

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

**Joe Mambretti, Director, ([j-mambretti@northwestern.edu](mailto:j-mambretti@northwestern.edu))**

**International Center for Advanced Internet Research ([www.icaair.org](http://www.icaair.org))**

**Northwestern University**

**Director, Metropolitan Research and Education Network ([www.mren.org](http://www.mren.org))**

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

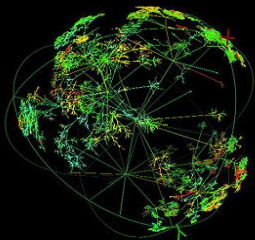
**PI: StarLight SDX, Co-PI Chameleon, PI-iGENI, PI-OMNINet**

**Global Research Platform**

**Co-Located With The 19<sup>th</sup> IEEE International Conference on eScience**

**Limassol, Cyprus**

**October 9-10, 2023**



# The GRP: A Platform For Global Science



## GLOBAL RESEARCH PLATFORM

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

# Global Research Platform: Global Lambda Integrated Facility Available Advanced Network Resources



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



[www.glif.is](http://www.glif.is)

**STARLIGHT**<sup>SM</sup>

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



Schematic overview of the GNA-G AutoGOLE



**AutoGOLE Open R&E Exchanges**

**STARLIGHT<sup>SM</sup>**

# 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

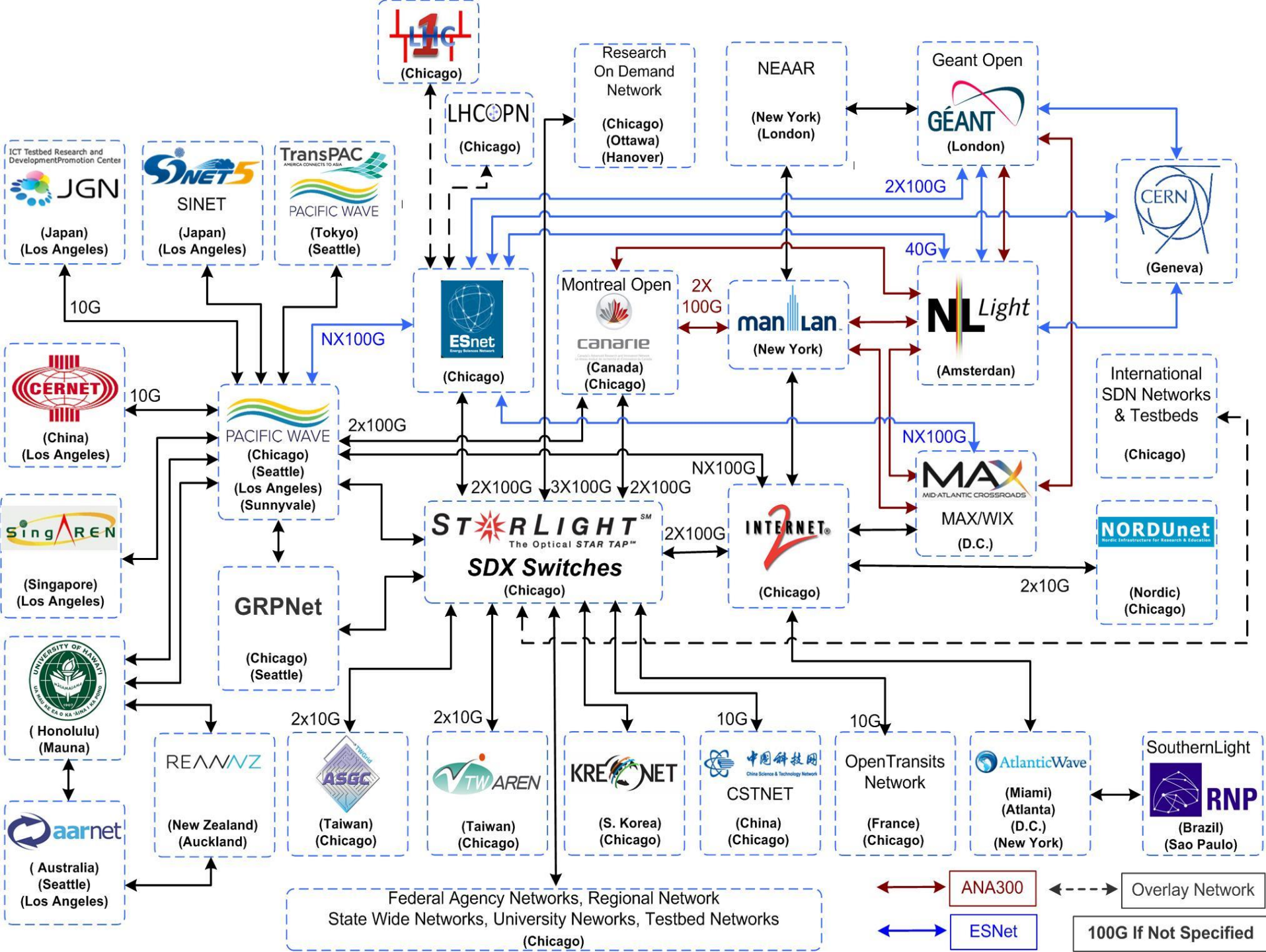


View from StarLight



Abbott Hall, Northwestern University's Chicago Campus

**Currently: 20+ 400 Gbps Paths Prototyping 800 Gbps** **STARLIGHT**<sup>SM</sup>

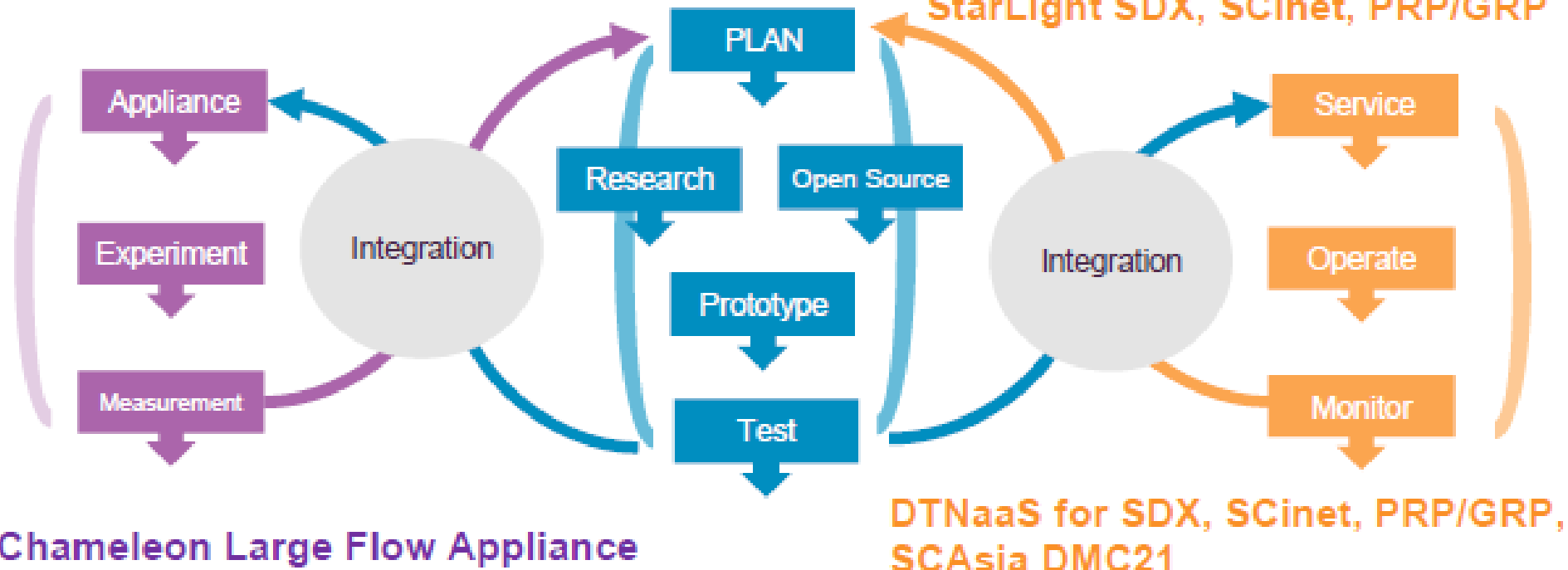


