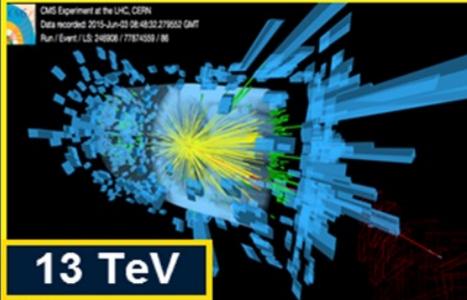


Global Network Advancement Group Next Generation Network-Integrated System for Data Intensive Sciences



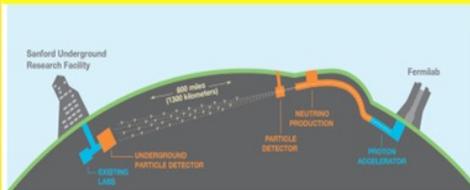
Booth 845



Rubin Observatory



LHC



LBNF/DUNE



SKA

**LHC Run 3
and HL-LHC**

**Rubin
Observatory**

SKA

Bioinformatics

**Earth
Observation**

**Gateways
to a New Era**

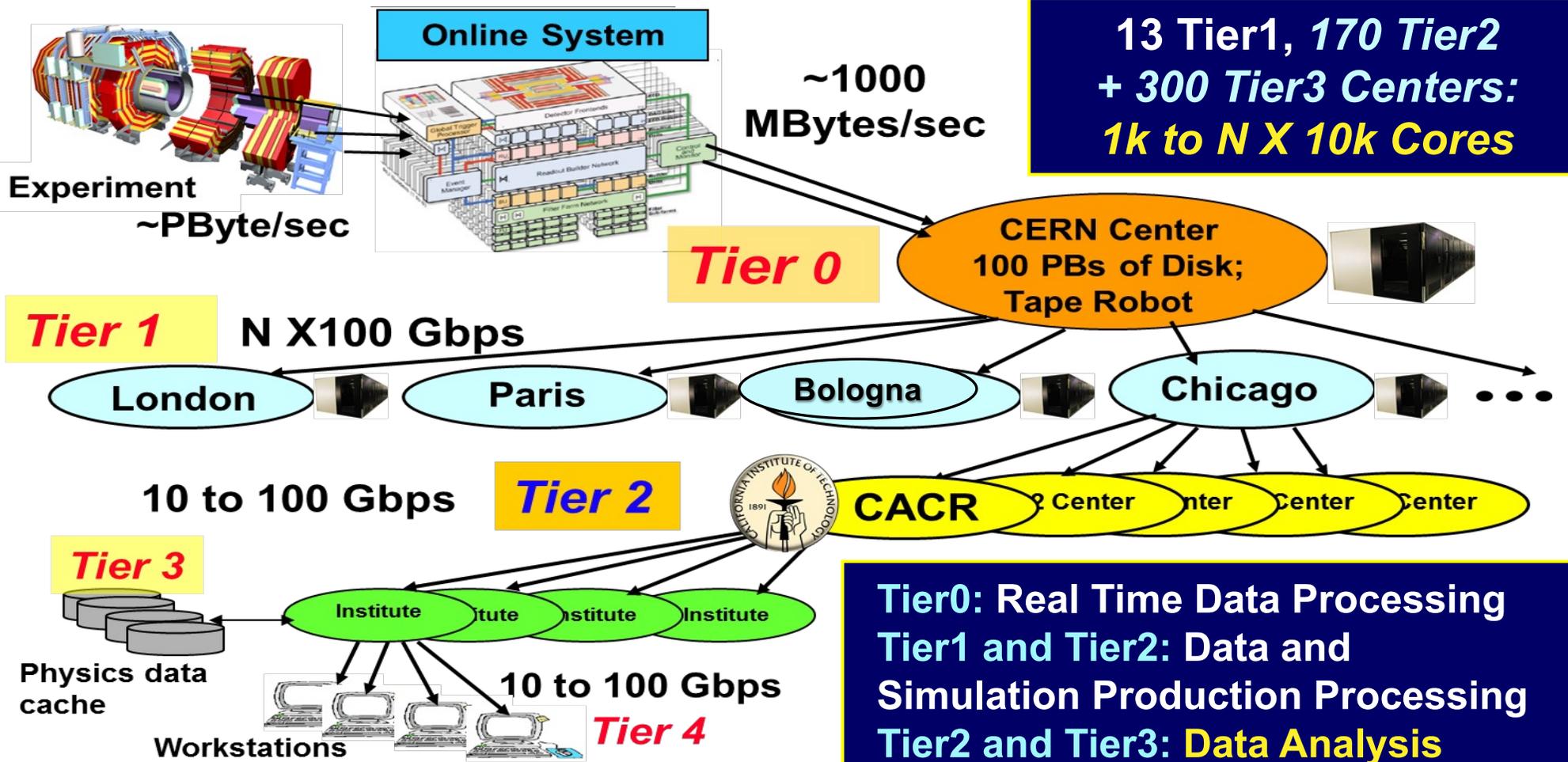


**SC24 Network Research Exhibition
NRE-13 and 9 Partner NREs
INDIS Workshop November 18, 2024**





Global Data Flow: LHC Grid Hierarchy A Worldwide System Invented by Caltech (1999)



A Global Dynamic System
Increased "Elastic" Use of Additional HPC and Cloud Resources
Fertile Ground for Control with Ai/ML

LHC: Discovery of the Higgs Boson and Beyond; 75 Years of Exploration !

