

StarLight Software Defined Exchange (SDX) And The Global Research Platform

Joe Mambretti, Director, (j-mambretti@northwestern.edu)

International Center for Advanced Internet Research (www.icaair.org)

Northwestern University

Director, Metropolitan Research and Education Network (www.mren.org)

**Director, StarLight International/National Communications Exchange Facility
(www.startap.net/starlight),**

PI: StarLight SDX, Co-PI Chameleon, PI-iGENI, PI-OMNINet

Global Research Platform

Co-Located With The 19th IEEE International Conference on eScience

Osaka, Japan

September 16-17, 2024



The GRP: A Platform For Global Science



*A Next Generation, Software Defined,
Globally Distributed, Multi-Domain
Computational Science Environment*

Global Research Platform: Global Lambda Integrated Facility

Available Advanced Network Resources



Next Generation eScience Ecosystem

Visualization courtesy of Bob Patterson, NCSA; data compilation by Maxine Brown, UIC.



www.glif.is

STARLIGHTSM

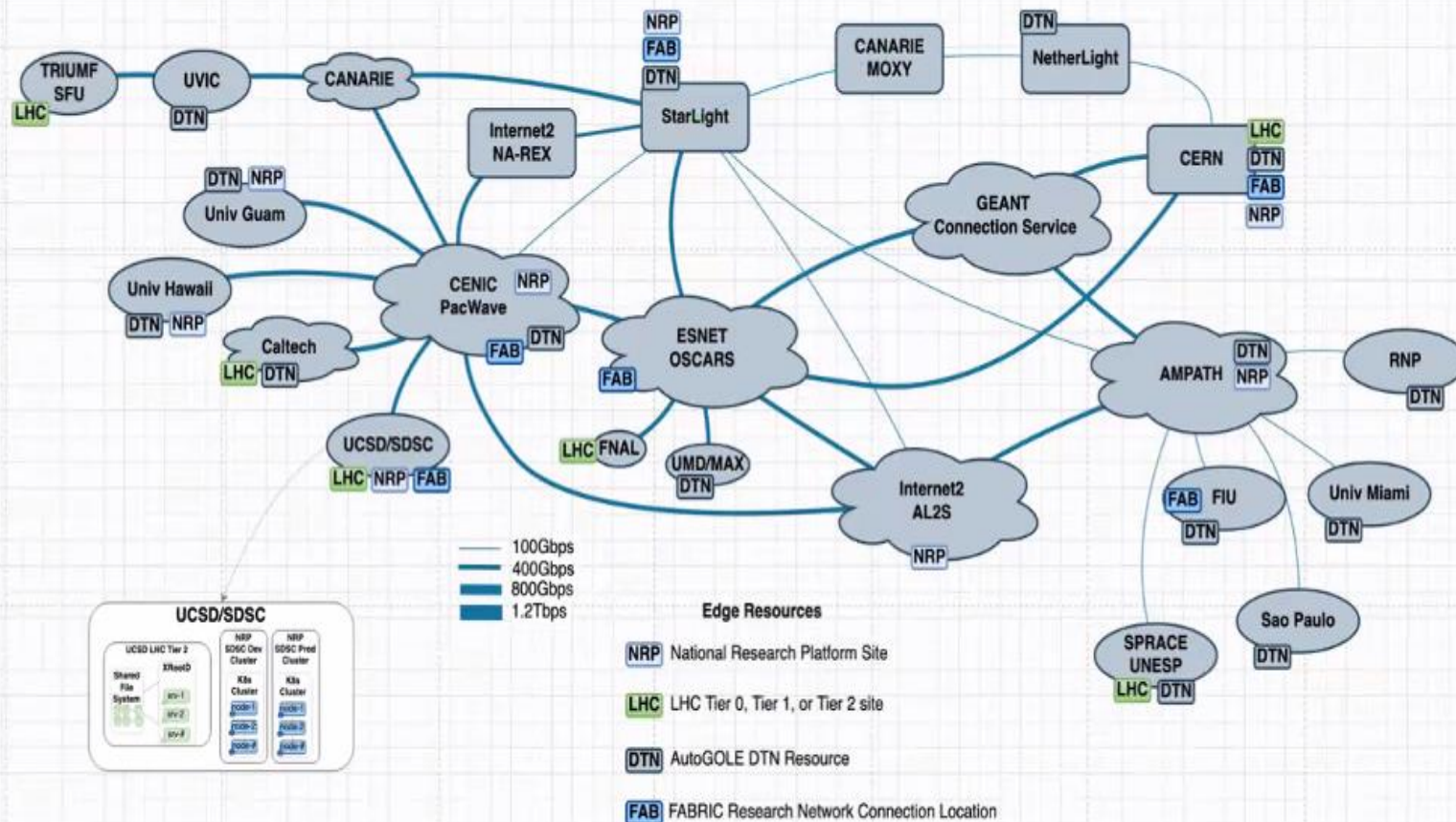
"The global advancement of science by realizing a multiresource infrastructure through international collaboration."

Schematic overview of the GNA-G AutoGOLE

GNA-G
Global Network Administration Group



STARLIGHTSM



NRP
Connections

P4 Testbed Site/End System Data

StarLight – “By Researchers For Researchers”

StarLight: Experimental Optical Infrastructure/Proving Ground For Next Gen Network Services
Optimized for High Performance Data Intensive Science
Multiple 100 Gbps
(110+ Paths)
StarWave
100 G Exchange
World's Most Advanced Exchange
Multiple First of a Kind
Services and Capabilities

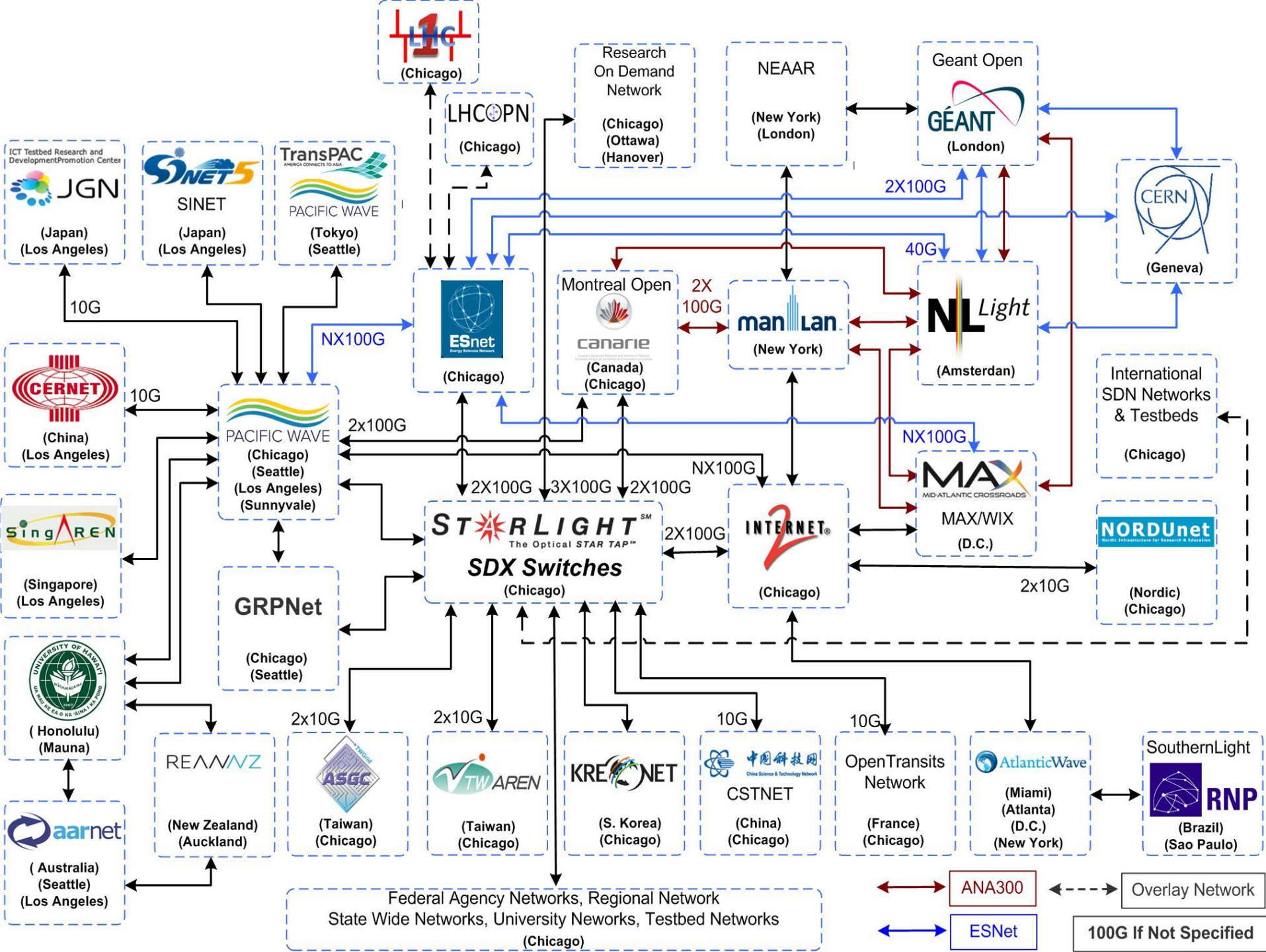


View from StarLight



Abbott Hall, Northwestern University's Chicago Campus

Currently: 26+ 400 Gbps Paths Prototyping 800 Gbps **StarLight** SM



ESnet6



ESnet



U.S. DEPARTMENT OF
ENERGY

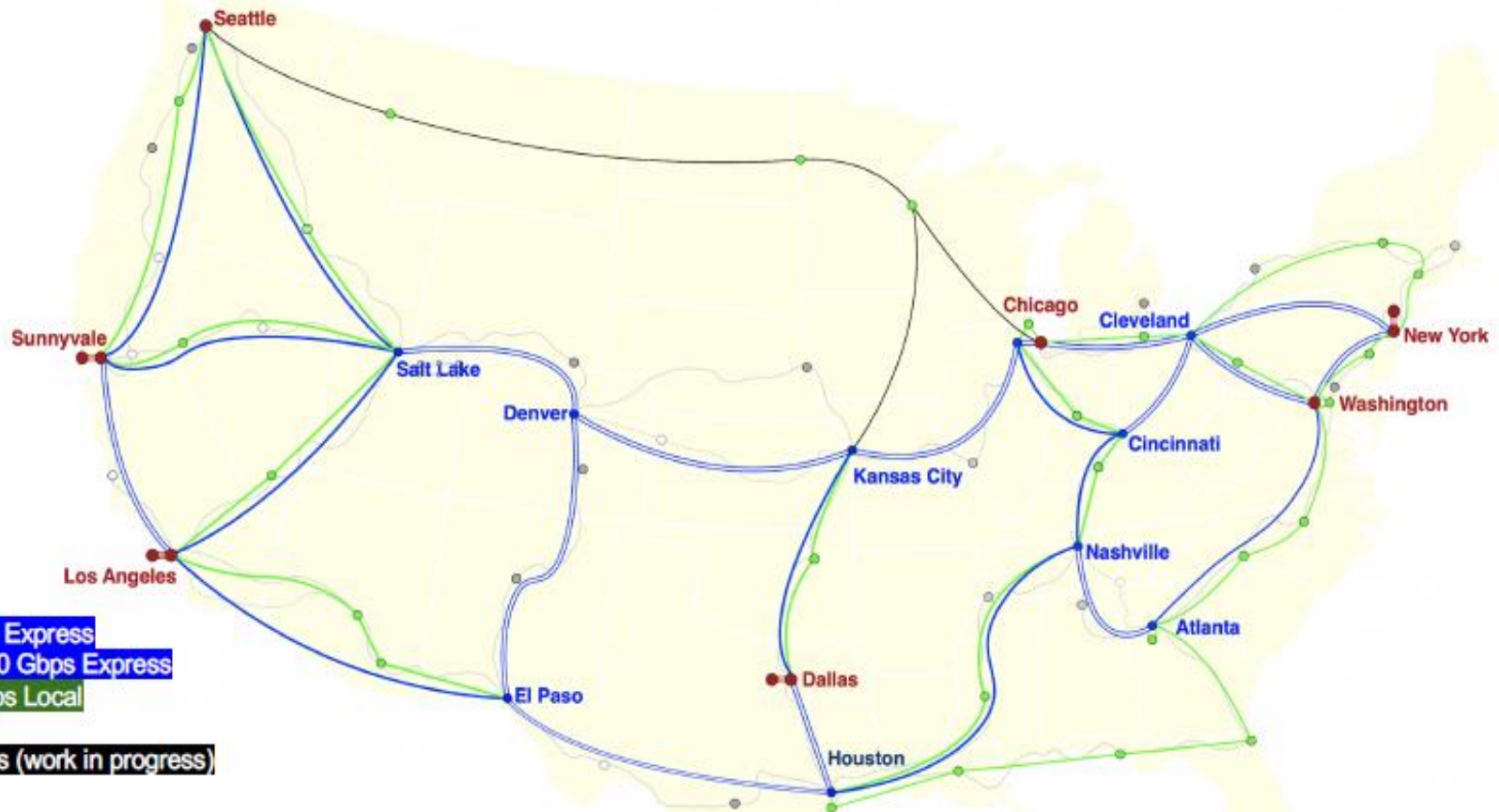
Office of
Science

STARLIGHTSM

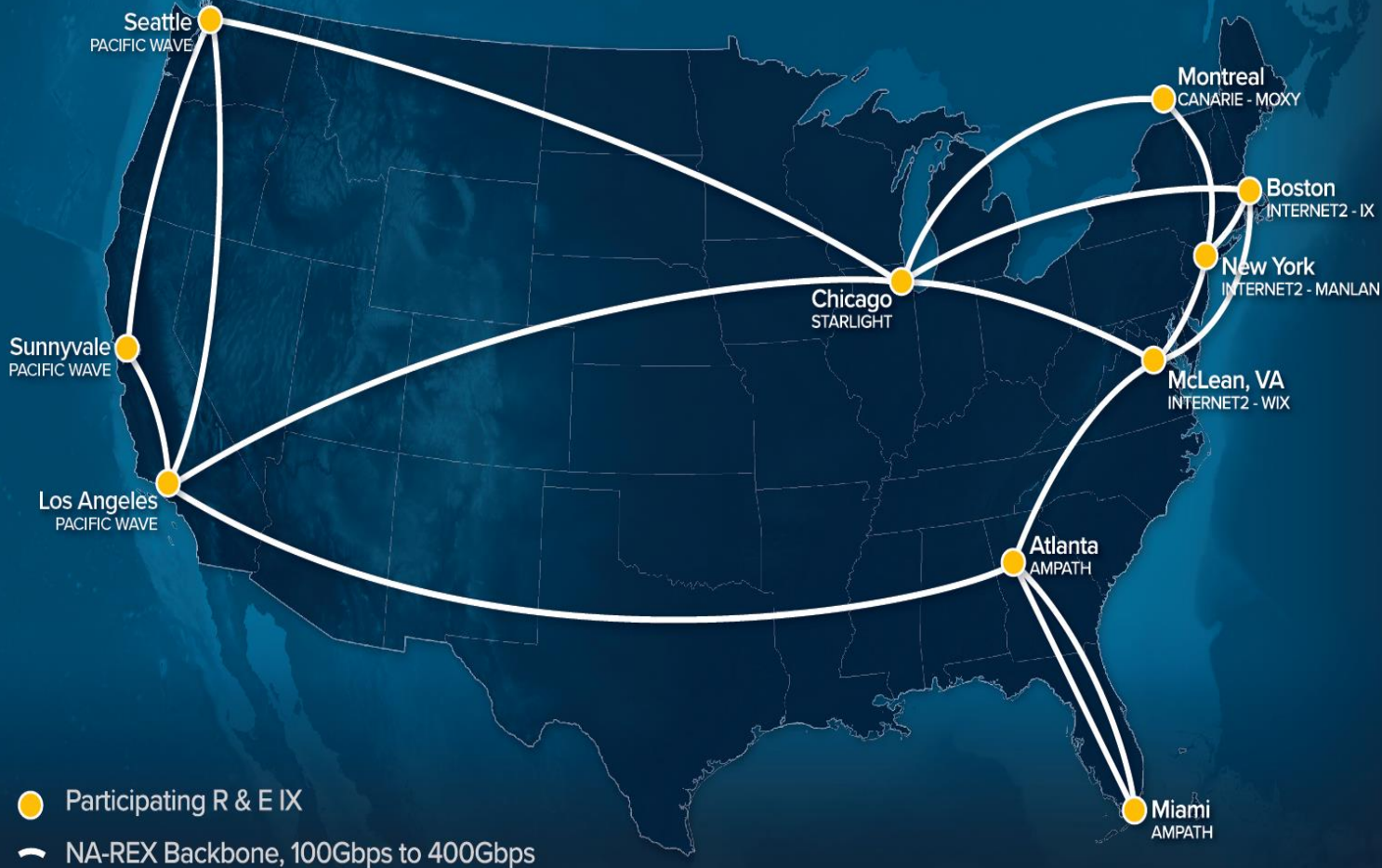
Internet2 Backbone Topology

Backbone Topology - Capacity and Traffic Management

Chris Wilkinson, Director of Planning and Architecture



NA-REX North America Research & Education Exchange Collaboration



International Federated Testbeds As Instruments for Computer Science/Network Science

- **The StarLight Communications Exchange Facility Supports ~ 28 Network Research Testbeds (Instruments For Computer Science/Networking Research)**
- **StarLight Supports Two Software Defined Exchanges (SDXs)**
- **An NSF IRNC SDX & A Network Research**
- **A Research SDX Supports National and International Federated Testbeds**