# StarLight Software Defined Exchange

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

StarLight Testbeds

StarLight SDX, SCinet, PRP/GRP



iCAIR

STARLIGHT<sup>SM</sup> SDX

Source: Jim Chen

STARLIGHT<sup>SM</sup>

# ESnet 6

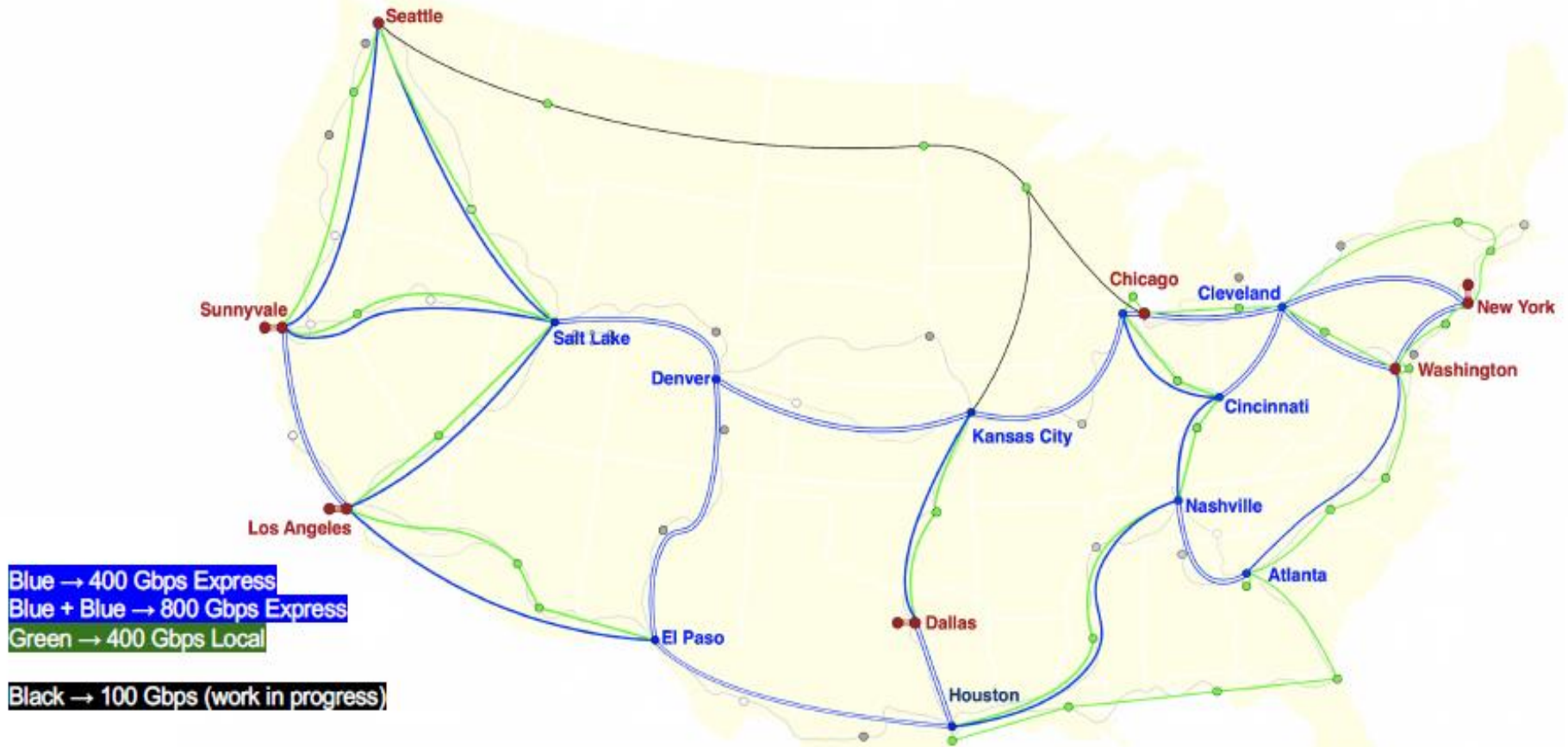




# Internet2 Backbone Topology

## Backbone Topology - Capacity and Traffic Management

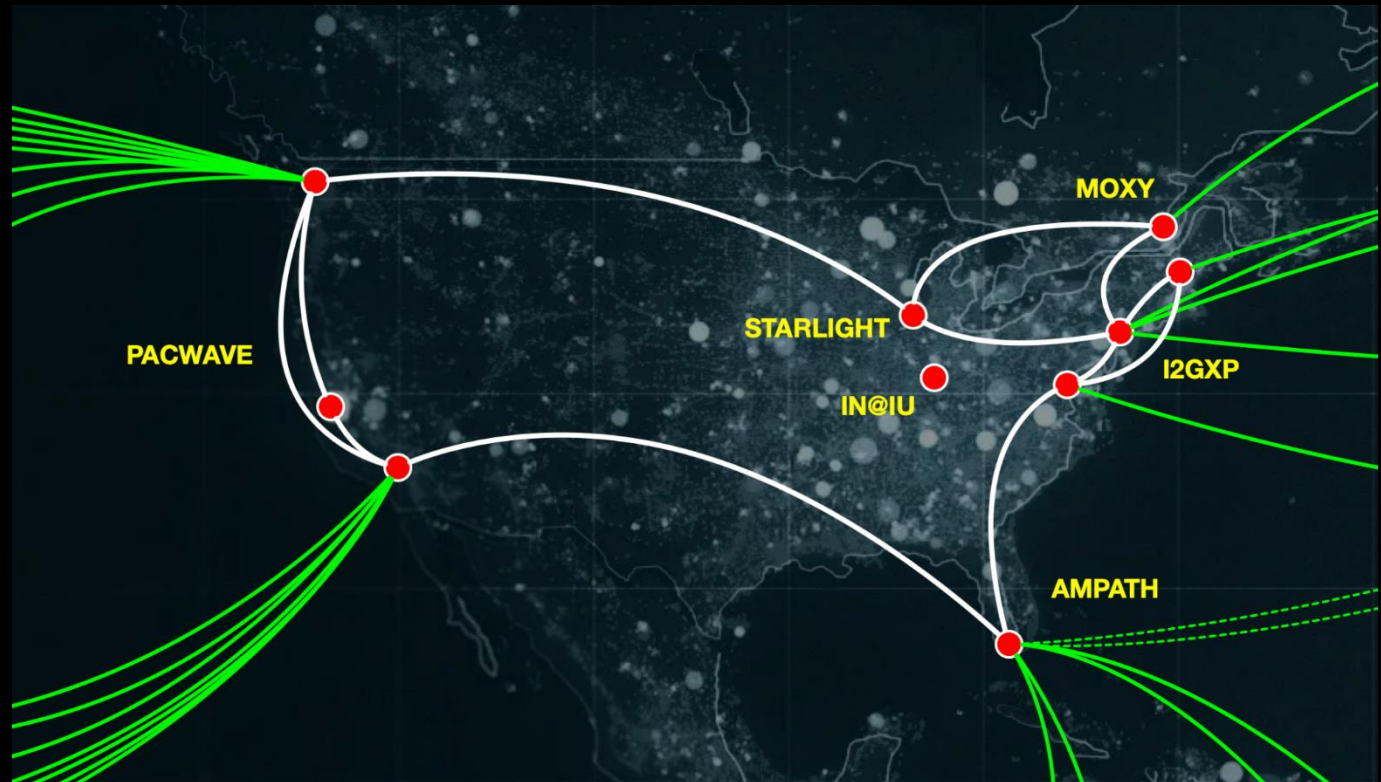
Chris Wilkinson, Director of Planning and Architecture



# AP-REX 2.0 – NA-REX

## Addition of partners:

- AMPATH / FIU
- CANARIE
- IU International
- MOXY
- StarLight International / National Communications Networks Exchange Facility



10

# **International Federated Testbeds As Instruments for Computer Science/Network Science**

- **The StarLight Communications Exchange Facility Supports ~ 25 Network Research Testbeds (Instruments For Computer Science/Networking Research)**
- **StarLight Supports Two Software Defined Exchanges (SDXs), An NSF IRNC SDX & A Network Research GENI SDX (Global Environment for Network Innovations)**
- **The GENI SDX Supports National and International Federated Testbeds**

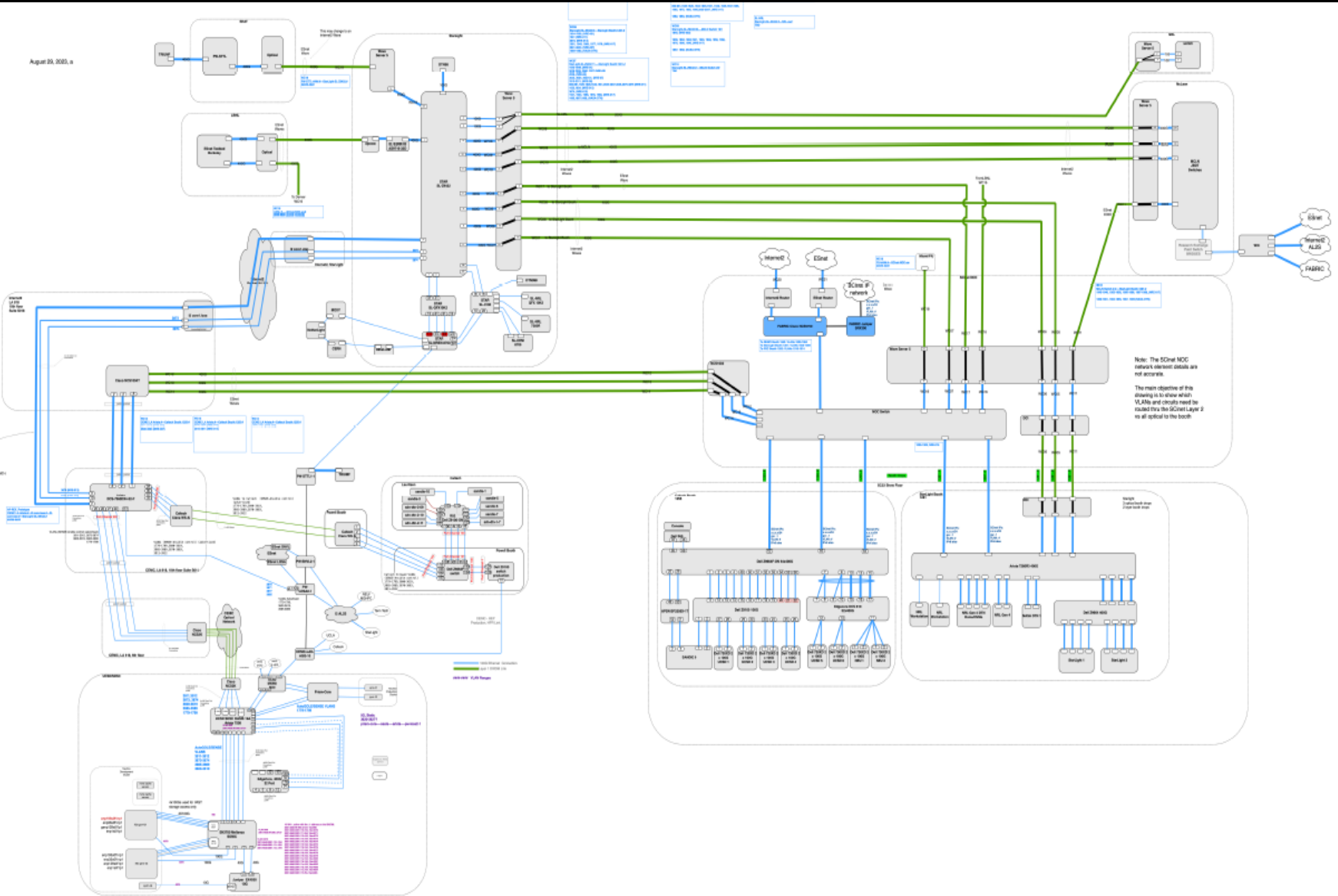


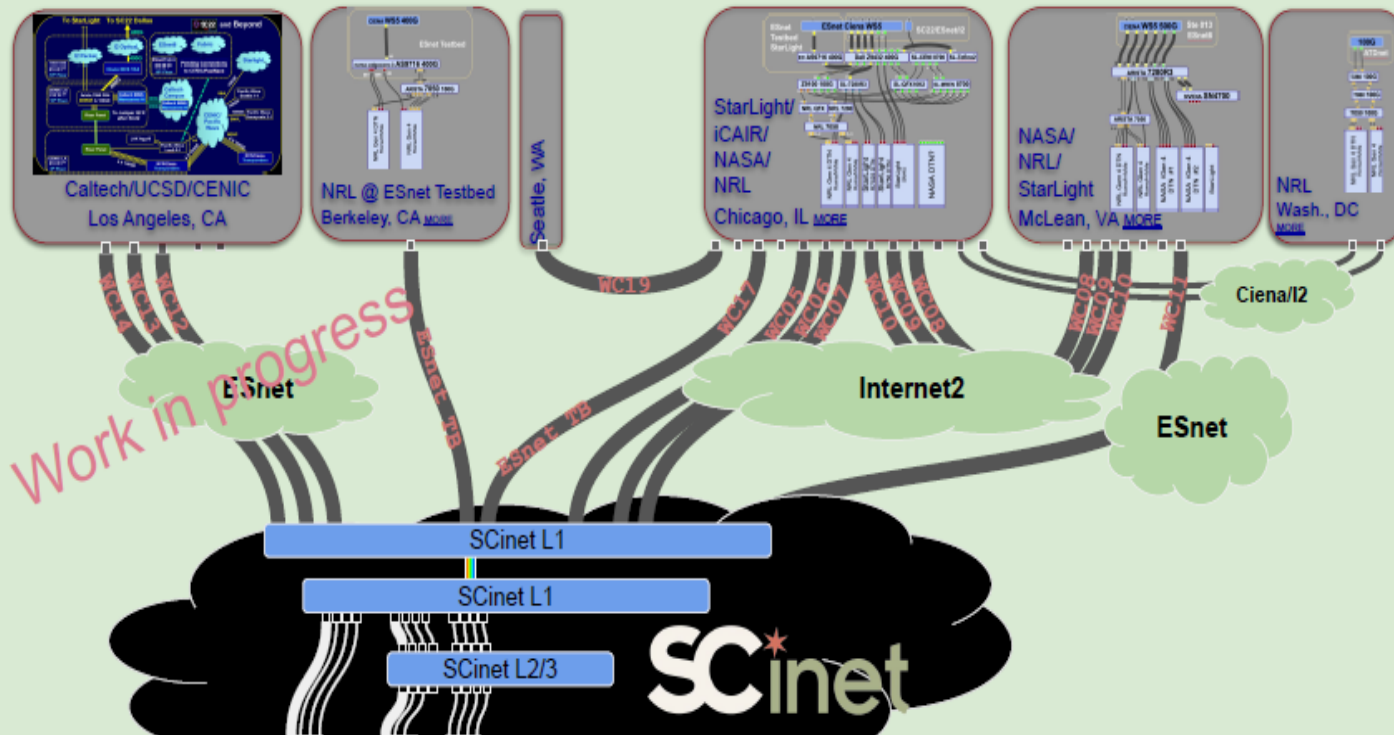
# SCinet National WAN Testbed

- **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**



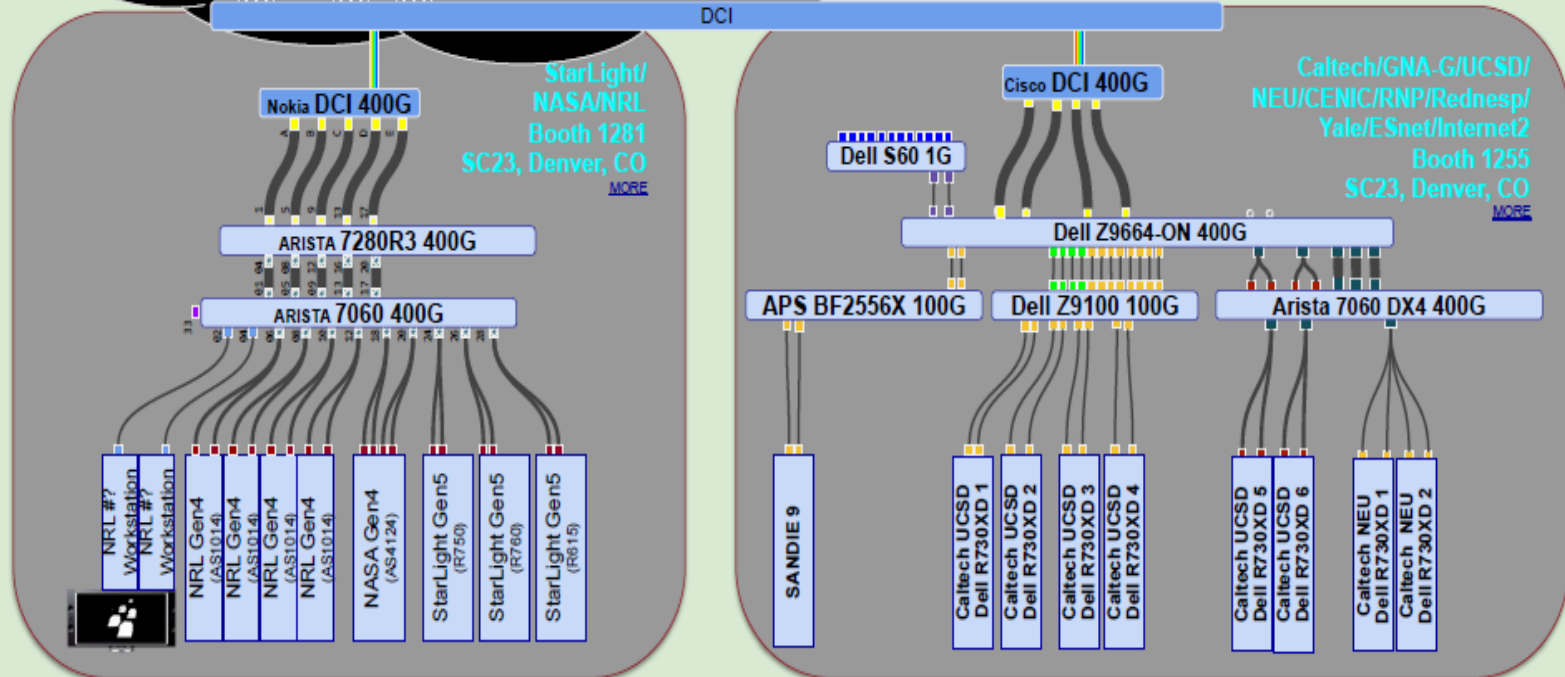
August 29, 2023, a





*Work in progress*

**JOINT  
BIG  
DATA  
TESTBED**



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

09/19/2023

Latest Version at:  
<https://tinynurl.com/SC23-JBDT>  
To request changes, please leave a comment

