Dynamic Network and Storage Resource Management

Alex Sim, Scientific Data Management Group
Computational Research Division, Lawrence Berkeley National Laboratory

U.S. DEPARTMENT OF ENERGY
Office of Science

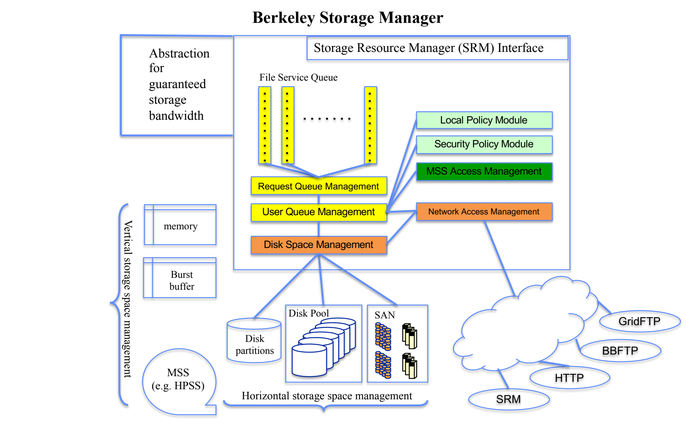
Helping scientists access data efficiently by developing resource management approaches

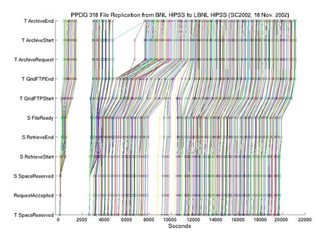
Dynamic Storage Resource Management

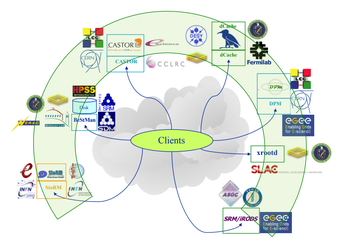
- Optimize data movement among vertical storage hierarchy and horizontal distributed storages
- Enable storage accessibility and improve science productivity
- Enable scientific collaborations with standard interface to the distributed storage resources - OGF GFD.129
- 43 BeStMan deployments worldwide and 5 backend deployments for CERN EOS system

Berkeley Storage Manager (BeStMan)

- Manages underlying storage space for files
- Supports multiple storage partitions and hierarchical storage spaces including HPSS
- Works on existing storages, and adaptable to special file systems and storages with customized plug-in
- Supports policy based quality of service
- Supports multiple transfer protocols for data distribution and load balancing for multiple transfer servers
- Supports co-scheduling with network bandwidth reservation
- Related products: Bulk Data Mover, DataMover-Lite, BeStMan Gateway, srm-lite, BASE library



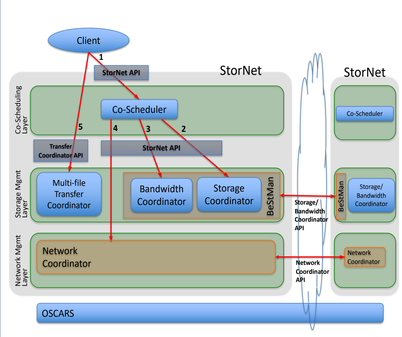


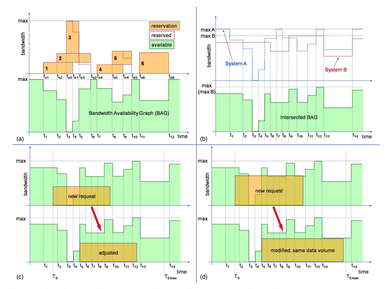


Collaboration with OSG, ESGF, FNAL, BNL, LLNL, ORNL, NCAR, SDSC, DESY, CERN, INFN, and many more

Co-scheduling Storage and Network Bandwidth

- Develop an integrated end-to-end resource provisioning system (StorNet) for high performance data transfers
- Improve resource utilization by co-scheduling network and storage resources; increase the reliability and predictability of data transfers while minimizing resource waste.
- Support negotiation for transfer completion timeline.
- Resource co-scheduling algorithm: U.S. Patent # 8,705,342





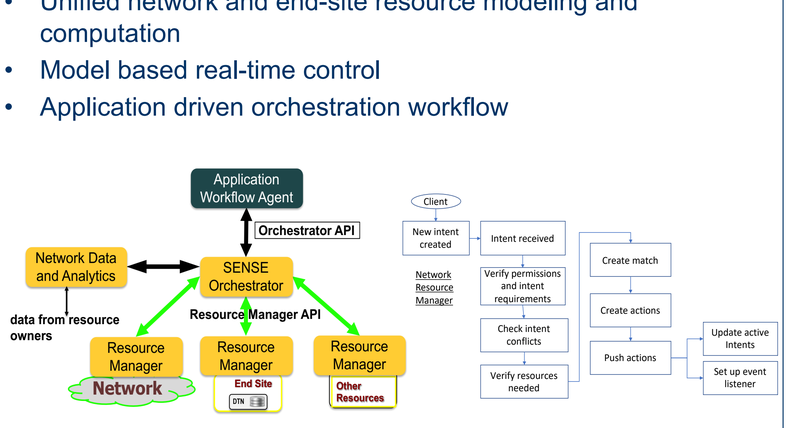
Bandwidth Availability Graph (BAG)
(a) reservations to BAGs; (b) BAG intersection; (c) fitting a request; (d) modifying and fitting a request

