

And Plans For SC23 Demonstrations

Edoardo Martelli, Carmen Misa Moreira, Joe Mambretti,
Bruno Hoefft, Tom Lehman, Shawn McKee, Marian Babik,
Vitaliy Kondratenko, Tristan Sullivan, et al,

LHCOPN-LHCONE MEETING #50

FZU INSTITUTE OF PHYSICS OF THE CZECH ACADEMY OF SCIENCES

PRAGUE CZECH REPUBLIC

APRIL 18-19, 2023

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



Schematic overview of the GNA-G AutoGOLE

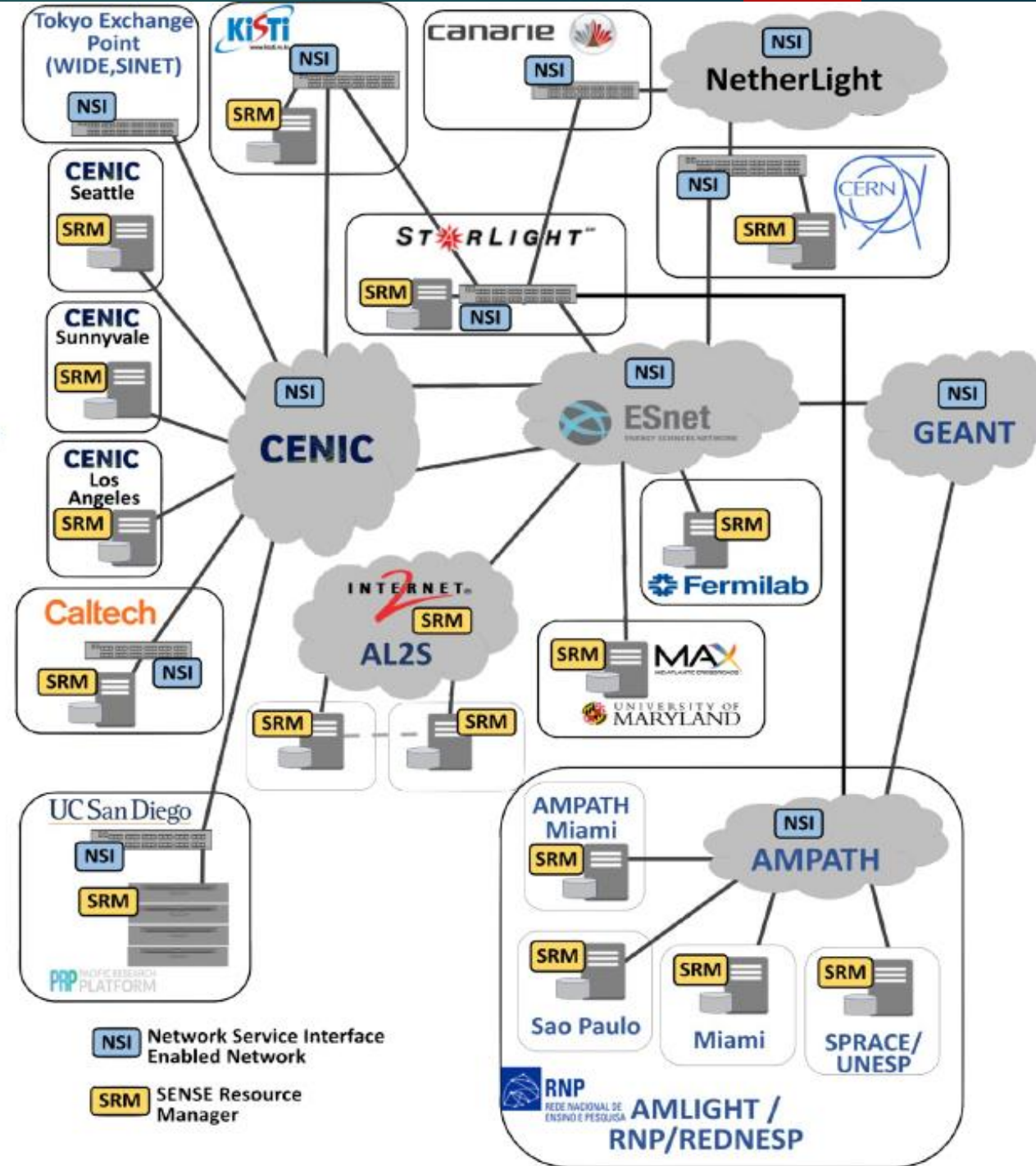


AutoGOLE Open R&E Exchanges

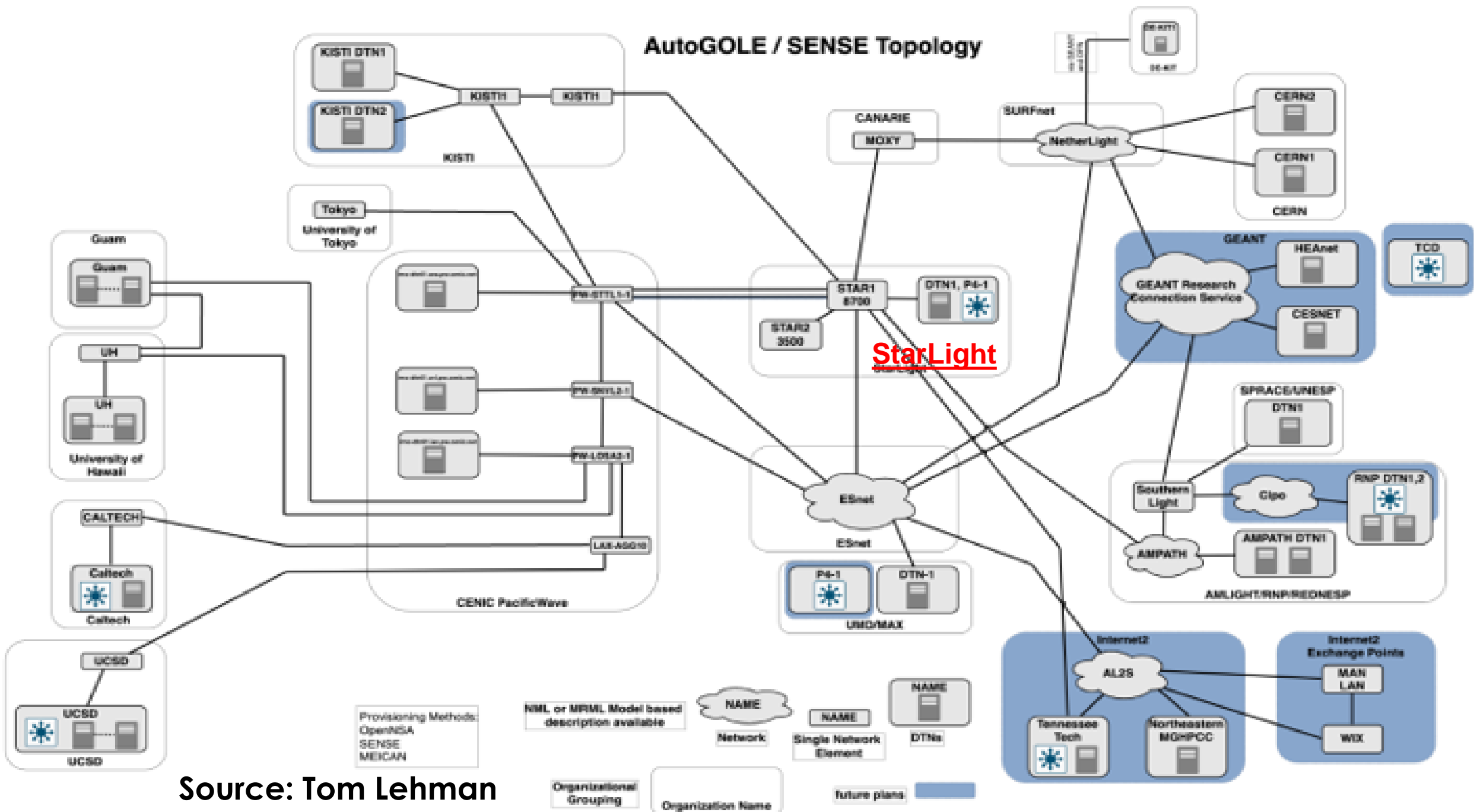
SENSE/AutoGole

- AutoGOLE, NSI, and SENSE working together provide the mechanisms for complete end-to-end services which includes the network and the attached End Systems (DTNs).

Source: Tom Lehman



AutoGOLE / SENSE Topology



Source: Tom Lehman

SENSE provisioning system

SENSE (SDN for E2E Networked Science at the Exascale): provision system that dynamically builds end-to-end virtual guaranteed networks across administrative domains without manual intervention.

- ❑ Provisioning automation: bring-up and management of services without human involvement.
- ❑ Multi-domain: multiple administrative domains, independent policies and AUP (Acceptable Use Policy).
- ❑ Resource orchestration: allocation and reservation of resources including compute, storage and network.
- ❑ End-to-end: DTN NIC to DTN NIC, across Science DMZ (Demilitarized zone), WANs, Open exchange points...