[SC23 floorplan](#)





# **Example SC23 SCinet Network Research Exhibitions**

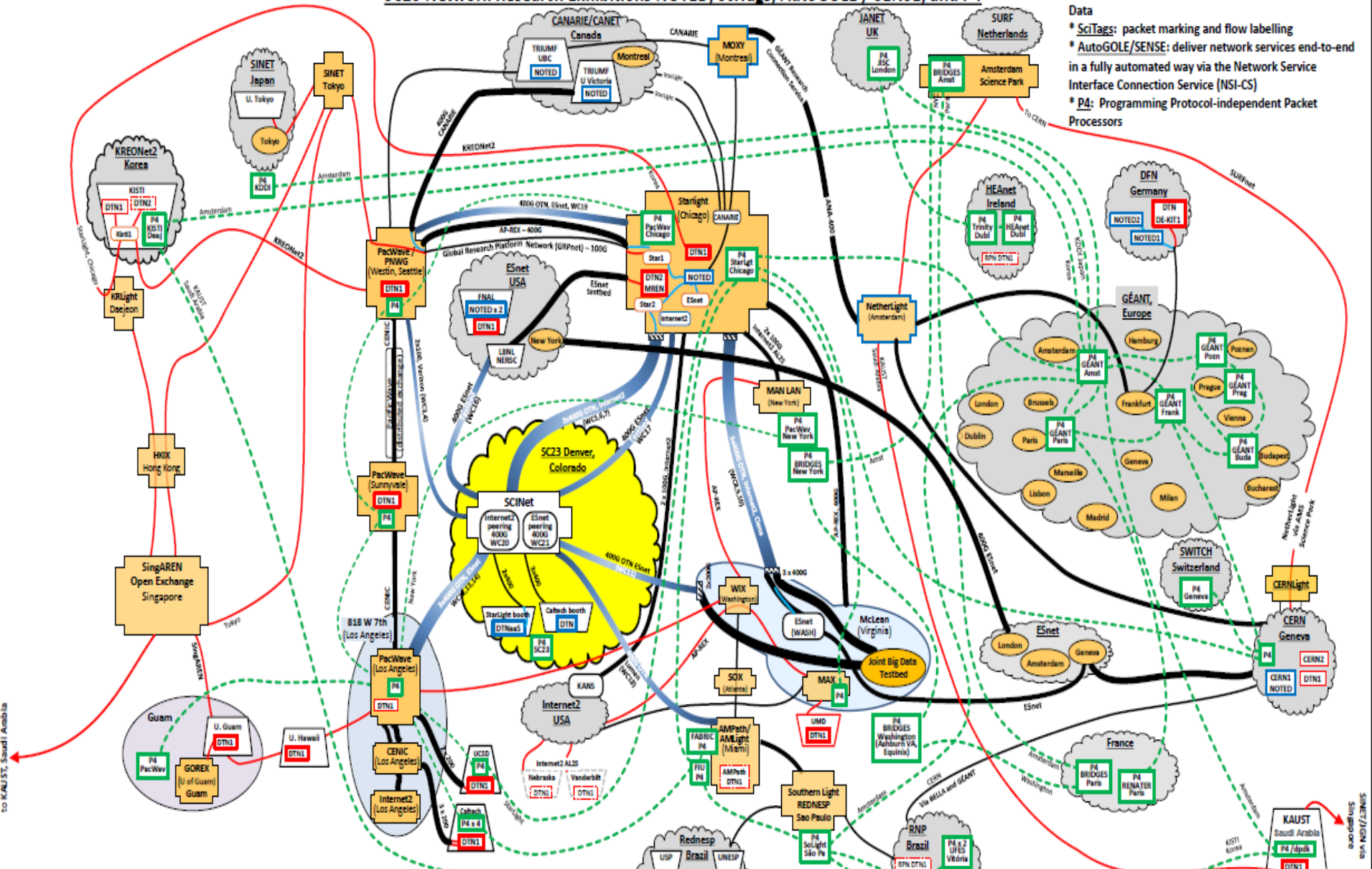
- **Global Research Platform (GRP)**
- **SDX 1.2 Tbps WAN Services**
- **SDX E2E 400 Gbps WAN Services**
- **400 Gbps DTNs & Smart NICs**
- **Network Optimized Transport for Experimental Data (NOTED) – With AI/ML Driven WAN Network Orchestration**
- **SDX International Testbed Integration**
- **StarLight SDX for Petascale Science**
- **DTN-as-a-Service For Data Intensive Science**
- **P4 Integration With Kubernetes**
- **PetaTrans Services Based on NVMe-Over-Fabric**
- **NASA Goddard Space Flight Center HP WAN Transport Services**
- **Resilient Distributed Processing & Rapid Data Transfer**
- **PRP/NRP Demonstrations**
- **Open Science Grid Demonstrations**
- **N-DISE Named Data Networking for Data Intensive Science**
- **Orchestration With Packet Marking (SciTag)**





# SC23 Network Research Exhibitions NOTED, SciTags, AutoGOLE / SENSE, and P4

\* **NOTED**: Network-Optimized Transfer of Experiment Data  
 \* **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



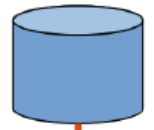
SC23 NRE map v. 9 - WJelohnston, Esetnet, [wej@es.net](mailto:wej@es.net)

<b>NOTED</b>	SC21 NOTED infrastructure is in blue	<b>GENS</b>	Pattern filled rectangles are DTNs or other computing/storage elements	<b>McLean</b>	Carrier hotels, etc	<b>BRP</b>	Cloud bubbles are mostly regional/national networks
<b>AG/SENSE</b>	AutoGOLE / SENSE infrastructure is in red	<b>ESnet</b>	Shared circuits supporting demonstrations	<b>Paris</b>	Ovals are points of presence in regional infrastructure	<b>100G</b>	100G
<b>P4</b>	P4 infrastructure is in green	<b>Starlight</b>	Circuits supporting AutoGOLE / SENSE	<b>ESnet</b>	Rounded rectangles are individual switch/router	<b>200G</b>	200G
<b>general</b>	Shared or general infrastructure is in black	<b>Calltech</b>	Sites	<b>SOX</b>	Exchange points (external or internal to a site)	<b>400G</b>	400G
		<b>SOX</b>	Exchange points (external or internal to a site)	<b>SOX</b>	Exchange points (external or internal to a site)	<b>800G</b>	800G
		<b>SOX</b>	Exchange points (external or internal to a site)	<b>SOX</b>	Exchange points (external or internal to a site)	<b>1 T/s</b>	1 T/s
		<b>SOX</b>	Exchange points (external or internal to a site)	<b>SOX</b>	Exchange points (external or internal to a site)	<b>SCINet managed</b>	SCINet managed

- NOTES**
- 1) Within exchange points, etc. line width does not usually indicate bandwidth
  - 2) Map files (JPEG, PDF, and PPTX) are at <https://www.dropbox.com/sh/p2wcyjyubel7q/AAAMU5F09xvUzFospm3RzLzLz1d0>
  - 3) P4 connections are only topological and are not associated with particular network links

to KAUST, Saudi Arabia

SINET/GEN via Singapore



Rucio

FTS

NOTED at KIT

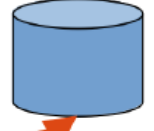
NOTED at CERN

AutoGOLE SENSE

Direct Dynamic Circuit  
LHCOPN default path via CERN



Dynamic Circuits  
LHCOPN default path



Data transfer

TRIUMF

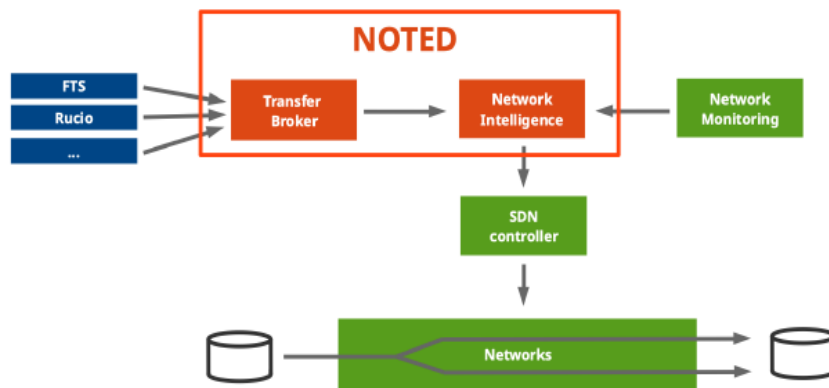


FTS  
File Transfer Service



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

**DTN-UVIC-100g**



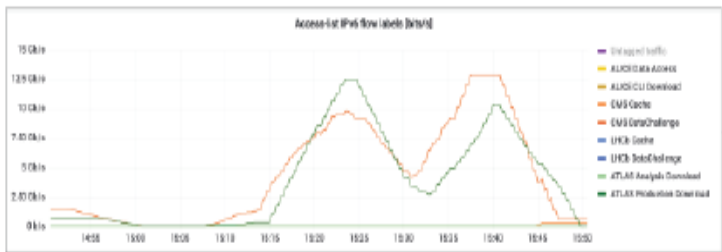
1. Clients requesting data transfers from/to DTN-SC22-400g while passing science domain and activity fields via transfer protocols.

**DTN-KIT-100g**



**canarie**

4. High performance tests using eBPF-TC filters to test encoding of the science domains and activity fields in the IPv6 flow label at scale.



3. P4 programmable switch at CERN collecting the science domains and activity bits encoded in the packets.

**CERN-LHCONE**

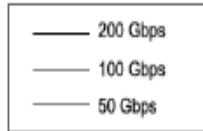
**P4 EdgeCore Wedge**

**CERN**

**DTN-CERN-100g**  
2.6GHz/32 cores  
SSD, 100Gbps



5. Sampling of the low level TCP/IP metrics, which can be used by sites and R&Es to better understand the scientific flows.



**DTN-SC22-400g**  
R7503 2.6 GHz  
NVMe 2.0  
2x200 Gbps

**XRootD**

**SC22 STARLIGHT™**



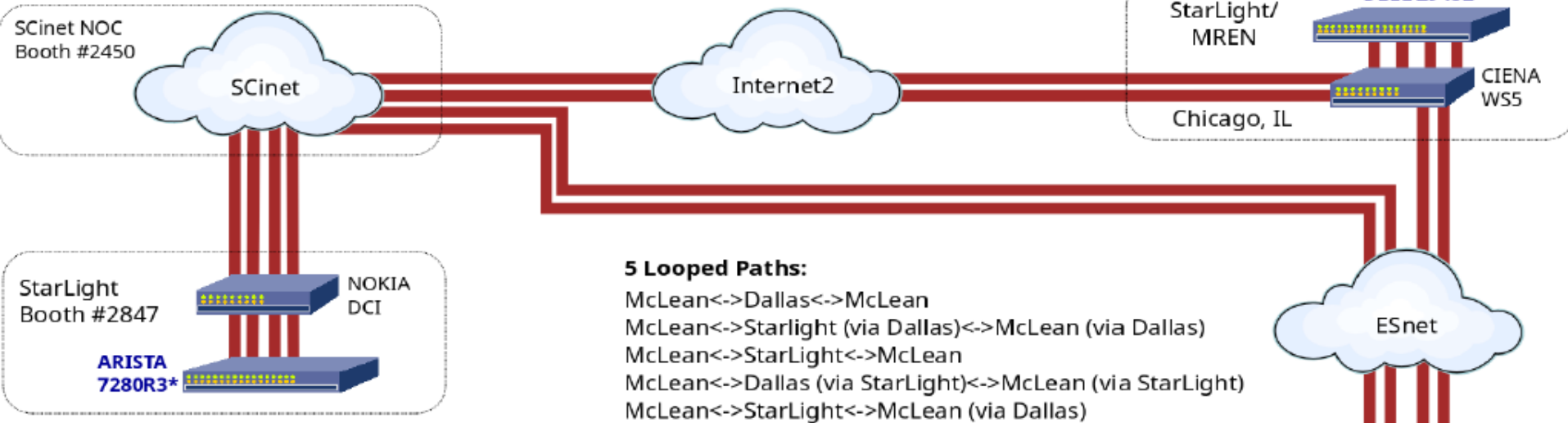
2. XRootD storage responds to the client requests and marks the data transfer packets with the corresponding science domain and activity.

