ESnet
ENERGY SCIENCES NETWORK

Building a Data Infrastructure

(A WAN Provider's Perspective)

Chin Guok
ESnet Planning & Architecture Group Lead
Lawrence Berkeley National Laboratory

Data on the Network
Computing and Software Roundtable
Virtual Meeting
Apr 6, 2021

U.S. DEPARTMENT OF ENERGY
Office of Science

BERKELEY LAB

What Do You Need for a Data (Network) Infrastructure?

- **Capacity** - Planning and deploying of the right amount of capacity in the right place, and at the right time
- **Capability** - Building services that will be integral to a Data Infrastructure that is beyond just networking
- **Coordination** - Enabling the orchestration of resources to build end-to-end solutions whose whole is greater than the sum of its parts

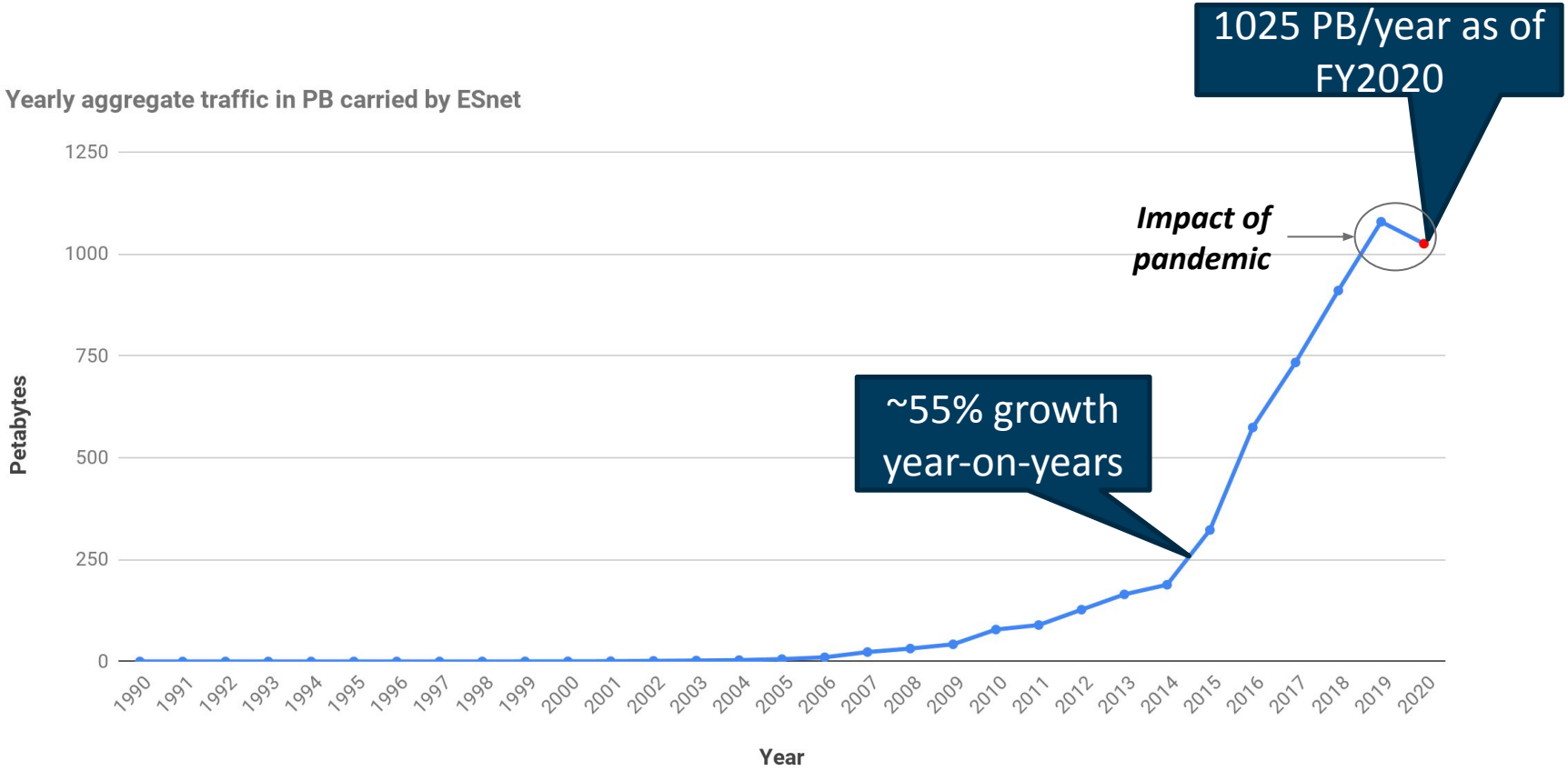
ESnet

ESnet Data Infrastructure Activities

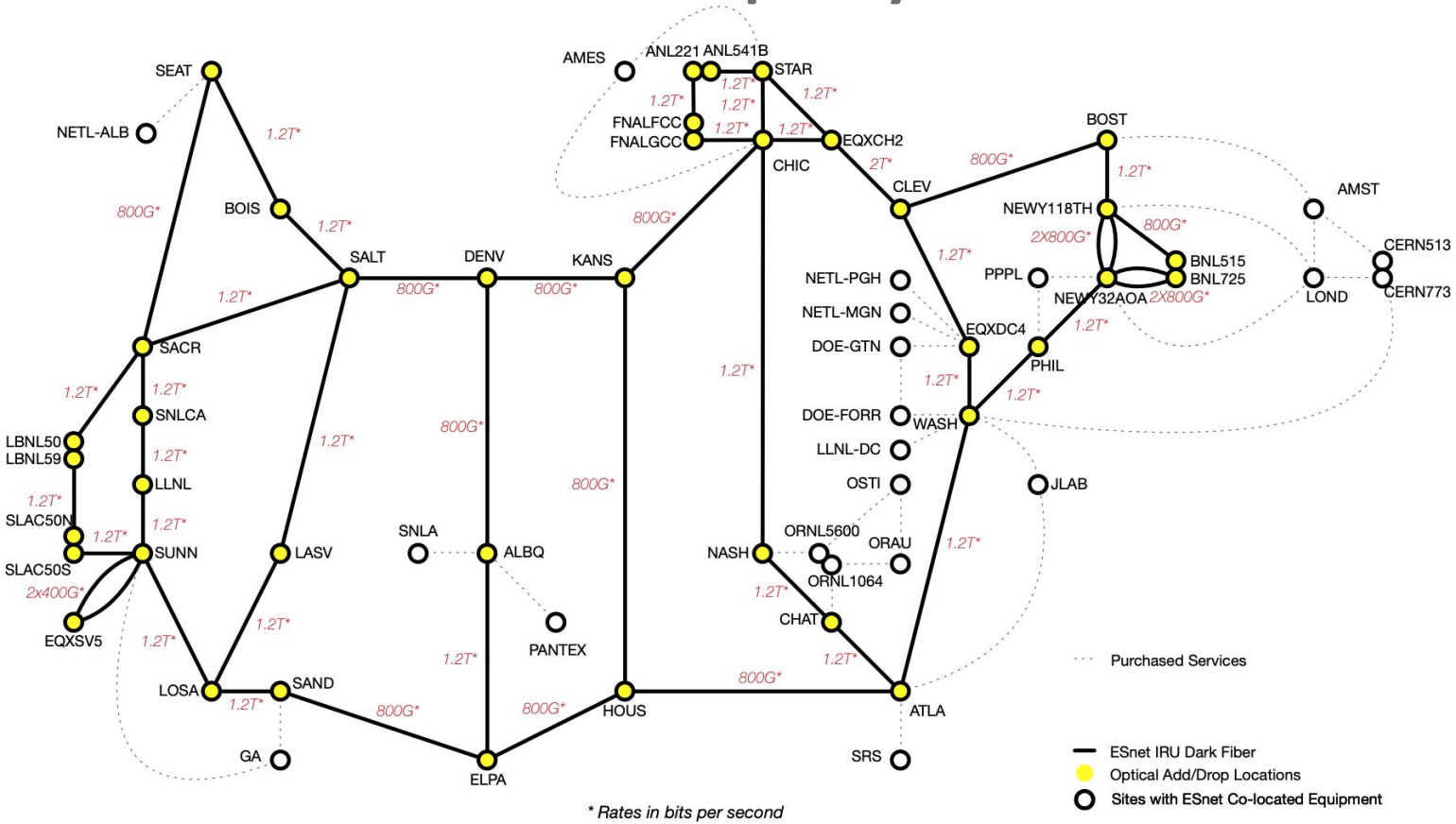
- **Capacity**
  - ESnet6 CY2022 Planned Capacity
- **Capability**
  - In-Network Caching Pilot
  - DTN-as-a-Service efforts
  - Edge Computing
- **Coordination**
  - SENSE: SDN for End-to-end Networking Science at the Exascale
