FERMILAB-SLIDES-19-054-SCD

Fermilab

U.S. DEPARTMENT OF ENERGY | Office of Science

BigData Express: Toward Predictable, Schedulable, and High-Performance Data Transfer

BigData Express Research Team

September 18, 2019

Acknowledgments

This work was supported by the U.S. DOE Office of Science ASCR network research program

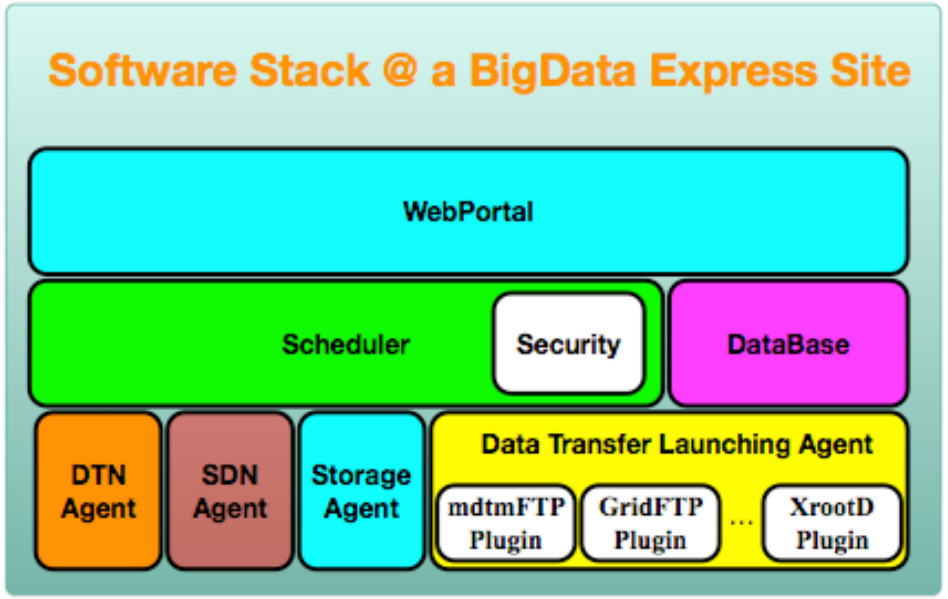
BigData Express

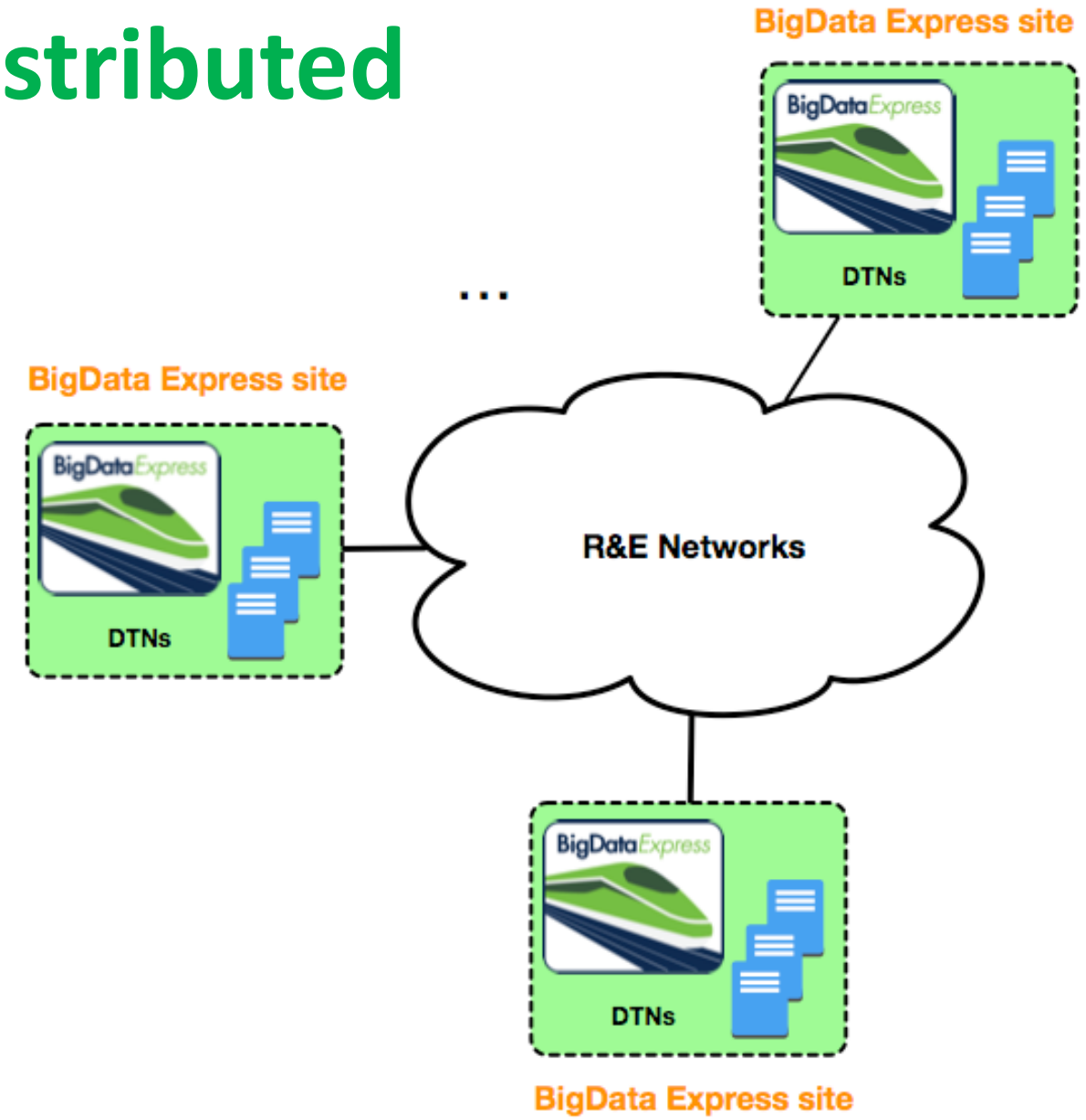
- BigData Express: a schedulable, predictable, and high-performance data transfer service
  - A peer-to-peer, scalable, and extensible data transfer model
  - A visually appealing, easy-to-use web portal
  - A high-performance data transfer engine
  - A time-constraint-based scheduler
  - On-demand provisioning of end-to-end network paths with guaranteed QoS
  - Robust and flexible error handling
  - CILogon-based security

BigData Express Major Components

- **BigData Express Web Portal**
  - Access to BigData Express services
- **BigData Express Scheduler**
  - Time-constraint-based scheduler
  - Co-scheduling DTN, storage, & network
- **AmoebaNet**
  - Network as a service
  - Rate control
- **mdtmFTP**
  - High-performance data transfer engine
  - http://mdtm.fnal.gov
- **DTN Agent**
  - Manage and configure DTNs
  - Collect & report DTN configuration and status
- **Storage Agent**
  - Manage and configure storage systems
  - I/O estimation
- **Data Transfer Launching Agent**
  - Launch data transfer jobs
  - Support different data transfer protocols