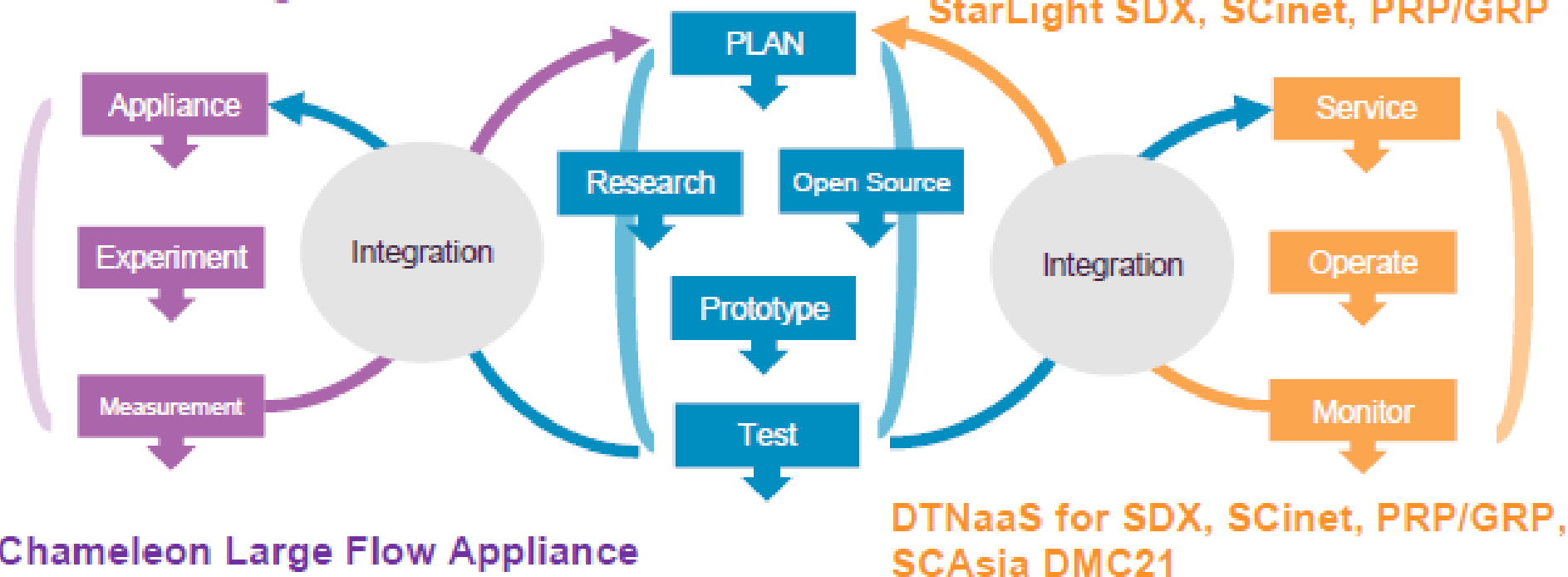


StarLight Software Defined Exchange

StarLight Software Defined Exchange (SDX) CD/CI/CD Innovation Workflow

StarLight Testbeds

StarLight SDX, SCinet, PRP/GRP



iCAIR

STARLIGHTSMSDX

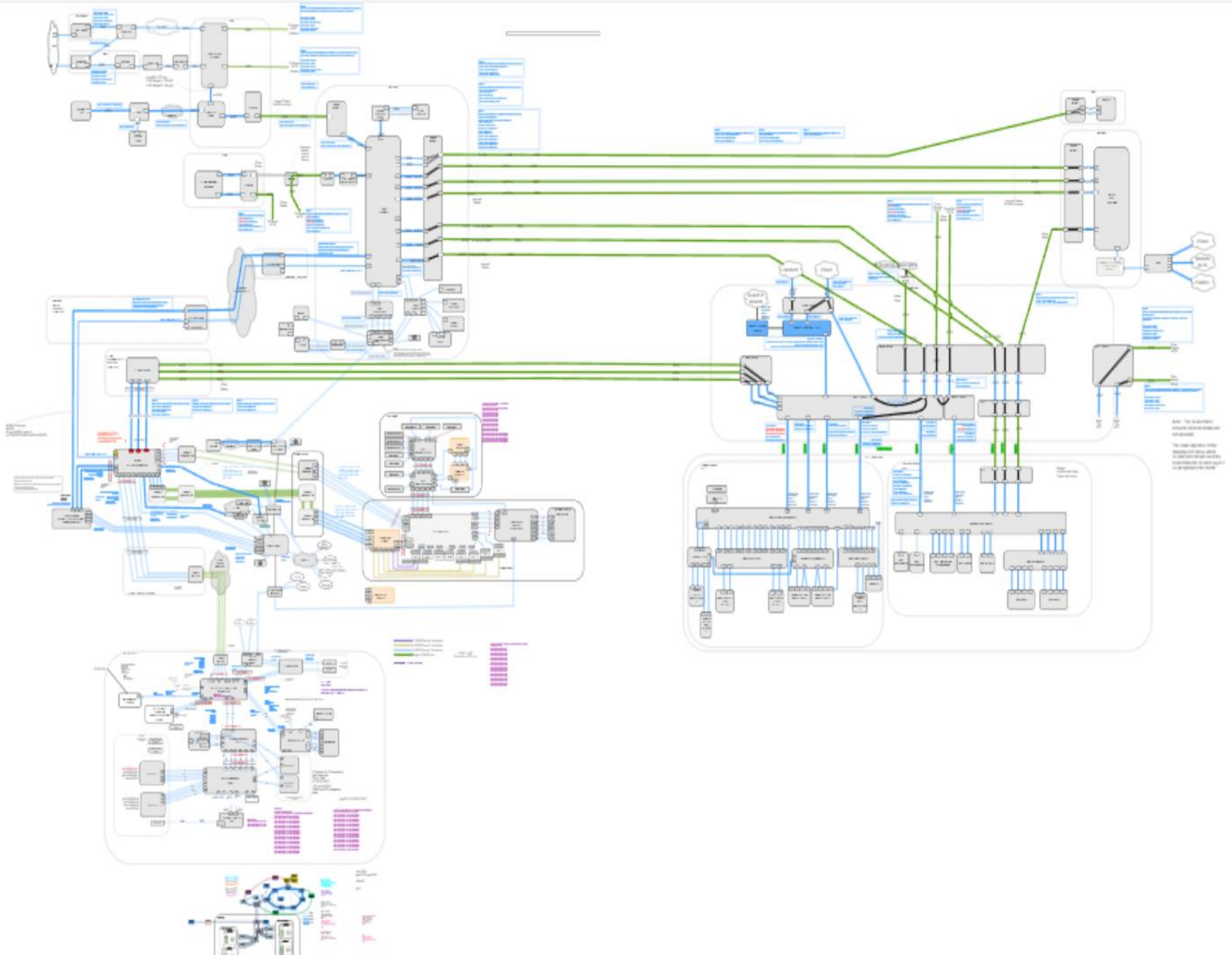
Source: Jim Chen

STARLIGHTSM

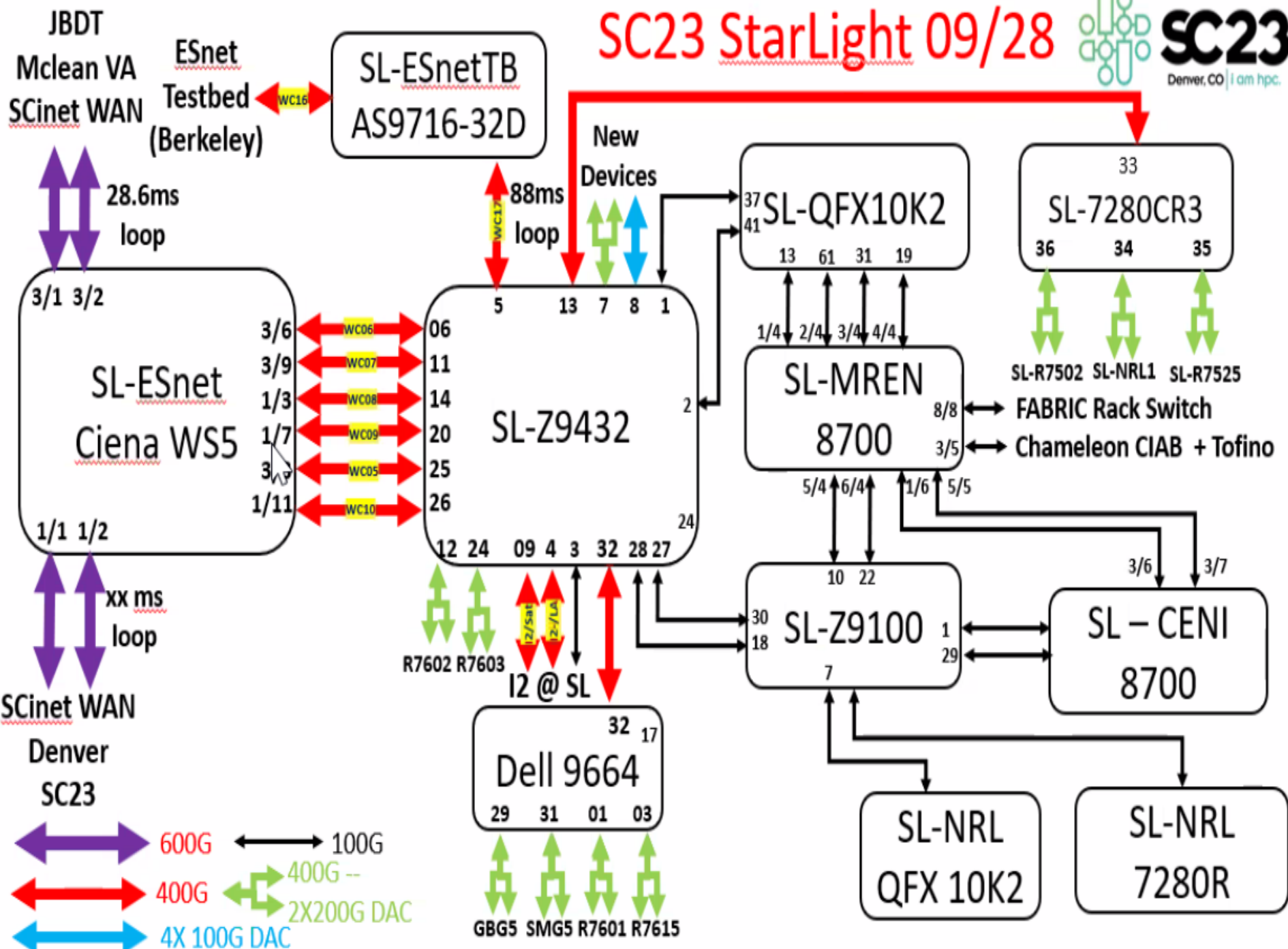
SCinet National WAN Testbed

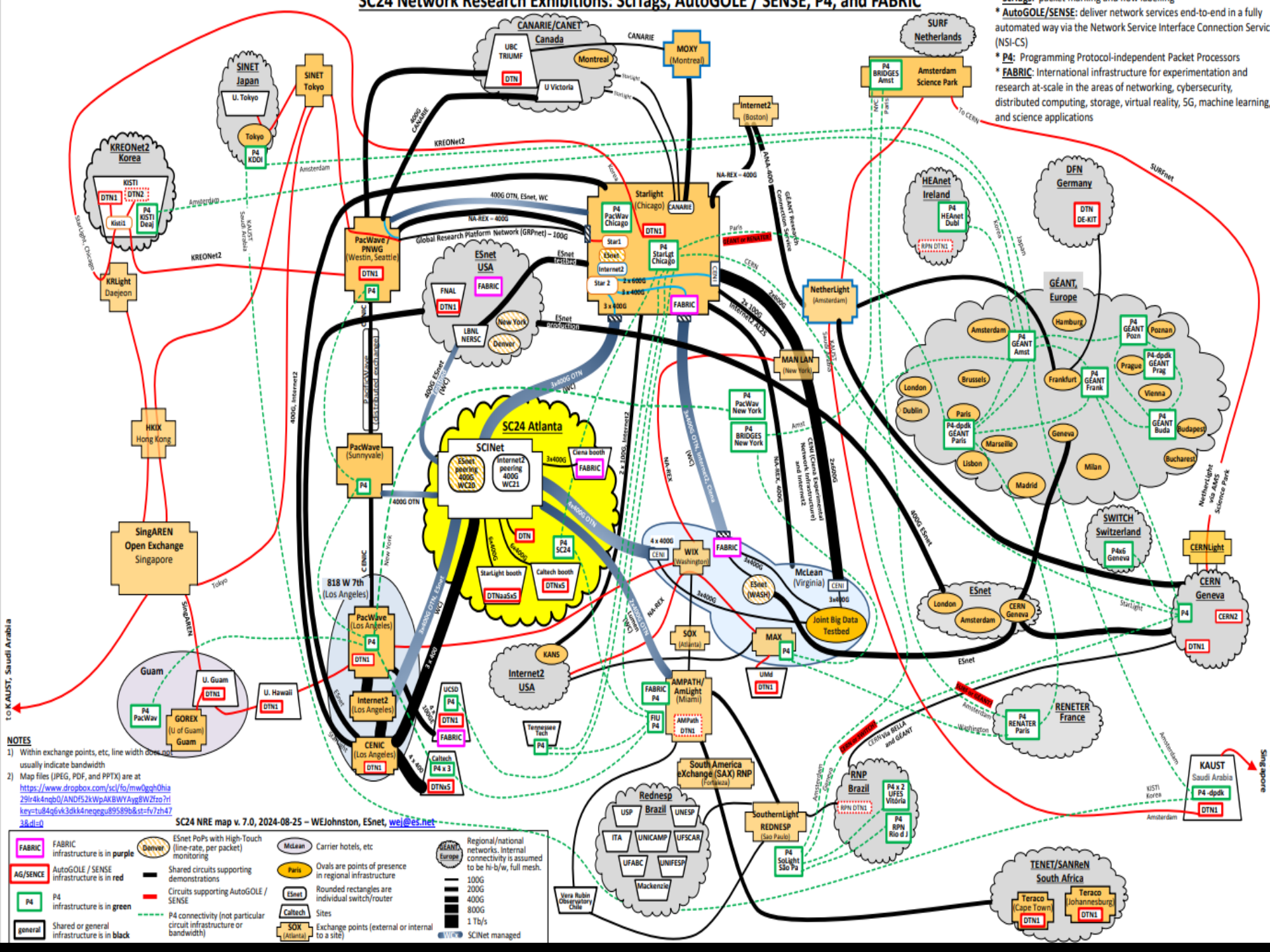
- **As In Previous Years, iCAIR Supports SCinet In Designing and Implementing a National WAN Testbed**
- **A Key Focus Is 400, 800, and 1.2 Tbps Path Services and Interconnections, Including Direct Connections To Edge Nodes, Primarily High Performance DTNs**
- **The SC23 National WAN Testbed Is Being Designed and Implemented To Support Demonstrations and Experiments Of Innovations Related To Data Intensive Science**





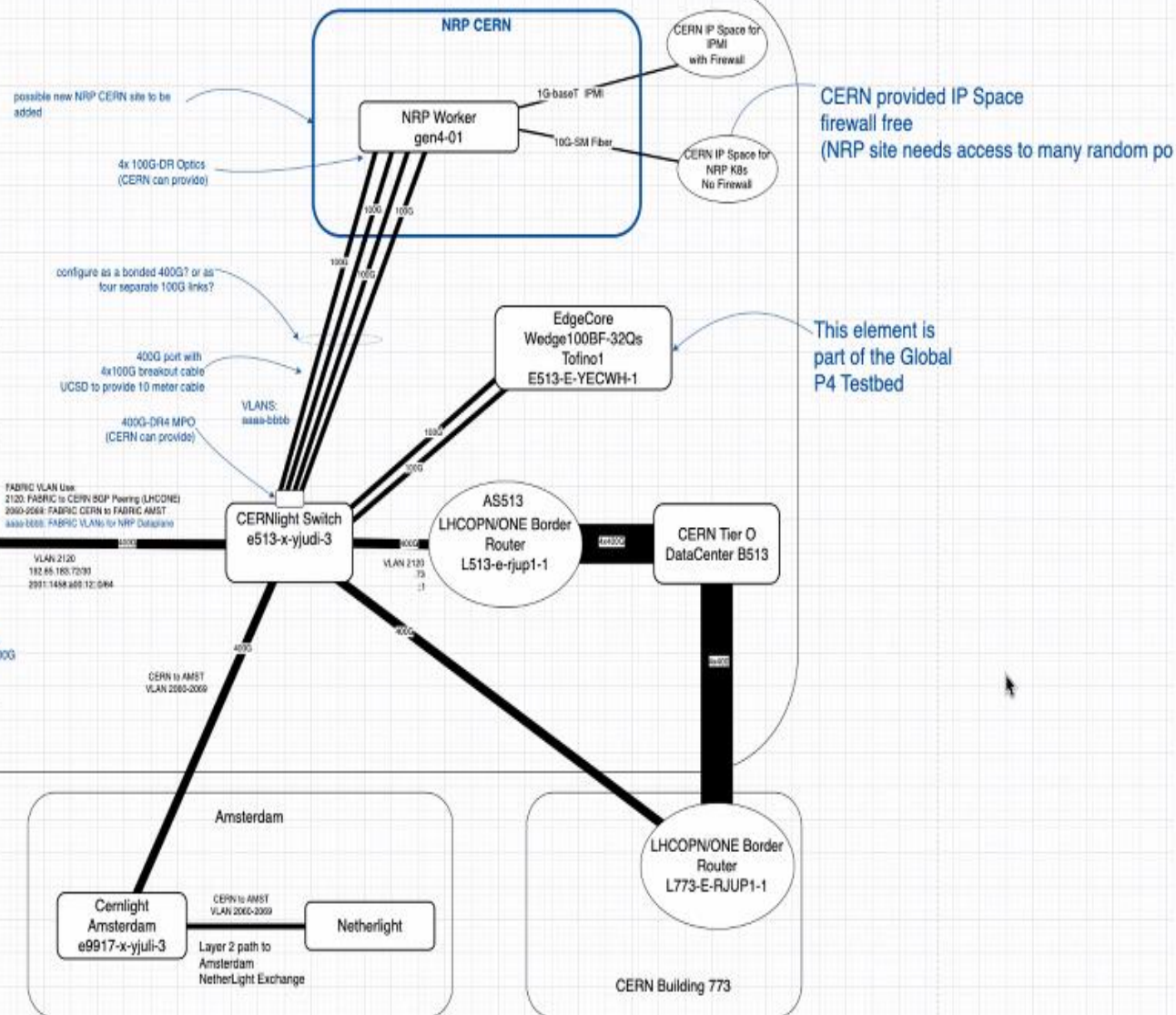
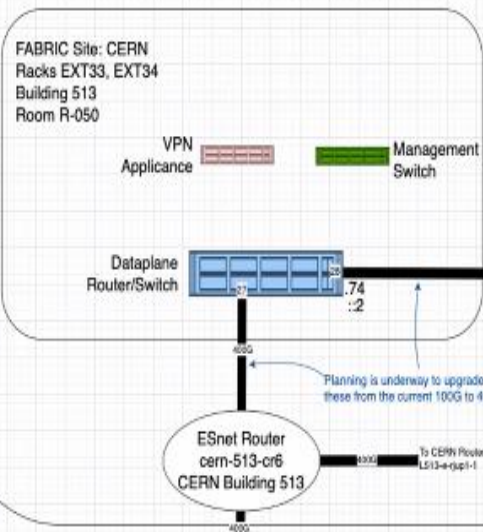
SC23 StarLight 09/28





