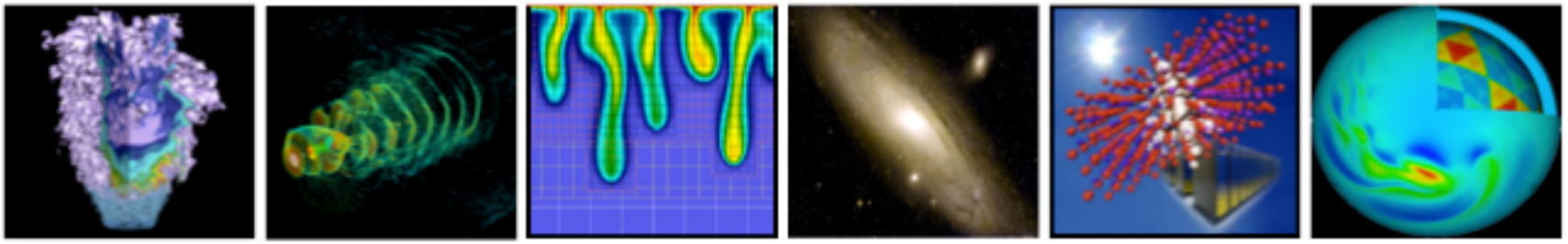


NERSC Global File System Futures

Greg Butler
Storage Systems Group

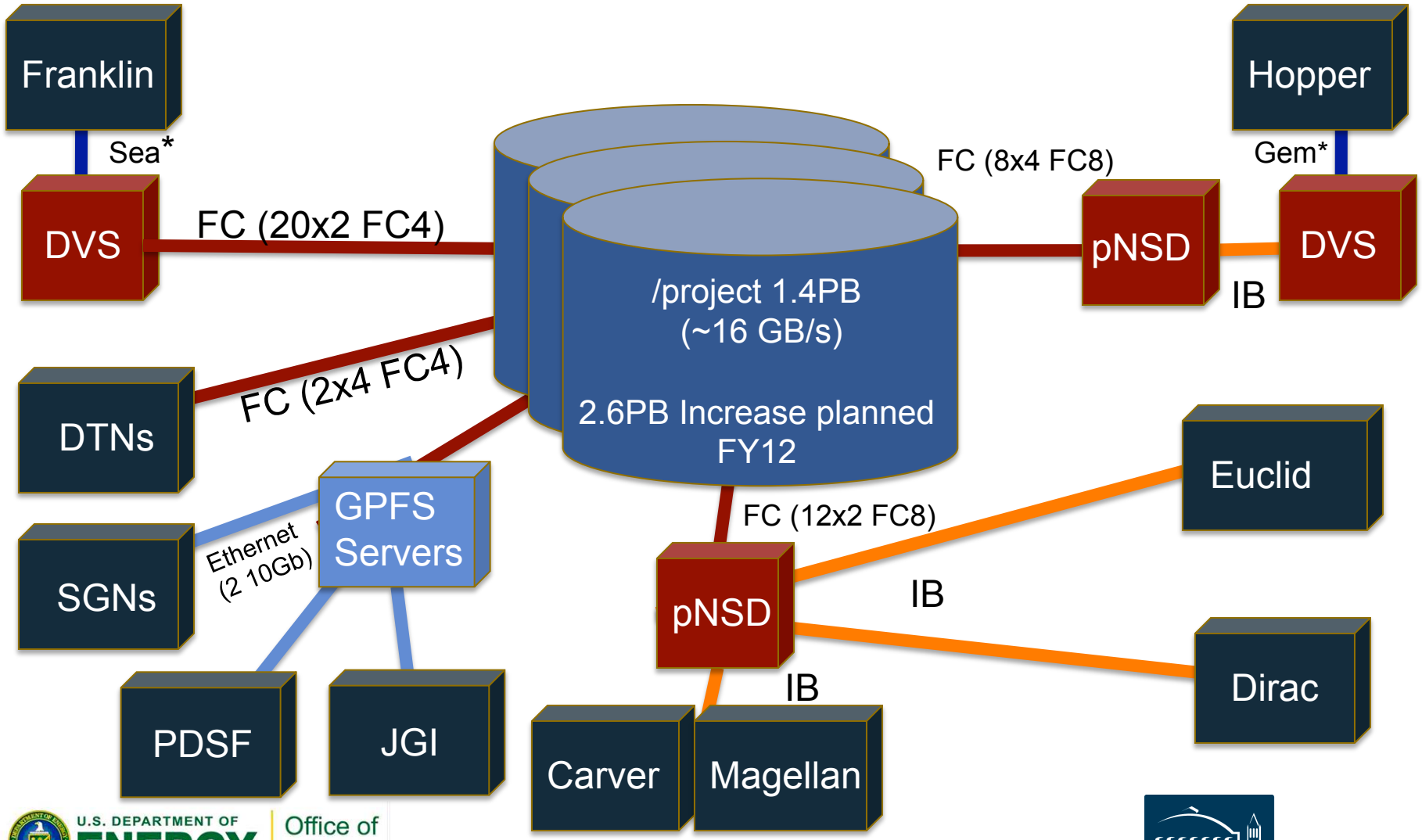
NUG April 26, 2012

U.S. DEPARTMENT OF ENERGY | Office of Science
NERSC National Energy Research Scientific Computing Center
BERKELEY LAB Lawrence Berkeley National Laboratory

