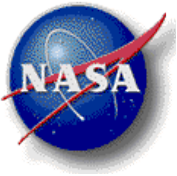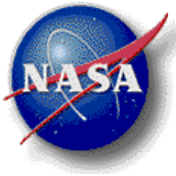# NCCS User Forum

June 19, 2012

Agenda

- Introduction

- Discover Updates

- NCCS Operations & User Services Updates

- Question & Answer

- Breakout Session:

   – Climate Data Analysis using Vistrails and UV-CDAT (Tom Maxwell)

Accomplishments

- Purchase Order for Discover SCU8 Augmentation

  – Sandy Bridge: additional ~160 TFLOPs, late summer

  – Many Integrated Core (MIC): Fall 2012

- ~4 Petabytes (usable) Discover disk arriving late June

- Five NCCS Brown Bag seminars so far, more to come

- Dirac Archive

  – Completed "High Availability" server implementation

  – Implemented single-tape-copy default

- New Earth System Grid Federation (ESGF) Node, serving:

  – NASA's IPCC AR5 climate simulation contributions

  – Obs4MIPS – observations formatted for use in climate/ocean/weather model intercomparison studies (e.g., CERES EBAF, TRMM, …)
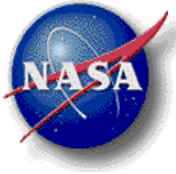
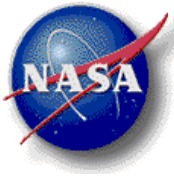**NCCS HPC Capacity Evolution**
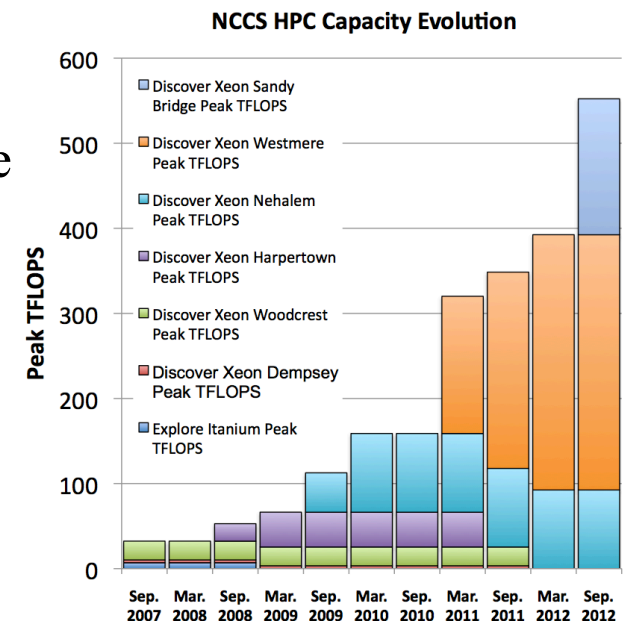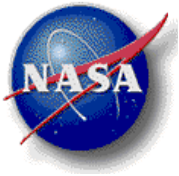
New Employee

- Al Settell, CSC Program Manager, NCCS Support, alan.d.settell@nasa.gov

Discover Update

Dan Duffy, NCCS Lead Architect

Discover SCU8 – Intel Xeon Sandy Bridge Nodes

- Adds ~160 TFLOPs (40%) to Discover's capacity

- 480 IBM iDataPlex Nodes, each configured with

  - Dual Intel Sandy Bridge 2.6 GHz processors (E5-2670) 20 MB Cache

  - 16 cores per node (8 cores per socket)

  - 32 GB of RAM (maintain ratio of 2 GB/core)

  - 8 floating point operations per clock cycle

  - Quad Data Rate InfiniBand

- Advanced Vector Extensions (AVX)

  - New instruction set (http://software.intel.com/en-us/avx/)

  - Just have to recompile

- Still working on actual SBU charge for Sandy Bridge: 1.5 – 1.8?