CERN Building 513 (Data Center)

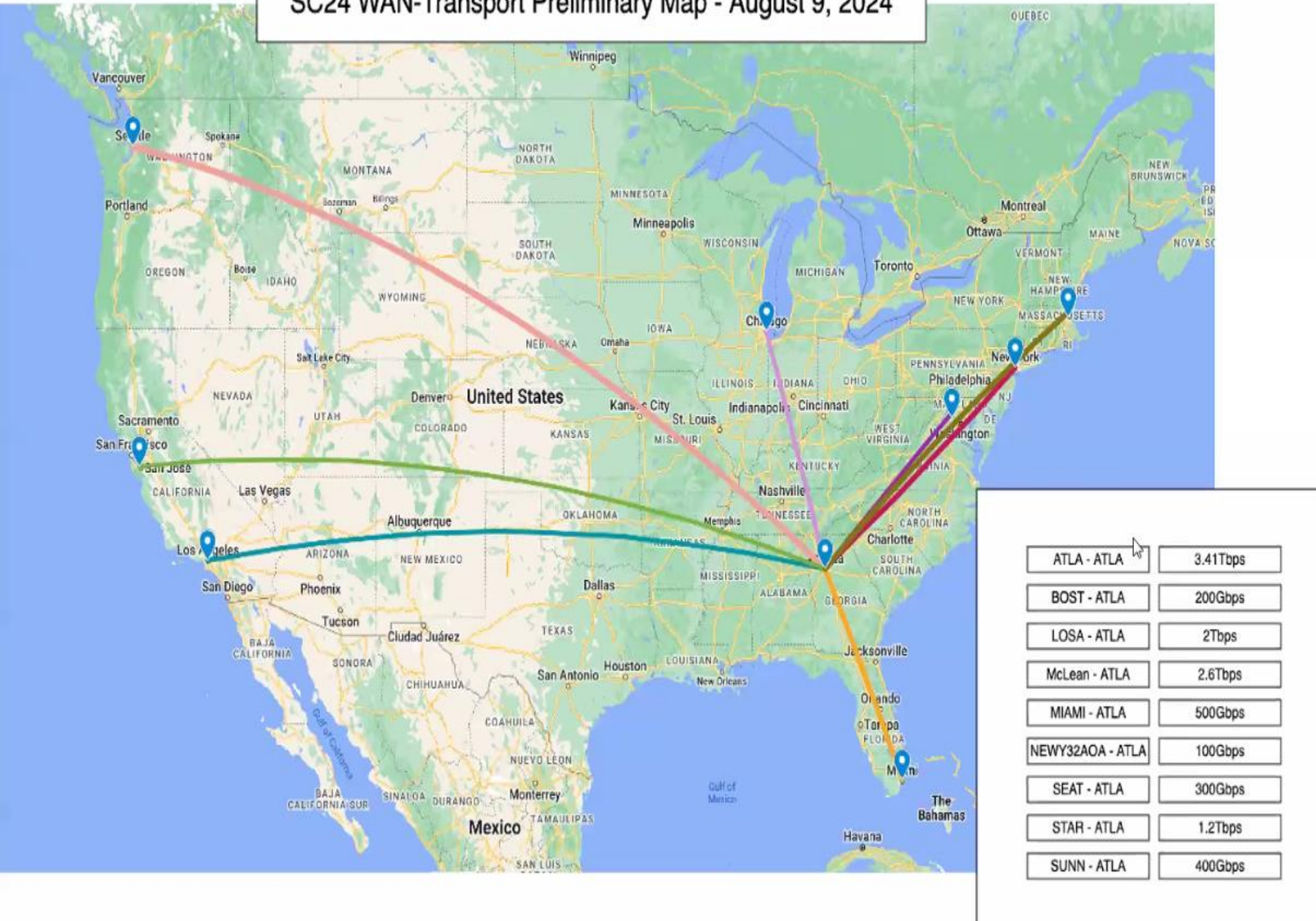
Network Cables



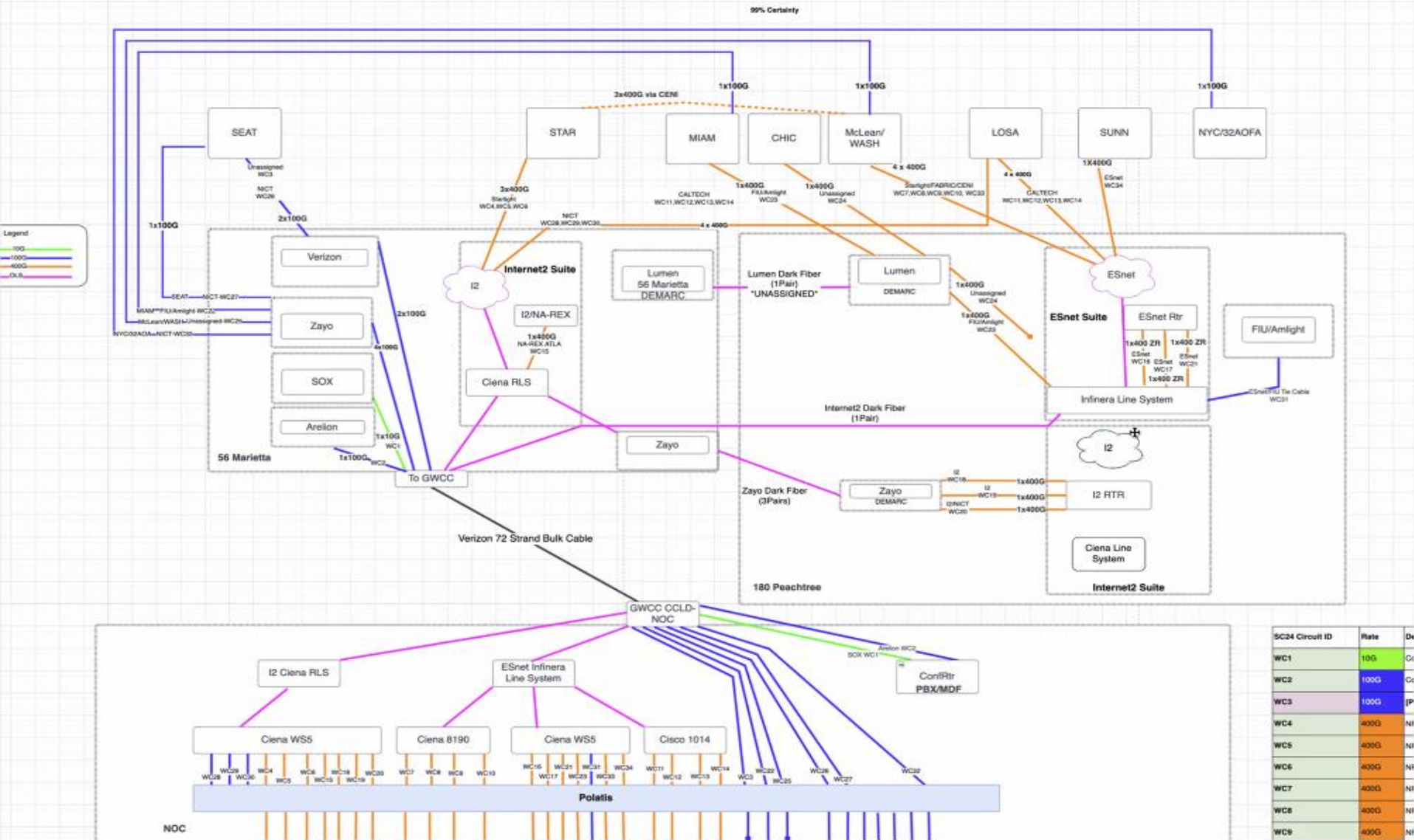
Example SC24 SCinet Network Research Exhibitions

- Global Research Platform (GRP)
- SDX 1.2 Tbps WAN Services
- SDX E2E 400 Gbps 800 Gbps WAN Services
- 400 Gbps DTNs & Smart NICs
- Orchestration With Packet Marking (SciTags)
- ESnet High Touch Network Measurements
- NA REX Continental Backbone For Data Intensive Science
- SDX International Testbed Integration
- StarLight SDX for Petascale Science
- DTN-as-a-Service For Data Intensive Science With Scitags
- P4 Integration With Kubernetes, P4 Global Lab
- NASA Goddard Space Flight Center HP WAN Transport Services (400 G Dsk-Dsk)
- Resilient Distributed Processing & Rapid Data Transfer
- AutoGOLE/SENSE E2E Orchestration Net Services And Workflow Integration
- Open Science Grid Demonstrations
- National Research Platform Demonstrations
- Chameleon FABRIC/FAB Integration
- SciStream Multi Site Data Streaming Orchestration
- Distributed Pipelines Over WANs For On-Line Data Analysis
- DTNs for Research Enhanced Environments (ONION-RED ONION)
- Distributed Hybrid Quantum Computing With PQC/QKD Secured Links

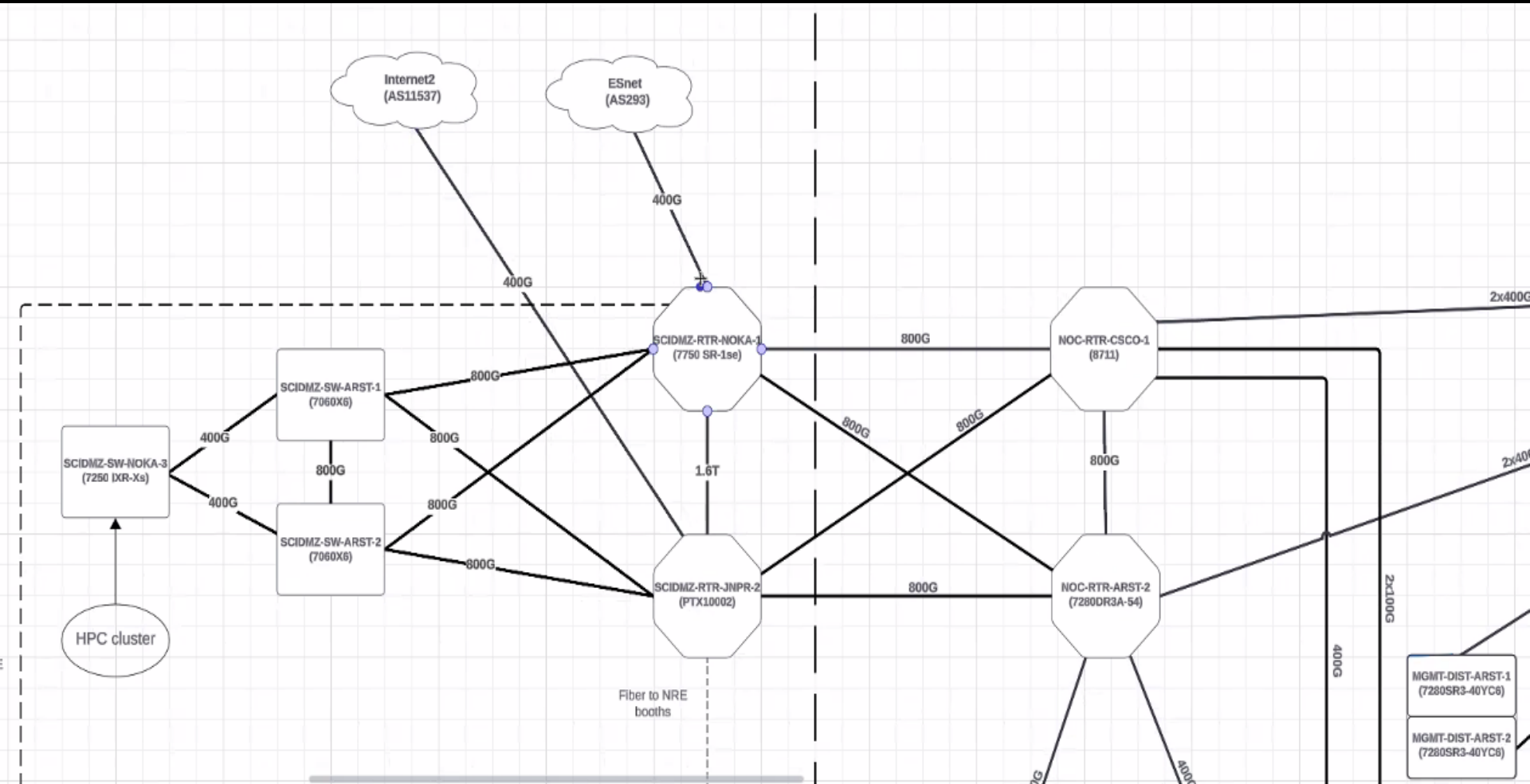
SC24 WAN-Transport Preliminary Map - August 9, 2024



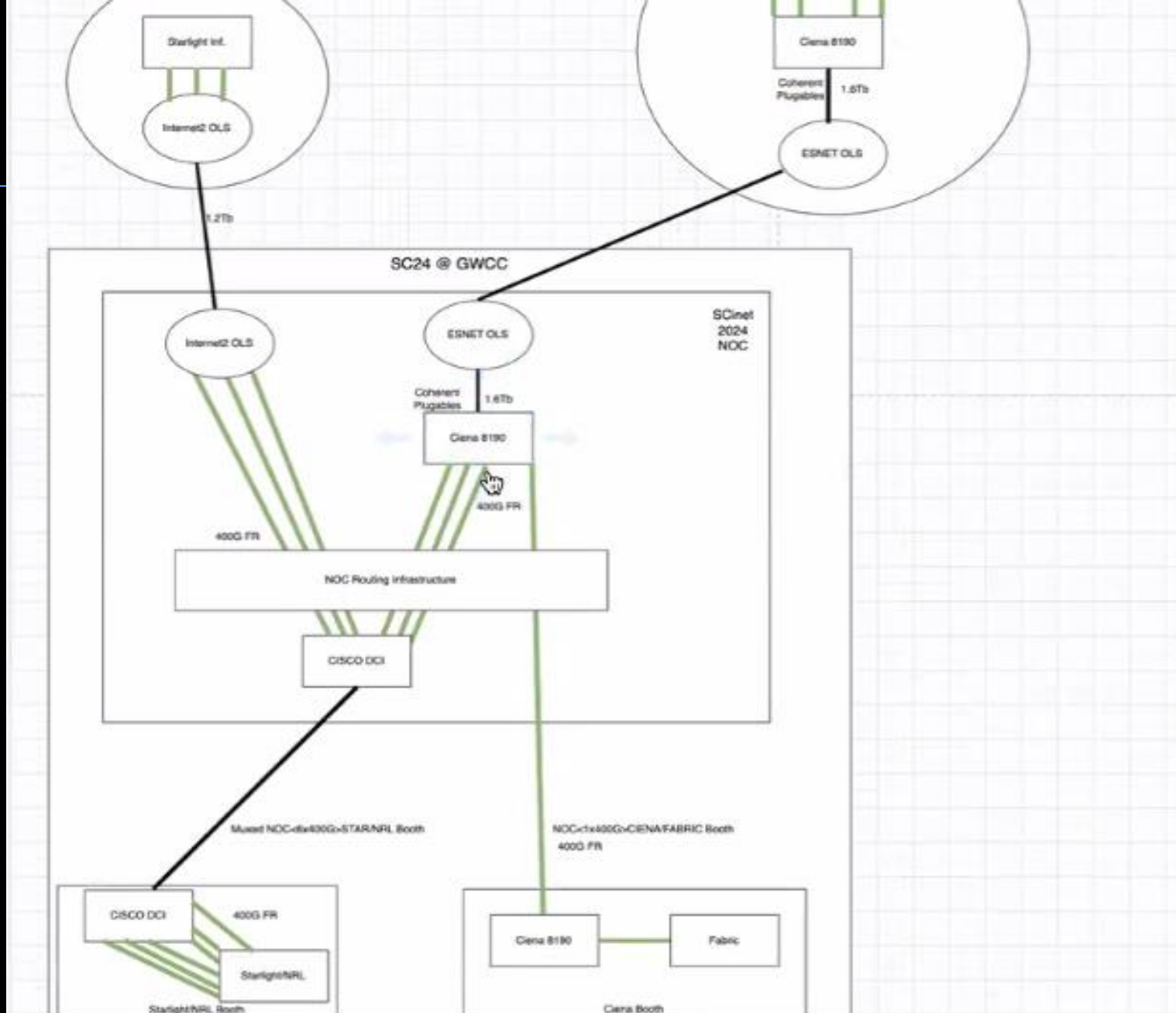
SC24 SCinet 2024



SCinet SC24 Venue 800 Gbps Science DMZ







SC24 Venue: 1.2 Tbps and 400 Gbps Paths **STARLIGHTSM**

Planned StarLight Booth SC24, Nov Atlanta

OPEN items for booth in callouts:

C WC05
D WC06
F WC07/WC16
E WC11
G WC16/17

CISCO DCI 400G

ARISTA 7280PR3 400G

7280SE

ARISTA 7050PX4 400G

NRL #?
Workstation
NRL #?
Workstation

NRL Gen5
(ASG-1115S)

NRL Gen5
(ASG-1115S)

NRL Gen5
(ASG-1115S)

NRL Gen5
(ASG-1115S)

StarLight Gen5
(R7615)

StarLight Gen5
(R7615)

StarLight Gen5
(R760)

NASA Gen5
(TBD)

NASA Gen5
(TBD)

ARISTA 7010T 1G

8/19/2024
Arista 7010TX MGMT
7280SE-48
NRL AS1115
NRL AS1115
NRL AS1115
NRL AS1115
7050PX4-32
7280CR3-24
SL-R7615
SL-R7615
SL-R760
NASA-TBD
NASA-TBD
24U rack

Please go here to see
VLAN info, etc

[JBDT team VLAN
spreadsheet](#)

[Tom's folder
Tom's diagram](#)