Dynamic Network Resource Management

- Optimize data movement with network resource orchestration, scheduling, and provision

SENSE: SDN for End-to-end Networked Science at the Exascale

- End-to-end model-based distributed resource reasoning and intelligent service orchestration
- Hierarchical service resource architecture
- Unified network and end-site resource modeling and computation
- Model based real-time control
- Application driven orchestration workflow



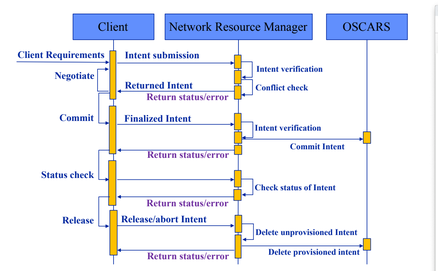
Network Resource Manager

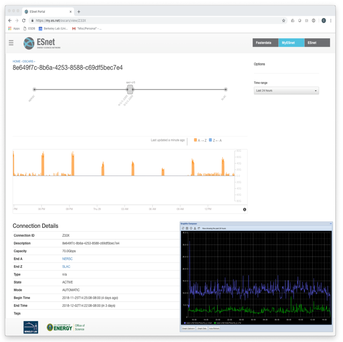
- Intent-based interface to accommodate network bandwidth scheduling requests
- Generate available topology model in real-time
- Evaluate and respond to SENSE Orchestrator information and service requests, including negotiation
- Provision network resources
- Provide status, monitoring, and debug functions

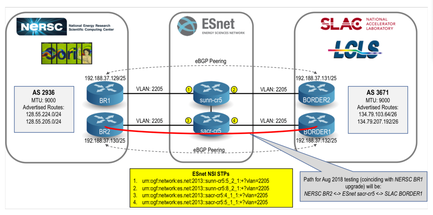
SENSE Use Cases

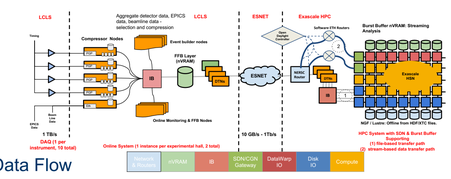
ExaFEL: Exascale for Free Electron Lasers

- Streaming data from the Linac Coherent Light Source (LCLS) at SLAC to storage resources at NERSC via network resources from ESnet
- Brings two (or more) DOE facilities together as one entity to ensure a workflow succeeds – superfacility
- Network RM functionalities
  - Resource Discovery Service
  - Data Plane Connectivity Service, including options for Quality of Service; scheduling; preemption; negotiation; and lifecycle monitoring and debugging
  - Real-time monitoring and debugging









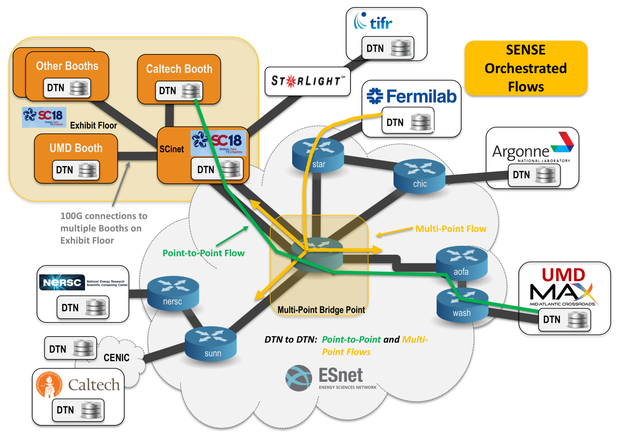
ExaFEL Data Flow

SC'18 Demo

- Data Transfer Node Priority Flow: Deterministic end-to-end data transfers
- Big Data Express: Intelligent selection of WAN paths based on user requirements

Network RM features

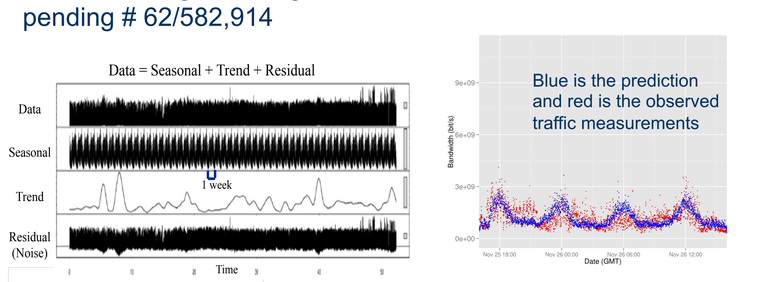
- Point-to-Point (Layer 2) network provision
- Multi-Point (Layer 2) network provision
- Quality of Service
- Negotiation
- Scheduling

