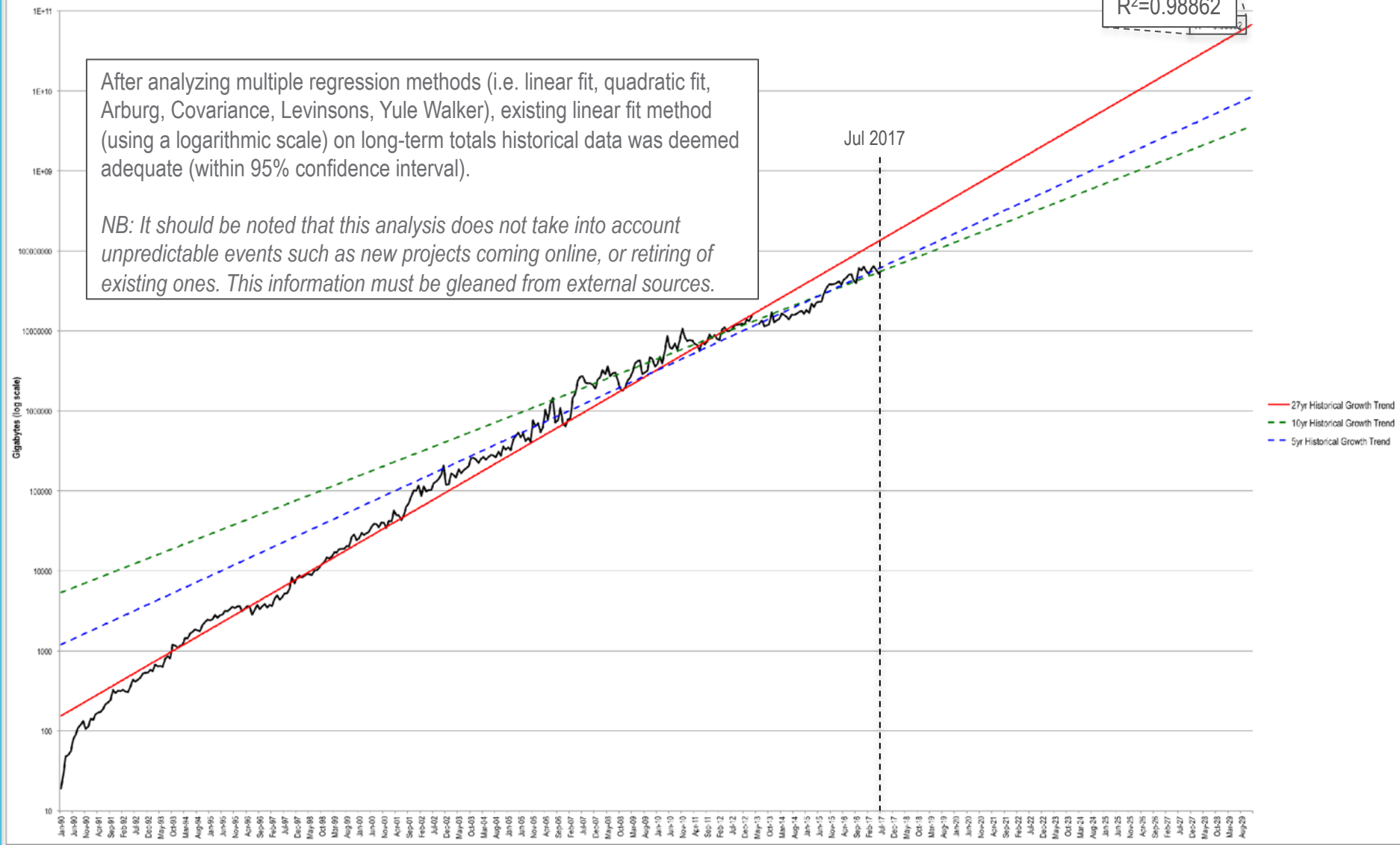
27-Year Totals Historical Trend Analysis

ESnet Accepted Traffic: 5yr, 10yr, 27yr Historical Growth Trends

$R^2=0.98862$

After analyzing multiple regression methods (i.e. linear fit, quadratic fit, Arburg, Covariance, Levinsons, Yule Walker), existing linear fit method (using a logarithmic scale) on long-term totals historical data was deemed adequate (within 95% confidence interval).

NB: It should be noted that this analysis does not take into account unpredictable events such as new projects coming online, or retiring of existing ones. This information must be gleaned from external sources.