  - SBU of 1 is based on the Westmere node (12 cores per node, 4 floating point operations per clock)

*SCU8 Sandy Bridge hardware arriving ~late August 2012*
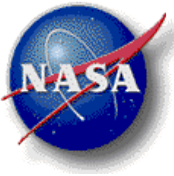
Discover SCU8 – Many Integrated Cores (MIC)

- In addition to the Sandy Bridge cores, the NCCS will be integrating 240 Intel MIC Processors

  Intel® MIC Architecture:
  An Intel Co-Processor Architecture

  - ~1 TF per processor
  - PCI-E Gen3 connected
  - To start: 1 per every other Sandy Bridge node

- How do you program for the MIC?
  - Full suite of Intel Compilers
  - Doris Pan and Hamid Oloso have access to a prototype version and have developed experience over the past 6 months or so
  - Run in two different types of modes: Offload versus Native
  - Better performance when running native mode
    - Typically one would run at least 1 MPI process in native mode and spawn OpenMP threads
  - NCCS/SSSO will host brown bags and training sessions soon!
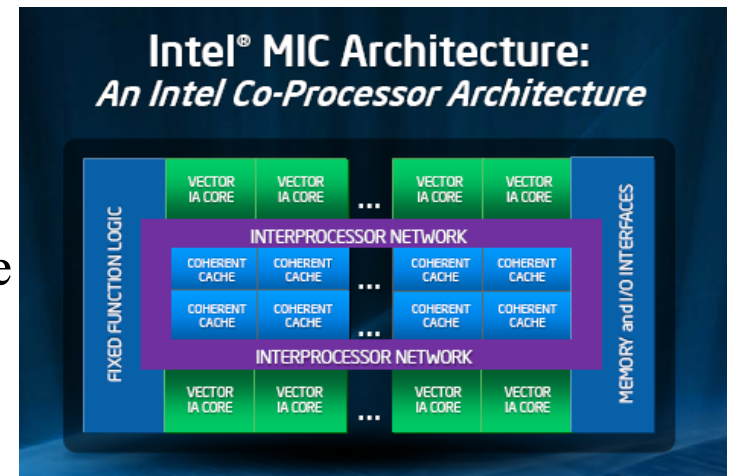
SCU8 MIC hardware arriving Fall 2012
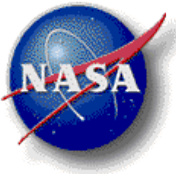
Discover: Large "nobackup" augmentation

- Discover NOBACKUP Disk Expansion
  - 5.4 Petabytes RAW (about 4 Petabytes usable)
  - Doubles the disk capacity in Discover NOBACKUP
  - NetApp 5400
    - http://www.netapp.com/us/products/storage-systems/e5400/
    - 3 racks and 6 controller pairs (2 per rack)
    - 1,800 by 3 TB disk drives (near line SAS)
    - 48 by 8 GB FC connections
    - ~18 GB/sec total bandwidth

- Architecture is better suited for smaller I/O operations (higher IOPS)

  - This is compared to our existing DDN arrays which are designed for large streaming I/O

- Delivered this week; integrated by next month

- Also focusing on improvements to the metadata subsystem

  - Hopefully will have some progress on that by the next user forum

NASA Center for Climate Simulation

NCCS Operations & User Services Update

Ellen Salmon
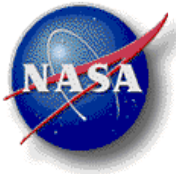
# Coming Discover Changes: Discover PBS 11 and Linux SLES 11 SP1

- Reason for upgrades:
  - PBS 11 has bugfix for very long startup of PBS system that occurs following Discover maintenance downtimes
  - SLES 11 SP1 is required to maintain current Linux security patches
- PBS 11: No user changes needed, per vendor Altair
  - Upgrade very soon, possibly during June 27 Discover maint.
- SLES 11 SP1:
  - NCCS staff testing & documenting the few needed user changes (due to updated libraries and Linux kernel)
  - Planning for phased, rolling deployments with minimal downtime

NASA Center for Climate Simulation
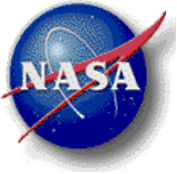
# Dirac / Archive Update

- Single-tape-copy default for archive implemented 31 May 2012.

