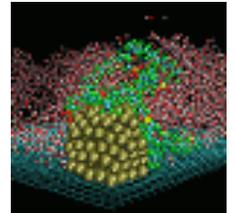
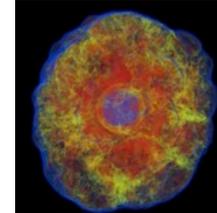
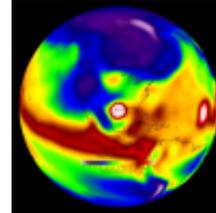
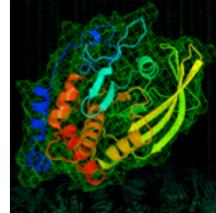
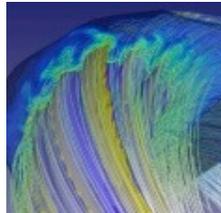
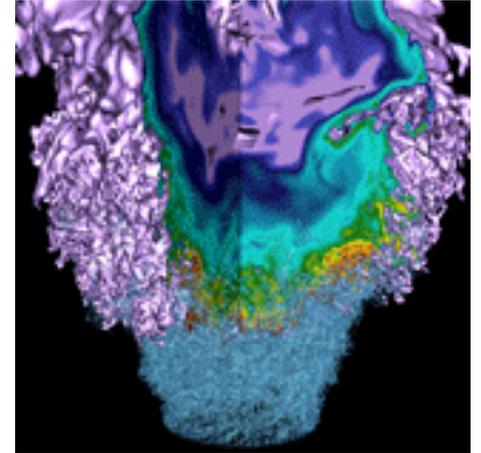


Lessons Learned From Running an HPSS Globus Endpoint



Lisa Gerhardt
Jeff Porter
Nick Balthaser
GlobusWorld
April 22, 2016

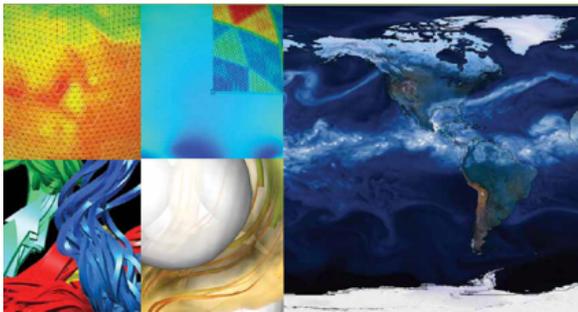
NERSC is the Production HPC & Data Facility for DOE



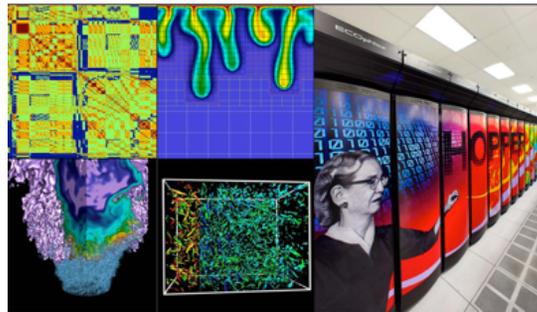
U.S. DEPARTMENT OF
ENERGY

Office of
Science

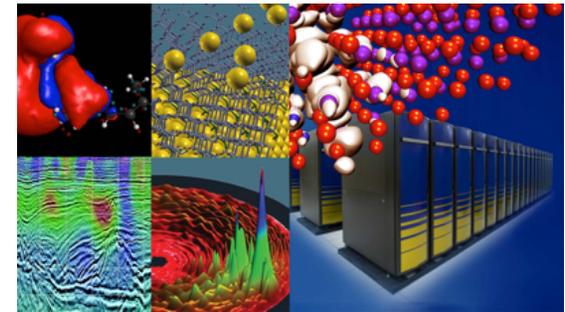
Largest funder of physical
science research in U.S.