Capacity prediction methodology

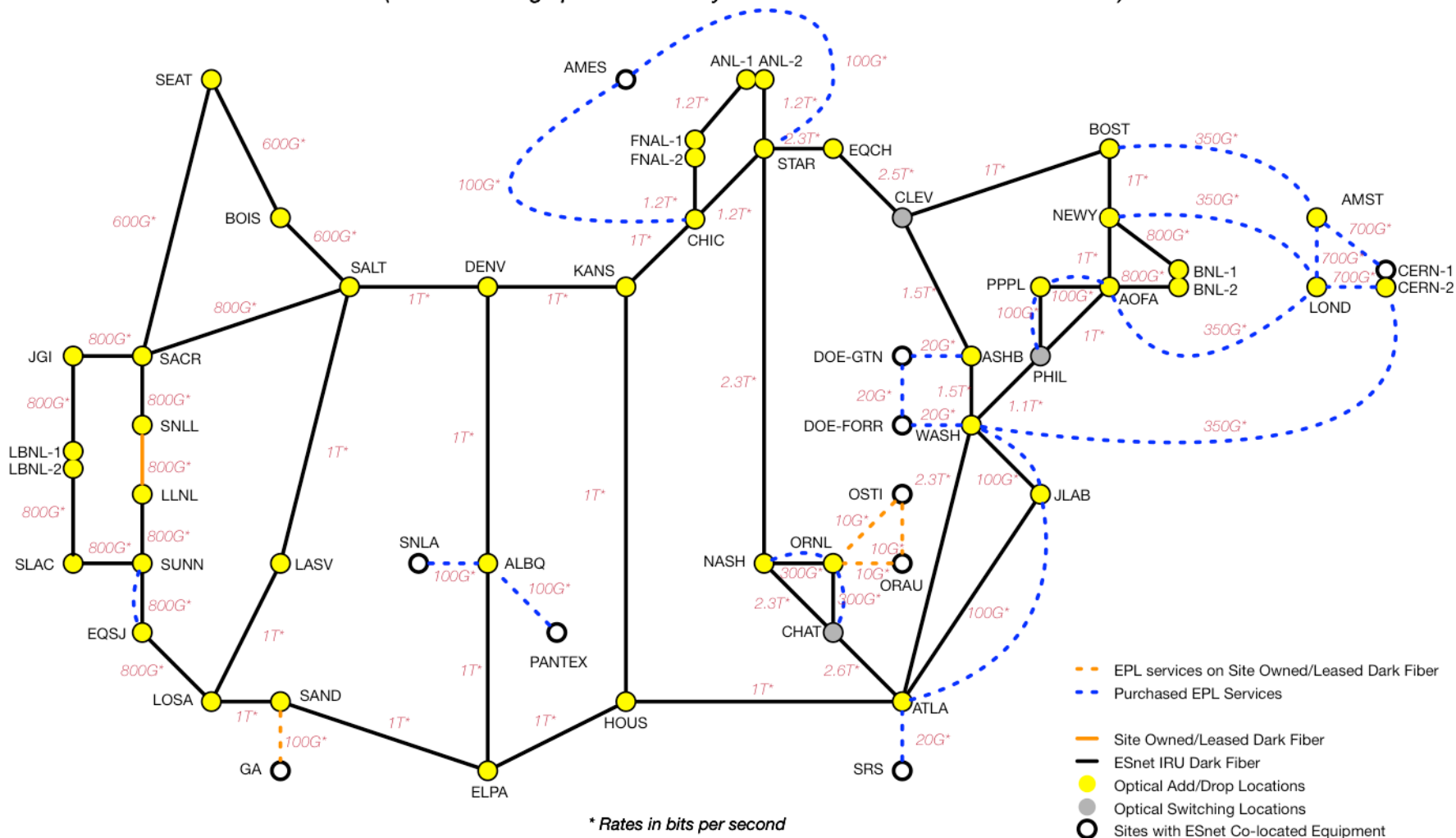
1. Determine predicted baseline usage (for 2020, 2025, and 2030)
 1. Perform best-fit growth curve of ingress traffic per router
 2. Adjust individual router predictions such that total of all router ingress traffic matches ESnet's 25+ year total traffic growth curve
 3. Using historical flow data and predicted ingress traffic data, perform full mesh path computation to determine per link utilization from edge to edge (PE to PE)
2. Strategic capacity planning* (for 2020 and 2025)
 1. Add burst overhead bandwidth per link based on historical knowledge
 2. Add additional bandwidth to paths based on resiliency strategy
 3. *Keep in view new projects on the horizon*

**NB: This is an iterative process as we continue to monitor growth trends as well as field requests for new requirements (e.g. new experiments, etc)*

ESnet6 – a terabit network (Jan 2021)

ESnet6 (Proposed) Footprint

Jan 2021 Bandwidth Capacity Plan
 (Based on usage prediction analysis which includes Jul 2021 traffic data)

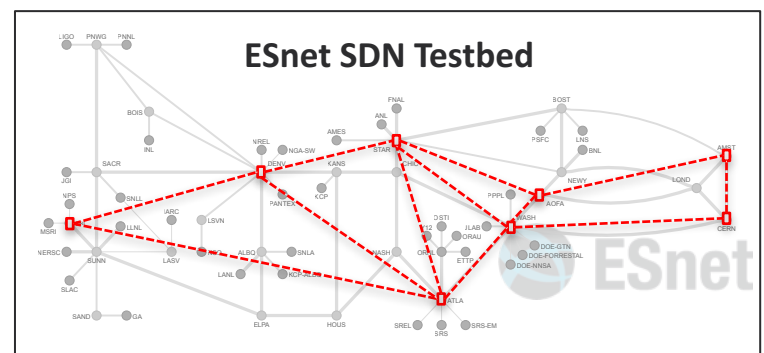


NB: Some sites without PE routers on local premises are not shown in this diagram