# 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

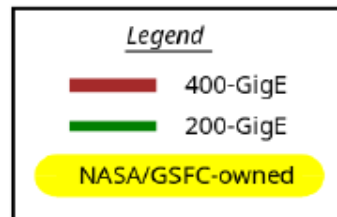
SC22 @ Dallas, TX



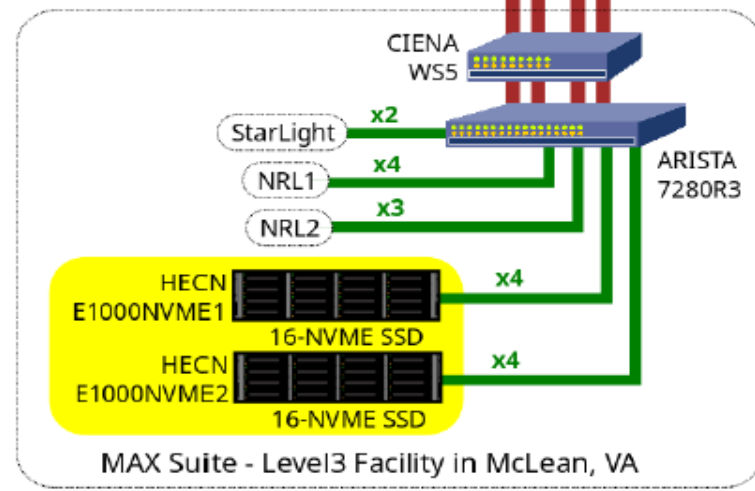
### 5 Looped Paths:

- 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



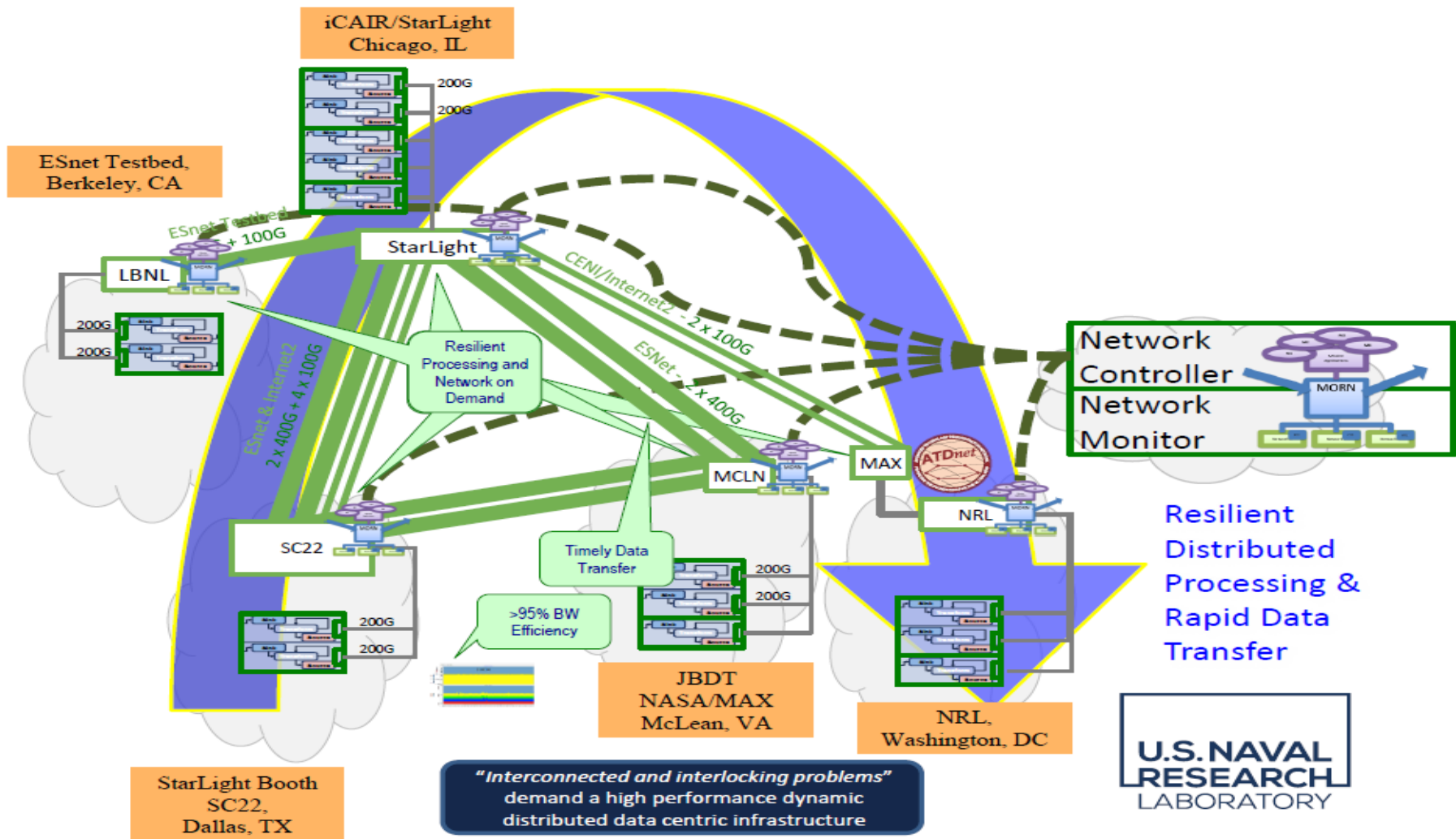
\*IP Routed Loopback Point  
Back To MAX Suite - Level3  
McLean, VA



