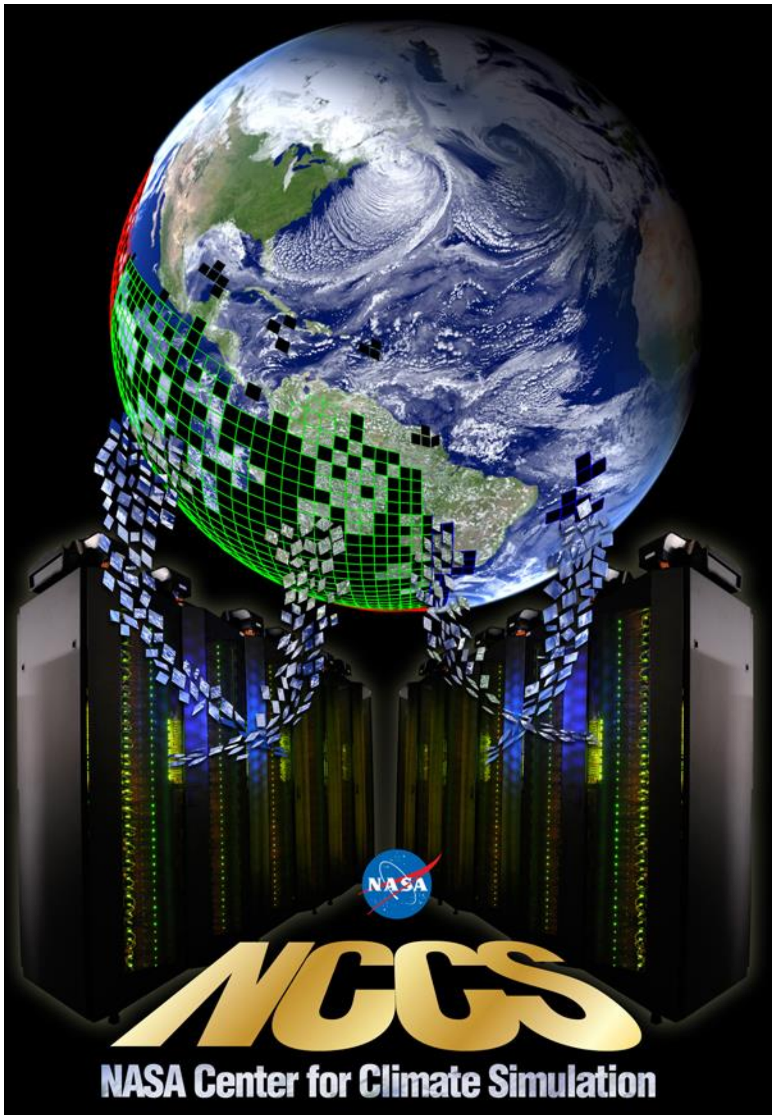


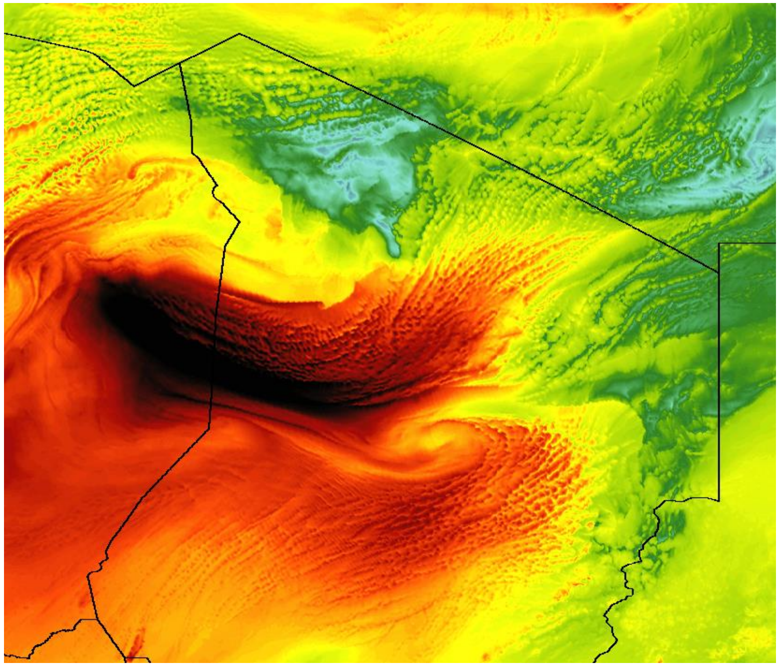
NASA Center for Climate Simulation: Can Disk Replace Tape?