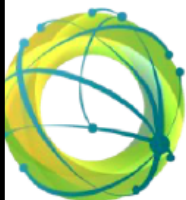
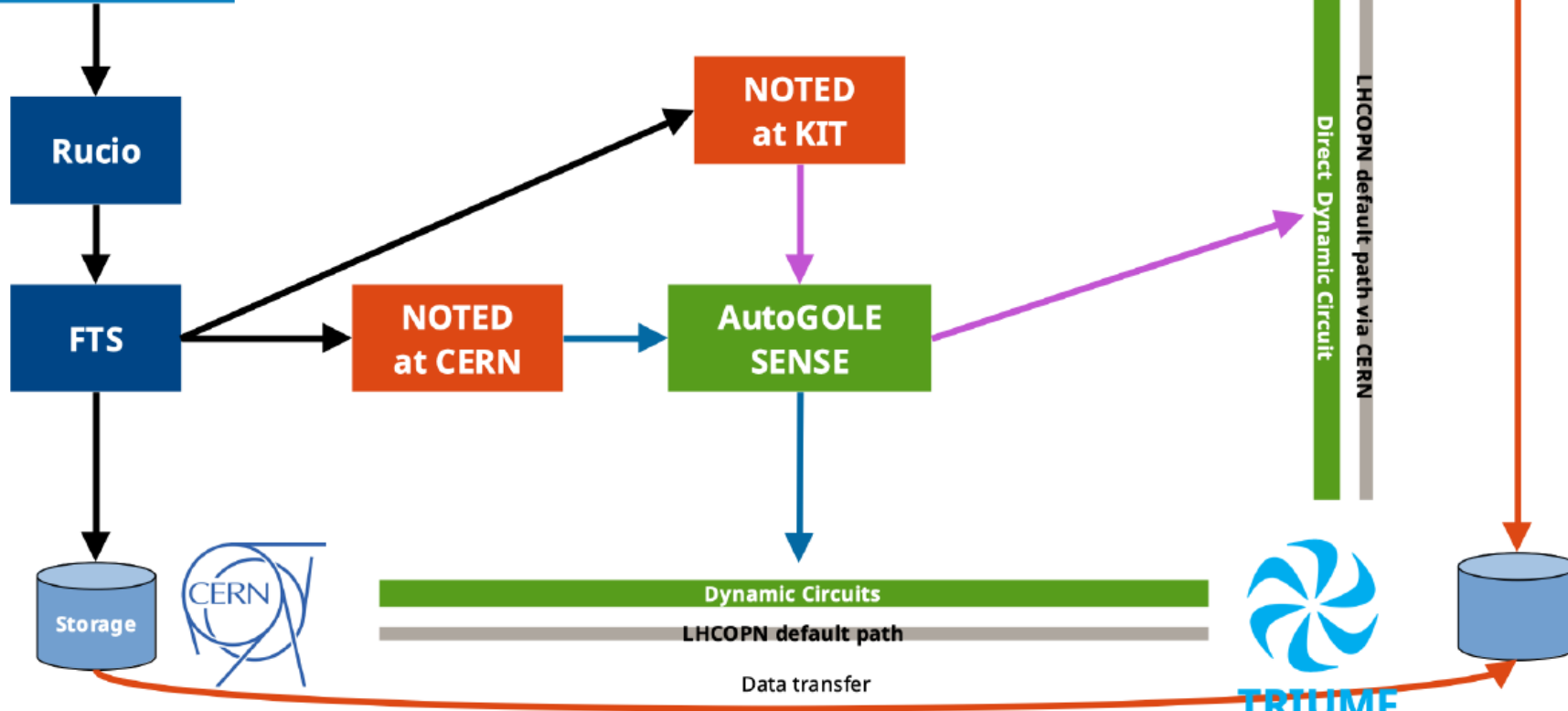
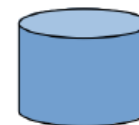
- 400G - LR4
- 400G - FR4
- 400G - DAC
- 200G - SR4
or DAC
- 100G - CLR4
- 100G - LR4
- 100G - SR4
- 100G - DAC
- 40G - SR4
- 40G - DAC
- 10G
- 1G

StarLight/
NASA/
NRL
Booth
2551
SC24,
Atlanta,
GA
[MORE](#)

08/19/2024

SC24-JBDT

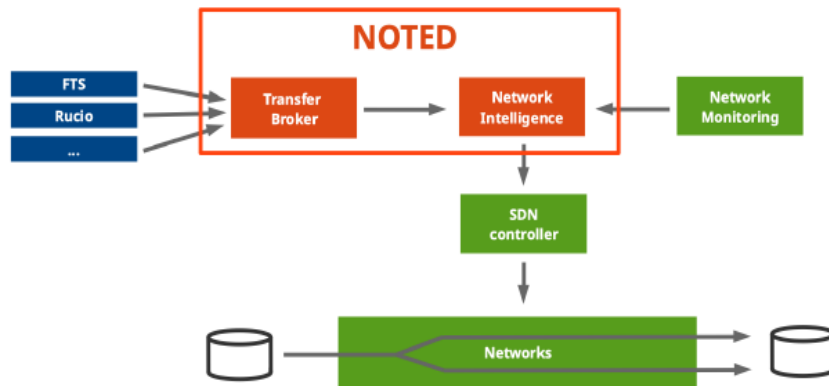
STARLIGHTSM



FTS
File Transfer Service



SKELETON AND ELEMENTS OF NOTED



FTS (File Transfer Service):

- Inspect and analyse data transfers to estimate if an action can be applied to optimise the network utilization → **get on-going** and **queued transfers**.

CRIC (Computing Resource Information Catalog):

- Enrichment to **get an overview** and knowledge of the **network topology** → get IPv4/IPv6 addresses, endpoints, rcsite and federation.

FLOWCHART AND DATASET STRUCTURE

- Input parameters: configuration given by the user
 - In noted/config/config.yaml → define a list of {src_rcsite, dst_rcsite}, maximum and minimum throughput threshold, SENSE/AutoGOLE VLANs UUID and user-defined email notification among others.
- Enrich NOTED with the topology of the network:
 - Query CRIC database → **get endpoints** that could be involved in the data transfers for the given {src_rcsite, dst_rcsite} pairs.
- Analyse on-going and upcoming data transfers:
 - Query FTS recursively → **get on-going** data transfers for each set of source and destination endpoints.
 - The total utilization of the network is the **sum** of on-going and upcoming **individual data transfers** for each source and destination endpoints for the given {src_rcsite, dst_rcsite} pairs.
- Network decision:
 - If NOTED interprets that the link **will be congested** → **provides a dynamic circuit** via SENSE/AutoGOLE.
 - If NOTED interprets that the link **will not be congested** anymore → **cancel the dynamic circuit** and the traffic is routed back.

scitags.org

Network Flow and Packet Marking for
Global Scientific Computing

View On
GitHub

Download
Tech. Spec

Join
scitags.org

Hosted on GitHub Pages — Theme by [orderedlist](#)

Scientific network tags (scitags) is an initiative promoting identification of the science domains and their high-level activities at the network level.

It provides an open system using open source technologies that helps *Research and Education (R&E) providers* in understanding how their networks are being utilised while at the same time providing feedback to the *scientific community* on what network flows and patterns are critical for their computing.

Our approach is based on a network tagging mechanism that marks network packets and/or network flows using the science domain and activity fields. These tags can then be captured by the *R&E providers* and correlated with their existing netflow data to better understand existing network patterns, estimate network usage and track activities.

The initiative offers an **open collaboration on the research and development of the packet and flow marking prototypes** and works in close collaboration with the scientific storage and transfer providers to enable the marking capability. The project is currently in the prototyping phase and is open for participation from any science domain that require or anticipate to require high throughput computing as well as any interested *R&E providers*.

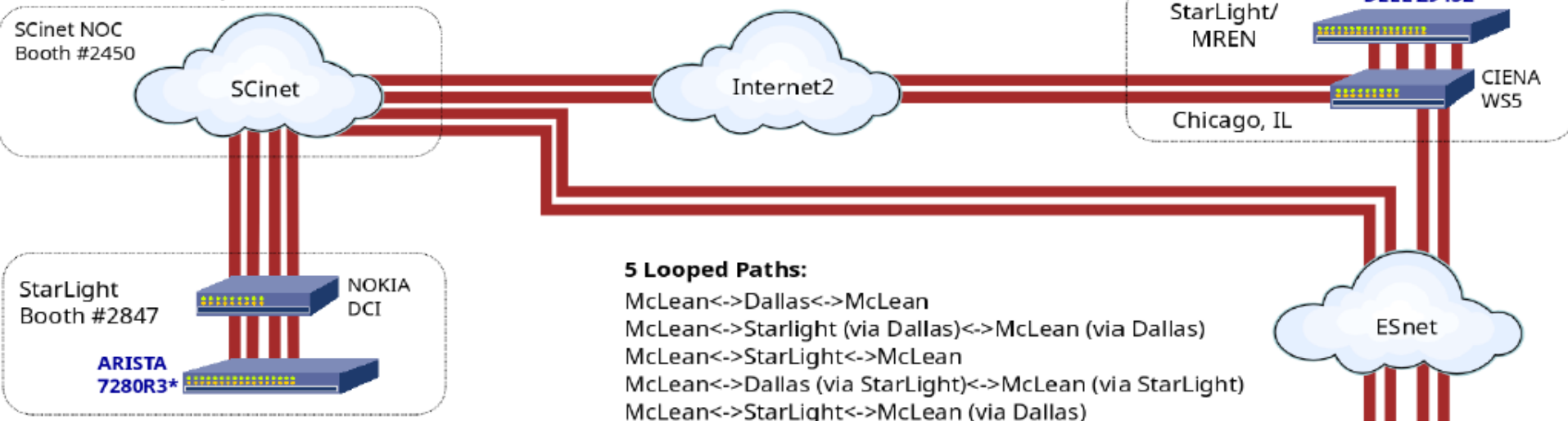
SC22

Joint Big Data Testbed

Demonstrations of 400 Gbps Disk-to-Disk WAN File Transfers using NVMe-oF/TCP

An SC22 Collaborative Initiative Among NASA and Several Partners

SC22 @ Dallas, TX



5 Looped Paths:

McLean<->Dallas<->McLean
 McLean<->Starlight (via Dallas)<->McLean (via Dallas)
 McLean<->StarLight<->McLean
 McLean<->Dallas (via StarLight)<->McLean (via StarLight)
 McLean<->StarLight<->McLean (via Dallas)

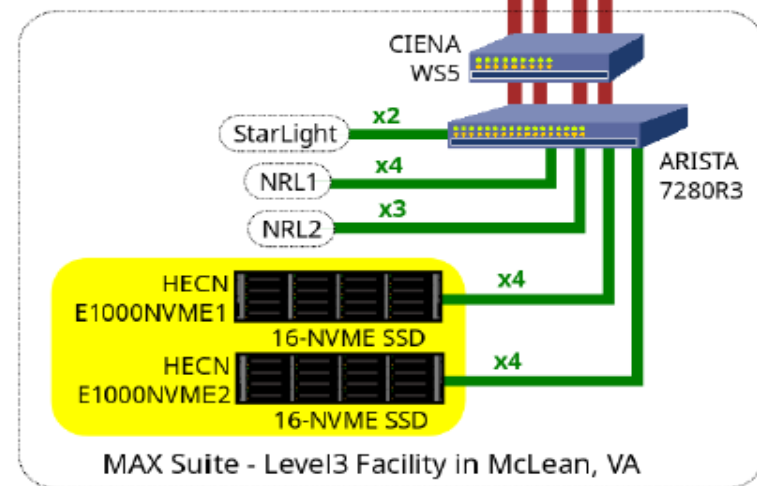
R&D Partners



Legend

400-GigE
 200-GigE
 NASA/GSFC-owned

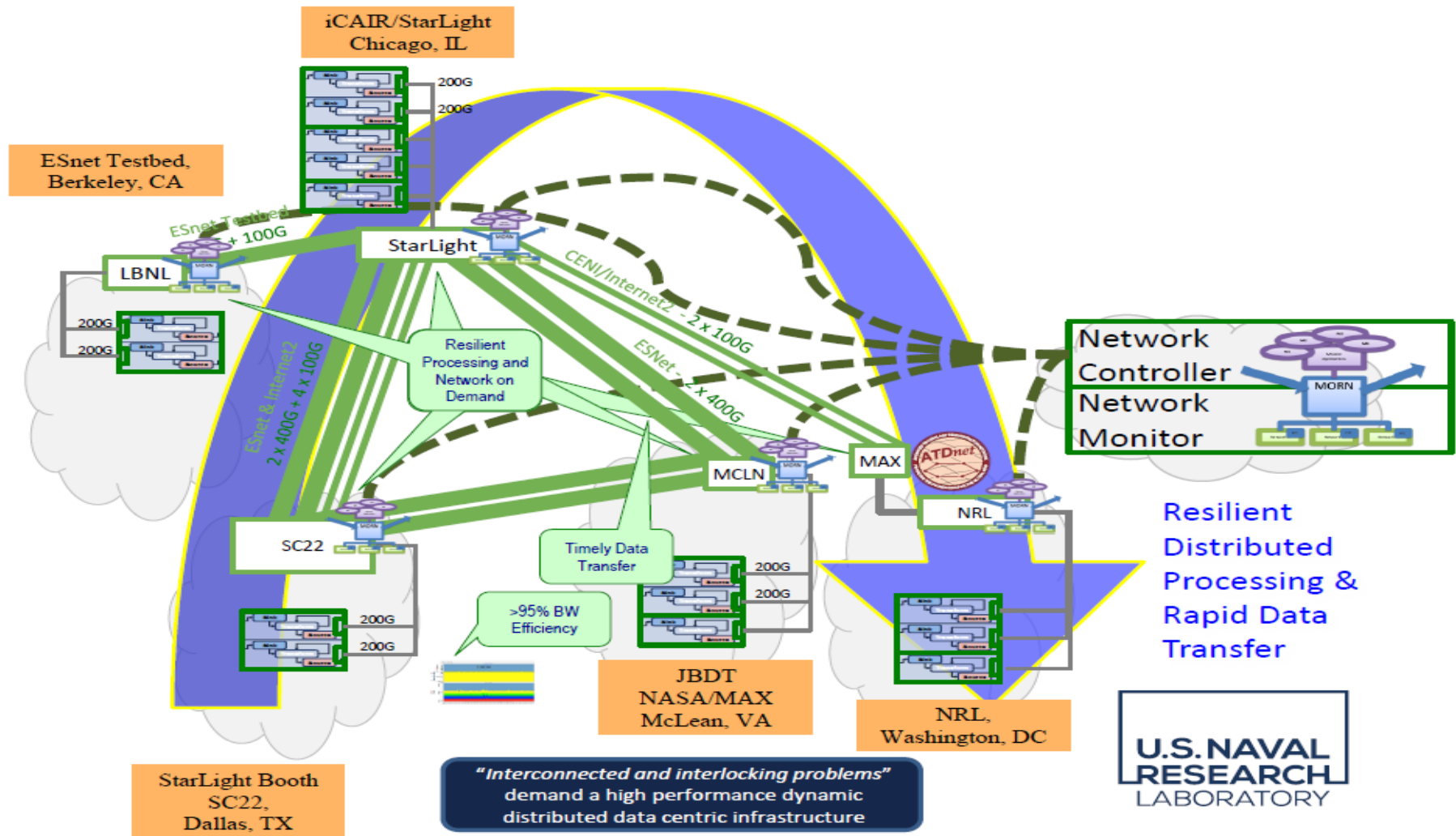
*IP Routed Loopback Point
 Back To MAX Suite - Level3
 McLean, VA



In memory of Paul Lang and Pat Gary

NASA/GSFC High End Computer Networking (HECN) Team
 Diagram by Bill Fink - 10/20/2022

