SC23 Network Research Exhibition:
Caltech Booth 1255 Demonstrations
Hosting a Wide Range of NREs

The Global Network Advancement Group:
A Next Generation System for Data Intensive Sciences

Submitted on behalf of the teams by:
Harvey Newman, Caltech, newman@hep.caltech.edu



Abstract

The Global Network Advancement Group and its Data Intensive Sciences and SENSE/AutoGOLE Working Groups, a worldwide collaboration bringing together major science programs, research and education networks, and advanced network R&D projects spanning the U.S, Europe, Asia, Latin America and Oceania, are developing a next generation network-integrated system designed to meet the challenges of exabyte and terabit/sec workflows supporting thousands of scientists, and to clear the path to the next round of discoveries in high energy and astrophysics, bioinformatics and many other fields of data intensive science.

This network research exhibition (NRE) will present and demonstrate, in partnership with several NREs covering specific areas, many aspects of the next generation system which the GNA-G is developing both for data intensive sciences and to help set the course for future R&E network operations.

In order to meet these high-level goals, the teams in this NRE are working towards a new dynamic and adaptive programmable software-driven system which coordinates worldwide networks as a first class resource along with computing and storage, across multiple domains. We are following a systems design approach to create a global dynamic fabric that flexibly allocates, balances and conserves the available network resources, while negotiating and working with the site-resident systems that aim to accelerate workflow.

Reinforcement and other model-based machine learning techniques are also planned to be used to optimize system operations, according to objective functions that take priority, policy, responses to network- and site-state changes, workflow objectives and other constraints into account.

An overarching concept is “Consistent Network Operations,” where stable load balanced high throughput workflows crossing optimally chosen network paths, up to preset *high water marks* to accommodate other traffic, are provided by autonomous site-resident services dynamically interacting with network-resident services, in response to demands from the science programs’ principal data distribution and management systems.

The demonstrations in this and partner NREs are based on a wide range of ongoing R&D projects: from regional caches/data lakes to intelligent control and data planes to machine-learning based based optimization, including: SENSE/AutoGOLE and its integration with the FTS and Rucio data management system, NOTED, GEANT/RARE freeRtr and the Global P4 Lab, ALTO/TCN, PolKA, ESNet High Touch, Qualcomm GradientGraph, AmLight, Fabric, Bridges; NetPredict and Hecate, among others.

Many of the aspects composing this NRE and its partners benefit from the Open Science Grid, National Research Platform (NRP) led by UCSD, as well as the Global Research Platform (GRP) led by iCAIR, and the Americas Research Platform (AmRP) led by FIU.

We are also leveraging the worldwide move towards a fully programmable ecosystem of networks and end-systems, bring together the use of P4, NPL and other programmable network services; segment routing including SRv6 and SR MPLS along with PolKA; and other emerging core concepts and components such as BGP Classful Transport (BGP-CT) that enables multidomain overlays with defined service levels and other attributes.

New themes this year include: (1) the Global P4 Lab, its architecture and capabilities, (2) a broader view of the programmable networking ecosystem (P4, SONIC, NPL, BGP-CT, gNMI and other mainstream programmable approaches, and (3) the synergy between GEANT/RARE freeRtr and the major vendors' OSes and tools. A key recent development associated with the Global P4 Lab effort is the development of a Digital Twin (See https://github.com/srl-labs/containerlab/blob/main/README.md) that incorporates high fidelity instances of containerized as well as VM-based instances of network OSes, and which thereby enables a smooth transition from a complex simulated network to an actual production network deployment. An additional powerful capability is to scrape an actual multidomain network topology using Netbox to form the Digital Twin automatically. The Digital Twin thus has a strategic role in bridging the gap between advanced developments of new features and functions, and progressive field deployments in production R&E networks.

These groundbreaking efforts build on the common global infrastructure, aimed at both production network operations and the progressive deployment of new advanced services, that the GNA-G is developing through its unique partnership of national, regional, continental and transoceanic networks, global exchange points and other R&E network providers, illustrated in Figure 1. The GNA-G organization and architecture shown in Figure 2 illustrates both the synergetic relationship between R&D and Production Operations, the joint mission to consistently serve and meet the needs of both the major science programs and the at-large academic and research communities, and a sustained commitment to deploy and leverage the latest

developments to meet the evolving goals of these communities, while advancing the missions of the network science and engineering communities alike.



Figure 1 The GNA-G mission and its transoceanic partner consortia spanning the Atlantic and Pacific regions.

Since 2020, the GNA-G and its partners have deployed the Global SENSE/AutoGOLE and Global P4 Lab programmable testbeds, and have established a persistent development trajectory harnessing the teams' own advancements in software defined and Terabit/sec network technologies and methods, intelligent global operations

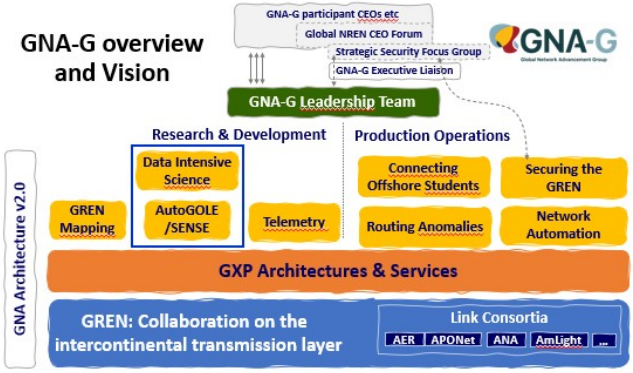


Figure 2 The GNA-G organizational architecture and working groups, emphasizing the Data Intensive Sciences and SENSE/AutoGOLE groups.