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

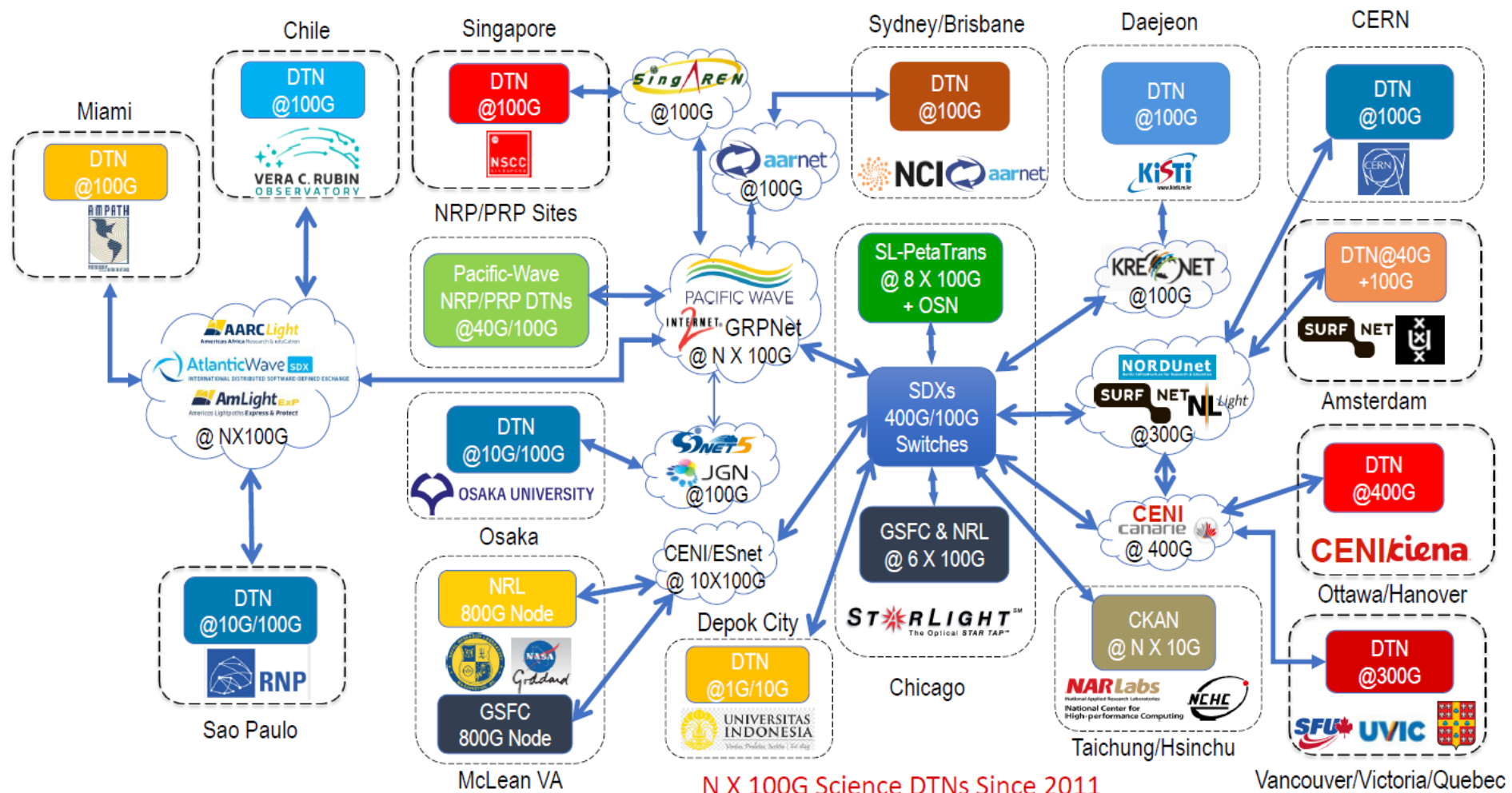


DISTRIBUTION A: Approved for Public Release

STARLIGHT

# GRP DTNaaS For Petascale Science

## GRP Service: DTNaaS for Petascale Sciences Data Movement



N X 100G Science DTNs Since 2011

Vancouver/Victoria/Quebec

# DTN-as-a-Service – Demonstrated At SC22

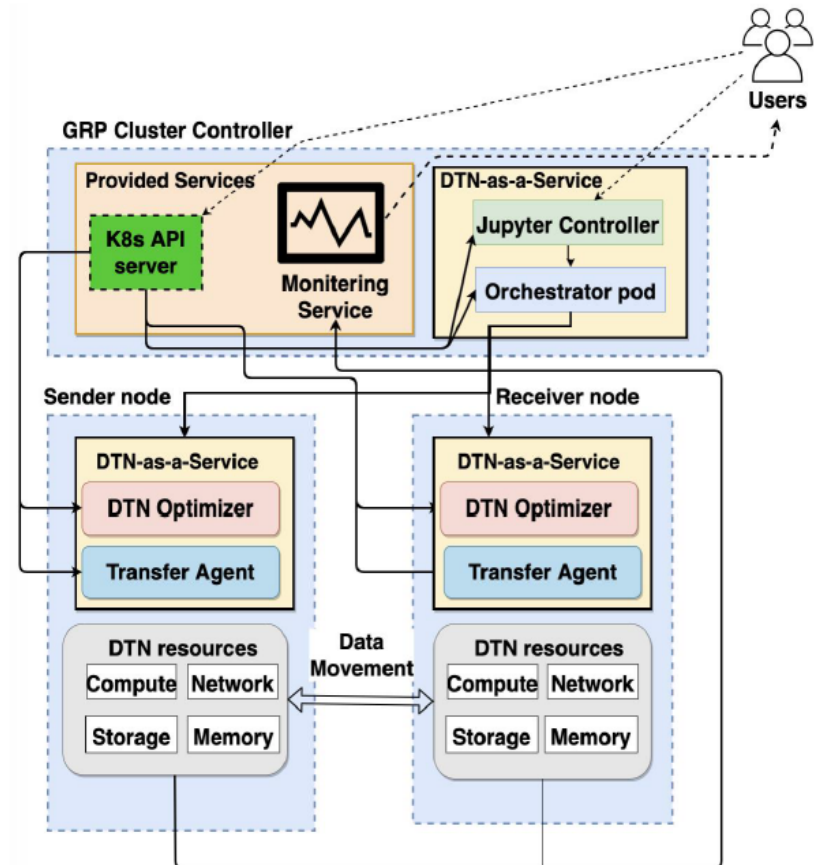
## GRP Cluster with DTN-as-a-Service

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

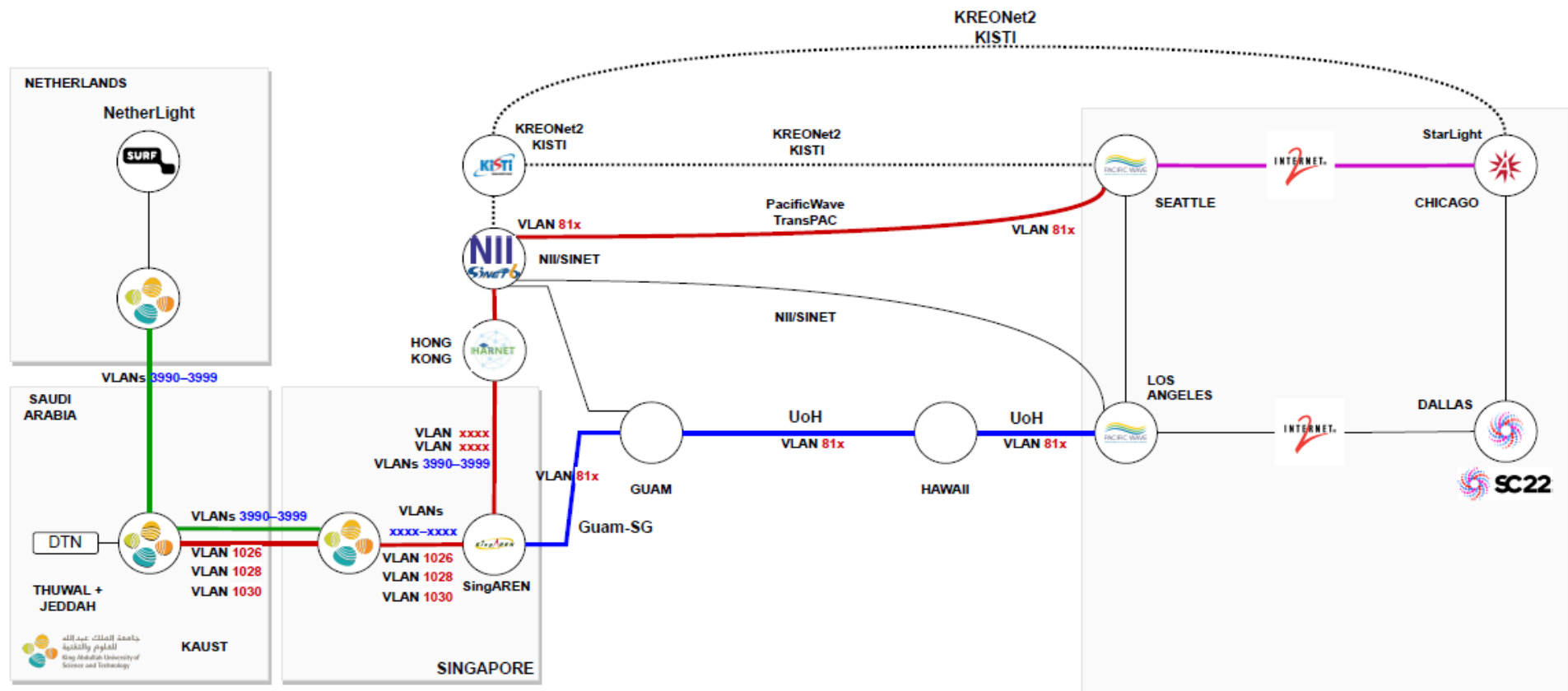
1. Deploy DTNaaS workloads via k8s API server
2. Use Jupyter to optimize and run transfers
3. Observe performance from monitoring service

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







**KAUST-US NEW PERMANENT CIRCUITS (PNWGP Proposal: use VLANs 811, 812, 813)**

- 1) StarLight (VLAN 81x - TBD)  
Amsterdam > Jeddah > Singapore > Hong Kong > Daejeon > Chicago  
NetherLight > KAUST > SingAREN > NICT/INSCC > KREONet2/KISTI > STARLIGHT
- 2) PNWGP / TransPAC (VLAN 81x - TBD)  
Amsterdam > Jeddah > Singapore > Tokyo > Seattle  
NetherLight > KAUST > SingAREN > TransPAC/PacWave
- 3) CENIC (VLAN 81x - TBD)  
Amsterdam > Jeddah > Singapore > Guam > Hawaii > Los Angeles  
NetherLight > KAUST > SingAREN > Guam-SG > UoH > Pacific Wave

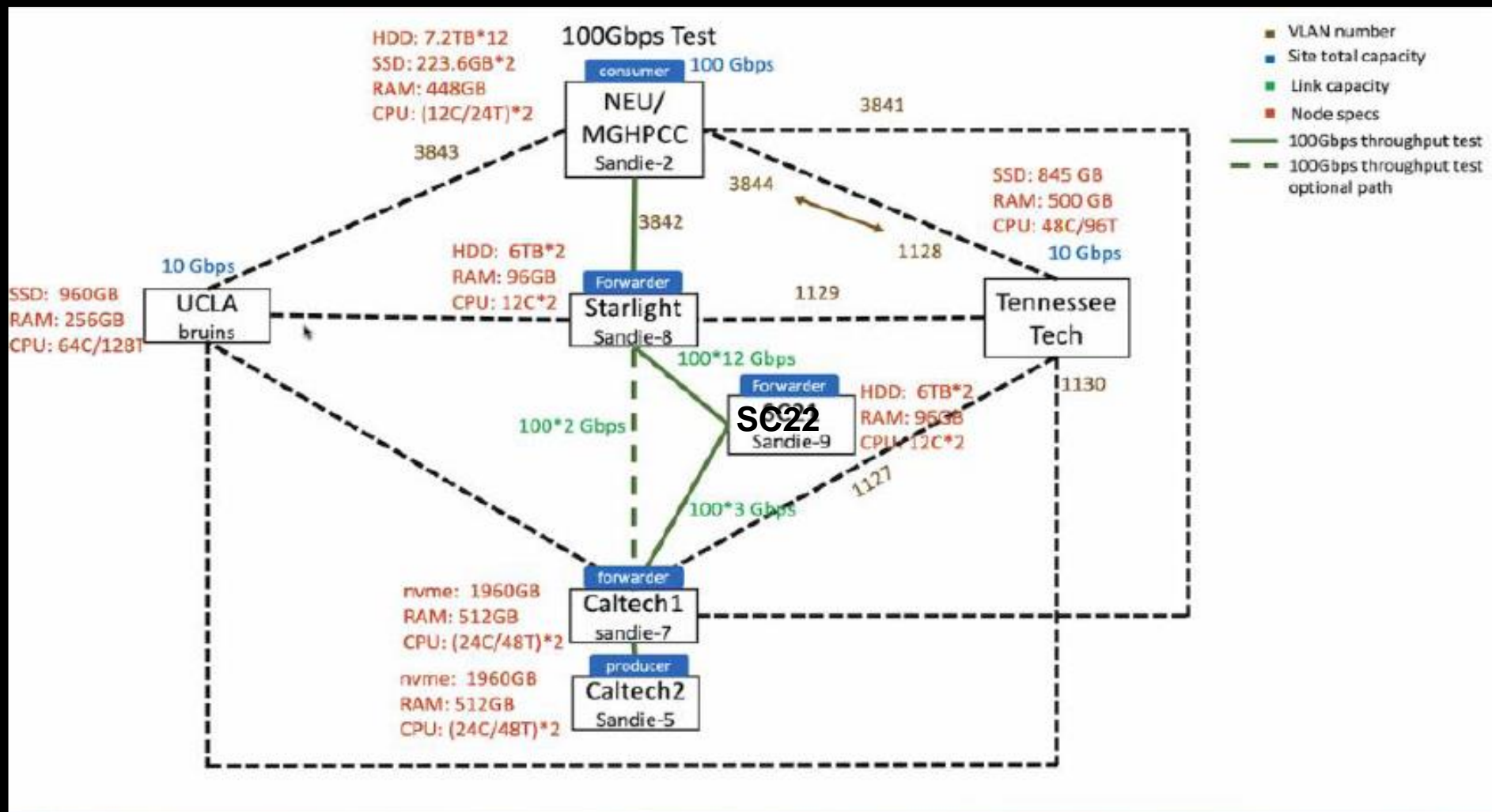
KAUST has deployed 3 circuits - VLANs 1026, 1028, and 1030 - to peer with PacificWave/Cenic, StarLight and SC22

The range 3990-3999 is also available between KAUST Routers in NetherLight PoP and in SingAREN PoP. The AutoGOLE / SENSE can be used to stitch circuits to each end in NetherLight and in SingAREN, and pass traffic through between Europe and Singapore.

Mike: -> I'm responsible specifically for the path between 'Tokyo > Seattle', our suggestion is to use the following VLANs for the permanent VLAN assignments: 811, 812, 813. I can configure the Tokyo side of device interfaces facing NII/SINET & TransPAC towards Seattle.

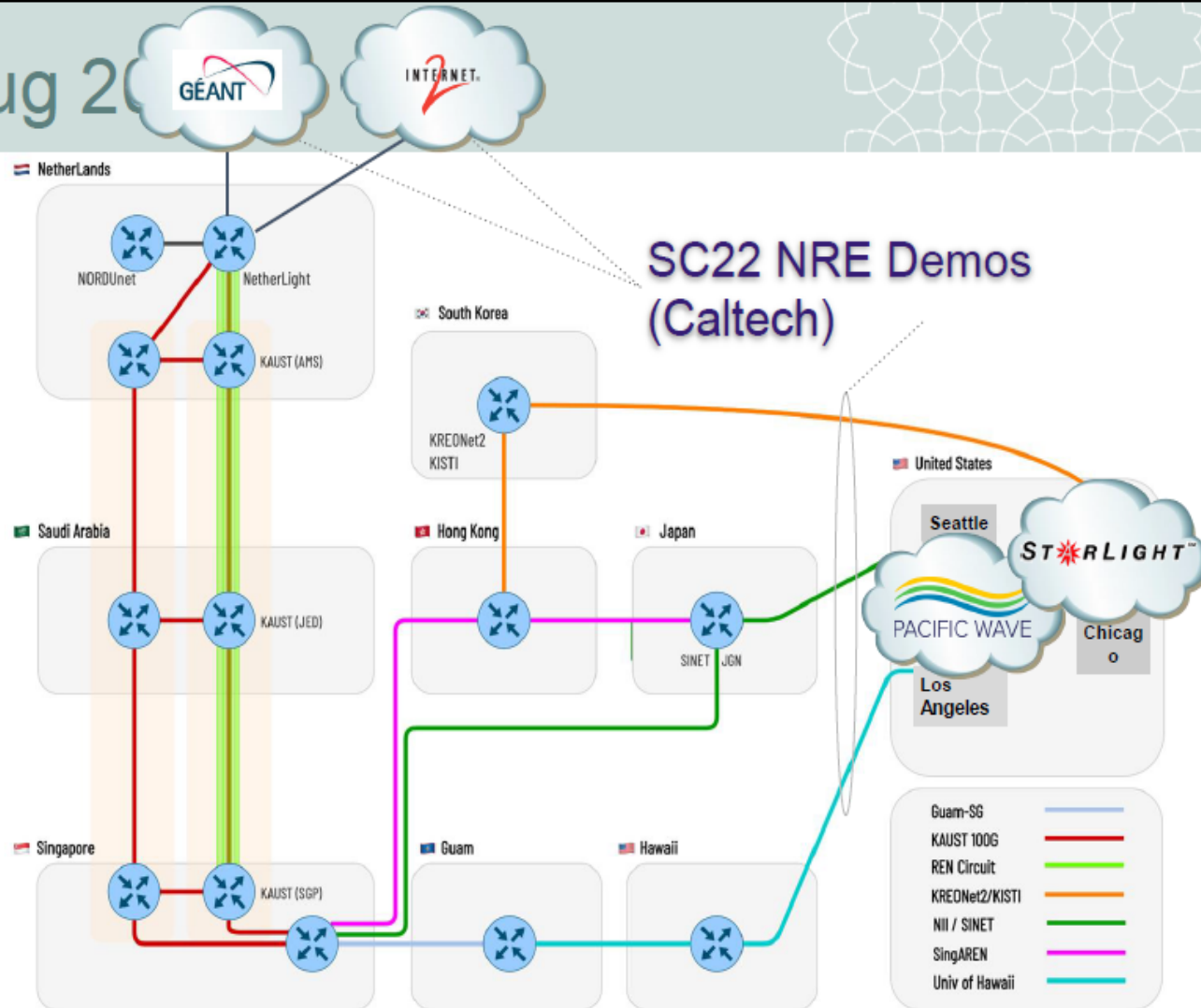
Please contact StarLight NOC for the circuit 1, and CENIC NOC for circuit 3