Resilient Distributed Processing & Rapid Data Transfer

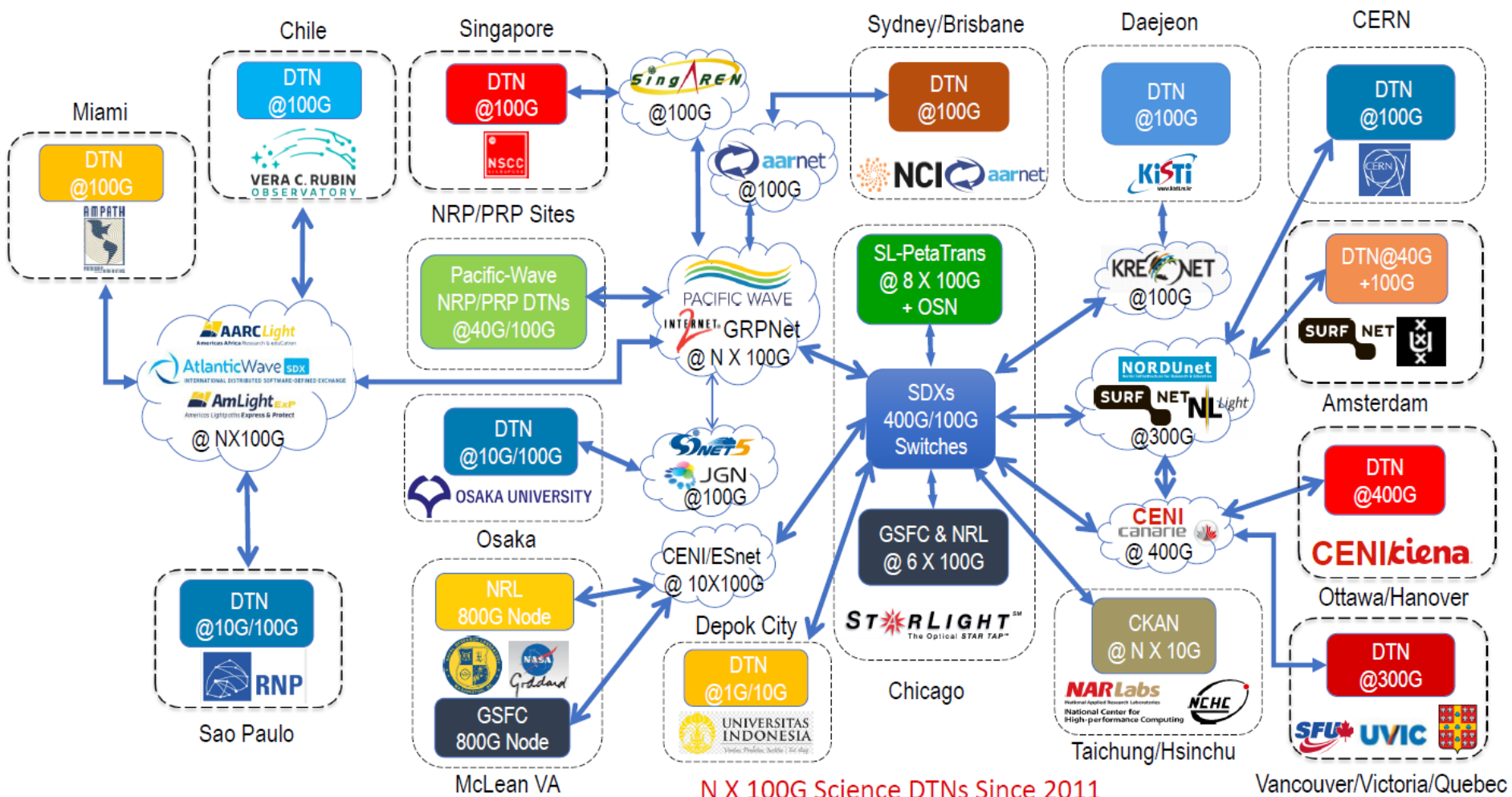


DISTRIBUTION A: Approved for Public Release

STARLIGHT

GRP DTNaaS For Petascale Science

GRP Service: DTNaaS for Petascale Sciences Data Movement



N X 100G Science DTNs Since 2011

DTN-as-a-Service – Demonstrated At SC22

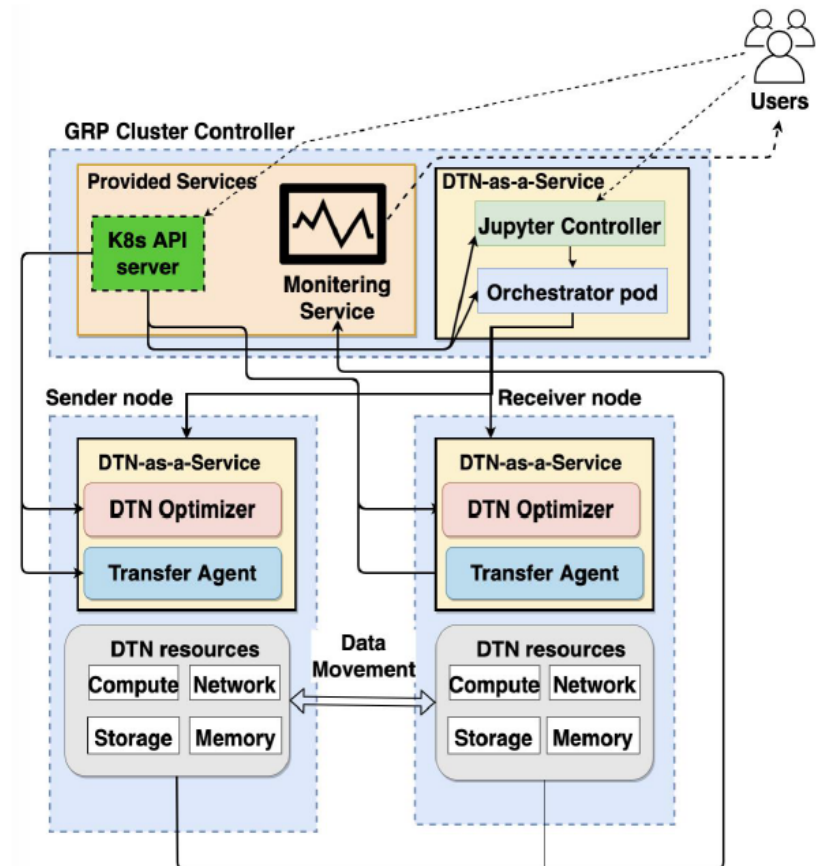
GRP Cluster with DTN-as-a-Service

DTN-as-a-Service(DTNaaS) provides a data movement workflow in GRP k8s cluster:

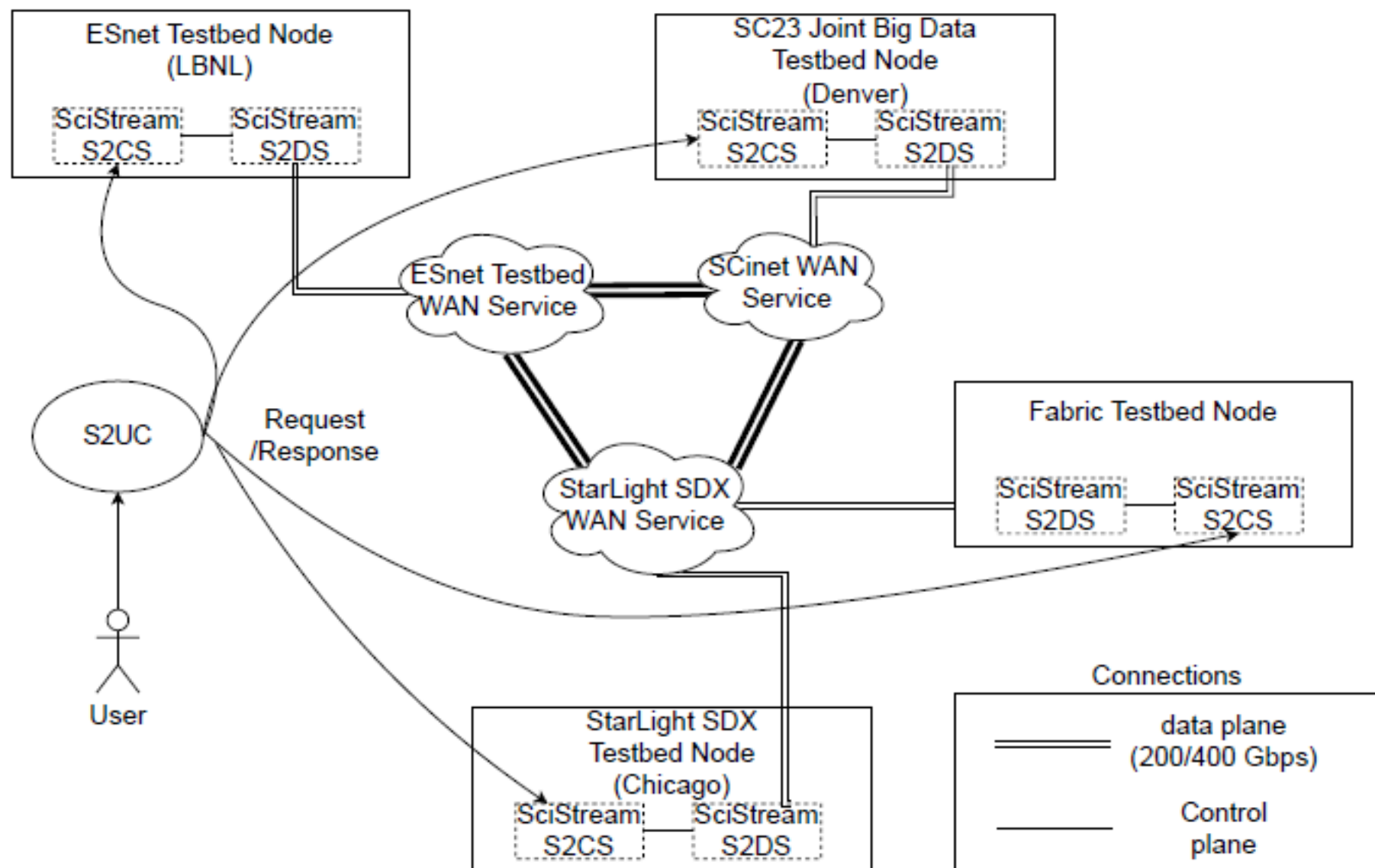
1. Deploy DTNaaS workloads via k8s API server
2. Use Jupyter to optimize and run transfers
3. Observe performance from monitoring service

GRP DTNaaS Components:

- Orchestrator: controller of DTNaaS to manage agent and optimizer pods via REST API.
- Transfer Agent: run transfer jobs
- DTN Optimizer: optimize the DTN resources for workflow
- Jupyter: web interface to run DTNaaS interactively

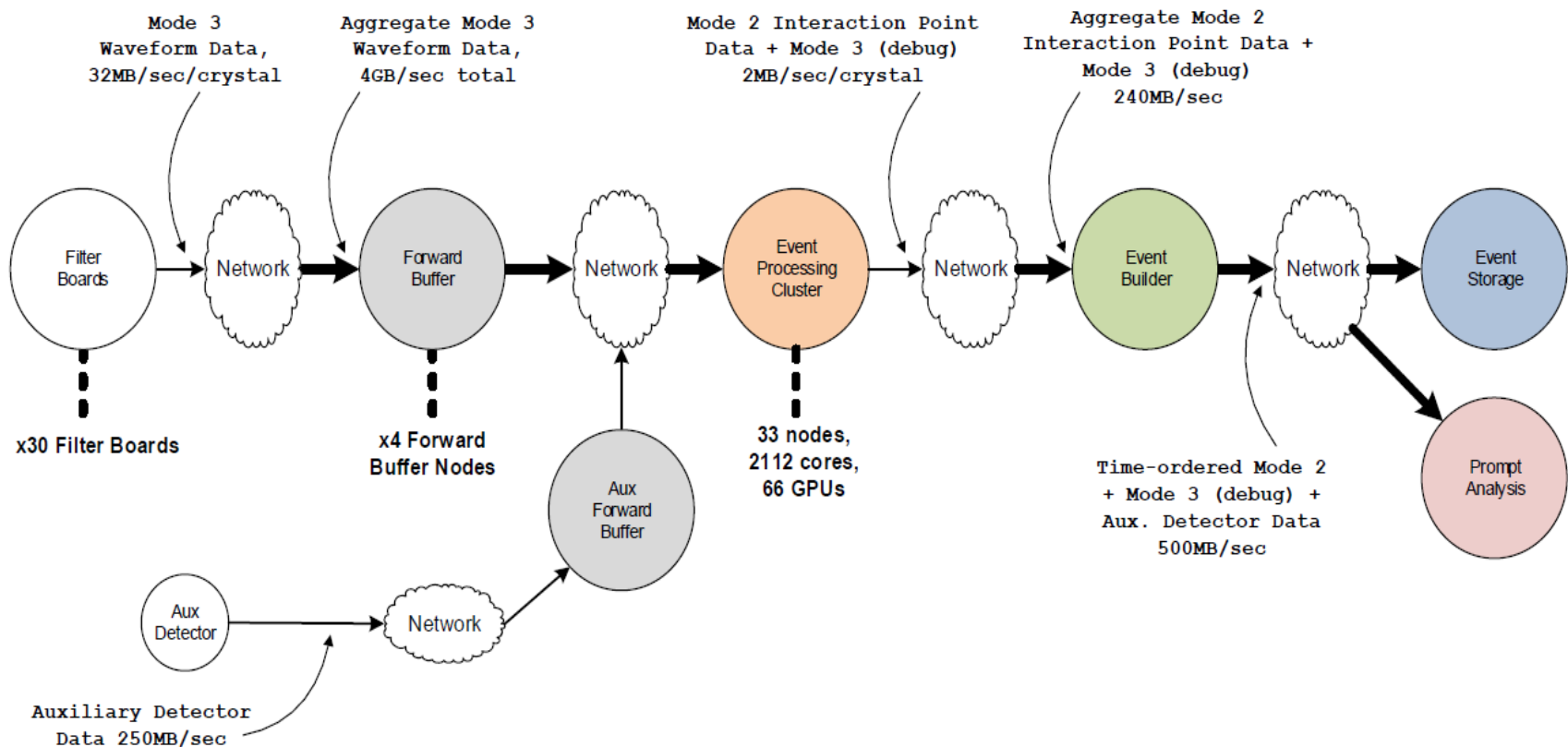


SC23 NRE-010: Multi-site data streaming orchestration with SciStream

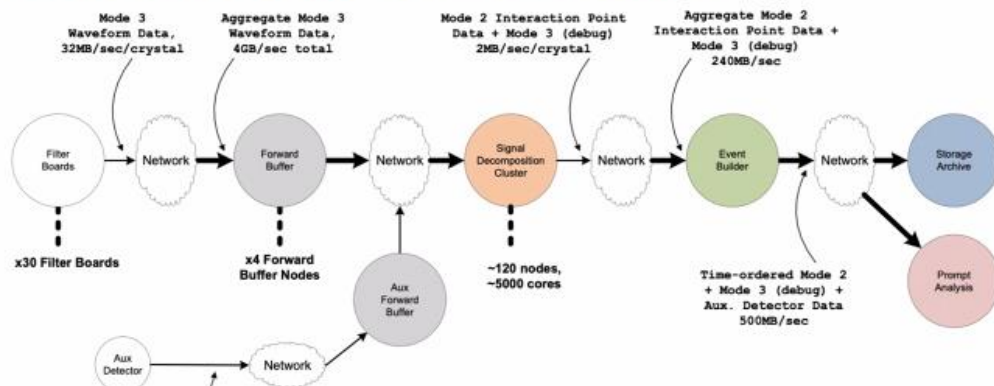


ESnet Gamma Ray Energy Tracking Array (GRETA) Prototype

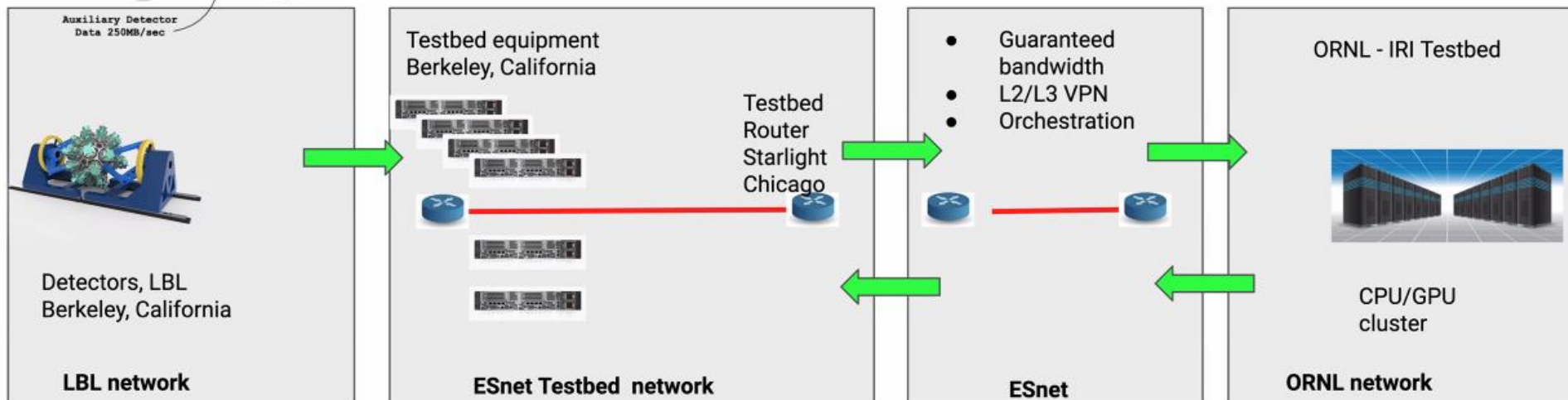
GRETA Data Pipeline



GRETA to DELERIA



Impact: scaling up
online processing using
user facilities



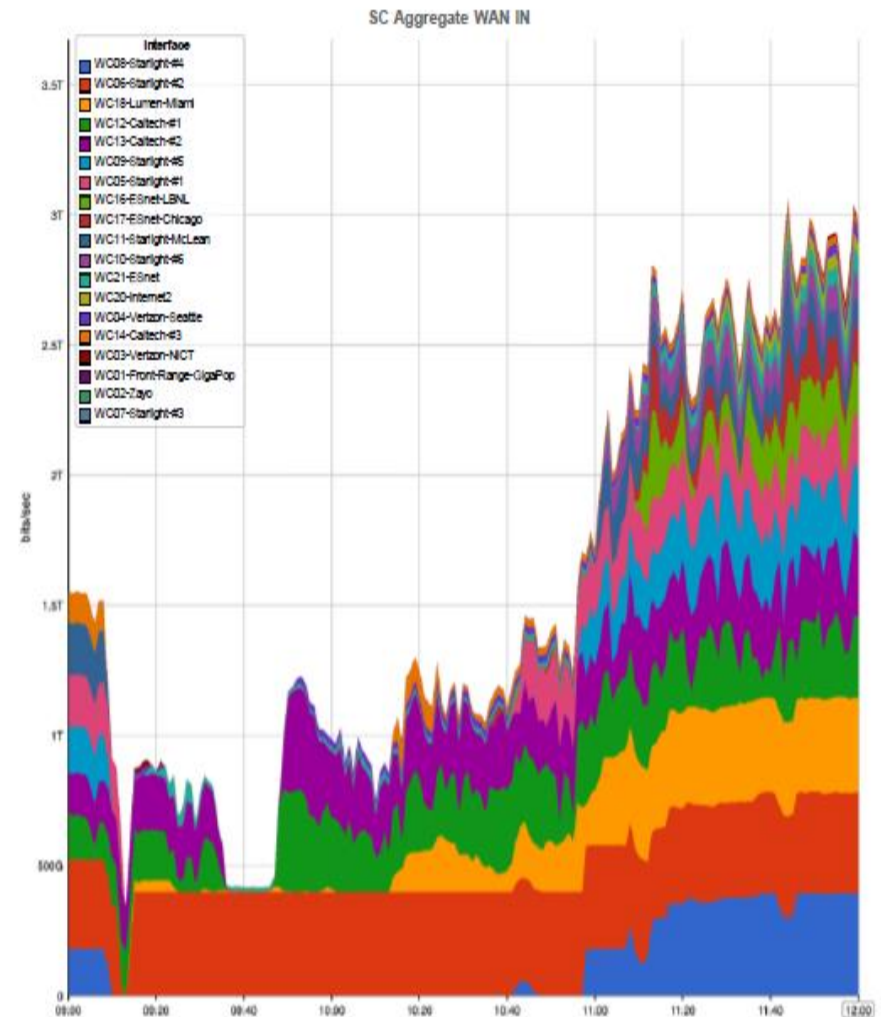
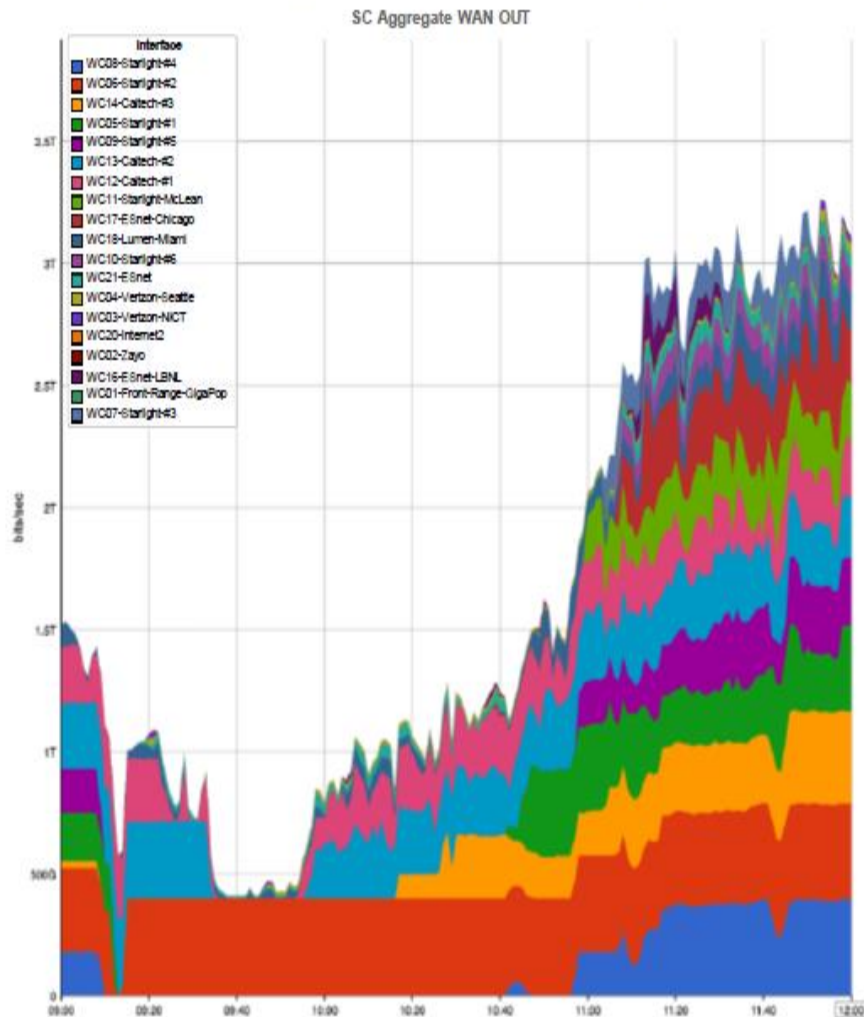
Source: ESnet Testbed