ESnet6 Services Definition Process



- Determine workflows based on requirements workshops
 - 6 DOE Office of Science program offices:
 - Advance Scientific Computing Research (ASCR) - Basic Energy Sciences (BES)
 - Biological and Environmental Research (BER) - Fusion Energy Sciences (FES)
 - High Energy Physics (HEP)
 - Nuclear Physics (NP)
 - Two requirements workshops a year
 - 3 years to rotate through all 6 programs
- ESnet testbed research activities



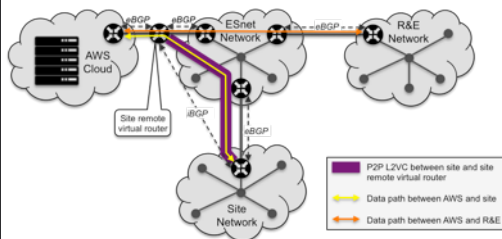
ESnet6 Workflows, Services, and Technical Requirements

Input on requirements are documented as workflows, which are then formalized as services, driving the technical requirements.

Workflows

Example workflow:

Use of Cloud compute resources as an extension of the site's resources.



Services

Example Service:

Virtual Private Cloud

- Service Description
- Service Attributes
 - Scale
 - Scope
 - Demarcation

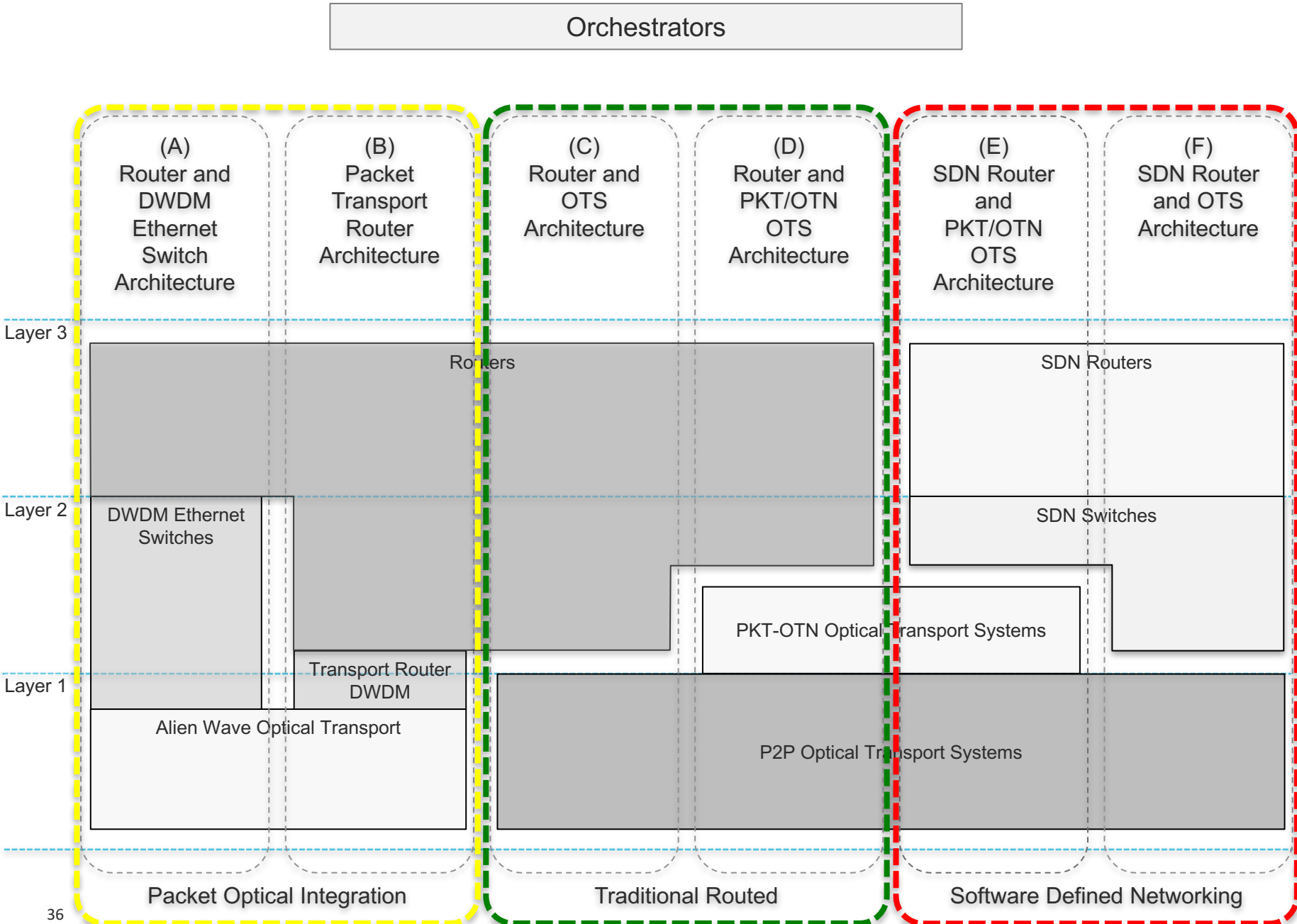


Technical Requirements

Example Requirements:

- L2VPN
 - EVPN
 - MPLS
 - ISIS-SR
 - BGP-SR-TE
- e/iBGP
- ...