# Named Data Networking (NDN) for Data Intensive Science Experiments (N-DISE)



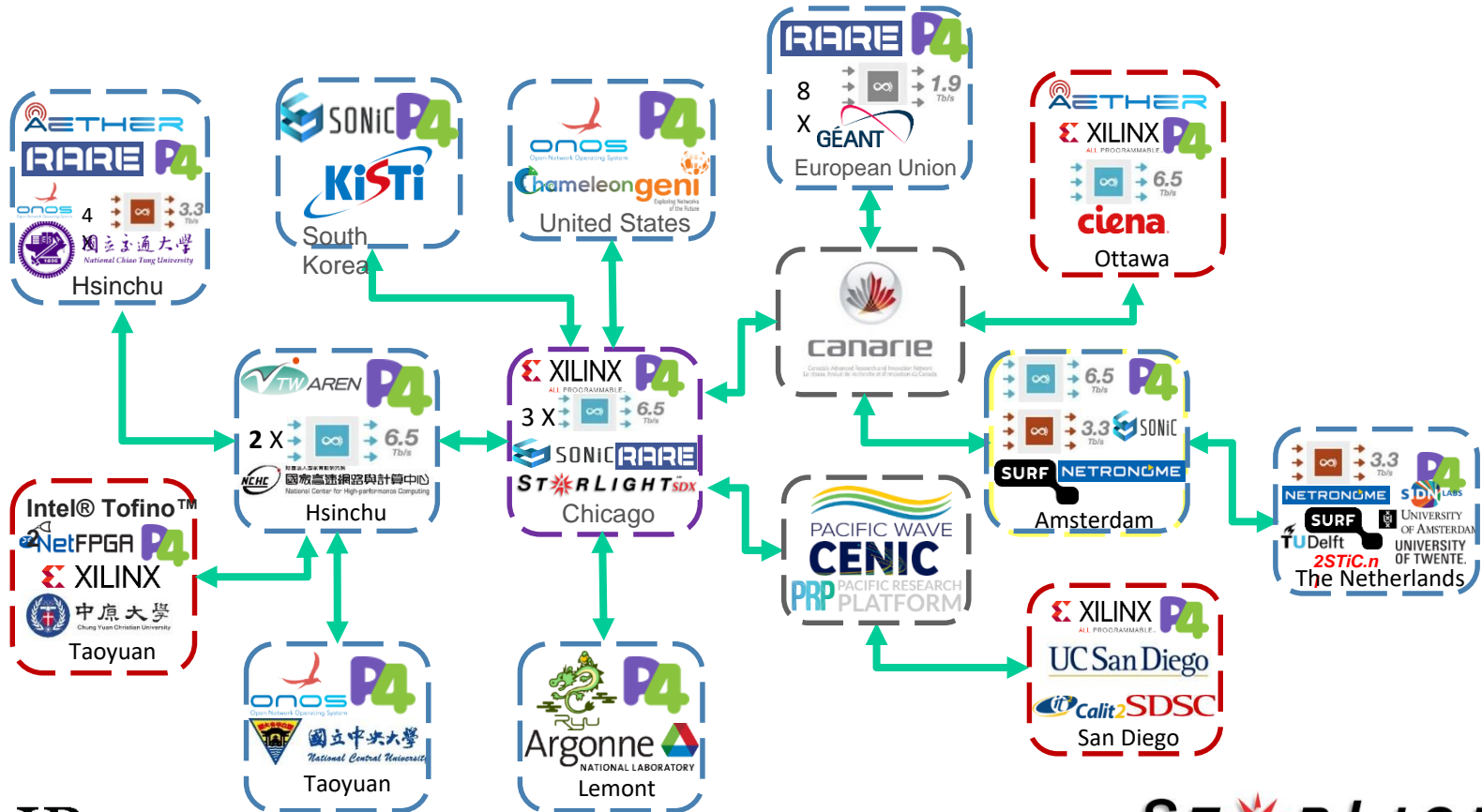
# AER Update - Aug 2010

- Since the AER MoU, KAUST is coordinating with REN partners deployment of sharing spare capacity
- KAUST is supporting the following partners by offering point-to-point circuits for submarine cable backup paths:
  - AARnet
  - GÉANT
  - NetherLight
  - NII/SINET
  - SingAREN
- The [SC22 NRE Demonstrations](#) will also be supported by KAUST closing the ring from Amsterdam to Singapore and back to the US
  - SC22 NRE



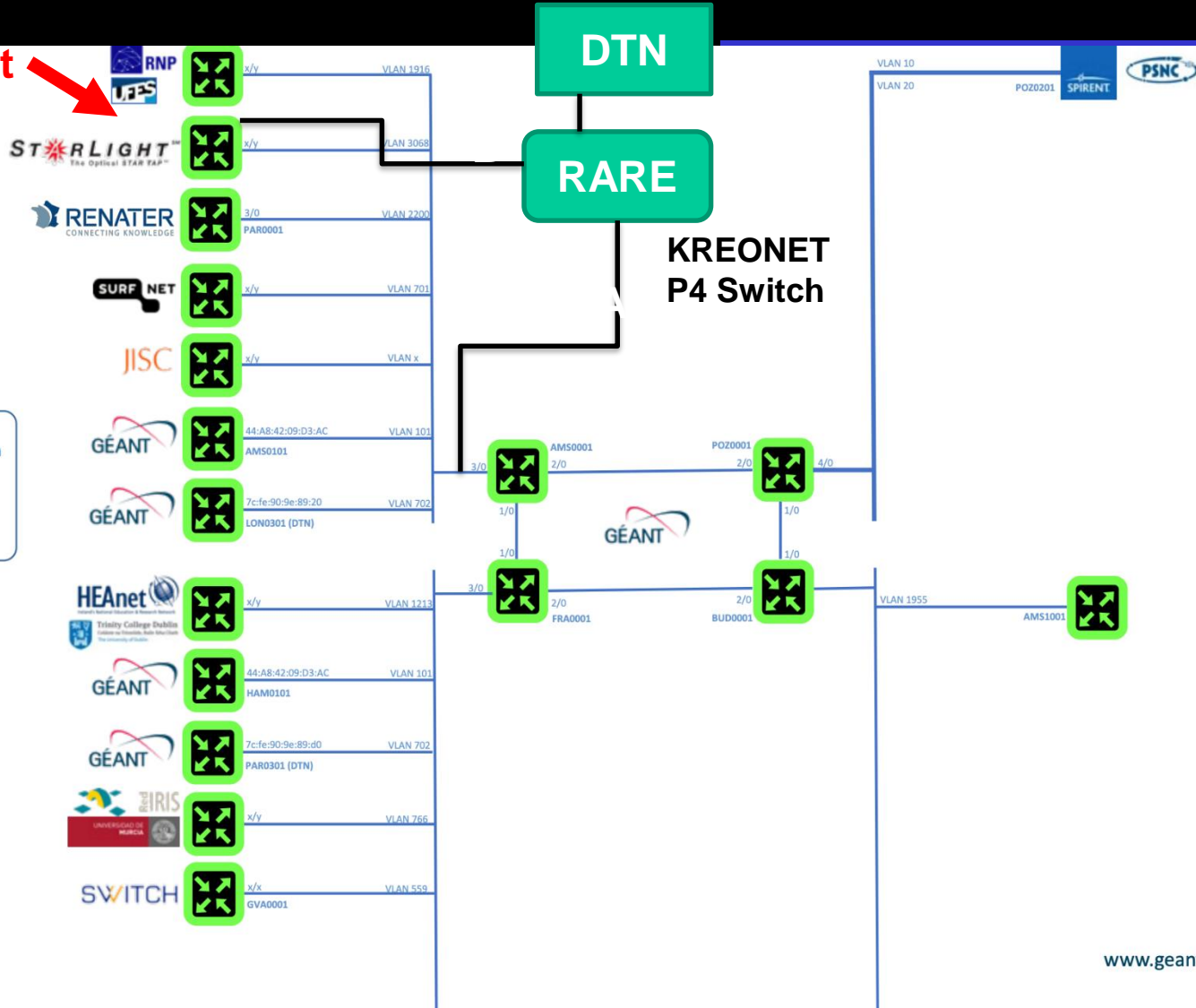
# International P4 Testbed Showcase at SC23

## GRP Service: International P4 Experimental Networks (iP4EN)



# Integration With GEANT P4 Testbed

StarLight





[www.chameleoncloud.org](http://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