STARLIGHTSM

SC23 Bandwidth Challenge

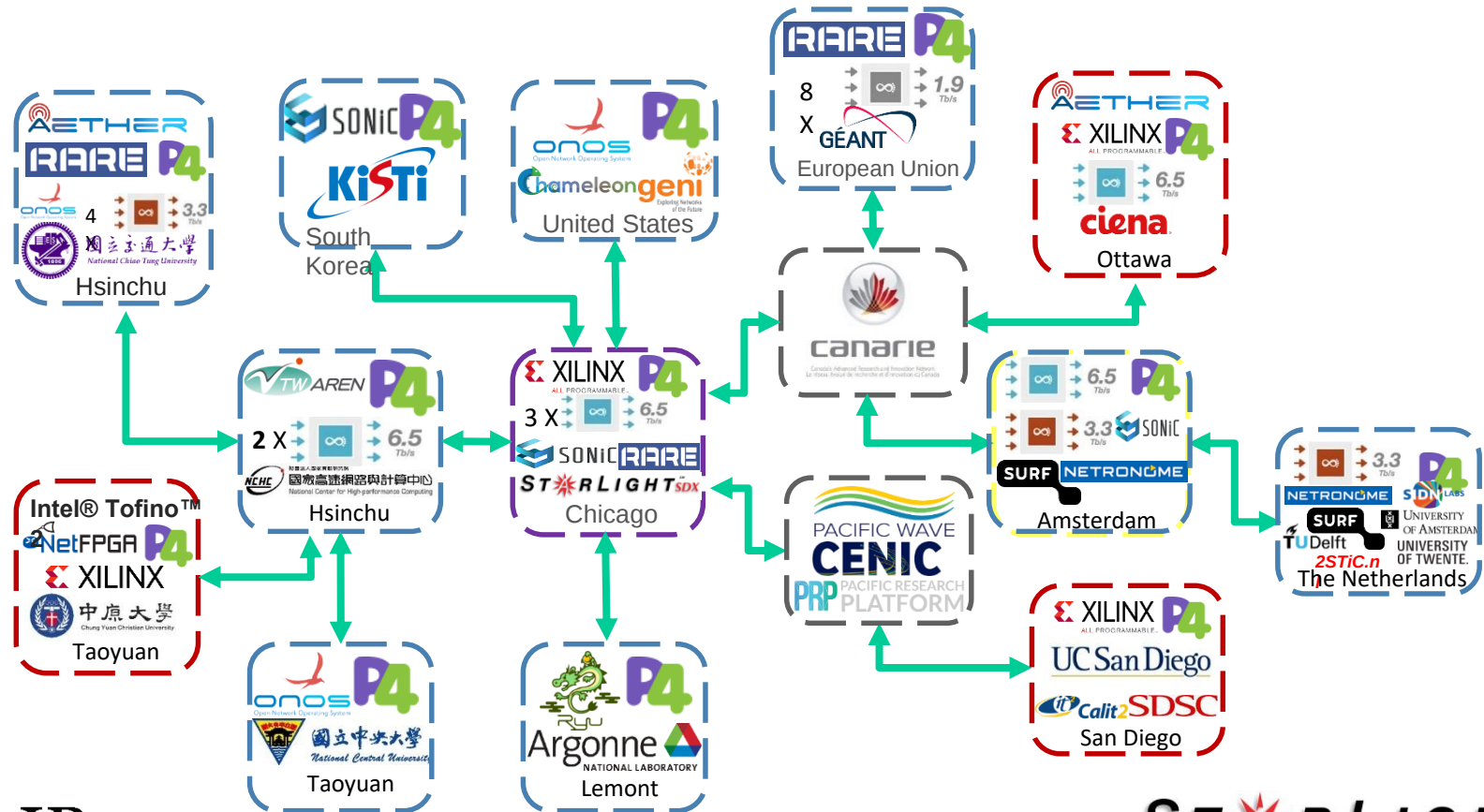
StarLight contributes 4 of Top 5

StarLight contributes 2 of Top 5



International P4 Testbed Showcase at SC23

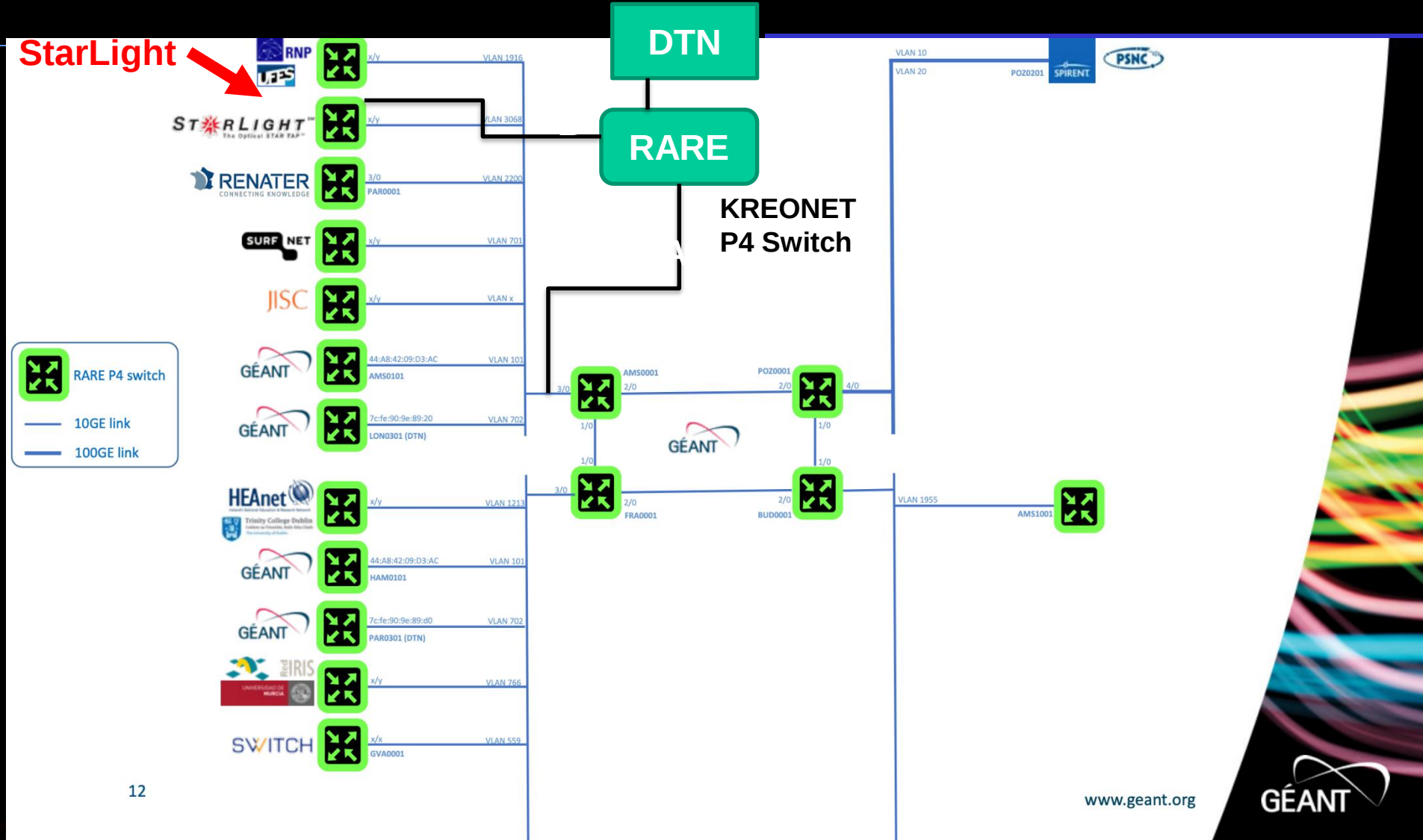
GRP Service: International P4 Experimental Networks (iP4EN)



iCAIR

STARLIGHTSM SDX

Integration With GEANT P4 Testbed





www.chameleoncloud.org

CHAMELEON: A LARGE SCALE, RECONFIGURABLE EXPERIMENTAL INSTRUMENT FOR COMPUTER SCIENCE

Kate Keahey

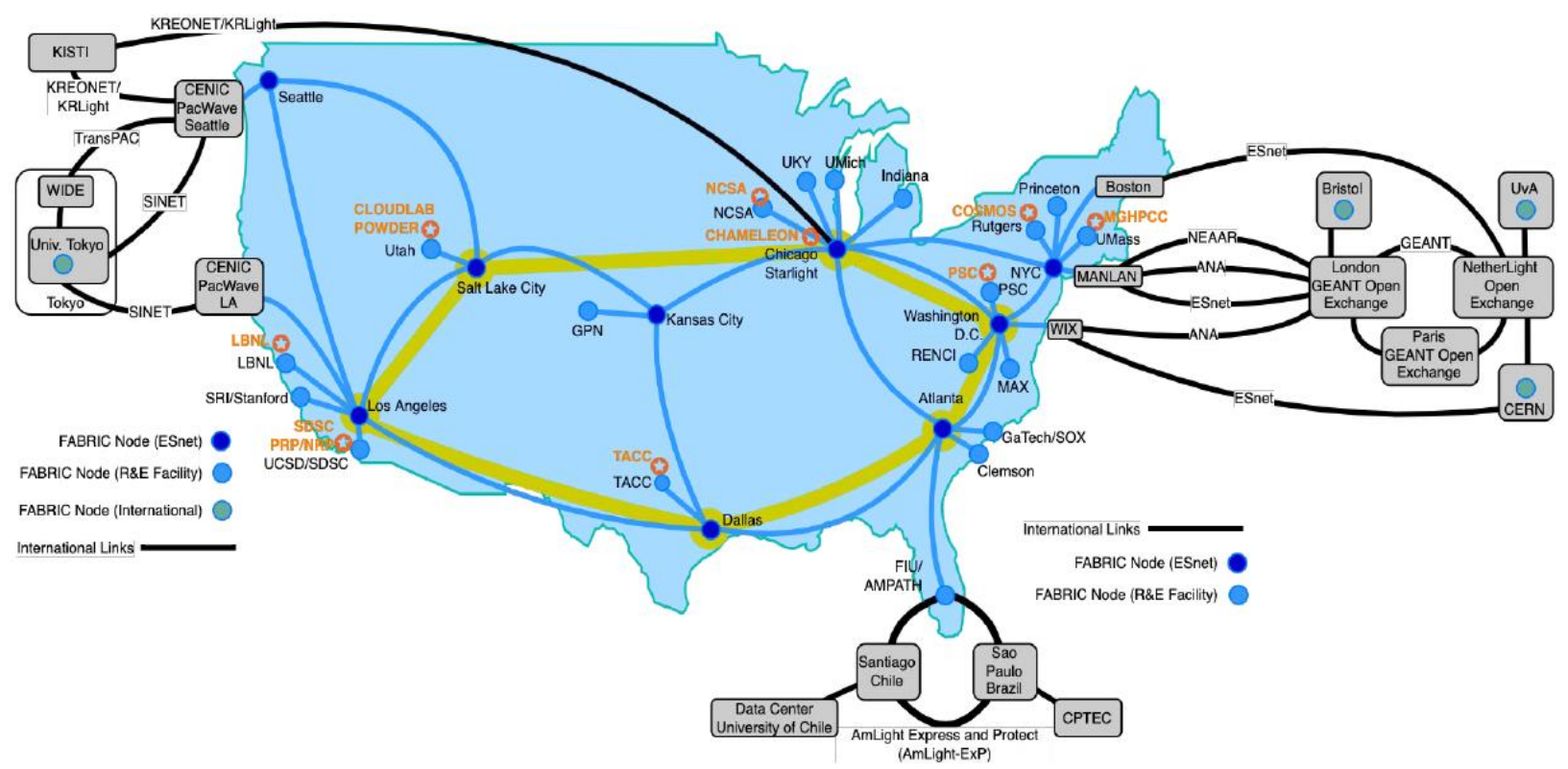
Joe Mambretti, Pierre Riteau, Paul Ruth, Dan Stanzione

SEPTEMBER 28, 2017

1

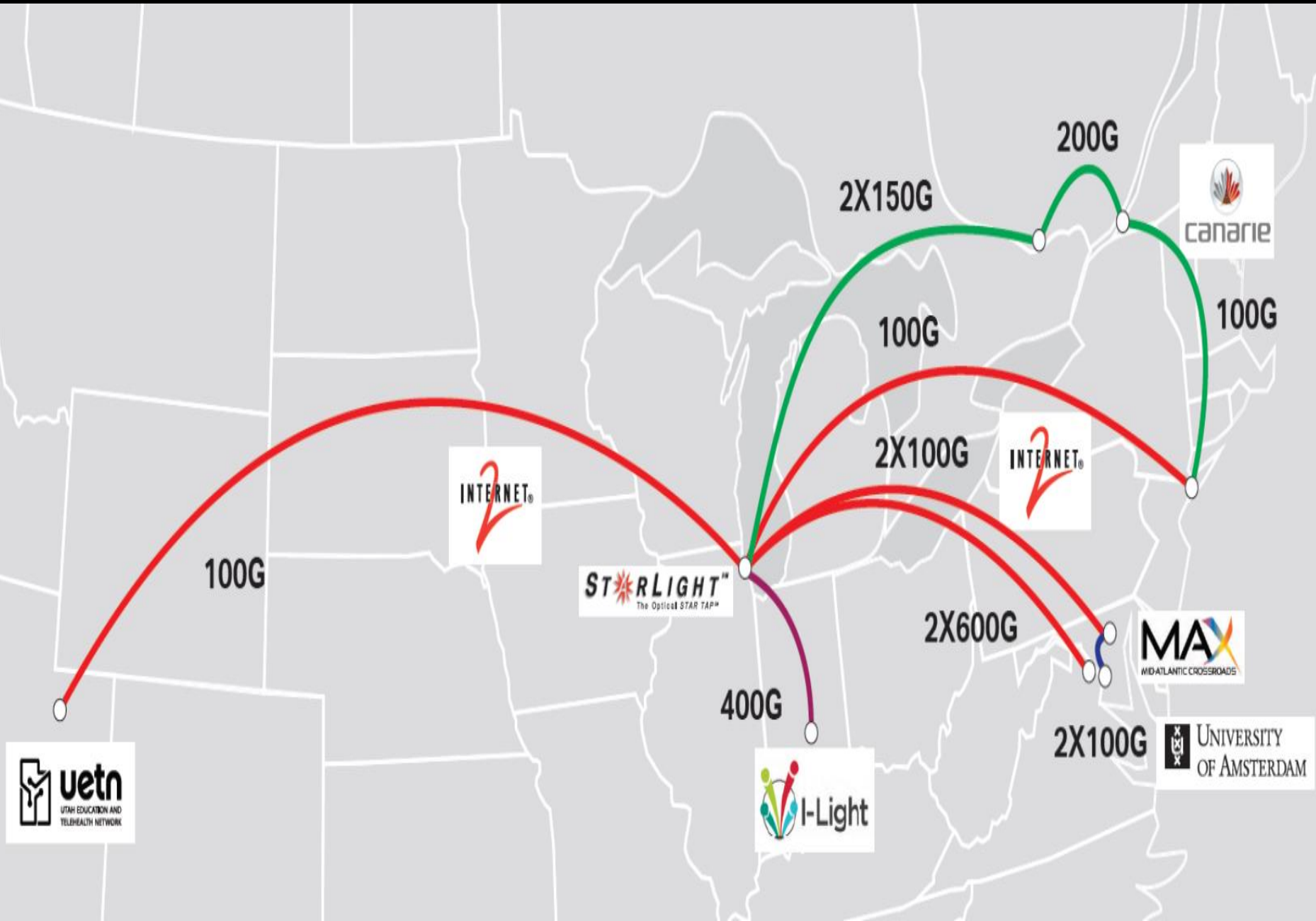


FABRIC Testbed (+FAB)