31st International Conference on Massive Storage Systems and Technology (MSST 2015)
NASA Center for Climate Simulation (NCCS)

Ellen.Salmon@nasa.gov
High Performance Computing, Code 606.2
NASA Goddard Space Flight Center
Greenbelt, MD 20771

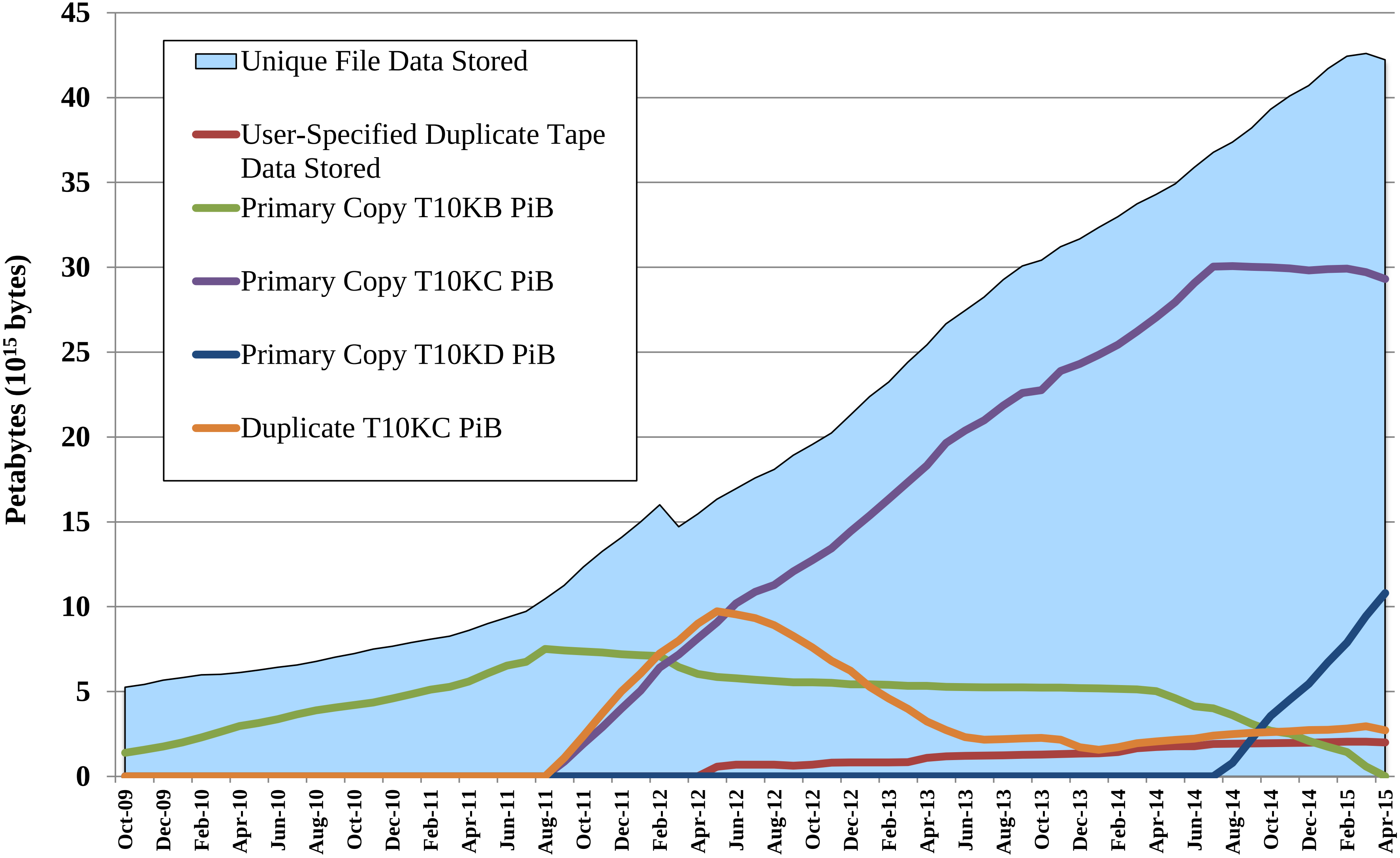
NASA Center for Climate Simulation (NCCS)