R&D Phase: Architecture and Technologies Matrix



ESnet6 R&D technology findings in a nutshell

- Architecture A&B (Packet Optical Integration)
 - Potentially highly scalable, cost compelling, power & space efficient
 - All-in-one box requires very careful resiliency planning
 - Separate transponder and amplifier chain (vendor) domains require some custom integration for optical end-to-end management
- Architecture C&D (Traditional Routed)
 - Clean layer separation provides simpler resiliency planning
 - Cost challenged, potentially unsustainable power and space requirements moving towards 2025
- Architecture E&F (Software Defined Networking)
 - Highly flexible for resource slicing and service creations (i.e. NFV)
 - Potentially highly scalable, cost compelling
 - Non-trivial design complexity (e.g. distinct requirements and designs for management, control, and data planes)
 - Potential production support complications (e.g. troubleshooting, multi-vendor components, etc)
 - Lack of maturity for WAN scale solutions

Conclusion: ESnet6 will be a hybrid, containing certain aspects of each architecture and balancing the benefits and trade-offs.

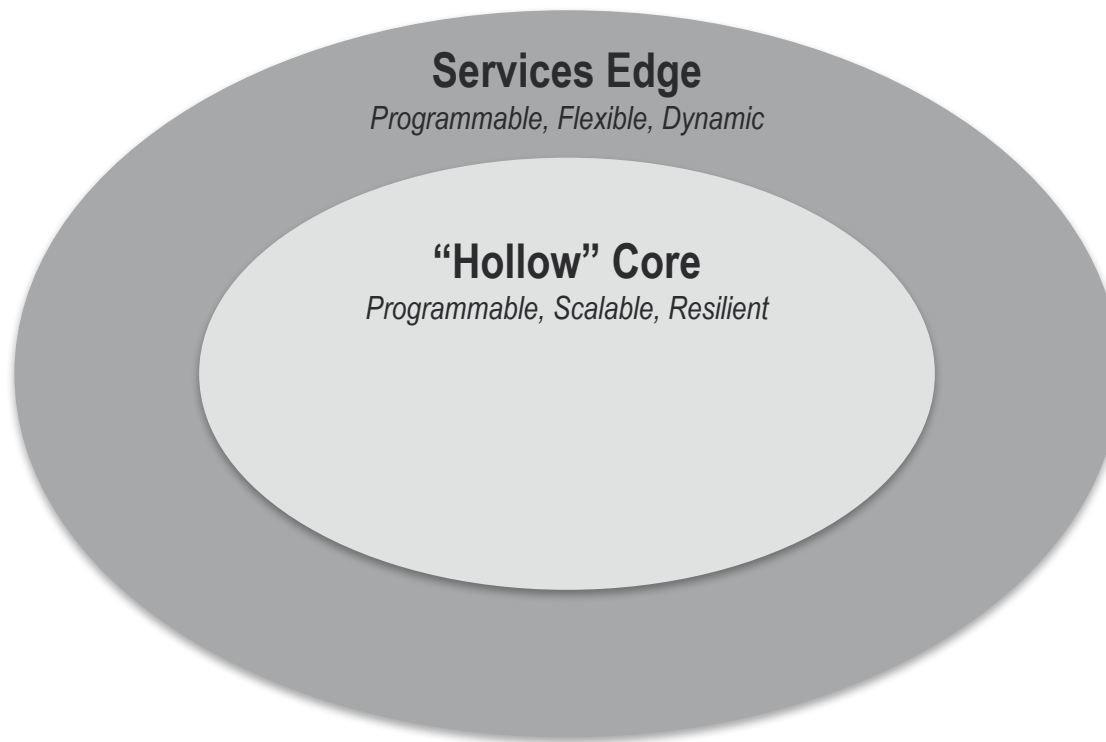
(Comments from the community will be taken as part of an external independent review process.)