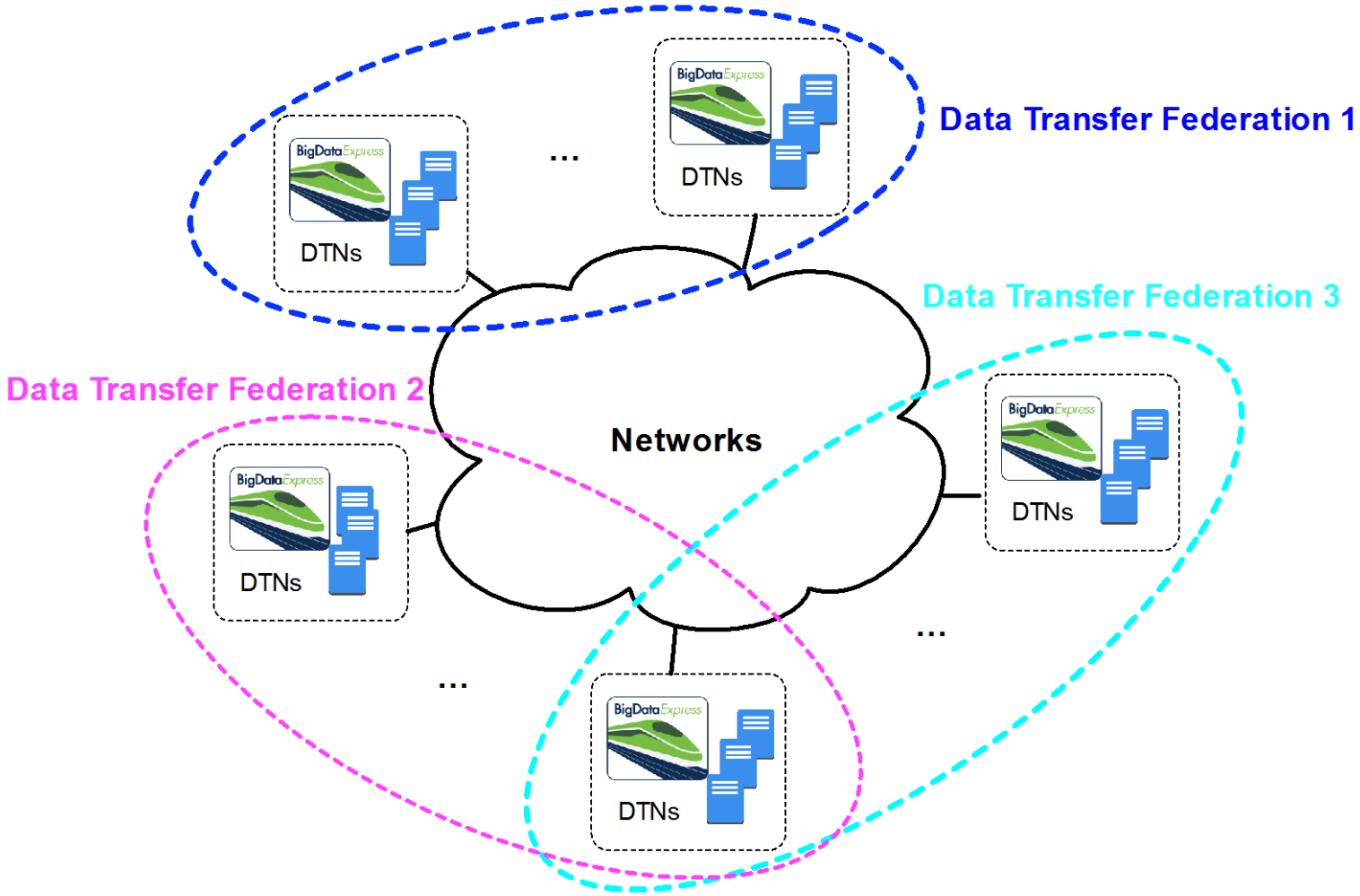
BigData Express -- Distributed





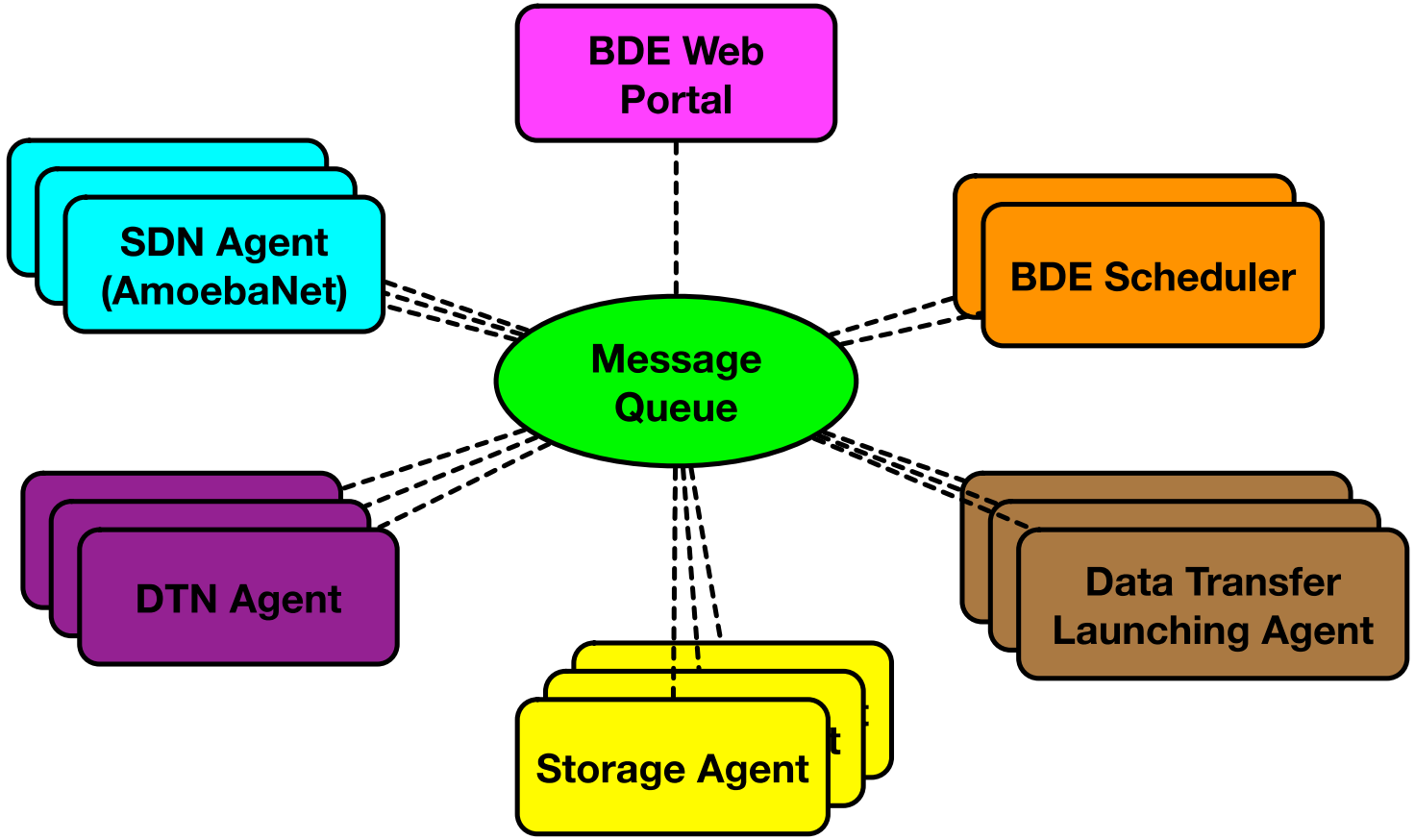
