The Global Network Advancement Group

A Next Generation System for the LHC Program and Data Intensive Sciences

Harvey Newman[1,*]

[1]California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125, United States

Abstract. This paper presents the rapid progress, vision and outlook across multiple state of the art development lines within the Global Network Advancement Group (GNA-G) and its Data Intensive Sciences and SENSE/AutoGOLE working groups, which are designed to meet the present and future needs and address the challenges of the Large Hadron Collider and other science programs with global reach. Since it was founded in the Fall of 2019 and the working groups were formed in 2020, in partnership with ESnet, Internet2, CENIC, GEANT, ANA, RNP, StarLight, NRP, N-DISE, AmLight, and many other leading research and education networks and network R&D projects, as well as Caltech, UCSD/SDSC, Fermilab, CERN, LBL, and many other leading universities and laboratories, the GNA-G working groups have deployed two virtual circuit and programmable testbeds spanning six continents which supports continuous developments aimed at the next generation of programmable networks interworking with the science programs' computing and data management systems. The talk covers examples of recent progress in developing and deploying new methods and approaches in multidomain virtual circuits, flow steering, path selection, load balancing and congestion avoidance, segment routing and machine learning based traffic prediction and optimization.

1 Introduction

The GNA-G and its Data Intensive Science and AutoGOLE/SENSE working groups continue to progress and develop multiple toolsets and services designed to meet the needs of the Large Hadron Collider and other data intensive science programs. We continue to evolve the concept of a next generation global system and architecture capable of meeting the challenges of scale and complexity of these programs, making best use of the available R&E network infrastructures worldwide while simultaneously accommodating the network traffic of the at large academic and research community. This includes several major areas where state of the art open source tools and services continue to be developed by our community, while leveraging the latest developments from industry.

2 The GNA-G and Its Data Intensive Sciences Working Group

The (GNA-G) [1] was formed in September 2019 as an open volunteer group devoted to developing the blueprint to make using the Global R&E networks simpler and more effective. Its primary mission is to support global research and education using the technology, infrastructures and investments of its participants. A second important mission of the GNA-G is to act as a data intensive research and 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

*email: newman@hep.caltech.edu

On behalf of the GNA-G and its Data Intensive Sciences and AutoGOLE/SENSE Working Groups. The partner groups and their members can be found at http://gna-g.net/join-working-group/data-intensive-science/partners

The GNA-G Data Intensive Sciences [2] and SENSE/AutoGOLE [3] Working Groups are 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. These groups and their many partners in academia and industry are working in concert to develop a next generation network-integrated system designed to meet the challenges of exabyte data volumes 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.

To meet these goals, the teams are developing 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.

3 GNA-G Data Intensive Sciences WG R&D Projects and Activities

The GNA-G Data Intensive Science working group activities 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.

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.

Themes this year include: (1) Worldwide deployment of the Global P4 Lab, its architecture and capabilities, (2) a broader view of the programmable networking ecosystem (including 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 [4] 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 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. The

GNA-G organization and architecture includes working groups focused on R&D and Production Operations, with the joint mission to 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.

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 and monitoring systems, along with workflow optimization methodologies with real-time analytics, and state of the art long distance data transfer methods and tools and server designs. This development path, composed of multiple branches and a process of progressive integration, is 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, and 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 GNA-G's target programs include the Large Hadron Collider (LHC), BioGenome and the human genome as well as SKA, the Vera Rubin Observatory and others. GNA-G developments include the latest developments in bottleneck structures and analysis for real-time congestion resolution as well as 5G "holodeck" applications running over a segment-routed transcontinental links between the UCSD and NYU campuses. The 5G applications also are helping to 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.