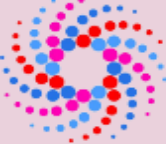
ESnet

ENERGY SCIENCES NETWORK

MEICAN: AutoGOLE front-end UI



Management Environment of Inter-Domain Circuits for Advanced Networks

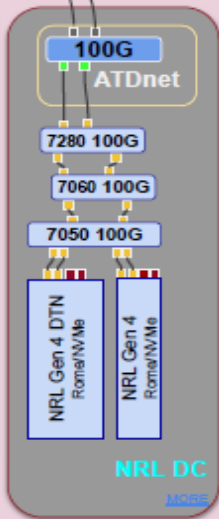
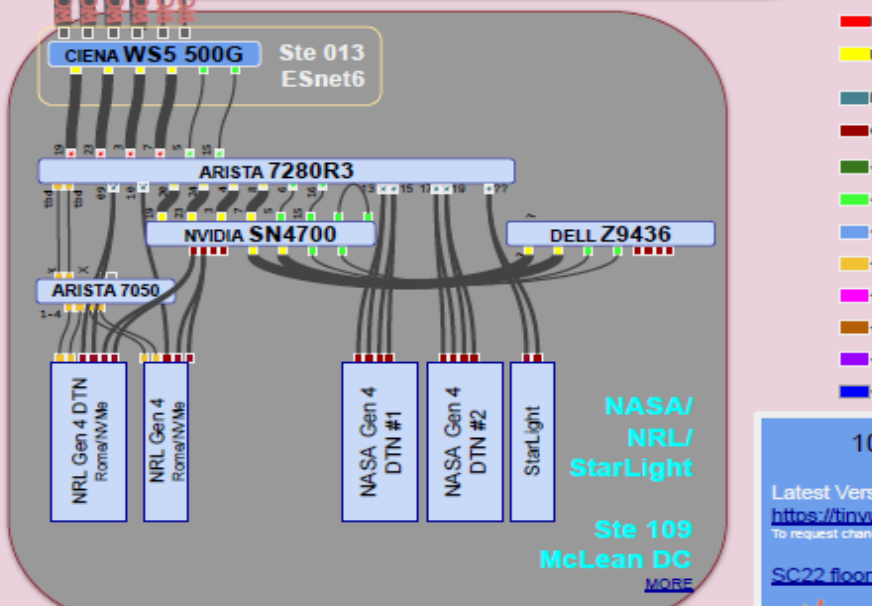
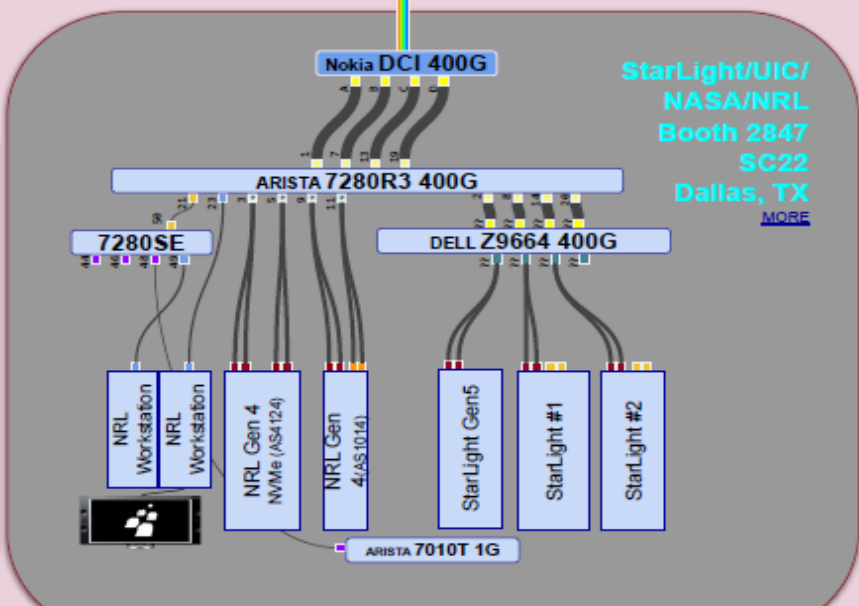
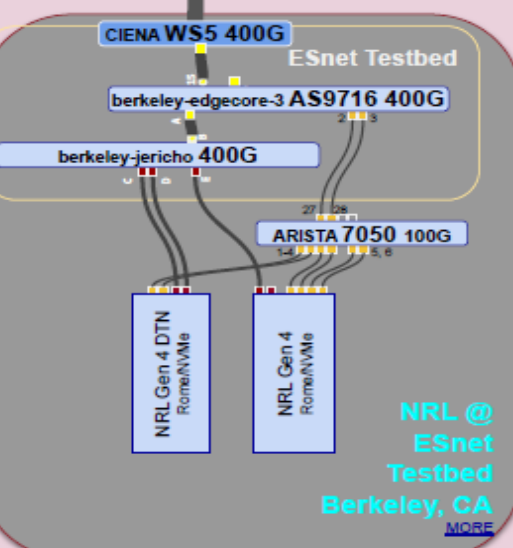
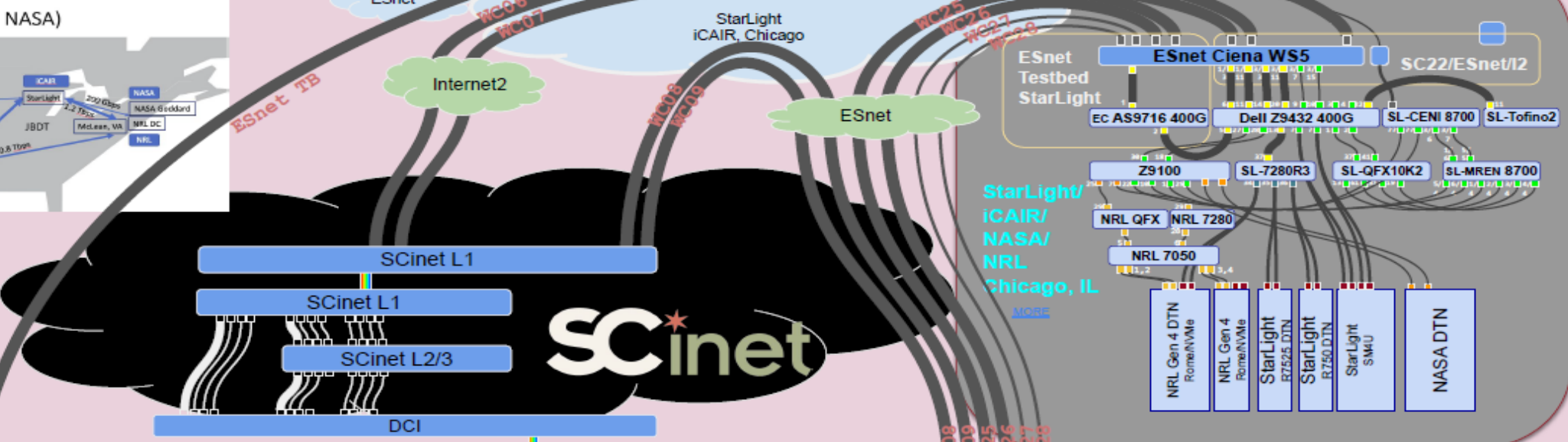


SC22

Dallas, TX | hpc accelerates.

JOINT BIG DATA TESTBED

JBDT SC22 WAN (NW iCAIR, NRL, NASA)



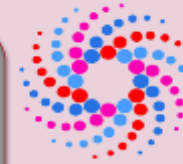
- 400G - LR4
- 400G - FR4
- 400G - DAC
- 200G - SR4 or DAC
- 100G - CLR4
- 100G - LR4
- 100G - SR4
- 100G - DAC
- 40G - SR4
- 40G - DAC
- 10G
- 1G

10/22/2022

Latest Version at:
<https://jinwul.com/SC22-JBDT>
 To request changes, please leave a comment

[SC22 floorplan](#)

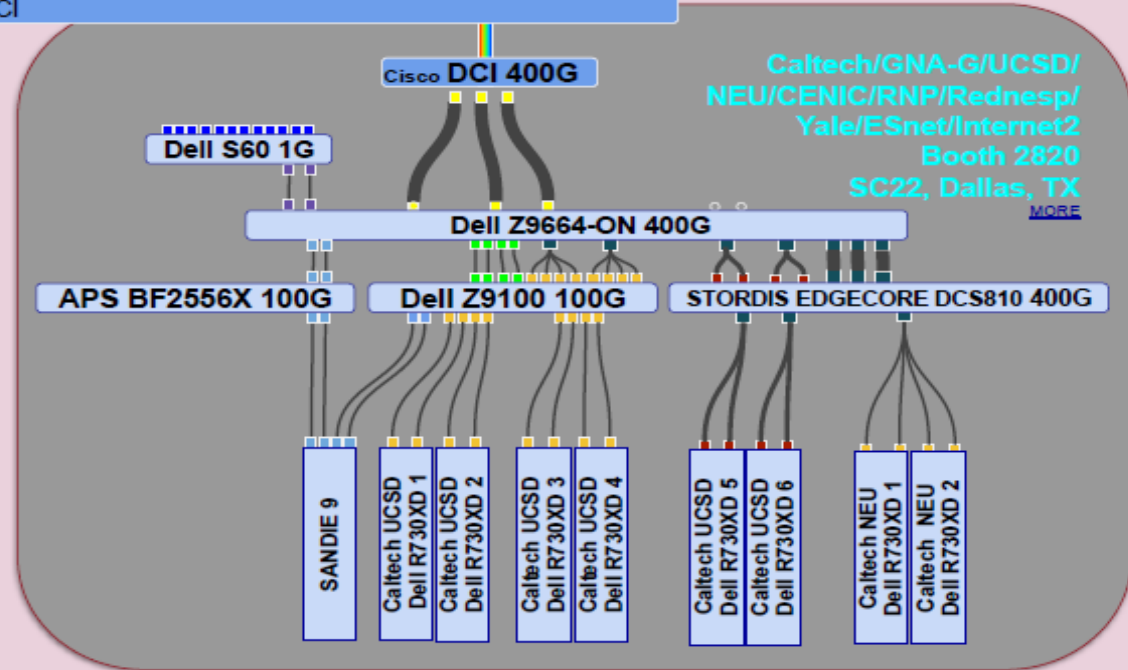
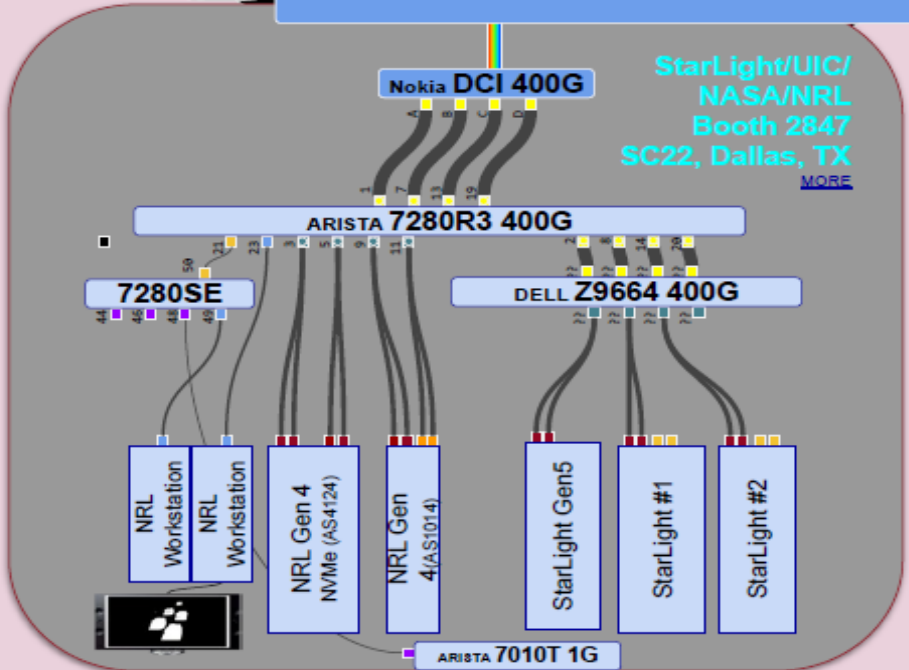
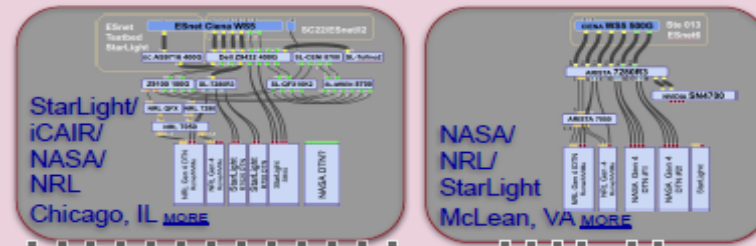
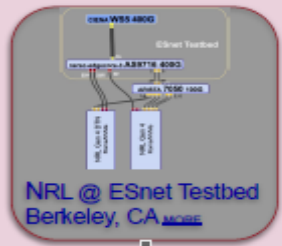
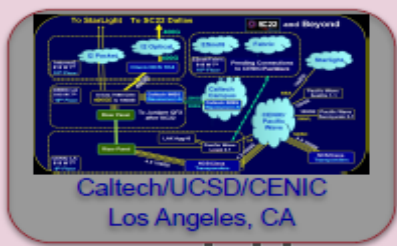
SC21 SC19



SC22

Dallas, TX | hpc accelerates.