A Peer-to-Peer model

BigData Express -- Flexible



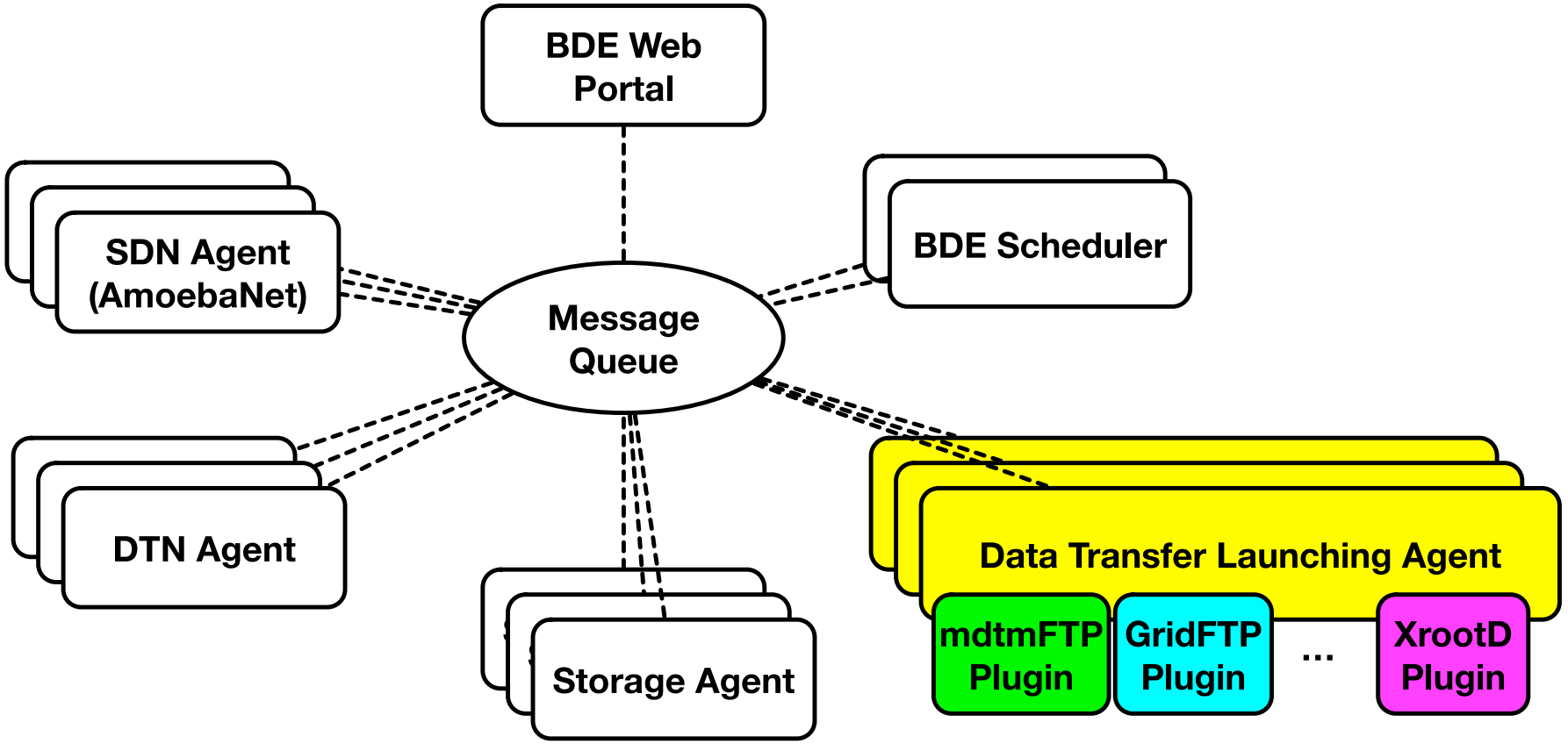
- **Flexible to set up data transfer federations**
- **Providing inherent support for incremental deployment**

