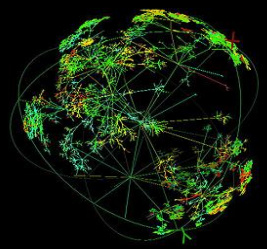
The Global Research Platform, StarLight Software Defined Exchange (SDX), SC24 NREs

Joe Mambretti, Director, (j-mambretti@northwestern.edu)
International Center for Advanced Internet Research (www.icair.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

iCAIR, Northwestern University, StarLight/MREN
IEEE/ACM International Conference on High Performance Computing, Networking, Storage, and Analytics
Atlanta, Georgia
November 17-21, 2024



iCAIR

STARLIGHT SDX

The GRP: A Platform For Global Science

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

STARLIGHT℠

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

STARLIGHT℠

Annual Global Research Platform Workshop – Co-Located With IEEE International Conference On eScience Sept 16-17, 2024

Other GRP Events

- GRP Demonstrations, IEEE/ACM International Conference on High Performance Computing, Networking, Storage and Analytics, Atlanta, Georgia, November 2024
- Mini-Global Research Platform Workshop, Co-Located With the Supercomputing Asia Conference, Singapore, March, 2025
- GRP Demonstrations, OFCnet, OFC Conference, San Francisco, California, March 2024
- Global Research Platform Workshop, Co-Located With IEEE International Conference on eScience, Chicago, Illinois, September, 2025

Selected Applications/Instruments

GENI
www.geni.net

Open Storage Network
www.openstorage network.org

OSG
www.openscience grid.org

Polar Geospatial Center
www.pgc.umn.edu

LSST
www.lsst.org

ISS
www.nasa.gov/ station

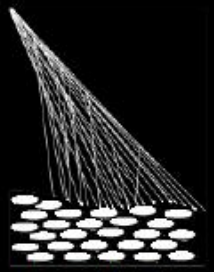
Virgo
www.virgo-gw.eu

LHC
home.cern/science/ accelerators/large-hadron-collider

GLEON
www.gleon.org

GRP
theglobalresearch platform.net/

IceCube
icecube.wisc.edu



Pierre Auger Observatory
www.auger.org

LIGO
www.ligo.caltech. edu

SKA
www.skatelescope. org

LHCONE
twiki.cern.ch/twiki/bin /view/LHCONE/ WebHome

USGS EROS
www.usgs.gov/ centers/eros

Blue Waters
bluewaters.ncsa. illinois.edu

Chameleon
www.chameleon cloud.org

Jetstream
www.jetstream-cloud.org

Belle II
www.belle2.org

XENON
xenon.astro. columbia.edu

SDSS
www.sdss.org

LHCOPN
twiki.cern.ch/twiki/bin /view/LHCOPN/ WebHome

PRAGMA
www.pragma-grid.net

CHASE-CI
www.calit2.net/ newsroom/artic le.php?id=2910

Genomic Science Program
genomicscience. energy.gov

LBNF/DUNE/ ProtoDUNE
lbnf.fnal.gov

NOvA
novaexperiment. fnal.gov

ALMA
www.alma observatory.org

NEON
www.neonscience. org

CENTRA
www.global centra.org

SAGE2
sage2.sagecommons. org

IVOA
www.ivoa.net

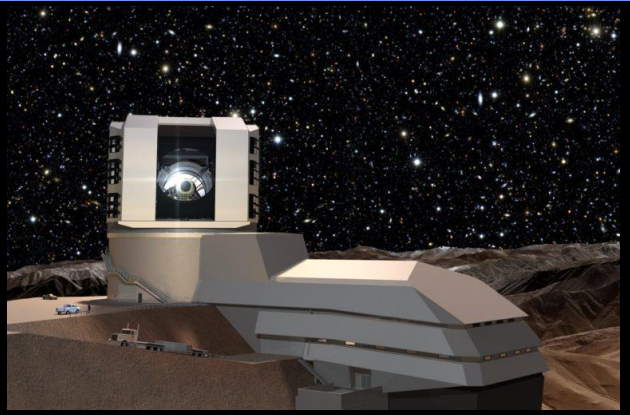
Compilation by Maxine Brown and Joe Mambretti

STARLIGHT℠

Instruments: Exebytes Of Data

High Luminosity LHC

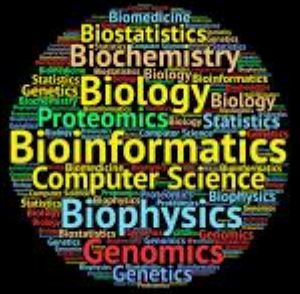
SKA Australia Telescope Facility



Vera Rubin Observatory

KSTAR Korea Superconducting Tokamak

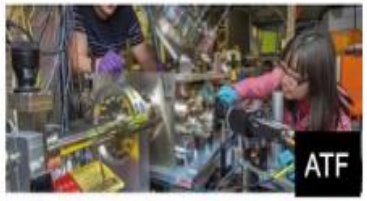
Next Gen Advanced Photon Source



Bioinformatics/Genomics

DOE EXPERIMENTAL USER FACILITIES

- DOE operates 24 experimental user facilities
- Similar to the computing facilities, some of them are undergoing upgrades
- Their data rates and their computing needs will increase accordingly



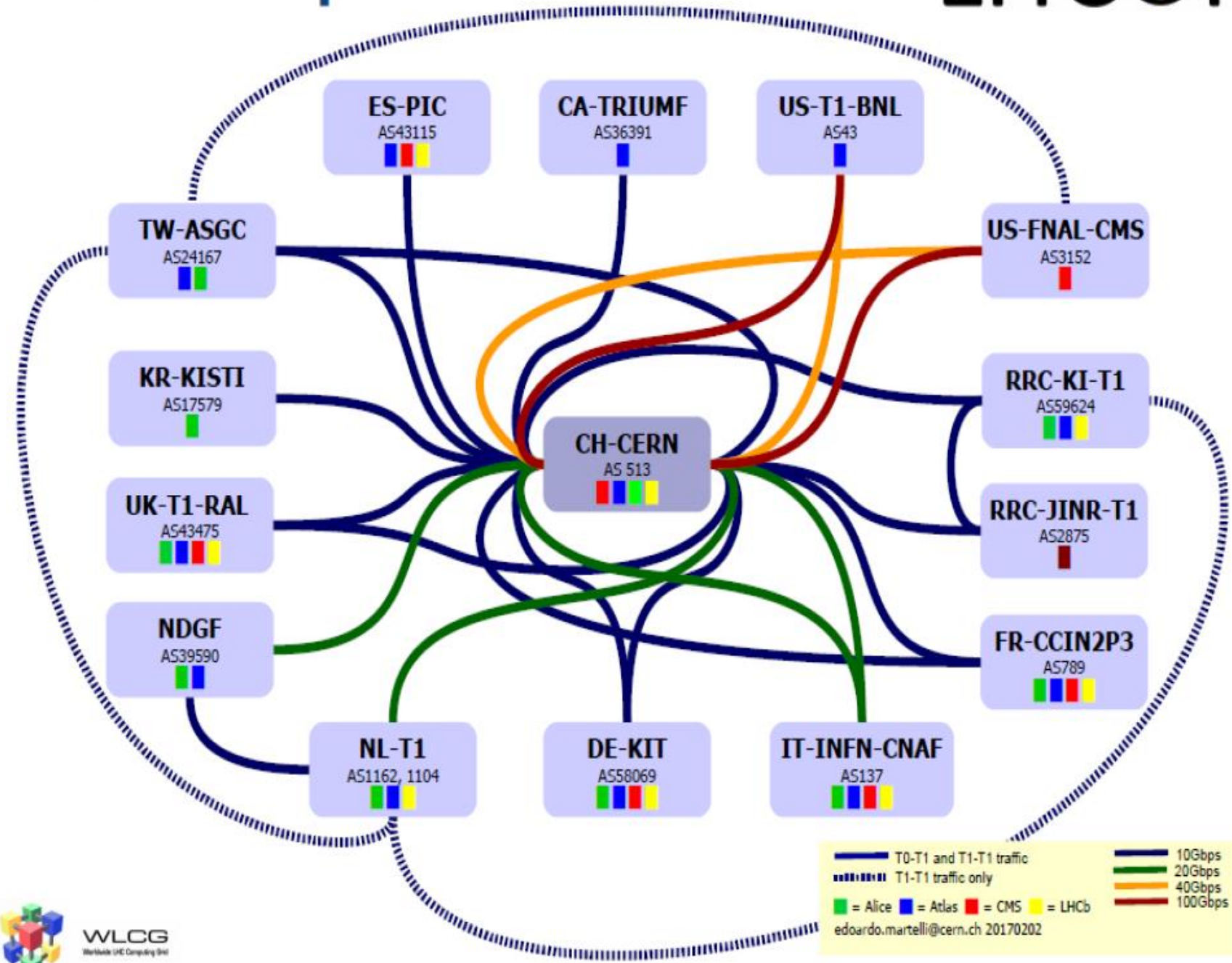
Next Generation Research Platforms

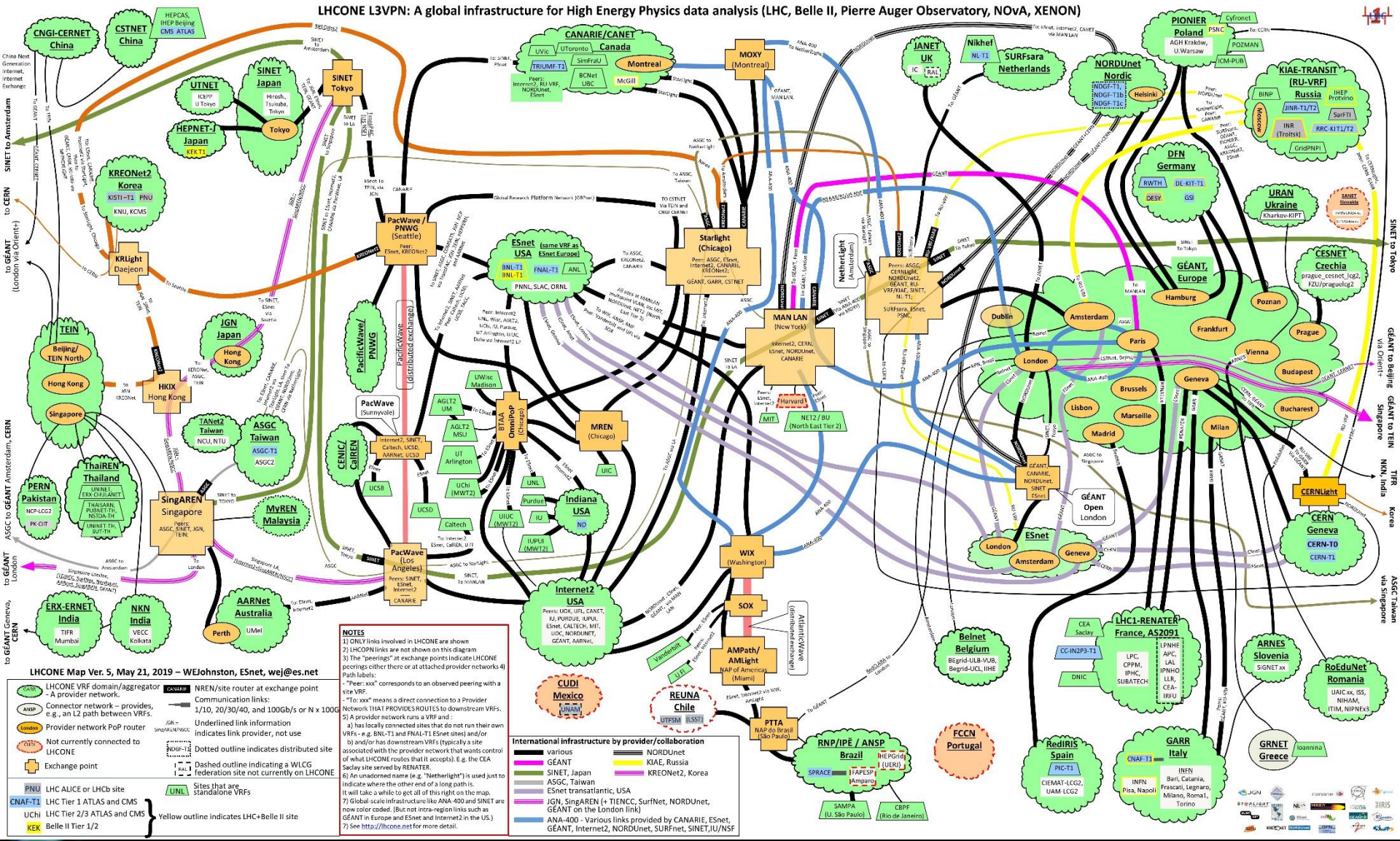
- US National Research Platform
- Asia Pacific Research Platform
- Korean Research Platform
- EU SLICES
- Worldwide LHC Computing Grid (WLCG)
- DOE Integrated Research Infrastructure (IRI)
- Open Science Grid
- Open Science Data Grid
- Et Al

Worldwide LHC Computing Grid

- The Worldwide LHC Computing Grid (WLCG): Global Collaboration Of ~ 170 Computing Centers In More Than 40 Countries, Integrating National and International Grid Infrastructures.
- WLCG Provides Global Resources To Gather, Store, Distribute and Analyze ~200 Petabytes of LHC Data Each Year, Planned To Significantly Increase When the High Luminosity LHC Is Implemented
- WLCG – Partnership of EGI (European Grid Infrastructure), OSG (Open Science Grid), and NeIC (Nordic e-Infrastructure Collaboration).

LHCOPN map


ES-PIC
AS43115
CA-TRIUMF
AS36391
US-T1-BNL
AS43
TW-ASGC
AS24167
US-FNAL-CMS
AS3152
KR-KISTI
AS17579
RRC-KI-T1
AS59624
CH-CERN
AS 513
UK-T1-RAL
AS43475
RRC-JINR-T1
AS2875
NDGF
AS39590
FR-CCIN2P3
AS789
NL-T1
AS1162, 1104
DE-KIT
AS58069
IT-INFN-CNAF
AS137
T0-T1 and T1-T1 traffic
T1-T1 traffic only
= Alice = Atlas = CMS = LHCb
edoardo.martelli@cern.ch 20170202
10Gbps
20Gbps
40Gbps
100Gbps



Major Service, Architecture, Technology Themes

- Orchestration Among Multiple Domains
- Large Scale High Capacity WAN Services
- High-Fidelity Data Flow Monitoring, Visualization, Analytics, Diagnostic Algorithms, Event Correlation AI/ML/DL
- International Testbeds for Data-Intensive Science
- Persistent Specialized Platfoms and Facilities (e.g., High Performance Networks, Open Exchange Points)

NA-REX North America Research & Education Exchange Collaboration

AMPATH PATHWAY TO THE AMERICAS

canarie

ESnet ENERGY SCIENCES NETWORK

INTERNET2

International Networks at Indiana University

STARLIGHT The Optical STAR TAP

CENIC

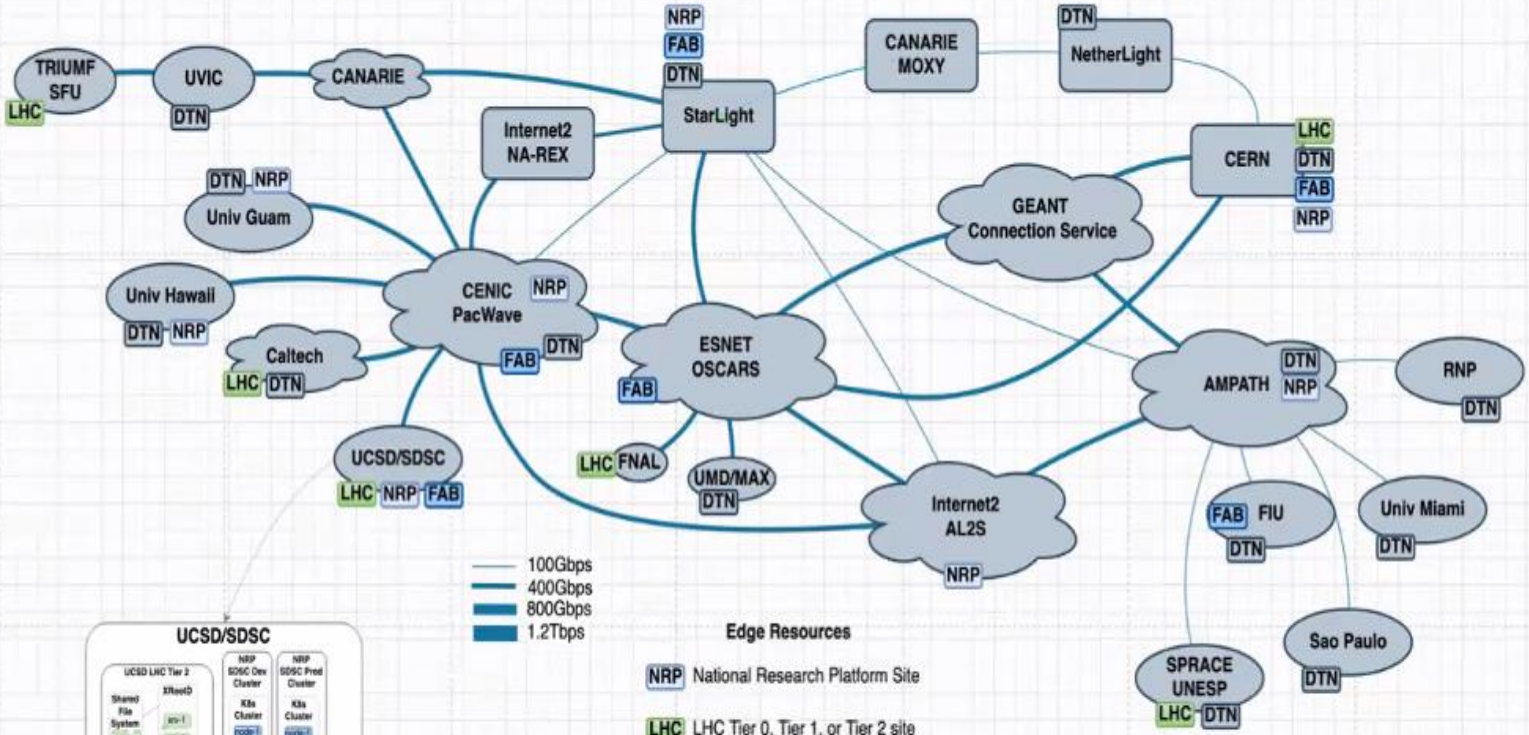
PACIFIC NORTHWEST GIGAPOP

PACIFIC WAVE

November 2023

AutoGOLE Open R&E Exchanges

STARLIGHT℠



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

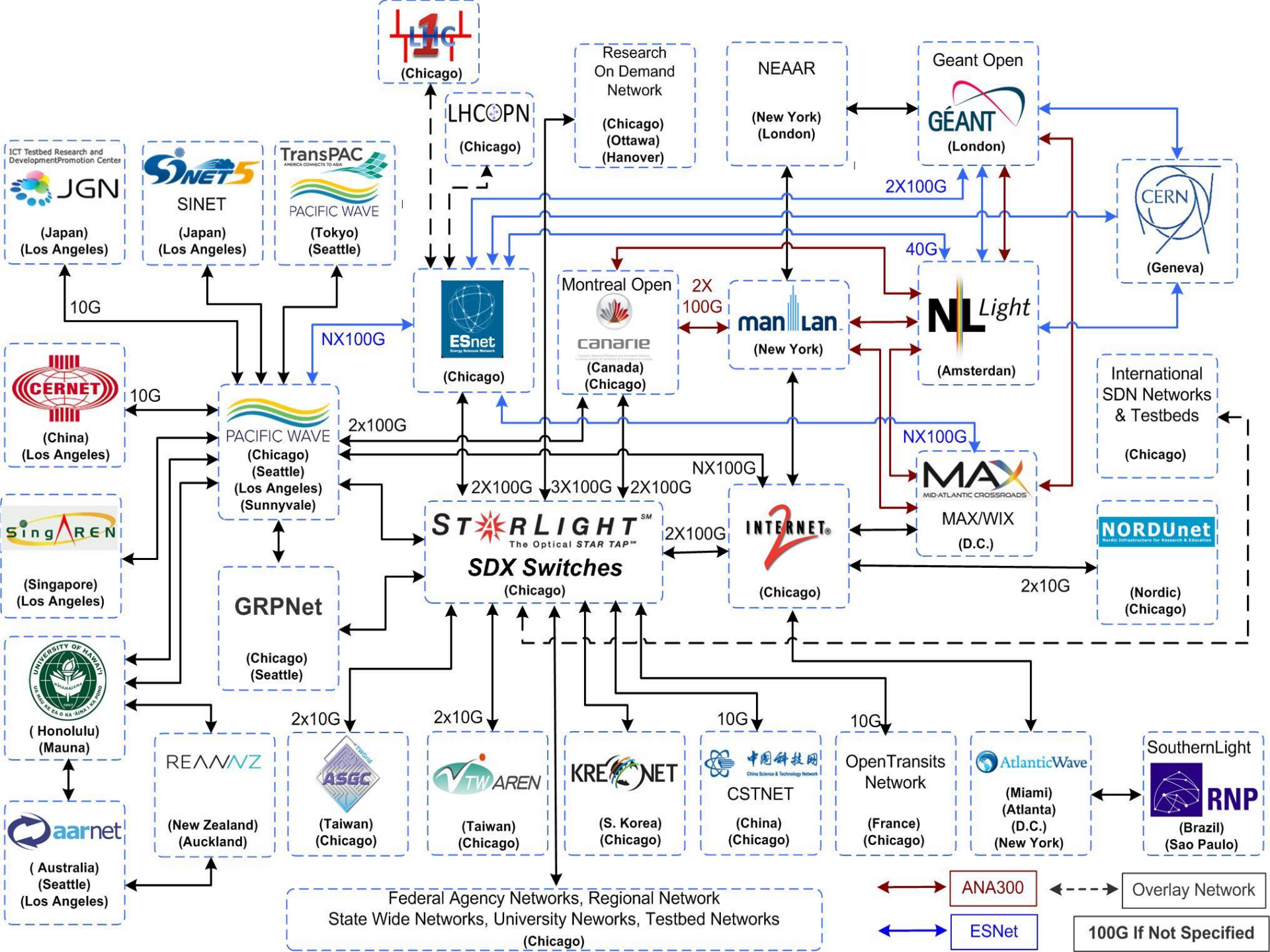
View from StarLight

Abbott Hall, Northwestern University's Chicago Campus

Currently: 26+ 400 Gbps Paths Prototyping 800 Gbps Tbps

iCAIR

STARLIGHT℠


LHC
(Chicago)
LHCOPN
(Chicago)
Research On Demand Network
(Chicago)
(Ottawa)
(Hanover)
NEAAR
(New York)
(London)
Geant Open
GÉANT
(London)
CERN
(Geneva)
2X100G
40G
ICT Testbed Research and DevelopmentPromotion Center
JGN
(Japan)
(Los Angeles)
SINET
(Japan)
(Los Angeles)
TransPAC
PACIFIC WAVE
(Tokyo)
(Seattle)
10G
NX100G
ESnet
(Chicago)
Montreal Open
canarie
(Canada)
(Chicago)
2X 100G
manLan
(New York)
NL Light
(Amsterdan)
CERNET
(China)
(Los Angeles)
10G
PACIFIC WAVE
(Chicago)
(Seattle)
(Los Angeles)
(Sunnyvale)
2x100G
International SDN Networks & Testbeds
(Chicago)
NX100G
MAX
MID-ATLANTIC CROSSROADS
MAX/WIX
(D.C.)
NX100G
2X100G
3X100G
2X100G
SingAREN
(Singapore)
(Los Angeles)
STARLIGHT
The Optical STAR TAP
SDX Switches
(Chicago)
2X100G
INTERNET2
(Chicago)
2x10G
NORDUnet
(Nordic)
(Chicago)
GRPNet
(Chicago)
(Seattle)
(Honolulu)
(Mauna)
2x10G
2x10G
10G
10G
REANNZ
(New Zealand)
(Auckland)
ASGC
(Taiwan)
(Chicago)
TWAREN
(Taiwan)
(Chicago)
KREONET
(S. Korea)
(Chicago)
CSTNET
(China)
(Chicago)
OpenTransits Network
(France)
(Chicago)
AtlanticWave
(Miami)
(Atlanta)
(D.C.)
(New York)
SouthernLight
RNP
(Brazil)
(Sao Paulo)
aarnet
(Australia)
(Seattle)
(Los Angeles)
Federal Agency Networks, Regional Network
State Wide Networks, University Neworks, Testbed Networks
(Chicago)
ANA300
ESNet
Overlay Network
100G If Not Specified

SCinet National WAN Testbed

- As In Previous Years, iCAIR Supported 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 SC24 National WAN Testbed Was Designed and Implemented To Support Demonstrations and Experiments Of Innovations, Many Related To Data Intensive Science