Physicists Find Elusive Particle Seen as Key to Universe



2013 Nobel Prize
Englert **Higgs**



	Energy Frontier	Intensity Frontier	Cosmic Frontier
Higgs Boson	●		
Neutrino Mass		●	●
Dark Matter	●	●	●
Cosmic Acceleration			●
★ Explore the Unknown	●	●	●

48 Year Search; 75 Year Exploration

Theory (1964): 1950s – 1970s;

LHC + Experiments Concept: 1984

Construction: 2001; Operation: 2009

Run1: Higgs Boson Discovery 2012

Run2 and Going Forward:

Precision Measurements and BSM

Exploration: 2013 - 2042



Advanced Networks Were Essential to Higgs Discovery and Every Ph.D Thesis; They will be Essential to All Future Discoveries

- **NOTE: ~85% of Data Still to be Taken**
- **Greater Intensity: Upgraded detectors for more complex events**
- **To 5X Data Taking Rate in 2029-40**

The Core of LHC Networking

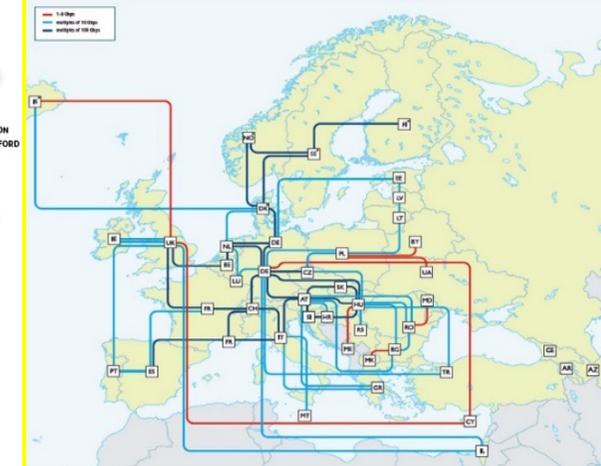
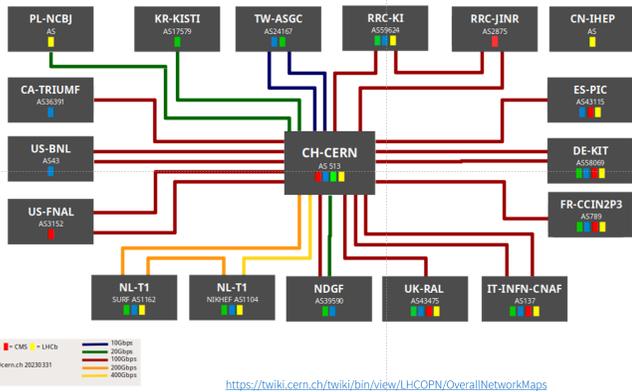
LHCOPN, LHCONE, GEANT, ESnet, Internet2, CENIC...

LHCOPN: Simple & Reliable Tier0+1 Ops

Internet2n

GEANT

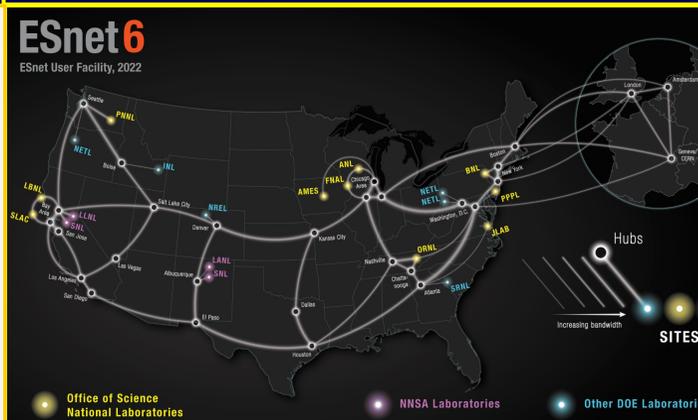
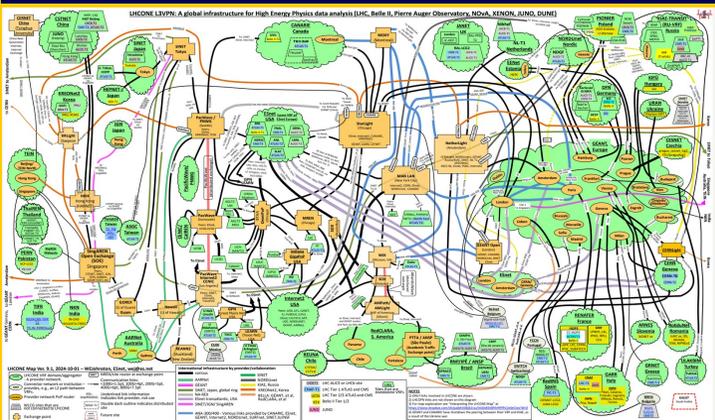
LHCOPN