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

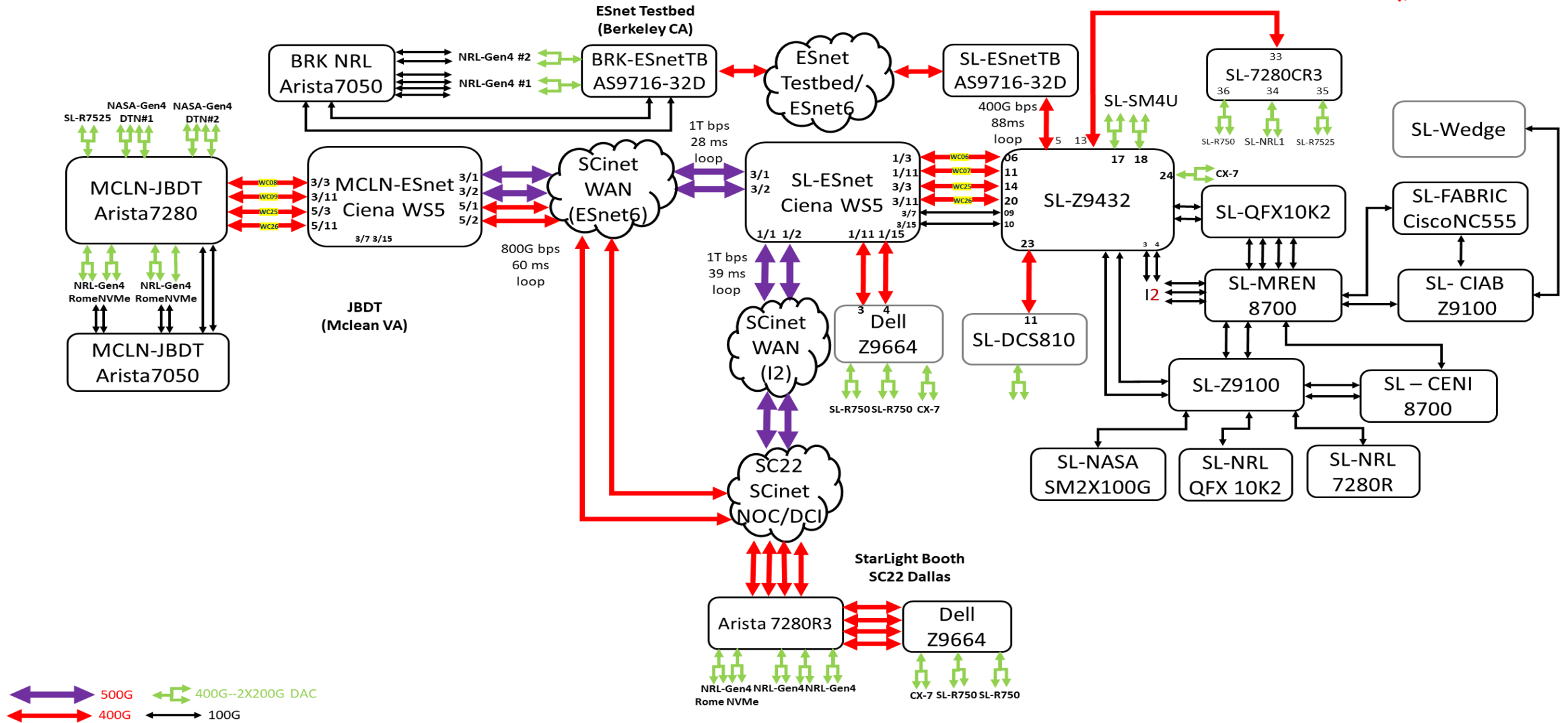
10/20/2022

Latest Version at:
<https://tinynrd.com/SC22-JBDT>
To request changes, please leave a comment

[SC22 floorplan](#)

SC22 SCinet WAN Testbed

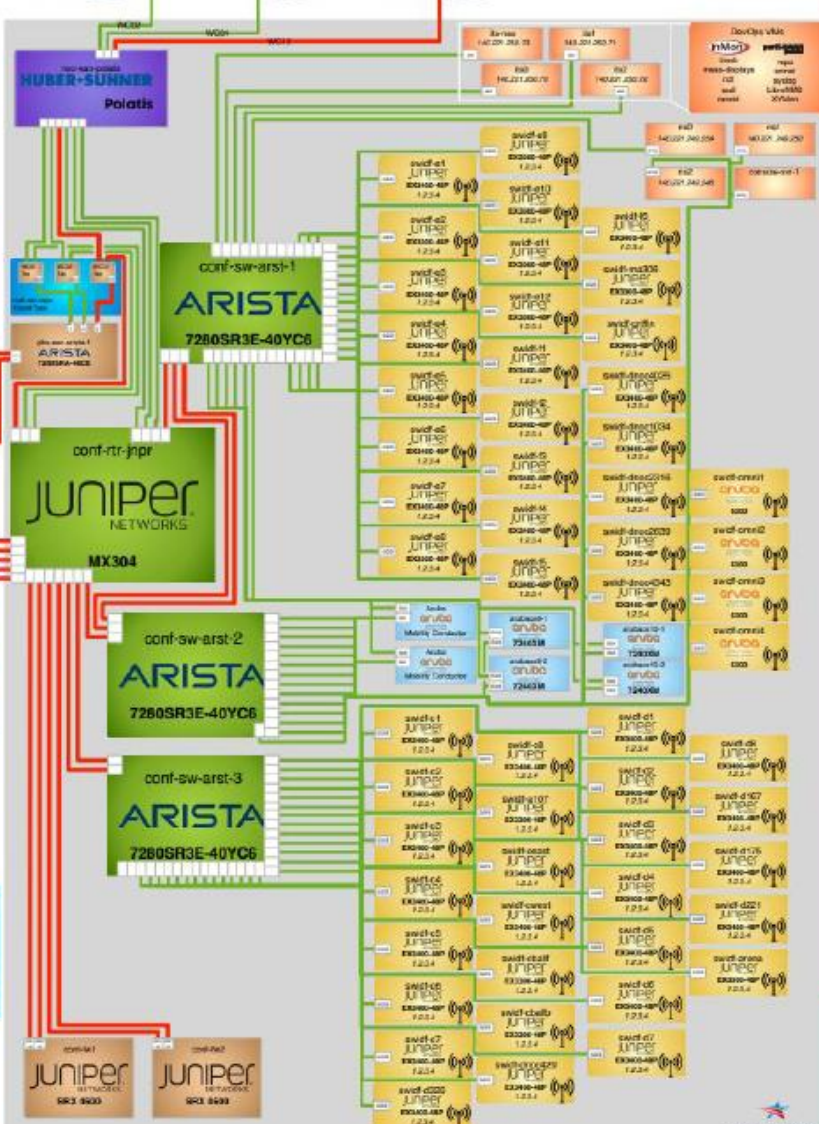
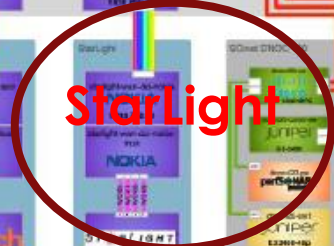
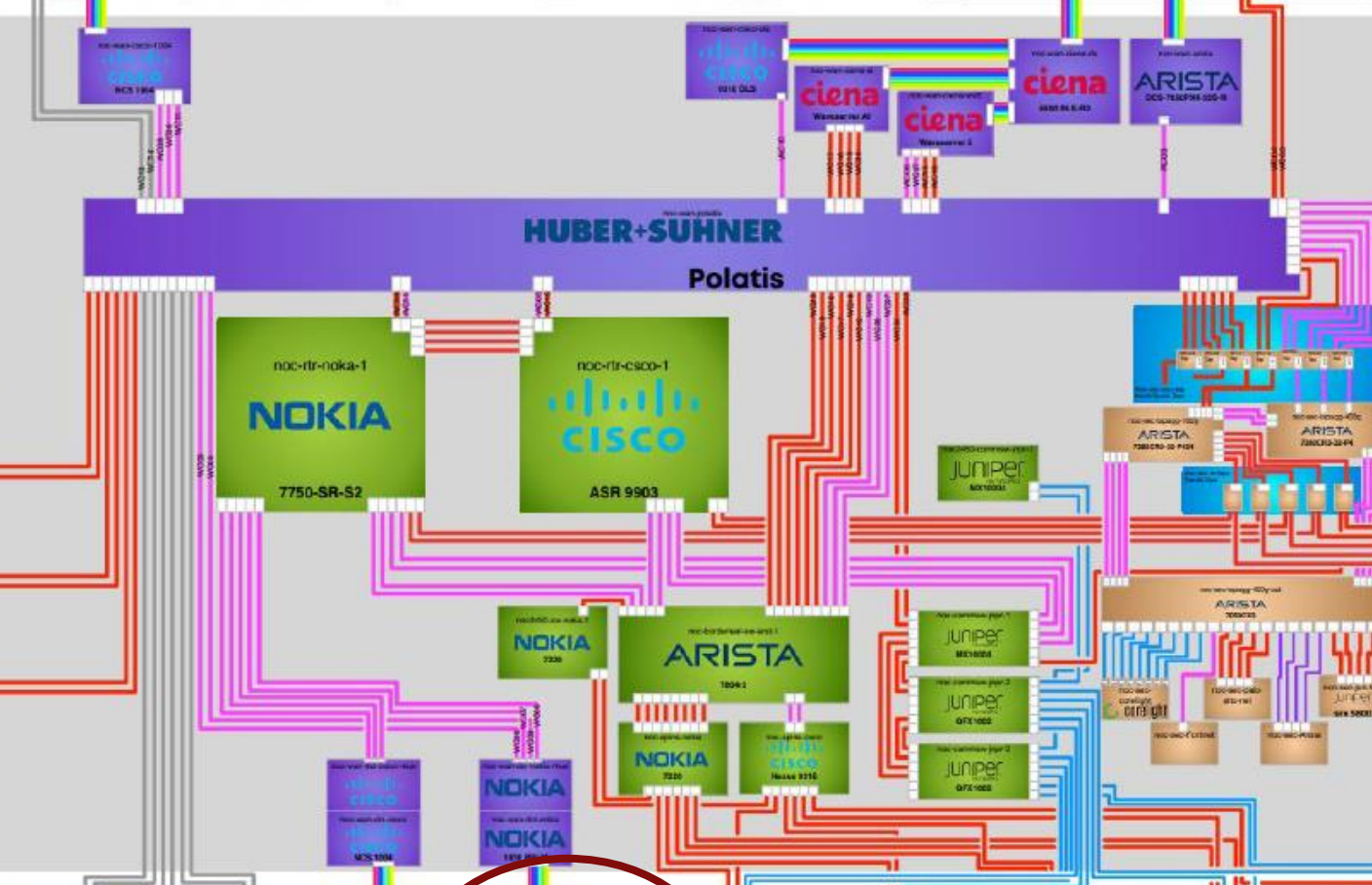
1T/800G/400G WAN Testbed by ESnet Testbed-I2-SCinet-StarLight-JBDT 11/14/2022