- Funded by NASA's Science Mission Directorate, located at Goddard Space Flight Center in Greenbelt, Maryland.
- Provide an integrated high-end computing environment designed to support the specialized requirements of Climate and Weather modeling.
  - State-of-the-art high-performance computing, data storage, and networking technologies
  - Advanced analysis and visualization environments
  - High-speed access to petabytes of Earth Science data
  - Collaborative data sharing and publication services.



Detail of the Total Aerosol Optical Depth for a dust storm over Chad from the 1.5-km GEOS-5 global simulation for 1200 GMT 15 June 2012 (forecast hour 15). Image source: William Putman/GMAO

NCCS Mass Storage by Media Type
October 2009 – April 2015



Some Thoughts and Observations

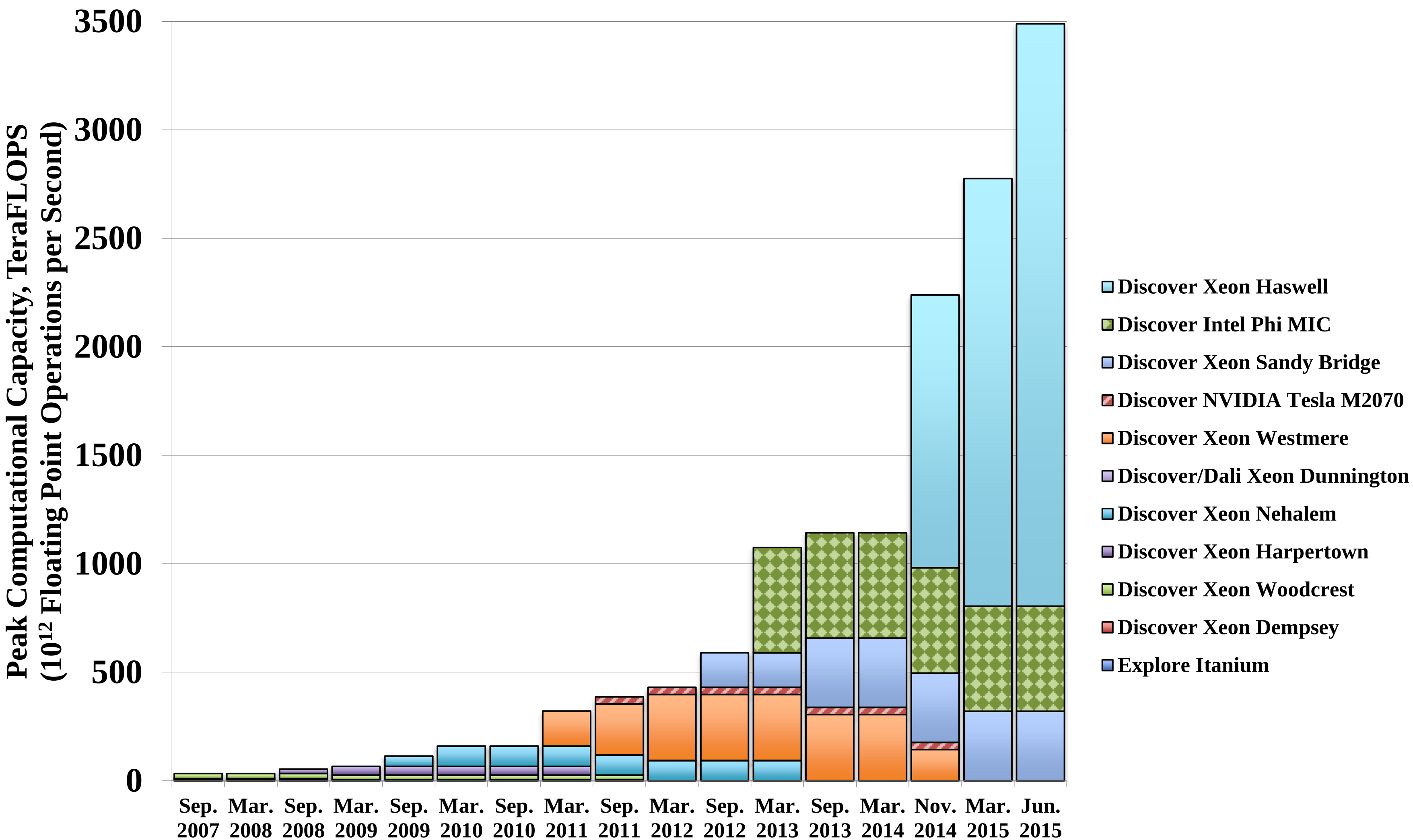
- At NCCS, HPC “nobackup” disk storage capacity (~33 PB) is approaching that of current data stored in NCCS tape-based mass storage system (~45 PB).
- “Disk storage attached to HPC should be actively used data.” Still true?
- Robotic tape-based mass storage can more quickly accommodate 1 Petabyte of new data – assuming a tape library, transports and media are already available.
  - Question: cost to provide headroom, disk vs. tape?
- We have a bimodal age distribution of mass storage files read (peaks for newest files, and those created more than a year ago) – but ~75% of files are never read.