  - ASCR Integrated Research Infrastructure Task Force

ESnet

An Exabyte Network Today



ESnet6 Planned Backbone Capacity for FY2022



* Rates in bits per second

ESnet

ESnet Data Infrastructure Activities

- Capacity
  - ESnet6 CY2022 Planned Capacity
- **Capability**
  - In-Network Caching Pilot
  - DTN-as-a-Service efforts
  - Edge Computing
- Coordination
  - SENSE: SDN for End-to-end Networking Science at the Exascale
  - ASCR Integrated Research Infrastructure Task Force

ESnet

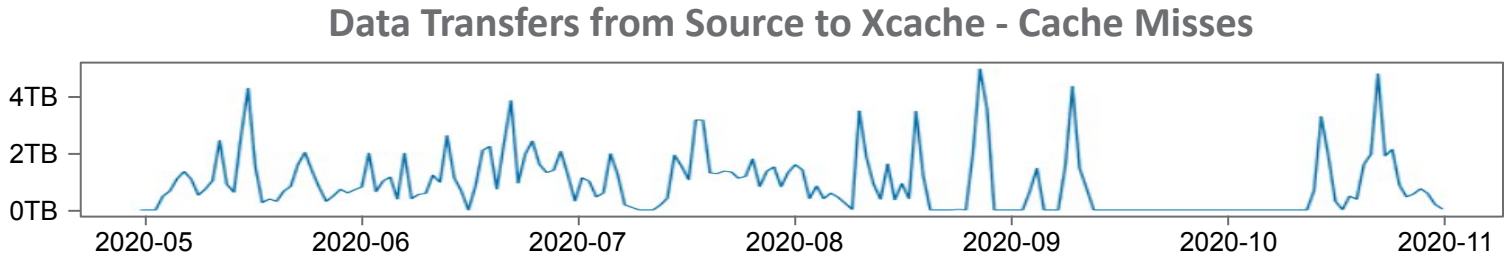
In-Network Caching Pilot Experiment

- In-network temporary data cache for data sharing
- Collaboration with UCSD, US CMS, Open Science Grid (OSG)
  - ESnet cache node as a part of SoCal Petabyte scale cache for US CMS
  - Petabyte scale cache is deployed/operated by UCSD and Caltech
  - ESnet: Provide a storage host to US CMS
  - UCSD: Operation support for the ESnet node
- Goals
  - Study how network cache storage helps network traffic performance and the overall application performance
  - Accumulate experience on how the US DOE scientific experiments and simulations share data among their users