Preferred ESnet6 architecture



ESnet6 (“Hollow-Core”) Conceptual Architecture Overview



“Hollow” Core

- Low cost, high bandwidth capacity with dynamic provisioning
- No per-hop L3 routing within the Core, packets will be (label) switched
- Protection and restoration for Edge-to-Edge resiliency

Services Edge

- Highly programmable data and control plane for flexibility
- Rich capability to provide multiple services
- Intelligence and Orchestration of network functions through central software control



Software is an integral part of ESnet6 design



Managing the network: guiding principles

- Zero touch self regulating network
 - Avoid manual intervention in the daily operations of the network.
 - Prioritize automation of frequent operations over infrequent ones.
- Must always have access to an accurate fault stream
 - A basic fault monitoring system must always be available.
 - Without it the operational state of the network is unknown.
- Must always be able to take corrective action
 - This may be using the EMS or it may be CLI access.
 - CLI access should be the method of last resort and only used in the case of an emergency.
- Leveraging cloud technologies to allow for scaling as appropriate
- Clearly defined source of truth for all configuration data

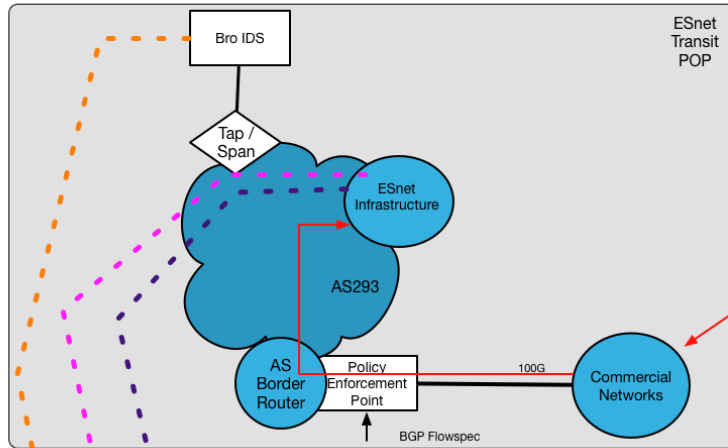
Zero touch network management

- Manage network via an declarative interface
 - Avoid manually configuring devices
- Network can detect security anomalies
 - Traffic pattern analysis
 - Bro data
- Network can detect large flows for redirection
 - Traffic pattern analysis
 - Flow data
- Network can predict failures and over-subscription scenarios
 - Analytic models analyze measurement and monitoring data

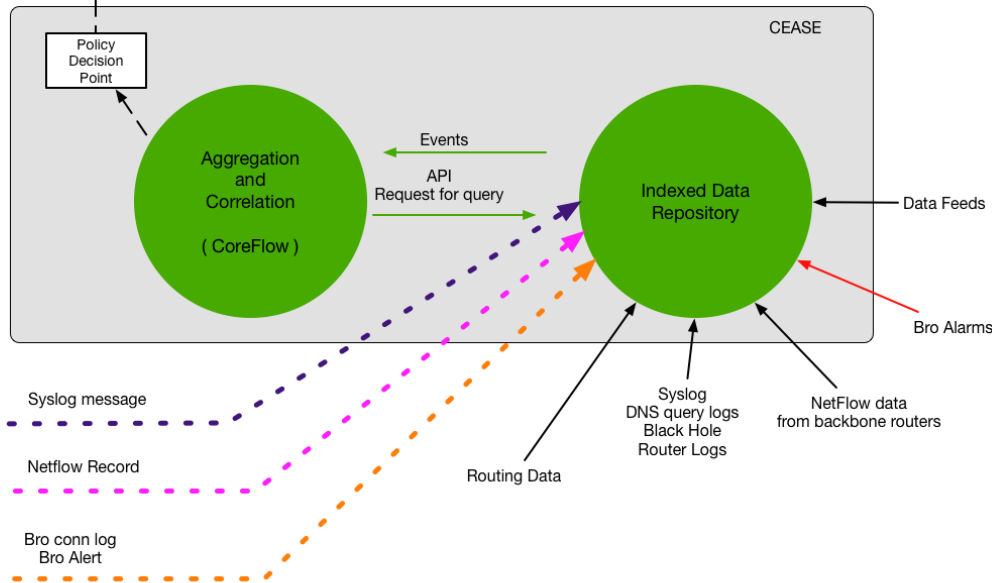
Software Defined Networking

- Direct Programmability
 - Standards-based programmatic access to all network device capabilities
 - Fast and responsive API with asynchronous notification capabilities
 - Programmable dataplane for control over individual data flows
 - Fine grained streaming telemetry
- Agility and Flexibility
 - Underlying infrastructure abstracted for applications and network services provides greater development velocity