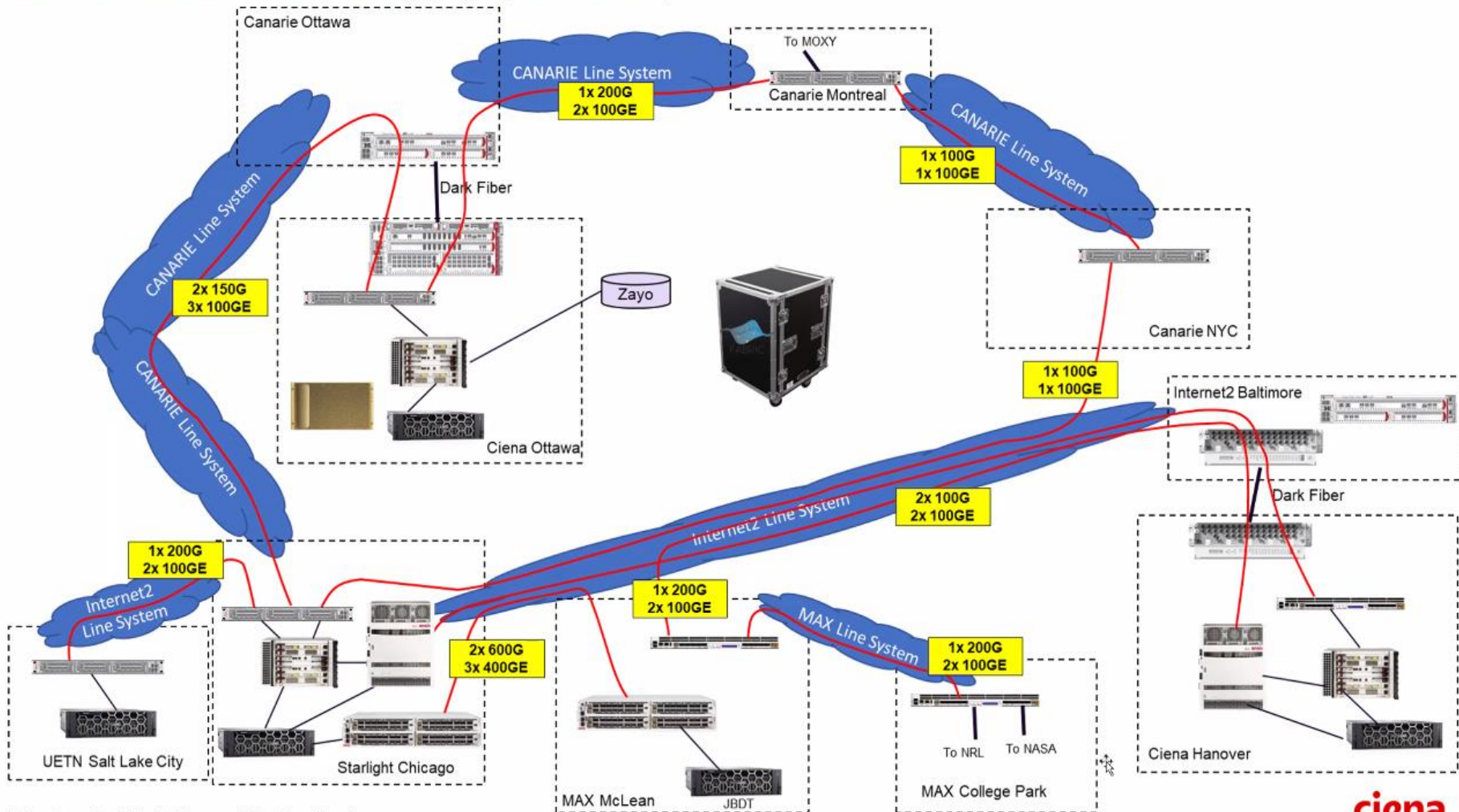
FABRIC Topology - with FAB Sites

Ciena CENI



Ciena CENI

Current CENI Topology (April 2024)



© Ciena Corporation 2024. All rights reserved. Proprietary Information.

ciena

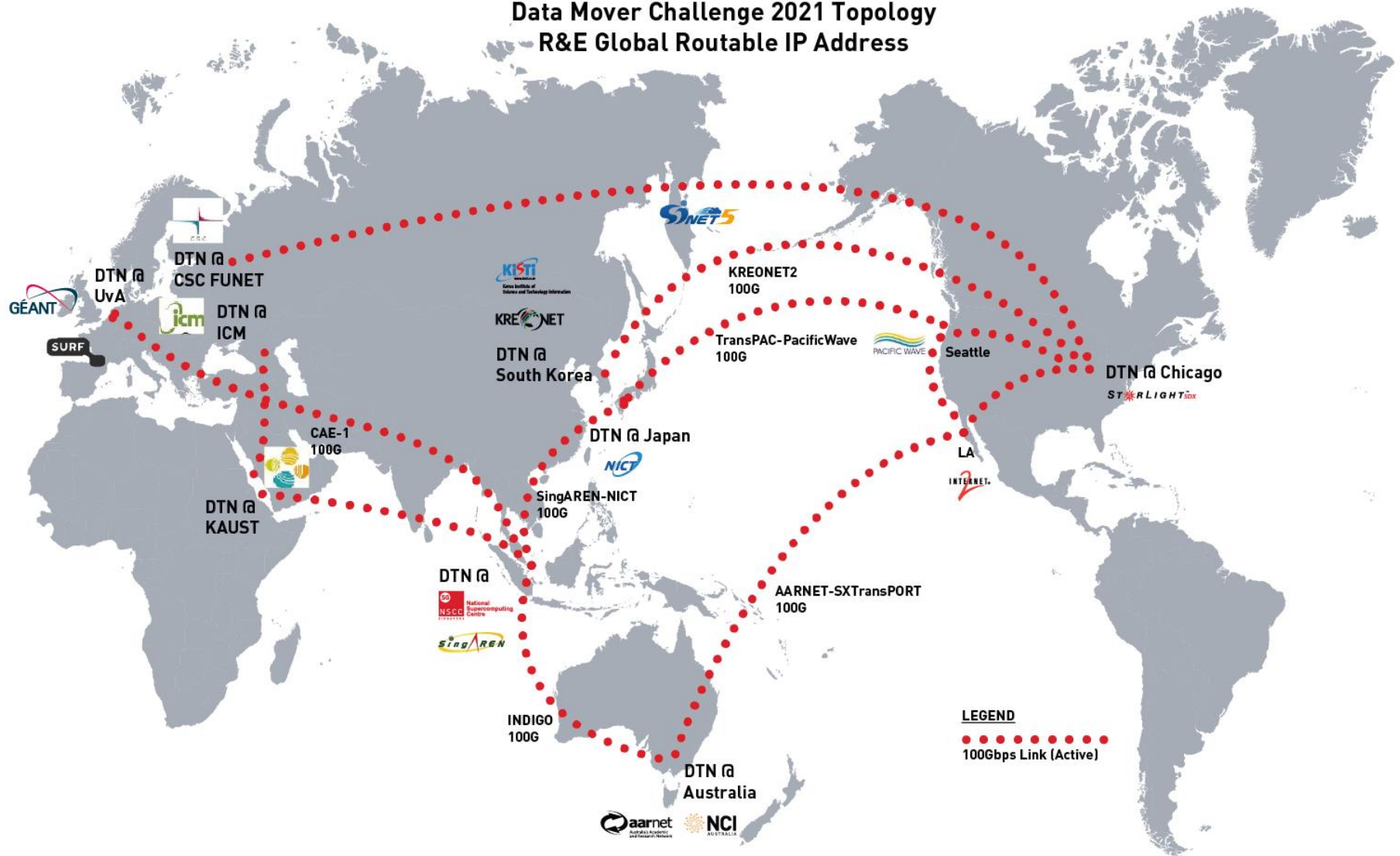
STARLIGHTSM

FabFed

FABRIC Federation Software Extension (FabFed)

- Goals
 - Provide a software stack for FABRIC users to run experiment with **tool-based federation** across **multiple testbeds and providers**
 - Integrate external network and cloud resources into a FABRIC slice in an **automated and orchestrated** fashion
 - Explore a “**testbed of testbeds**” paradigm for the future
- Targets
 - NSF Cloud Testbeds (Chameleon & CloudLab)
 - Internet2 AL2S L2, L3 and Cloud Connect Services
 - Public Cloud (AWS, GCP, Azure)
 - AutoGOLE/SENSE Testbed (most regional and international RENS)
 - Other NSF Testbeds such as POWDER, and more

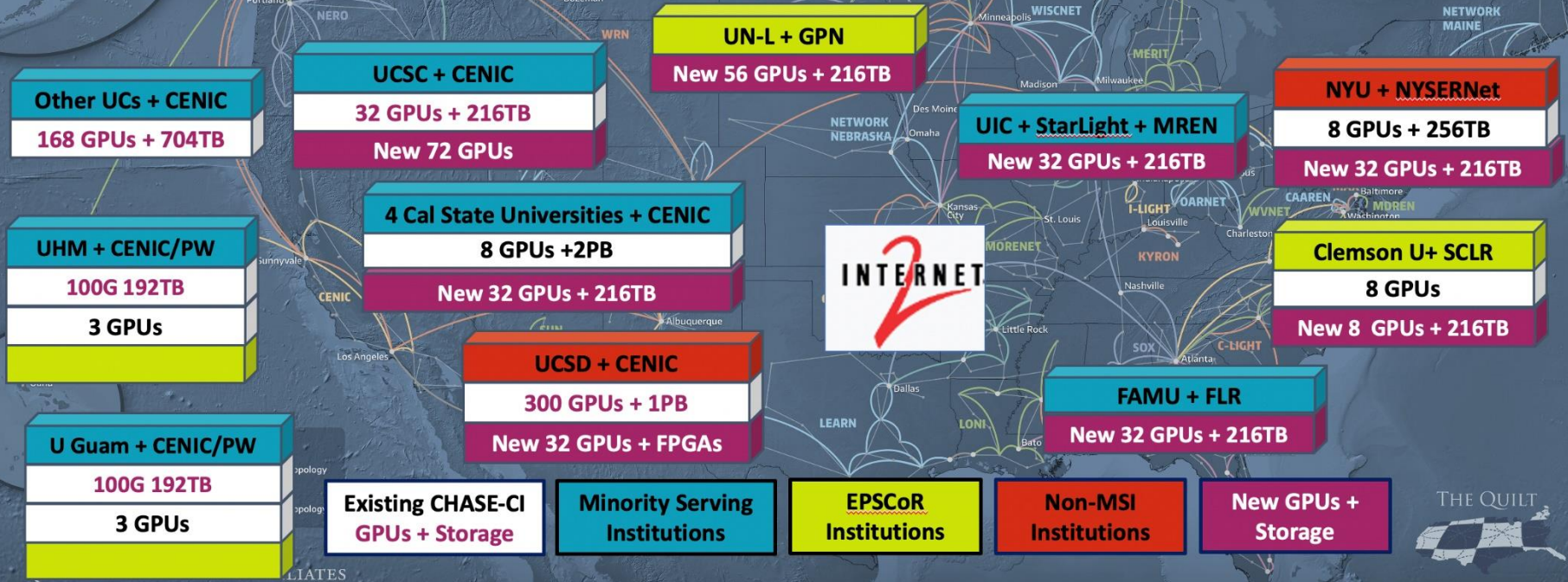
Data Mover Challenge 2021 Topology



National Research Platform

REGIONAL RESEARCH AND EDUCATION NETWORKS IN THE UNITED STATES

Proposed Extension of Nautilus 2021-2024



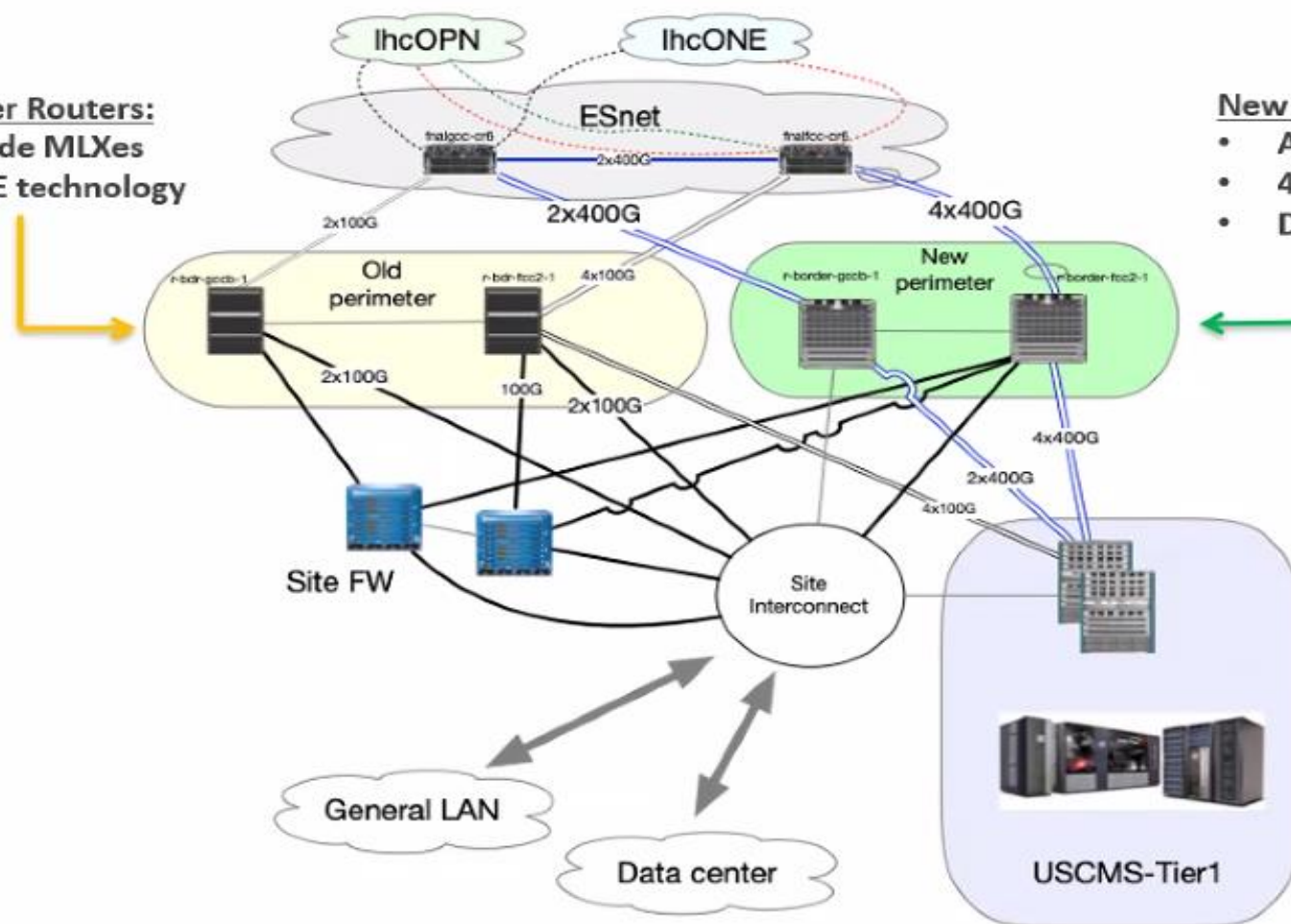
+ Open Science Grid & Open Science Data Grid

STARLIGHTSM

WAN Connectivity Transition! (to 2.4Tb/s...)

Old Border Routers:

- Brocade MLXes
- 100GE technology



New Border Routers:

- Arista 7808R3s
- 400-800GE technology
- Deep buffers

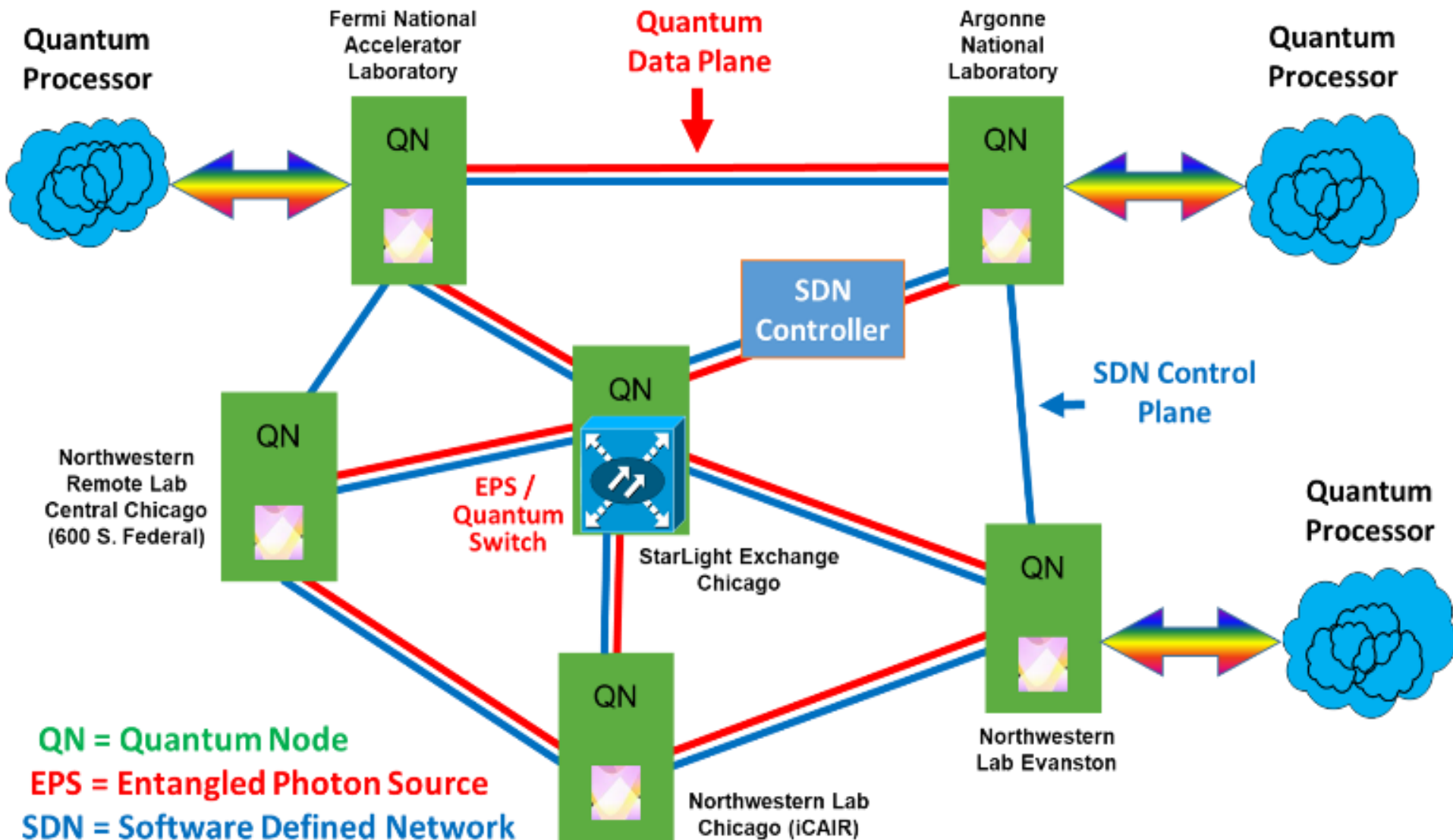
Also: Large Scale Network Research Testbed On Site

