SC21 Network Research Exhibition: Demonstration Preliminary Abstract

LHC Networking And NOTED

Edoardo Martelli (edoardo@cern.ch), Joanna Waczynska (joanna.waczynska@cern.ch)

Abstract

This NRE abstract describes an experimental technique, NOTED (Network Optimized for Transfer of Experimental Data), being developed by CERN for potential use by the Large Hadron Collider (LHC) networking community. This SC21 NRE will demonstrate the capabilities of NOTED using an international networking testbed.

Overview

The goal of the NOTED project is to optimize transfers of LHC data among sites by addressing problems such as saturation, contention, congestion, and other impairments.

The Worldwide LHC Computing Grid (WLCG - a global collaboration of approximately 200 interconnected computing centers) provides global computing, storage, distribution, and analytic resources supporting physics experiments using data generated by the (LHC) experiments at CERN. The WLCG's three tier structure (Tier 0 at CERN, Tier 1, and Tier 2 sites are interconnected by global high performance multi-domain networks, the LHC Optical Private Network (LHCOPN) and the LHC Open Network Environment (LHCONE).

The NOTED optimization method employs a combination of a) a deep understanding of the network traffic acquired by an analysis of the data flows with b) an appropriate response (e.g., dynamic allocation of additional capacity) to specific patterns detected among those flows.

For the WLCG, these data flows are generated by the File Transfer Service (FTS). NOTED has information about (and an interface to) job queues using FTS and Rucio. NOTED obtains FTS information about these flows from the Computing Resource Information Catalog (CRIC) database (http://cms-cric.cern.ch), specifically to identify the site network prefixes (IPv4/IPv6) of the storage elements generating and/or receiving the flows. No specific information on network topologies or configurations is required.

Aggregation is a key next step, the grouping of transfer information. This aggregation stage is a critical element for the decision making process. This step enables combining multiple totally independent FTS decisions to provide a comprehensive overview that allows an assessment of potential impacts by all flows on one or more network path segments, i.e., allowing anticipation of future contention/congestion.

This process also allows for the classification of transfers and detecting those that will generate or are generating large volumes of network traffic and group endpoints by specific sites (prefixes), aggregating sites. Because this process has its own custom designed controller, this information can be used for decision making (match action/response), e.g., via dynamic controller configurations allowing potentially for full automation of decision-making regarding network allocations and reconfigurations. Decision making can involve an almost unlimited number of variables, singly or in combination, such as sites, specific flows, priorities, capacity required, capacity available, impairment issues, e.g., congestions, timing/scheduling/duration considerations, and others.

To optimize such decision making, this project is using the AI Conv-LSTM model to forecast traffic, using behavior patterns, within a time window, e.g., 5 min. As an initial experiment, the model was trained to predict traffic based on information about transfers from FTS (from TRIUMF-SFU to Tier0/Tier1). 2. from TRIUMF-SFU to Tier0/Tier1 and from Tier0/Tier1 to TRIUMF-SFU. This forecasting has been based on aggregated information about transfers from the previous 20 minutes.

Currently, consideration is being given integrating this capability with the LHCONE Point-To-Point Service (LHCONE P2P). Also, this project is investigating using this service with the AutoGOLE, a worldwide collaboration of Open eXchange Points and Research & Education networks that deliver end-to-end automated network services, relying on connection requests using the Network Service Interface Connection Service (NSI-CS). Increasingly, R&E networks across the globe use NSI to provision and release international network services to facilitate multi-domain network service requests, provisioning circuits within minutes.

Goals

The SC21 goal of this NRE experiment/demonstration is to showcase the capabilities of NOTED for optimizing data intensive science.

Resources

This NRE demonstration will be conducted using resources of the collaborating domains, including:

Global high-bandwidth paths among Open Exchange Points and the SC21 venue.

Open Exchange Points, including CERNLight, NetherLight, the Montreal Open Exchange (MOXY), and the StarLight International Exchange

The SCinet WAN

Data Transfer Nodes (DTNs), which are important resources for data intensive science such as the high energy physics (HEP) community. (DTNs across the world allow for fast and efficient transport of data over long distances, in part via caching mechanisms.)

Involved Parties

- Edoardo Martelli, CERN, edoardo@cern.ch),
- Joanna Waczynska (joanna.waczynska@cern.ch)
- Sofia Vallecorsa, CERN IT-DI-OPL, Sofia.Vallecorsa@cern.ch
- Edward Karavakis, IT-ST-PDS edward.Karavakis@cern.ch
- Florent Parent, Compute Canada,
- Thomas Tam, CANARIE Thomas.Tam@canarie.ca
- Bruno Hoeft (SCC)DE-KIT, bruno.hoeft@kit.edu
- Vitaliy Kondratenko, TRIUMF, vitaliyk@triumf.ca
- Joe Mambretti, StarLight/MREN j-mambretti@northwestern.edu
- Fei Yeh, StarLight, fyeh@northwestern.edu