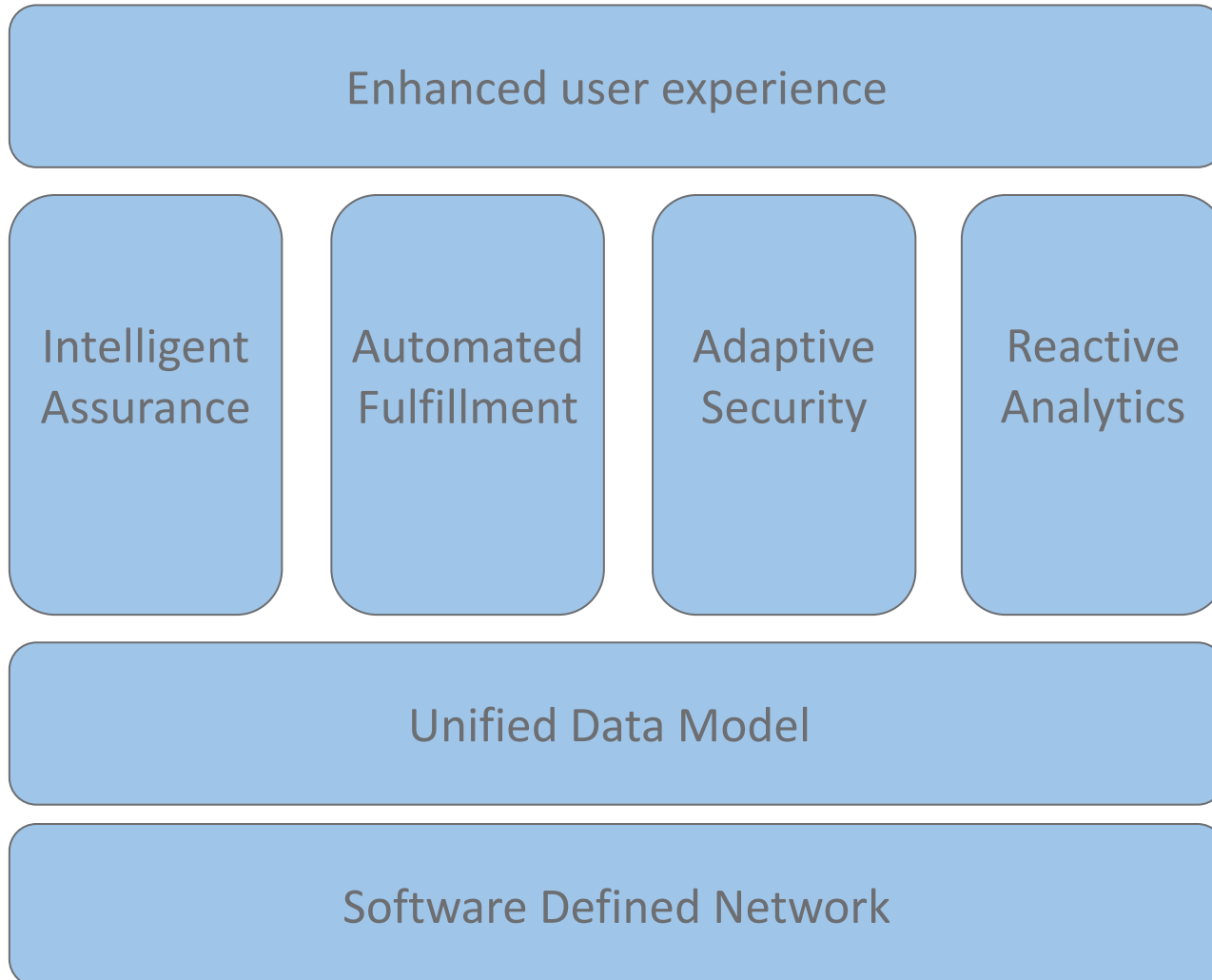
Future security capabilities are based on software and network telemetry data



Correlation Evaluation And Security Enforcement




ESnet6 Software Pillars



ESnet6: a 413.3b project



ESnet6: 413.3b Critical Decision project*

CD Phases	Fiscal Year
CD-0, Approve Mission Need	FY2017 
CD-1/3a, Approve Alternative Selection and Cost Range, Approve Long Lead Procurement	FY2018
CD-2/3b, Approve Performance Baseline, Approve start of construction	FY2018/2019
CD-4, Approve Project Completion	FY2022

* Depends on Congressional funding for this project which is uncertain for FY18 at this time