LHCONE VRF: 170 Tier2s ++

ESnet6

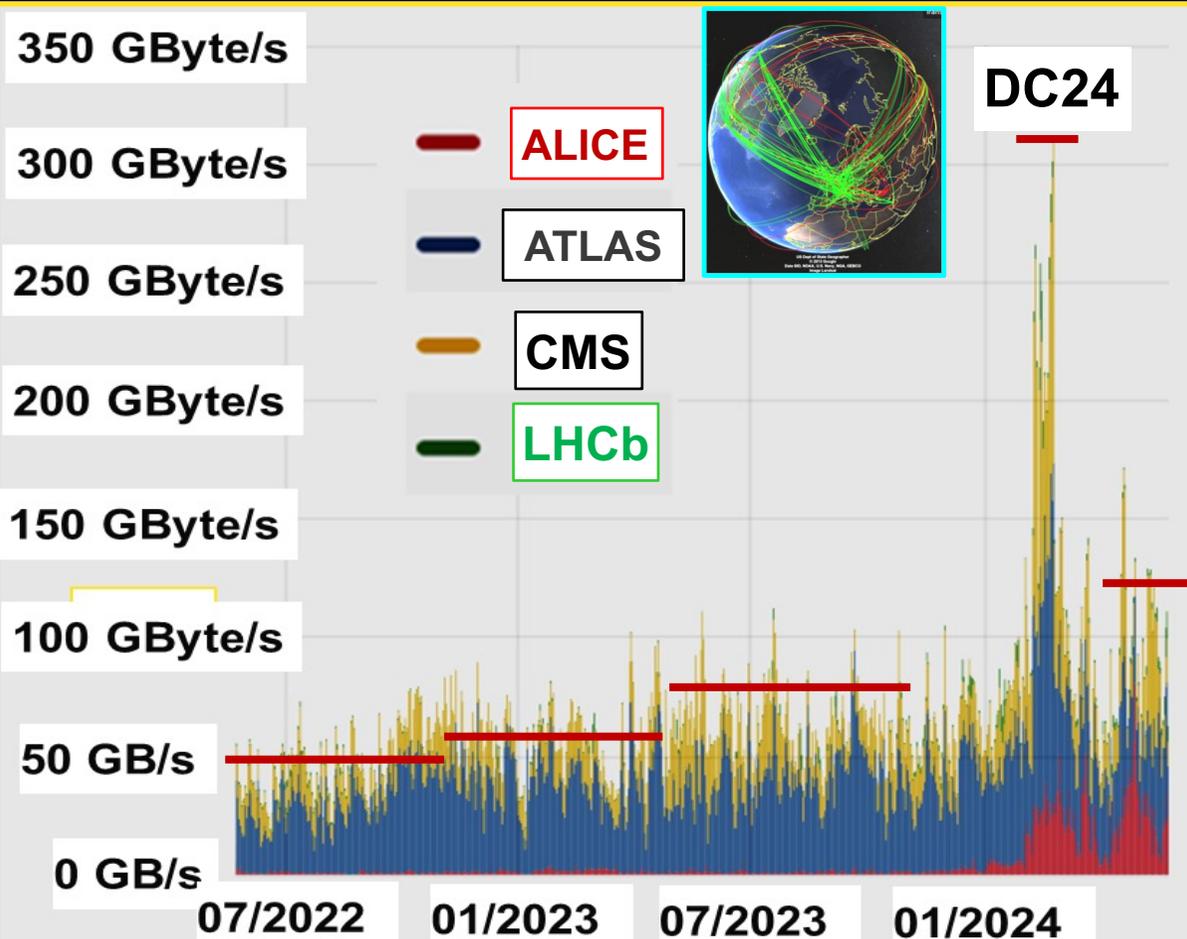
CENIC and NRP



+ NRENs in Europe, Asia, Latin America (RNP, Rednesp), Au/NZ; US State Networks

LHC Data Flows Increase in Scale and Complexity: Another Burst Upward in 2023-4

WLCG Transfers Dashboard: Throughput June 2022 – May 2024



*70-150 GBytes/s Weekly Avg
To 170+ GBytes/s Daily Avg*

Complex Workflow

- To ~2 M jobs (threads) simultaneously
- Multi-TByte to Petabyte Transfers
- To ~75 M File Transfers/Day
- Millions of remote connections

Another Sea Change in 2023-4

- 2X in Transfer Rates and Files Transferred
- **DC24 (25% HL LHC): 300+ GB/s**

~1.8 to 2X Growth in 12 Months: 100 to 1000X Per Decade Equivalent (?)

<https://monit-grafana.cern.ch/d/AfdonlyGk/wlcg-transfers?orgId=20&from=now-2y&to=now>

Global Network Advancement Group (GNA-G) Leadership Team: Since September 2019

leadershipteam@lists.gna-g.net



Buseung Cho
KISTI (Korea)



Ivana Golub
PSNC, GEANT
(Europe)



Harvey Newman
Caltech (US)



David Wilde,
Chair
Aarnet (Australia)



Alex Moura
KAUST
(Saudi Arabia)

- An open volunteer group devoted to developing the blueprint to make using the Global R&E networks both simpler and more effective
- Its primary mission is to support global research and education using the technology, infrastructures and investments of its participants.
- The GNA-G is a data intensive research & science engager that facilitates and accelerates global-scale projects by (1) enabling high-performance data transfer, and (2) acting as a partner in the development of next generation intelligent network systems that support the workflow of data intensive programs

See <https://www.dropbox.com/s/qsh2vn00f6n247a/GNA-G%20Meeting%20slides%20-%20TechEX19%20v0.8.pptx?dl=0>

GNA-G - strategic pillars

GREN innovations

- Global test facility
 - *eg SC demos; P4 testbed*
- Innovative network services
- Data movement

GREN operations

- Global network architecture & services
- Coordinated operations
- Automation, orchestration
- Monitoring, reporting, GREN map
- Secure the GREN; MANRS

A forum for discussion.

**Support for global big science & data intensive projects.
Engagement with international R&E network consortia.**

Structure



GNA-G Participant CEOs/Directors

Global NREN CEO Forum

GNA-G Executive Liaison

NomCom

GNA-G Leadership Team

Research and Development

Operations

Securing the GREN

Data-intensive Science

Smart Sensor Cables

GREN Risk Review

Advancing GREN Operations

GREN Mapping

AutoGOLE/SENSE

Telemetry

Routing Anomalies

Network Automation

GXP Architectures & Services

Link consortia

AER

APR

ANA

AmLight

APOnet

...

GREN: Collaboration on the intercontinental transmission layer

GNA Architecture 2.0

- Total 600 Gbps capacity between Brazil and the USA:
- Rednesp (formerly ANSP) has 3 links to the USA, connecting Sao Paulo to Florida.
 - Atlantic 100G link: direct São Paulo- Miami
 - Pacific 100G link: São Paulo to Santiago (Chile), from there to Panama, San Juan and Miami
 - 200G link: São Paulo to Florida through Fortaleza, on the Monet cable (Angola cables)
- Upgrade of the 400G AmLight Link (Rednesp + RNP) on Monet to 800G expected by March 2025