And monitoring systems, workflow optimization methodologies with real-time analytics, and state of the art long distance data transfer methods and tools and server designs. together with emerging technologies and concepts in programmable networks which are designed to meet the challenges faced by leading edge data intensive experimental programs in high energy physics, astrophysics, genomics and other fields of data intensive science to clear the path to the next round of discoveries.

These advances, being developed within the Global Network Advancement Group (GNA-G) framework and its SENSE/AutoGOLE and Data Intensive Sciences (DIS) working groups, the Global Research Platform, and its many science, computer science and R&E network teams aim to address the key challenges including: (1) global data distribution, processing, access and analysis, (2) the coordinated use of massive but still limited computing, storage and network resources, (3) coordinated operation and collaboration within global scientific enterprises each encompassing hundreds to thousands of scientists, and (4) enabling the science programs to make efficient use of the available network and site infrastructures, while simultaneously accommodating and in concert with the network operations required to supporting the worldwide academic and research community across national, regional and transoceanic boundaries.

The major programs being highlighted include the Large Hadron Collider (LHC), BioGenome and the human genome as well as astrophysics projects such as the SKA and the Vera Rubin Observatory and others. We will also highlight the latest developments in bottleneck structures and analysis for real-time congestion resolution others well as state of the art 5G "holodeck" applications running over a segment-routed transcontinental links between the UCSD and NYU campuses. The 5G applications also will help pave the way for other real-time highly differentiated services and applications including extended reality (XR) and vehicle to anything (V2X) as well as the future transition to 6G wireless applications.

Several of our teams' demonstrations include state of the art programmable open source network operating systems, layer 2 multi-domain virtual circuit overlays coupled to extensive structured use of IPv6 address spaces, segment routing approaches to efficient flow distribution, steering and load balancing, and stateful traffic engineering and workflow acceleration taking network segment and site state, policy and priority and real-time SLAs into account.

The teams' several branches of advanced development, driven by a diverse set of challenging use cases within the GNA-G framework and associated R&D projects, and enabled by the National Research Platform and the Global Research Platform, are embedded and inter-operate within an emerging "composable architecture" of subsystems, components and interfaces organized into several areas:

- **Visibility:** monitoring and information tracking and management including IETF ALTO/OpenALTO, BGP-LS, sFlow/NetFlow, Perfsonar, Traceroute, Qualcomm GradientGraph congestion information, Kubernetes statistics, Prometheus, P4/Inband telemetry
- **Intelligence:** NetPredict, Hecate, RL-Gradient Graph, Yale Bilevel optimization, stateful decisions using composable metrics (policy, priority, network- and site-state, SLA constraints,

responses to ‘events’ at sites and in the networks, ...), Coral, Elastiflow/Elastic Stack
- **Controllability:** SENSE/AutoGOLE, SUPA, P4/PINS, segment routing (SRv6, SR MPLS, PolKA), BGP/CT
- **Network OSes and Tools:** SONIC, GEANT RARE/freeRtr, Calico VPP, Bstruct-Mininet environment, ...
- **Orchestration**: SENSE, Kubernetes, dedicated code and APIs for interoperation and progressive integration

The cornerstone system concepts, components and developments demonstrated will include:

- Integrated operations and orchestrated management of resources: interworking with and advancing the site (Site-RM) and network resource managers (Network-RM) developed in the SENSE program.
- Fine-grained end-to-end monitoring and data collection, with a focus on the edges and end sites, enabling data analytics-assisted intelligent and automatic decisions driven by applications supported by optimized path selection and load balancing mechanisms driven by machine learning.
- An ontological model-driven framework with integration of an analytics engine, API and workflow orchestrator extending work in the SENSE project, enhanced by efficient multi-domain resource state abstractions and discovery mechanisms.
- The integration of Qualcomm Technology’s GradientGraph (G2) with 5G / OpenALTO / SRv6 will be used to demonstrate how applications including XR, auto/V2X, and IoT can benefit from the intelligent routing, rate limiting, and service placement decisions computed by G2. We will also demonstrate how the integration of G2 with science networks through the integration of the LHC Rucio / FTS data management system, AutoGOLE / SENSE virtual circuit and orchestration services / and OpenALTO comprehensive monitoring may be used towards optimizing large-scale data transfers to share data from the Large Hadron Collider at CERN and scientists across the globe.
- Adapting NDN for data intensive sciences including advanced cache design and algorithms and parallel code development and methods for fast and efficient access over a global testbed, leveraging the experience in the NDN for Data Intensive Science Experiments (N-DISE)
- A paragon network at several our partners’ sites Composed of P4 programmable devices, including Tofino and Tofino2-based switches, Smart NICs and Xilinx FPGA-based network interfaces providing packet-by- packet inspection, agile state tracking, real-time decisions and rapid reaction as needed.
- High throughput platform demonstrations in support of workflows for the science programs mentioned. This will include reference designs of NVMe server systems to match a 400G network core, as well as servers with multi-GPUs and programmable smart NICs with FPGAs.
- Integration of edge-focused extreme telemetry data (from P4 switches and end hosts) and end facility/application caching stats and other metrics data to facilitate automated decision-making process.
- Development of dynamic regional caches or “data lakes” that treat nearby as a unified data resource, building on the successful petabyte cache currently in operation between Caltech and UCSD and in ESNet based on the XRootD federated access protocol; extension of the cache concept to more remote sites such as Fermilab, Nebraska and Vanderbilt.
- We will introduce how PolKA (Polynomial Key-based Architecture) Source Routing (SR) brings Better, Faster, and Stronger functionalities to support the TE challenges for DIS networks. We plan to divide the demonstration of PolKA capabilities into two scenarios: i) in a data-intensive transfer over 100G and 400G network by using PolKA underlay tunnels; ii) in a comparison between segment routing and PolKA over a high-speed intercontinental testbed composed of P4-enabled programmable switches that interconnect DIS research facilities (Europe, South, and North America) employing its controllability of flows to support TE. The flows can be classified, balanced, and steered at the edge using a Policy-Based Routing (PBR) so that TE decisions can be guided by the Quantitative Theory of Bottleneck Structures (QTBS) for optimization.
- We will use an overlay network with PolKA tunnels forming virtual circuits for this demo. The overlay will be built by integrating persistent resources from GNA-G AutoGOLE/SENSE and GEANT RARE testbeds. Underlay congestion can be detected by tunnel monitoring and signalized to the overlay so that the traffic is steered from congested tunnels to other paths. Comparisons between segment routing and PolKA regarding controllability and performance metrics are also planned. From the application's point of view, PolKA full deployment enables us to meet the extreme traffic engineering demanded by data-intensive sciences by offering a new range of network functionalities such as: multipath routing, in-network telemetry and proof-of-transit with path attributes to support higher level stateful traffic engineering decisions.