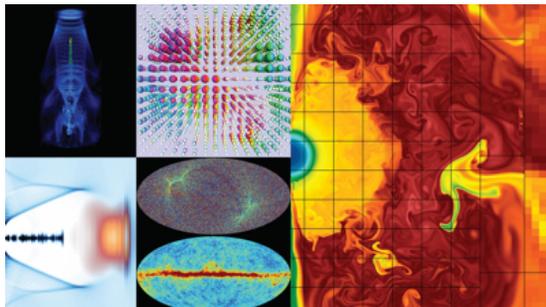
Biological and
Environmental Systems



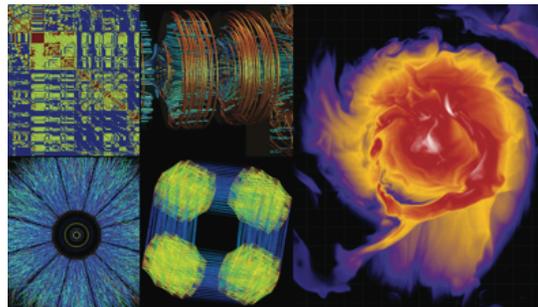
Applied Math, Exascale



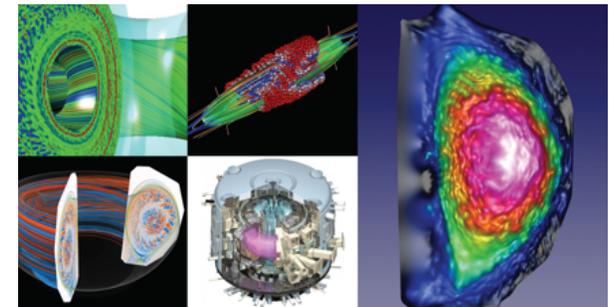
Materials, Chemistry,
Geophysics



Particle Physics,
Astrophysics

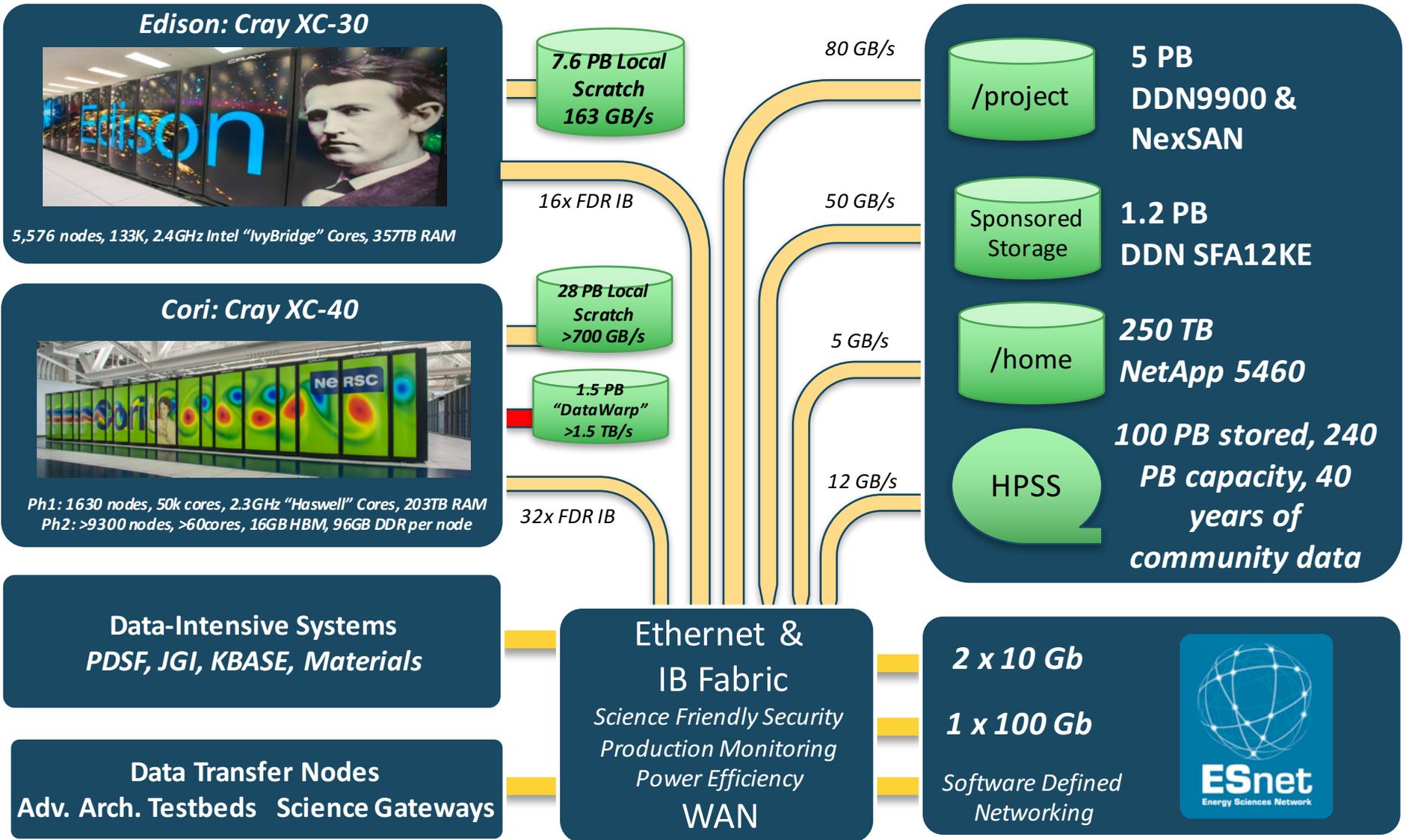


Nuclear Physics



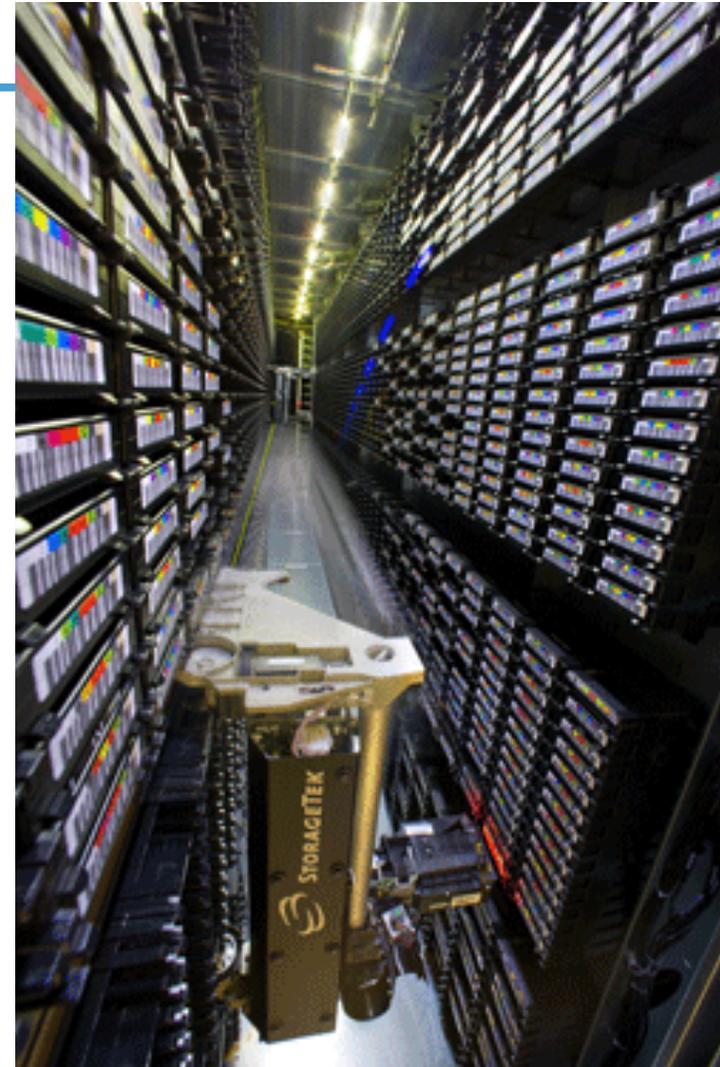
Fusion Energy,
Plasma Physics

NERSC - 2016



The HPSS Archive at NERSC

- The NERSC HPSS archive is a Hierarchical Storage Management system (HSM) that has been at NERSC since 1998
- Critical part of many groups data management and retention strategies
- Currently holds 100 PB of data
- Roughly 100 TB of I/O per day
- About 150 TB / month transferred via Globus, running since 2012
- Also use hsi, htar, pftp to transfer



Two Different Worlds



Globus: Data in motion, splits into multiple streams.
Goal: Get where you're going ASAP



HPSS: Data at rest, slow accrual over time.
Goal: Preserve important data indefinitely

Lesson 1: Death By A Thousand Cuts

- “Drag & Drop” transfers are a breeze with Globus’ GUI and users frequently use this to store or retrieve 1,000s – 10,000s of files simultaneously
- In HPSS files can end up scattered across hundreds of tapes, retrieving without ordering can cause major slow downs
- Using hsi, can order retrievals to get all files off of each tape in order