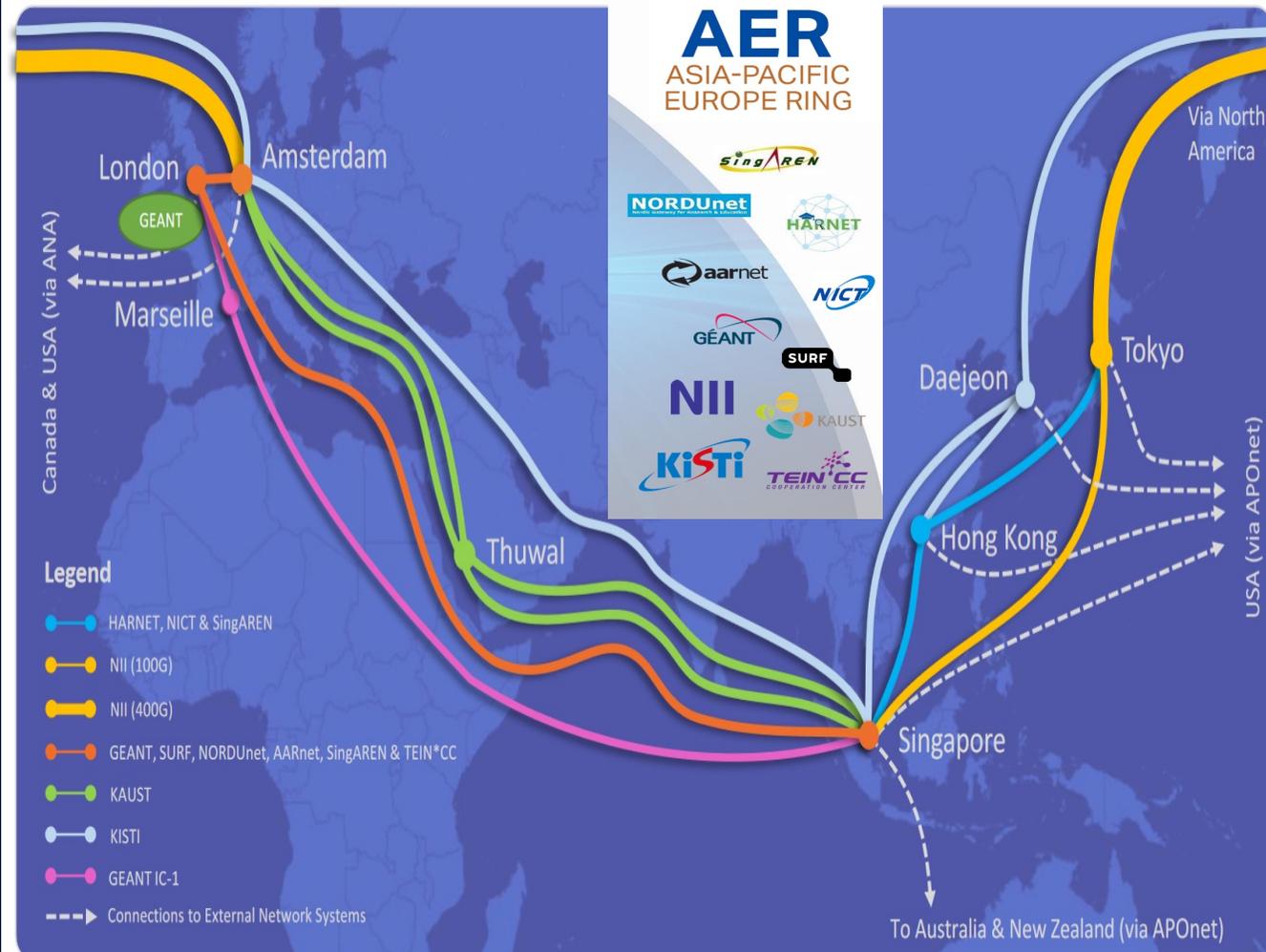
- Transatlantic Links: Bella link between São Paulo and Sines in Portugal, and a link connecting São Paulo, Angola and South Africa.
- "Backbone SP" connecting most universities in São Paulo State with 100G links was completed in May 2023
 - ★ UNICAMP (University of Campinas),
 - ★ UNIFESP (Federal University at São Paulo)
 - ★ UNESP (State University of São Paulo)
 - ★ USP (University of São Paulo)
 - ★ UFSCAR (Federal University of São Carlos)
 - ★ Mackenzie University
 - ★ ITA (Aeronautics Institute of Technology)
 - ★ UFABC (Federal University of ABC) to SP4
- ★ The SIRIUS synchrotron light source is now connected to the Backbone via UNICAMP in 2024

**USP dark fiber infrastructure nearing completion: 3 campuses already connected at 200G
Four main campuses are expected to be connected at 400G-800G by the end of 2025**

KAUST and the Asia Pacific Oceania Network (APONet): Closing the Global Ring East and West



- Since the AER MoU, KAUST is coordinating with REN partners on shared deployment of 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 SC24 NRE Demonstrations are also supported by KAUST closing the ring from Amsterdam to Singapore and back to the US



Charter: https://www.dropbox.com/s/4my5mjl8xd8a3y9/GNA-G_DataIntensiveSciencesWGCharter.docx?dl=0

- **Principal aims of the GNA-G DIS WG:**

- (1) To meet the needs and address the challenges faced by major data intensive science programs**

- **In a manner consistent and compatible with support for the needs of individuals and smaller groups in the at large A&R communities**

- (2) To provide a forum for discussion, a framework and shared tools for short and longer term developments meeting the program and group needs**

- **To develop a persistent global testbed as a platform, to foster ongoing developments among the science and network communities**

- **While sharing and advancing the (new) concepts, tools & systems needed**
- **Members of the WG partner in joint deployments and/or developments of generally useful tools and systems that help operate and manage R&E networks with limited resources across national and regional boundaries**
- **A special focus of the group is to address the growing demand for**
 - **Network-integrated workflows**
 - **Comprehensive cross-institution data management**
 - **Automation, and**
 - **Federated infrastructures encompassing networking, compute, and storage**
- **Working Closely with the AutoGOLE/SENSE WG**

- **Top Line Message:** To realize the physics discovery potential and meet the challenges of data intensive sciences, we need a new dynamic system which:
 - ★ **Coordinates worldwide networks** as a first class resource along with computing and storage, across and among world regions
 - ★ **Follows a systems design approach:** A global fabric that flexibly allocates, balances and makes best use of the available network resources
 - ★ **Negotiating with site systems** that aim to accelerate workflow
 - ★ **Builds on ongoing R&D projects:** from regional caches/data lakes to intelligent control and data planes to ML-based optimization
 - ★ **Leverages the worldwide move towards a fully programmable ecosystem of networks and end-systems** (P4, SONIC; PoIKA, SRv6), plus operations platforms (OSG, NRP; global SENSE Testbed, Global P4 Lab, FABRIC)
 - ★ **Simultaneously supports the LHC experiments, VRO other data intensive programs** and the larger worldwide academic and research community
 - ★ **The LHC experiments together with the GNA-G and its Working Groups, the WLCG and the worldwide R&E network community** are key players
 - ★ **Together with the major programs:** LHC, LBNF/DUNE, VRO, SKA
 - ★ **SC24 is a Major Milestone, and a Leap Forward Towards this Goal**