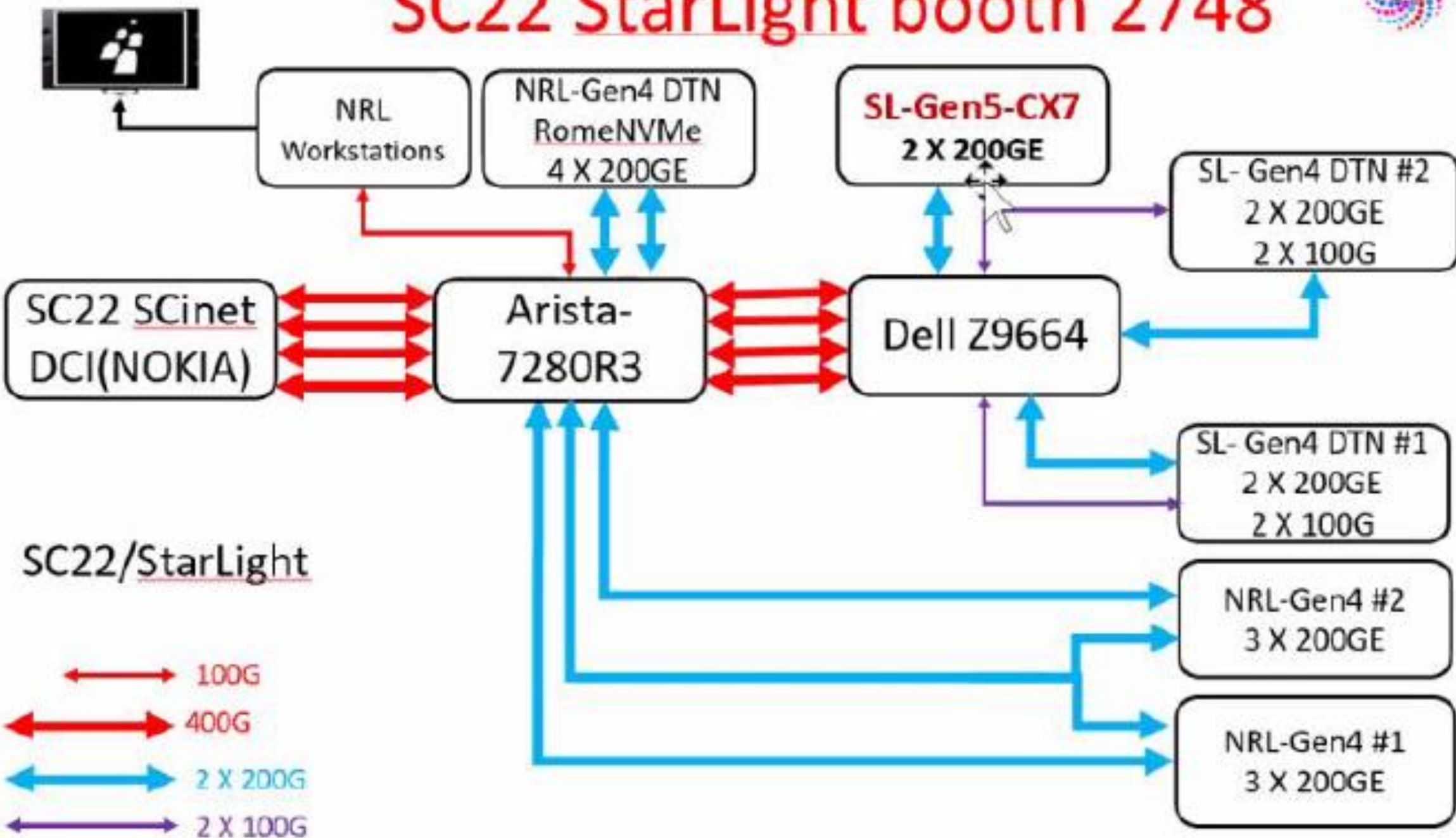
SCinet
Network Architecture
V7 - November 9th, 2022
Alexander Barnes TA00

- 1 Gigabit Ethernet
- 10 Gigabit Ethernet
- 40 Gigabit Ethernet
- 100 Gigabit Ethernet
- 400 Gigabit Ethernet
- 400G DCI
- Disk Fber
- DWDM OTN
- Wi-Fi Access Point

- Routing
- WAN
- Edge
- DevOps
- Security
- Wireless



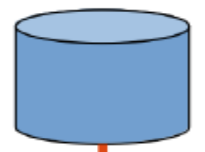
SC22 StarLight booth 2748



SC22/StarLight

Example SC22 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 (SciTags)**
- ▶ **Smart Amplified Group Environment Enhanced with Artificial Intelligence for Global Collaboration (SAGE3)**
- ▶ **JANUS Container Orchestration**



Rucio

FTS

NOTED at KIT

NOTED at CERN

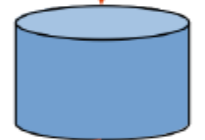
AutoGOLE SENSE



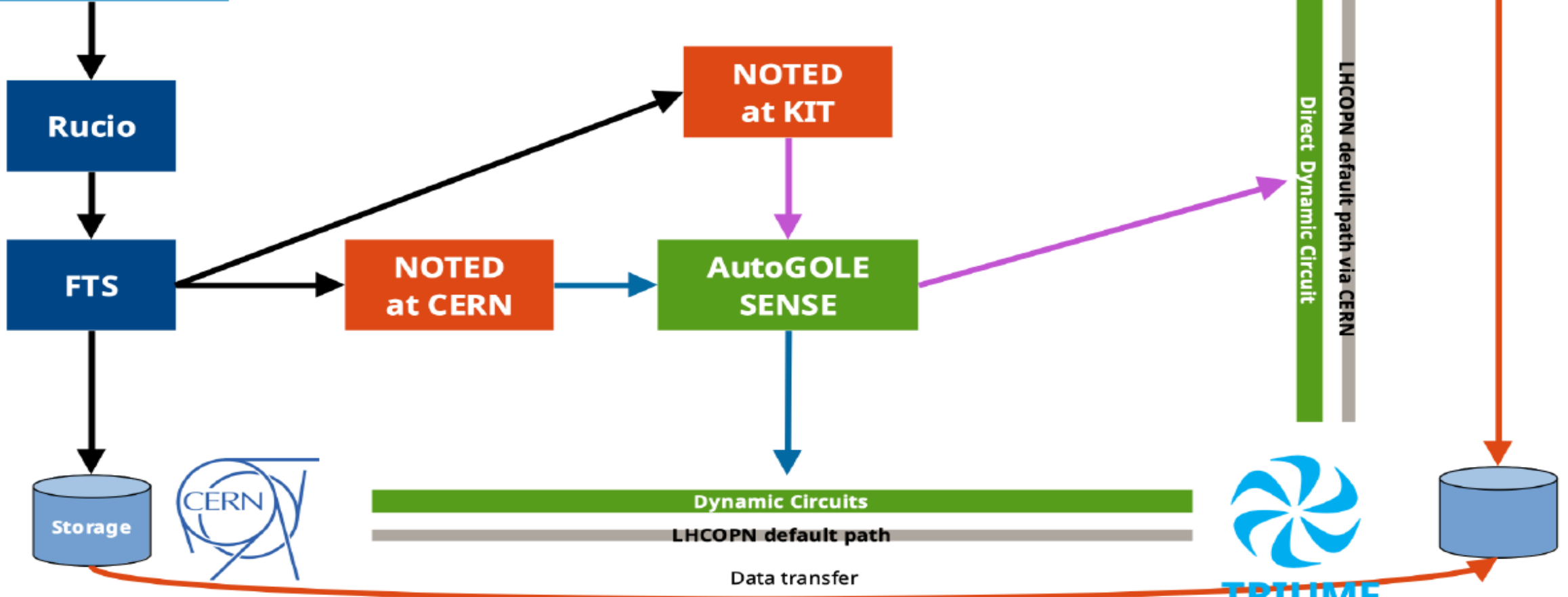
Data transfer



TRIUMF



FTS File Transfer Service



Components and participants

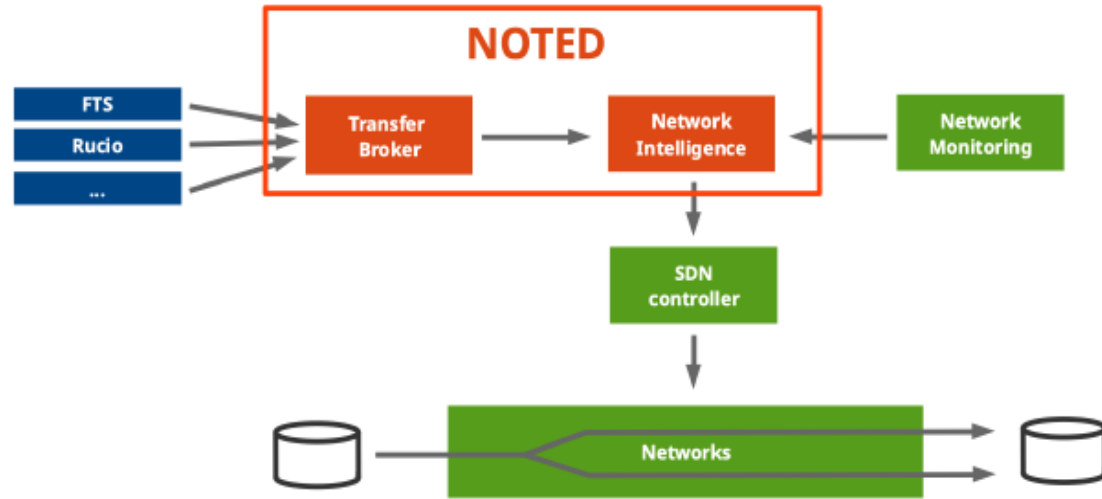
Components:

- ❑ NOTED controller and FTS at CERN.
- ❑ NOTED controller at KIT.
- ❑ Data storage at CERN, TRIUMF, KIT.
- ❑ AutoGOLE/SENSE circuits between CERN-TRIUMF and KIT-TRIUMF SENSE circuits are provided by ESnet, CANARIE, STARLIGHT, SURF.