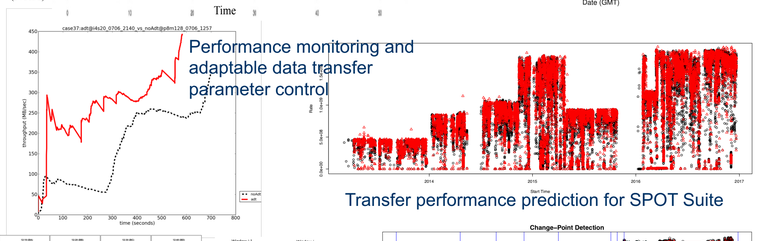


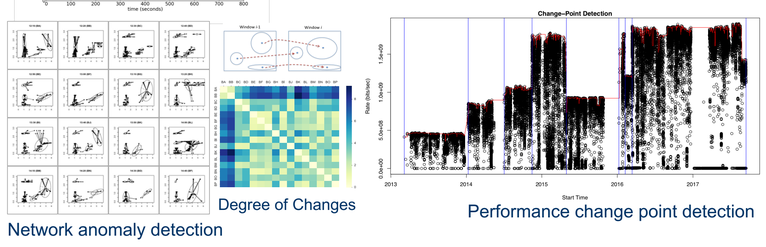
Lead PI: Inder Monga (Scientific Networking Division)
Involves ESnet, NERSC, FNAL, ANL, Caltech, Univ. of Maryland, and Mid-Atlantic Crossroads.
This poster only shows work at LBNL/ESnet.

Performance Monitoring and Prediction

- Network traffic performance analysis and prediction model
- Validated with the actual network traffic measurements
- Methods extended to other types of performance analysis, prediction, and anomaly detection
- Anomalous signal timing detection method - U.S. Patent pending # 62/582,914



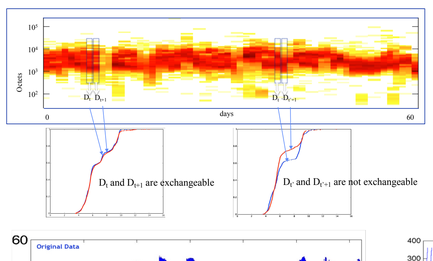


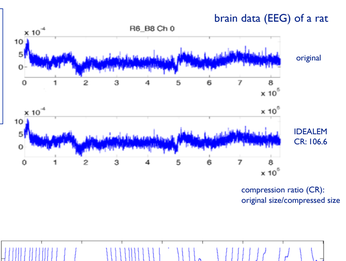


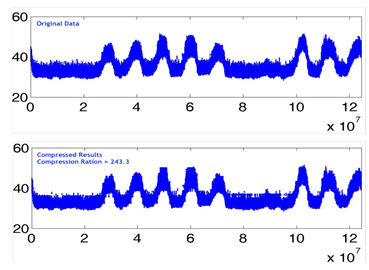
Data Reduction Technique for Storage

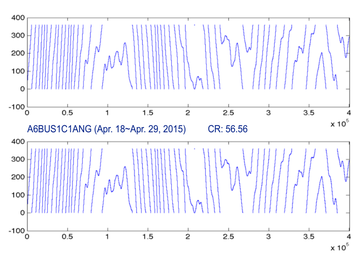
IDEALEM

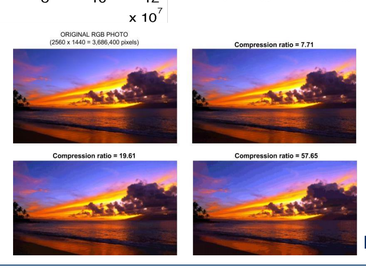
- Developed a data reduction method based on statistical similarity - U.S. Patent pending # 14/555,365
- Extended applications to network measurement data, power grid measurements, and others











IDEALEM to images

Summary

Dynamic network and storage resource management

- Enable performance aware intelligent resource management
- Increase data access efficiency
- Optimize data movement among vertical storage hierarchy and horizontal distributed storages
- Optimize data movement with network resource orchestration, scheduling and provision

Current Work

- Automated performance diagnostic method and failure detection method
- Machine learning based network data throughput performance prediction and anomaly detection
- Machine-learning based HPC performance anomaly detection and prediction
- Intent-based resource management
- Autonomous storage management

LBNL team: Suren Byna, Junmin Gu, Craig Tull, John Wu, Evangelos Chaniotakis, Chin Guok, John MacAuley, Inder Monga

Collab. Institutions: ANL, BNL, FNAL, LLNL, ORNL, SLAC, Caltech, NCAR, TAMUC, UCSD/SDSC, UMD, USC/ISI, CERN, DESY, KISTI, INFN, ESGF, LHC, OSG