SC15-24: SDN Next Generation

Terabit/sec Ecosystem for Exascale Science

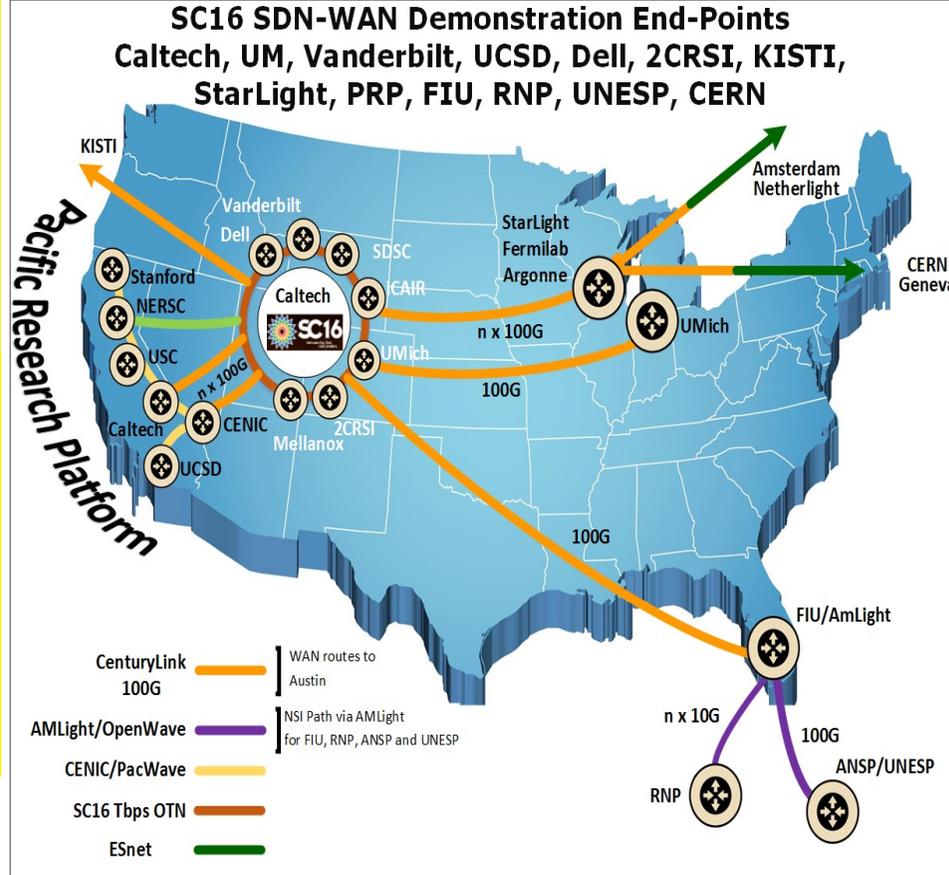
supercomputing.caltech.edu



SDN-driven flow steering, load balancing, site orchestration Over Terabit/sec Global Networks

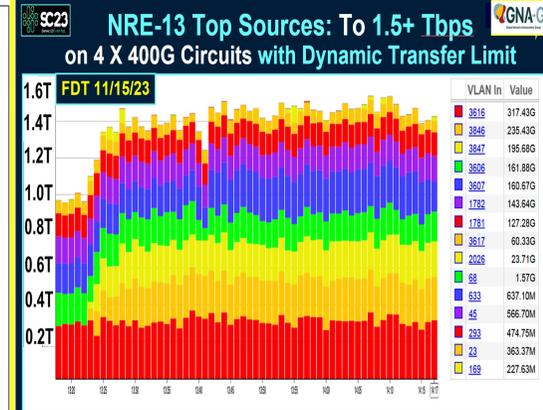
SC16+: Consistent Operations with Agile Feedback Major Science Flow Classes Up to High Water Marks

Preview PetaByte Transfers to/from Sites With 100G - 1000G DTNs



LHC at SC15: Asynchronous Stageout (ASO) with Caltech's SDN Controller

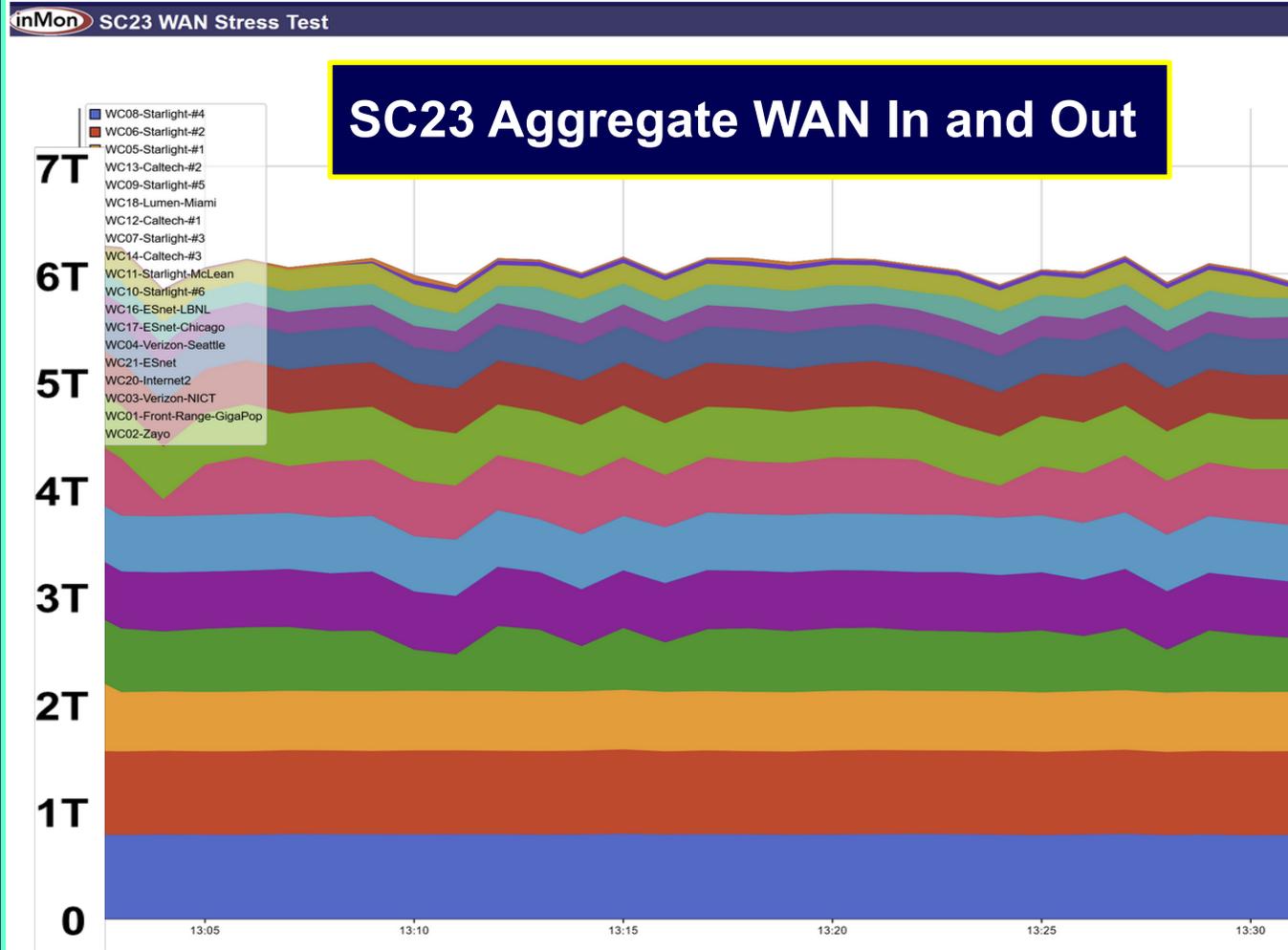
Tbps Rings for SC18-24: Caltech, Ciena, SCInet, StarLight + Many HEP, Network, Vendor Partners