The teams' advanced developments, 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 Gradient-Graph congestion information, Kubernetes statistics, Prometheus, and 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, etc.
- **Orchestration:** SENSE, Kubernetes, dedicated code and APIs for interoperation and progressive integration

3.1 Cornerstone System Concepts and Subsystems Under Development

The cornerstone system concepts, and subsystems under development include:

- Integrated operations and orchestrated management of resources through the site (Site-RM) and network resource managers (Network-RM) developed in the SENSE project.
- Fine-grained end-to-end monitoring and data collection, with a focus on the edges and end sites, enabling analytics-assisted automatic decisions supporting applications with optimized path selection and load balancing mechanisms.
- 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.
- Integration of Qualcomm Technology's GradientGraph (G2) with 5G wireless, OpenALTO, and PolKA-based source routing. These are being used to demonstrate how applications including XR, auto/V2X (Vehicle to anything), and IoT can benefit from the intelligent routing, rate limiting, and service placement decisions computed by G2.
- 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) project.
- A paragon network composed of more than 30 sites 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 platforms in support of workflows for the science programs mentioned. This includes reference designs of NVMe server systems to match a 400G (or multiples of 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 an 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.
- The use of PolKA (Polynomial Key-based Architecture) source routing and flow-control capabilities to provide the to address the challenging traffic engineering (TE) needs of the by the most data intensive science programs. This includes i) data-intensive transfers over 100G and 400G network using PolKA underlay tunnels; and ii) PolKA-based source routing as compared to segment routing methods (such as SR-MPLS or SRv6) over a high-speed intercontinental testbed composed of P4-enabled programmable switches that interconnect data intensive research facilities in North and South America, Europe and Asia. The flows are 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; implemented through GradientGraph) for optimization.
- A dynamic overlay network with PolKA tunnels forming virtual circuits, built by integrating persistent resources from GNA-G AutoGOLE/SENSE and GEANT RARE testbeds. Underlay congestion is detected by tunnel monitoring and signaled 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.

- 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.

4 GNA-G Inter-WG activities

The work in the Data Intensive Sciences WG involves close collaboration with the LHC experiments' network R&D efforts, the GNA-G AutoGOLE/SENSE WG and its global testbed, the GEANT/RARE freertr team and Global P4 testbed, and the National (NRP), Americas (AmRP) and Global (GRP) research platforms, as well as: the major R&E networks notably ESnet, Internet2, CENIC, GEANT, RNP, AmLight, SURFnet, AarNet, rednesp, KREONet and many other national, regional and transoceanic networks as well as the ANA and APONet consortia; advanced network R&D projects and testbeds such as StarLight, Fabric/Fab, AP-Rex 2.0 and Bridges; leading university and laboratory sites including Fermilab, LBNL, BNL, CERN, KISTI, Caltech, Nebraska, Vanderbilt, Yale, UNESP, UFES, Northeastern, UCLA, Tennessee Tech and George Mason. The WG also maintains strong vendor partnerships with Ciena, Arista, Dell, Edgecore, Cisco, Juniper, NVidia and others.

Major themes include:

- The use of multi-domain virtual circuits with bandwidth guarantees provisioned by the SENSE Network and Site Resources Managers and Orchestrator

- High throughput at-scale demonstrations leading to production-ready services interfaced to and in some cases integrated with the science programs' mainstream services for data transfers and overall data management. The first examples are the integration between SENSE, Rucio and FTS in the ATLAS and CMS experiments at the LHC. 400G disk-to-disk throughput between Caltech and UCSD using CMS production software together with SENSE has recently been accomplished.

- Incorporation of the ALTO/TCN (Application Level Traffic Optimization/Transport Control) system being developed for efficient and flexible data transport control using deep infrastructure visibility, as part of the integration between SENSE and Rucio/FTS. This development also uses Qualcomm's GradientGraph to identify, handle and route around or avoid network bottlenecks

- Automation services implemented in Kubernetes and the k8s namespace at both SENSE sites and in the National Research Platform

- Development of a multi-network OS programmable network platform using systems from academia and industry, based on Tofino, SONIC, and modern open source core network operations and management tools such as BGP-CT, gNMI and the gNMIc client, and including smart edge devices such as the Mellanox Bluefield2 and Xilinx U55c

- Development of new segment routing techniques, both the new PolKA approach using a single encoded user-defined label developed by the UFES team, and SRv6 Micro-SIDs.