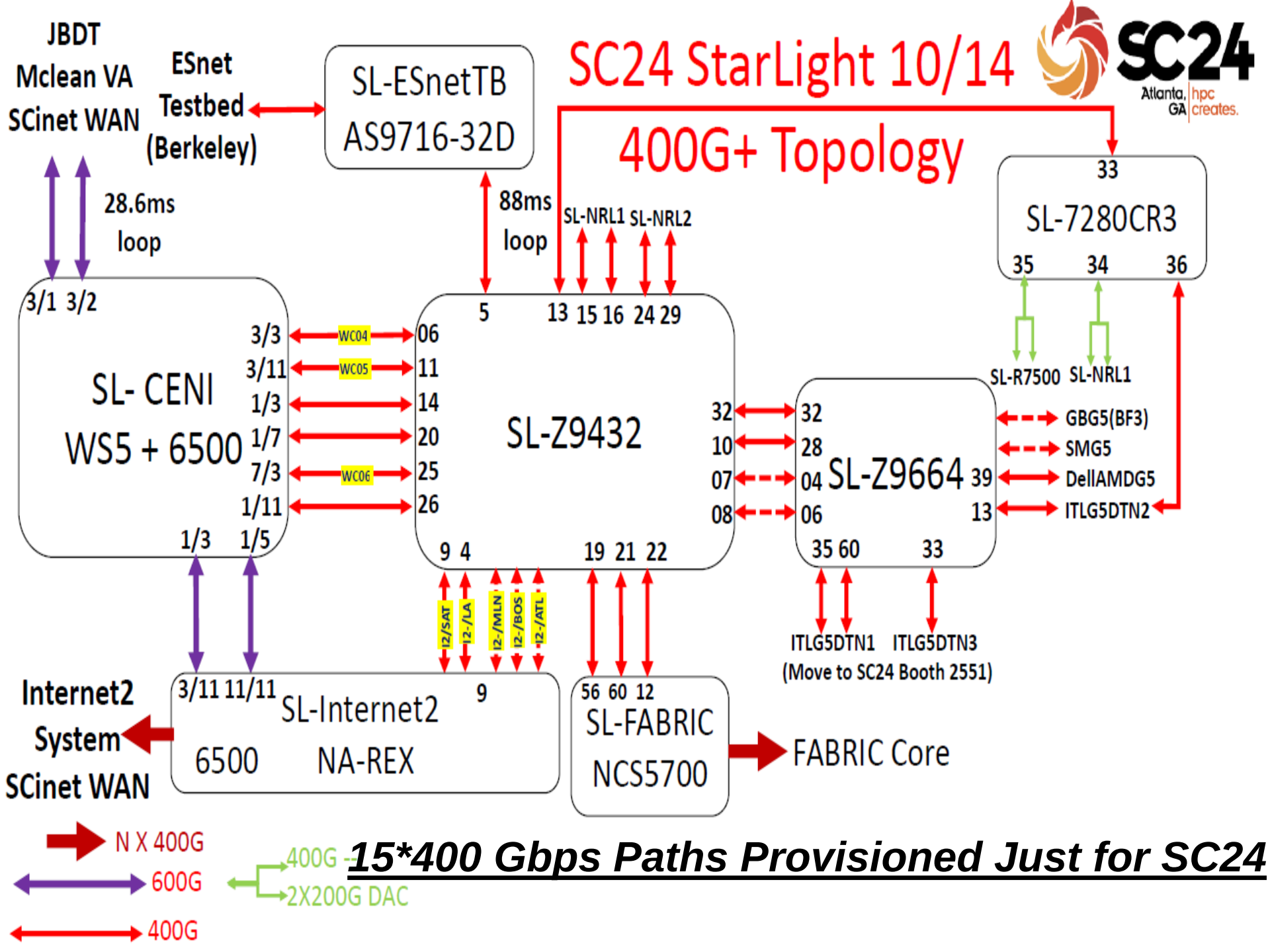
Example SC24 SCinet Network Research Exhibitions

- 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

STARLIGHT℠

Tbps WAN Services For Data Intensive Science

- NRE Demonstrations Leverage Experimental Research In the Optimal Design, Configuration, Components, and Integration Of DTNs
- Technologies: Software Defined Networking (SDN), 400 Gbps Smart NICs, NVMe over Fabric, RDMA,Techniques For Kernel Bypass Using Zero-Copy for Memory and Disk Copy (To Avoid Bottlenecks in Large Scale Data Transfers Over 1.2 Tbps WANs), GPU Acceleration, P4 And Optimal Affinity Bindings for NUMA Architecture for Higher Resource Utilization
- Demonstrations Also Showcase Middleware Required to Orchestrate infrastructure Resources for Reliable, Optimized High-Performance WAN Data Transfers.
- Demonstrations Showcase Measurement Techniques for Real-Time monitoring, benchmarking and evaluation including an AI Toolset.
- WAN Service Investigations Include Examining Implications for Production Operations.

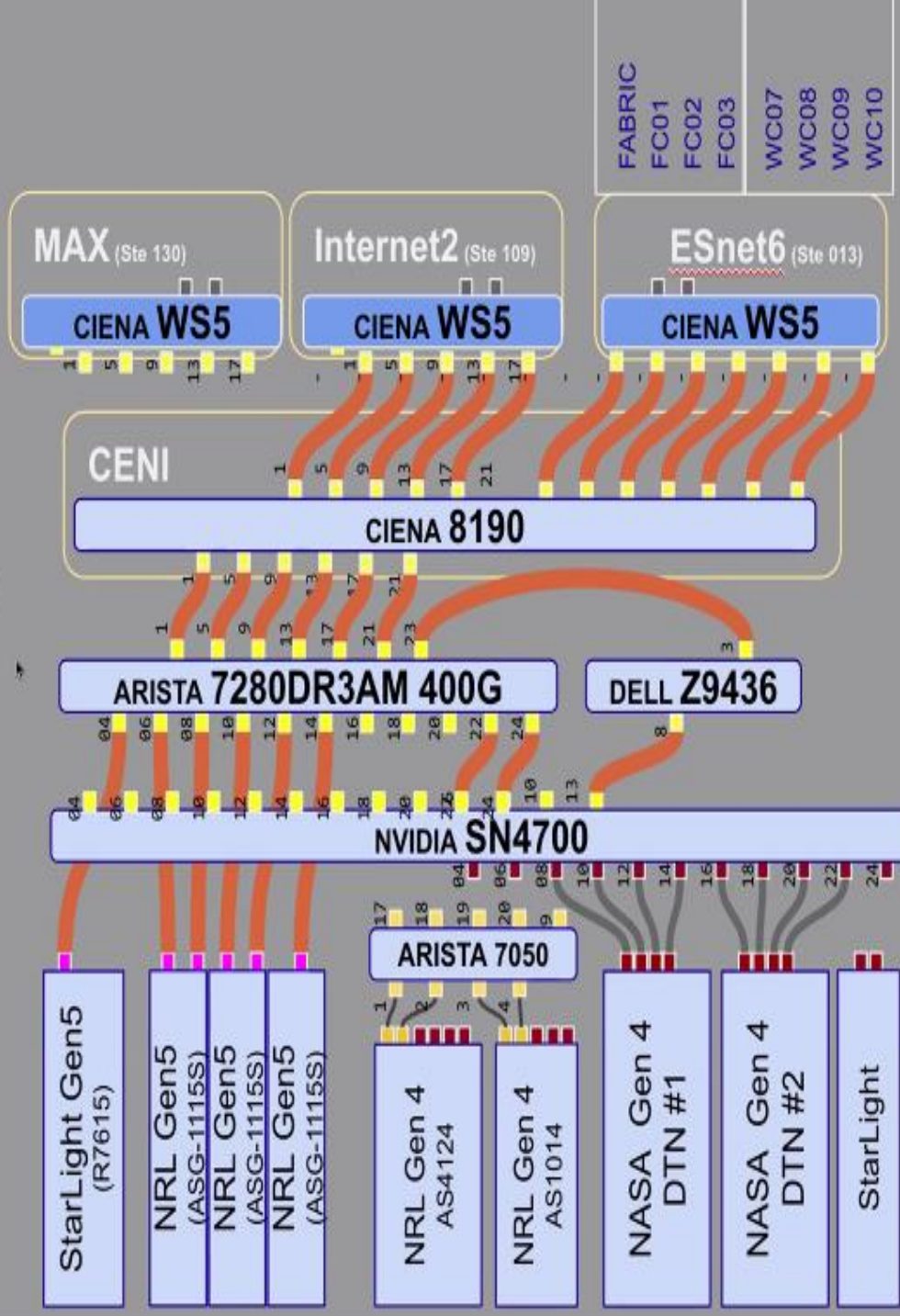

JBDT
Mclean VA
SCinet WAN
ESnet
Testbed
(Berkeley)
SL-ESnetTB
AS9716-32D
SC24 StarLight 10/14
400G+ Topology
SC24
Atlanta, GA
hpc creates.
28.6ms
loop
88ms
loop
SL-NRL1
SL-NRL2
SL-7280CR3
33
35
34
36
SL- CENI
WS5 + 6500
3/1
3/2
3/3
3/11
1/3
1/7
7/3
1/11
1/3
1/5
WC04
WC05
WC06
SL-Z9432
5
13
15
16
24
29
06
11
14
20
25
26
9
4
19
21
22
32
10
07
08
SL-Z9664
32
28
04
06
35
60
33
39
13
SL-R7500
SL-NRL1
GBG5(BF3)
SMG5
DellAMDG5
ITLG5DTN2
I2/SAT
I2-/LA
I2-/MLN
I2-/BOS
I2-/ATL
ITLG5DTN1
ITLG5DTN3
(Move to SC24 Booth 2551)
Internet2
System
SCinet WAN
3/11
11/11
SL-Internet2
6500
NA-REX
9
56
60
12
SL-FABRIC
NCS5700
FABRIC Core
N X 400G
600G
400G
400G -
2X200G DAC
15*400 Gbps Paths Provisioned Just for SC24

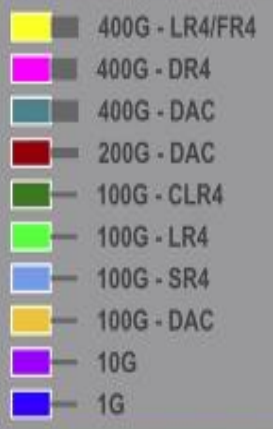
Need

- Sort out BW for CENI + FABRIC + SCinet
-

Use
7280 towards the WAN

7280 - for server connections (1 spare)