- Use 'dmtag –t 2 <archive_file>' to request 2 tape copies of your critical, hard-to-replace archive files.

  - E.g., input files from many disparate sources, results published in journals, important collections of files…

- To date, four NCCS users have used dmtag to request 2 tape copies, for about 5% of total archive file data (mostly MERRA-related).

- NCCS staff is closely tracking and analyzing tape-related errors.
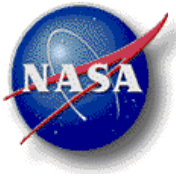
NCCS Brown Bag Seminars

- Twice a month in GSFC Building 33 (as available)
- Content is available on the NCCS web site following seminar
- Upcoming topics:
  - Monitoring PBS Jobs and Memory
  - Scientific Computing with Python
  - Programming with Intel MIC (coming in Fall 2012)
  - ….
- Suggest topics of interest on today's feedback or signup sheet, or via email to support@nccs.nasa.gov
- *We will repeat seminars upon request.*
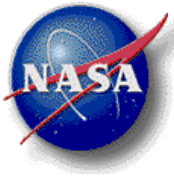
Twice-a-Month NCCS Brown Bag Seminars: Delivered & Proposed Topics

- ✓ Intel Many Integrated Core (MIC) Prototype Experiences

- ✓ Using Totalview on Discover, Parts 1 & 2

- ✓ NCCS Archive Usage & Best Practices Tips, & dmtag

- Monitoring PBS Jobs and Memory

- Introduction to Using Matlab with GPUs

- Using the TAU Profiling Tool

- Best Practices for Using Matlab with GPUs

- Introduction to the NCCS Discover Environment

- Scientific Computing with Python

- Using CUDA with NVIDIA GPUs

- Using OpenCL

- Using GNU Octave
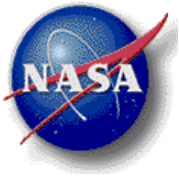
- Using Database Filesystems for Many Small Files
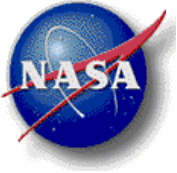
Questions & Answers

NCCS User Services:
support@nccs.nasa.gov
301-286-9120

https://www.nccs.nasa.gov
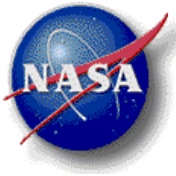
Contact Information

NCCS User Services:

support@nccs.nasa.gov

301-286-9120

https://www.nccs.nasa.gov

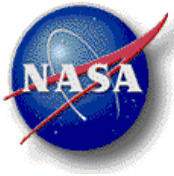*Thank you*

Supporting Slides –
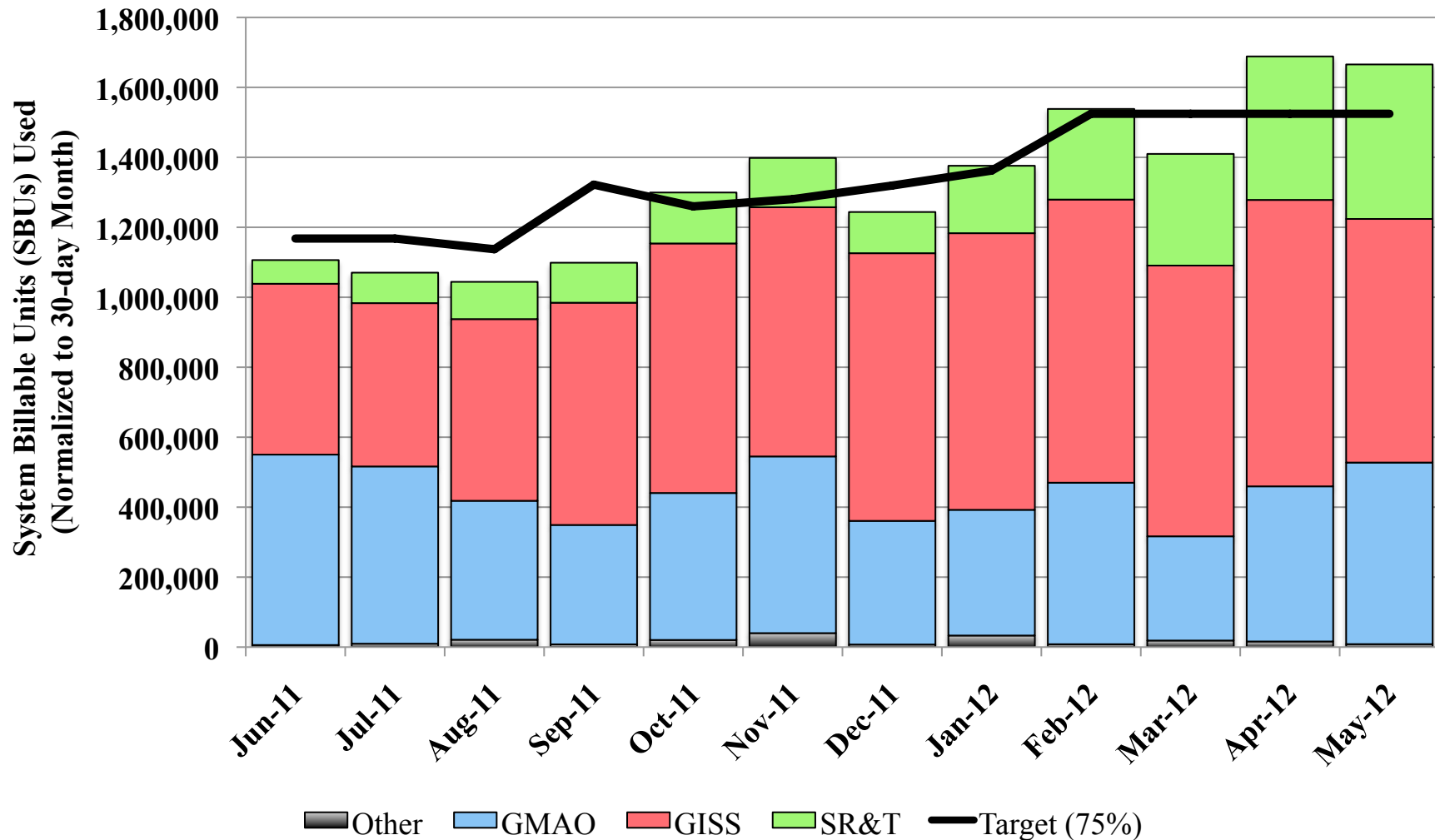# Discover & Dirac / Archive