29 100G NICs; Two 4 X 100G and Two 3 X 100G DTNs; 1.5 Tbps Capability in one Rack; 9 32 X100G Switches

SC23 Stress Test 11/16/23

Caltech Booth providing 2.3 Tbps of 6.2 Tbps



Going Forward

- Latest kernels: full use of all PCIe slots
- 400GE with CX7 NICs and DR4 Transceivers
- Multi-User: Scheduled stable N X 100G flows with FDT & SENSE
- SENSE 400G paths: ESnet production, NA-REX via StarLight; Links to CERN
- NVMe SSD Front End Operations + HSM
- PCIe 6.0 and CXL DTN tests by ~SC24

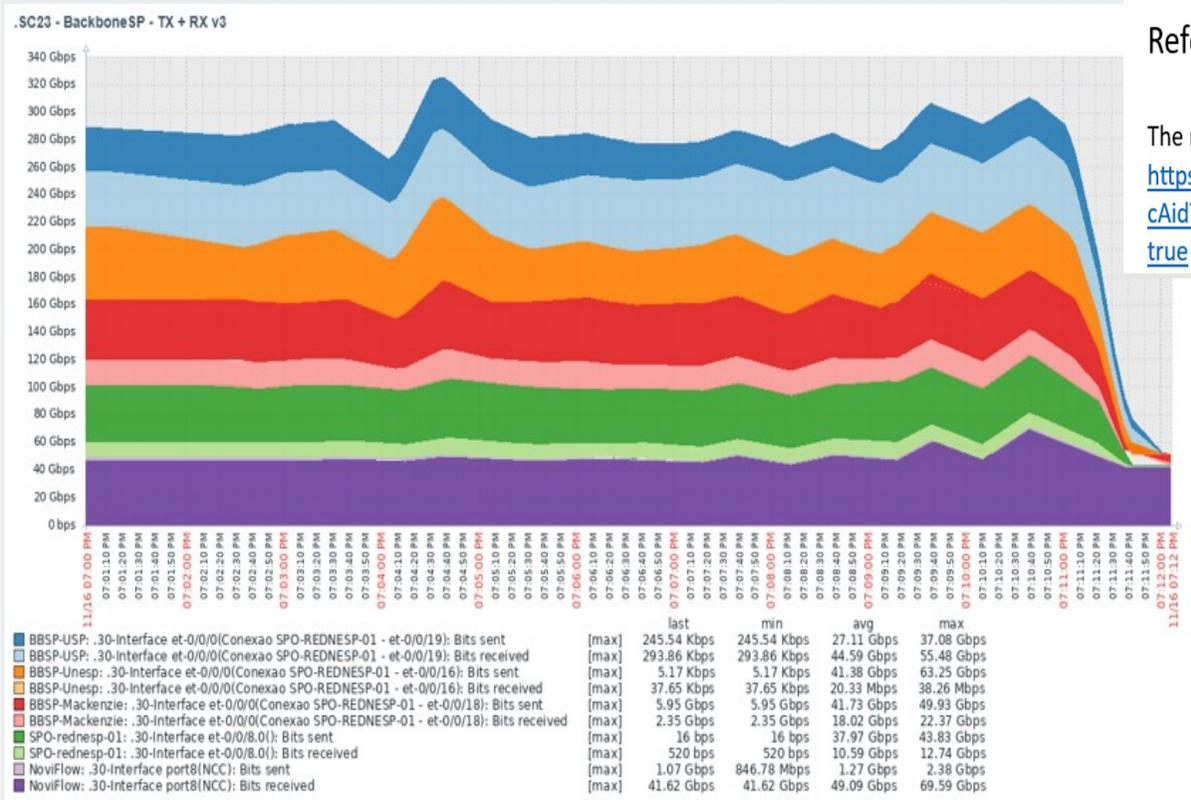
With 2 Gen5 + Gen3 Servers at SC23 and 3 Gen5 Servers at Caltech

FDT 11/16/23

This Just In (11/23) Rednesp Backbone: Record US ↔ Brazil Results

Two networking tools were used to generate traffic: **iperf3** and **fdt**.
 During SC23 data tsunami, on November the 16th, a peak of 330 gbps (considering data from Brazil to the USA and vice-versa) was achieved and can be seen in the next figure.

These results are very good, considering that the 100 gbps links also carry production traffic. However, it is certainly possible to achieve higher bandwidths with more tuning and with a more controlled bandwidth allocation in the links. Rednesp is now trying to optimize its infrastructure to achieve a more efficient use of the intercontinental links connecting Sao Paulo, Brazil, to the USA, to Europe and to other countries in South America.



