Networking Areas of Interest for HEP

High Energy Physics (HEP) experiments, as well as many similar data intensive, global science domains, rely on networks as one of the critical components of their infrastructure both within the participating laboratories and sites as well as globally to interconnect those sites, data centres and experiments instrumentation. Recent work by the HEPiX Network Function Virtualization working group has created a Phase I report[1] about activities underway both within HEP and more broadly within industry. The report surveyed ongoing activities relevant to HEP from the datacenter to the WAN and suggested possible areas of interest that HEP may want to explore with the various NRENs and network researchers. The report was presented and discussed at the January 13-14th 2020 LHCONE/LHCOPN meeting[2] at CERN and resulted in a recommendation to pursue (at least) three areas of network development considered useful to at least one or more WLCG and HEP experiments.

- **Making our network use visible** - Understanding the HEP traffic flows in detail is critical for understanding how our complex systems are actually using the network. With a standardized way of marking traffic, any NREN or end-site could quickly provide detailed visibility into HEP traffic to and from their site, a benefit for NRENs and users.
- **Shaping data flows** - It remains a challenge for HEP storage endpoints to utilize the net-work efficiently and fully. Shaping flows via packet pacing to better match the end-to-end usable throughput results in smoother flows which are much friendlier to other users of the network by not bursting and causing buffer overflows.
- **Network orchestration to enable multi-site infrastructures** - Within our data centers,technologies like OpenStack and Kubernetes are being leveraged to create very dy-namic infrastructures to meet a range of needs. Critical for these technologies is a level of automation for the required networking using both software defined networking and network function virtualization. As we look toward HL-LHC, the experiments are trying to find tools, technologies and improved workflows that may help bridge the anticipated gap between the resources we can afford and what will actually be required to extract new physics from the massive data we expect to produce. To support this type of resource organiza-tion evolution, we need to begin to prototype and understand what services and interactions are required from the network. We would suggest a sequence of limited scope proof-of-principle activities in this area would be beneficial for all our stakeholders.

In addition, our workflow systems need increased network awareness and orchestration moving toward the HL-LHC era. There are significant challenges to deliver a workflow system capable of effectively interacting with the network. In a previous NSF project (ANSE project[3]) a few of us

[1] https://doi.org/10.5281/zenodo.3565562
[2] https://indico.cern.ch/event/828520/
[3] https://www.nsf.gov/awardsearch/showAward?AWD_ID=1246133

View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Clemson University: TigerPrints

instrumented the workflow and data management systems in ATLAS and CMS with network “hooks”, preparing for future production quality SDN systems. It will be interesting to explore what is possible to augment our LHC experiments with SDN capability and it is important to start discussing this with the broader research community.