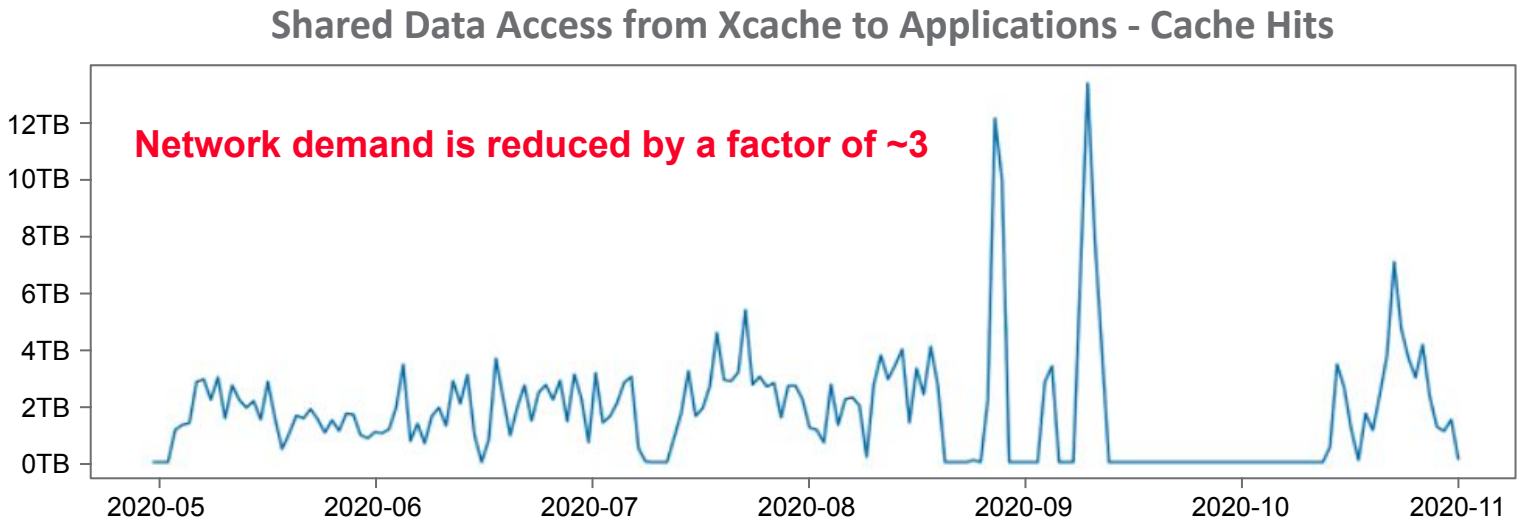
ESnet

Sim et al.

In-Network Caching Pilot Data Movements



Total Data Transfer
= 168.083TB

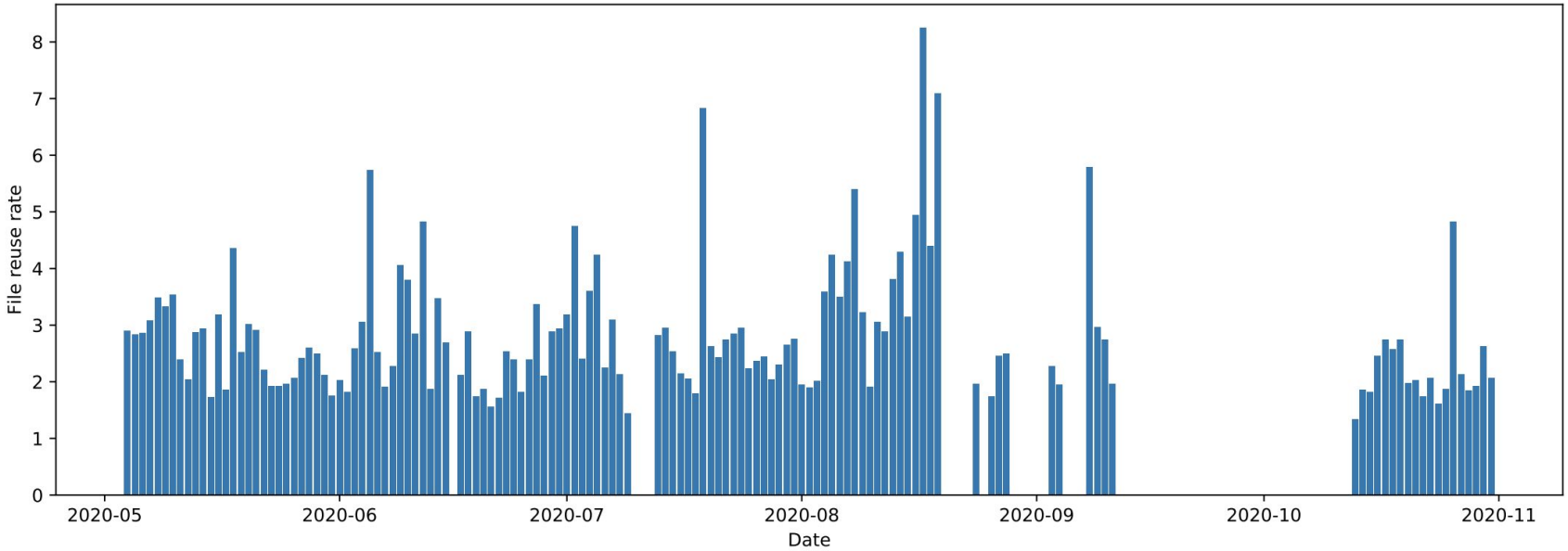


Total Shared
Data Access
= 322.748TB

ESnet

Sim et al.