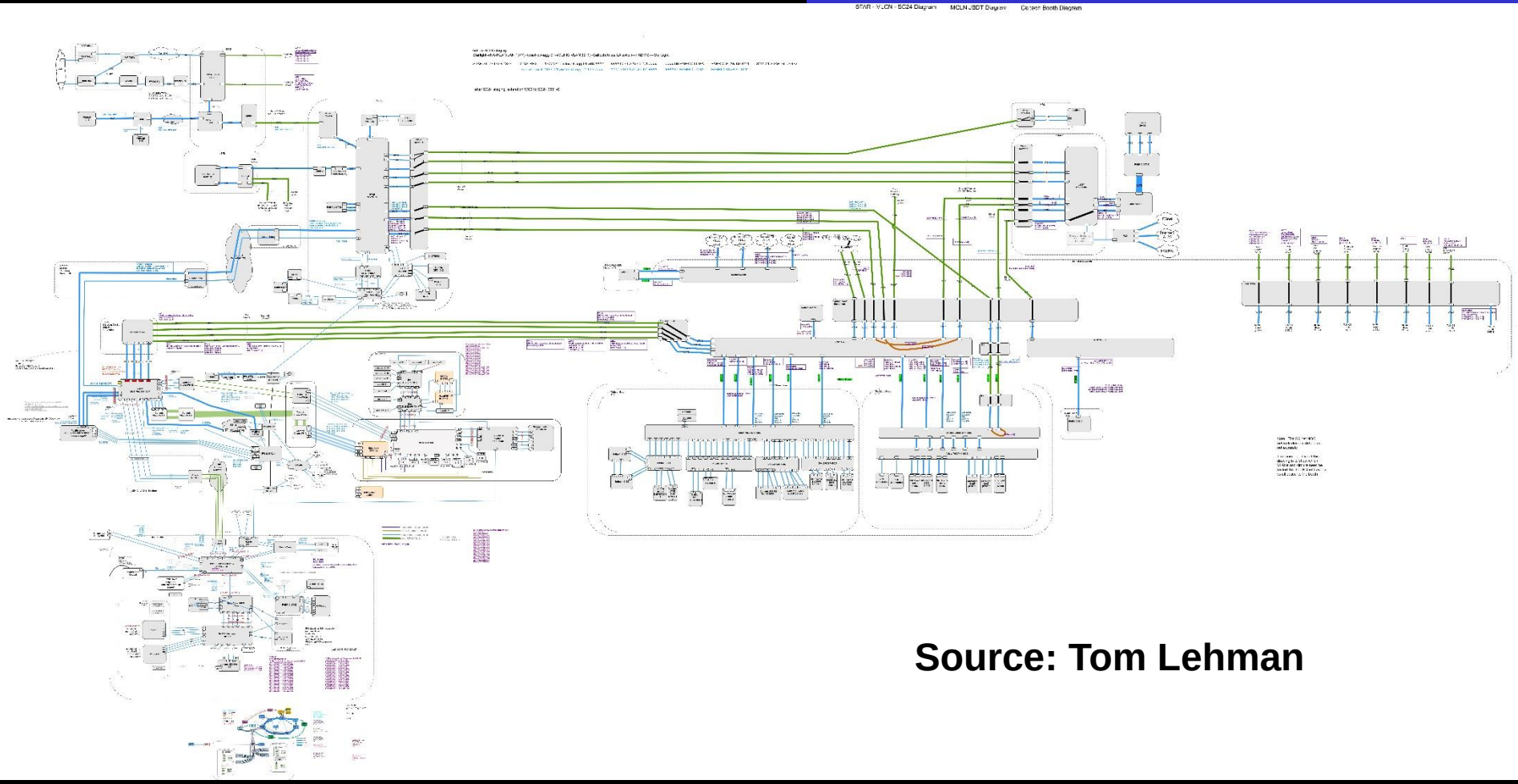
McLean now and after SC

VLAN info

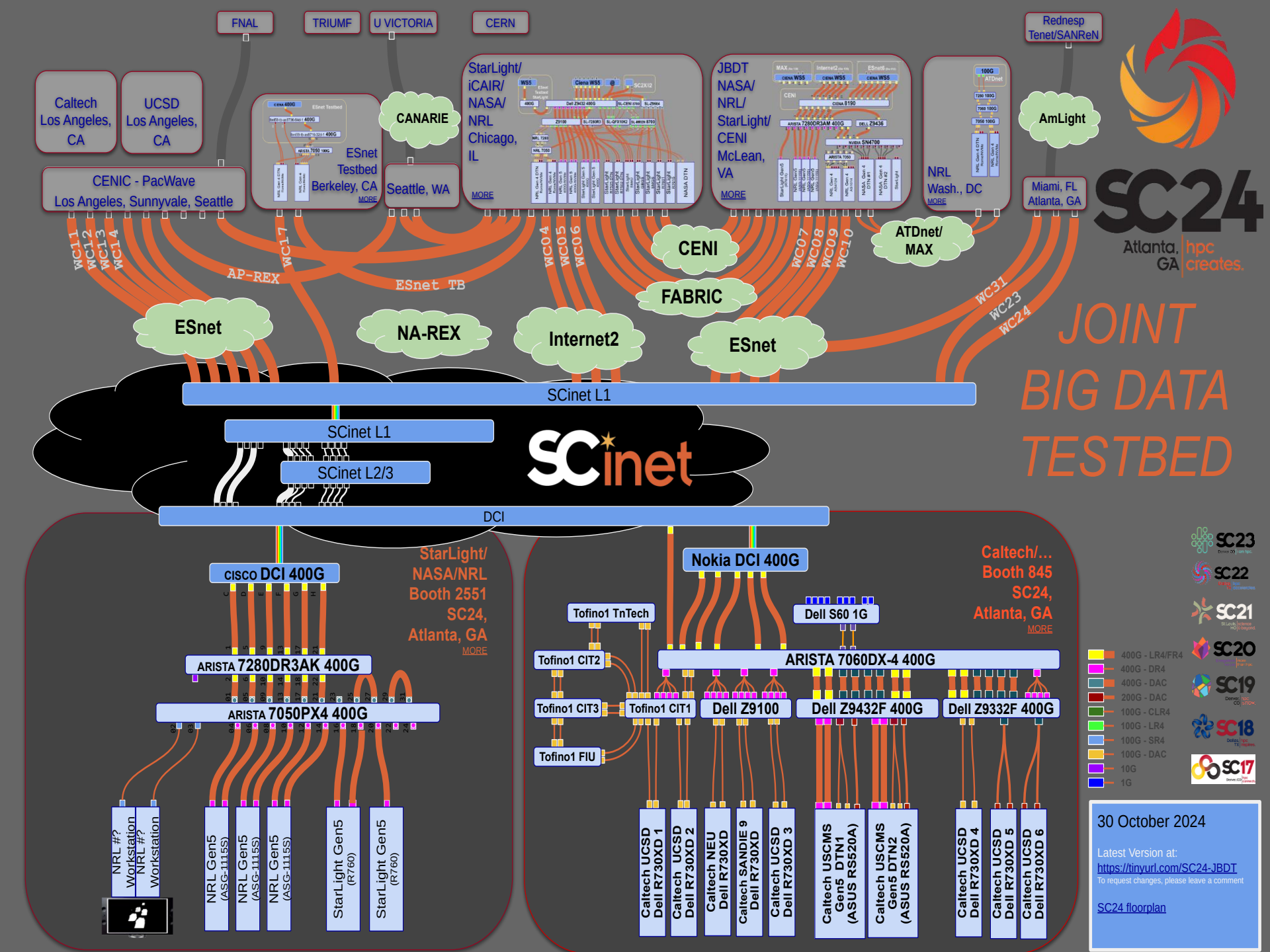
VLAN info SC24

10/16/2024

International Research Infrastructure Testbed

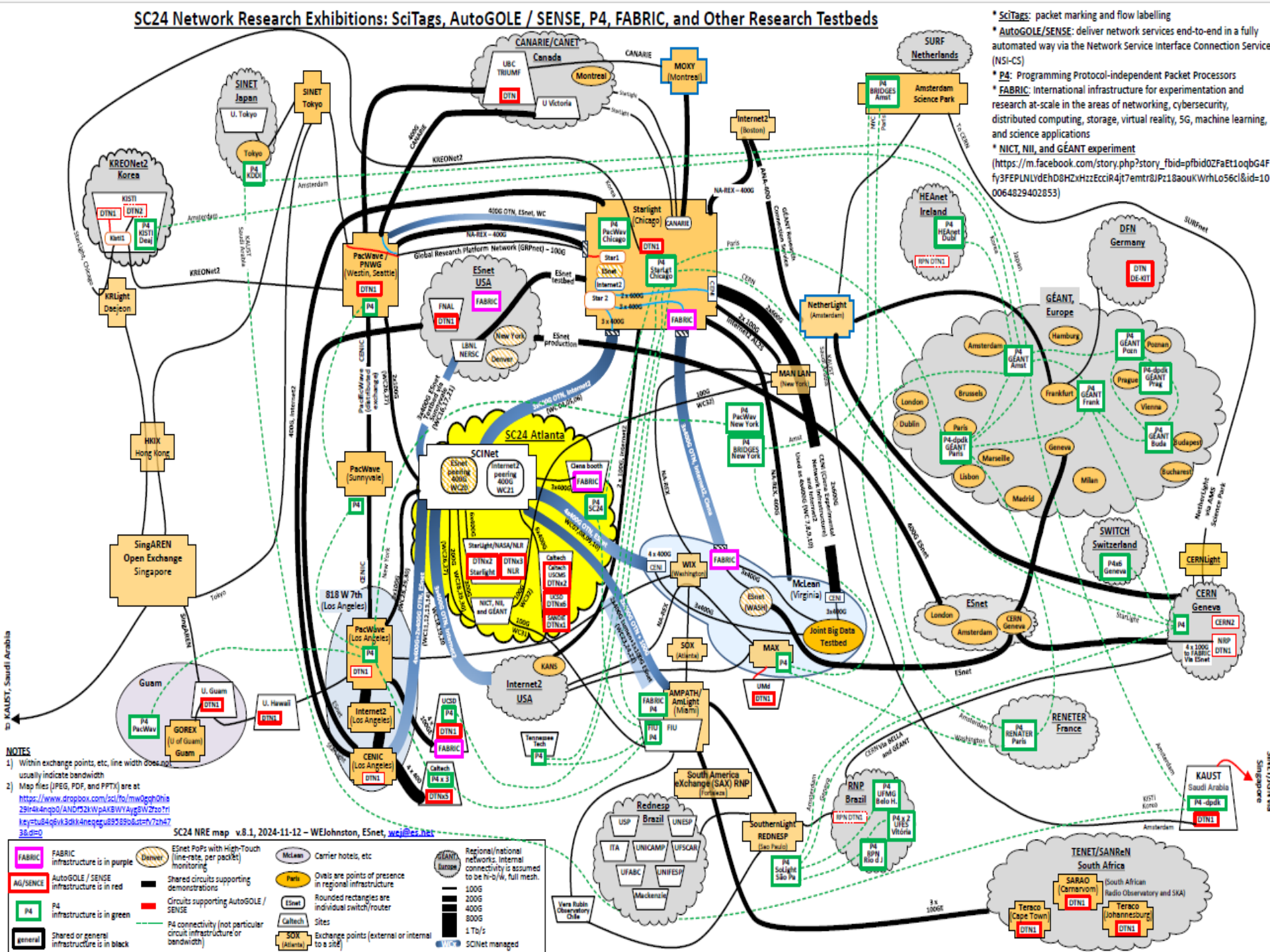


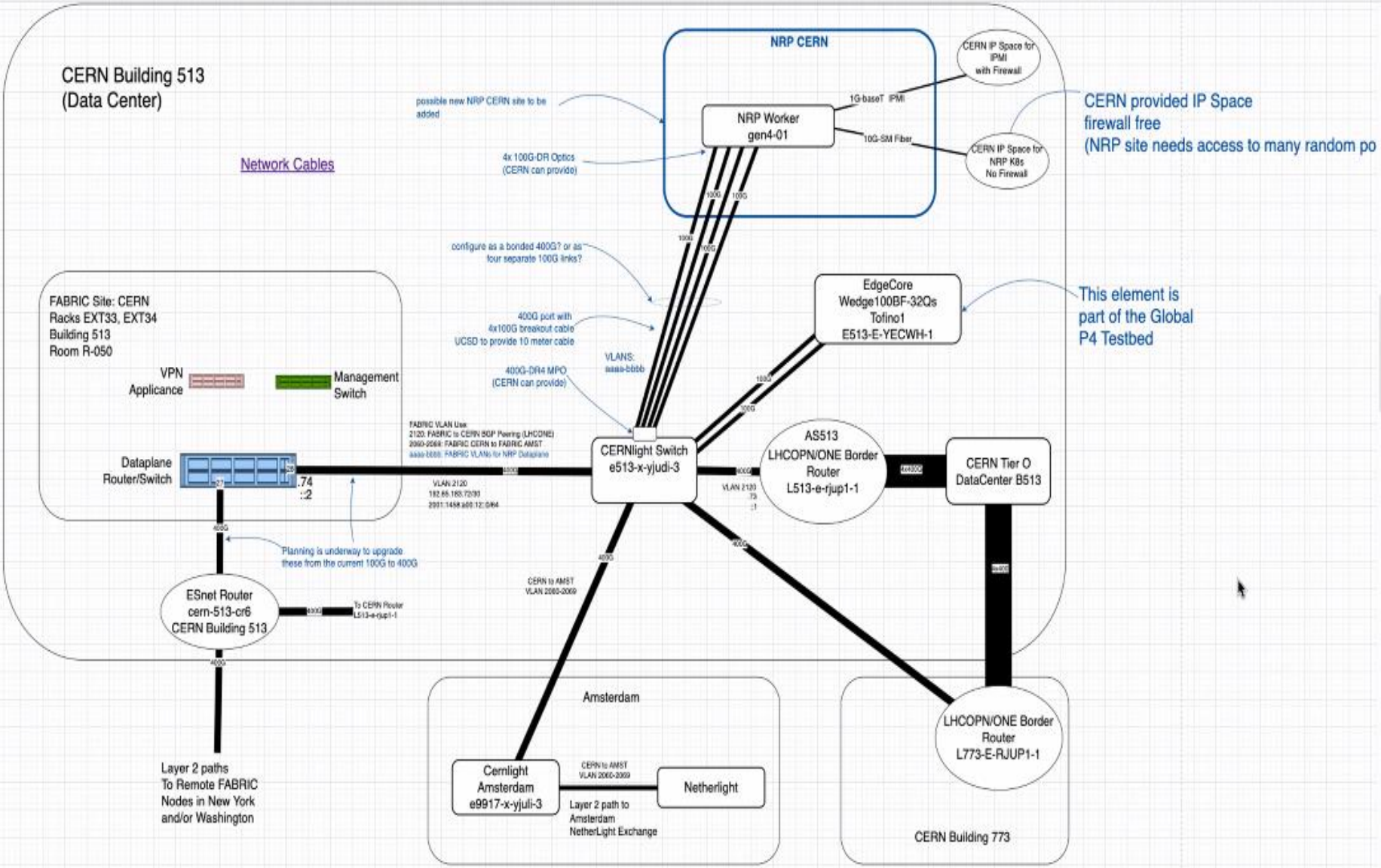
Source: Tom Lehman

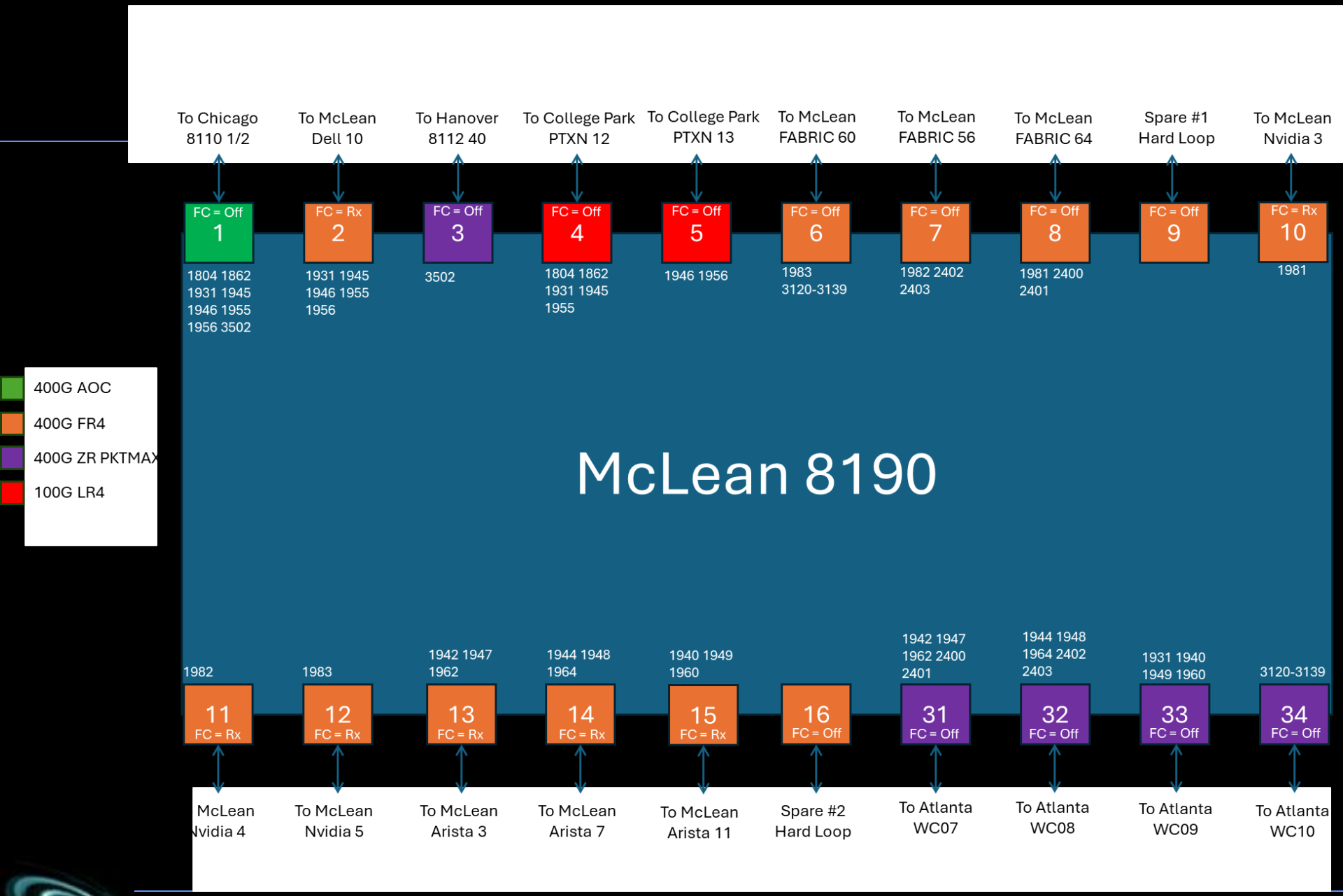

FNAL
TRIUMF
U VICTORIA
CERN
Rednesp
Tenet/SANReN
Caltech
Los Angeles, CA
UCSD
Los Angeles, CA
CENIC - PacWave
Los Angeles, Sunnyvale, Seattle
ESnet
Testbed
Berkeley, CA
MORE
CANARIE
Seattle, WA
StarLight/
iCAIR/
NASA/
NRL
Chicago, IL
MORE
JBDT
NASA/
NRL/
StarLight/
CENI
McLean, VA
MORE
NRL
Wash., DC
MORE
AmLight
Miami, FL
Atlanta, GA
SC24
Atlanta, GA
hpc creates.
JOINT
BIG DATA
TESTBED
WC11
WC12
WC13
WC14
AP-REX
WC17
ESnet TB
WC04
WC05
WC06
CENI
FABRIC
ATDnet/
MAX
WC07
WC08
WC09
WC10
WC31
WC23
WC24
ESnet
NA-REX
Internet2
ESnet
SCinet L1
SCinet L1
SCinet L2/3
SCinet
DCI
StarLight/
NASA/NRL
Booth 2551
SC24,
Atlanta, GA
MORE
cisco DCI 400G
ARISTA 7280DR3AK 400G
ARISTA 7050PX4 400G
NRL #? Workstation
NRL #? Workstation
NRL Gen5 (ASG-1115S)
NRL Gen5 (ASG-1115S)
NRL Gen5 (ASG-1115S)
StarLight Gen5 (R760)
StarLight Gen5 (R760)
Nokia DCI 400G
Caltech/…
Booth 845
SC24,
Atlanta, GA
MORE
Tofino1 TnTech
Tofino1 CIT2
Tofino1 CIT3
Tofino1 FIU
Tofino1 CIT1
Dell S60 1G
ARISTA 7060DX-4 400G
Dell Z9100
Dell Z9432F 400G
Dell Z9332F 400G
Caltech UCSD Dell R730XD 1
Caltech UCSD Dell R730XD 2
Caltech NEU Dell R730XD
Caltech SANDIE 9 Dell R730XD
Caltech UCSD Dell R730XD 3
Caltech USCMS Gen5 DTN1 (ASUS RS520A)
Caltech USCMS Gen5 DTN2 (ASUS RS520A)
Caltech UCSD Dell R730XD 4
Caltech UCSD Dell R730XD 5
Caltech UCSD Dell R730XD 6
400G - LR4/FR4
400G - DR4
400G - DAC
200G - DAC
100G - CLR4
100G - LR4
100G - SR4
100G - DAC
10G
1G
SC23
SC22
SC21
SC20
SC19
SC18
SC17
30 October 2024
Latest Version at:
https://tinyurl.com/SC24-JBDT
To request changes, please leave a comment
SC24 floorplan

SC24 Network Research Exhibitions: SciTags, AutoGOLE / SENSE, P4, FABRIC, and Other Research Testbeds

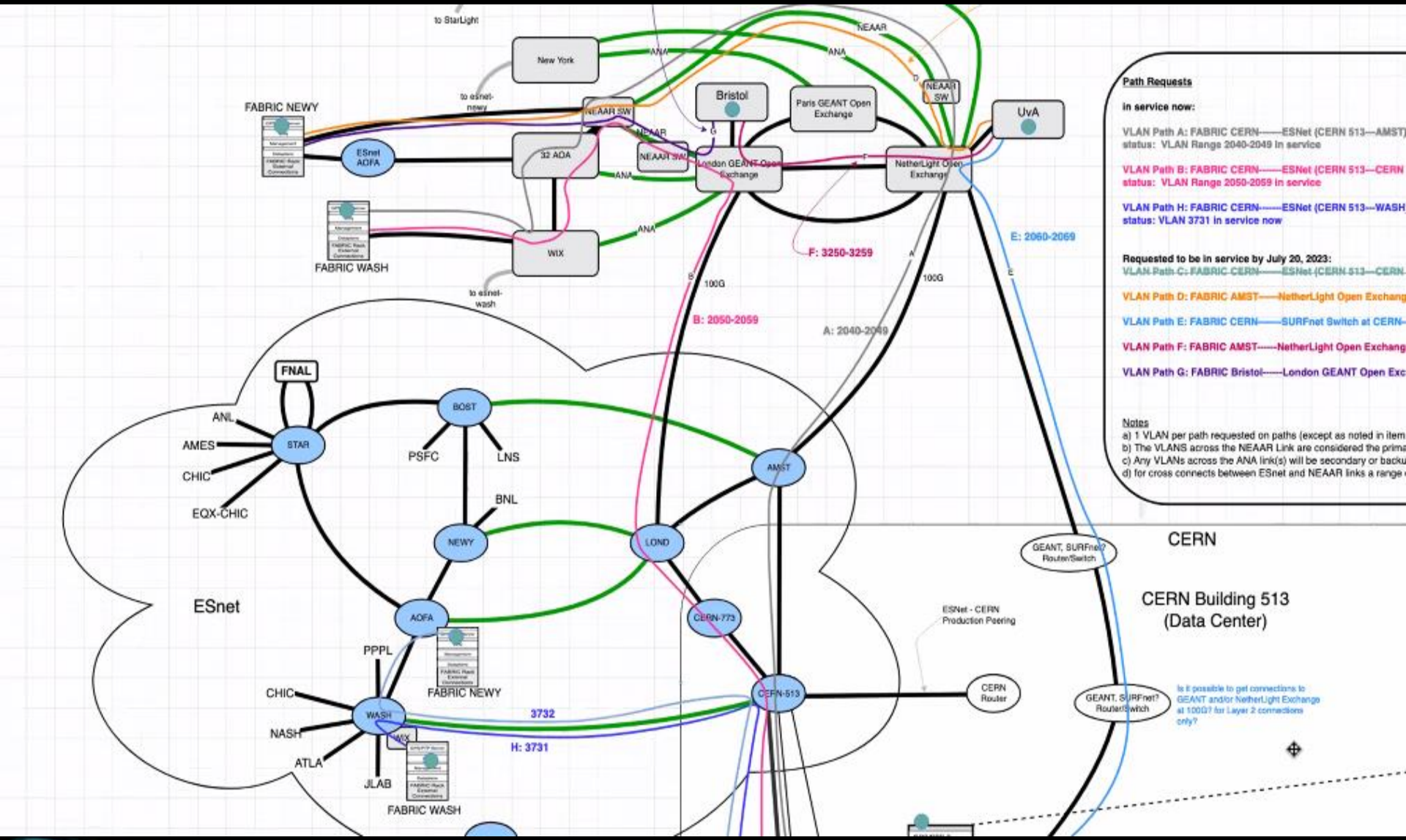
* SciTags: packet marking and flow labelling
* AutoGOLE/SENSE: deliver network services end-to-end in a fully automated way via the Network Service Interface Connection Service (NSI-CS)
* P4: Programming Protocol-independent Packet Processors
* FABRIC: International infrastructure for experimentation and research at-scale in the areas of networking, cybersecurity, distributed computing, storage, virtual reality, 5G, machine learning, and science applications
* NICT, NII, and GÉANT experiment (https://m.facebook.com/story.php?story_fbid=pfbid0ZFaEt1oqbG4Ffy3FEPLNLYdEhD8HZxHzzEcciR4jt7emtr8JPz18aouKWrhLo56cl&id=100064829402853)



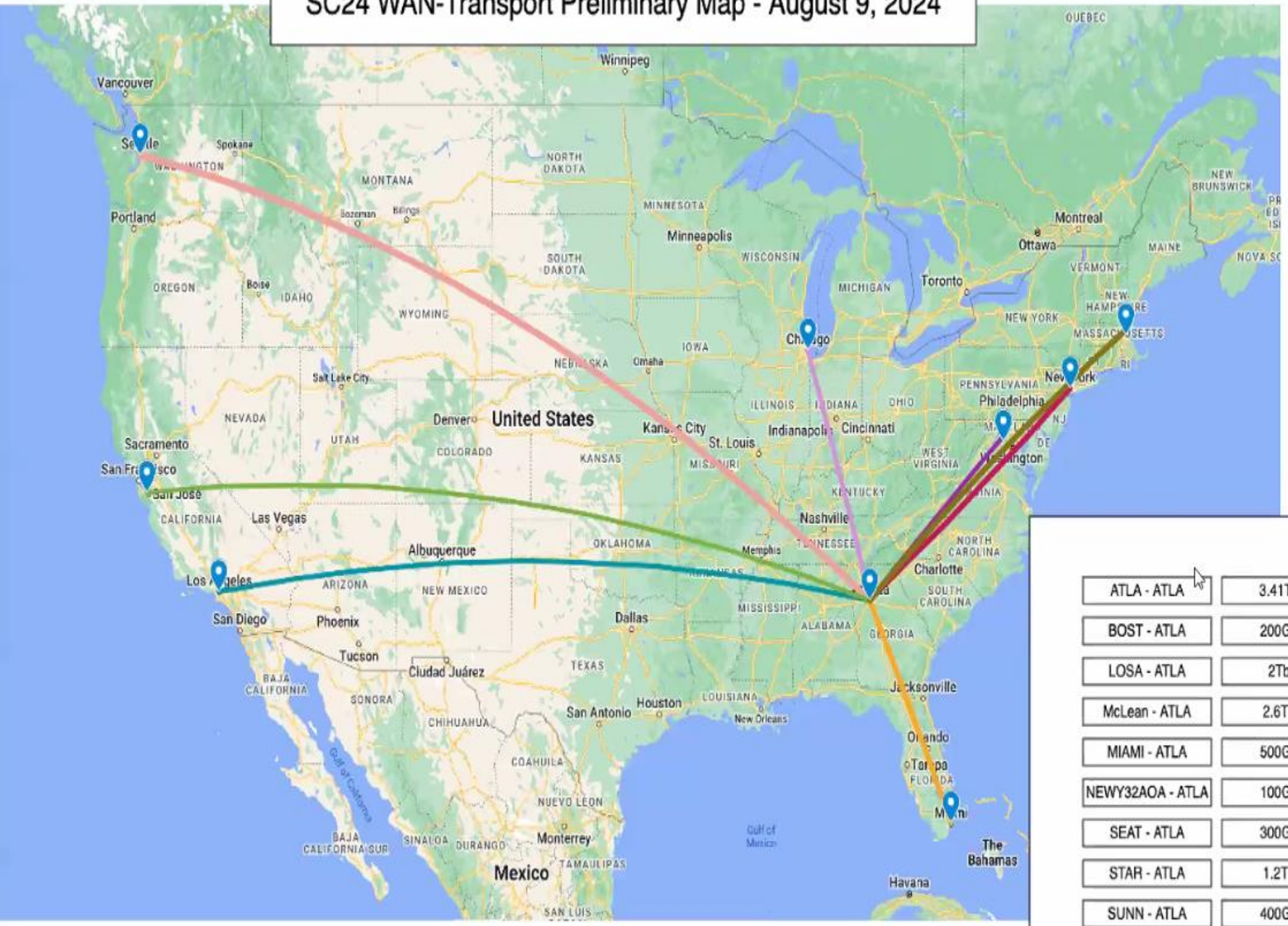

CERN Building 513
(Data Center)
Network Cables
NRP CERN
NRP Worker
gen4-01
CERN IP Space for
IPMI
with Firewall
CERN IP Space for
NRP K8s
No Firewall
1G-baseT IPMI
10G-SM Fiber
possible new NRP CERN site to be added
4x 100G-DR Optics
(CERN can provide)
CERN provided IP Space
firewall free
(NRP site needs access to many random po
configure as a bonded 400G? or as four separate 100G links?
400G port with
4x100G breakout cable
UCSD to provide 10 meter cable
400G-DR4 MPO
(CERN can provide)
VLANS:
aaaa-bbbb
EdgeCore
Wedge100BF-32Qs
Tofino1
E513-E-YECWH-1
This element is
part of the Global
P4 Testbed
FABRIC Site: CERN
Racks EXT33, EXT34
Building 513
Room R-050
VPN
Applicance
Management
Switch
Dataplane
Router/Switch
.74
:2
FABRIC VLAN Use:
2120: FABRIC to CERN BGP Peering (LHCONE)
2060-2069: FABRIC CERN to FABRIC AMST
aaaa-bbbb: FABRIC VLANs for NRP Dataplane
VLAN 2120
CERNlight Switch
e513-x-yjudi-3
AS513
LHCOPN/ONE Border
Router
L513-e-rjup1-1
CERN Tier O
DataCenter B513
Planning is underway to upgrade these from the current 100G to 400G
ESnet Router
cern-513-cr6
CERN Building 513
To CERN Router
L513-e-rjup1-1
CERN to AMST
VLAN 2060-2069
Layer 2 paths
To Remote FABRIC
Nodes in New York
and/or Washington
Amsterdam
Cernlight
Amsterdam
e9917-x-yjuli-3
CERN to AMST
VLAN 2060-2069
Netherlight
Layer 2 path to
Amsterdam
NetherLight Exchange
LHCOPN/ONE Border
Router
L773-E-RJUP1-1
CERN Building 773


To Chicago 8110 1/2
To McLean Dell 10
To Hanover 8112 40
To College Park PTXN 12
To College Park PTXN 13
To McLean FABRIC 60
To McLean FABRIC 56
To McLean FABRIC 64
Spare #1 Hard Loop
To McLean Nvidia 3
FC = Off 1
1804 1862 1931 1945 1946 1955 1956 3502
FC = Rx 2
1931 1945 1946 1955 1956
FC = Off 3
3502
FC = Off 4
1804 1862 1931 1945 1955
FC = Off 5
1946 1956
FC = Off 6
1983 3120-3139
FC = Off 7
1982 2402 2403
FC = Off 8
1981 2400 2401
FC = Off 9
FC = Rx 10
1981
400G AOC
400G FR4
400G ZR PKTMAX
100G LR4
McLean 8190
1982
11 FC = Rx
1983
12 FC = Rx
1942 1947 1962
13 FC = Rx
1944 1948 1964
14 FC = Rx
1940 1949 1960
15 FC = Rx
16 FC = Off
1942 1947 1962 2400 2401
31 FC = Off
1944 1948 1964 2402 2403
32 FC = Off
1931 1940 1949 1960
33 FC = Off
3120-3139
34 FC = Off
McLean Nvidia 4
To McLean Nvidia 5
To McLean Arista 3
To McLean Arista 7
To McLean Arista 11
Spare #2 Hard Loop
To Atlanta WC07
To Atlanta WC08
To Atlanta WC09
To Atlanta WC10

STARLIGHT℠

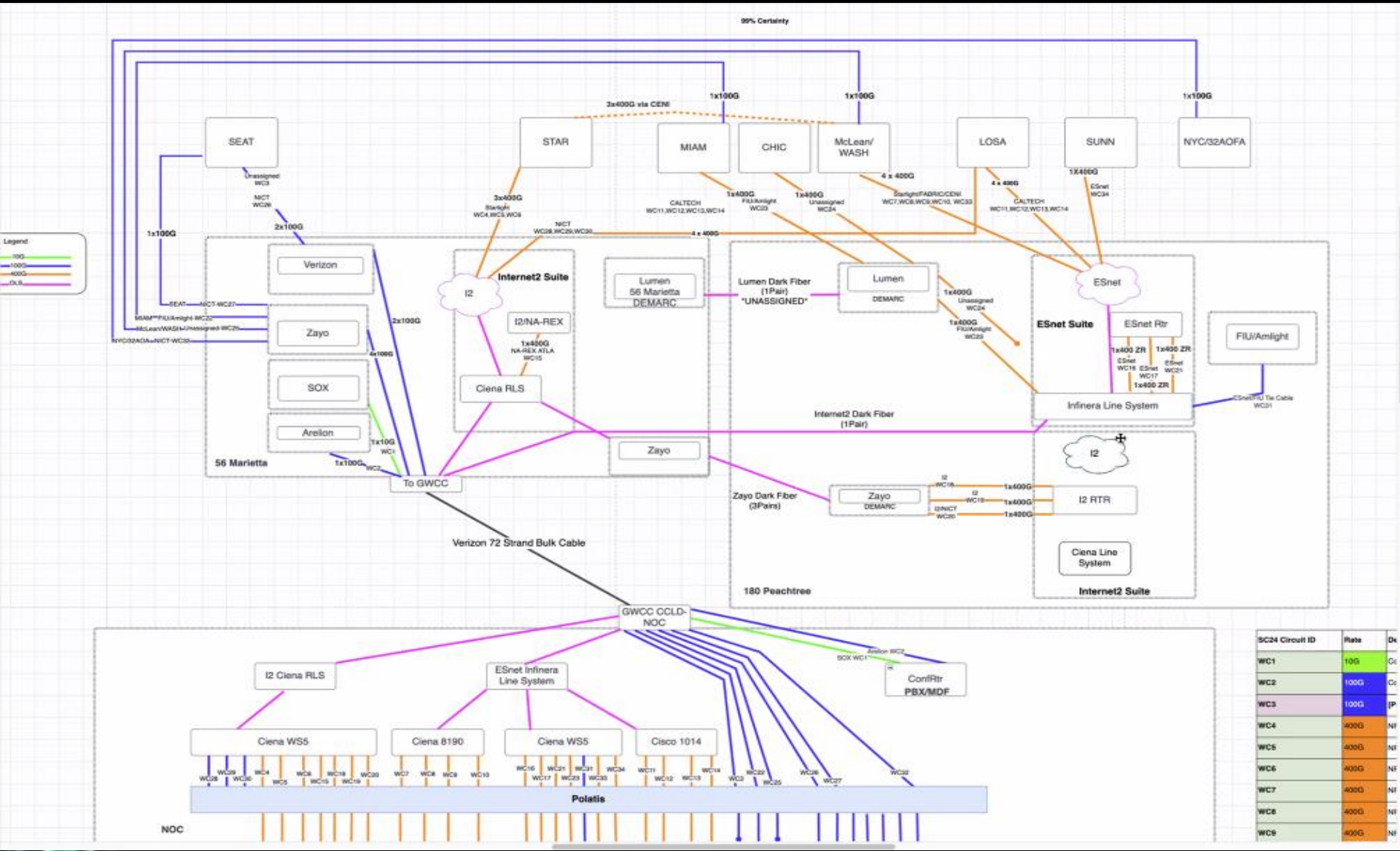

to StarLight
New York
NEAAR
ANA
FABRIC NEWY
to esnet-newy
NEAAR SW
Bristol
Paris GEANT Open Exchange
NEAAR SW
UvA
ESnet AOFA
32 AOA
NEAAR SW
London GEANT Open Exchange
NetherLight Open Exchange
ANA
FABRIC WASH
WIX
to esnet-wash
E: 2060-2069
F: 3250-3259
100G
100G
B: 2050-2059
A: 2040-2049
Path Requests
in service now:
VLAN Path A: FABRIC CERN------ESNet (CERN 513---AMST)
status: VLAN Range 2040-2049 in service
VLAN Path B: FABRIC CERN------ESNet (CERN 513---CERN
status: VLAN Range 2050-2059 in service
VLAN Path H: FABRIC CERN-------ESNet (CERN 513---WASH
status: VLAN 3731 in service now
Requested to be in service by July 20, 2023:
VLAN Path C: FABRIC CERN------ESNet (CERN 513---CERN
VLAN Path D: FABRIC AMST------NetherLight Open Exchang
VLAN Path E: FABRIC CERN------SURFnet Switch at CERN-
VLAN Path F: FABRIC AMST------NetherLight Open Exchang
VLAN Path G: FABRIC Bristol------London GEANT Open Exc
Notes
a) 1 VLAN per path requested on paths (except as noted in item
b) The VLANS across the NEAAR Link are considered the prima
c) Any VLANs across the ANA link(s) will be secondary or backu
d) for cross connects between ESnet and NEAAR links a range
FNAL
ANL
AMES
CHIC
EQX-CHIC
STAR
BOST
PSFC
LNS
BNL
NEWY
AMST
LOND
ESnet
AOFA
PPPL
CHIC
NASH
ATLA
JLAB
WASH
FABRIC NEWY
WIX
FABRIC WASH
3732
H: 3731
CERN-773
CERN-513
ESNet - CERN Production Peering
CERN Router
GEANT, SURFnet? Router/Switch
CERN
CERN Building 513 (Data Center)
GEANT, SURFnet? Router/Switch
Is it possible to get connections to GEANT and/or NetherLight Exchange at 100G? for Layer 2 connections only?