References

The rednesp presentation slides can be seen at [https://docs.google.com/presentation/d/1qUX1mvP3Ohb5zMuP18-cAidTvCcFEWF6/edit?usp=drive link&oid=114820778254083128813&rtpof=true&sd=true](https://docs.google.com/presentation/d/1qUX1mvP3Ohb5zMuP18-cAidTvCcFEWF6/edit?usp=drive_link&oid=114820778254083128813&rtpof=true&sd=true)



Global Network Advancement Group: Next Generation Network-Integrated System for Data Intensive Sciences

Network Research Exhibition NRE-113

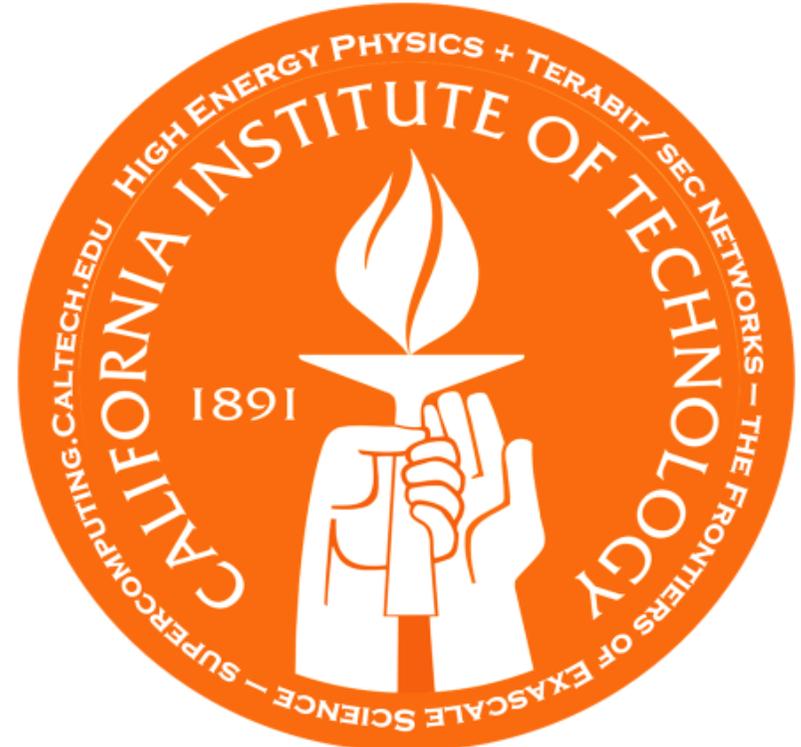
- **A Vast Partnership** of Science and Computer Science Teams, R&E Networks and R&D Projects; **Convened by** the GNA-G DIS WG; with GRP, AmRP, NRP
- **Mission: Demonstrate the road ahead**
 - **Meet the challenges** faced by leading-edge data intensive programs in HEP, astrophysics, genomics and other fields of data intensive science;
 - ★ **Compatible with other use**
 - **Clearing the path** to the next round of discoveries
- **Demonstrating a wide range of latest advances in:**
 - Software defined and Terabit/sec networks
 - Intelligent global operations and monitoring systems
 - Workflow optimization methodologies with real time analytics
 - State of the art long distance data transfer methods and tools, local and metro optical networks and server designs
 - Emerging technologies and concepts in programmable networks and global-scale distributed systems
- **Hallmarks:** Progressive multidomain integration; **compatibility internal + external;** **A comprehensive systems-level approach**



Worldwide Partnership at SC24 and Beyond



Global Petascale to Exascale Workflows for Data Intensive Sciences



Accelerated by Next Generation Programmable SDN Architectures and Machine Learning Applications