Mass Storage Challenges, Regardless of Media

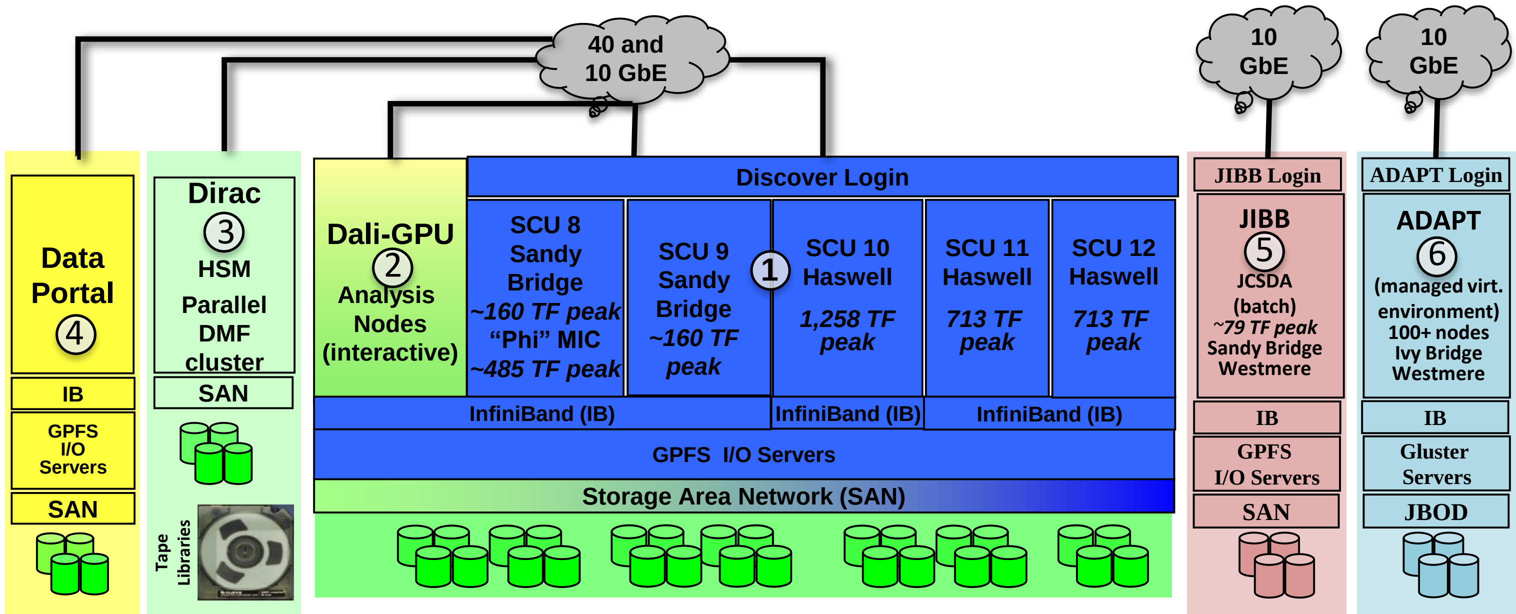
- How to help users figure out what they no longer need?
  - Cost/benefit of "just buying more storage" vs. users' and system administrators' time to find/remove unneeded data.
- How to implement low-impact movement of data from older media to new media.
- How to dealing with hardware/software failure modes at scale
  - Power/cooling: "Disk-B-Ques"
  - NFS complications with filesystem consistency
  - Media transport issues
- In current POSIX-based filesystems, growth in number of files and directories (not just data volume) is a significant challenge.