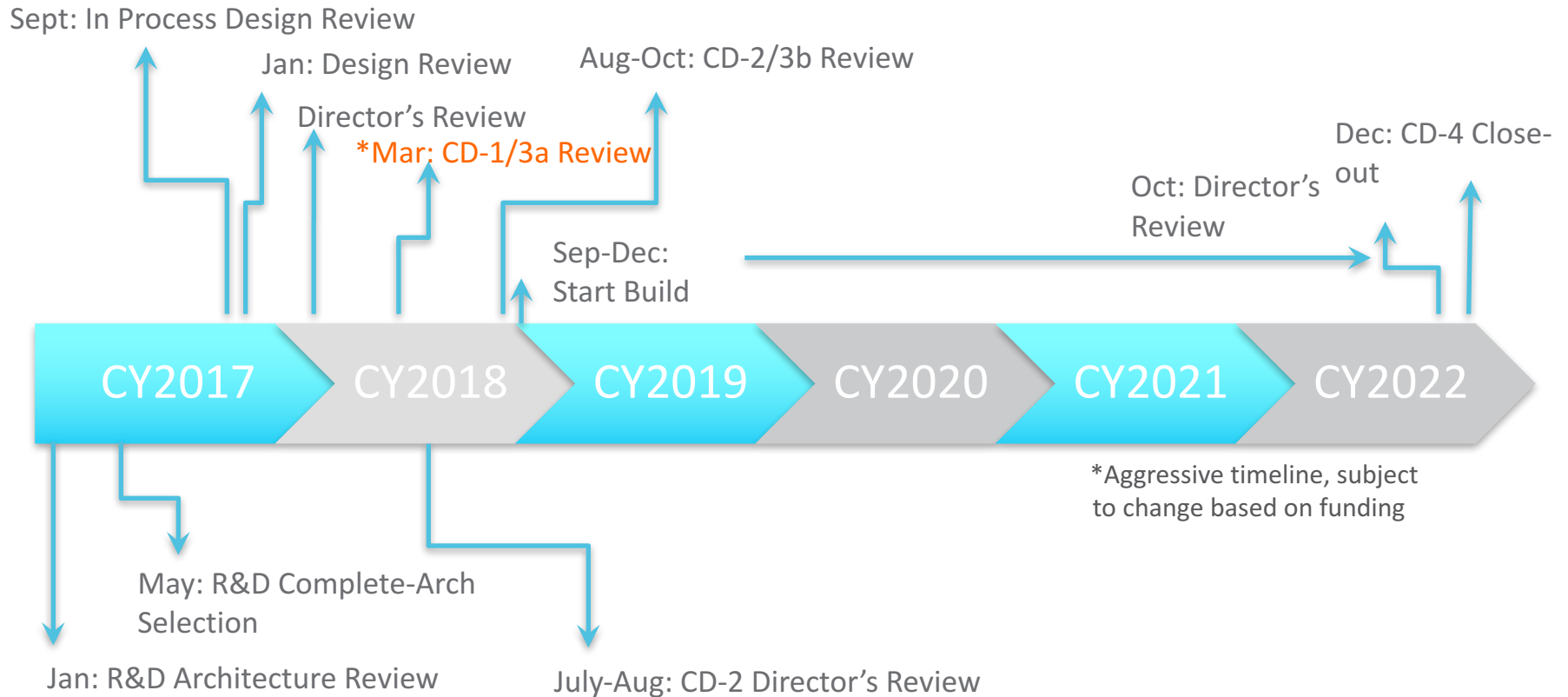
ESnet6 Scope

- ESnet6, "The Project", has a distinct start and end date
- ESnet aka. "The Facility", has to continue without service interruption
 - ESnet5 runs in parallel to ESnet6 until all services are transitioned

ESnet6 Project Scope (Threshold)	<ul style="list-style-type: none"> • Capability to meet 2025 traffic projections <ul style="list-style-type: none"> • Dark Fiber route, MANs and EU Rings • Capacity delivered to 2021 traffic projections • Security infrastructure supporting threat detection, analysis and mitigation at line rate at the commercial peering points (EQIX, HE) • Limited dark fiber extensions • All currently offered ESnet5 services
ESnet6 Project Scope (Objective)	As we free up contingency funding, we expect to apply it: <ul style="list-style-type: none"> • Enhanced dark fiber extensions • Enhanced security infrastructure • Targeted new services
Not in Scope	<ul style="list-style-type: none"> • Transatlantic Network capacity via purchased 100G circuits



ESnet6 Project Timeline



***We are reviewing various project tailoring strategies that may effect some of the CD milestones.**



ESnet6 Risk Management



- ESnet6 developed a plan and process to identify and manage project risks
- The plan has been reviewed by ESnet, ASCR and the DOE FPD
- Monthly risk review meetings are conducted to ensure timely risk management

- 28 Project risks have been identified and are under management
- Project risks also lead to Program risks, as they are interdependent