BigData Express -- Scalable



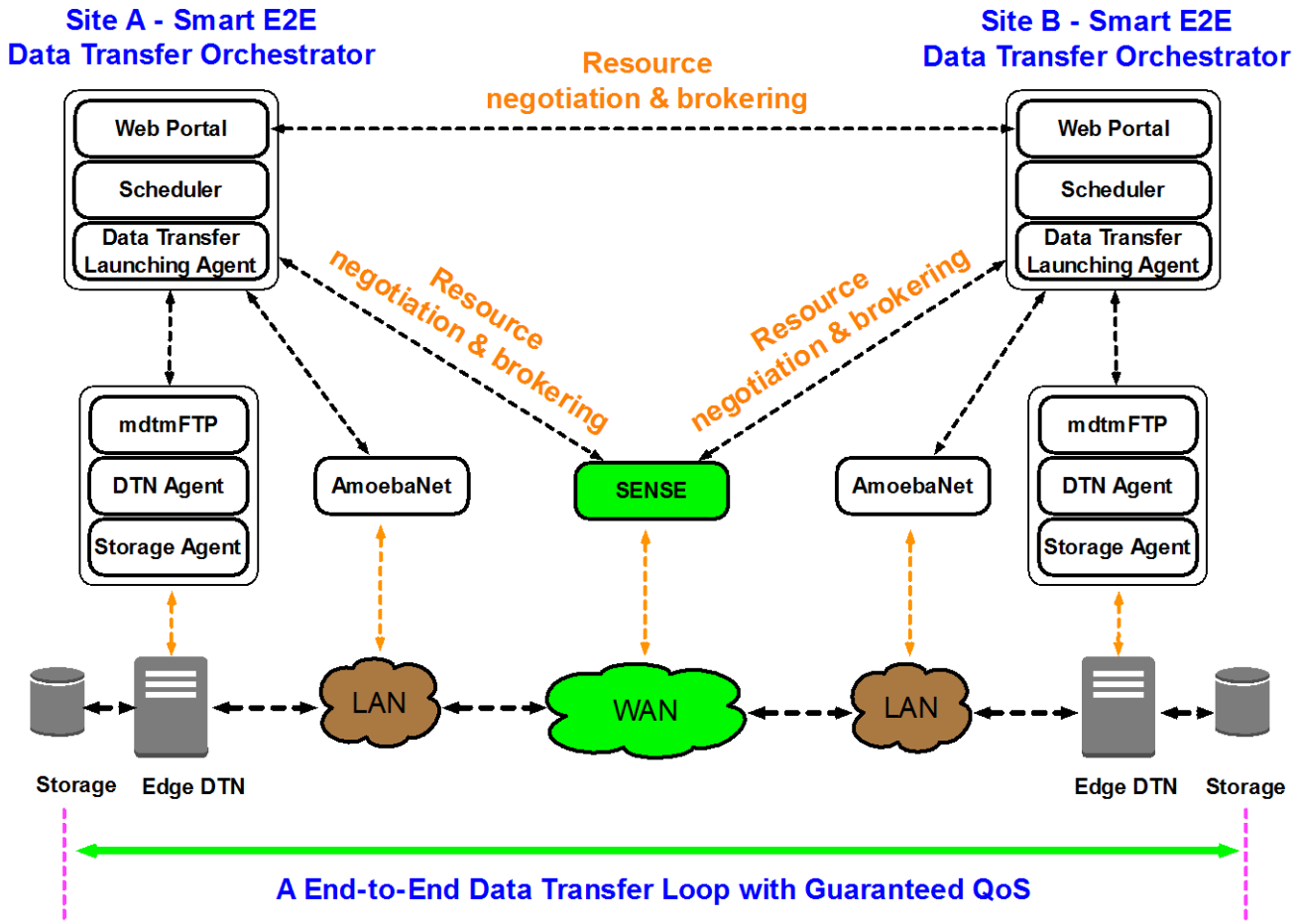
- BigData Express scheduler manages site resources through agents
- Use MQTT as message bus