SC24 WAN-Transport Preliminary Map - August 9, 2024



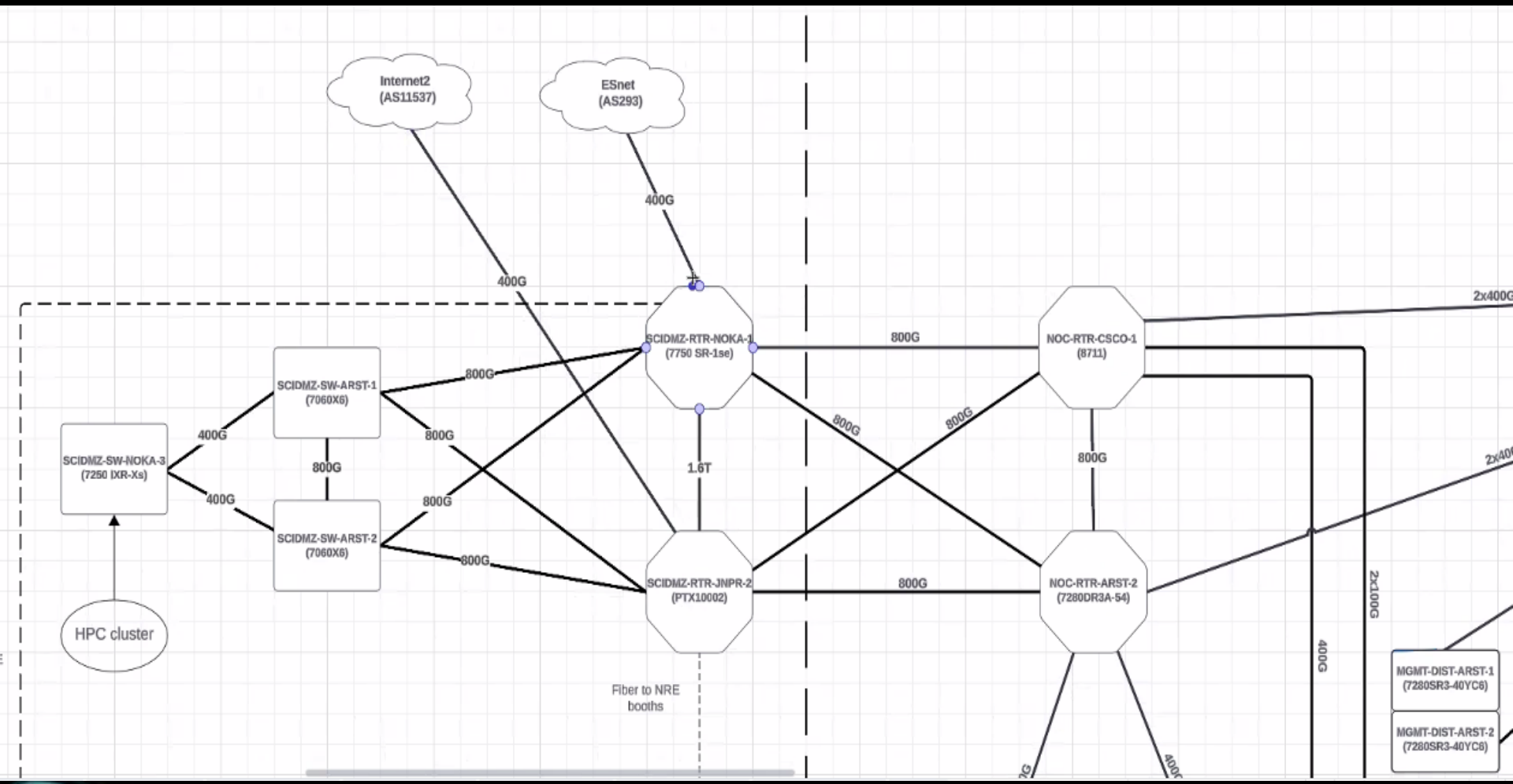
ATLA - ATLA	3.41Tbps
BOST - ATLA	200Gbps
LOSA - ATLA	2Tbps
McLean - ATLA	2.6Tbps
MIAMI - ATLA	500Gbps
NEWY32AOA - ATLA	100Gbps
SEAT - ATLA	300Gbps
STAR - ATLA	1.2Tbps
SUNN - ATLA	400Gbps

SC24 SCinet 2024



STARLIGHT SM

SCinet SC24 Venue 800 Gbps Science DMZ



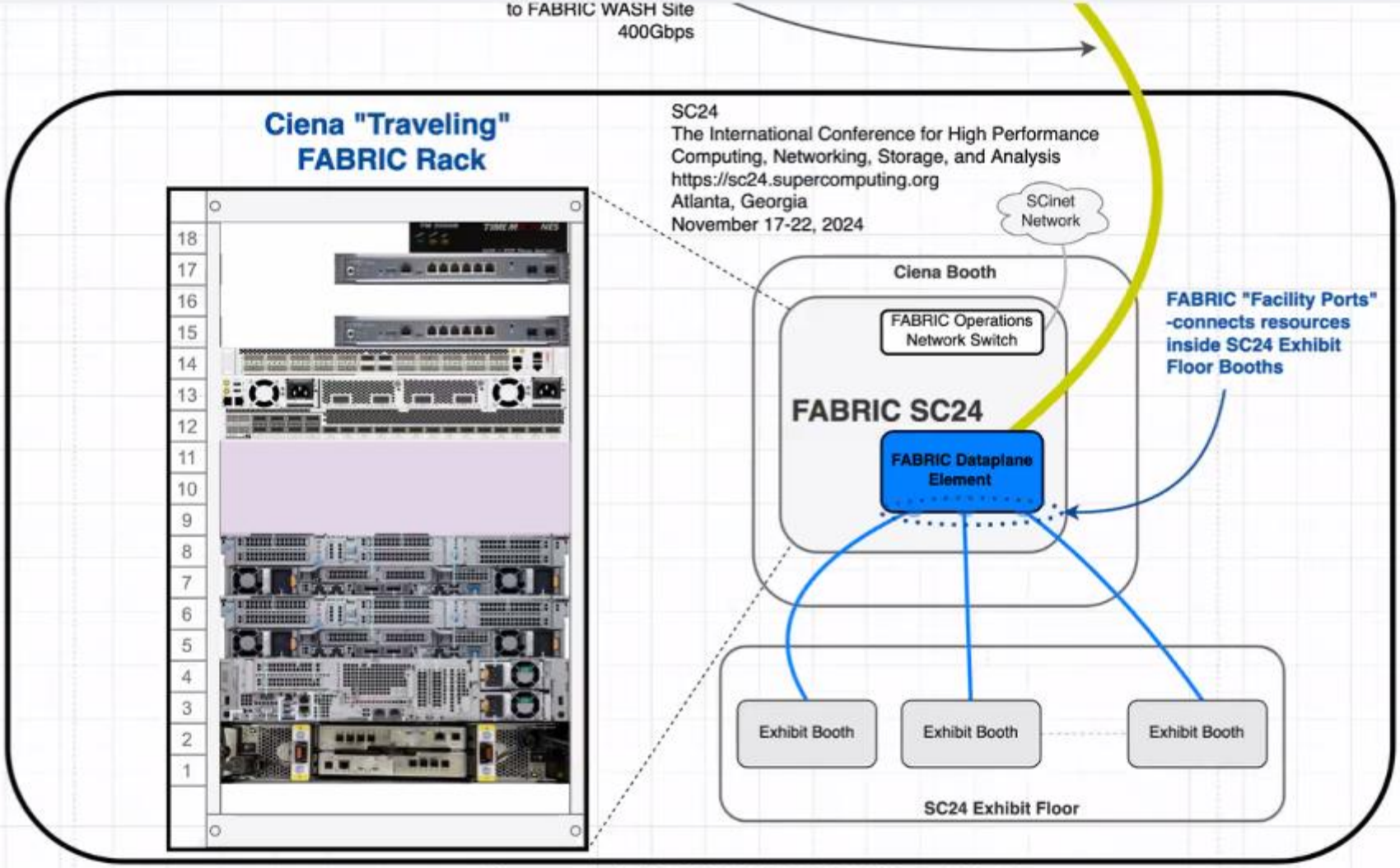
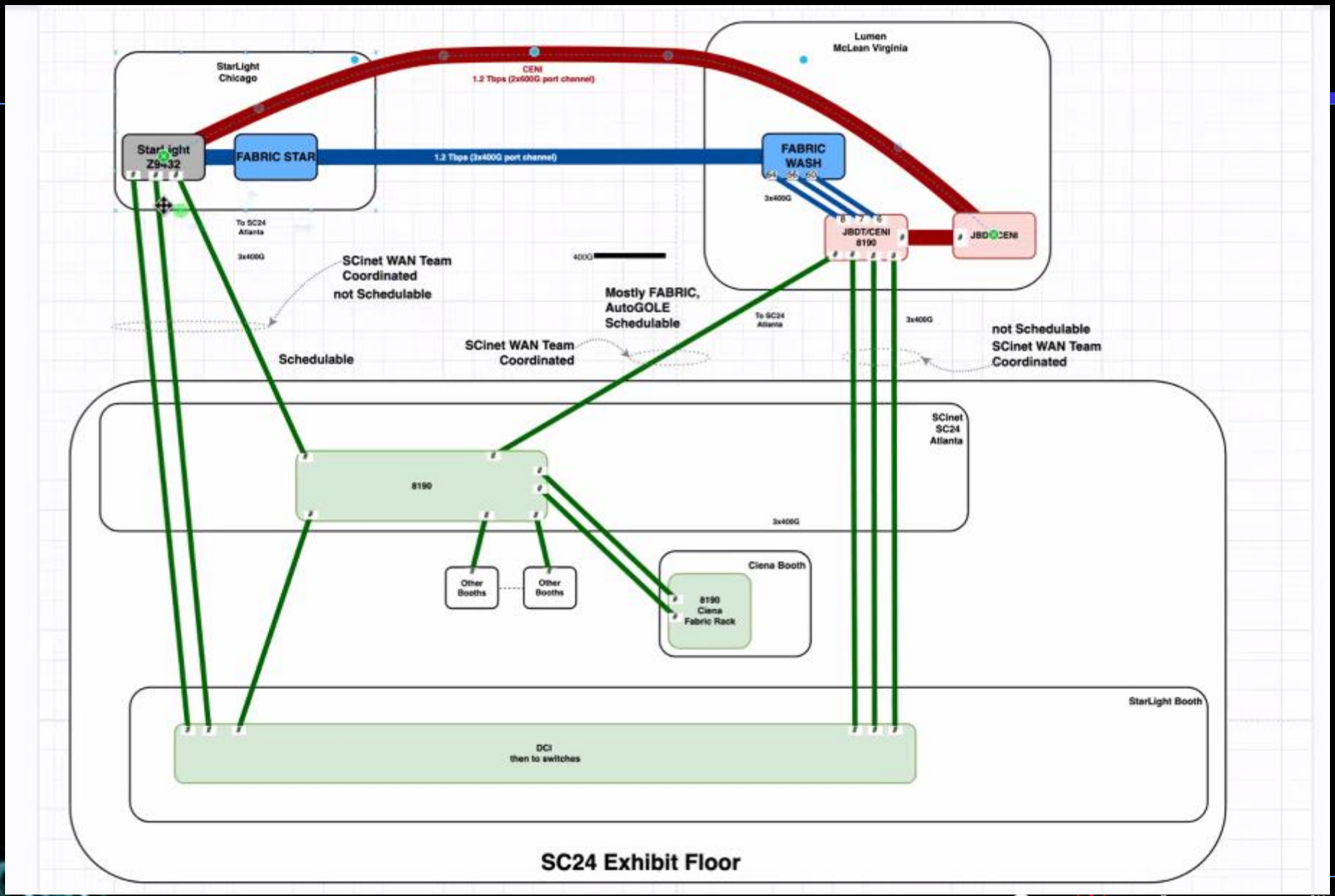
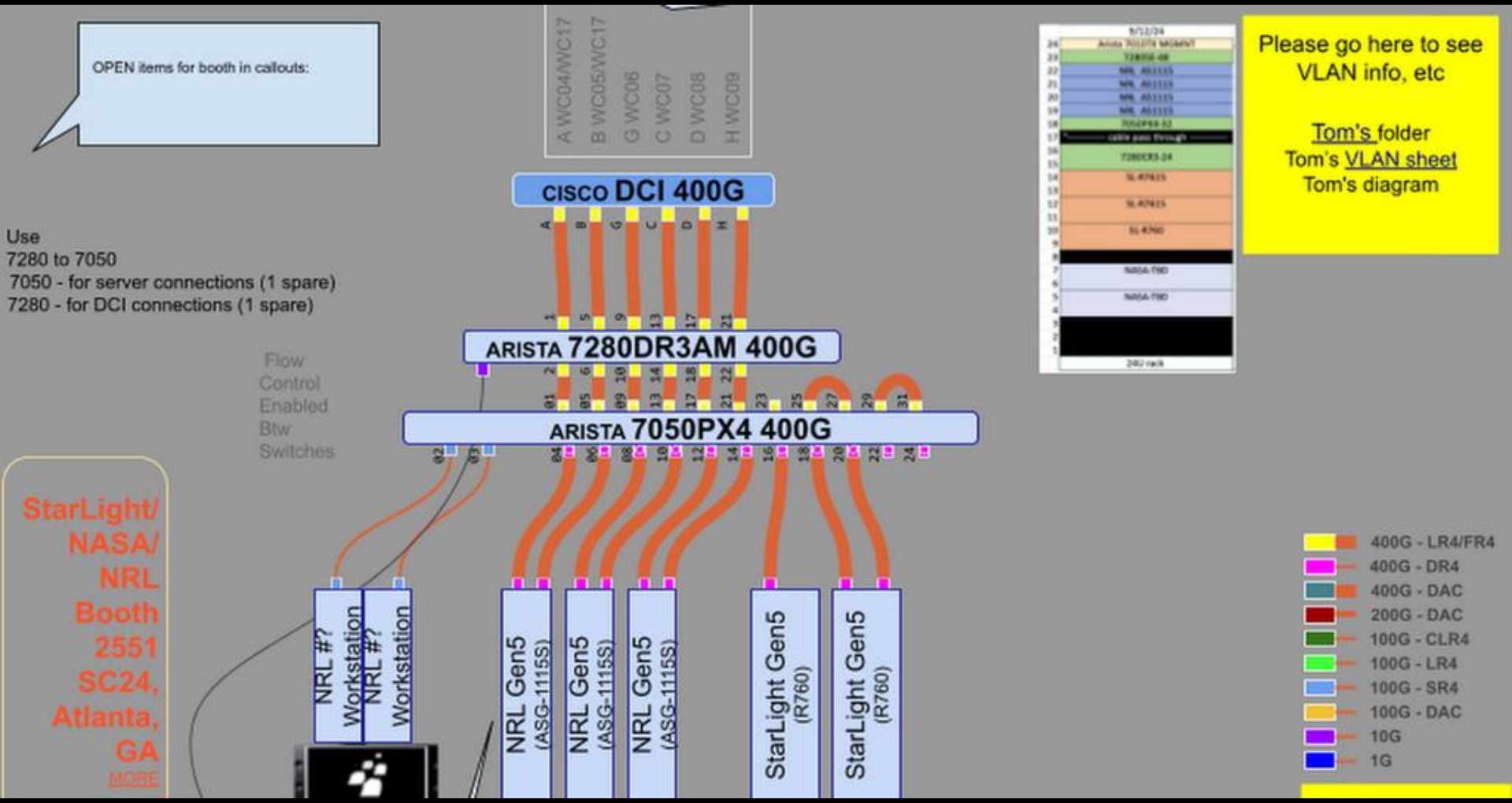


Figure 2 FABRIC SC24 SCinet Node Deployment and WAN Connections

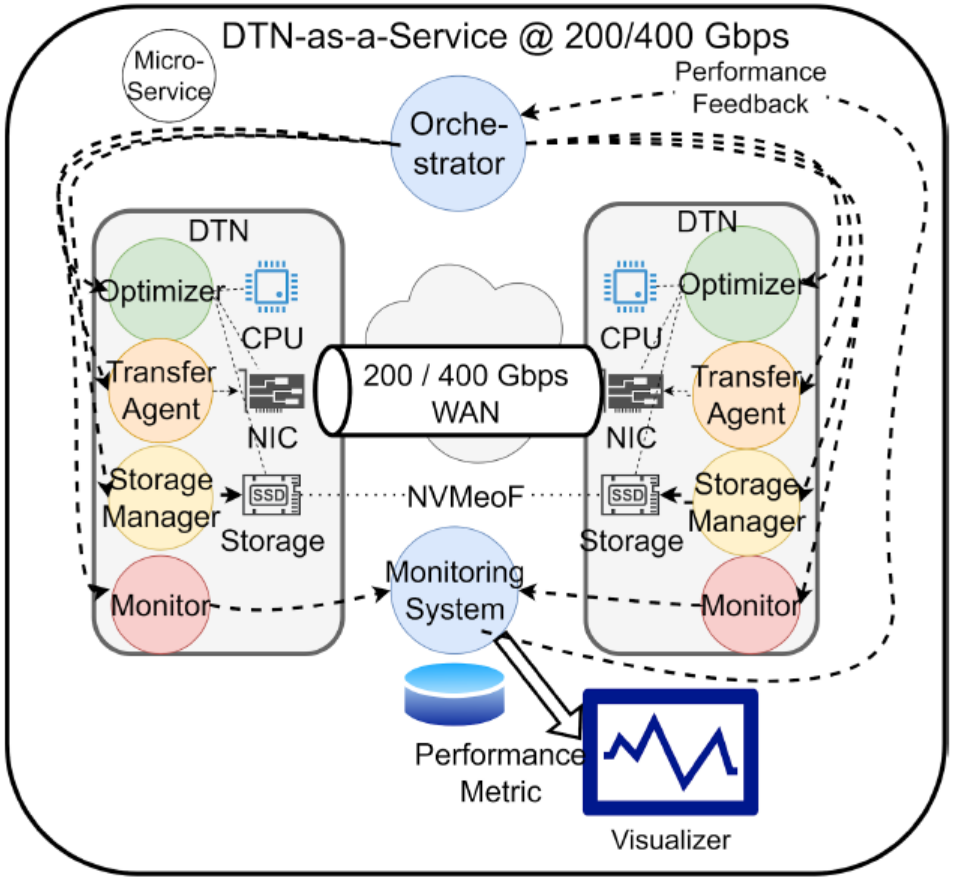

StarLight
Chicago
CENI
1.2 Tbps (2x400G port channel)
Lumen
McLean Virginia
StarLight
Z9-32
FABRIC STAR
1.2 Tbps (3x400G port channel)
FABRIC
WASH
3x400G
To SC24
Atlanta
3x400G
400G
JBDT/CENI
8190
JBD/CENI
SCinet WAN Team
Coordinated
not Schedulable
Mostly FABRIC,
AutoGOLE
Schedulable
To SC24
Atlanta
3x400G
not Schedulable
SCinet WAN Team
Coordinated
Schedulable
SCinet WAN Team
Coordinated
SCinet
SC24
Atlanta
8190
3x400G
Other
Booths
Other
Booths
Ciena Booth
8190
Ciena
Fabric Rack
StarLight Booth
DCI
then to switches
SC24 Exhibit Floor


OPEN items for booth in callouts:
A WC04/WC17
B WC05/WC17
G WC06
C WC07
D WC08
H WC09
CISCO DCI 400G
Use
7280 to 7050
7050 - for server connections (1 spare)
7280 - for DCI connections (1 spare)
ARISTA 7280DR3AM 400G
Flow Control Enabled Btw Switches
ARISTA 7050PX4 400G
Please go here to see VLAN info, etc
Tom's folder
Tom's VLAN sheet
Tom's diagram
StarLight/ NASA/ NRL Booth 2551 SC24, Atlanta, GA
MORE
NRL #? Workstation
NRL #? Workstation
NRL Gen5 (ASG-1115S)
NRL Gen5 (ASG-1115S)
NRL Gen5 (ASG-1115S)
StarLight Gen5 (R760)
StarLight Gen5 (R760)
400G - LR4/FR4
400G - DR4
400G - DAC
200G - DAC
100G - CLR4
100G - LR4
100G - SR4
100G - DAC
10G
1G

Tbps WAN Service R&D

- SDX 1.2 Tbps WAN Services
- SDX 1.2 Tbps WAN Infrastructure
- SDX E2E 800 Gbps WAN Services
- SDX E2E 400 Gbps WAN Services
- 400 Gbps DTNs & Smart NICs
- Software Defined Networking (SDX)
- Middleware
- Protocols
- Hardware and Software Configurations
- Transceivers, Fiber, etc

200/400 Gbps DTN-as-a-Service in High-Performance Research Platform



- 200/400 Gbps end-to-end high-performance data transfer over WAN
- DTN-as-a-Service with microservice architecture, optimizing and transferring using containers
- NVMeoF with streaming support
- Performance monitoring and visualization using opensource platforms (Prometheus, Grafana, and sFlow)

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

中原大學
Chung Yuan Christian University

KANSAI UNIVERSITY

StarLight 400G Real-Time Packet Processing Prototype

400G-100G Packet Header Processor Over WAN With Tofino2 + FPGA

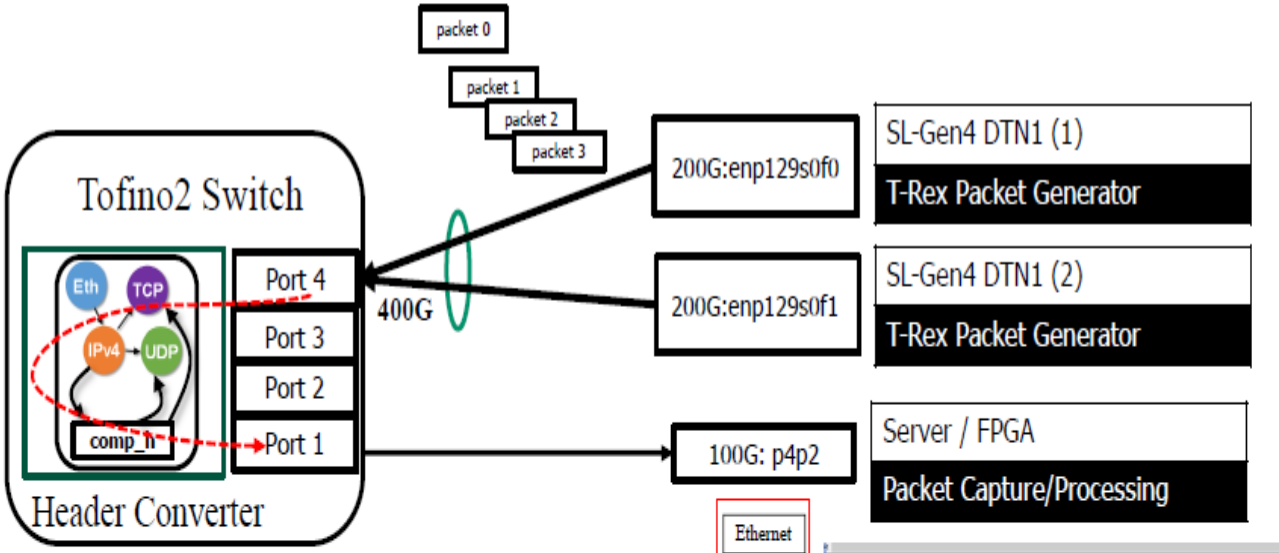
Yu-Kuen Lai, Miura Hideyoshi*, Yi-Siang Lin, Yu-Jui Chang, Kouji Hirata

Department of Electrical Engineering, Chung-Yuan Christian University, Chungli, Taiwan