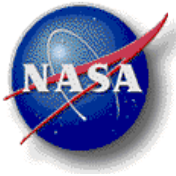NASA Center for Climate Simulation
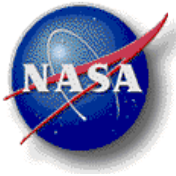
NCCS HPC Capacity Evolution
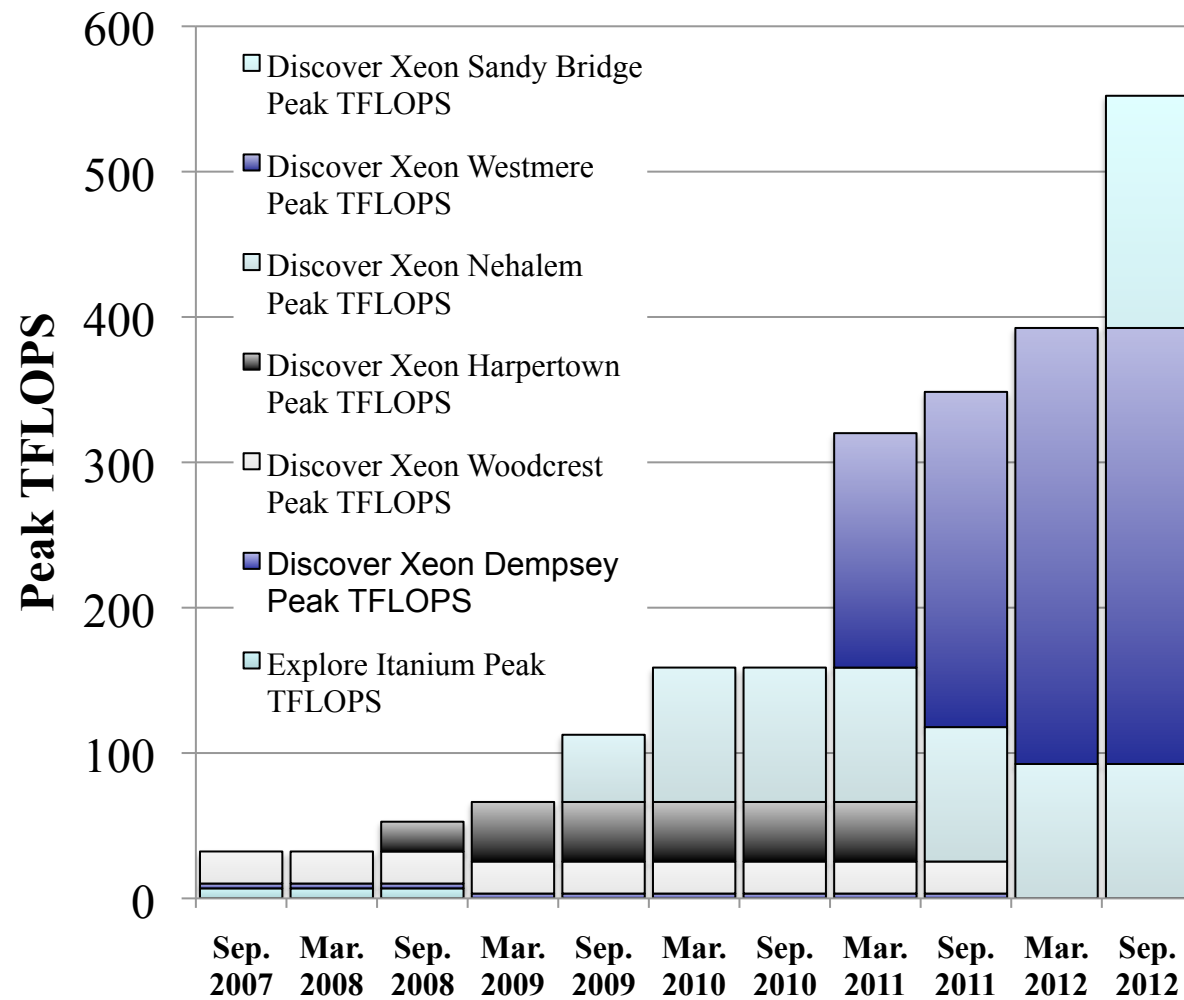
NCCS Discover Linux Cluster Utilization Normalized to 30-Day Month