- Network traffic prediction and engineering optimizations using the latest graph neural network and other emerging deep learning methods, developed by ESnet's Hecate /DeepRoute project.

Elements and Goals of the Demonstrations

- **LHC:** End to end workflows for large scale data distribution and analysis in support of the CMS experiment's LHC workflow among Caltech, UCSD, LBL, Fermilab, Nebraska, Vanderbilt, and GridUNESP (Sao Paulo) including automated flow steering, negotiation and DTN autoconfiguration; bursting of some of these workflows to the NERSC and other HPC facilities and the cloud; use of both edge and in-network caches to increase data access and processing efficiency.
- **SENSE/AutoGOLE:** The Global Network Advancement Group (GNA-G)'s worldwide collaboration of Open exchange Points and Research & Education networks delivers end-to-end network services with fully API driven automated methods. The GNA-G AutoGOLE/SENSE Working Group demonstration will present key technologies, methods and a system of dynamic Layer 2 and Layer 3 network services to meet the challenges and address the requirements of the largest data intensive science programs and workflows. The services are designed to support multiple petabyte transactions across a global footprint, represented by a persistent testbed spanning the US, Europe, Asia Pacific and Latin American regions. The SC23 demonstrations will focus on new features in the area of programmatic driven domain science workflow integration and realtime monitoring/troubleshooting. A key theme this year is the integration of SENSE virtual circuits with the LHC data transfer and management systems FTS and Rucio.
- **SENSE/Rucio:** SENSE is a multi resource, multi domain orchestration system which provides an integrated set of network and end-system services. The Rucio/File Transfer Service (FTS)/XRootD data management and movement system is the key infrastructure used by LHC experiments and more than 30 other programs in the Open Science Grid. This demonstration will show the interoperation of SENSE with the Rucio/FTS/XRootD data management/movement system to enable a new set of services for domain science workflows. The new features include an ability for science workflows to define priority levels for data movement operations through a Data Movement Manager (DMM) that translates Rucio-generated priorities into SENSE requests and provisioning operations. Additional features include full-lifecycle monitoring, evaluation, and adjustment of associated network services. The focus will be on a prototype deployment at UCSD, Caltech, Fermilab, and CERN, along with other US CMS Tier2 sites.

  Emerging features in the DMM include the ability to include policy by program and by transaction class, balance existing SLAs and transfers in progress against incoming requests, and the development of success metrics that take these factors into account.
- **Global Ring and KAUST:**

  These demonstrations will also showcase the power of collaboration in the global research and education network (REN) community. The demo will have KAUST as a collaborator in the Asia-Pacific Global Ring (AER [*]), closing the global ring by interconnecting Amsterdam to Singapore, and then onto Los Angeles with support from partners including SingaREN, JGN-X and APOnet.

[*] See Fast and stable Network Ring between Asia-Pacific and Europe | AARNet

The King Abdullah University of Science and Technology (KAUST) is making strides to collaborate with the global Research and Education community by harnessing the power of advanced networking infrastructures. Recognizing the potential for synergistic development and efficient data transfer in large-scale scientific projects, KAUST has formally joined hands with important international initiatives like the Asia-Pacific Europe Ring (AER) and became an active participant in the Global Network Architecture (GNA-G) initiative.

These associations are not merely symbolic; they hold profound implications for the improvement of research services operations on a global scale. KAUST's strategic affiliation with the AER provides a robust platform to facilitate data-intensive research that demands reliable, high-speed, and efficient connectivity. On the other hand, by participating in the GNA, KAUST is actively contributing to the advancement of the Global Research and Education ecosystem.