BigData Express -- Extensible



- Extensible Plugin framework to support various data transfer protocols
  - mdtmFTP, GridFTP, XrootD, …

BigData Express -- End-to-End Data Transfer Model

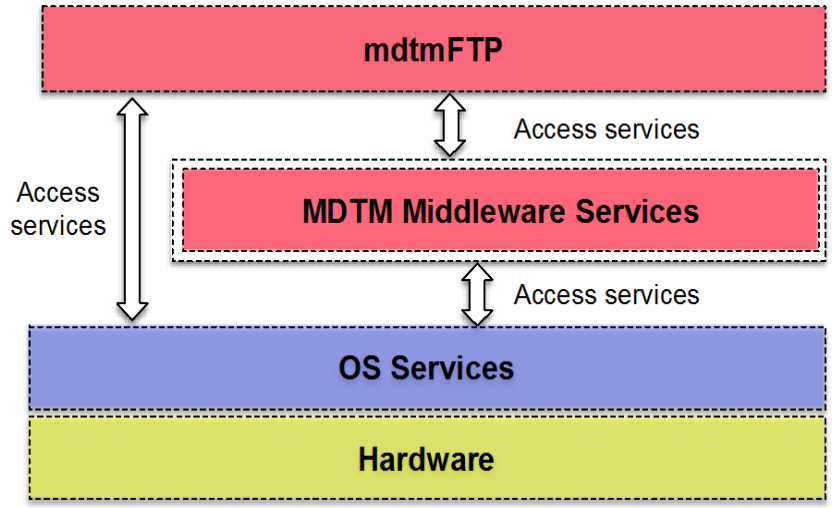


- Application-aware network service
  - On-demand programming
- Fast-provisioning of end-to-end network paths with guaranteed QoS
- Distributed resource negotiation & brokering