Participants:



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.

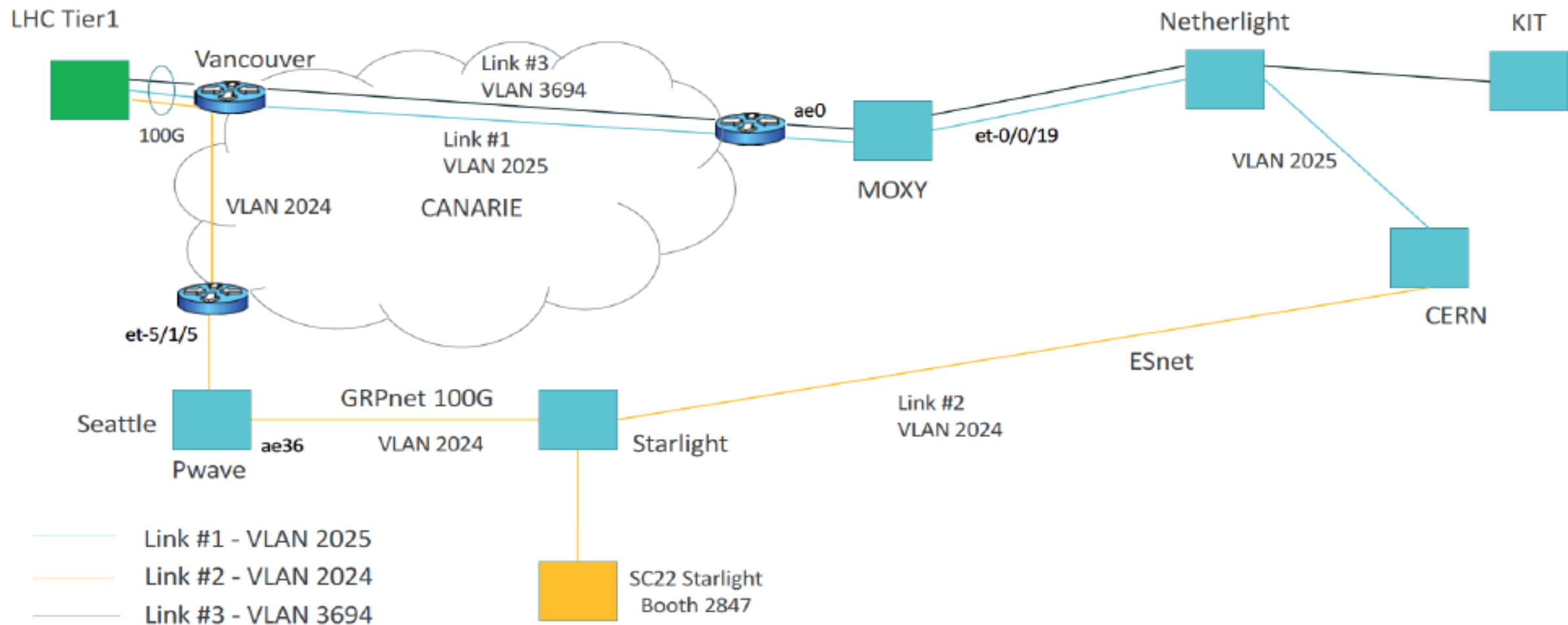
CRIC (Computing Resource Information Catalog):

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

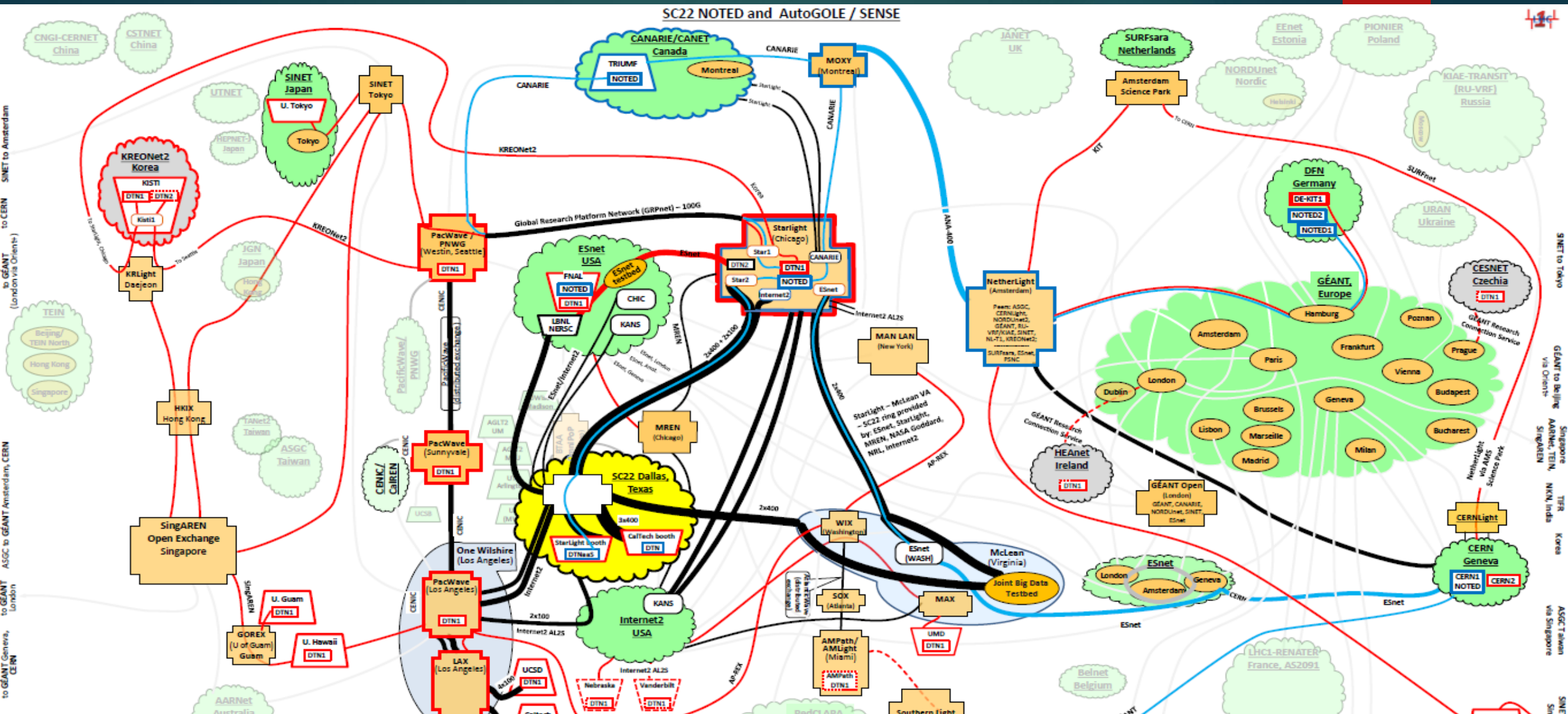
FLOWCHART AND DATASET STRUCTURE

- 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.

SC22 demo logical connections



SC22 NOTED and AutoGOLE / SENSE



Network Optimized Transport for Experimental Data (NOTED)

SC22 map v. 8 based on LHCONe Map Ver. 5.4, September 2020 – WEJohnston, ESnet, wej@es.net

SC22 NOTED infrastructure is in blue	AutoGOLE / SENSE infrastructure is in red	Shared or general infrastructure is in black	Pattern filled rectangles are computing/storage Elements	Solid fill rectangles indicate a DTN	Circuits supporting NOTED	Circuits supporting AutoGOLE / SENSE	Multi color lines indicate shared facility	"Carrier hotels"	Ovals are mostly regional infrastructure / testbeds	Rounded rectangles are individual switch/router	Sites	Exchange points (external or internal to a site)	Cloud bubbles are collections of associated infrastructure	100G	200G	400G	800G	1.2 Tb/s	Planned facility
--------------------------------------	---	--	--	--------------------------------------	---------------------------	--------------------------------------	--	------------------	---	---	-------	--	--	------	------	------	------	----------	------------------

NOTES
 1) Files are at <https://www.dropbox.com/sh/p2wcyrvubek7g/AAAMJGF509sxUFospm3pRLzLa?dl=0>



NOTED: AN INTELLIGENT NETWORK CONTROLLER TO IMPROVE THE THROUGHPUT OF LARGE DATA TRANSFERS IN FILE TRANSFER SERVICES BY HANDLING DYNAMIC CIRCUITS

Carmen Misa Moreira¹, Edoardo Martelli¹, Bruno Hoelt², Vitaliy Kondratenko³, Joel J. Mambretti⁴

¹CERN (Conseil Européen pour la Recherche Nucléaire), IT department CS-NE section, email: firstname.lastname@cern.ch
²KIT Karlsruhe Institute of Technology, Steinbuch Center for Computing, email: Bruno.Hoelt@kit.edu
³TRIUMF ATLAS Tier-1 Computing Center, email: vitaliy.k@triumf.ca
⁴International Center for Advanced Internet Research, Northwestern University, email: j.mambretti@northwestern.edu



SKELETON AND ELEMENTS OF NOTED

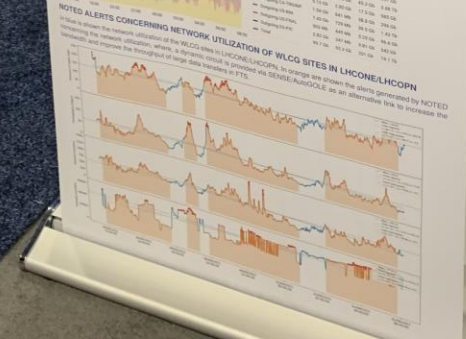
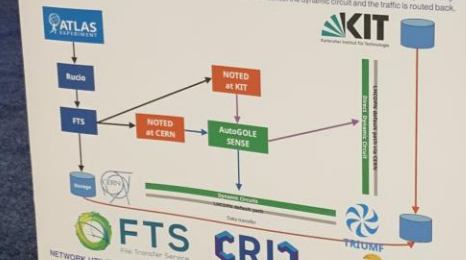


FTS (File Transfer Services):
 - Inspect and analyze data transfers to estimate if an action can be applied to optimise the network utilisation → get on-going and queued transfers.