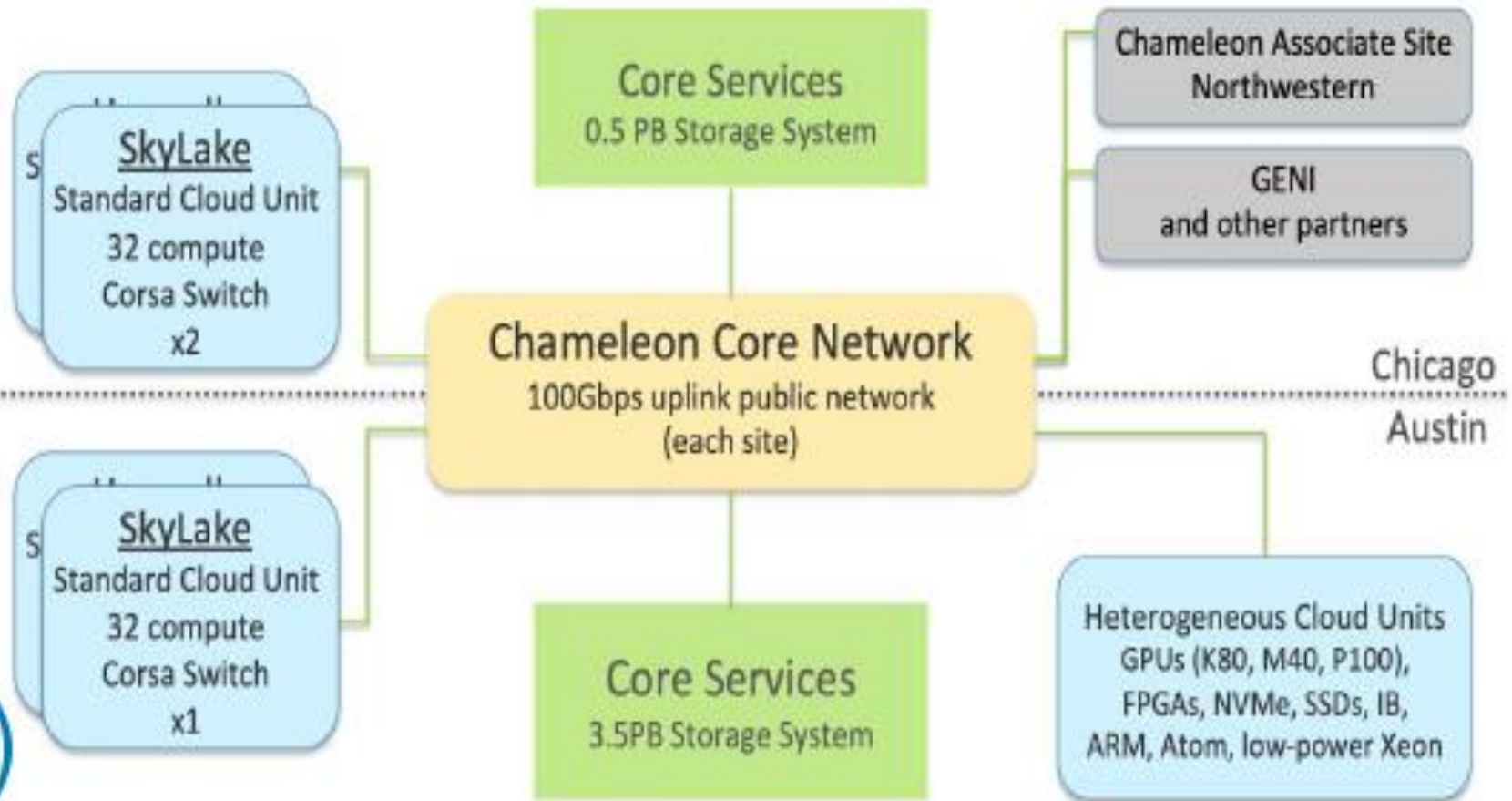


**TACC**

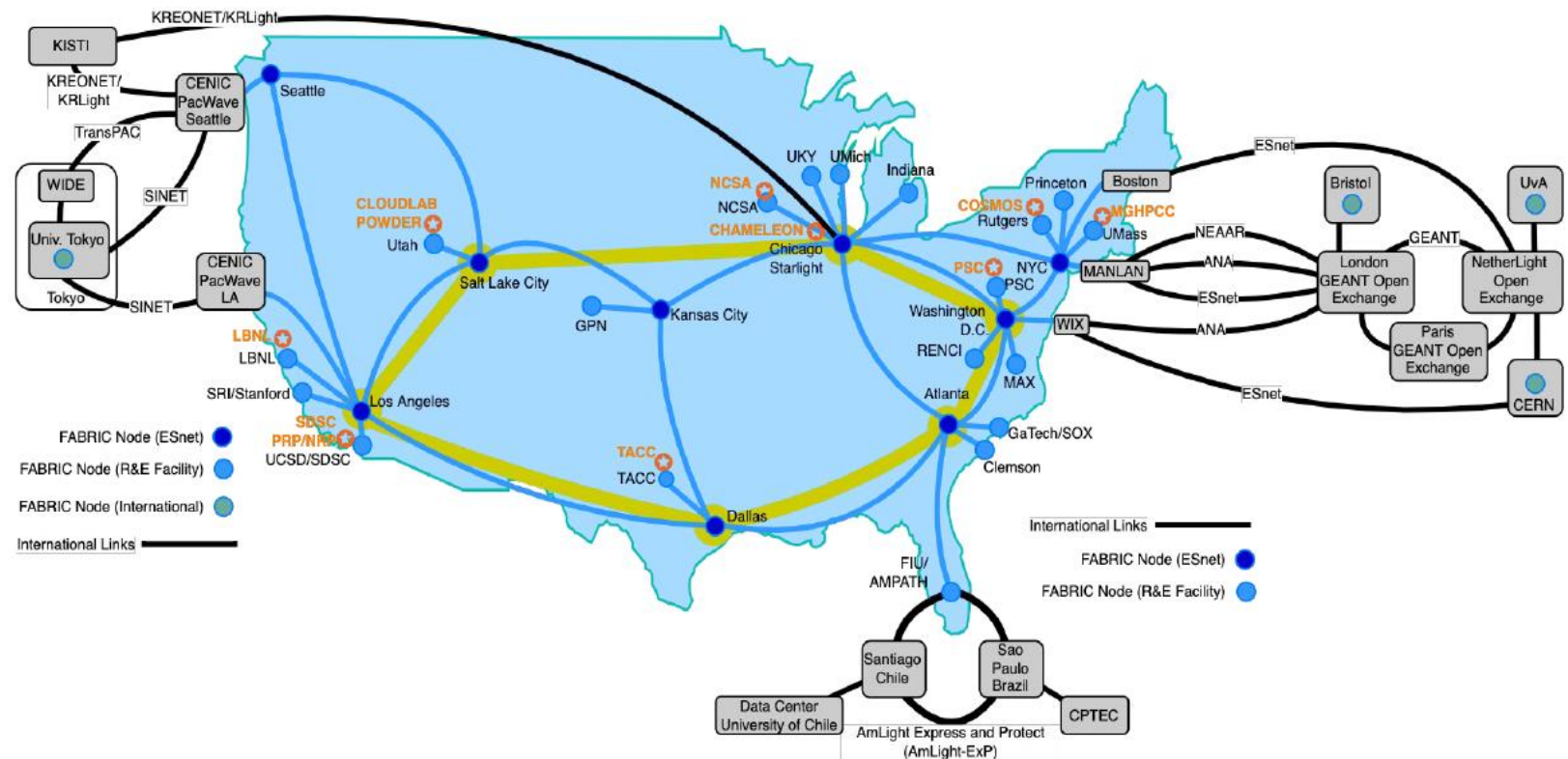
renci



# Chameleon CHI In A Box(CIAB) at StarLight



# FABRIC Testbed (+FAB)



FABRIC Topology - with FAB Sites





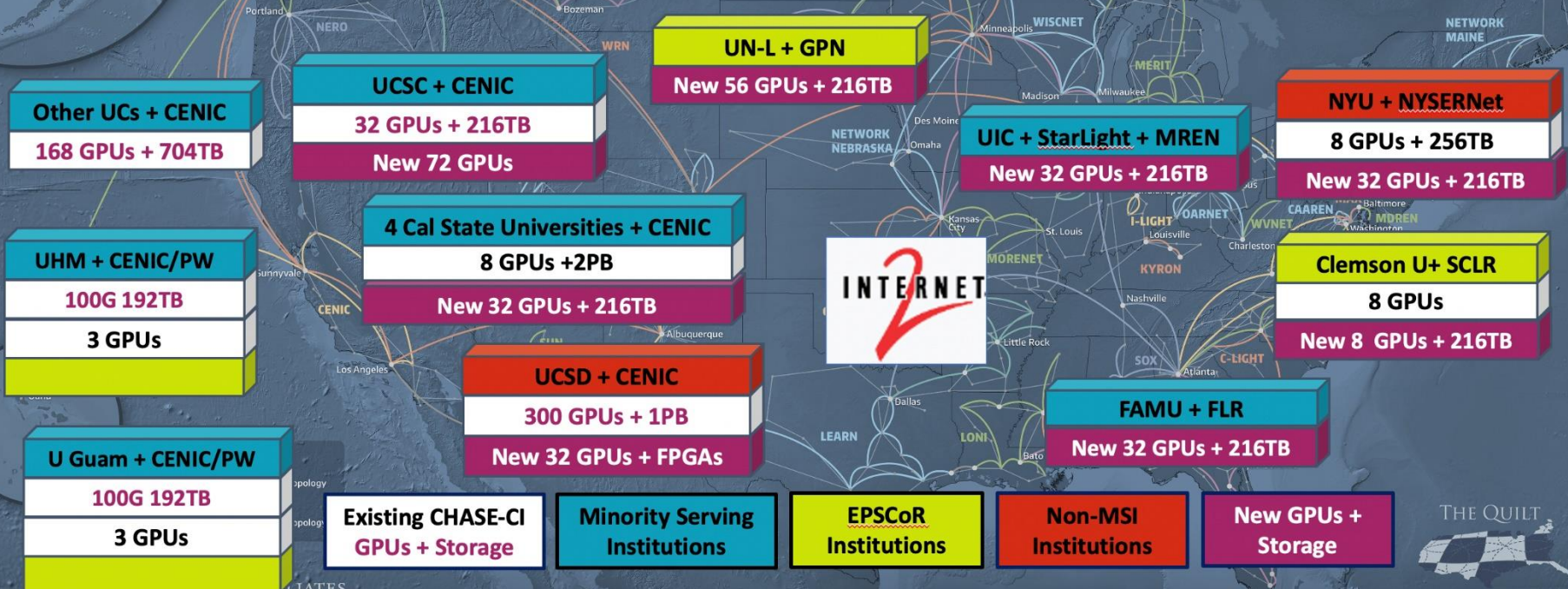
The image displays a stacked area chart with approximately 15 distinct colored layers, representing different data flows or network components. The layers are stacked from bottom to top in the following order: blue, red, orange, green, purple, cyan, pink, light green, red, blue, purple, cyan, yellow-green, orange, and dark purple. The total height of the stack increases significantly from left to right, indicating a substantial increase in total data volume. A white mouse cursor is positioned over the orange layer in the lower right quadrant of the chart.

**SCinet “Data Tsunami”  
Near 5 Tbps**



# REGIONAL RESEARCH AND EDUCATION NETWORKS IN THE UNITED STATES

## Proposed Extension of Nautilus 2021-2024



12-6-2017

# 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)

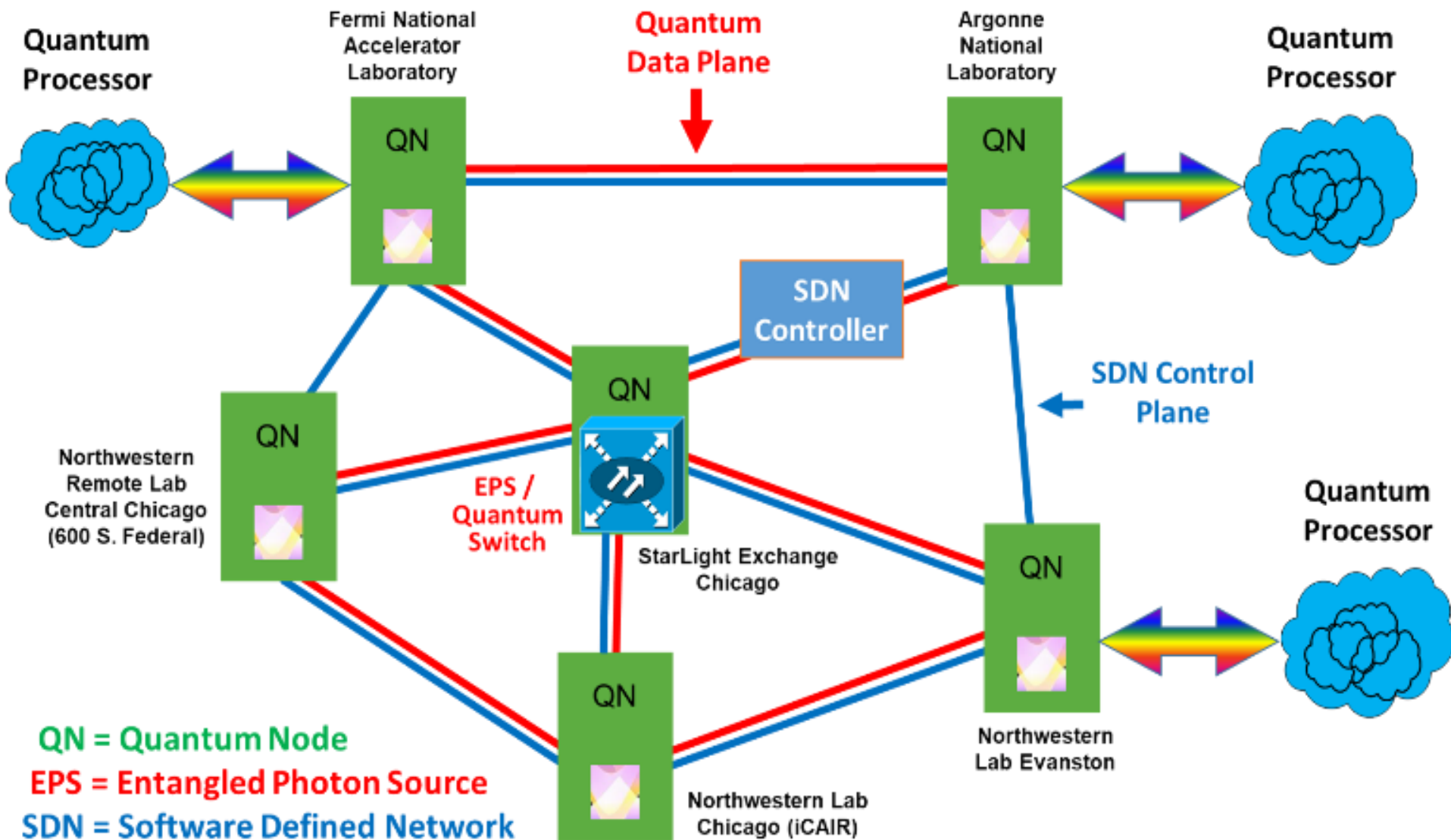


# Complexity Of Challenges Requires Consortia

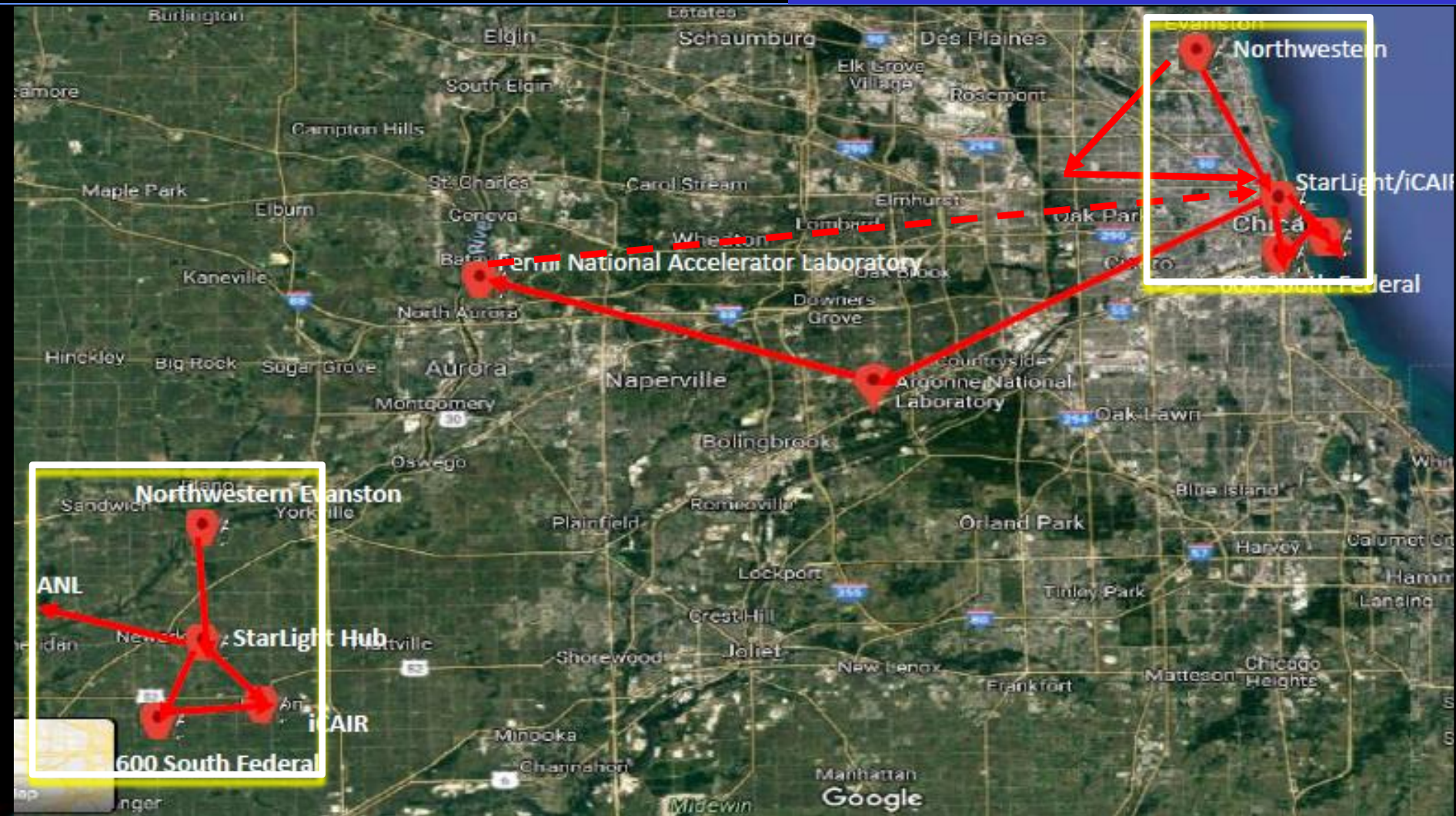
- **Northwestern University Established INQUIRE (Initiative at Northwestern for Quantum Information Research and Engineering), For Quantum Science Research**
- **This Initiative Participates in the Chicago Quantum Exchange and The Illinois Express Quantum Network, which includes the U.S. Department of Energy's Argonne National Laboratory, Fermi National Accelerator Laboratory, Multiple Research Universities, and Several Corporations.**
- **These National Laboratories, Northwestern University, Including the International Center for Advanced Internet Research (iCAIR), the StarLight International/National Communications Exchange Facility Consortium, the Metropolitan Research and Education Network (MREN), the Illinois Quantum Information Science and Technology Center (IQUIST) at the University of Illinois at Urbana-Champaign, And Other Research Partners, Including Internationally, Are Collaborating On This initiative.**