mdtmFTP

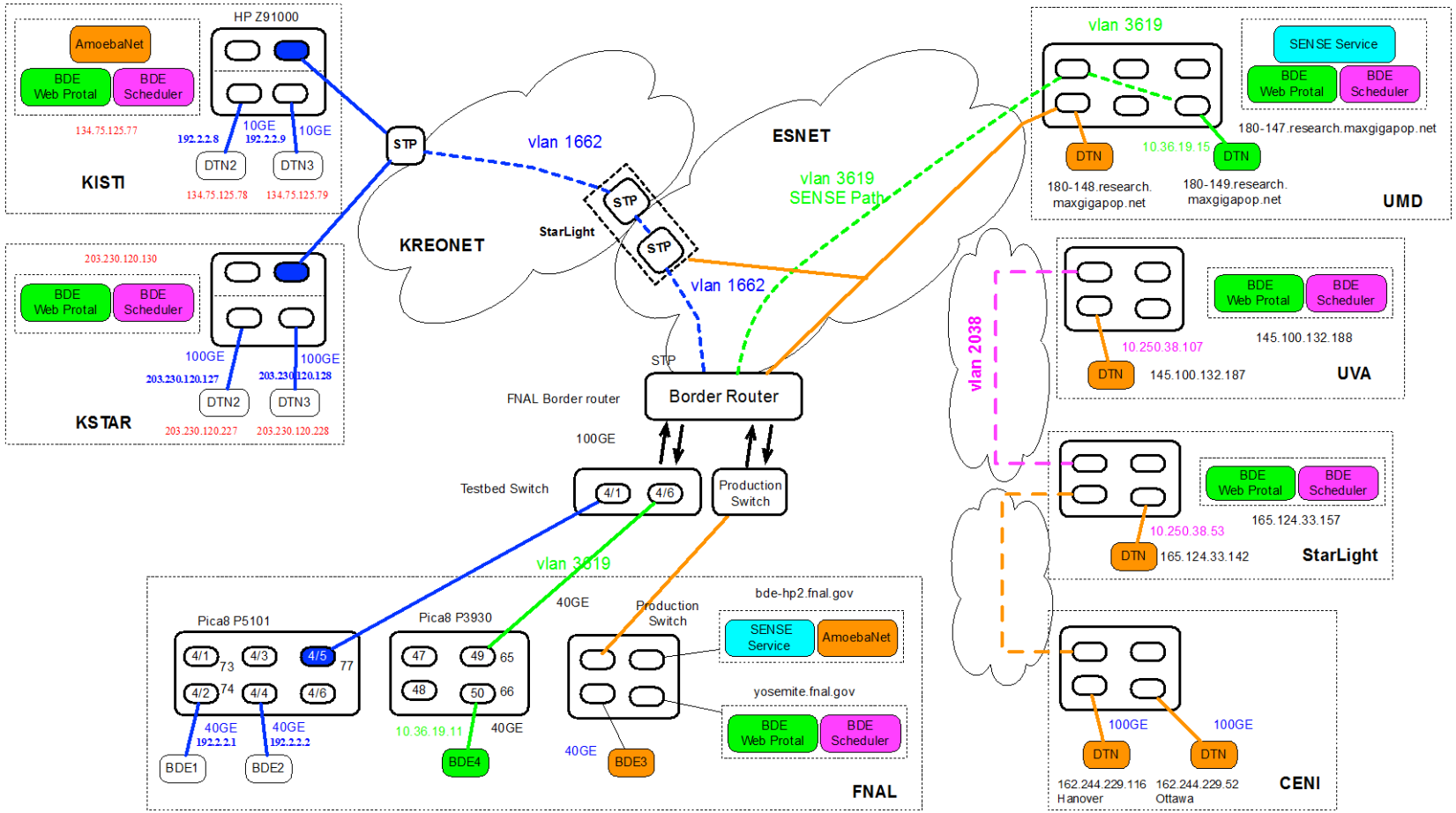
A high-performance data transfer tool

- Pipelined I/O-centric design to streamline data transfer
- Multicore-aware data transfer middleware (MDTM) optimizes use of underlying multicore system
- Extremely efficient in transferring of Lots Of Small Files
- Various optimization mechanisms
  - Zero copy
  - Asynchronous I/O
  - Batch processing



A DOE/SC/ASCR-sponsored research project
Software is available at: http://mdtm.fnal.gov

BigData Express SC18 DEMO



BigData Express – Deployment

- **Asia**
  - KISTI, South Korea
  - KSTAR
- **Europe**
  - University of Amsterdam, Netherlands
- **North America**
  - Fermilab
  - StarLight, Northwestern University
  - UMD/MAX, University of Maryland, College Park
  - Ciena (Canada)
    - US East
    - CA East
- **Australia & Pacific areas**
  - National Computational Infrastructure (NCI)

BigData Express Release

- License
  - **Apache 2.0**
- Docker release
  - http://xxxxx
- Source code package release
  - Available upon request

BigData Express Documents

- BigData Express Admin Manual
  - https://bigdataexpress.fnal.gov/bde_install_manual/index.html
- BigData Express Design and Architecture
  - https://bigdataexpress.fnal.gov/bde_design_manual/index.html
- BigData Express Web Portal User Manual
  - https://bigdataexpress.fnal.gov/bde_portal_manual/index.html

BigData Express Roadmap

- REST APIs support for scientific workflows
- Support IPv6
- Support Kubernetes
  - Automating docker-based deployment

More information about BigData Express

http://bigdataexpress.fnal.gov

Contact: wenji@fnal.gov

This document was prepared by BigData Express using the resources of the Fermi National Accelerator Laboratory (Fermilab), a U.S. Department of Energy, Office of Science, HEP User Facility. Fermilab is managed by Fermi Research Alliance, LLC (FRA), acting under Contract No. DE-AC02-07CH11359.