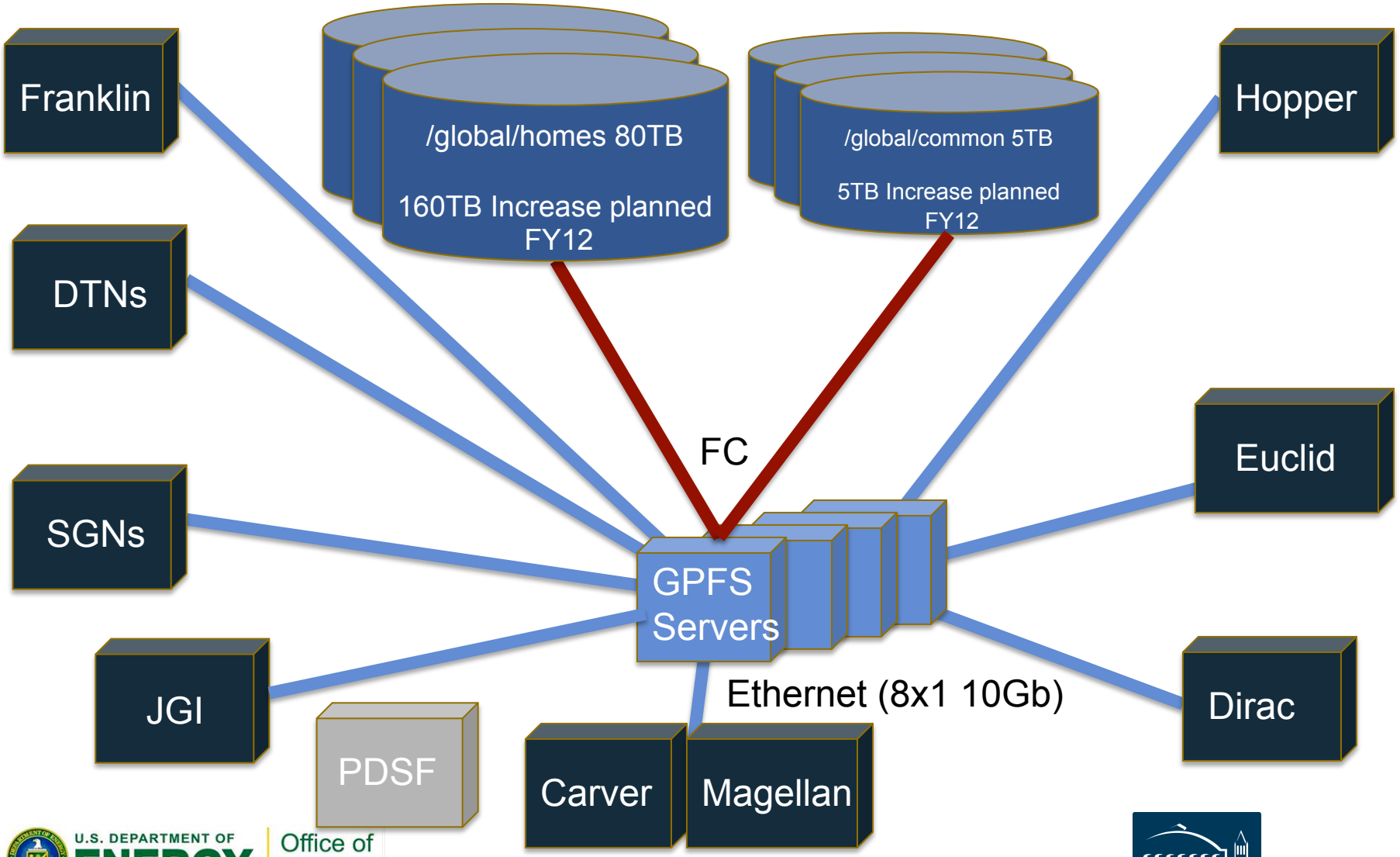
The NERSC Global File Systems

- **/global/project is for sharing and long-term residence of data on all NERSC computational systems.**
  - Not purged, quota enforced (4TB default per project), projects under 5TB backed up daily
  - Serves 280 projects (from 200 in 2010) over FC8 primarily, 10Gb Ethernet alternatively
  - Bandwidth is about 16GB/s, with 1.4PB total capacity, increasing to ~4PB total in 2012
- **/global/homes provides a common login environment for users across systems.**
  - Not purged but archived, quota enforced (40GB per user), backed up daily
  - Served 7000+ users (from 4000 users in 2010) over 10Gb Ethernet
  - 80TB total capacity, increasing to ~240TB total in 2012
- **/global/common provides a common installed software environment across systems.**
  - 5TB total capacity, increasing to ~10TB total in 2012