Supporting Material

NASA Center for Climate Simulation
High Performance Computing Capacity Evolution



Notional NCCS HPC Platform Summary



① *Discover* Linux Supercomputer

- Intel Xeon nodes
  - 3,264 nodes
  - 79,872 cores
  - Peak 3 PFLOPS (gen. purpose)
  - 341 TB memory
- Coprocessors:
  - Intel Phi MIC
    - 480 units
    - ~485 TFLOPS
- Shared disk: 33 PB

② *Dali-GPU* Analysis

- 12 12-core nodes
- 16 GB memory per core
- Dali-GPU has NVIDIA GPUs

③ *Mass Storage System*

- 4.4 PB disk
- ~70 PB robotic tape library
- Data Management Facility (DMF) storage management

④ *Dataportal* Data Sharing Services

- Earth System Grid
- OPeNDAP, THREDDS
- Data download: http, https, ftp
- Web Mapping Services (WMS) server

⑤ *JIBB*

- Supports Joint Center for Satellite Data Assimilation community
- 288 Intel Xeon Westmere and 120 Sandy Bridge nodes

⑥ *ADAPT* Advanced Data Analytics Platform

- Managed Virtual Machine Environment
- Intel Xeon Westmere and Ivy Bridge nodes
- Virtualized InfiniBand to Virtual Machines

Questions to Answer

Abstract: Due to the wide-spread use of disks in massive-scale systems, they are de-facto replacing tape for long-term archives in some installations. In this panel, participants will discuss power consumption, space, migration and tape/disk archive management challenges as disks strive for dominance in long-term storage and archival applications.

1) Power consumption: spinning disks eat up a lot of power relative to tapes. Can disk spin down reduce power adequately to make them effective as a long-term archive medium? What other mechanisms could be used to reduce disk power consumption? [Can’t comment.]
2) Space: disk volumetric efficiency is lower than tape's. Can techniques such as compression and very high disks-per-unit-volume ratios (a la Copan and Backblaze) make up some of the difference. Does this really matter in today's large data centers?
3) Migration: how will data be migrated from older generation, lower-density drives to newer, higher-density drives, and is disk better than tape for these kinds of migrations? If so, why? As the disk bandwidth-to-capacity ratio continues to decline, will disks run up against the same low-bandwidth-to-capacity issues tape faced?
4) Tape challenges: what tape issues are causing users to reconsider it as the long-term archive medium of choice? Management complexity?Too few vendors building the technology? Migration challenges of large archives due to low bandwidth-to-capacity ratios? Or is tape fine for another decade and more? What applications still absolutely require tape or are a good fit for tape?