Moreover, KAUST has moved beyond passive participation and is actively supporting improvement of the kingdom's international connectivity in tandem with Maeen, the Saudi Arabia National Research and Education Network (NREN), and other RENs around the globe. This collaboration is aimed at enhancing the performance of data transfers, with the current scale of operations involving the transfer of

multiple gigabytes of data every day. This high-volume data handling capability not only supports the vast data requirements of large-scale scientific research projects but also enables seamless international collaborations.

Stepping further, KAUST's affiliation with the Global Network Architecture (GNA) underscores its strategic alignment with Saudi Vision 2030. This involvement positions KAUST at the forefront of global digital integration, fostering digital transformation, innovation, and research. By developing new technologies and services, KAUST enhances its global research capacity, echoing the Vision 2030's call for technological advancement and collaboration. Ultimately, KAUST's commitment reflects its dedication to shaping a progressive, cooperative research ecosystem.

- **Toward 1.2 Tbps Services WAN Services: Architecture, Technology and Control Systems**

  Large scale data production within and among science research collaborations and sites continues to increase, a long term trend that continues to accelerate, especially because of the deployment of new science instrumentation, including planned high luminosity research infrastructure. Consequently, the science networking community has begun to prepare for service paths beyond 100Gbps, with a thematic focus on 400 Gbps LANs and WANs (represented in other NRE demonstrations by this consortium). However, this consortium is also investigating capabilities for WAN services beyond 400 Gbps, including those approaching 800 Gbps, 1 Tbps, and multi-Tbps WAN and LAN services.

- **400 Gbps E2E WAN Services: Architecture, Technology and Control Systems**

  Data production among science research collaborations continues to accelerate, a long term trend that in part is propelled by large scale science instrumentation, including high luminosity research instruments. Consequently, the networking community is preparing for service paths beyond 100 Gbps, including 400 Gbps, 800 Gbps and 1 Tbps WAN and LAN services. In this progression, 400 Gbps E2E WAN services are a key building block. Consequently, the requirements and implications of 400 Gbps WAN services are being explored at scale, including 400 Gbps E2E over thousands of miles. These demonstrations will showcase 400 Gbps E2E WAN services from the StarLight International/National Communications Exchange Facility in Chicago to the SC23 venue, between StarLight and the multi-agency Joint Big Data Testbed (JBDT) Facility in McLean, Virginia, and between the JBDT Facility and the SC23 venue.

- **Software Defined Exchange (SDX) Multi-Services for Petascale Science**

  With multiple national and international partners, iCAIR is designing, developing, implementing, and operating an international Software Defined Exchange (SDX) at the StarLight International/National Communications Exchange Facility, which integrates multiple services designed for large scale global data intensive science. The StarLight SDX is based on a flexible, scalable, programmable platform. This SDX, which is managed by a multi-organizational consortium, has been proven able to integrate many different multi-domain services and to insure services isolation. Services include those based on 100 Gbps Data Transfer Nodes (DTNs) for Wide Area Networks (WANs), including trans-oceanic WANs. Currently, a key focus is scaling to 400 Gbps WAN and LAN E2E technologies that provide high performance transport services for petascale science, controlled using Software Defined Networking (SDN) techniques. SDN enabled DTN services are being designed specifically to optimize capabilities for supporting large scale, high capacity, high performance, reliable, high quality, sustained individual data streams for science research.

- **Global Research Platform (GRP)**

  The Global Research Platform (GRP) is an international scientific collaboration that is creating innovative advanced ubiquitous services that integrate resources around the globe at speeds of gigabits and terabits per second, especially for data-intensive science research. GRP focuses on design, implementation, and operation strategies for next-generation distributed services and infrastructure to facilitate high-performance data gathering, analytics, transport, computing, and storage among multiple science sites at 100 Gbps or higher (e.g., 400 Gbps WAN streams0. GRP community partners are located in North America, Asia, Europe, and South America and work together to customize international fabrics and distributed cyberinfrastructure to support optimal data-intensive scientific workflows. Essentially, the GRP is a worldwide Science DMZ, a distributed environment for data-intensive research. The GRP leverages optical circuits and open exchange facilities provided by its collaborators.

- **400 Gbps Data Transfer Nodes**

  With its national and international partners, iCAIR and the StarLight consortium members, as well as Caltech, are designing and prototyping a 400 Gbps Data Transfer Node (DTN) capable of generating and receiving single 400 Gbps streams, based on PCI Gen 5 technology and next gen smart NICs. For previous SC NRE demonstrations, iCAIR and Caltech used several 200 Gbps DTNs and iCAIR bonded streams into a 400 Gbps flow. The 400 Gbps will showcase capabilities for streaming WAN/LAN E2E 400 Gbps science data flows.

- **AmLight Express and Protect (AmLight-ExP).** The AmLight network uses a hybrid network strategy that combines optical spectrum and leased capacity to build a reliable, leading-edge network infrastructure for research and education. AmLight supports high-performance network connectivity required by international science and engineering research and education collaborations involving the National Science Foundation (NSF) research community, with expansion to South America and West Africa. These capabilities will be used in support of the SENSE/AutoGOLE demonstrations and the LHC, astrophysics and genomics-related use cases. High-throughput low latency experiments and demonstrations of auto- recovery from network events will be shown, using optical spectrum on the Monet submarine cable, and its 100G ring network that interconnects the research and education communities in the U.S. and South America, as well as the SACS cable to West Africa.
- **Amlight-EXP/SANReN/TENET/CSIR Demonstrations**

  This collaboration offers a unique opportunity to work alongside CHPC and SANREN, encompassing African involvement in the experiments while closely monitoring the performance of the AmLight-ExP's SACS link through a diverse array of experiments arranged for SC23.