  - Provides software packages common across platforms
- **/global/scratch provides high bandwidth and capacity data across systems.**
  - Purged, quota enforced (20TB per user), not backed up
  - Serves 7000+ users over FC8 primarily, 10Gb Ethernet alternatively
  - Bandwidth is about 16GB/s, with 1.1PB total capacity, looking to increase to several PB total in 2012
- **/global/projectb is a new file system for JGI users.**
  - Contains both project space and scratch space over 10Gb Ethernet (for Genepool) and IB (for Carver)
  - /global/projectb/scratch: purged, quota enforced (20TB per user), not backed up
  - /glboal/projectb/projectdirs: not purged, quota enforced (5 TB per user), to be backed up soon
  - Bandwidth is about 15GB/s, with 2.6PB total capacity

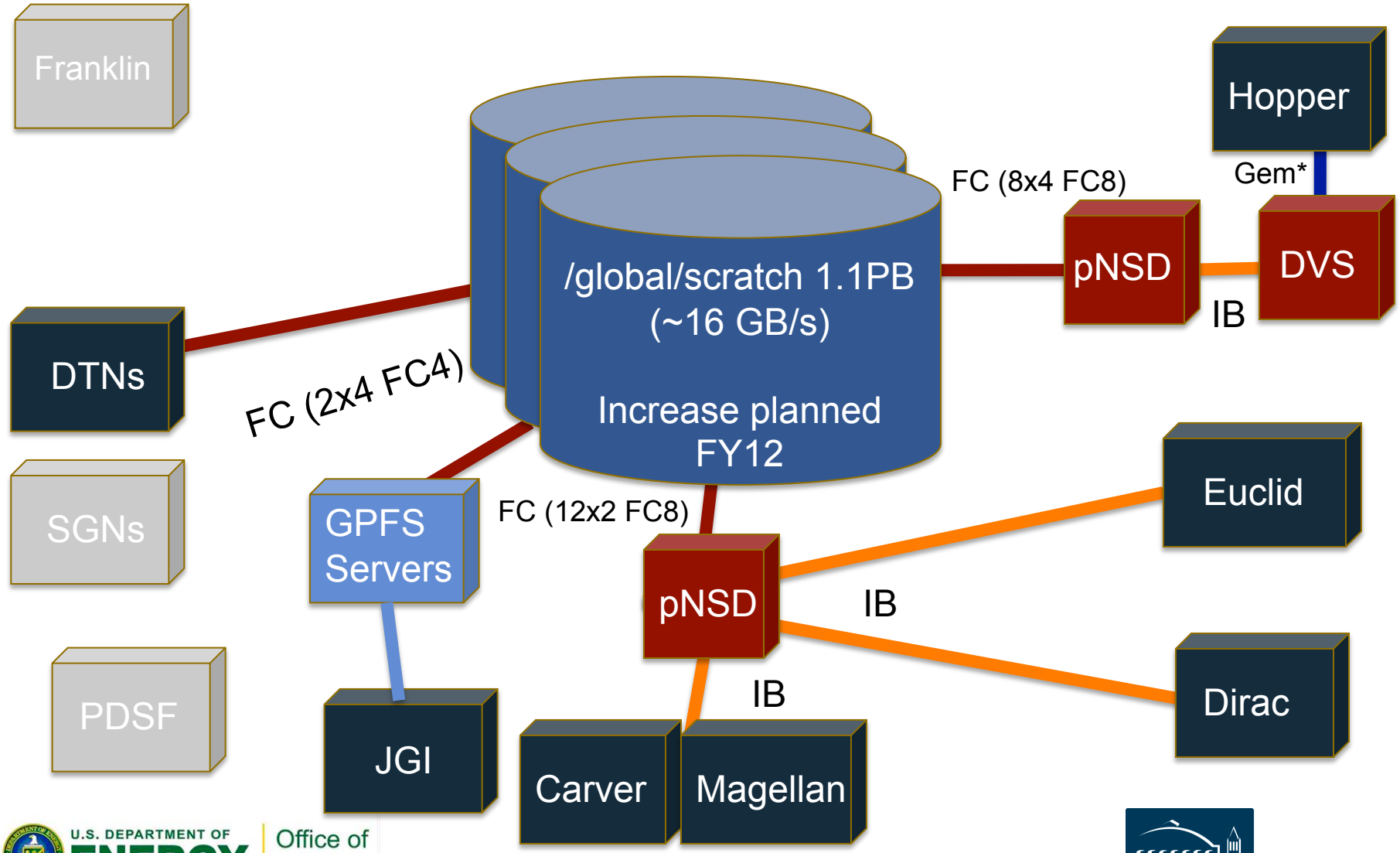
NGF Project (/global/project)