In-Network Caching Pilot File Reuse Rates



- Data file reuse rate = (sum of reuses) / (total number of unique files)
  - Sum of reuses = all shared data access counts of the day (cache hits)
  - Total number of unique files = number of unique files for the day

ESnet

Sim et al.

In-Network Caching Pilot Summary

- 1,286,748 accesses from May 2020 to Oct 2020
  - Total 490.831 TB of client data access (first time reads and repeated reads)
  - Transferred/cached 168.08 TB (from remote sites to cache)
  - Saved 322.748 TB of network traffic volume (repeated reads only)
  - **Network demand reduced by a factor of ~3**
- Future work
  - Deployment of 2 additional caches in summer 2021
  - Cache miss rate study
  - Cache utilization study

ESnet

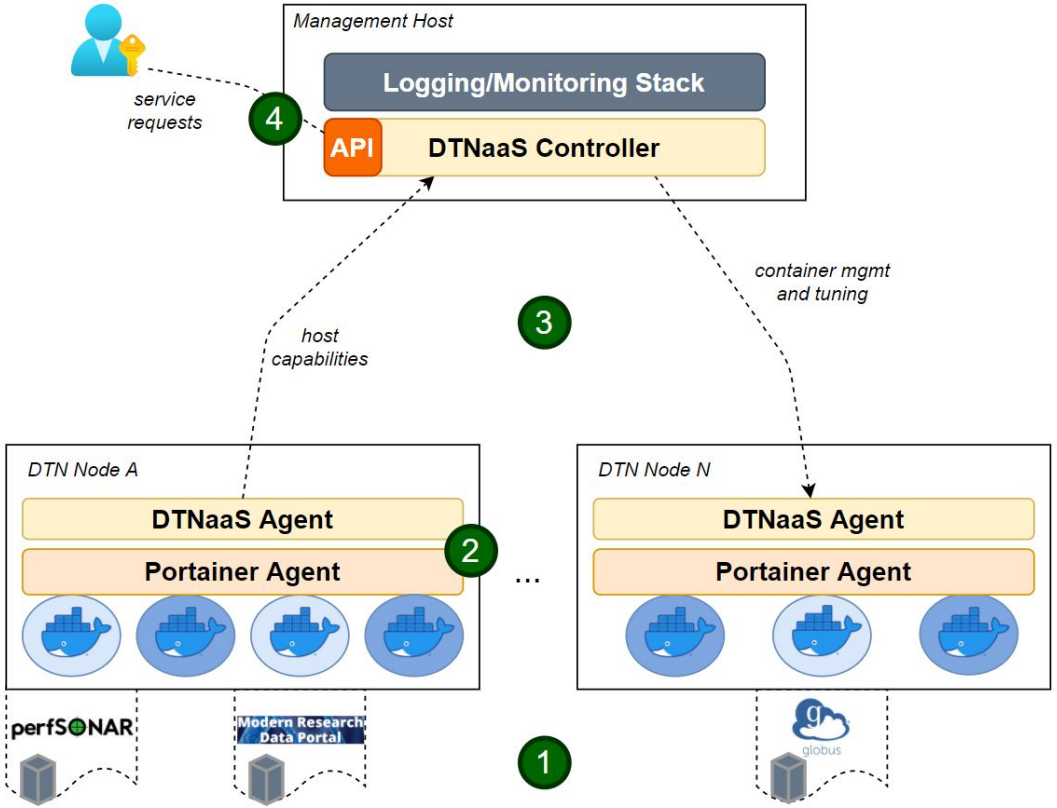
Sim et al.

DTN-as-a-Service

- Exploring an ESnet-managed data movement service
  - Support a pool of transfer software images that “just work” across the ESnet footprint
  - Reduce reliance on varying levels of (site) network/system expertise for DTN deployments
- Investigate opportunities to expand data movement software availability and to support more flexible service deployments
- A potential candidate for supporting the ESnet6 smart services edge

ESnet

DTN-as-a-Service Prototype Service



1. Data mover software in containers
2. Network and storage performance optimization
3. Configuration and tuning flexibility
4. Lightweight service orchestration

ESnet

Kissel et al.