*Graduate School of Science and Engineering, Kansai University, Japan

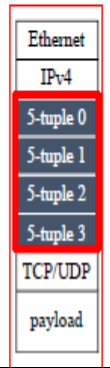
Jim Hao Chen, Joe Mambretti

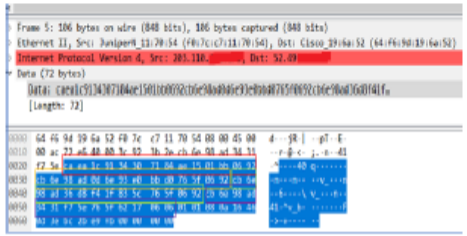
International Center for Advanced Internet Research Northwestern University, USA



```
header comp_h {
    bit<32> src_addr;
    bit<32> dst_addr;
    bit<16> src_port;
    bit<16> dst_port;
    bit<8> protocol;
    bit<8> comp_type;
}
```

```
struct switch_header_t {
    ethernet_t   ethernet;
    ipv4_t       ipv4;
    comp_h       comp_0;
    comp_h       comp_1;
    comp_h       comp_2;
    comp_h       comp_3;
    tcp_t        tcp;
    udp_t        udp;
}
```





iCAIR

STARLIGHT SDX

StarLight DTN 400G TCP WAN Service Prototypes

25 TCP streams between SL IntelG5 and SL AMDG5 Servers

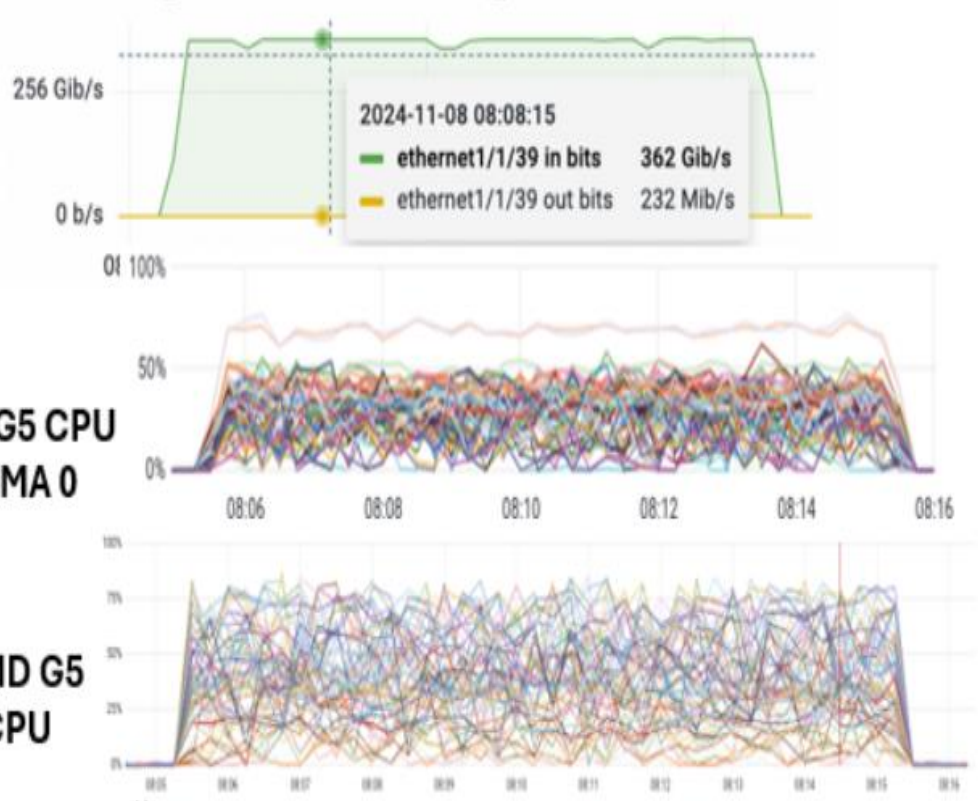
Nov 8 2024 Test Results

L.A. loop (NA-REX) RTT 89.6 ms

Avg: 378Gb/s Range: 376-382 Gb/s

McLean loop (CENI) RTT 27.8 ms

Avg: 386Gb/s Range: 381-391 Gb/s



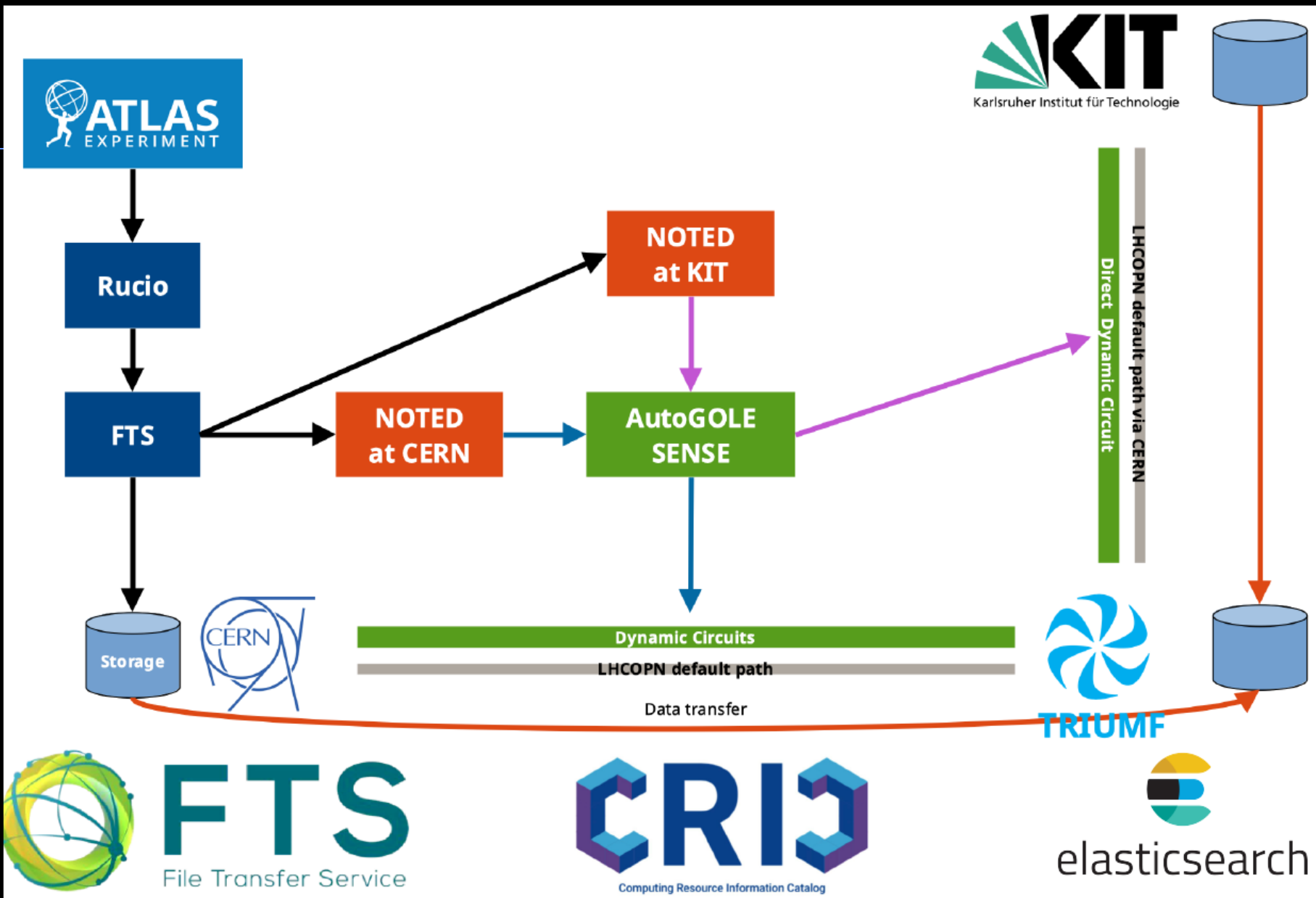
Intel G5 CPU NUMA 0

AMD G5 CPU

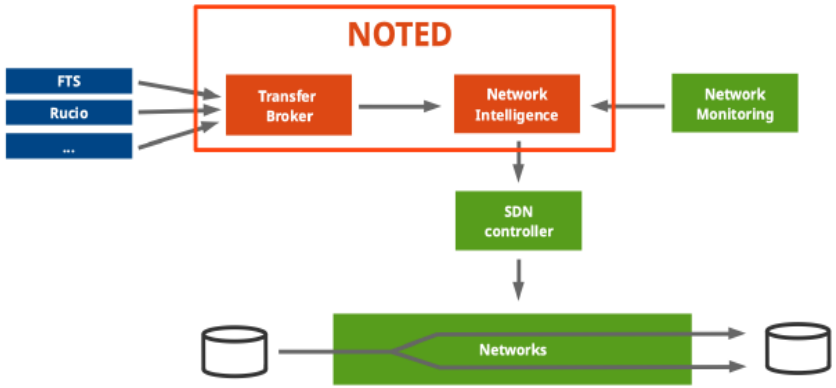
iCAIR

國立成功大學 National Cheng Kung University

STARLIGHT SDX


ATLAS EXPERIMENT
Rucio
FTS
NOTED at KIT
NOTED at CERN
AutoGOLE SENSE
Direct Dynamic Circuit
LHCOPN default path via CERN
Storage
CERN
Dynamic Circuits
LHCOPN default path
Data transfer
KIT
Karlsruher Institut für Technologie
TRIUMF
FTS
File Transfer Service
CRIC
Computing Resource Information Catalog
elasticsearch

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 be congested anymore → cancel the dynamic circuit and the traffic is routed back.

StarLight Science Data Exchange Service Prototypes

GEN3 https://gen3.org/

- Gen3 is a **data platform** for building **data commons and data ecosystems** .It consists of several **open-source software services.**
- It support the **healthy data ecosystems** by enabling the interoperation and creation of **cloud-based data resources**, including data commons and analysis workspaces.
- *Powered by CTDS, University of Chicago*
  - *http://ctds.uchicago.edu*
- *Apache-2.0 license*

scitags https://scitags.org/

- **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 utilized while at the same time providing feedback to the *scientific community* on what network flows and patterns are critical for their computing.
- 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.

iCAIR

STARLIGHT SDX

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.

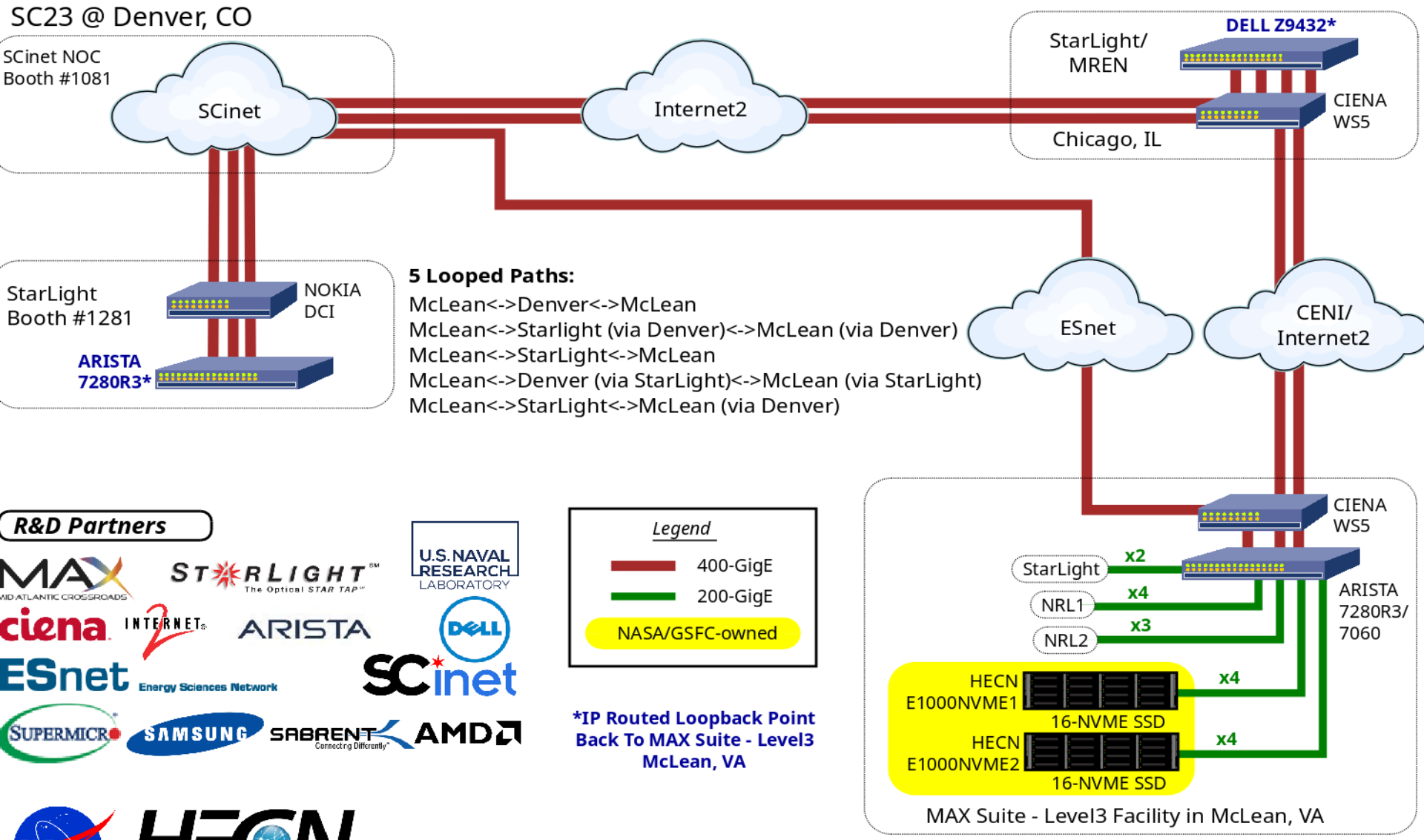
Our approach is based on a network tagging mechanism that marks network packets and/or network flows using the science domain and activity fields. These tags can then be captured by the *R&E providers* and correlated with their existing netflow data to better understand existing network patterns, estimate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

SC23 Joint Big Data Testbed

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

An SC23 Collaborative Initiative Among NASA and Several Partners



NASA

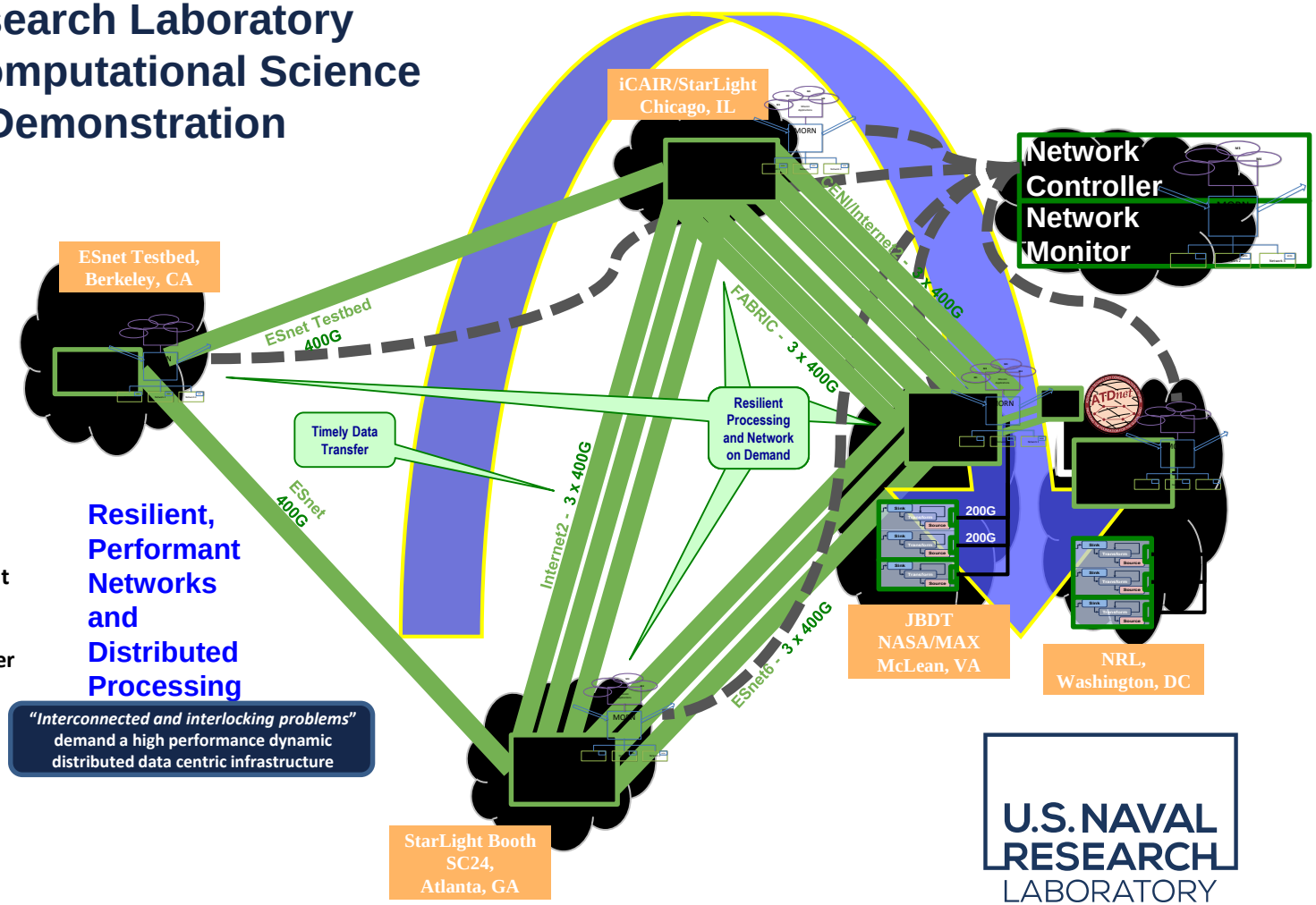
HECN
High End Computer Networking

In memory of Paul Lang and Pat Gary

NASA/GSFC High End Computer Networking (HECN) Team
Diagram by Bill Fink - 10/13/2023

Naval Research Laboratory
Center for Computational Science
SC24 Demonstration

2. Tbps RDMA data movement over global distance for timely Terabyte transfers (goal << 1 min Tbyte transfer on N by 400G network).



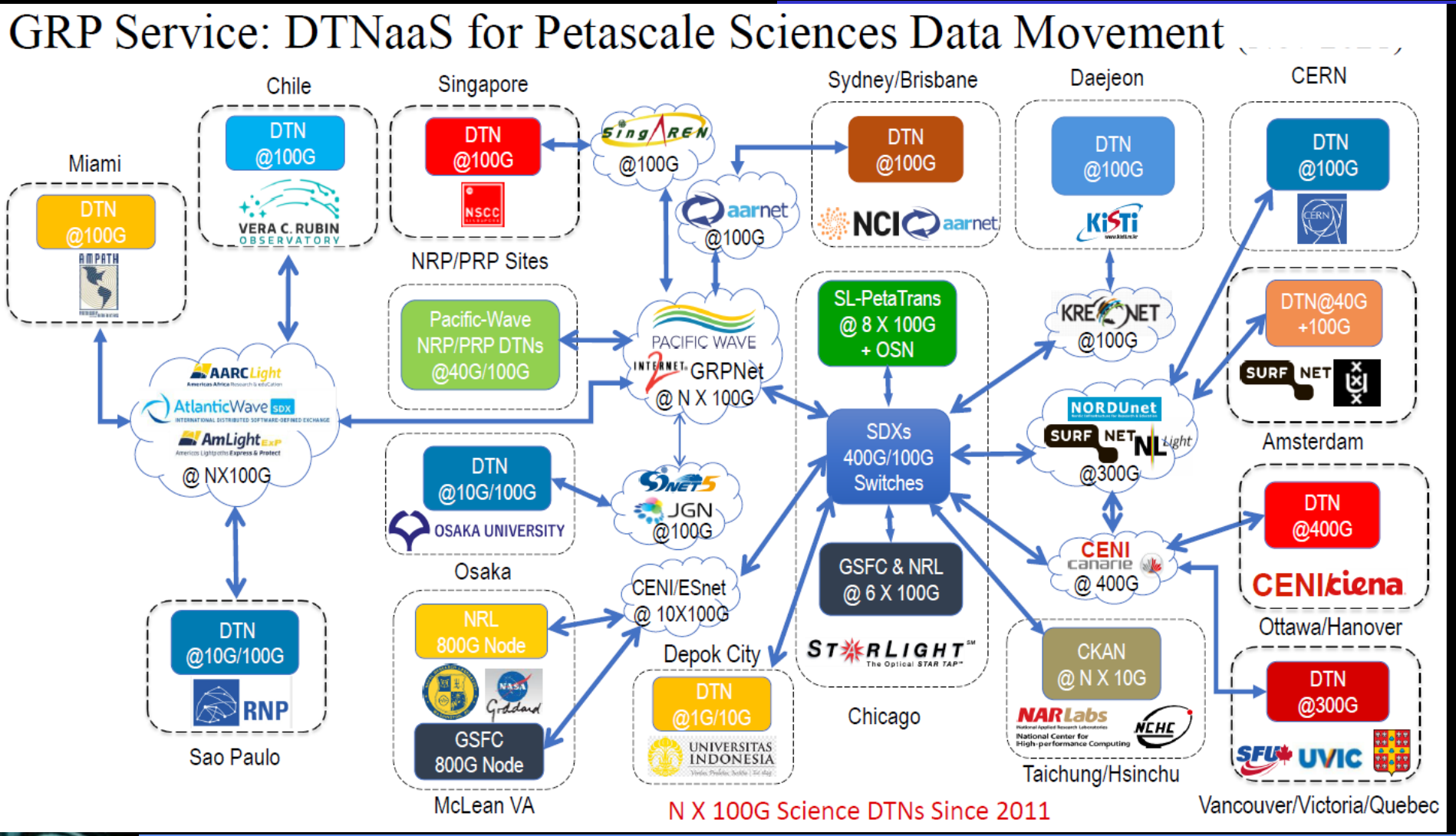
Resilient, Performant Networks and Distributed Processing

"Interconnected and interlocking problems" demand a high performance dynamic distributed data centric infrastructure

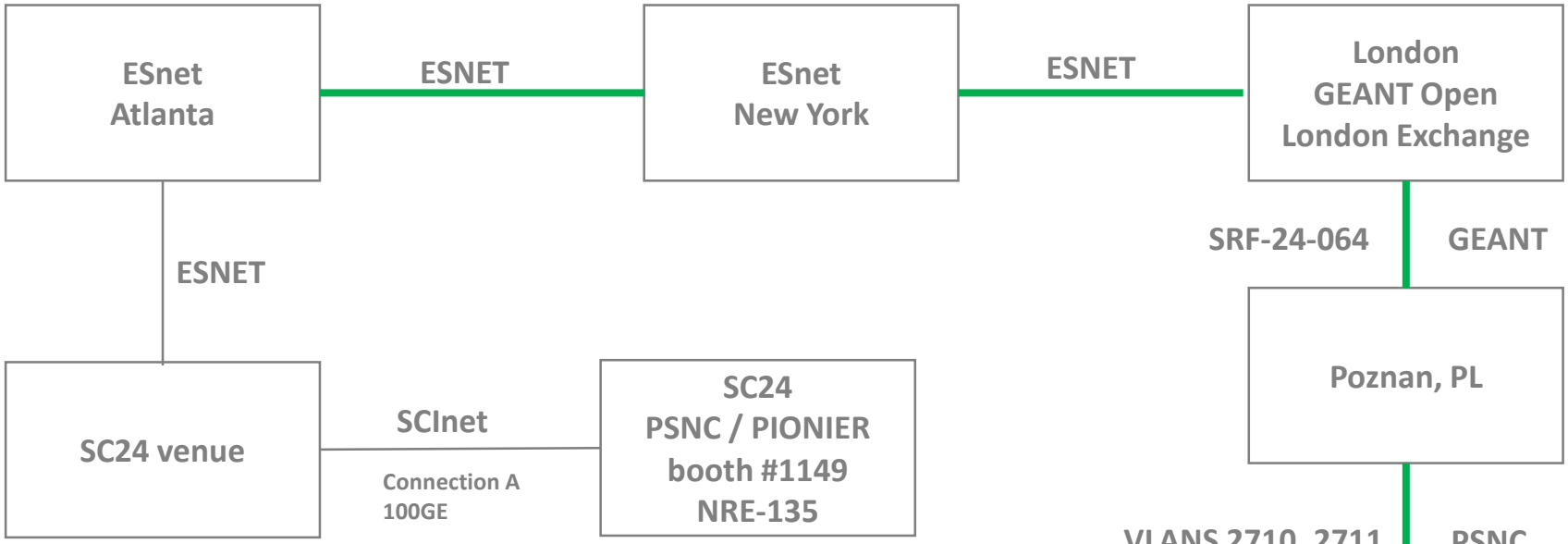
.