NGF Global Homes & Common



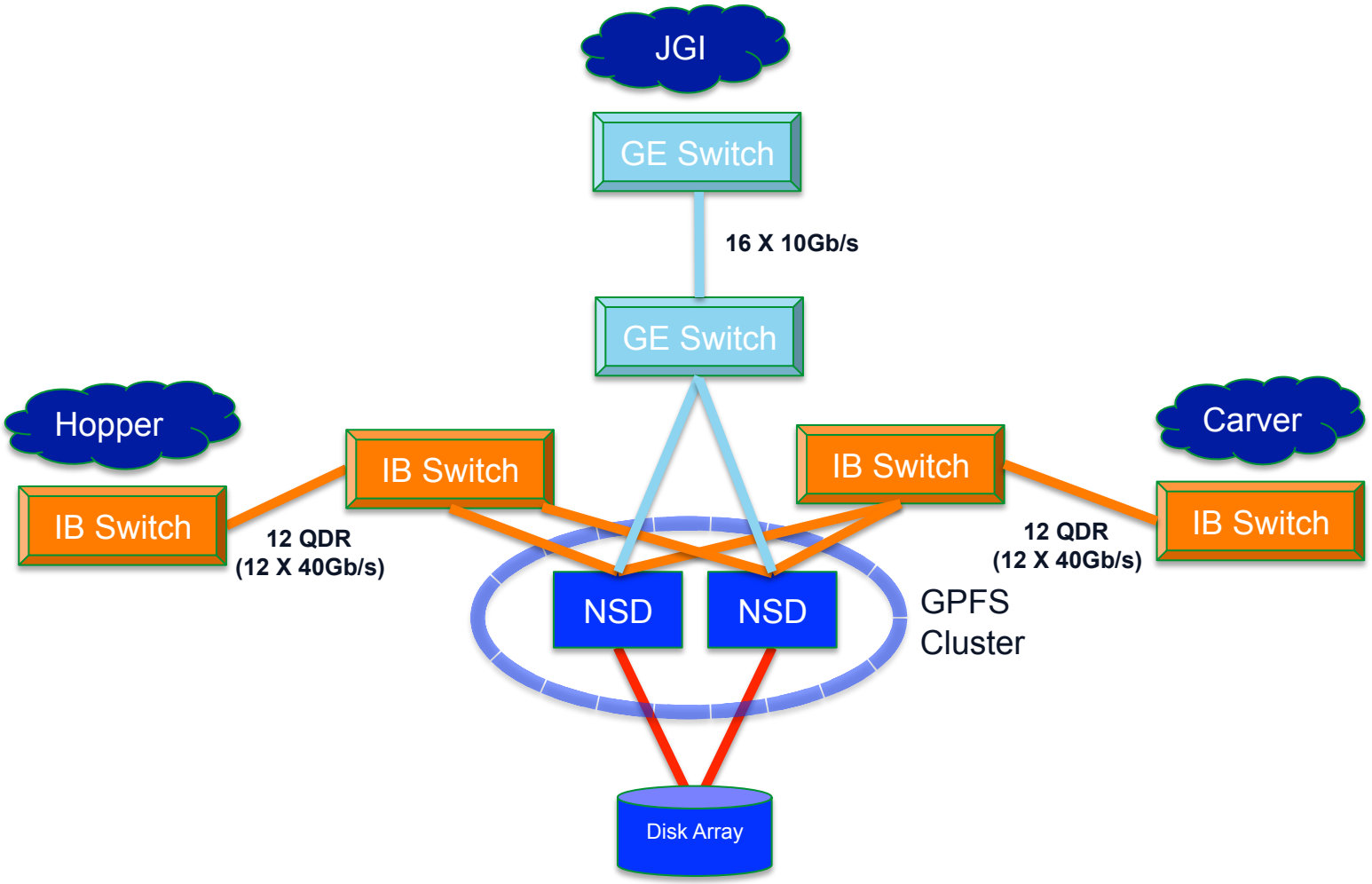
NGF Global Scratch



New NGF connectivity

- **Moving to IB/10GE based client connections to the file systems**
  - Connect storage directly via FC to GPFS NSD servers
  - Connect GPFS servers to each cluster's IB or 10GE network, if it doesn't have one, create a small IB network to scale with cluster's proprietary network (Seastar, Gemini, ...)
  - Eventual goal to eliminate the FC SAN
- **Move from pNSD deployment to NSD deployment**
  - Scalability of server and data network resources

NGF /global/projectb



NGF Stability

- **NERSC has seen a number of NGF problems lately. These can be categorized into 4 broad classes.**
  - Zombies
  - Storage hardware failures
  - Network communication issues
  - Users running near their quotas

Zombies

- **We've seen numerous conditions occur that can cause GPFS on a host to be unable to properly participate in the GPFS protocol.**
  - OOM – anything that kills a single GPFS thread/process may leave GPFS unable to complete pending operations, and release locks. Leading to GPFS hangs across the center.
  - Kernel bugs – we recently hit a bug which caused the linux scheduler to break after 208+ days of system uptime.(fixed in 2.6.32.50+)
  - Repeated Packet loss – this is more of a network infrastructure issue, but we've seen conditions occur where some but not all packets over a particular link, or switch are dropped. In this case, gpfs can get stuck in a live-lock condition where GPFS thinks the affected nodes are alive, and participating in the GPFS protocol, but never succeed in operations that require passing an un-passable packet(dropping jumbo frames is one example)

Storage Failures

- **Storage hardware failures – (multiple failures in redundant disk subsystems)**
  - The NGF implementation can tolerate any single point of failure but multiple failures could cause file system outages
  - We have seen an exceptional number of multiple hardware failures recently that affect the file system availability
  - We are actively working with our vendor to identify the cause(s) of these failures, and address them.

Network Failures

- **Network communication issues**