Quantum Communications And Networks: Motivation

- **Quantum Enables Many New Applications**
 - Security – e.g., Quantum Key Distribution (QKD), Highly Secure Information Transmission, Quantum Encryption
 - Quantum Sensors
 - Quantum – e.g., Precise Clocks
 - New Applications Derived From Unique Properties (e.g., Superposition) And Novel Quantum Devices
 - Communications Among Quantum Computers, e.g., To Address Complex Computational Science Problems Through Distributed Quantum Environments (iCAIR's Quantum Research Focus)



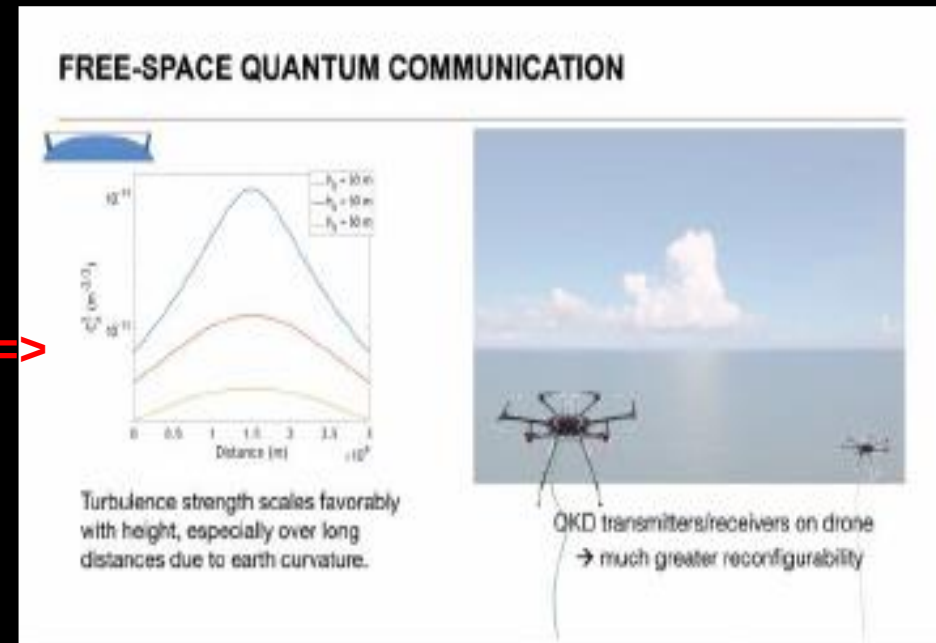
Emerging Chicago Quantum Exchange Testbed



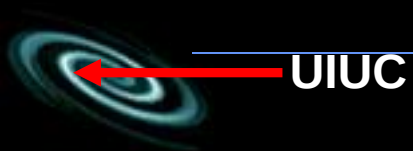
The map shows the Chicago metropolitan area with several red pins and lines indicating connections. A red line connects a pin at Northwestern University (Evanston) to a pin at StarLight/iCAIR (Chicago). Another red line connects a pin at Fermilab (Batavia) to a pin at Argonne National Laboratory (Lisle). A third red line connects a pin at StarLight/iCAIR to a pin at Argonne National Laboratory. Two inset maps are included: one in the top right showing a closer view of the Northwestern and StarLight/iCAIR area, and another in the bottom left showing a closer view of the StarLight/iCAIR and Argonne National Laboratory area. The map is labeled with various city names and highway numbers.

IQUIST Quantum Network Testbed: QUIUC-NET

(Hyper)Entangled Sources
Photon Detectors
Quantum Memories
Processing Nodes
Net Aps
Protocols
Distributed Processing
Sensing Net Verification
Repeater Enhanced Quantum Links
Free Space Quantum Communications=>



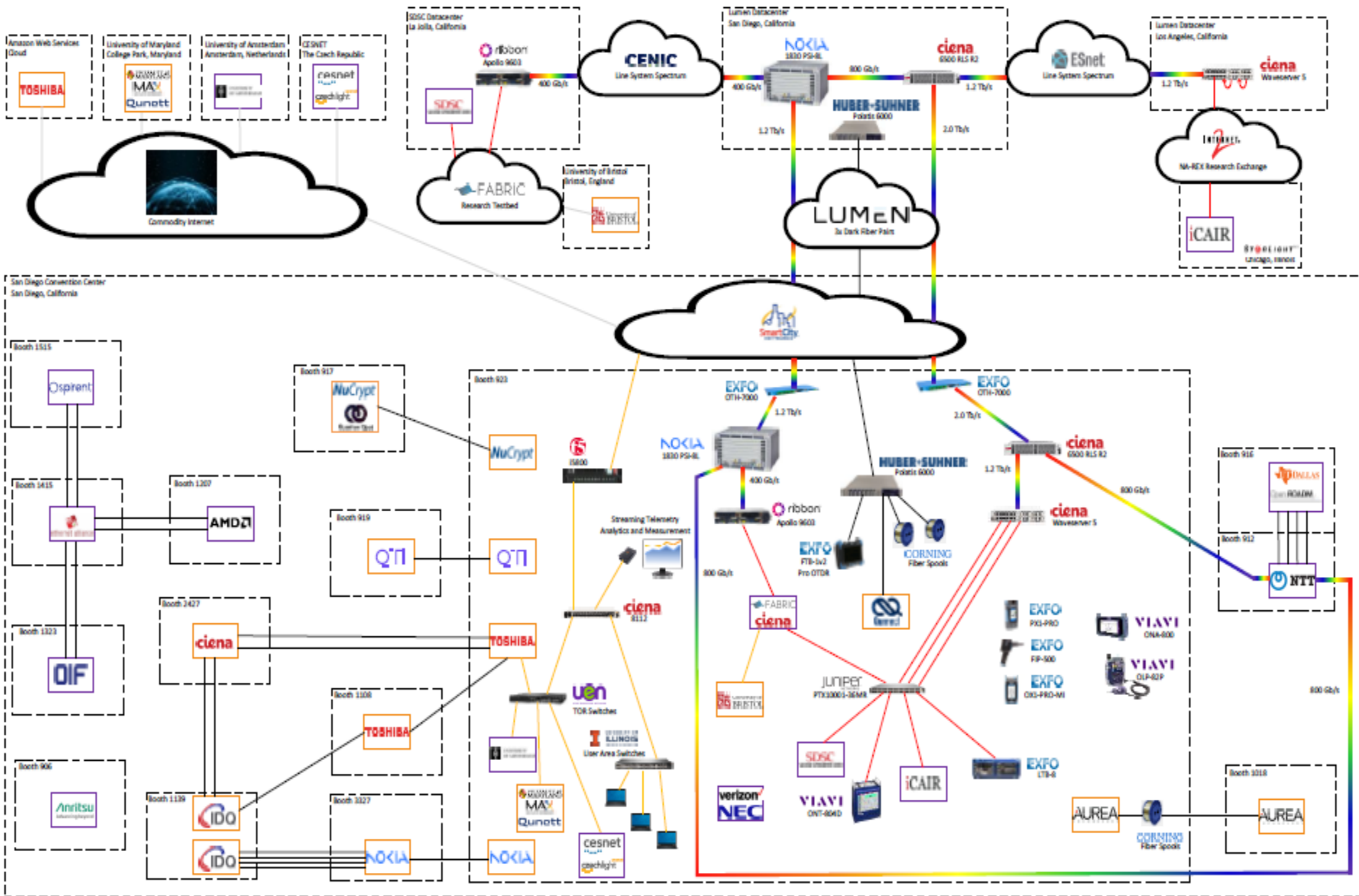
Source: Paul Kwiat, Director,
IQUIST



ECOC Teleportation Paper

- ***Quantum Teleportation Over Optical Fibers Carrying Conventional Classical Communications Traffic***
- **Jordan M. Thomas, Gregory S. Kanter, Fei I. Yeh, Jim Hao Chen, Joe J. Mambretti, Scott J. Kohlert, Prem Kumar**





OFC

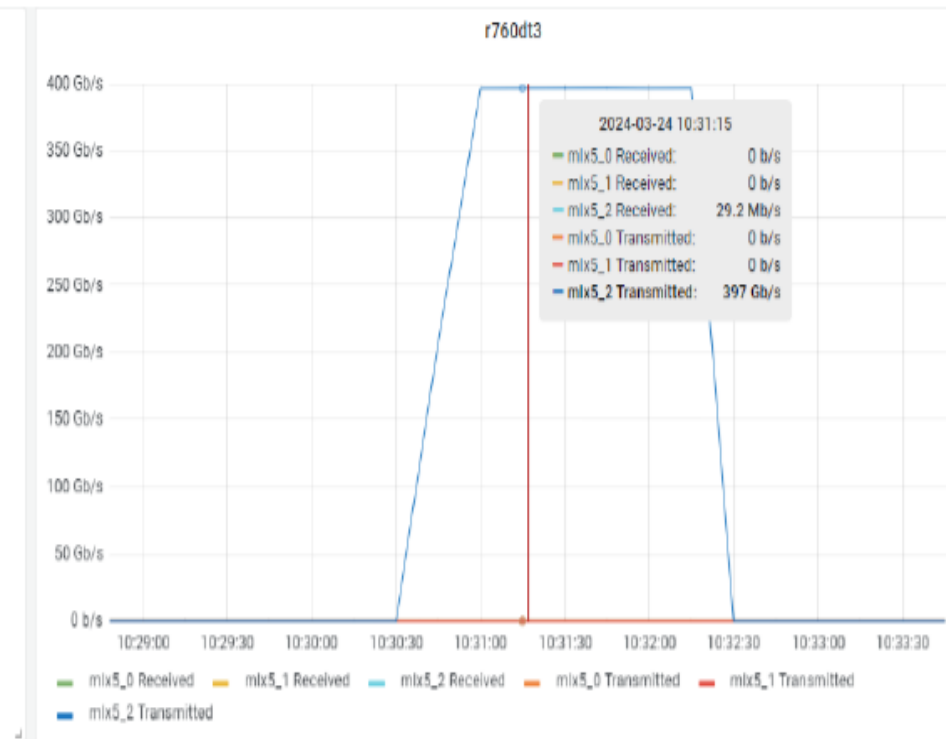
OFC 2024 - OFCnet Architecture





Extending Data Center Services Over 400G WAN

Prototype Solution Initial Results: Single Stream RDMA/RoCE Over 400G Network



StarLight-San Diego OFCnet loopback: Rtt 96.4 ms, Peak @ 397G

Exploring OFCnet-land

397 Gbps

OFCnet WAN Overview

OFCnet Traffic Last 1 minute



FABRIC



SDSC



Viavi



EXFO



ICAIR

EMC Input Optics



XCVR-1-1



XCVR-1-2

OFCnet Traffic rate

Bitrate

1.58 Tb/s

Simultaneous Phone calls

24.7 Mil

DVDs/sec

42.1

Library of Congress (digital)/sec

0.00267

OFCnet Data transferred Last 1 hour

Bytes

746 TB

Total Hour-long calls

25.9 Mil

DVDs

159 K

Library of Congress (digital)

10.1

OFCnet Management Traffic



OFCnet Kiosk Traffic



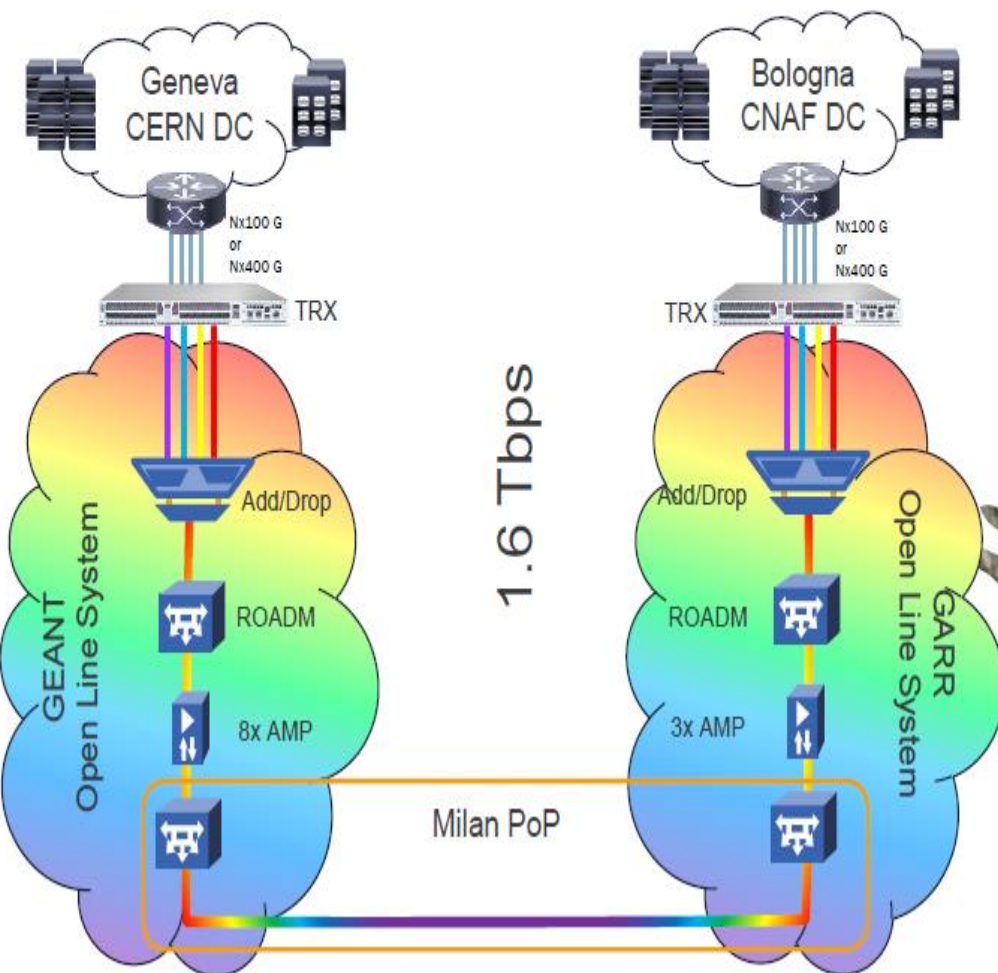
OFC 2025 San Francisco

- **OFC 50 – Celebrating 50 Years Of Optical Networking and Communications**
- **Moscone Center, San Francisco, California**
- **Technical Conference: 30 March – 03 April 2025**
- **Exhibition: 01 – 03 April 2025**
- **Expected: Over 13,500 Registrants From 83+ Countries, Showcase of More Than 700 Exhibiting Global Companies**
- **Topics: 1.6 Terabit, AI, Coherent PON, Linear Pluggable Optics (LPO), Multicore Fiber, Data Center Technology, Quantum Networking, etc.**
- **OFCnet – Supported By CENIC, et al**



CERN - CNAF DCI

As presented last year we have a pilot link based on multidomain Spectrum sharing Connection.
We managed to activate a 4x100Gbps to be used as LHCOPN link for INFN-T1 during the Data Challenge 2024



CERN, Esplanade des Particules, Meyrin, Svizzera

INFN CNAF, Viale Carlo Bertini Pichat, Bologna, BO

Also Ref LHC Data Challenge

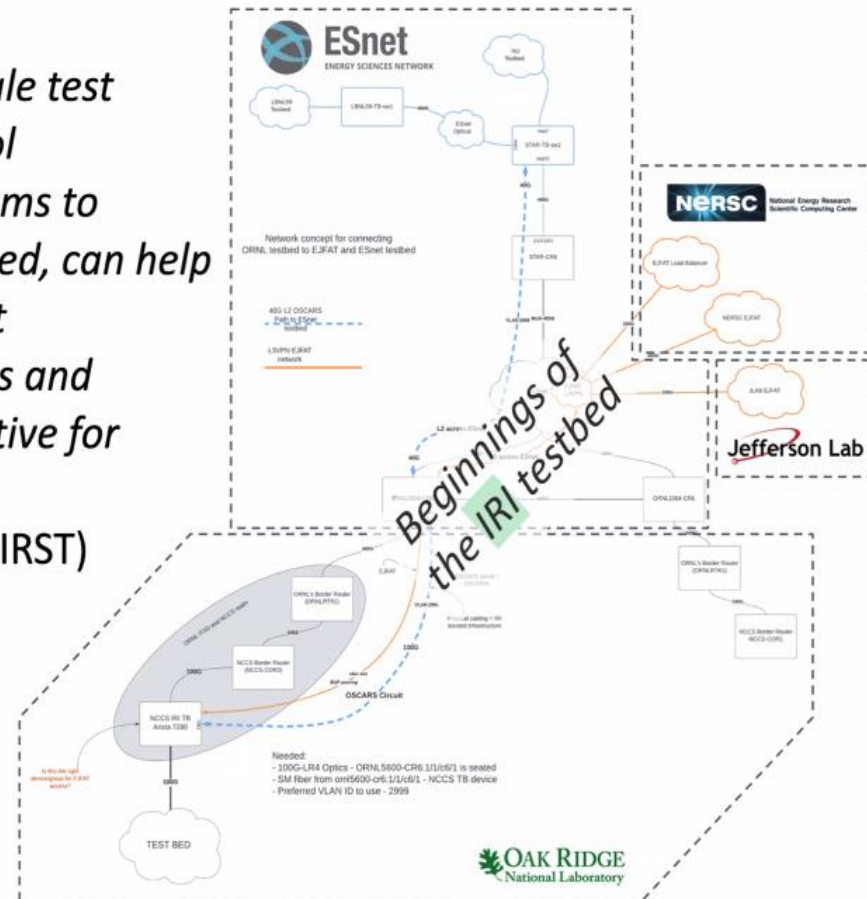
IRI Testbed - A shared playground for IRI developers and pilot applications will accelerate transition to operations

“A best practice in systems design is to create an at-scale test environment in which to build, assess, and improve tool functionality before transitioning technology and systems to production. Such a common test environment, or testbed, can help the broader complex move swiftly to realize an IRI built

on a collectively strong set of tools and functionality already shown effective for science.”

- Federated IRI Science Testbed (FIRST) report

I. Monga, M. Shankar, C. Guok, "Federated
IRI Science Testbed (FIRST): A Concept
Note", December 2023, doi:10.2172/2205149,
<<https://doi.org/10.2172/2205149>>



www.startap.net/starlight

Thanks to the NSF, DOE, NASA,
NIH, DARPA
Universities, National Labs,
International & Industrial
Partners,
and Other Supporters



STARLIGHTSM