- Development of new optical technologies both at the edge and in R&E core networks. This includes the use of bidirectional (BiDi) transceivers to increase the connectivity in campuses with minimal impact on long-installed infrastructure and the use of pluggable (currently 400ZR) transceiver/transponders to reduce the cost of bringing multiple 400G connections to campuses. Inclusion of 400ZR+ pluggables in wide-ranging testbed setups also is planned, notably following SC23.

5 The GNA-G DIS WG Next Generation System: Elements, Development Trajectory and Goals

- **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 HPC facilities and the cloud; use of both edge and in-network caches to increase data access and processing efficiency.
- **SENSE/AutoGOLE:** The GNA-G AutoGOLE/SENSE Working Group is deploying 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 which has been installed that spans the US, Europe, Asia Pacific and Latin American regions. The SENSE developments follow a comprehensive approach to requesting and provisioning end-to-end network services across multiple domains that combines deployment of infrastructure across multiple labs/campuses and wide area networks, 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, PolKA and/or SRv6 routing, and Qualcomm Technologies GradientGraph
- **SENSE/Rucio/FTS Integration:**
  SENSE is a multi-resource, multidomain 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. The interoperation of SENSE with the Rucio/FTS/XRootD data management/movement system enables 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.
- **Global Ring and KAUST:**
  The King Abdullah University of Science and Technology (KAUST) is making strides to collaborate with the global Research and Education community by harnessing advanced networking infrastructures.This year's developments highlight KAUST's role 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. KAUST's strategic affiliation with the AER provides a robust platform to facilitate data-intensive research that demands reliable, high-speed, and efficient connectivity. By participating in the GNA-G, KAUST is actively contributing to the advancement of the Global Research and Education ecosystem.

- **400 Gbps End to End Wide Area Network Services**
  As data production among science research collaborations continues to accelerate, the networking community is preparing for service paths beyond 100 Gbps, including 400 Gbps 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. The SC23 demonstrations by Caltech, StarLight and partners will demonstrate 400 Gbps end-to-end WAN services among StarLight, the multi-agency Joint Big Data Testbed (JBDT) Facility in McLean, Virginia, and the SC23 venue.

  Building on the 400G network infrastructures provided by CENIC and Pacific Wave, together with ESnet and Internet2, among Los Angeles, Caltech and UCSD, the Caltech campus and HEP network teams and Ciena are installing two 800G waves between the Caltech campus and the CENIC PoP in LA, to extend the three 400G links being provided by ESnet between the SC23 venue in Denver and LA, using Waveserver Ai and Waveserver 5 platforms carrying four 400G links to the campus and the Caltech CMS SDN testbed and Tier2 facility. This will be the first demonstration of 1+ Tbps wide area and regional link coming to an LHC research team's Tier2 facility, as a forerunner of network and computing operations of the High Luminosity LHC era. This setup may also be extended to CERN once the 400G ESnet transatlantic 400G links are in service, either by SC23 or shortly afterward.

- **400 Gbps Data Transfer Nodes**
  With its national and international partners, iCAIR and the StarLight consortium members and 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 generation smart NICs such as the NVIDIA Bluefield2 and 3, and Xilinx U55c.

- **AmLight Express and Protect (AmLight-ExP)**
  is working in support of the SENSE/AutoGOLE developments and LHC-related use cases, with a focus on the physics groups at FIU and in Latin America who are parts of the LHC physics program or the Vera Rubin Observatory, while also working in close partnership with the Brazilian national network RNP and Sao Paulo regional network Rednesp and its consortium of associated universities. Developments led by AmLight-EXP include high-throughput low latency experiments, and demonstrations of auto- recovery from network events, 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.

- **Integration of OpenALTO, SRv6 and Qualcomm Technologies' GradientGraph:**
  These developments are designed to show 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.