New explorations in Machine Learning being applied to network telemetry data

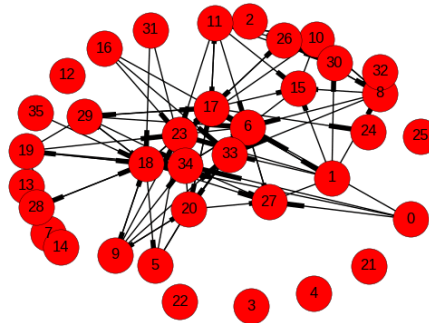


Mariam Kiran
DOE early career grant, 2017

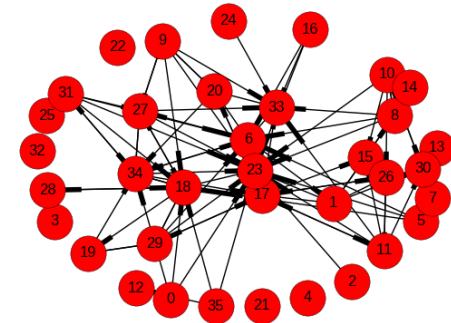


Anshuman Chabbra
Undergrad intern

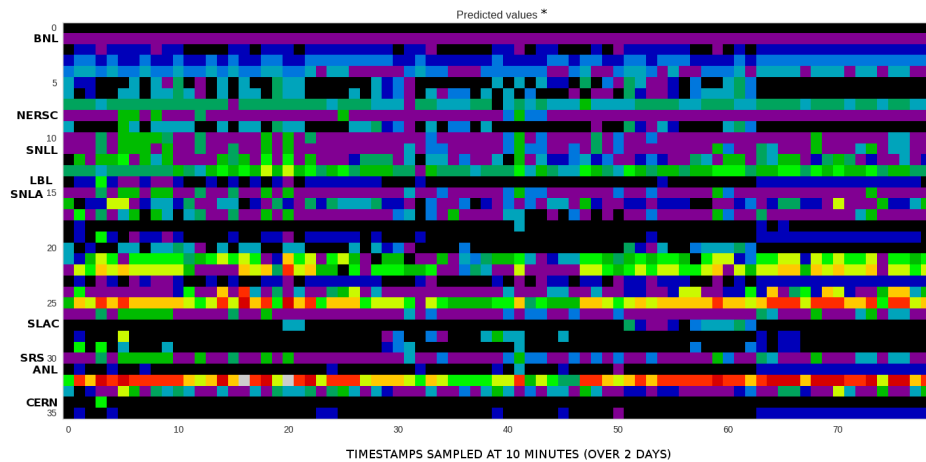
Original Link Topology - After 10 minutes



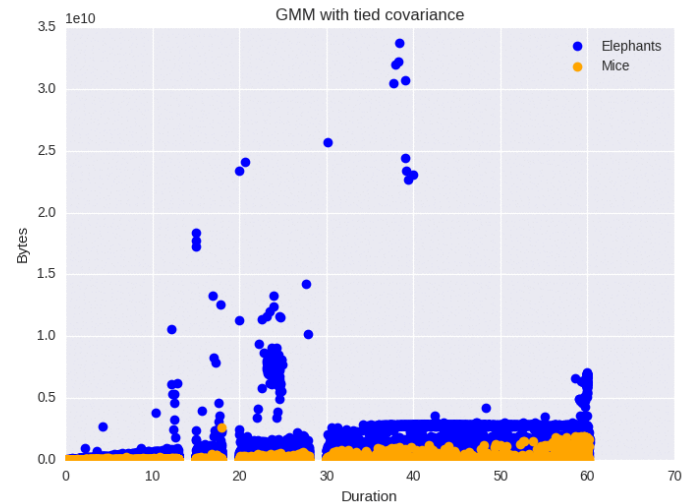
Predicted Link Topology - After 10 minutes



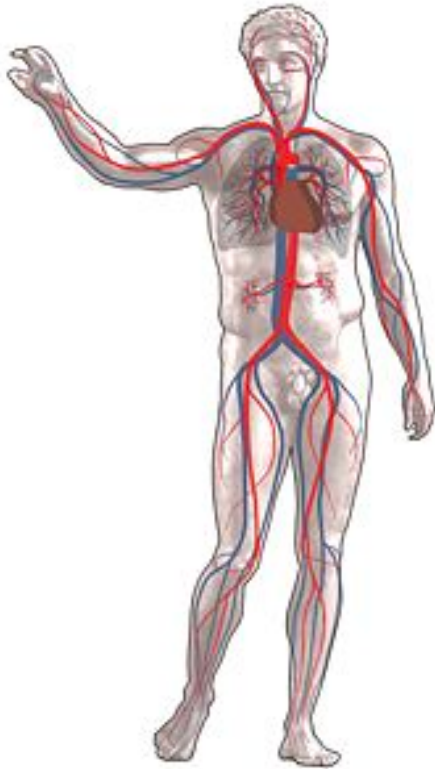
0: 'GA', 1: 'BNL', 2: 'PNNL', 3: 'JAMES', 4: 'ESNETW', 5: 'ORAU', 6: 'NSO', 7: 'LANL', 8: 'NERSC', 9: 'LNS', 10: 'ESNETE', 11: 'SNLL', 12: 'FNAL', 13: 'PANTEX', 14: 'LBL', 15: 'SNLA', 16: 'OSTI', 17: 'DOEG', 18: 'PSFC', 19: 'UGO', 20: 'NREL', 21: 'ORNL', 22: 'MRC', 23: 'DOEF', 24: 'KCF', 25: 'PPPL', 26: 'NGA', 27: 'SLAC', 28: 'Y12', 29: 'JLAB', 30: 'SRS', 31: 'ANL', 32: 'LLNL', 33: 'INL', 34: 'CERN', 35: 'KCF'



*92.12% Accurate when compared with actual packet transfers between links



ESnet is the circulatory system of DOE Office of Science



1. ESnet facility is **engineered and optimized** to meet the diverse needs of DOE Science
2. ESnet6 is architected with **capability to meet FY2025** Science needs
3. ESnet6 is being designed as a **flexible, innovation-enabling platform**

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

We are hiring!!

- Do interesting things!
- Work on a one of a kind, global scale network!
- Learn from smart people!

Network Engineers!

<https://lbl.taleo.net/careersection/engineer/jobdetail.ftl?job=83959>

Software Engineers!

<https://lbl.taleo.net/careersection/engineer/jobdetail.ftl?job=84046>

Thank You and Questions?