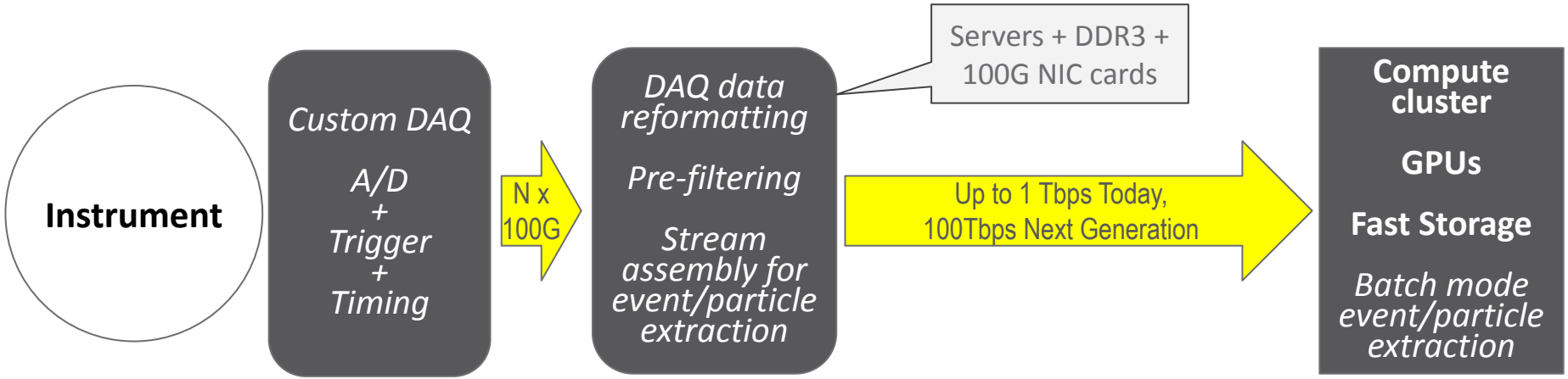
DTN-as-a-Service Use Cases and Ongoing Engagements

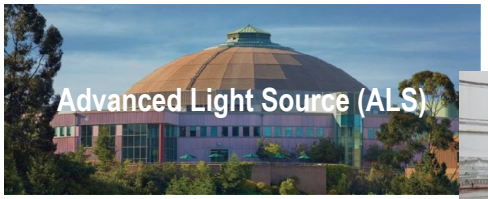
- Transitioning the ESnet 10G/40G testing DTNs to DTNaaS
- Deployed in the ESnet 100G Testbed to manage experimenter images
- Evaluation of data mover software stacks supported by DTNaaS
  - For example - Zettar/zx[1]
- Investigating a deployment model to support In-Network Data Caching effort
- Seeking partnerships for service trial
  - Using DTNaaS to accelerate an existing data workflow challenge
  - Outreach to JGI, FRIB, EMSL

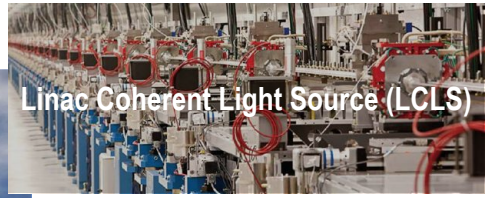
ESnet

[1] E. Kissel, C. Fang, Zettar zx Evaluation for ESnet DTNs
https://www.es.net/assets/Uploads/zettar-zx-dtn-report.pdf

Edge Computing for Real Time Workflows - Today







ESnet

Yatish et al.

Edge Computing for Real Time Workflows - Future

Low cost scalable FPGA packet processors to replace expensive CPU+DRR solutions

100% link load balancing
(Out of order UDP protocols)
Real time data compression

FPGA HW laid balancing into cluster. Group DAQ data by timestamp and frame for parallel data processing

Instrument

Custom DAQ

A/D
+
Trigger
+
Timing

N x 100G

DAQ data reformatting

Pre-filtering

Stream assembly for event/particle extraction

Up to 1 Tbps, 100Tbps NxGen

Out of order reassembly

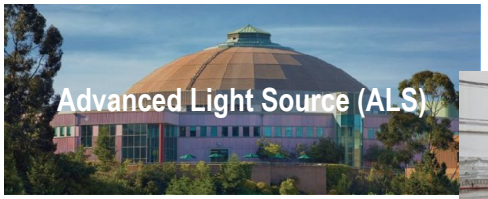
Local compute load balancing

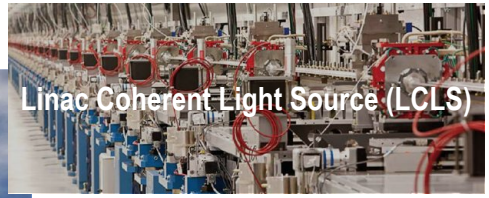
Compute cluster

GPUs

Fast Storage

Batch mode event/particle extraction





ESnet

Yatish et al.

