BigData Express: A scalable and high-performance data transfer platform

Abstract

Big data has emerged as a driving force for scientific discoveries. To meet data transfer challenges in big data era, DOE's Advanced Scientific Computing Research (ASCR) office has funded the BigData Express project (http://bigdataexpress.fnal.gov). BigData Express is targeted at providing *schedulable, predictable,* and *high-performance* data transfer service for DOE's large-scale science computing facilities and their collaborators.

In this demo, we use BigData Express software to demonstrate bulk data movement over wide area networks. The following features in BigData Express will be demonstrated:

- A peer-to-peer, scalable, and extensible model for data transfer services;
- A visually appealing, easy-to-use web portal;
- A high-performance data transfer engine;
- Orchestrating and scheduling of system (DTN), storage, and network (SDN) resources involved in the file transfers.
- On-Demand provisioning of end-to-end network paths with guaranteed QoS;
- Robust data transfer services provisioning through strong error handling mechanisms;
- Safe and secure data transfer services by using multiple security mechanisms;
- The interoperation between BigData Express and SENSE;
- Integration of BigData Express with scientific workflows.

I. Overview

Large scientific instruments (e.g., colliders, light sources, and telescopes) generate exponentially increasing volumes of data. To enable scientific discovery, science data must be collected, indexed, archived, shared, and analyzed, typically in a widely distributed, highly collaborative manner. Data transfer has become an essential function for science discoveries, particularly within big data environments.

The emergence of distributed, extreme-scale science applications has generated significant challenges regarding data transfer. We assert the data transfer challenges of the extreme-scale era are characterized by two relevant dimensions:

(1) High-performance challenges. It is becoming critical to transfer data at the highest possible throughputs because the volumes of science data are growing exponentially. DOE is deploying extreme-scale supercomputer facilities in support of extreme-scale science applications. To fully utilize these expensive computing resources, ultra-high-throughput data transfer capabilities will be required to move data in or out of them.

(2) Time-constraint challenges. Scientific applications typically have explicit or implicit time constraints on data transfer. Providing real-time and deadline-bound data transfer is a challenging task in the extreme-scale era.

DOE's Office of Science Advanced Scientific Computing Research (ASCR) office has funded the BigData Express Project (http://bigdataexpress.fnal.gov) to address these challenges. BigData Express is targeted at providing schedulable, predictable, and high-performance data transfer service for DOE's large-scale science computing facilities and their collaborators.

BigData Express software is currently deployed and being evaluated at multiple research institutions, including UMD, StarLight, FNAL, KISTI, KSTAR, SURFnet, Ciena, and AmLight. The BigData Express research team is currently collaborating with StarLight to deploy BigData Express on emerging research platforms, including Pacific Research Platform, National Research Platform, and Global Research Platform. We are working toward utilizing BigData Express to build high-performance data transfer federations within big data science.

In this demo, we use BigData Express software to demonstrate bulk data movement over wide area networks. The following features in BigData Express will be demonstrated:

- A peer-to-peer, scalable, and extensible model for data transfer services;
- A visually appealing, easy-to-use web portal;
- A high-performance data transfer engine;
- Orchestrating and scheduling of system (DTN), storage, and network (SDN) resources involved in the file transfers.
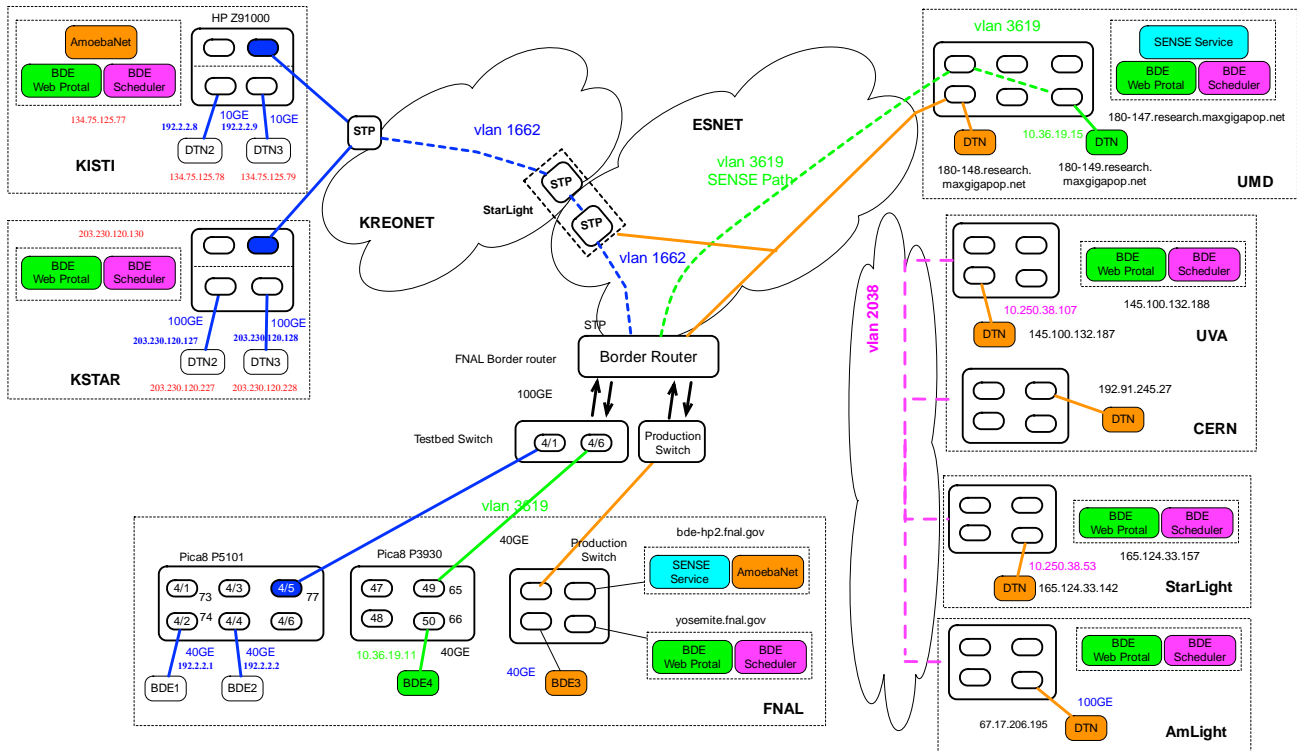
Figure 1 BigData Express SC'19 Demo