U.S. NAVAL RESEARCH LABORATORY

GRP DTNaaS For Petascale Science



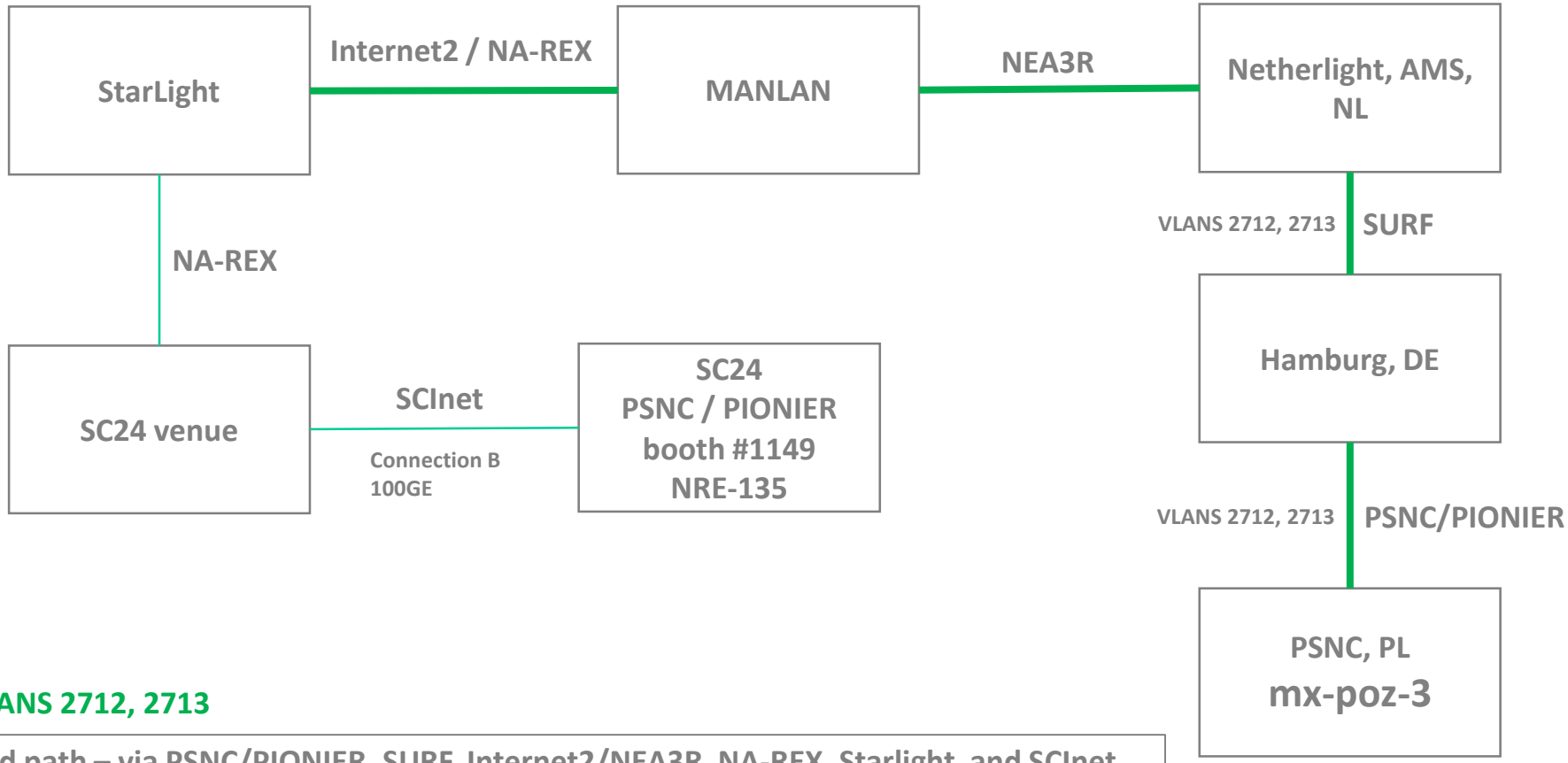
Quantum Communications NRE



VLANS 2710, 2711

1st path – via PSNC/PIONIER, GEANT, ESNET and SCInet

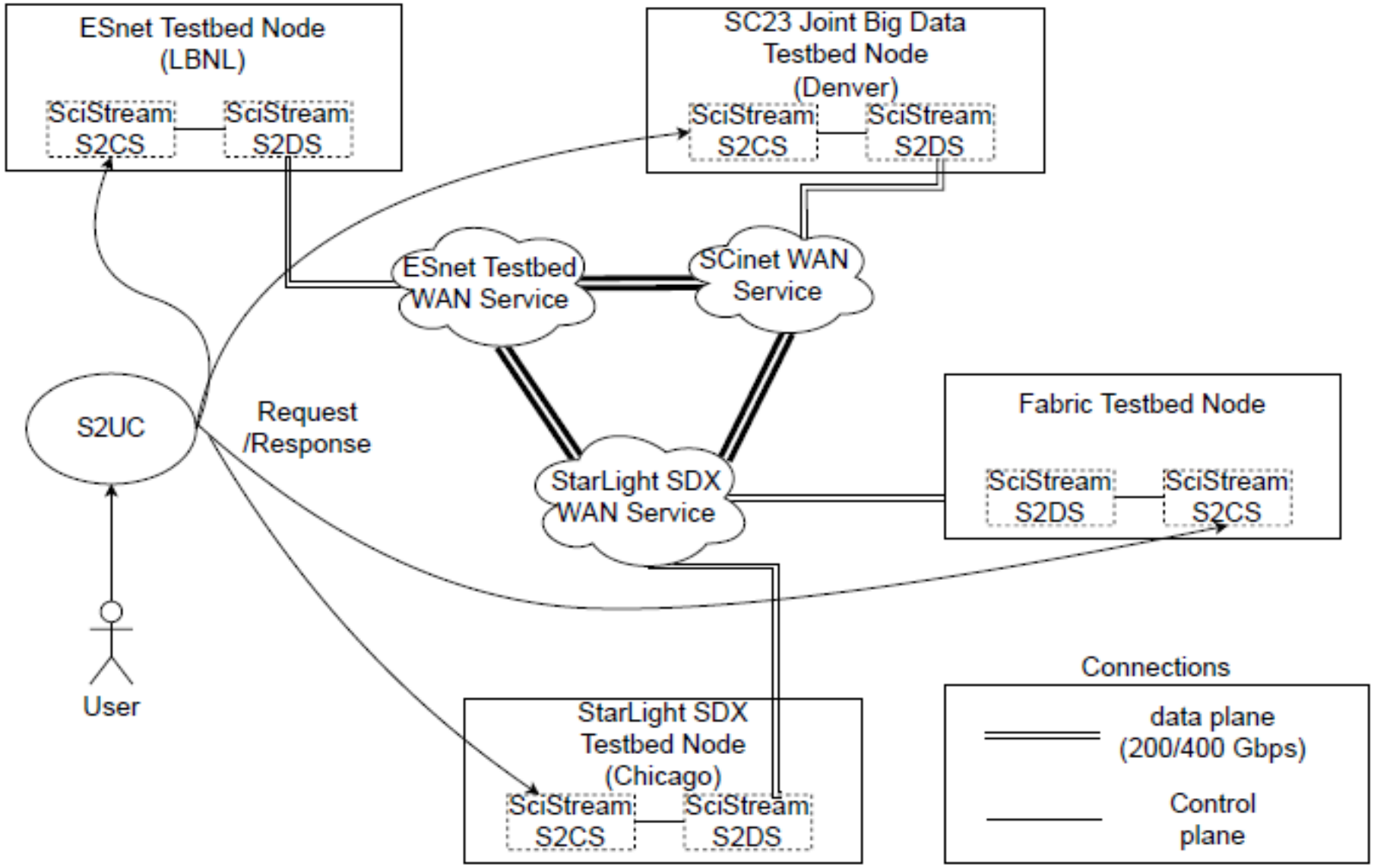
Quantum Communications NRE



VLANS 2712, 2713

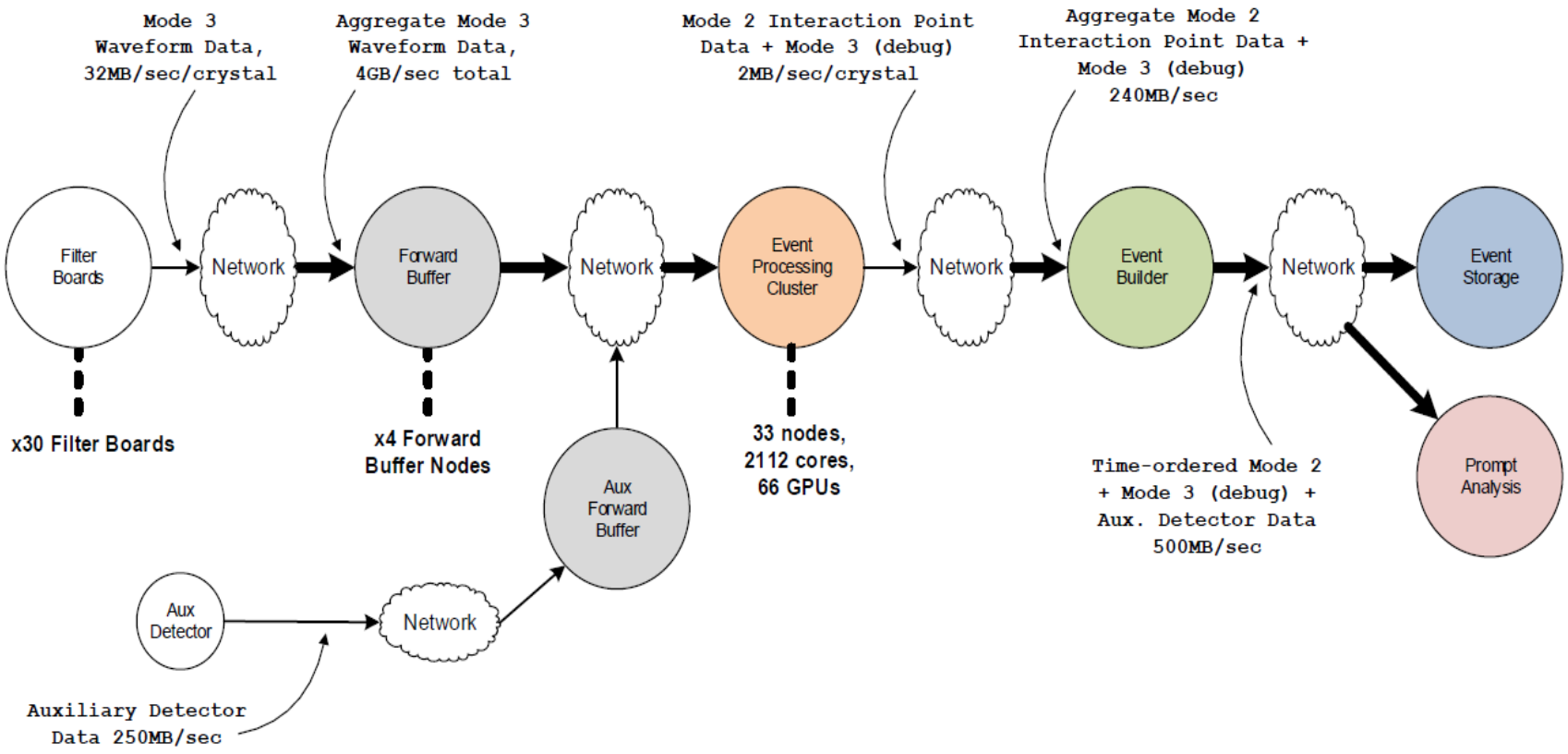
2nd path – via PSNC/PIONIER, SURF, Internet2/NEA3R, NA-REX, Starlight and SCInet

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



ESnet Gamma Ray Energy Tracking Array (GRETA) Prototype

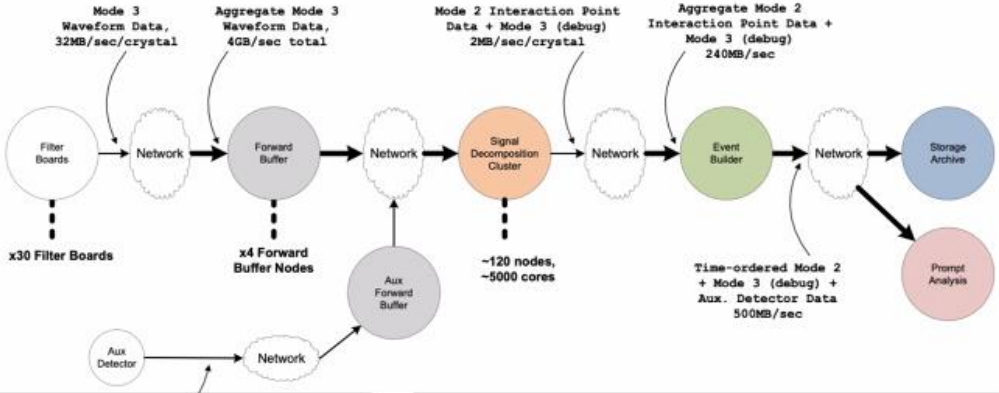
GRETA Data Pipeline



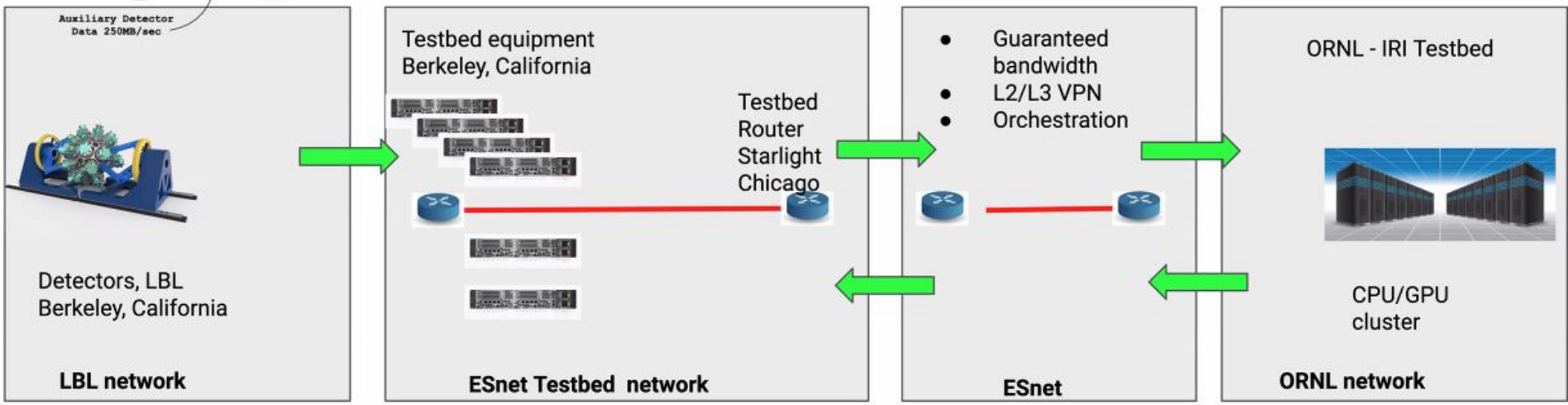
Source: ESnet Testbed

STARLIGHT SM

GRETA to DELERIA



Impact: scaling up online processing using user facilities



Source: ESnet Testbed

STARLIGHT SM

StarLight Is Provding Support For Multiple Other NREs, e.g.,

Revisiting the Science DMZ

SCinet Network Research Exhibition

Corey Eichelberger, Ezra Kissel, George Robb, Jason Zurawski

Goal

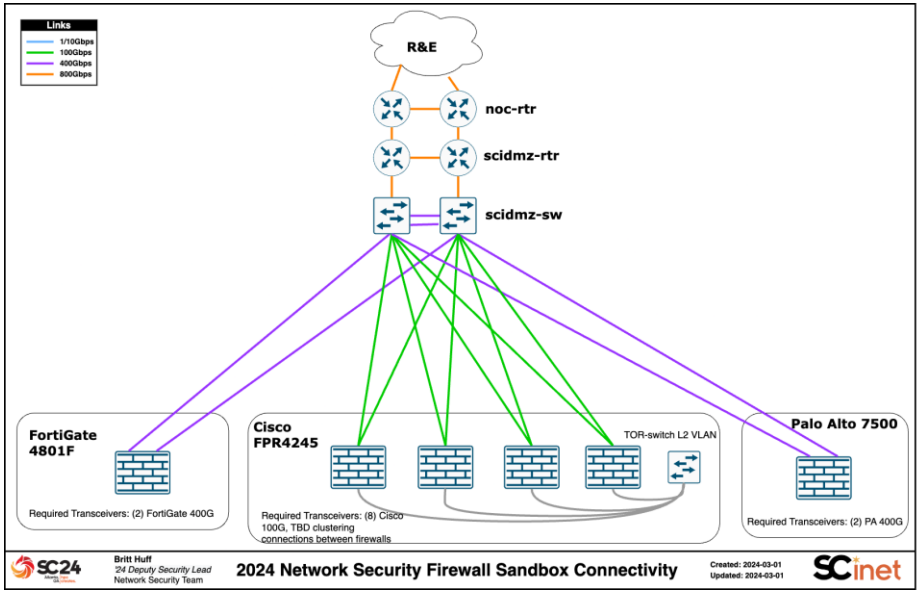
Understand the causes and impact of data movement friction in modern cyberinfrastructure.

Approach

1. **Test data mobility with and without multi-vendor network security devices.**
2. **Model performance expectations for secure network transfers.**
3. **Incorporate detailed telemetry and monitoring capabilities.**

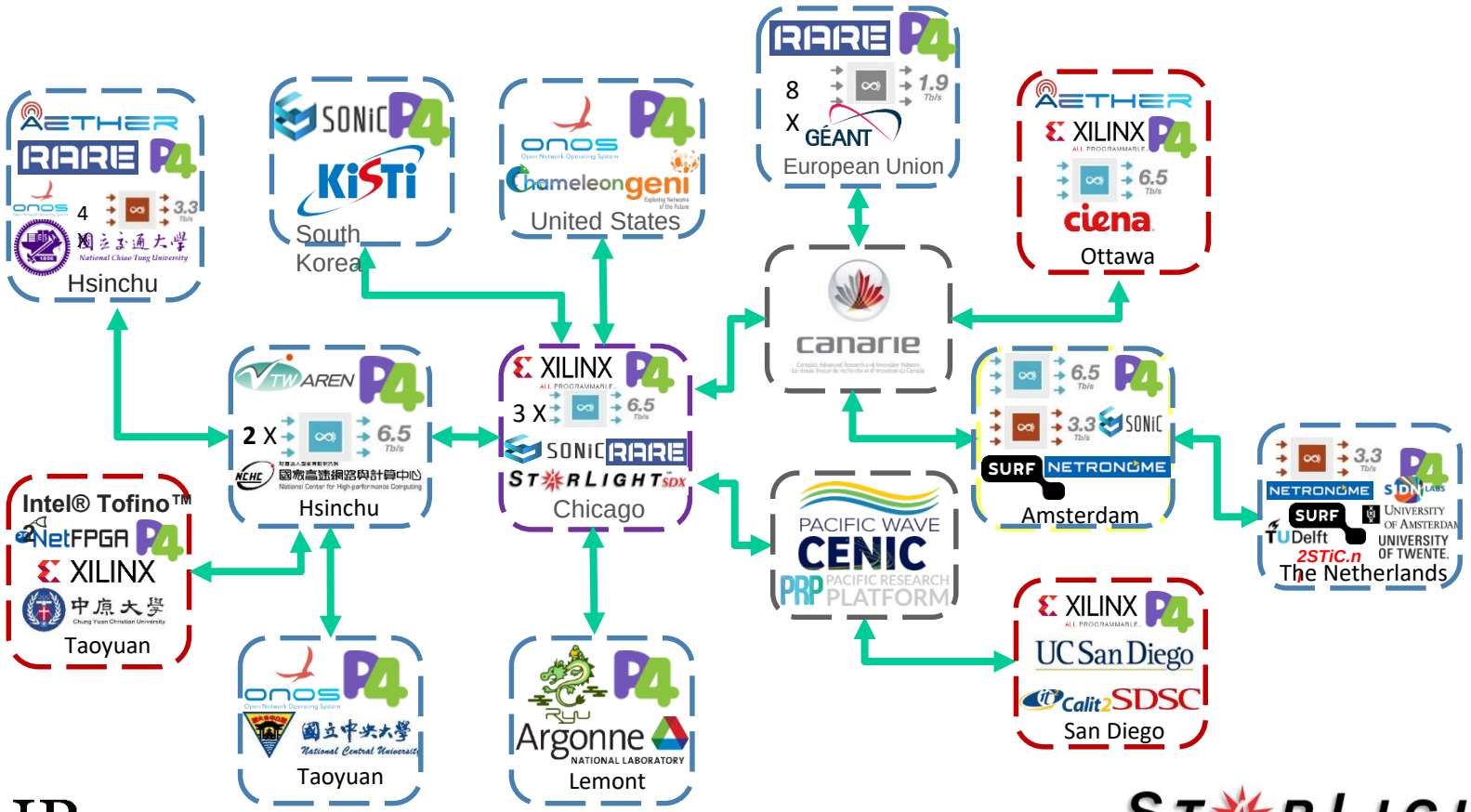
Expected Outcomes

- **Develop and publish documentation.**
- **Update best practices.**



International P4 Testbed Showcase at SC23

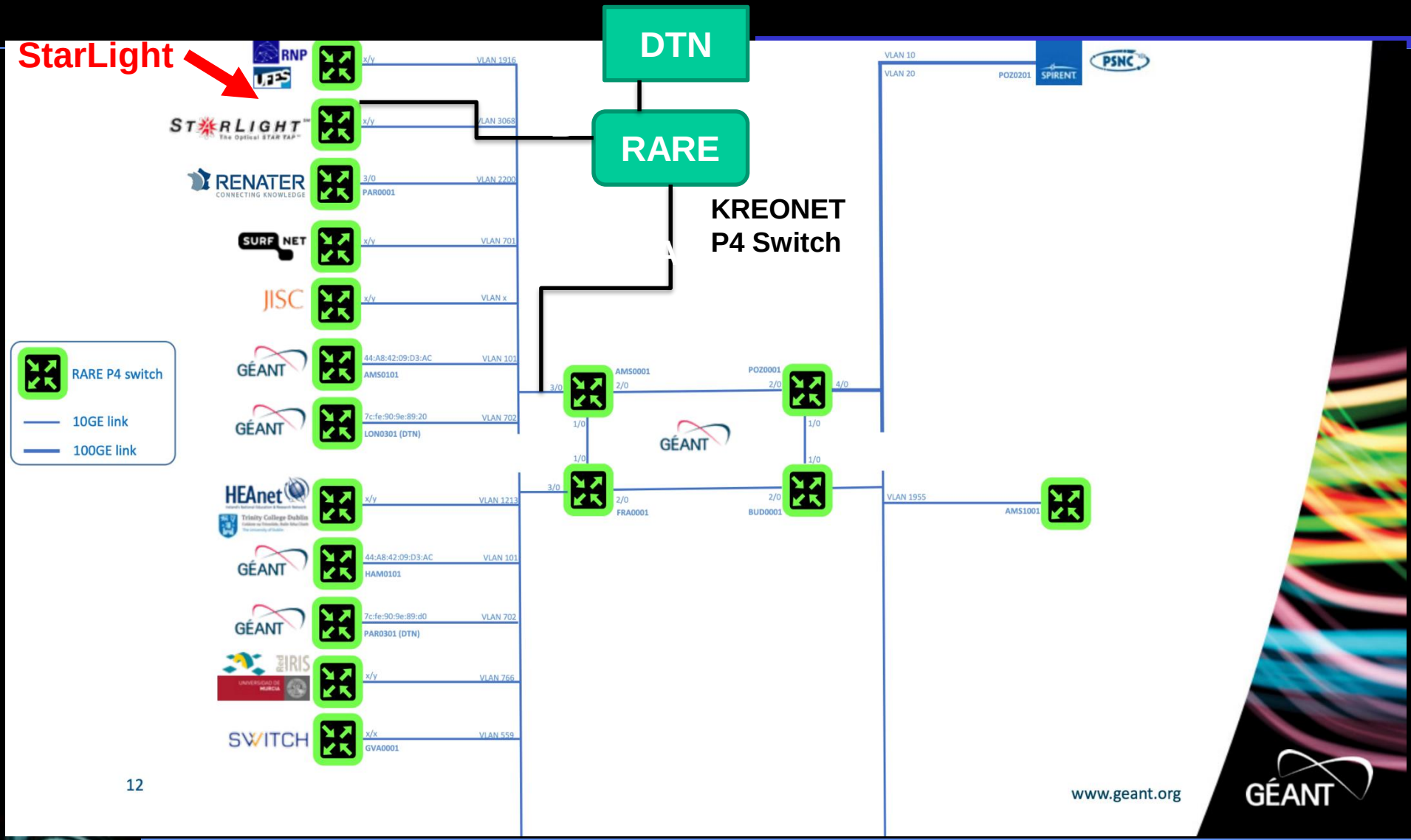
GRP Service: International P4 Experimental Networks (iP4EN)