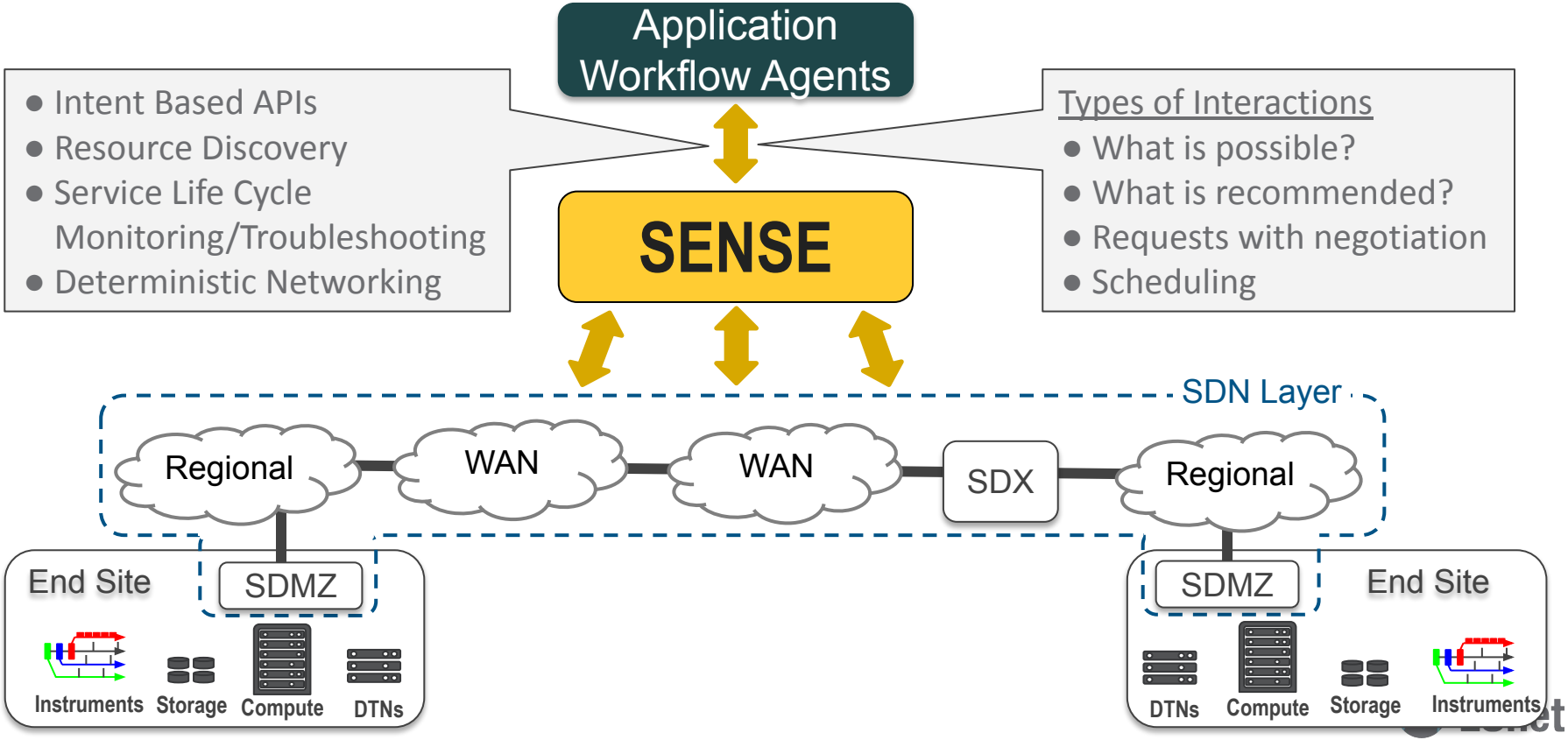
ESnet Data Infrastructure Activities

- **Capacity**
  - ESnet6 CY2022 Planned Capacity
- **Capability**
  - In-Network Caching Pilot
  - DTN-as-a-Service efforts
  - Edge Computing
- **Coordination**
  - SENSE: SDN for End-to-end Networking Science at the Exascale
  - ASCR Integrated Research Infrastructure Task Force

ESnet

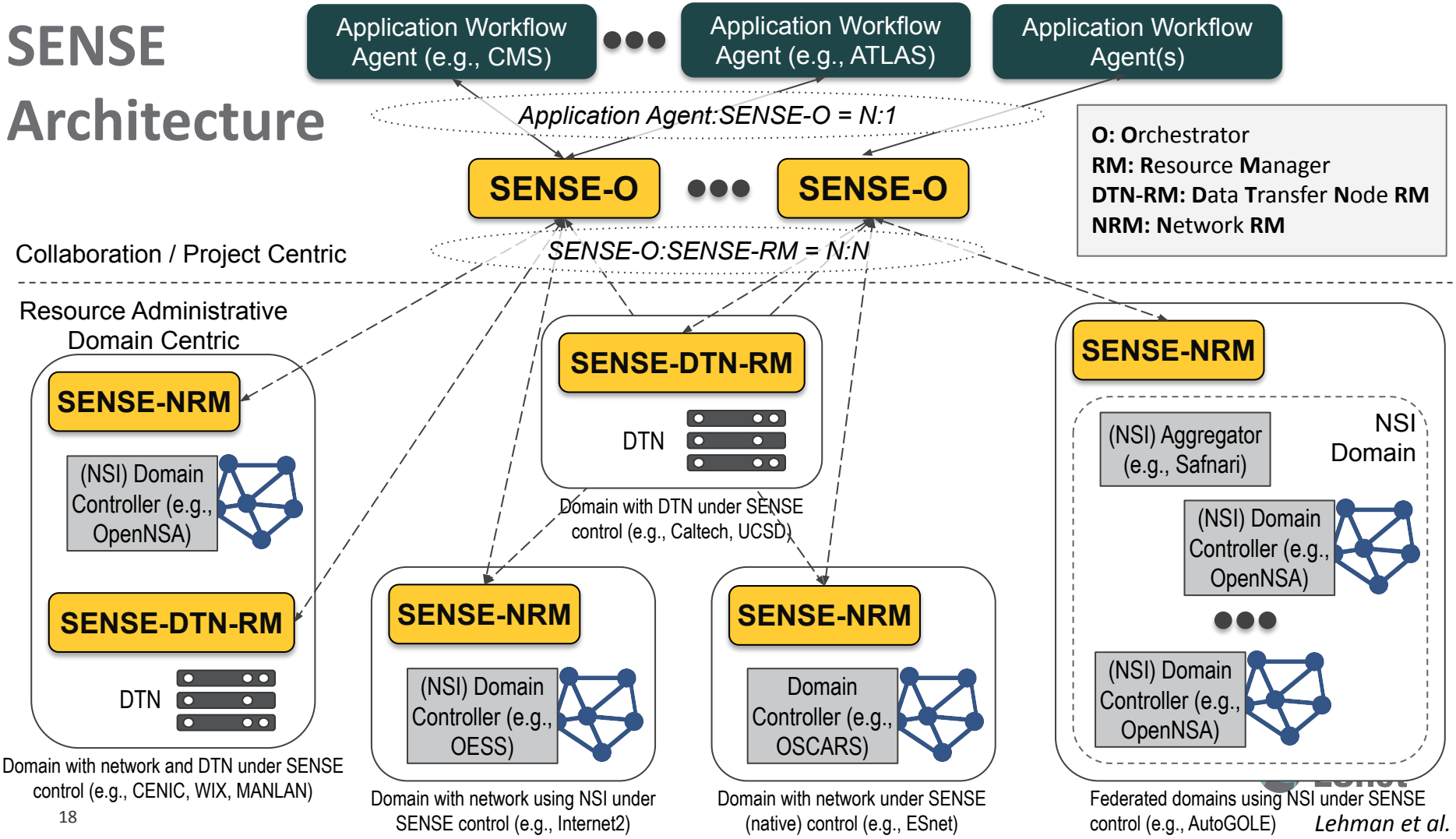
SENSE: SDN for End-to-End Networked Science at the Exascale