  The aim of these is demonstrate AmLight-ExP's 100G link over SACS submarine cable system, which established a vital connection between South Africa and the United States via Miami. This demo includes the CHPC and SANREN/TENET within the network of NRENs and science teams that participate in the GNA-G Data Intensive Sciences Working Group. This will showcase the turbocharging of the WACS/SACS/AmLight 100G link, utilizing 100G DTNs, using anonymized Genomics and/or Astrophysics data from a DTN at CHPC to DTNs in Miami and SC23 show floor.

  About the partners in these demonstrations:

  - The South African National Research Network (SANReN) group is responsible for the design, acquisition and roll-out of national and international capacity for the SA NREN, as well as the development and incubation of advanced services. It forms part of a comprehensive South African government approach to cyberinfrastructure, geared at ensuring the successful participation of South African researchers in the global knowledge production endeavor.
  - SANReN is managed and implemented by the CSIR's Next Generation Enterprises and Institutions (NGEI) cluster and is a key component of the **National Integrated Cyberinfrastructure System (NICIS)**, alongside the Centre for High Performance Computing (CHPC) and the Data Intensive Research Initiative of South Africa (DIRISA).
  - The Tertiary Education and Research Network of South Africa (TENET) main purpose is to secure, for the benefit of South African universities and associated research and support institutions, Internet and Information Technology services. TENET is a service organization and is committed to service excellence and to services that are strongly aligned and consistent with the organizational requirements of the user community. TENET operates the SANReN network under the terms of a collaboration agreement with the CSIR.

- **SENSE:** The Software-defined network for End-to-end Networked Science at Exascale (SENSE) research project is building smart network services to accelerate scientific discovery in the era of 'big data' driven by Exascale, cloud computing, machine learning and AI. The SENSE SC23 demonstration showcases a comprehensive approach to request and provision end-to-end network services across domains that combines deployment of infrastructure across multiple labs/campuses, SC booths and WAN with a focus on usability, performance and resilience through:
  - Intent-based, interactive, real time application interfaces providing intuitive access to intelligent SDN services for Virtual Organization (VO) services and managers;
  - Policy-guided end-to-end orchestration of network resources, coordinated with the science programs' systems, to enable real time orchestration of computing and storage resources.
  - Auto-provisioning of network devices and Data Transfer Nodes (DTNs);
  - Real time network measurement, analytics and feedback to provide the foundation for full lifecycle status, problem resolution, resilience and coordination between the SENSE intelligent network services, and the science programs' system services.
  - Priority QoS for SENSE enabled flows
  - Multi-point and point-to-point services

- **Integration of OpenALTO, SRv6 and Qualcomm Technologies' GradientGraph**

  We intend to demonstrate the integration of OpenALTO, SRv6 and Qualcomm Technologies' GradientGraph, showing how applications including XR, auto/V2X, and holographic telepresence can benefit from the intelligent routing, rate limiting, and service placement decisions computed by the GradientGraph platform. We also intend to demonstrate the integration of OpenALTO and GradientGraph with science networks, Rucio, FTS, AutoGOLE, SENSE and OpenALTO towards optimizing large-scale data transfers from the LHC at CERN to scientists across the globe.
- **5G/Edge Computing Application Performance Optimization**. UCSD, the Pacific Research Platform, Caltech, Yale and Qualcomm Technologies will attempt to demonstrate the first end-to-end integrated traffic optimization framework based on bottleneck

structure analysis for 5G applications. This intended demonstration will focus on showing the integration of Qualcomm technology's GradientGraph with the IETF ALTO open standard, to support the optimization of edge computing applications such as XR, holographic telepresence, and vehicle networks. Holodecks at UCSD and NYU will be interconnected across the CENIC and NYSERNet regional networks via a transcontinental AP-REX link.

- **High-Performance Routing of Science Network Traffic**. LHC/CERN, Caltech, the Pacific Research Platform, ESnet, Yale and Qualcomm Technologies will attempt to demonstrate the first integration of the IETF ALTO open standard with the Rucio, FTS applications and AutoGOLE/SENSE to optimize the steering of large-scale data transfers through global science networks. Based on the open-source project OpenALTO, the IETF standard exposes highly detailed network state information that the application can use to optimize its performance. This intended demonstration will show how Qualcomm Technologies' GradientGraph platform leverages this information to compute optimized application placement and flow routing strategies.
- **The ALTO-TCN** NRE will demonstrate both a basic capacity: deep visibility into the networking infrastructures supporting data-intensive sciences, and its novel integration into the DOMA control framework at CERN: data orchestration (Rucio) and data transport scheduling (FTS). The visibility demonstration introduces cascading ALTO servers to compose networking visibility across multiple networking domains, addressing the multi-domain visibility challenges in important settings such as LHCONE. The Rucio integration introduces a novel workflow programming framework on top of the basic infrastructure-provider-distance ALTO abstraction, to guide source and destination selections to achieve resource-efficient data orchestration. The FTS integration updates the default semi-gradient FTS optimizer to use efficient zero-order, full gradient optimization, to achieve efficient, fair resource control, which is an important exploration of application-defined networking. The NRE plans to explore opportunities for interaction with both the AUTOGole/SENSE project and the NOTED project, to achieve full-stack (networking and transport layers) infrastructure control.
- **N-DISE: Named Data Networking for Data Intensive Science Experiments:** The NDN for Data Intensive Science Experiments (N-DISE) project aims to accelerate the pace of breakthroughs and innovations in data-intensive science fields such as the Large Hadron Collider (LHC) high energy physics program and the BioGenome and human genome projects. Based on Named Data Networking (NDN), a data-centric future Internet architecture, N-DISE will deploy and commission a highly efficient and field-tested petascale data distribution, caching, access and analysis system serving major science programs. The N-DISE project builds on recently developed high-throughput NDN caching and forwarding methods, containerization techniques, leverage the integration of NDN and SDN systems concepts and algorithms with the mainstream data distribution, processing, and management systems of CMS, as well as the integration with Field Programmable Gate Arrays (FPGA) acceleration subsystems, to produce a system capable of delivering LHC and genomic data over a wide area network at throughputs approaching 100 Gbits per second, while dramatically decreasing download times. N-DISE will leverage existing infrastructure and build an enhanced testbed with high performance NDN data cache servers at participating institutions.

