SC21 Network Research Exhibition: Demonstration Abstract

N-DISE: NDN for Data Intensive Science Experiments

Submitted by the N-DISE Team

Abstract

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 will build 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.

Goals

This 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) containerization techniques to simplify the deployment and configuration of the NDN-DPDK forwarder, (3) joint caching and multipath forwarding capabilities of the VIP algorithm, (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. The demonstration activities will take place over an upgraded N-DISE WAN testbed connecting participating institutions (Northeastern, Caltech, UCLA, and Tennessee Tech) including several 100G links, and 4 X 100G connectivity between Caltech and the CENIC PoP in Los Angeles.

Integration of NDN with CMS Software Stack

We will present a transparent integration of NDN with the current CMS software stack via an NDN based XRootD Open Storage System plugin [1]. This includes developing containerized NDN C++ consumer and producer applications able to communicate with the NDN-DPDK forwarder developed by NIST [2], using the CISCO developed shared memory interface (memif) library. We will showcase real-time CMS job completion using the NDN XRootD OSS plugin on a testbed running applications in Docker container using the NDN-DPDK forwarder.

Containerization

As part of the NRE demo, we will utilize containerization techniques to package NDN-DPDK forwarder along with the HEP producer and consumers. We will then utilize these deployments to move HEP data over wide area networks. NDN-DPDK forwarders are time intensive to configure and deploy. Our goal is to utilize Docker and similar technologies to simplify this deployment and configuration step. After this containerization, any community that aims to deploy NDN-DPDK will be able to simply deploy a docker container. However, such deployments come with a slight performance penalty. We have tested docker containers between multiple machines on local testbeds. Through this demo we plan to demonstrate the feasibility of NDN-DPDK deployment using Docker and evaluate forwarding performance over wide area networks.

Joint Caching and Multipath Forwarding

We will demonstrate the performance of the adaptive, distributed VIP joint caching and forwarding algorithm [3], which has been implemented with the NDN-DPDK forwarder over the N-DISE WAN testbed. Particular emphasis will be placed on evaluating the performance of joint caching and multipath forwarding capabilities of the VIP algorithm, over more complex network scenarios enabled by the N-DISE WAN testbed. Specifically, path diversity and multipath forwarding is expected to yield more caching opportunities, leading to improved performance in terms of throughput, delay, and cache hit rates.

SDN Support for N-DISE

SDN support for NDN in a multi-domain testbed is advancing rapidly through the work of the Global Network Advancement Group (GNA-G) and its AutoGOLE/SENSE and Data Intensive Sciences Working Groups. We will use a set of SENSE-enabled virtual circuits with bandwidth guarantees spanning multiple network domains, with nodes at the N-DISE collaborating sites (Caltech, Northeastern, Tennessee Tech, UCLA) and others including StarLight and SCinet, and an intelligent data plane and control plane using P4 and the Reservoir Labs gradient graph (G2) software. This installation will support decisions on impactful flow (group) identification, gradient graph calculations providing recommendations for congestion resolution or avoidance by moving selected flows to alternative paths and/or adjusting virtual circuit capacities. The P4 user-defined label parsing and matching capabilities will be used to enable flow-group identification and express flow handling requirements and decisions, and in-band telemetry (INT) will be used to track the end-to-end state and progress of flow-group transfers.

Demonstration Testbed

The N-DISE WAN testbed builds on the testbed of the SANDIE (SDN-Assisted NDN for Data Intensive Experiments) project, and connects N-DISE nodes at Northeastern University (NEU), Caltech, and UCLA, and Tennessee Tech, which are equipped with high-performance intelligent NICs, including Mellanox ConnectX-5 and ConnectX-6. NEU nodes will be relocated at the Massachusetts Green High-Performance Computing Center (MGHPCC), upgrading the connection to 100 Gb level. The path from Caltech to NEU will be controlled by SENSE, which can provide virtual circuits with a 100 Gb bandwidth guarantee. The layer 2 demo topology is provided by SCinet in collaboration with Internet2, ESnet, and CENIC and other regional providers to the nodes at NEU, Caltech, UCLA, Tennessee Tech and Starlight (Chicago). The nodes are connected to all other nodes and to the StarLight booth at SC21 at layer 2 through the use of VLANs. The NDN forwarder used for the demo, NDN-DPDK, runs on top of this layer 2 topology as the layer 3 forwarder.

Involved Parties

- Edmund Yeh, Northeastern Univ., eyeh@ece.neu.edu
- Harvey Newman, Caltech, newman@hep.caltech.edu
- Lixia Zhang, UCLA, lixia@cs.ucla.edu
- Jason Cong, UCLA, cong@cs.ucla.edu
- Susmit Shannigrahi, Tennessee Tech, sshannigrahi@tntech.edu
- Yuanhao Wu, Northeastern, wu.yuanh@husky.neu.edu
- Volkan Mutlu, Northeastern, fvmutlu@ece.neu.edu
- Yuezhou Liu, Northeastern, liu.yuez@husky.neu.edu
- Catalin Iordache, Caltech, catalinn.iordache@gmail.com
- Justas Balcas, Caltech, justas.balcas@cern.ch
- Raimondas Sirvinskas, Caltech, raimis.sirvis@gmail.com
- Sichen Song, UCLA, songsichen123@gmail.com
- Michael Lo, UCLA, milo168@g.ucla.edu
- Sankalpa Timilsina, Tennessee Tech, sankalpatimilsina12@gmail.com
- Davide Pesavento, davide.pesavento@nist.gov
- Chengyu Fan, Colorado State University, chengy.fan@gmail.com

References

[1] HEP Applications: Iordache, Cătălin, et al. "Named Data Networking based File Access for XRootD." *EPJ Web of Conferences*. Vol. 245. EDP Sciences, 2020.

[2] NDN-DPDK: Shi, Junxiao, Davide Pesavento, and Lotfi Benmohamed. "NDN-DPDK: NDN Forwarding at 100 Gbps on Commodity Hardware." *Proceedings of the 7th ACM Conference on Information-Centric Networking*. 2020.

[3] Yeh, et al. "VIP: A framework for joint dynamic forwarding and caching in named data networks." *Proceedings of the 1st ACM Conference on Information-Centric Networking*. 2014.