- Coordination of End-to-End Network Cyberinfrastructure



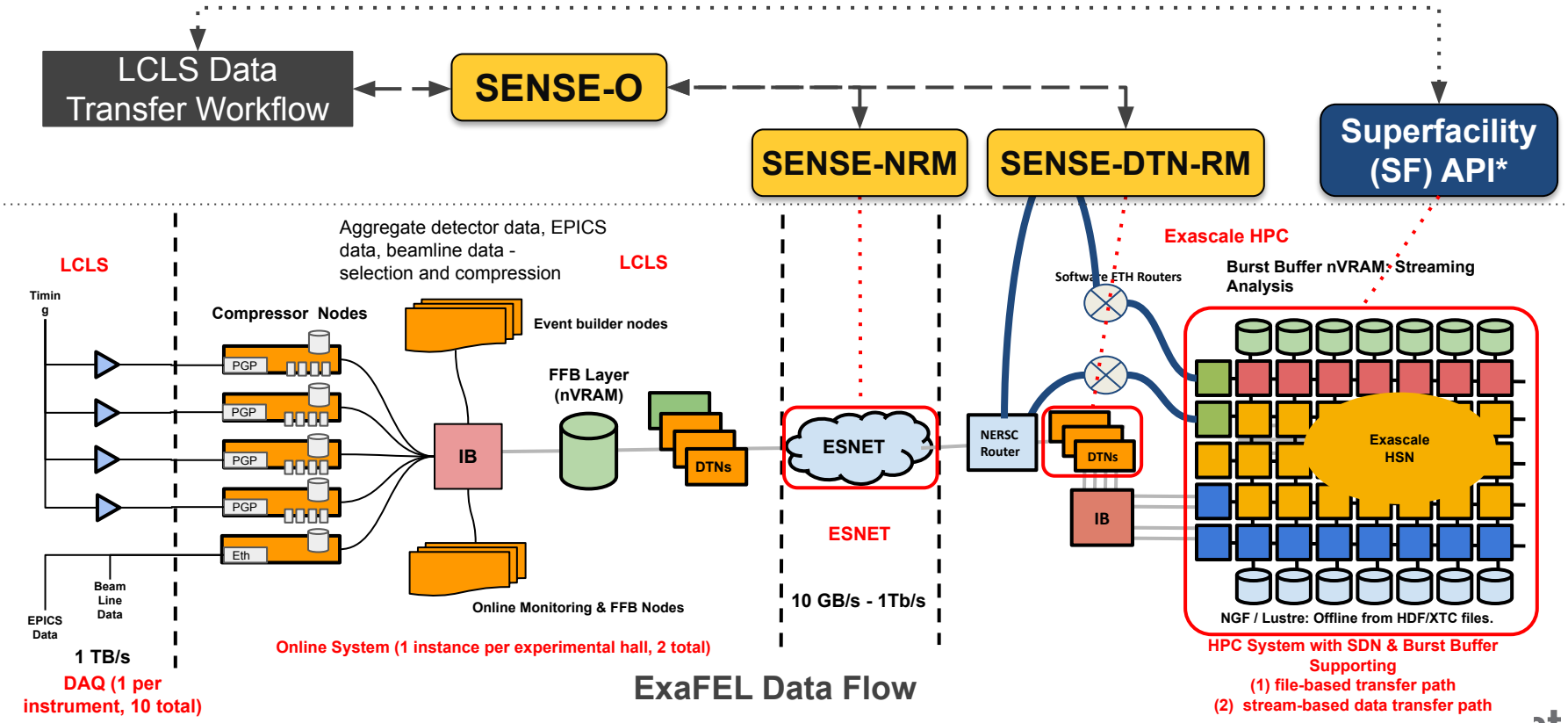
Lehman et al.

SENSE Architecture



Lehman et al.