  The N-DISE demonstration is designed to exhibit improved performance of the N-DISE system for workflow acceleration within large-scale data-intensive programs such as the LHC high energy physics, BioGenome and human genome programs. To achieve high performance, the demonstration will leverage the following key components: (1) the transparent integration of NDN with the current CMS software stack via an NDN based XRootD Open Storage System plugin, (2) joint caching and multipath forwarding capabilities of the VIP algorithm, (3) integration with FPGA acceleration subsystems, (4) SDN support for NDN through the work of the Global Network Advancement Group (GNA-G) and its AutoGOLE/SENSE and Data Intensive Sciences Working Group.

- **High performance networking with the Sao Paulo Backbone SP linking 8 universities (Rednesp, RNP, UNESP and USP) and the Bella Link:**

  The research and education network at Sao Paulo (rednesp), formerly ANSP (Academic Network at Sao Paulo), connects dozens of research and education institutions in the State of Sao Paulo, Brazil also providing international connections to the USA and to Europe. After designing it in 2022, rednesp started deploying a new backbone, known as "Backbone SP" connecting 8 major research and education institutions with 100 Gbps links: University of Sao Paulo (USP), University of Sao Paulo State (UNESP), University of Campinas (UNICAMP), Aeronautics Institute of Technology (ITA), Federal University of São Carlos (UFSCAR), Mackenzie University, Federal University of ABC (UFABC) and Federal University at São Paulo (UNIFESP). All links are now fully operational and some additional links are expected soon, including a 100 Gbps link to UNICAMP from the Brazilian Synchrotron Light Laboratory (LNLS) which is part of the Brazilian Center for Research in Energy and Materials (CNPEM), in Campinas (SP). Backbone SP is connected to the USA through two 100 Gbps and two 200 Gbps links (all shared with Amlight and RNP) and to Europe through a 100 Gbps link using the Ellalink cable. There

is also a connection to Chile (to the Large Synoptic Survey Telescope (LSST)), and a new link to Buenos Aires in Argentina (RNP) is expected soon.

The rednesp team will demonstrate technical characteristics such as the latency, bandwidth and jitter of the Backbone SP and its integration with international links. Also, the usefulness of the network in supporting data intensive science applications in areas like high energy physics and cryo-electron microscopy shall be demonstrated.

With the recent availability of an L2 circuit linking the University of São Paulo State (UNESP) to CERN in Geneva, through the Bella Link, Brazil (and São Paulo State) now has three 100 gbps academic links to the USA and to Europe. This new circuit is supposed to present low latency, and that should open new opportunities for collaboration among Brazilian and European academic institutions. All links pass through the SP4 Equinix datacenter in the greater Sao Paulo area, where rednesp has a node called SouthernLight which is part of the GNA Autogole/SENSE testbed.

In this demo, we intend to compare properties such as effective bandwidth, latency to different continents and jitter using different paths of the 3 intercontinental links as well as the integration of the "Backbone SP" to international connections. We also expect to show results and metrics of parallel computations using some of the national and international computing facilities connected to this infrastructure.

- **Toward Fully-Automated Network Configuration Management for Large-Scale Science Networks**

  Configuring large-scale science networks is a time-consuming, labor-intensive and error-prone process. Although substantial progress has been made in automatically verifying the correctness of network configurations, how to effectively diagnose and repair configuration errors remain a less investigated area. Over the past few years, we have developed a series of tools that can automatically and efficiently verify, diagnose and repair network configuration errors. These tools include (1) Coral, a distributed, on-device dataplane verification framework (ACM HotNets'23, ACM SIGCOMM'23), (2) Scalpel, a network configuration diagnosis system using symbolic program analysis (ACM APNet'23), and (3) IVeri, a privacy-preserving, interdomain configuration verification system (IEEE INFOCOM'21, ACM APNet'23). During SC'23, we will demonstrate, through both a testbed reconstruction of the Internet2 WAN environment and large-scale simulations of WAN/LAN/ DC with real-world datasets, that these tools can efficiently automate the network configuration management for large-scale networks.