Facility Defense: Automated script

- **No Globus solution available, so devised a stop-gap**
- **Automated script that detects ingest of small files**
 - Mines HPSS logs to count files transferred by users in a sliding time threshold based on per day transfers
 - Automatically opens a user help ticket in our Service Now system alerting the user that they are using HPSS inefficiently
 - Depending on rate of transfer, can also ban user from further HPSS transfers, user must respond before access is restored

Lesson 2: Death By Repetition

- Interrupted transfers are automatically resumed and Globus tries to pick up file where it left off
- This causes file corruption in HPSS
- This feature is turned off by the Globus HPSS endpoint
- Get endless streams of retries as Globus tries to send a partial file and HPSS rejects it:

```
2016-01-17 03:55 pm    File system does not allow append to files    Hide Details
Error (transfer)
Endpoint: NERSC HPSS (9cd89cfd-6d04-11e5-ba46-22000b92c6ec)
Server: garchive.nersc.gov:2811
File: /~/AQ015_NSF_EaSM/NC_Nested/12km_wrfout/wrf_output_12km_A1B_2050_apr_22_30.tar.gz
Command: REST 0-17825792
Message: Non-zero offsets are not supported
---
500 Command not supported.\r\n

2016-01-17 03:54 pm    An operation timed out    Hide Details
Error (transfer)
Endpoint: NCAR GLADE (d33b3614-6d04-11e5-ba46-22000b92c6ec)