CRIC (Computing Resource Information Catalog):
 - Enrichment to get an overview and knowledge of the network topology → get IPv4/IPv6 addresses, endpoints, route and latency.

FLOWCHART AND DATASET STRUCTURE

Input parameters: configuration given by the user
 In: noted.config.yaml → define a list of (src_rscale, dst_rscale), 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_rscale, dst_rscale) 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_rscale, dst_rscale) 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 Initiative

Leads= Shawn McKee, Marian Babik

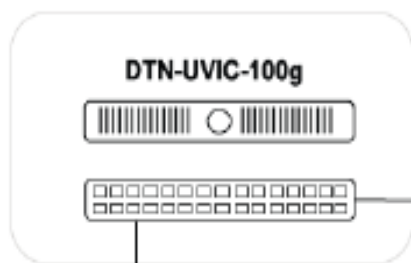
- **Scientific Network Tags** (scitags) is an initiative promoting identification of the science domains and their high-level activities at the network level.



- Enable tracking and correlation of our transfers with Research and Education Network Providers (R&Es) network flow monitoring
- Experiments can better understand how their network flows perform along the path
 - Improve visibility into how network flows perform (per activity) within R&E segments
 - Get insights into how experiment is using the networks, get additional data from R&Es on behaviour of our transfers (traffic, paths, etc.)
- Sites can get visibility into how different network flows perform
 - Network monitoring per flow (with experiment/activity information)
 - E.g. RTT, retransmits, segment size, congestion window, [etc.](#) all per flow

SC22 Packet/Flow Marking NRE

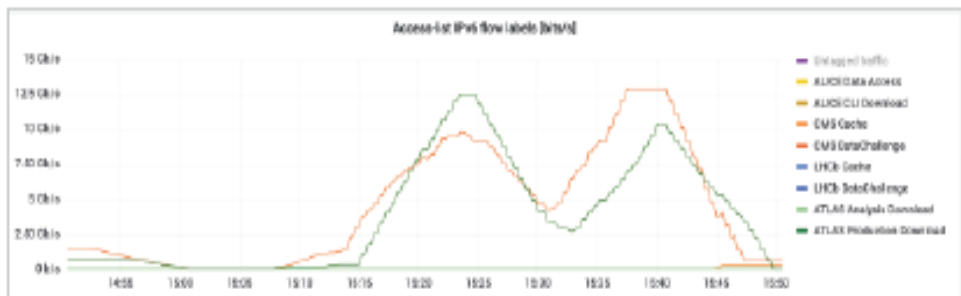
- ▶ The Goals of the SC22 Packet and Flow Marking NRE Demonstrations Were To Showcase The Capabilities of The Scitags Architecture And Methods For Optimizing Data Intensive Science
- ▶ Three Demonstrations Were Staged
 - ▶ IPv6 Packet Marking With eBPF-TC (100 Gbps)
 - ▶ XRootD Packet marking with Flowd+eBPF-TC
 - ▶ Accounting For Flow Labeled Packets Using a P4 Programmable Switch
- ▶ Participants:
 - ▶ CERN, StarLight, University of Victoria, KIT, ESnet, CANARIE



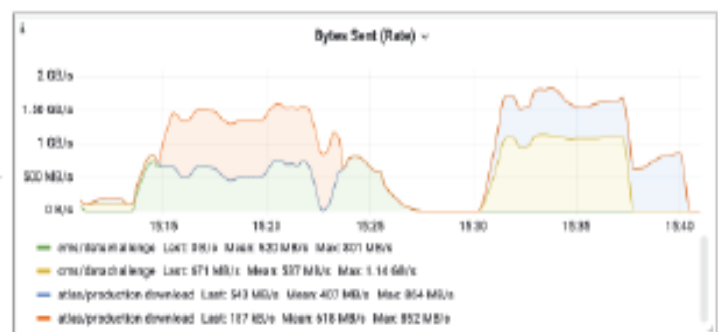
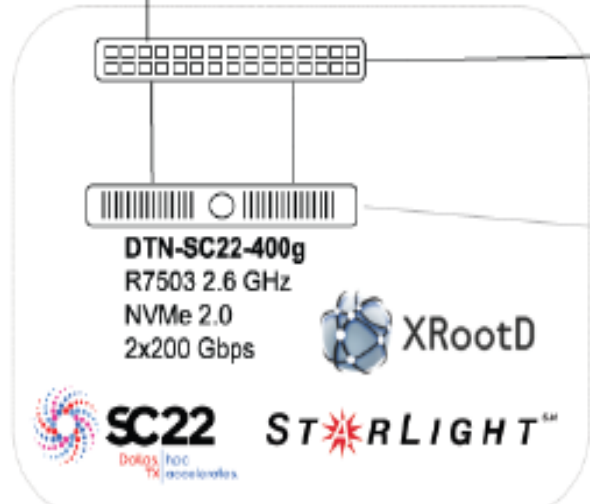
1. Clients requesting data transfers from/to DTN-SC22-400g while passing science domain and activity fields via transfer protocols.



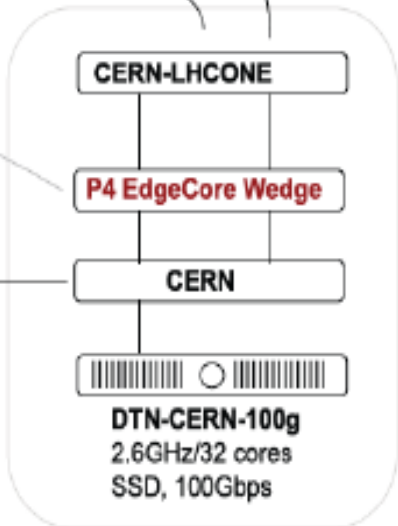
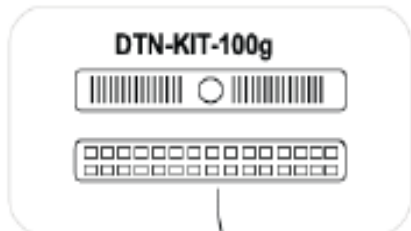
4. High performance tests using eBPF-TC filters to test encoding of the science domains and activity fields in the IPv6 flow label at scale.



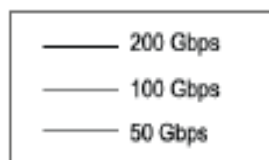
3. P4 programmable switch at CERN collecting the science domains and activity bits encoded in the packets.



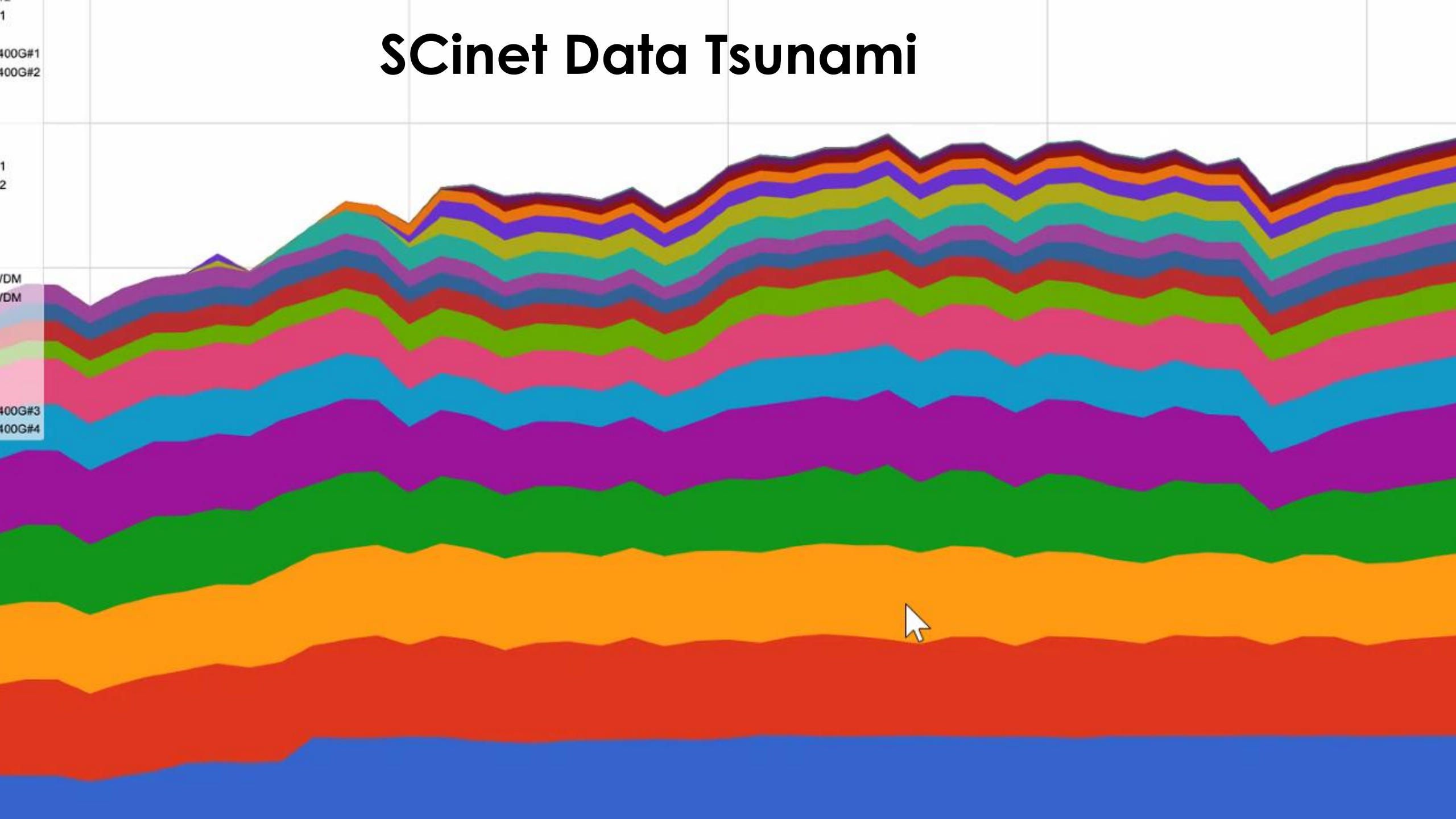
2. XRootD storage responds to the client requests and marks the data transfer packets with the corresponding science domain and activity.