- **ALTO-TCN Deep Visibility: FTS and Rucio Integration at the Transport Layer** These developments bring deep visibility into the networking infrastructures supporting data-intensive sciences, as well as incorporating a novel activity-based allocation approach into the control framework which has been developed by the Data Organization Management and Access (DOMA) group at CERN. This includes the data orchestration services provided by Rucio, and the data transport scheduling provided by FTS. Plans underway include interaction with both the AutoGOLE/SENSE project and the NOTED project, to achieve full-stack (networking and transport layers) infrastructure control.

- **N-DISE:** 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 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 is deploying and commissioning a highly efficient 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, and leverages 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.

  The N-DISE developments are designed to exhibit improved performance of 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 leverages 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**
  linking 8 universities through Rednesp and RNP, and the Bella (Brazil-Portugal) 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). The Rednesp team is studying, field-testing and improving the technical characteristics of the new regional and transatlantic network, such as the latency, bandwidth and jitter of the Backbone SP, and its integration with international links. With the recent availability of a Layer 2 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 strategically placed to provide low latency, and thus open new opportunities for collaboration among Brazilian and European academic institutions.

- **Towards 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).

6 Resources

The GNA-G uses a global footprint of 100G and an increasing number 400G regional and wide area links, encompassed in the AutoGOLE/SENSE and Global P4 testbeds described above. Following the upcoming Supercomputing 2023 (SC23) conference, plans include the deployment of a multi-Tbps wide area network including 400GE switches from Arista, Edgecore, Dell, Ciena and Juniper, and links among Starlight, McLean, Caltech, and the Ciena and ESnet testbeds, in addition to the Tbps FABRIC testbed. Together with servers with 100G, 200G and some 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 developments and leading edge demonstration s and deployments, as summarized above.

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, to be first shown at SC23 as part of the iCAIR demonstrations led by StarLight and NRL, will subsequently be shown in pre-production prototypes, leading to production-style deployments in 2024 and 2025.

7 Conclusions

The mission and driving principles of the GNA-G and its Data Intensive Sciences and AutoGOLE/SENSE working groups, and the motivation for the vast body of ongoing work reviewed above can be summarized as follows:

To realize the physics discovery potential and meet the challenges of the HL LHC and several other data intensive science programs with similar needs, 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
- Negotiates with site systems that aim to accelerate workflow
- Builds on multiple ongoing R&D projects: from regional caches/data lakes to intelligent control and data planes to machine learning-based optimization
- Leverages the worldwide move towards a fully programmable ecosystem of networks and end-systems (P4, SONIC; PolKA, SRv6), plus operations platforms including the Open

Science Grid, National Research Platform, ans the global SENSE/AutoGOLE and Global P4 Lab testbeds

- Simultaneously supports the LHC experiments, other data intensive programs and the larger worldwide academic and research community

The LHC experiments and other major science programs such as LBNF/DUNE, VRO, and the SKA together with the WLCG, the GNA-G and its working groups, and the worldwide R&E network community are key players in this ongoing process, which has been deigned to meet the otherwise daunting technological and operational challenges, and clear the way for the next round of scientific discoveries.

8 Acknowledgements

This work has been partially supported by the US National Science Foundation (NSF) Grants OAC-2030508, OAC-2019012, OAC-1841530, OAC-1836650, OAC-1659403, 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, DE-SC0015528, 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, and the generous support of industry partners most notably Arista, Ciena, STORDIS and Dell.

References

[1] The Global Network Advancement Group, https://www.gna-g.net/

[2] The Global Network Advancement Group's Data Intensive Sciences Working Group, https://www.gna-g.net/join-working-group/data-intensive-science/

[3] The Global Network Advancement Group's AutoGOLE/SENSE Working Group, https://www.gna-g.net/join-working-group/autogole-sense/

[4] Containerlab (https://containerlab.dev) launches network laboratories including a topology that includes interconnections among high-fidelity representations of containerized network operating systems of all the major network vendors.