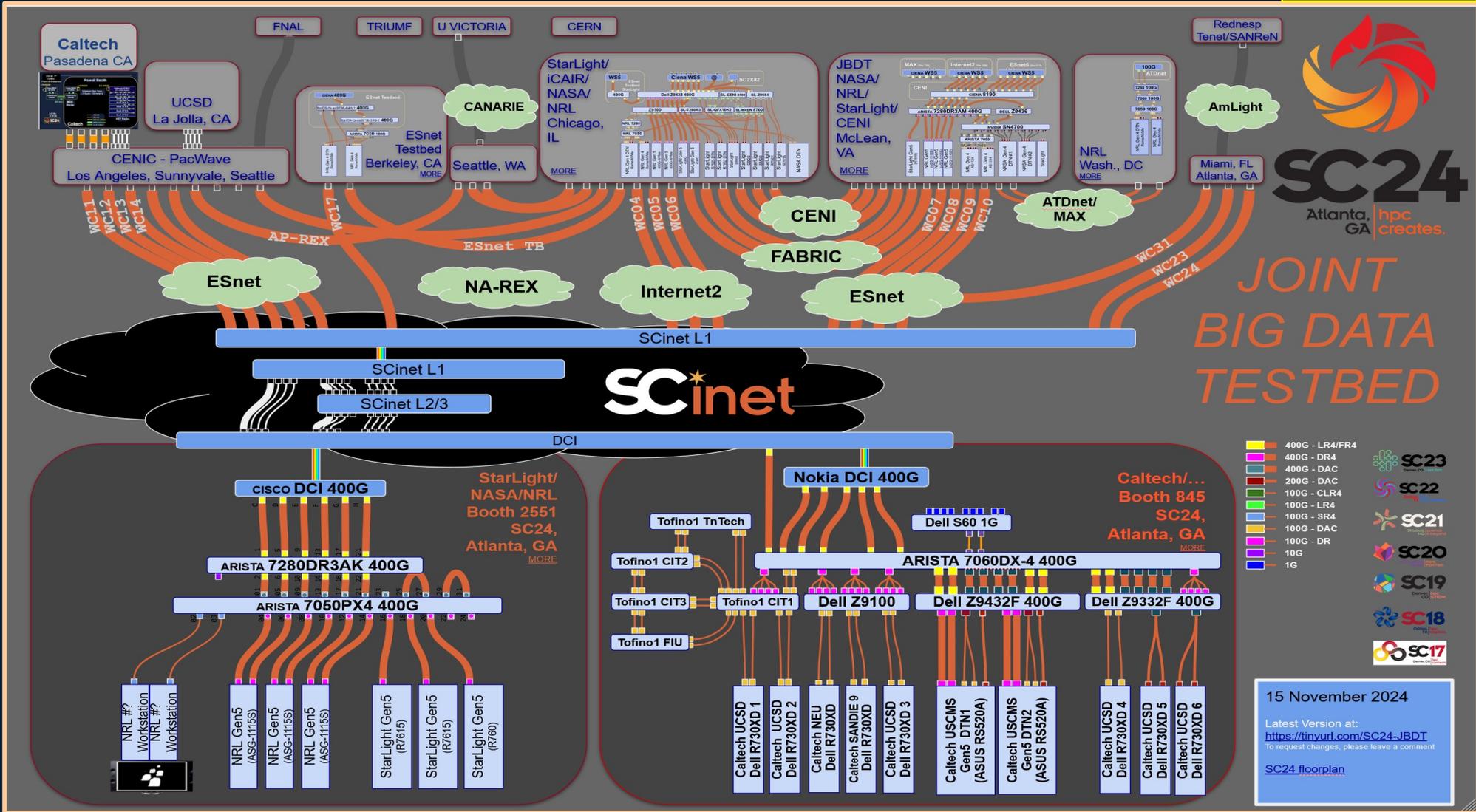


Global Petascale to Exascale Workflows for Data Intensive Sciences



- ★ **Advances Embedded and Interoperate within a ‘composable’ architecture of subsystems, components and interfaces, organized into several areas; coupled to rising Automation**
- ★ **Visibility:** Monitoring and information tracking and management including IETF ALTO/OpenALTO, BGP-LS, sFlow/NetFlow, Perfsonar, Traceroute, Qualcomm Gradient Graph congestion information, Kubernetes statistics, Prometheus, P4/Inband telemetry, *InMon*
- ★ **Intelligence:** Stateful decisions using composable metrics (policy, priority, network- and site-state, SLA constraints, responses to ‘events’ at sites and in the networks, ...), using NetPredict, Hecate, GradientGraph, Yale Bilevel optimization, Coral, Elastiflow/Elastic Stack
- ★ **Controllability:** SENSE/AutoGOLE/SUPA, P4, segment routing with SRv6, SR/MPLS and/or PoIKA, BGP/PCEP
- ★ **Network OSeS and Tools:** GEANT RARE/freeRtr, SONIC; Calico VPP, Bstruct-Mininet environment, ...
- ★ **Orchestration:** SENSE, Kubernetes (+k8s namespace), dedicated code and APIs for interoperation and progressive integration

Caltech and StarLight/NRL Booths at SC24



SC24: Global footprint. Terabit/sec Triangle Starlight – McLean – Atlanta; 6 X 400G to the Caltech Booth: 2X 400G LA-ATL; 4 X 400G to the Caltech Campus, 2 X 400G + 100G to Latin America and South Africa; 400G to Fermilab; 400G to CERN with CENI, Ciena, Internet2, ESnet, StarLight, US CMS and Network Partners

15 November 2024
 Latest Version at:
<https://tinyurl.com/SC24-JBDT>
 To request changes, please leave a comment
[SC24 floorplan](#)

NRE-04	Joe Mambretti (Northwestern) et al.	1.2 Tbps Services WAN Services: Architecture, Technology and Control Systems
NRE-05	Joe Mambretti (Northwestern) et al.	Global Research Platform
NRE-16	Chris Wilkinson (Internet2) et al.	NA-REX Prototype Demonstration
NRE-13	Harvey Newman (Caltech) et al.	The Global Network Advancement Group: A Next Generation System for Data Intensive Sciences
NRE-13a	Alex Moura (KAUST) et al.	Exploring FDT, QUIC, BBRv2 and HTTP/3 in High Latency WAN Paths
NRE-13b	Edmund Yeh (Northeastern) et al.	N-DISE: NDN for Data Intensive Science Experiments
NRE-15	Kasandra Pillay (SANReN/CSIR) et al.	High Speed Data Transfers from South Africa to USA !
NRE-19	Carlyn Ann-Lee (JPL) et al.	A Federated Learning Guide to Cosmic Ray Events
NRE-22	Tom Lehman (ESnet) et al.	AutoGOLE/SENSE: End-Site Resource Integration with Network Services
NRE-23	Tom Lehman (ESnet) et al.	SENSE and Rucio/FTS/XRootD/dCache Interoperation
NRE-25	Mariam Kiran (ORNL), M. Martinello (UFES) et al.	HECATE merges with PoKA AI-enabled traffic engineering For data-intensive science
NRE-29	John Graham (UCSD/SDSC), Marcos Schwarz (RNP) et al.	Multi Domain experiments using ESnet SENSE on the National Research Platform / PacWave / FABRIC
NRE-30	Marcos Schwarz (RNP), Carlos Ruggiero (USP/Rednesp) et al.	Global P4 Lab: Programmable Networking with P4, GEANT RARE/freeRtr and SONIC; Digital Twin
NRE-31	Carlos Ruggiero (USP/Rednesp) et al.	High Performance Networking with the Sao Paulo Backbone SP Linking 8 Universities and the Bella Link
NRE-32	Everson Borges (IFES), Magnos Martinello (UFES) et al.	PoKA Routing Approach to Support Traffic Engineering for Data-intensive Sciences
NRE-33	Y. Richard Yang (Yale) et al.	XTS: Scaling and Optimizing Efficiency and Control of Data Transport for Data-intensive Networks

Acknowledgements

This ongoing work is partially supported by the US National Science Foundation (NSF) Grants OAC-2030508, OAC1841530, OAC-1836650, MPS-1148698, and PHY-1624356, along with research grants from many international funding agencies and direct support from the many regional, national, and continental network and industry partners mentioned. The development of SENSE is supported by the US Department of Energy (DOE) Grants DE-SC0015527, DESC0015528, DE-SC0016585, and FP-00002494.

Finally, this work would not be possible without the significant contributions and the collaboration of the many HEP, network and computer and research teams partnering in the Global Network Advancement Group, in particular the GNA-G Data Intensive Sciences and AutoGOLE/SENSE Working Groups and the Global P4 Lab led by GEANT and the RNP Brazilian National Network, together with many industry partners, most notably Ciena, Dell and Arista