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



SCinet Data Tsunami



Scientific Network Tags

Packet Marking for Data Intensive Scientific Workflows



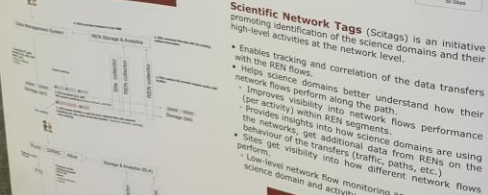
Managing large scale scientific workflows over networks is becoming increasingly complex, especially as multiple science projects share the same foundation resources simultaneously yet are governed by multiple divergent variables: requirements, constraints, configurations, technologies, etc. A key method to address this issue is to employ techniques that provide high fidelity visibility into exactly how science flows utilize network resources end-to-end. This demonstration showcases one such method, Scientific Network Tags (SciTags), an initiative that is promoting identification of the science domains and their high-level activities at the network level. This open system initiative provides open source technologies to help Research and Education Networks (REN) understand resource utilization while providing information to scientific communities on the behavior of their network flows.

SC22 scitags.org
 Packet Marking for Data Intensive Scientific Workflows

DFN **KIT**



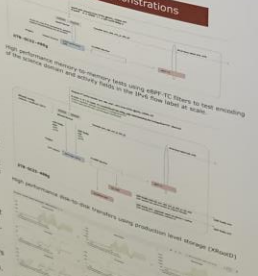
How SciTags Work?



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

- Enables tracking and correlation of the data transfers with the REN flows.
- Helps science domains better understand how their network flows perform along the path.
- Improves visibility into network flows performance (per activity) within REN segments.
- Provides insights into how science domains are using the networks; get additional data from RENs on the behaviour of the transfers (traffic, paths, etc.)
- Since get visibility into how different network flows perform.
 - Low-level network flow monitoring aggregated by the science domain and activity.

SC22 Demonstrations



What's next?

- Deploying and testing XRootD and iCache implementations.
- Organizing validation in collaboration with the WLCG experimental section of functionalities across a cross-domain proven and proven networks (Container/Future Data Distribution Systems).
- Develop and test storage systems, understand their requirements and work on SciTags implementation.
- Cooperate with storage systems, understand their requirements and work on SciTags implementation.
- Expanding available technologies for collecting SciTags at the network level.
- Testing and validating ways to generate flow identifiers for packet capture, flow analysis and monitoring systems.
- Plan to submit IETF informational RFC.



Planning for SC23 (Denver, Col)

- ▶ Network Research Exhibition (NRE) Descriptions (Due June 1)
- ▶ Defined Demonstration
- ▶ Assessment of Required Resources
- ▶ Implementation of Services/Resources
- ▶ Pre-Conference Staging Facilities
- ▶ Planned 1.2 Tbps StarLight Facility To StarLight Venue Booth,
1.2 Tbps StarLight Facility To Joint Big Data Testbed Facility
McLean Virginia

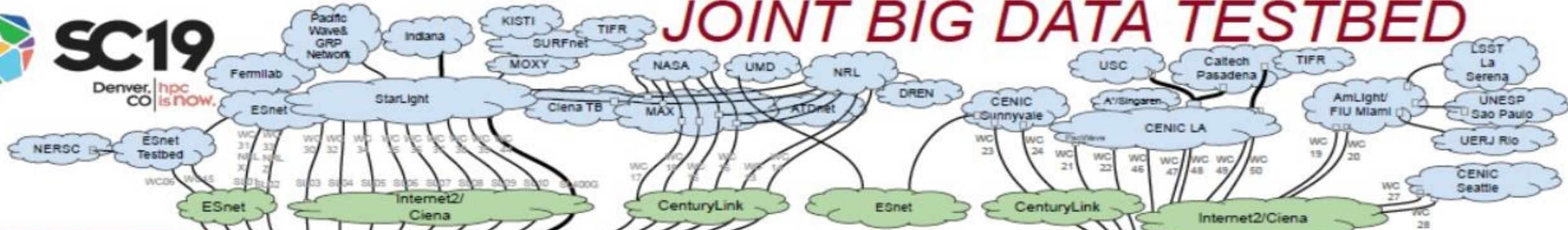




SC19

Denver, CO
hpc is now.

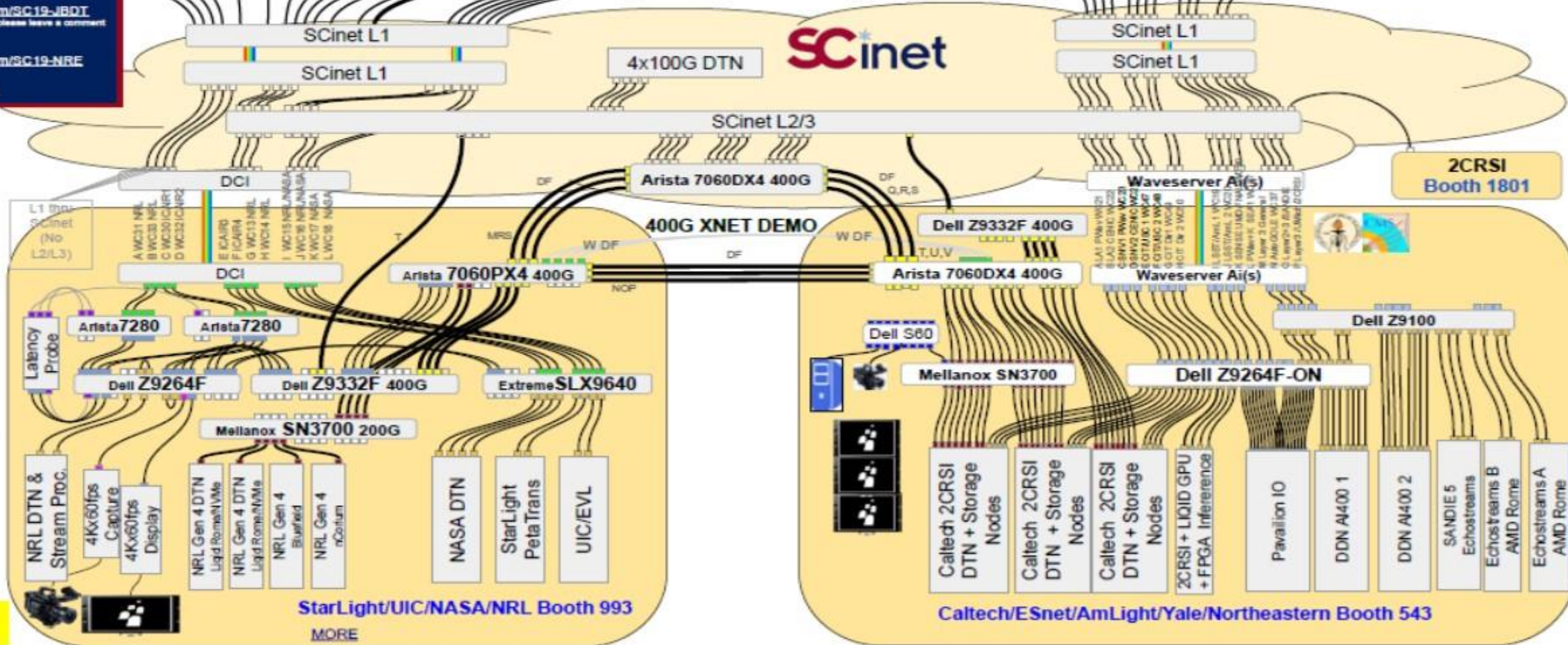
JOINT BIG DATA TESTBED



Latest Version at:
<http://tinyurl.com/SC19-IBDT>
 To request changes, please leave a comment

See also:
<http://tinyurl.com/SC19-NRE>

SC19 floorplan



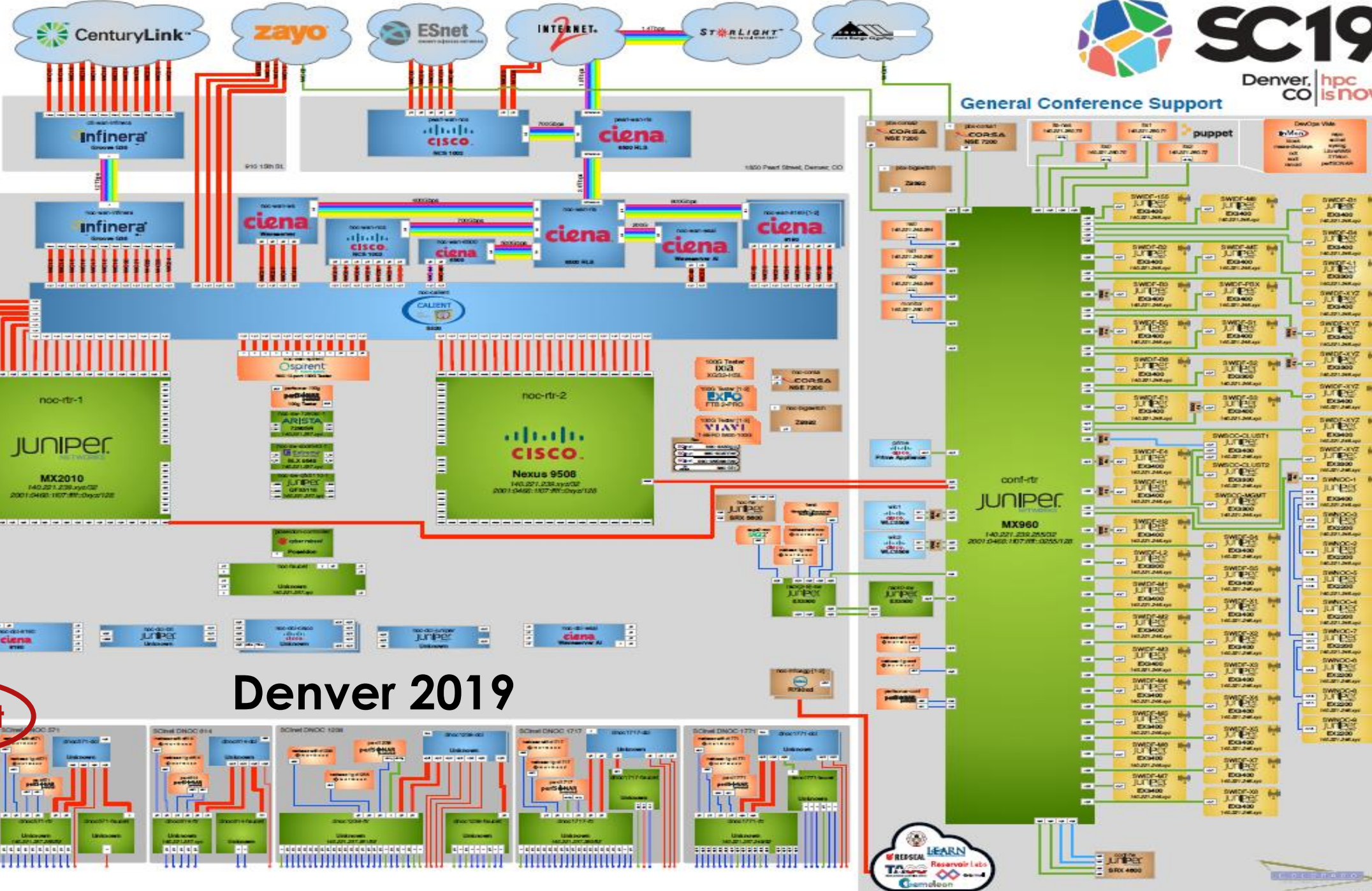
- 400G - FR4
- 200G - BR4 or DAC
- 100G - CLR4
- 100G - LR4
- 100G - SR4
- 100G - DAC
- 40G - SR4
- 40G - DAC
- 10G
- 10

