Co-Design of Software Defined Network and Exascale Science Flows

NAME: I. Foster[1], D. Katramatos[2], R. Kettimuthu[1], Q. Liu[4], S. Sen[4], B. Settlemyer[3], D. Yu[2]
AFFILIATION: 1: ANL, 2: BNL, 3: LANL, 4: ORNL
EMAIL:foster@anl.gov, dkat@bnl.gov, kettimut@anl.gov, qiangliu@ece.sunysb.edu, sens@ornl.gov, bws@lanl.gov, dtyu@bnl.gov
EXPERTISE: High performance networks, SDN, workflow.

KEY DOE NETWORK/TRANSPORT CHALLENGE

The DOE has made significant investments in infrastructure and facilities to support the extreme-scale science workflows that utilize large experimental facilities and/or high-performance computing systems. Such workflows are expected to generate large volumes of data in the order of exabytes and higher. The experimental, computing, archival, analysis, and visualization facilities that support these workflows are distributed across the DOE science facilities and involve users from national laboratories, universities, and international partners. The data generated, during or after the experiments and simulations at these science facilities, must often be transported within data center networks and over wide-area networks for the purpose of storage, visualization, analysis, and other tasks. The high-performance network flows needed for these science capabilities, referred to as science network flows, require simultaneous orchestration across multiple network segments and technologies to ensure that the available networking resources are used efficiently. These flows are quite complex to design and implement, typically requiring manual composition and configuration by teams of experts, but their construction needs to be simple and easy to achieve by the science users and applications. The challenge is to transition from complicated, labor-intensive, and costly manual configurations, which often result in over-provisioning and thus waste of resources, to the automatic, optimized provisioning without placing undue demands on users to understand and account for the layout and architecture of the underlying systems and networks. Clearly, this challenge is expected to become even more important and complicated in future due to the increasing sophistication of hosts, storage systems, and network technologies that constitute the exascale trajectory.

DOE RESEARCH CHALLENGE DIRECTION

Software Defined Networking (SDN), Network Function Virtualization (NFV), and related technologies hold a big promise in addressing the mentioned challenge to develop fast and robust automatic provisioning of the underlying network capabilities. Higher software layers can be built on top of these technologies to enable a sophisticated composition of science flows together with dynamic parameter optimizations at scales well beyond the current (mostly manual) operations. Expanding the SDN control plane and services into computing, storage and IO systems leads to the concept of Software Defined Network Science Flows that hold the promise of unprecedented speed, performance, and capabilities for the DOE science workflows. The science network flows represent a completely different set of challenges and the DOE facility environments are quite unlike the mainstream Internet and data center environments where the current SDN and NFV technologies are being developed. Thus, it will be necessary to research and develop a custom software networking infrastructure that enables high-performance science flows from a limited number of users at known sites connected over high bandwidth optical networks, and is also able to handle the predominant features of science network flows, i.e., a small number of high volume and precision flows over complex network, IO, and computation paths. Clearly, the required solutions are outside the planned trajectory of pre-packaged SDN and NFV technologies which mainly focus on big numbers of smaller flows, including flows from mobile devices. In addition to developing the technologies, robust transition plans are needed to facilitate their integration into DOE sites, including satisfying the strict cyber security requirements which are understood only to a very limited extent within these SDN and NFV contexts.