Server: gridftp01.ucar.edu:2811
Command: STOR ~/AQ015_NSF_EaSM/NC_Nested/12km_out/2050_A1B/wrf_output_12km_A1B_2050_apr_22_30.tar.gz
Message: The operation timed out
---
Timeout waiting for response
```

Facility Defense: Use Globus API to Automatically Block

- **Script queries Globus API to get number of failures per transfer**
- **If over threshold, uses Globus Administrator powers to pause transfer**
- **Gets NERSC user name from API and automatically opens a ticket so user can investigate the problem**
 - Still working on this last bit, Globus just implemented this functionality

Lesson 3: Don't Cross the Streams

- **Transfers via Globus to HPSS a factor of 2 – 5 slower than using hsi / htar to transfer**
 - Reason for this not well understood
 - Expect some slow down for the DSI layer to translate from gridFTP to HPSS, but not a factor of 2

A Two-Factor Solution

- **For large data transfers into HPSS, use Globus to transfer to intermediate spinning disk at NERSC**
- **Once there use hsi / htar to put data into HPSS**
- **Fastest, but very inconvenient**

Globus and HPSS Will Be Crucial

- Long term data retention is becoming a must for scientists
- The need for a smooth connection between HPSS and Globus will be crucial
- Fruitful collaboration with Globus to develop and improve the Globus / HPSS interface, hope it can continue