# Emerging Chicago Quantum Exchange Testbed



# Energizing IEQnet Testbed Topology



**BOUGHTON ROAD  
TOLL PLAZA**



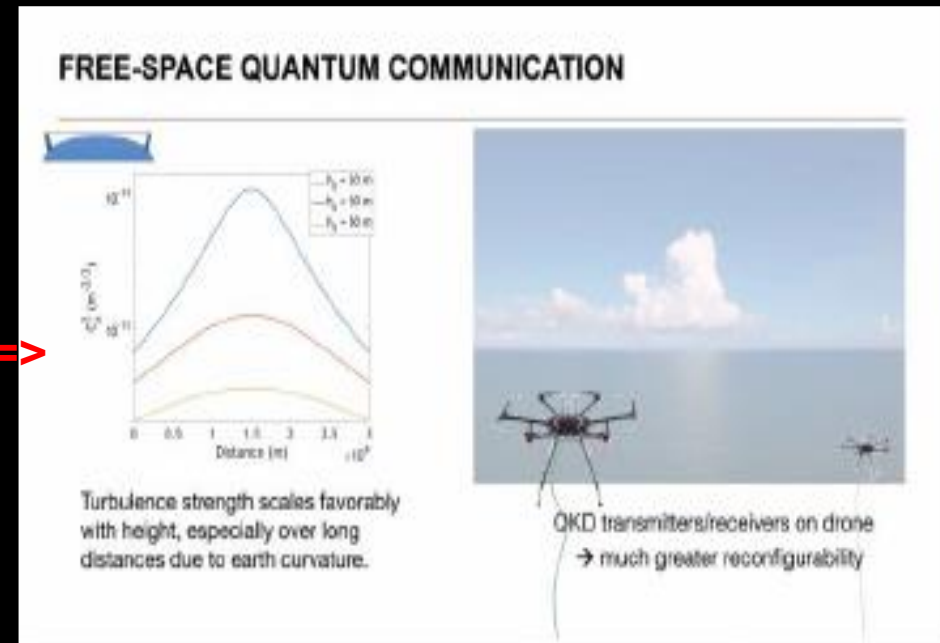
**Argonne**   
NATIONAL LABORATORY



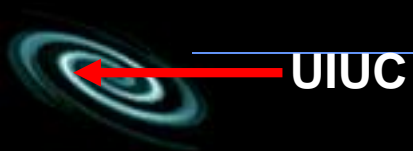


# IQUIST Quantum Network Testbed: QUIUC-NET

(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

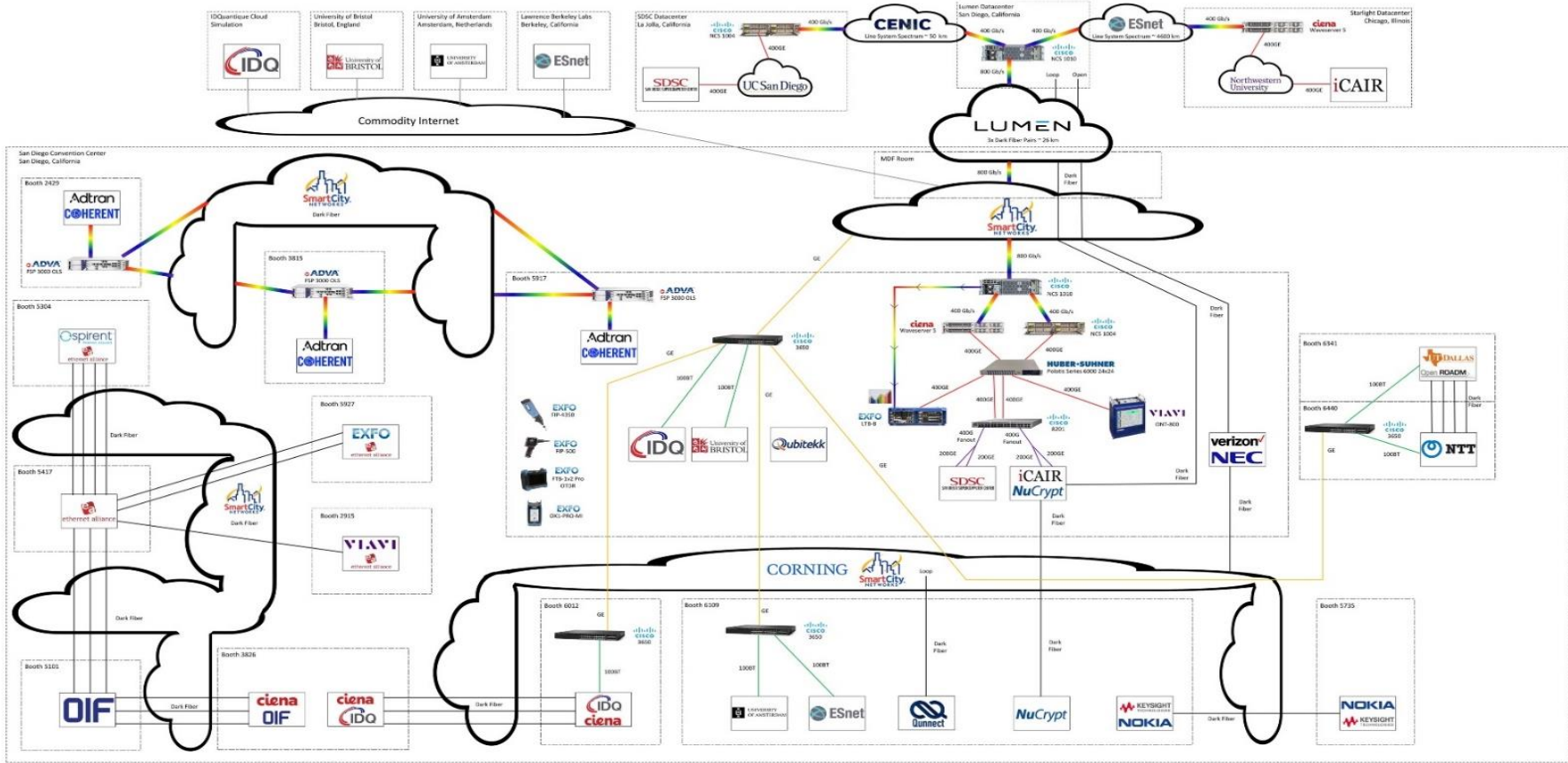




**Demo Lead Partner - NuCrypt (1) - Distribution of Quantum Entanglement Through Fiber With Co-Propagating Classical Data**

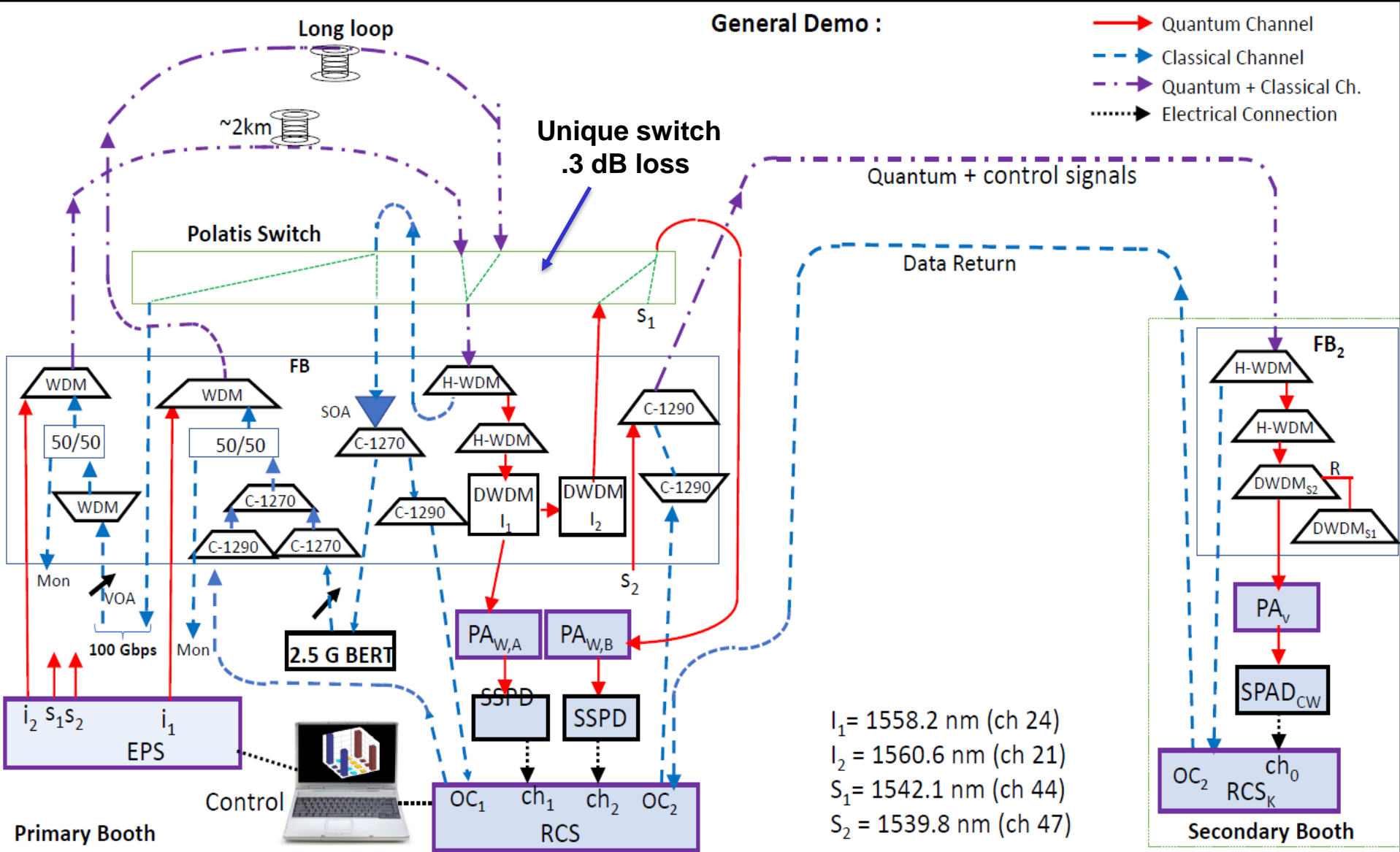
**(1) Spin Off From Northwestern University's Center for Photonic Communications and Computing, Which Was Also A Partner for the OFC 2023 Demonstrations (Prem Kumar, Director)**





- Dark Fiber
- 100Base-T
- Gigabit Ethernet
- 200 Gigabit Ethernet
- 400 Gigabit Ethernet
- DWDM
- OFCnet Demonstration

## OFC 2023 – OFCnet Architecture Diagram





**Co-Propagation  
And 400 Gbps WAN  
Demonstrations  
OFCnet Booth  
March 2023**

**STARLIGHT<sup>SM</sup>**



[www.startup.net/starlight](http://www.startup.net/starlight)

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

STARLIGHT<sup>SM</sup>