  - We have identified several issues with one of the models of 10 GE Ethernet switches used at NERSC that in combination with a misbehaving host (network driver bug, kernel oops, 208-day bug, etc.) can cause the switch to stop passing traffic even for other hosts.
  - We have worked with the network driver vendor to pinpoint the condition that have caused the switch to hang. This issue has been greatly reduced since we installed a new driver that the vendor has released to address this issue.
  - We are planning to replace these switches to alleviate other issues, and performance problems due to packet loss seen on these switches under congestion conditions.

Quotas

- **Users who are operating near their GPFS quota can cause excessive GPFS communication, and locking**
  - This manifests as hangs when performing quota commands such as myquota, and general filesystem slowness
  - This will likely cause an extreme drop in performance for the user running near their quota, as well as impacting performance of other users to a (usually) lesser degree.
  - We believe we have a workaround for this problem, and are in the process of getting ready to roll it out.

NGF Near-Term Plan

- Rebalance the NGF 10GE switches across 2 center routers (i10g and i10g-2) for higher network performance
- Mount /global/projectb on Carver (DONE) and Hopper via IB
- Expand /global/project by 2.6 PB, 15 GB/s peak
- Replace NGF servers for homes, global common, and sys commons with faster servers
- Upgrade homes with higher capacity and performance storage (by 120 TB, 3 GB/s per home file system)
- Replace NGF 10GE switches and add more uplinks for higher network performance and better reliability
- Upgrade NGF to GPFS PTF22 to fix a few outstanding issues
- Implement a new quota strategy to minimize file system trashing when users get close to their quota limit

NGF May 8 Outage Plan

- **All NGF file systems will be down except /project (all NERSC systems except PDSF will go down at 8 AM on 5/8)**
- **Offline fsck of home1 to clear FSSTRUCT errors per IBM recommendation**
- **Offline fsck of other file systems as a proactive maintenance to check for file system errors**
- **Move GSCRATCH DDN storage from Carver area to NGF area**
  - to simplify the SAN architecture
  - to reduce # of disk failures possibly caused by vibration
- **GPFS configuration change to reduce file system access "hangs" during NGF server restart**