iCAIR

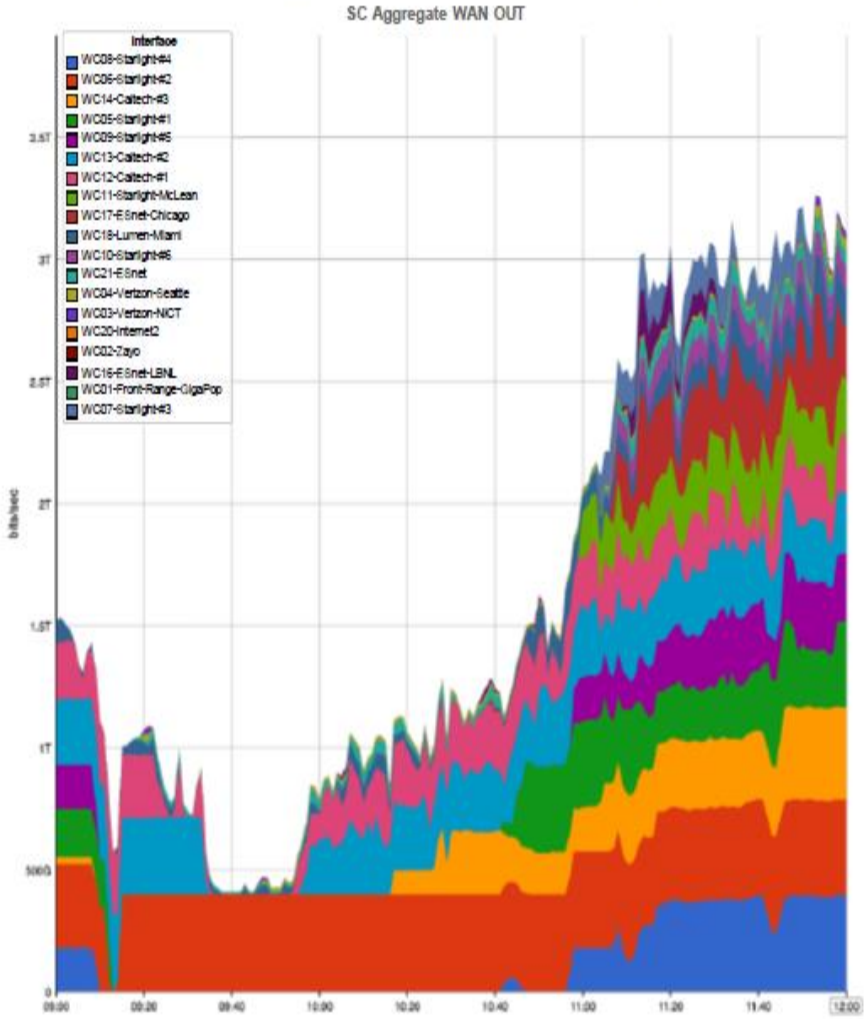
STARLIGHT SDX

Integration With GEANT P4 Testbed

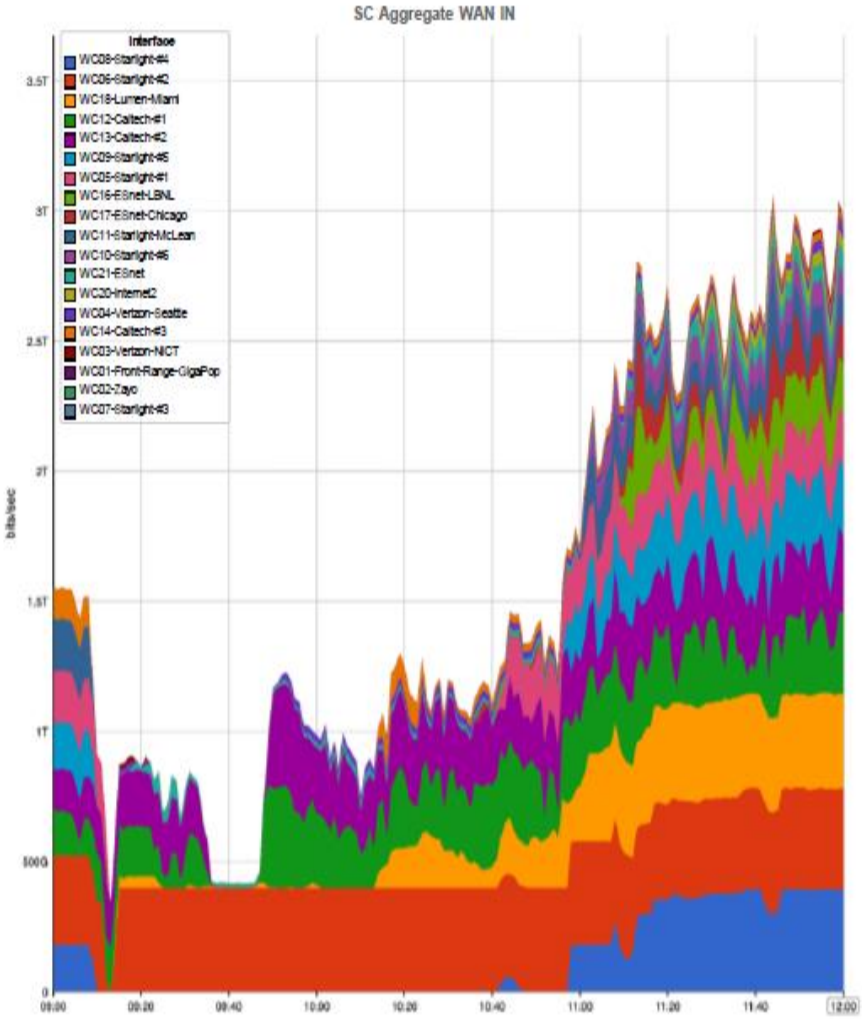


SC23 Bandwidth Challenge

StarLight contributes 4 of Top 5



StarLight contributes 2 of Top 5



Chameleon

www.chameleoncloud.org

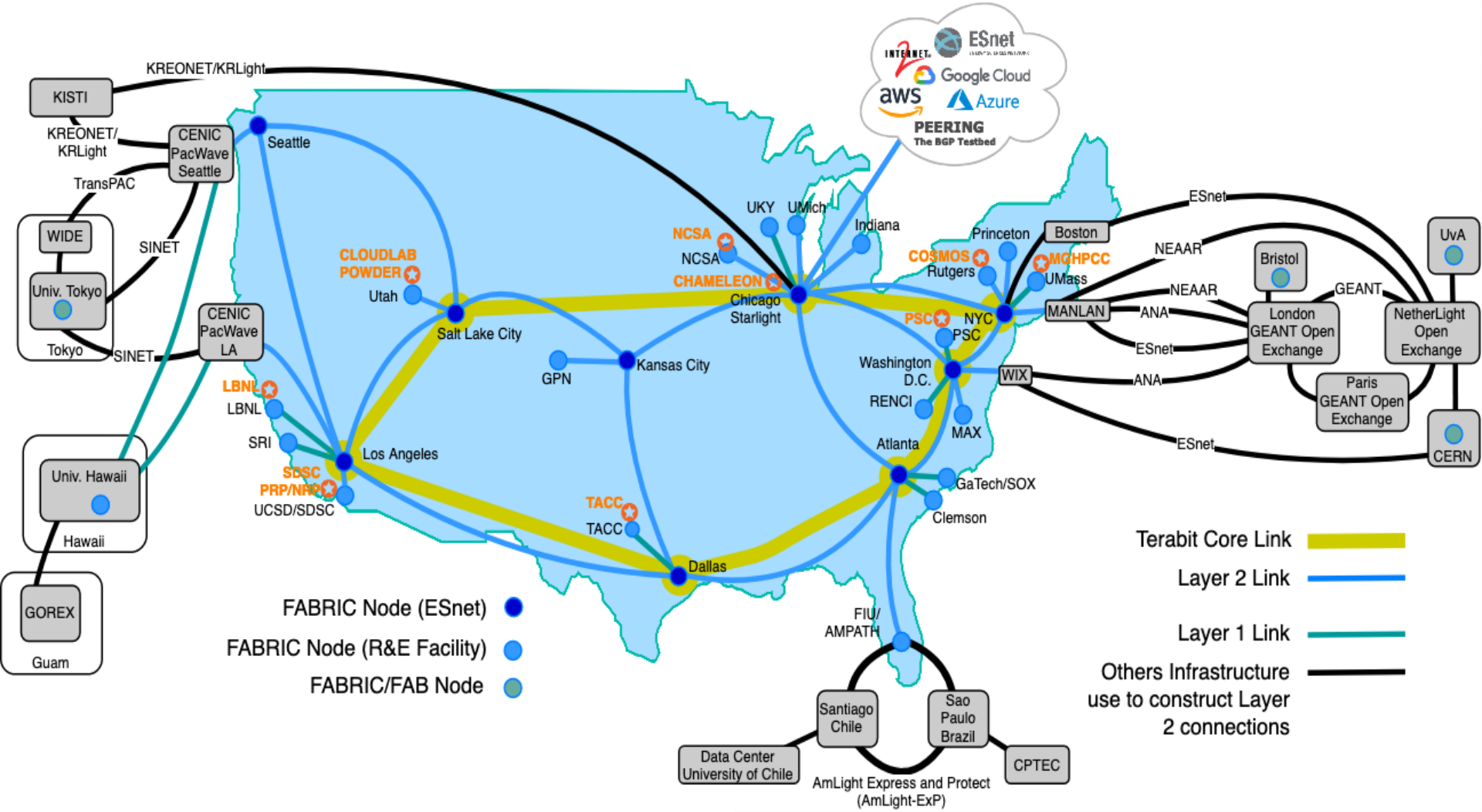
CHAMELEON: A LARGE SCALE, RECONFIGURABLE EXPERIMENTAL INSTRUMENT FOR COMPUTER SCIENCE

Kate Keahey

Joe Mambretti, Pierre Riteau, Paul Ruth, Dan Stanzione

THE UNIVERSITY OF CHICAGO TACC renci NORTHWESTERN UNIVERSITY NSF

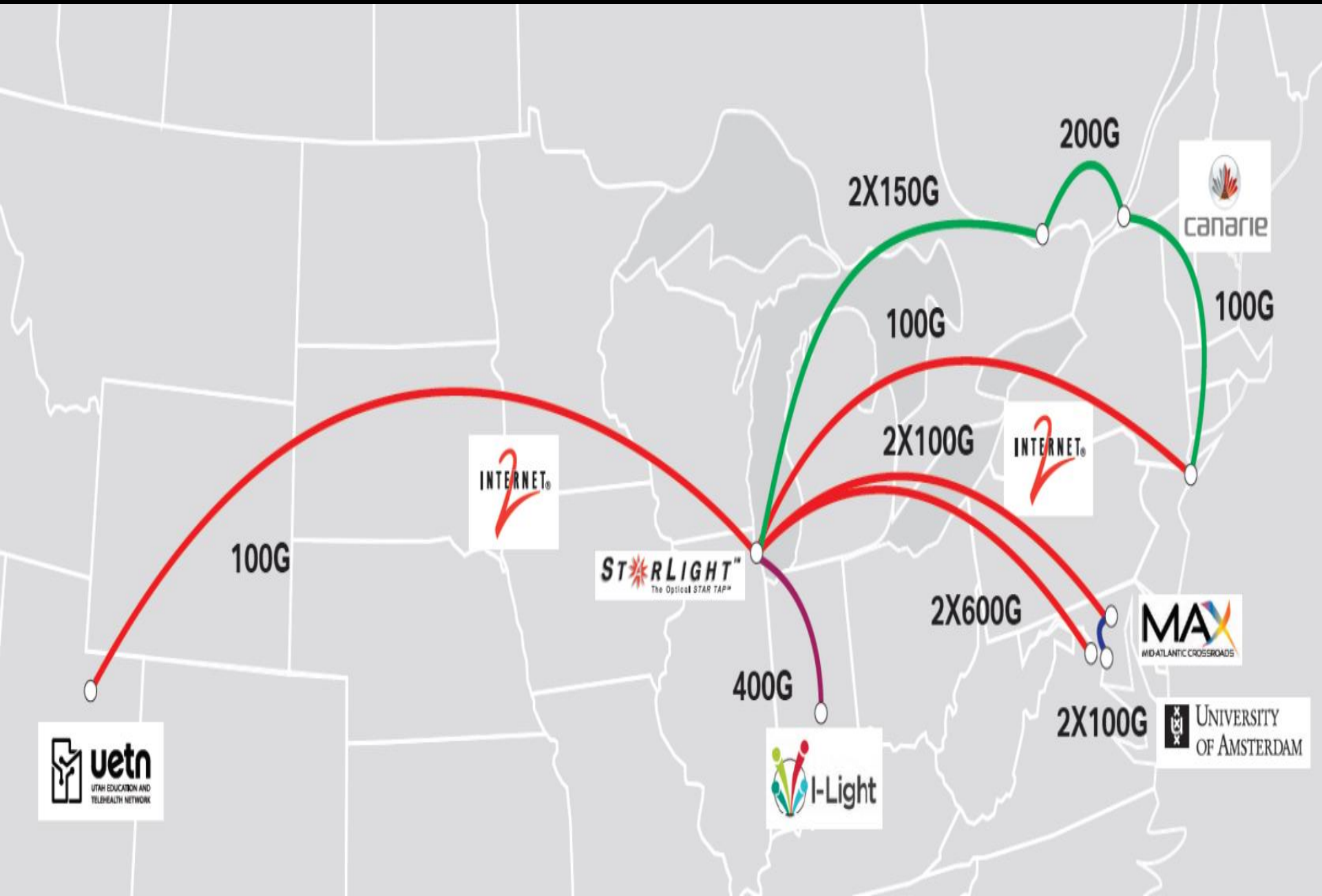
Paul Ruth PI, RENCI: FABRIC



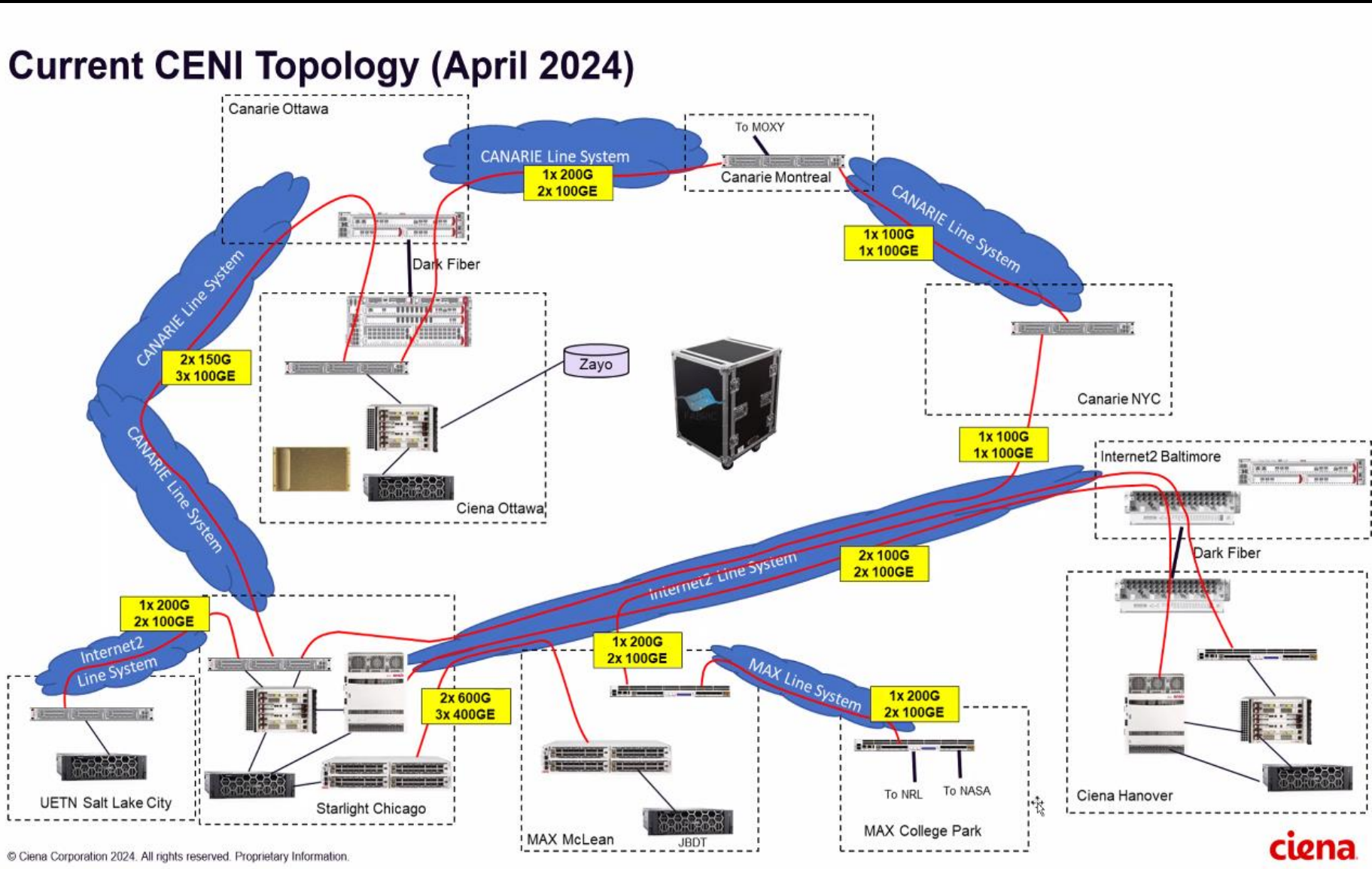
Core = 1.2 Tbps

STARLIGHT℠

Ciena CENI


200G
2X150G
canarie
100G
100G
2X100G
INTERNET2
INTERNET2
100G
STARLIGHT
The Optical STAR TAP
2X600G
MAX
MID-ATLANTIC CROSSROADS
400G
2X100G
UNIVERSITY OF AMSTERDAM
uetn
UTAH EDUCATION AND TELEHEALTH NETWORK
I-Light

Ciena CENI

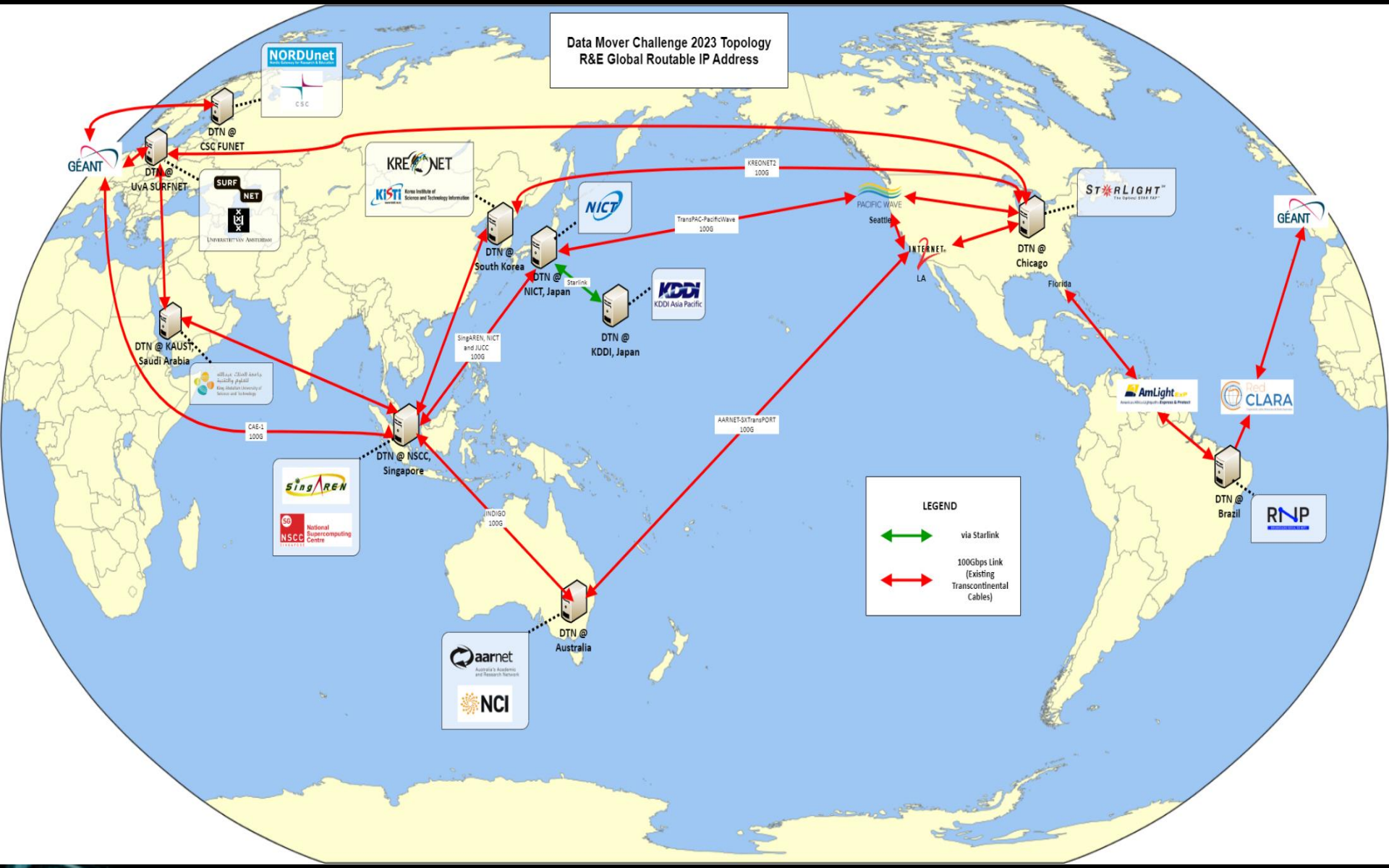


STARLIGHT℠

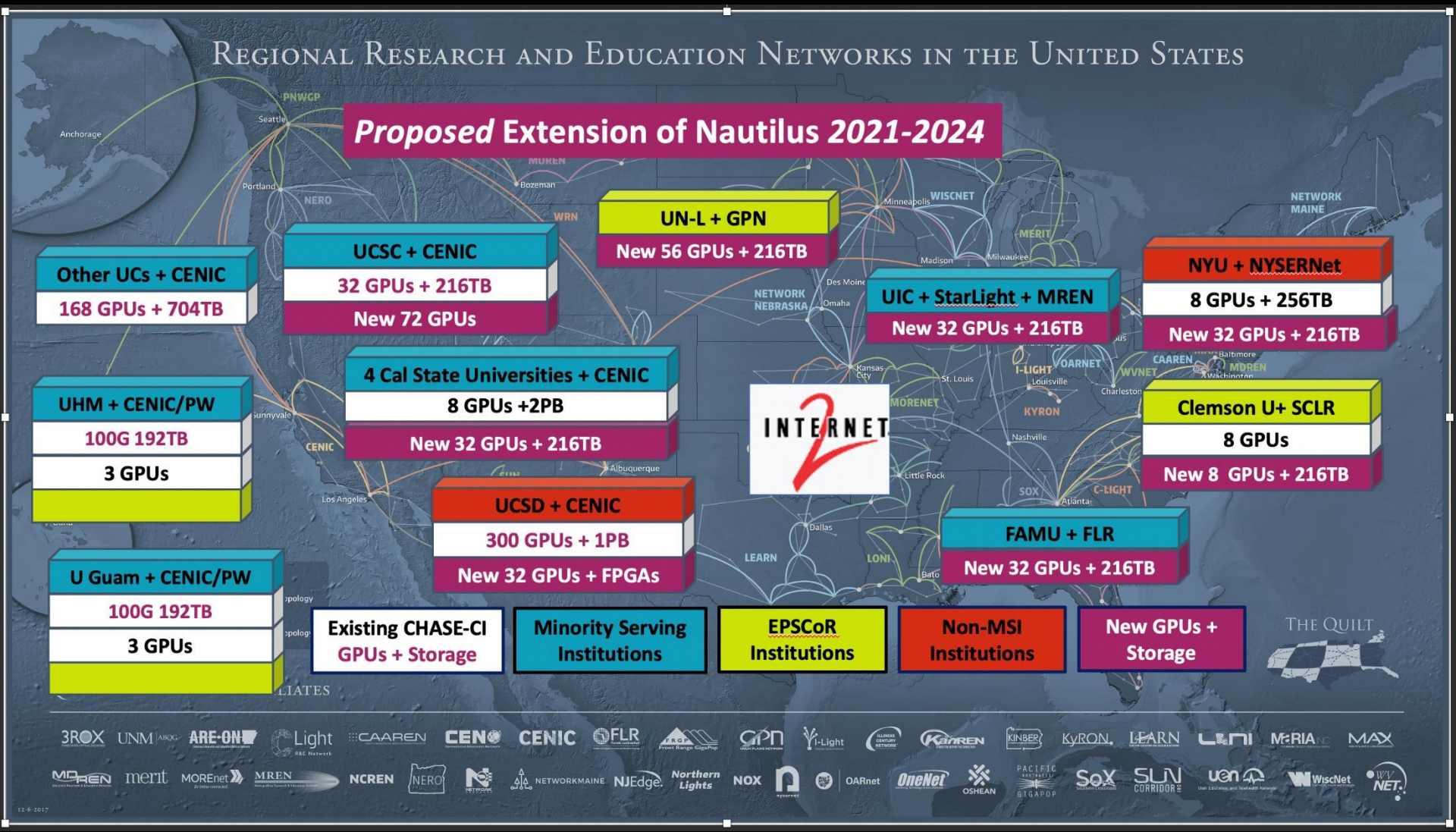
FabFed

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 2023 Topology
R&E Global Routable IP Address
NORDUnet
DTN @ CSC FUNET
GÉANT
DTN @ UvA SURFNET
SURFNET
KREONET
KISTI
DTN @ South Korea
NICT
DTN @ NICT, Japan
Starlink
DTN @ KDDI, Japan
KDDI Asia Pacific
KREONET2 100G
TransPAC-PacificWave 100G
PACIFIC WAVE
Seattle
INTERNET2
LA
DTN @ Chicago
STARLIGHT
Florida
GÉANT
DTN @ KAUST, Saudi Arabia
CAE-1 100G
SingAREN, NICT and JUCC 100G
DTN @ NSCC, Singapore
SingAREN
NSCC National Supercomputing Centre
INDIGO 100G
AARNET-SXTransPORT 100G
AmLight
RedCLARA
DTN @ Brazil
RNP
DTN @ Australia
aarnet
NCI
LEGEND
via Starlink
100Gbps Link (Existing Transcontinental Cables)