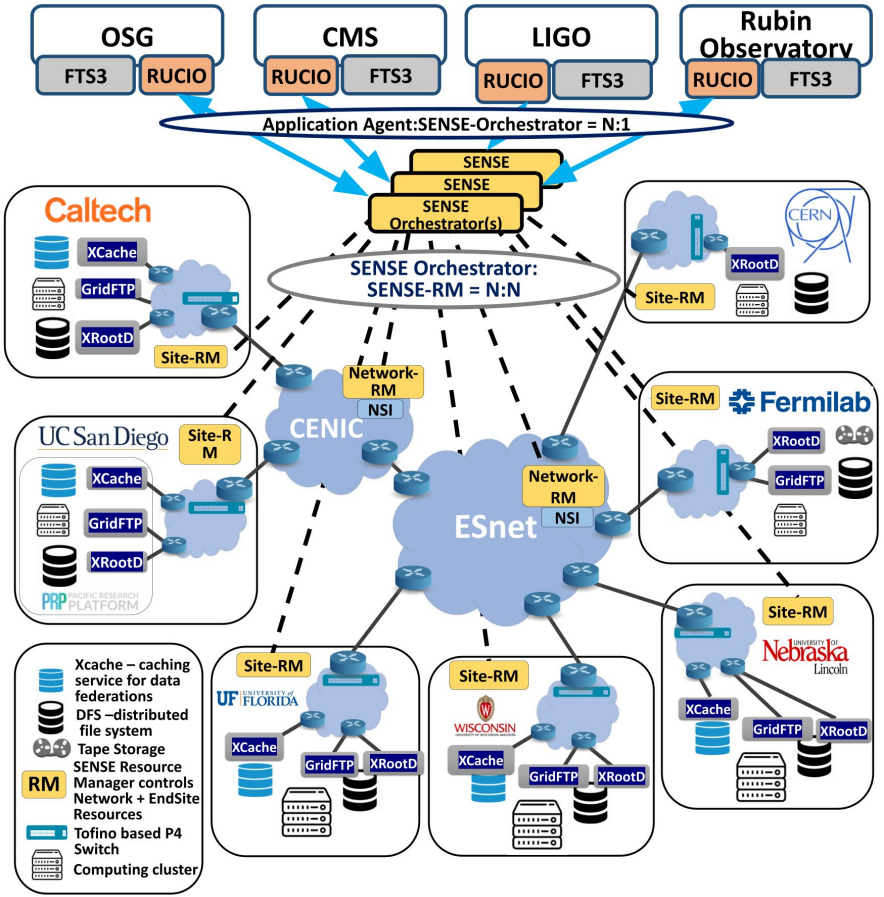
SENSE ExaFEL Use Case - Superfacility Automation Prototype



*NB: Superfacility (SF) API under development by NERSC

Lehman et al.

SENSE Future Domain Science Integrations



- RUCIO/FTS expected to be used in multiple domain science workflows
- SENSE Service and Orchestration model designed for adaptation and customization for different infrastructures, service requirements, and workflows

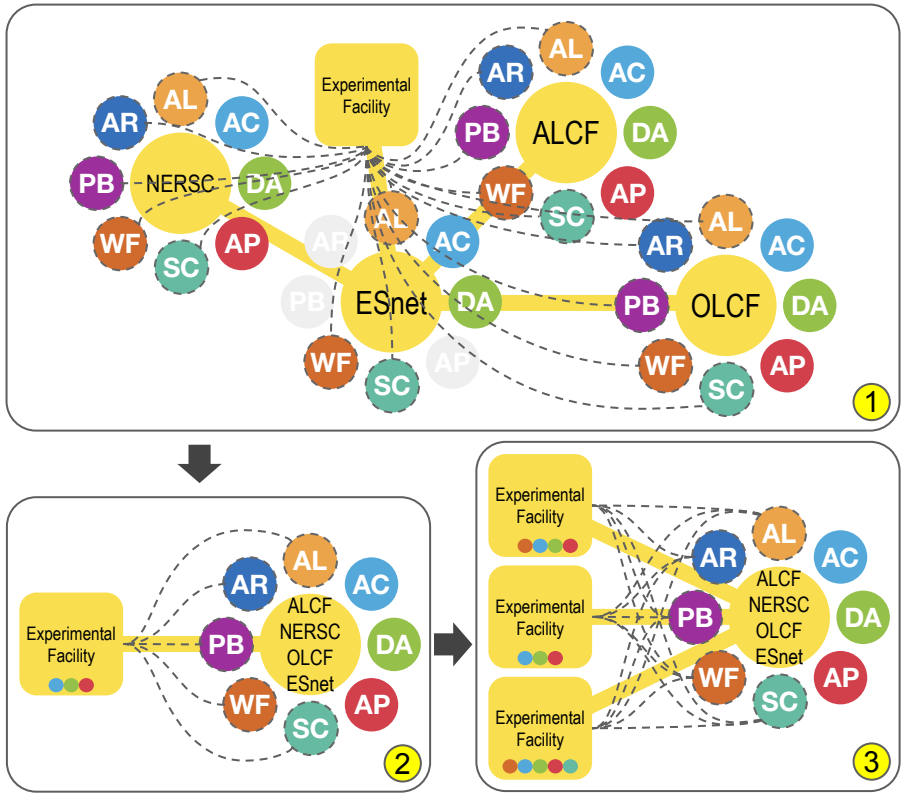
ESnet

Lehman et al.

ASCR Integrated Research Infrastructure Task Force

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

1. Today, an experimental facility must arrange separate bespoke interactions with individual HPC/HPN facilities.
2. A future paradigm with common interfaces could simplify integration of an experimental facility with multiple HPC/HPN facilities.
3. In turn, these common interfaces could support expansion and integration across multiple experimental facilities and HPC/HPN facilities



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

ESnet

*NB: Whitepaper to be released soon.

Brown et al.

Questions...

ESnet