StarLight/UIC/NASA/NRL Booth 993
 MORE

Caltech/ESnet/AmLight/Yale/Northeastern Booth 543

10/15/2019

- 1 Gigabit Ethernet
- 10 Gigabit Ethernet
- 40 Gigabit Ethernet
- 100 Gigabit Ethernet
- 400 Gigabit Ethernet
- 400G DCI
- 800G+ DCI
- DWDM OTN
- Wi-Fi Access Point



Denver 2019

StarLight

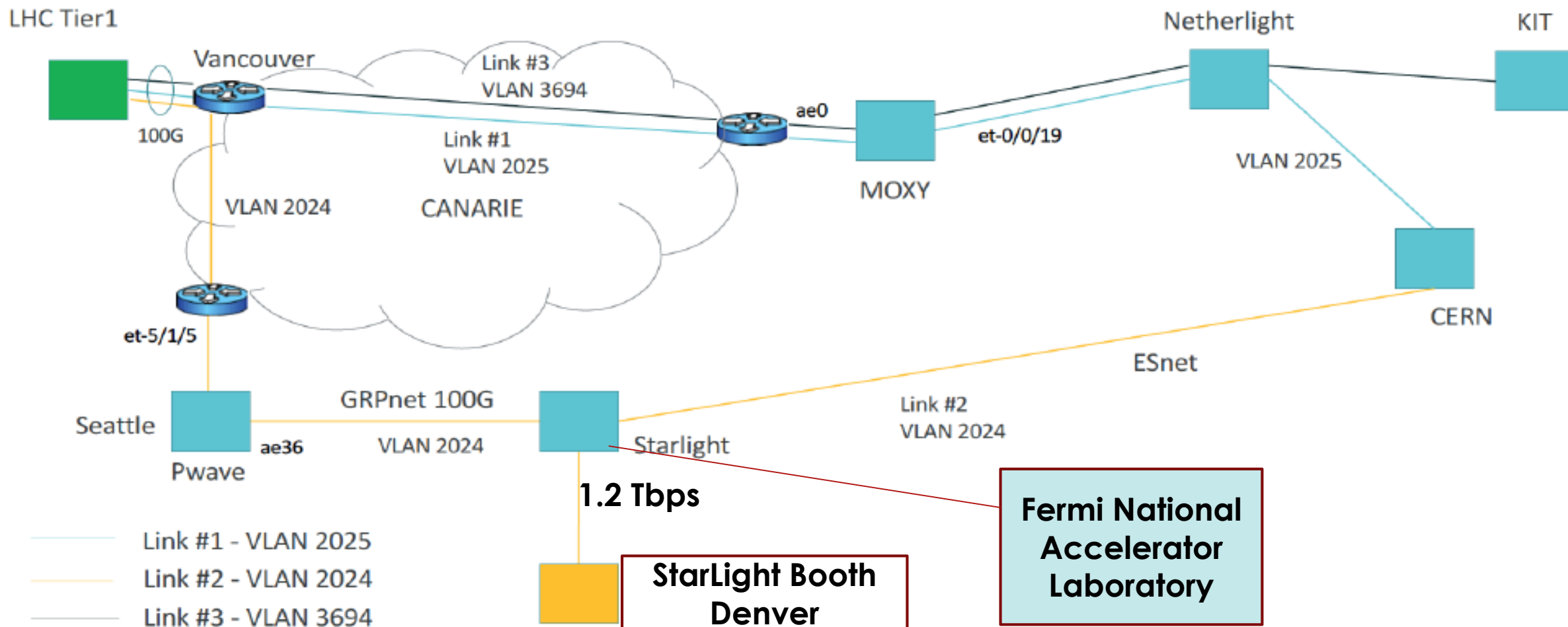
General Conference Support



SC23

demo logical connections

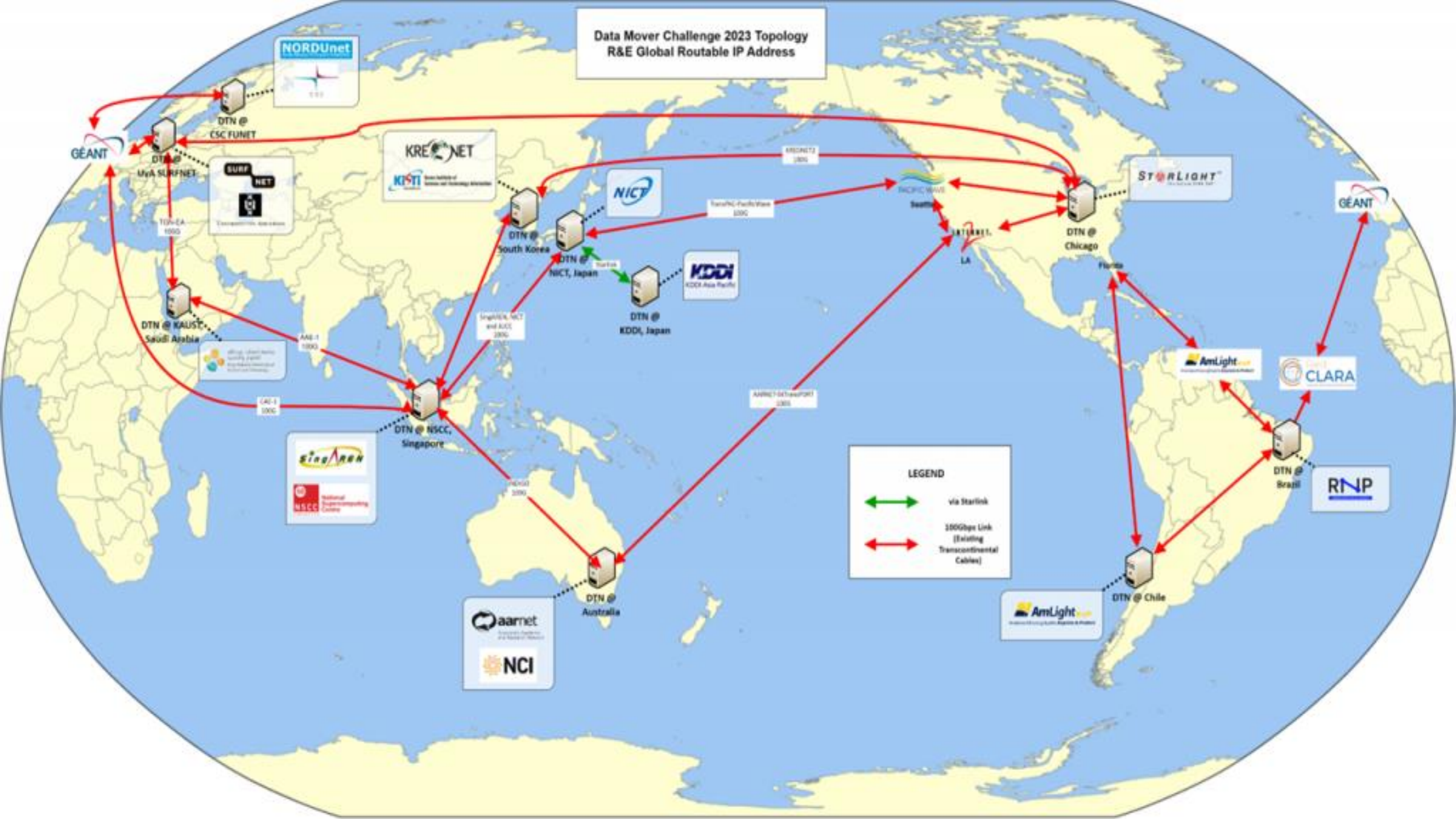
NOTED
AutoGOLE/SENSE



Proposed SC23 Packet/Flow Marking NRE

- ▶ Concept: The Goals of the SC23 Packet and Flow Marking NRE Demonstrations Will Be Building On the SC22 Demonstrations To Showcase The Capabilities of The Scitags Architecture And Methods For Optimizing Data Intensive Science
- ▶ Three Demonstrations Will Be Staged
 - ▶ IPv6 Packet Marking With eBPF-TC (100 Gbps)
 - ▶ XRootD Packet Marking with Flowd+eBPF-TC
 - ▶ Accounting For Flow Labeled Packets Using a P4 Programmable Switch
- ▶ Participants:
 - ▶ CERN, University of Victoria, KIT, ESnet, StarLight, CANARIE, Fermi National Accelerator Laboratory, SCInet, Digital Alliance?, Others?

Data Mover Challenge 2023 Topology
R&E Global Routable IP Address



Annual Global Research Platform Workshop – Co-Located With IEEE International Conference On eScience Oct 9-10, 2023



CALLS - PROGRAM - TRAVEL

'23 eScience

October 9-13, 2023

Limassol, Cyprus

IEEE eScience 2023 brings together leading interdisciplinary research communities, developers and users of eScience applications and enabling IT technologies. The objective of the eScience Conference is to promote and encourage all aspects of eScience and its associated technologies, applications, algorithms and tools with a strong focus on practical solutions and challenges. eScience 2023 interprets eScience in its broadest meaning that enables and improves innovation in data- and compute-intensive research across all domain sciences ranging from traditional areas in physics and earth sciences to more recent fields such as social sciences, arts and humanities, and artificial intelligence for a wide variety of target architectures including

Important Dates

- ~~February 10, 2023~~ **Friday, February 24, 2023**
Workshop Submissions
- ~~February 24, 2023~~ **Friday, March 10, 2023**
Workshop Acceptance Notification
- Friday, May 26, 2023**
Paper Submissions
- Friday, June 30, 2023**
Notification of Paper Acceptance



Thanks!

▶ **Questions?**