- **KISTI-SCION: Scalability, Control and Isolation on Next Generation Networks (https://scion-architecture.net)**

  KREONET is collaborating with multiple National Research and Education Networks (NRENs) such as GEANT, SWITCH, BRIDGES, and SingAREN to establish a next-generation Internet infrastructure based on SCION technology. The SCION-based Secure Research and Education Network is in the process of building a global 10G/100G ring network connecting North America, Europe, and Asia. Through SCION's path-based routing technology, routing transparency is achieved, resulting in advanced network security features like hijacking resilience, DDoS defense, and source authentication. Furthermore, by utilizing multi-path communication, data can be transmitted from Daejeon, Korea, to Geneva, Switzerland, simul-taneously using both east- and west-bound network links, harnessing idle resources and achieving twice the performance compared to traditional BGP-based single-path communication.

Resources

The NRE-13 collaborting groups and partner NREs will use optical Waveserver DCI and DWDM platforms in Denver, Los Angeles, and Cisco and Nokia DCIs along with use many 400G continental and some transoceanic links, together with many 100G continental and intercontinental wide area links coming into SC23, and the available on- floor and DCI links to the StarLight and Caltech booths. A 2.4 Tbps wide area network including 400GE switches from Arista, Edgecore, Dell and Juniper, and links among Starlight, McLean, Caltech and the SC23 venue in Denver will be deployed. Together with servers with 100, 200 and potentially 400G smart NICs and NVMe storage systems, and programmable switch routers at the sites running SONIC and/or GEANT/RARE freeRtr as well as fixed function switches in a global network architecture, this will host the wide range of demonstrations summarized in this and its partner NREs as well as others.

As of this writing the GEANT/RARE freeRtr is running in production mode on 100GE Tofino and several 400GE Tofino2 Edgecore switches in the Global P4 Lab that encompasses 38 sites. Server to server transfers between 400G DTNs also will be shown as part of the iCAIR demonstrations led by StarLight and NRL.

The AmLight EXP/CHPC/CHPC/TENET/SANRen/ CaRCC demonstrations hosted by Caltech will be conducted using resources of the collaborating domains, and will involve SACS/WACS/AmLight 100G link and the 400Gbps links from AmLight/Miami to the Caltech booth on the showfloor.