- On-Demand provisioning of end-to-end network paths with guaranteed QoS;
- Robust data transfer services provisioning through strong error handling mechanisms;
- Safe and secure data transfer services by using multiple security mechanisms;
- The interoperation between BigData Express and SENSE;
- Integration of BigData Express with scientific workflows.

II. Innovation

BDE is a high-performance data transfer service for big data. It contains a number of new and innovative features:

- A data-transfer-centric architecture to seamlessly integrate and effectively coordinate computing resources in an end-to-end data transfer loop.
- A distributed peer-to-to model for data transfer services, making it very flexible for the establishment of data transfer federations.
- A scalable software architecture. BigData Express makes use of MQTT as message bus to support communication among its components.
- An extensible plugin framework to support different data transfer protocols, including mdtmFTP, GridFTP, and XrootD.
- An end-to-end data transfer model with fast provisioning of end-to-end network paths for guaranteed QoS. Specifically, the use of an SDN-

enabled BigData-Express LANs and SDN-enabled WAN path services to reduce or eliminate network congestion.
- A high-performance data transfer engine. BigData Express adopts mdtmFTP as its default data transfer engine. mdtmFTP is specifically designed for optimization of data transfer performance on multicore systems (DTNs).
- A rich set of REST APIs to support scientific workflows.

III. Demonstration Resources

BigData Express software is currently deployed and evaluated at multiple sites, including FNAL, StarLight, UMD, KISTI, KSTAR, SURFnet, and AmLight. We are working with involved parties to implement high-bandwidth layer-2 circuits between these sites for our demonstrations. The planned network topology is illustrated in Figure 1.

- High-end Data Transfer Nodes (DTNs) are deployed at each site. Each DTN is back-ended with high-performance storage, such as SSD &RAID, and configured with one or multiple 40/100 GigE NICs.
- FNAL and KISTI have deployed an SDN-enabled LAN.
- ESnet provides programmable, on-demand WAN connection services among sites via OSCARs and SENSE, which support guaranteed bandwidth and designated time slot reservation.

IV. Involved Parties

- Qiming Lu, Fermilab, qlu@fnal.gov
- Liang Zhang, Fermilab, liangz@fnal.gov
- Wenji Wu, Fermilab, wenji@fnal.gov
- Phil DeMar, Fermilab, demar@fnal.gov
- Jim Chen, iCAIR, jim-chen@northwestern.edu
- Joe Mambretti, iCAIR, j-mambretti@northwestern.edu
- Chin Guok, ESnet, chin@es.net
- Inder Monga, ESnet, imonga@es.net
- Tom Lehman, Virnao, tom.w.lehman@gmail.com
- Xi Yang, UMD, maxyang@umd.edu
- Jin Kim, KISTI, jkim@kisti.re.kr
- Buseung Cho, KREONET, bscho@kisti.re.kr
- Julio Ibarra, Julio@fiu.edu
- Heidi Morgan, hlmorgan@isi.edu
- Jeronimo Bezerra, jbezerra@fiu.edu
- Vasilka Chergarova, vchergar@fiu.edu
- Adil Zahir, azahir@fiu.edu
- Gerben van Malenstein, Gerben.vanMalenstein@surfnet.nl

V. Related NRE Demonstrations

Other NRE demonstration which include collaborations with the BigData Express project include the following:

- SC19-NRE-019 - Global Petascale to Exascale Workflows for Data Intensive Science
- SC19-NRE-020 - LHC Multi-Resource, Multi-Domain Orchestration via AutoGOLE, SENSE, and TIFR
- SC19-NRE-013 - SENSE: Intelligent network services for science workflows