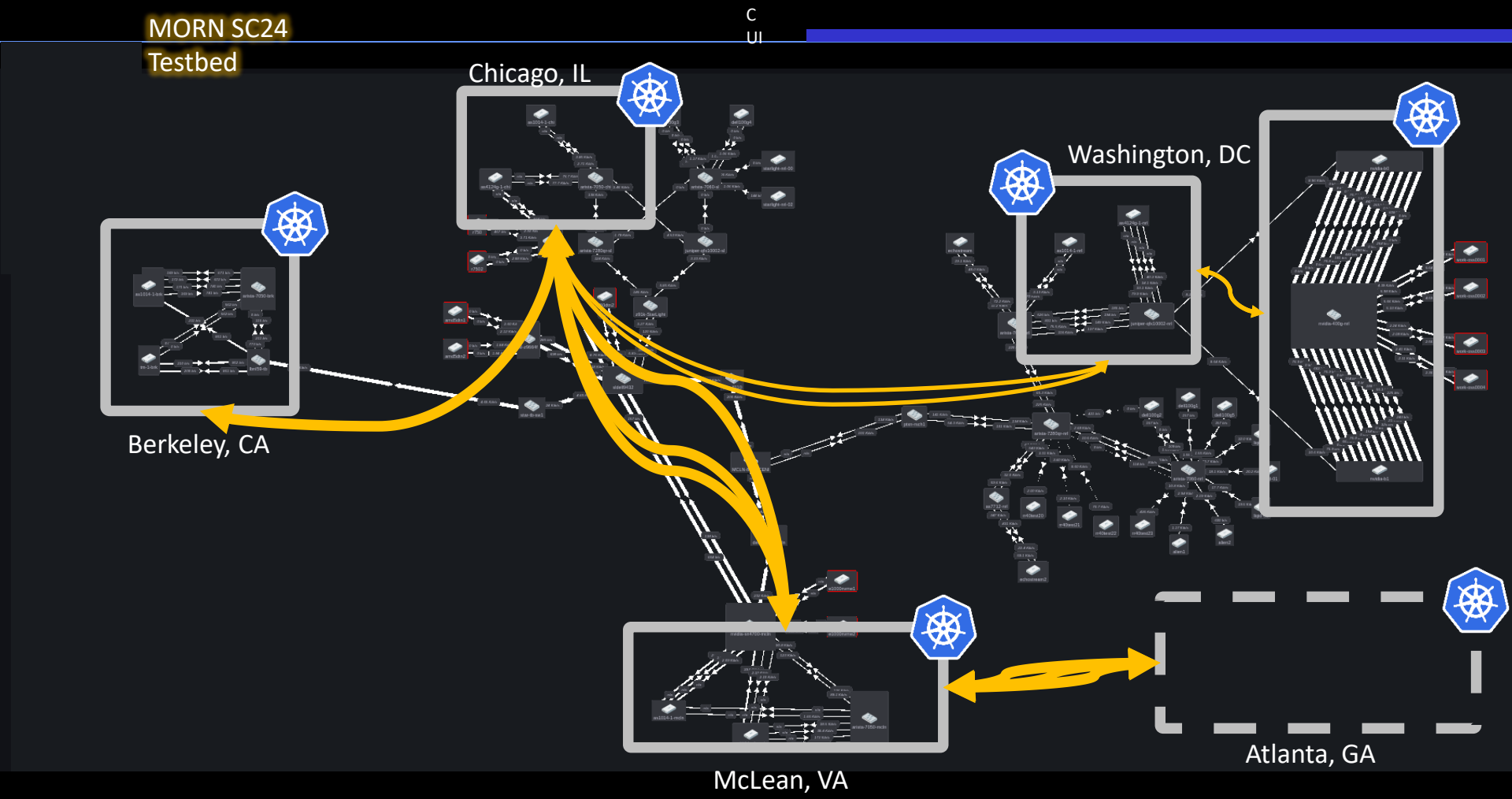
National Research Platform



+ Open Science Grid & Open Science Data Grid

STARLIGHT℠

MORN SC24 Testbed

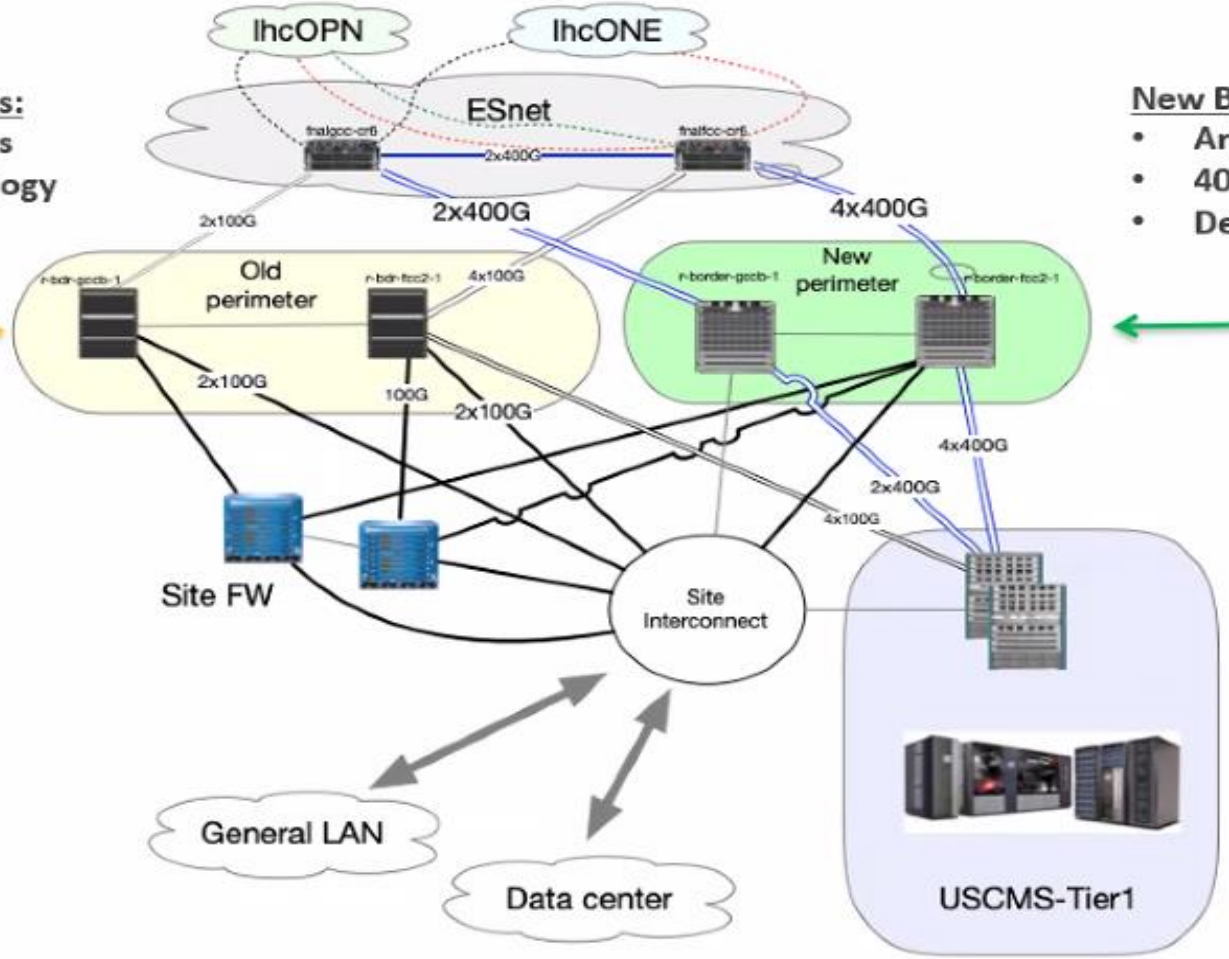


STARLIGHT℠

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

Old Border Routers:
- Brocade MLXes
- 100GE technology

New Border Routers:
- Arista 7808R3s
- 400-800GE technology
- Deep buffers

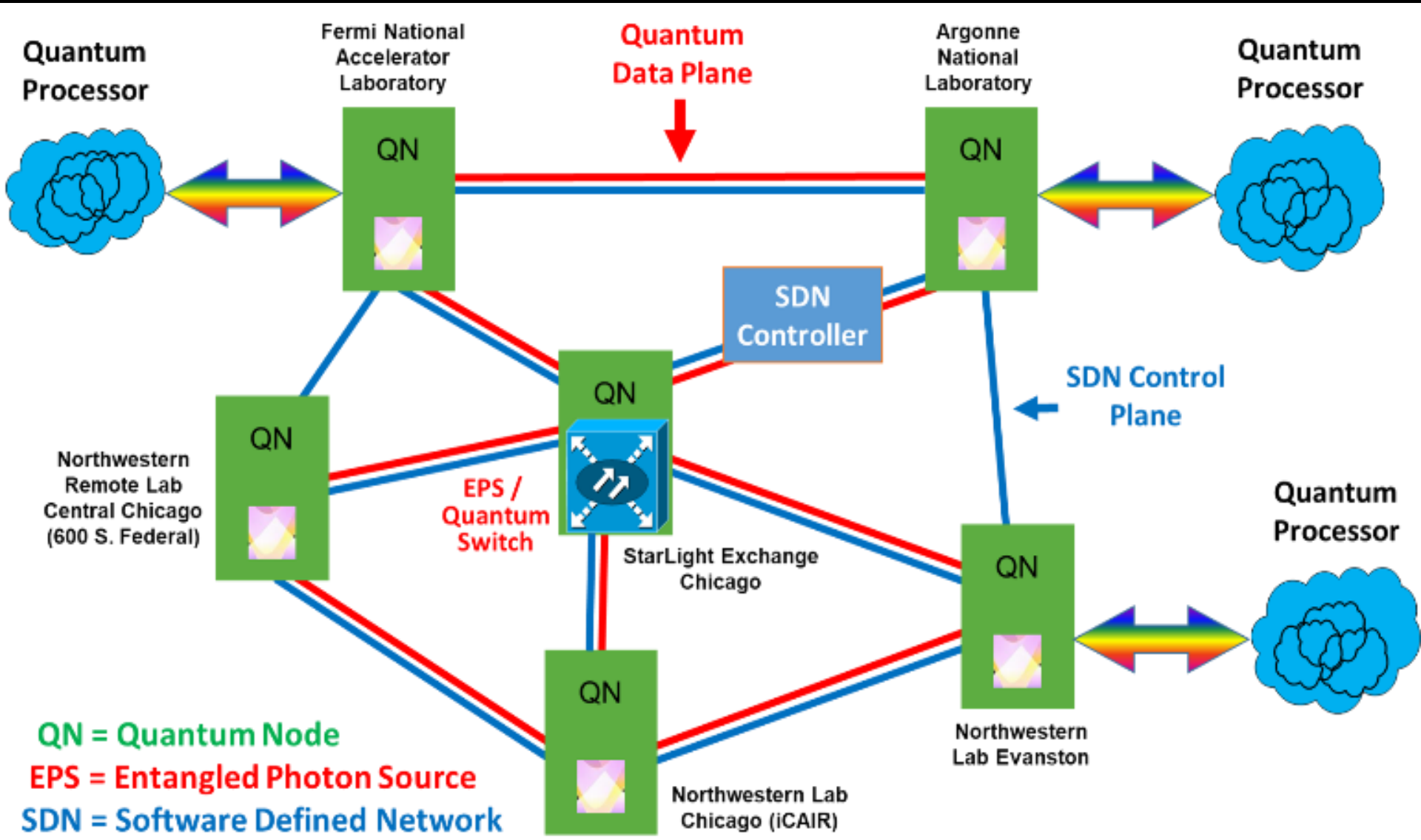


Also: Large Scale Network Research Testbed On Site

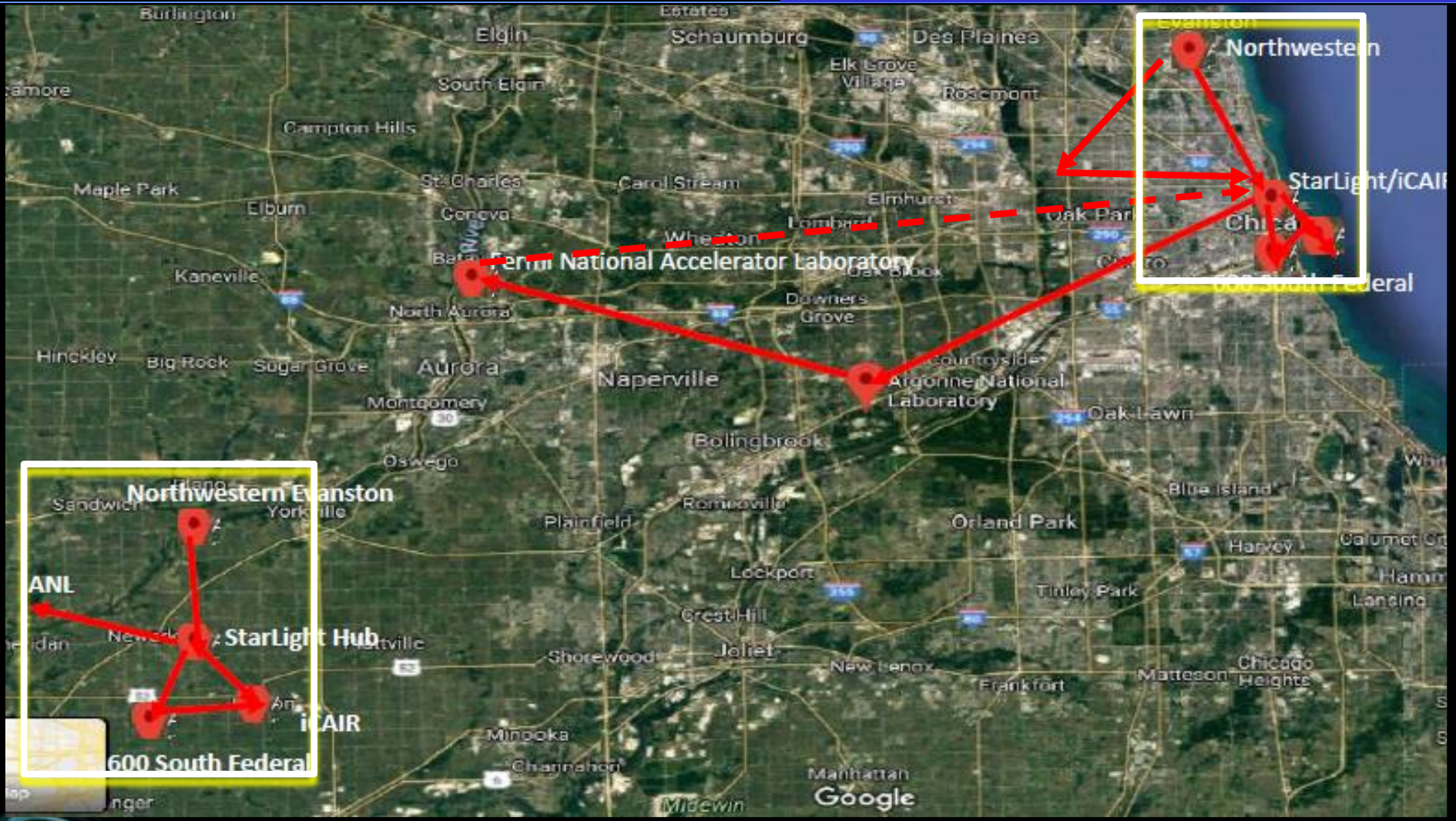
Quantum Communications And Networks: Motivation

- **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

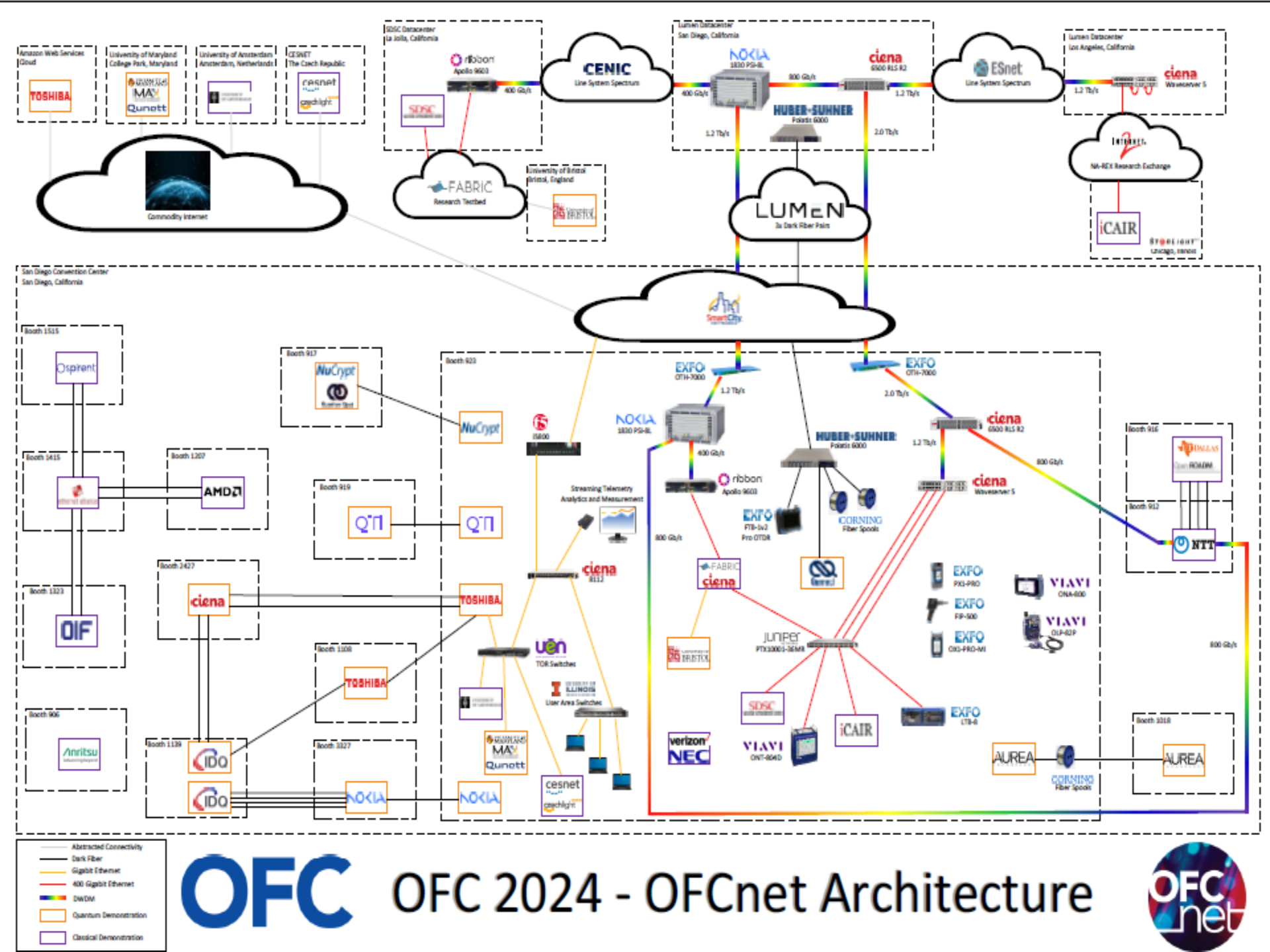


Energing IEQnet Testbed Topology



Illinois Express Quantum Network

STARLIGHT℠

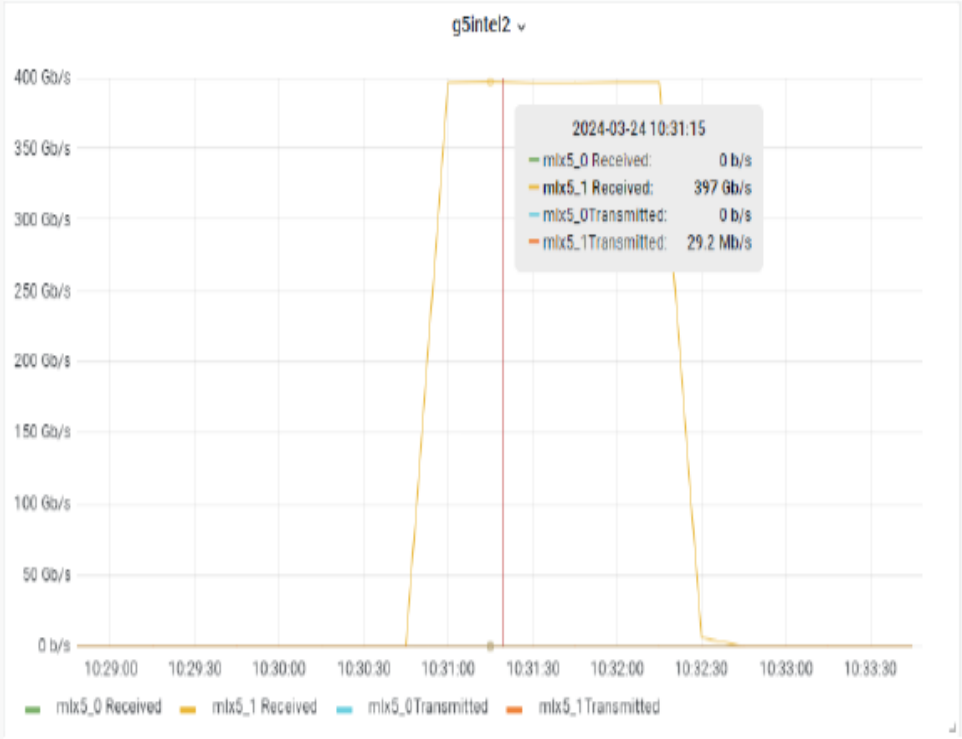


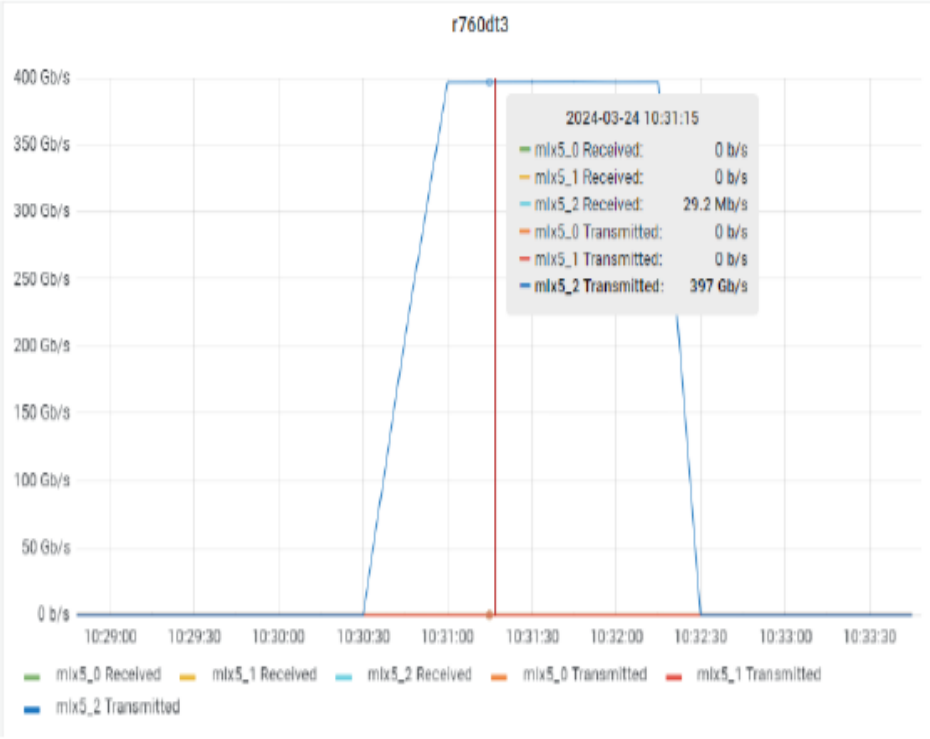
OFC 2024 - OFCnet Architecture

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

iCAIR

March 24, 2024

STARLIGHT SDX

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

DOE's Integrated Research Infrastructure (IRI) Vision:

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

Advanced Networking

Edge Sensors

Advanced Computing

Researchers

Local Campus Computing

Computing Testbeds

Cloud Computing

Data Management

High Performance Data Facility

Data Repositories PuRE Data Assets

Software

Software and Applications

AI Tools Digital Twins

New modes of integrated science

Rapid data analysis and steering of experiments

Novel workflows using multiple user facilities

AI-enabled insight from integrating vast data sources

U.S. DEPARTMENT OF ENERGY | Office of Science

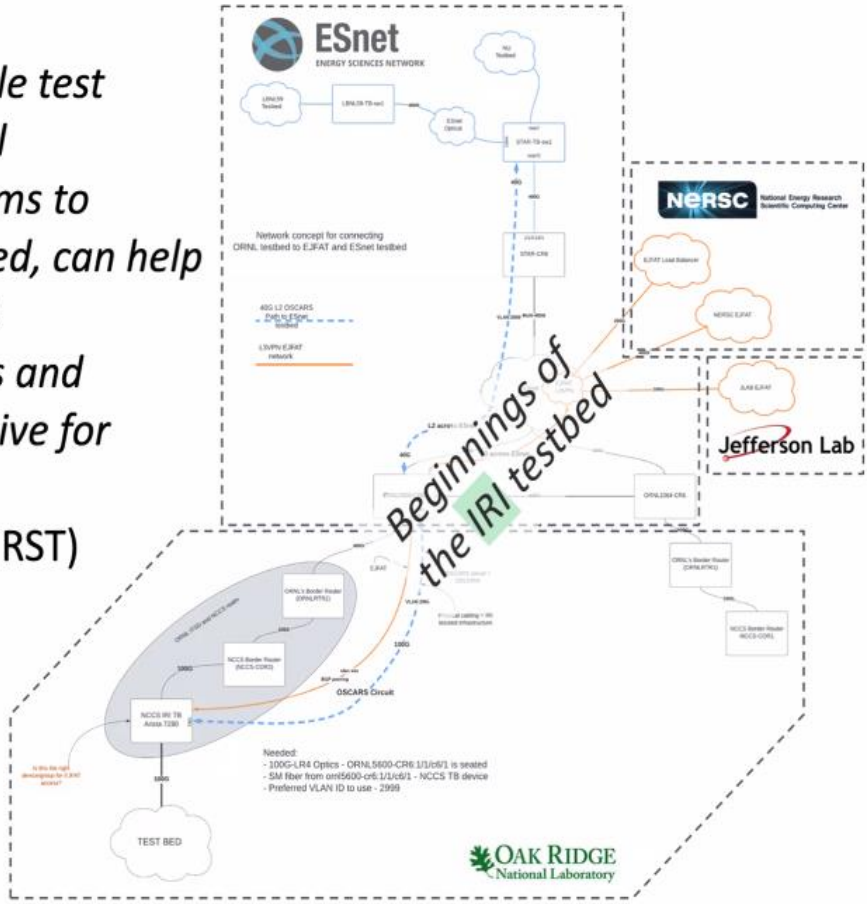
Slide from Ben Brown, DOE, at ECP IAC – Jan 2024

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**

STARLIGHT℠

"Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Networking and Information Technology Research and Development Program."

The Networking and Information Technology Research and Development (NITRD) Program

Mailing Address: NCO/NITRD, 2415 Eisenhower Avenue, Alexandria, VA 22314

Physical Address: 490 L'Enfant Plaza SW, Suite 8001, Washington, DC 20024, USA Tel: 202-459-9674, Fax: 202-459-9673, Email: nco@nitrd.gov, Website: https://www.nitrd.gov