Partners: **Group Leads and Participants, by Team**

- **Caltech HEP:** Harvey Newman (newman@hep.caltech.edu), Justas Balcas (jbalcas@caltech.edu), Raimondas Sirvinskas (raimis.sirvis@gmail.com), Catalin Iordache, Preeti Bhat, Andres Moya, Sravya Uppalapati
- **Caltech IMSS:** Jin Chang (jin.chang@caltech.edu), Azher Mughal (azher@caltech.edu), Dawn Boyd, Don S. Williams
- **UCSD/SDSC/NRP:** Frank Wuerthwein (fkw888@gmail.com), Tom deFanti (tdefanti@eng.ucsd.edu), Larry Smarr, John Graham, Tom Hutton (hutton@ucsd.edu), Dima Mishin, Jonathan Guiang, Diego Davila, Igor Sfiligoi, Aashay Arora
- **Yale:** Richard Yang (yry@cs.yale.edu), Ryan Yang, Jensen Zhang
- **Northeastern University:** Edmund Yeh (eyeh@ece.neu.edu), Yuanhao Wu, Volkan Mutlu, Yuezhou Liu
- **Tennessee Tech:** Susmit Shannigrahi (sshannigrahi@tntech.edu), Sankalpa Timilsina
- **UCLA:** Lixia Zhang (lixia@cs.ucla.edu), Jason Cong (cong@cs.ucla.edu), Michael Lo, Sichen Song
- **Fermilab:** Oliver Gutsche (gutsche@fnal.gov), Phil Demar (demar@fnal.gov)
- **ESnet:** Inder Monga (imonga@es.net), Chin Guok (chin@es.net), Tom Lehman (tlehman@es.net), John MacAuley, Xi Yang, Justas Balcas, Mariam Kiran
- **LBNL/NERSC:** Alex Sim (asim@lbl.gov)
- **Nebraska/UNL:** Garhan Attenbury (garhan.attebury@unl.edu)
- **Vanderbilt:** Andrew Melo, andrew.m.melo@accre.vanderbilt.edu
- **CERN:** Edoardo Martelli (edoardo.martelli@cern.ch), Carmen Misa (carmen.misa@cern.ch)
- **Qualcomm Gradient Graph:** Jordi Ros-Giralt (jros@qti.qualcomm.com)
- **UFES:** Magnos Martinello, Moises R.N. Ribeiro (moises@ele.ufes.br), Christina Dominicini (cristina.dominicini@ifes.edu.br), Everson Borges (everson@ifes.edu.br), Rafael Guimaraes
- **RNP:** Marcos Schwarz (marcos.schwarz@rnp.br), Leandro Ciuffo (leandro.ciuffo@rnp.br)
- **RENATER/GEANT/RARE:** Frédéric LOUI (frederic.loui@renater.fr)
- **UNESP (SPRACE NCC UNESP):** Sergio Novaes (Sergio.Novaes@cern.ch),Rogerio Iope (rogerio.iope@unesp.br)
- **Rednesp:** Antonio J F Francisco, Ney Lemke **(UNESP)** (ney.lemke@unesp.br), Carlos Antonio Ruggiero **(USP)** (toto@ifsc.usp.br), Jorge Marcos de Almeida (USP) (jorge@usp.br), Bruno Baldim (bruno@rednesp.br), Paulo Simionato (simionato@rednesp.br)
- **UERJ:** Alberto Santoro (Alberto.Santoro@cern.ch), **Eduardo Revoredo**
- **UNICAMP:** Sandro Rigo (srigo@unicamp.br)
- **George Mason/BRIDGES:** Bijan Jabbari (bjabbari@gmu.edu), Jerry Sobieski, Liang Zhang
- **Xiamen:** Qiao Xiang (xiangq27@gmail.com), Chenyang Huang, Ridi Wen, Yuxin Wang, Jiwu shu
- **CENIC:** Louis Fox (lfox@cenic.org), Christopher Bruton Sana Bellamine, Tony Nguyen,
- **Pacific Wave/USC:** Celeste Anderson (celestea@usc.edu)
- **Starlight/MREN/iCAIR:** Joe Mambretti (j-mambretti@northwestern.edu), Jim Chen, Fei Yeh
- **Internet2:** Chris Wilkinson (cwilkinson@internet2.edu), Matt Zekauskas (matt@internet2.edu), Christian Todorov
- **AmLight/FIU:** Julio Ibarra (Julio@fiu.edu), Jeronimo Bezerra, Vasilka Chergarova, Anthony Ulloa
- **Amlight/ISI:** Heidi Morgan (hlmorgan@isi.edu)
- **Ciena:** Scott Kohlert (skohlert@ciena.com), Marc Lyonnais, Rod Wilson
- **KISTI/KREONET:** Buseung Cho bscho@kisti.re.kr, Mazahir Hussain, Changjin Park
- **CANARIE:** Thomas Tam (Thomas.Tam@canarie.ca)
- **KAUST:** Alex Moura (alex.moura@kaust.edu.sa), Kevin Sale (kevin.sale@kaust.edu.sa), Jason Roos
- **DE-KIT:** Bruno Hoeft (bruno.hoeft@kit.edu)
- **JPL:** Lee, Carlyn-Ann (Carlyn-Ann.Lee@jpl.nasa.gov)
- **NIST:** Davide Pasavento (davide.pasavento@nist.gov)
- **Hawaii:** Chris Zane (czane@hawaii.edu)
- **SURFNet:** Hans Trompert (hans.trompert@surfnet.nl)
- **CESNET:** Michal Hažlinský, (hazlinsky@cesnet.cz)
- **Clemson:** Cole McKnight (cbmckni@g.clemson.edu)
- **NCHC/TAWREN:** Li-Chi Ku, lku@narlabs.org.tw
- **ETH Zürich:** Adrian Perrig (adrian.perrig@inf.ethz.ch), Jonghoon Kwon (jong.kwon@inf.ethz.ch), François Wirz (wirzf@inf.ethz.ch)
- **OVGU Magdeburg:** David Hausheer (hausheer@ovgu.de), Marten Gartner (marten.gartner@ovgu.de)
- **CSIR:** Kasandra Pillay (kpillay@csir.co.za), Brian Johnston, Johann Hugo, Sibelo Diamini
- **TENET:** Duncan Greaves (ceo@tenet.ac.za), Paul Sullivan
- **SANREN:** Ajay Makan, John Hay, Renier Pelser van Heerden, Thokozani Khwela (thokozani@sanren.ac.za)
- **GNA-G/AArNet:** David Wilde (David.Wilde@aarnet.edu.au)
- **GNA-G AutoGOLE / SENSE WG Members**: https://www.gna-g.net/join-working-group/autogole-sense
- **GNA-G Data Intense Science WG Members:** https://www.gna-g.net/join-working-group/data-intensive-science/
- **STORDIS:** Waldemar.Scheck@stordis.com

Partner NRE Demonstrations:

The NRE demonstrations hosted at or partnering with the Caltech Booth 1255 include:

NRE-001 Joe Mambretti
1.2 Tbps Services WAN Services: Architecture, Technology and Control Systems

NRE-002 Joe Mambretti
End-to-End 400G WAN Services

NRE-003 Joe Mambretti
AP-REX 2.0 Prototype Demonstration

NRE-004 Joe Mambretti
Global Research Platform (GRP)

NRE-005 Edoardo Martelli
LHC Networking And NOTED

NRE-009 Qiao Xiang
Fully-Automated Network Configuration Management for Large-Scale Science Networks

NRE-014 Tom Lehman
AutoGOLE/SENSE: End-to-End Network Services and Workflow Integration

NRE-015 Tom Lehman
SENSE and Rucio/FTS/XRootD Interoperation

NRE-016 Tom Lehman
FABRIC

NRE-019 Jeronimo Bezerra
AmLight 2.0: Flexible control, deep visibility, and programmability @ Tbps!

NRE-020 Marcos Schwarz
Global P4 Lab

NRE-021 Y. Richard Yang
ALTO-TCN: Application-Defined Network Control for Data Intensive Sciences Through Deep Network Visibility

NRE-022 Mariam Kiran
5G on the Showfloor

NRE-024 Alex S. Moura Exploring FDT, QUIC, BBRv2 and HTTP/3 protocols in High Latency WAN paths for use on Data Transfer Nodes

NRE-025 Edmund Yeh
N-DISE: Named Data Networking for Data Intensive Science Experiments

NRE-032 Magnos Martinello
PolKA routing approach to support traffic engineering for data-intensive science