NASA Center for Climate Simulation

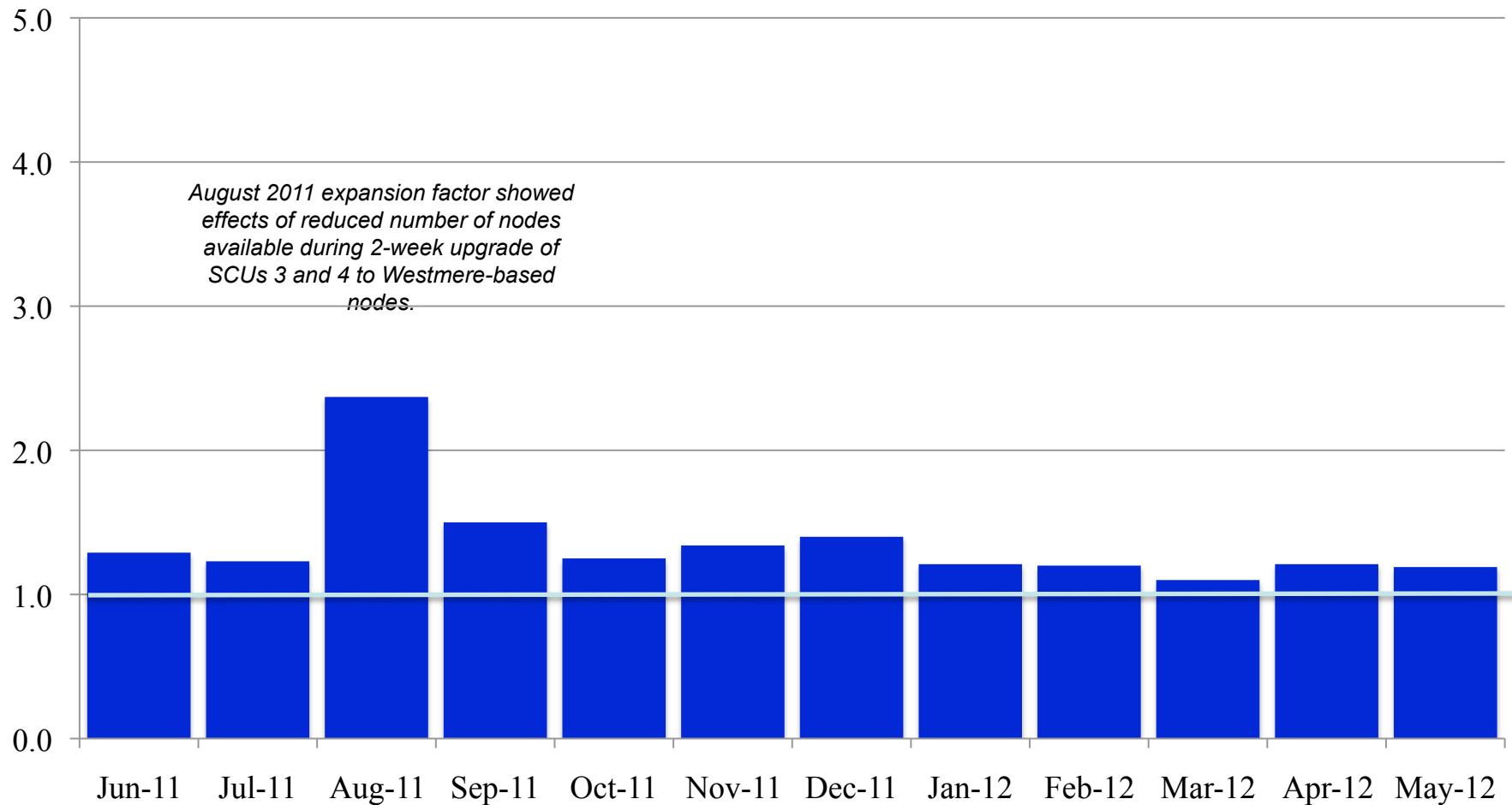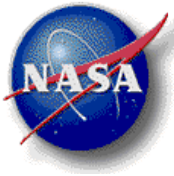# Discover Linux Cluster Expansion Factor

Expansion Factor = (Queue Wait + Runtime) / Runtime

August 2011 expansion factor showed effects of reduced number of nodes available during 2-week upgrade of SCUs 3 and 4 to Westmere-based nodes.
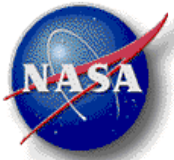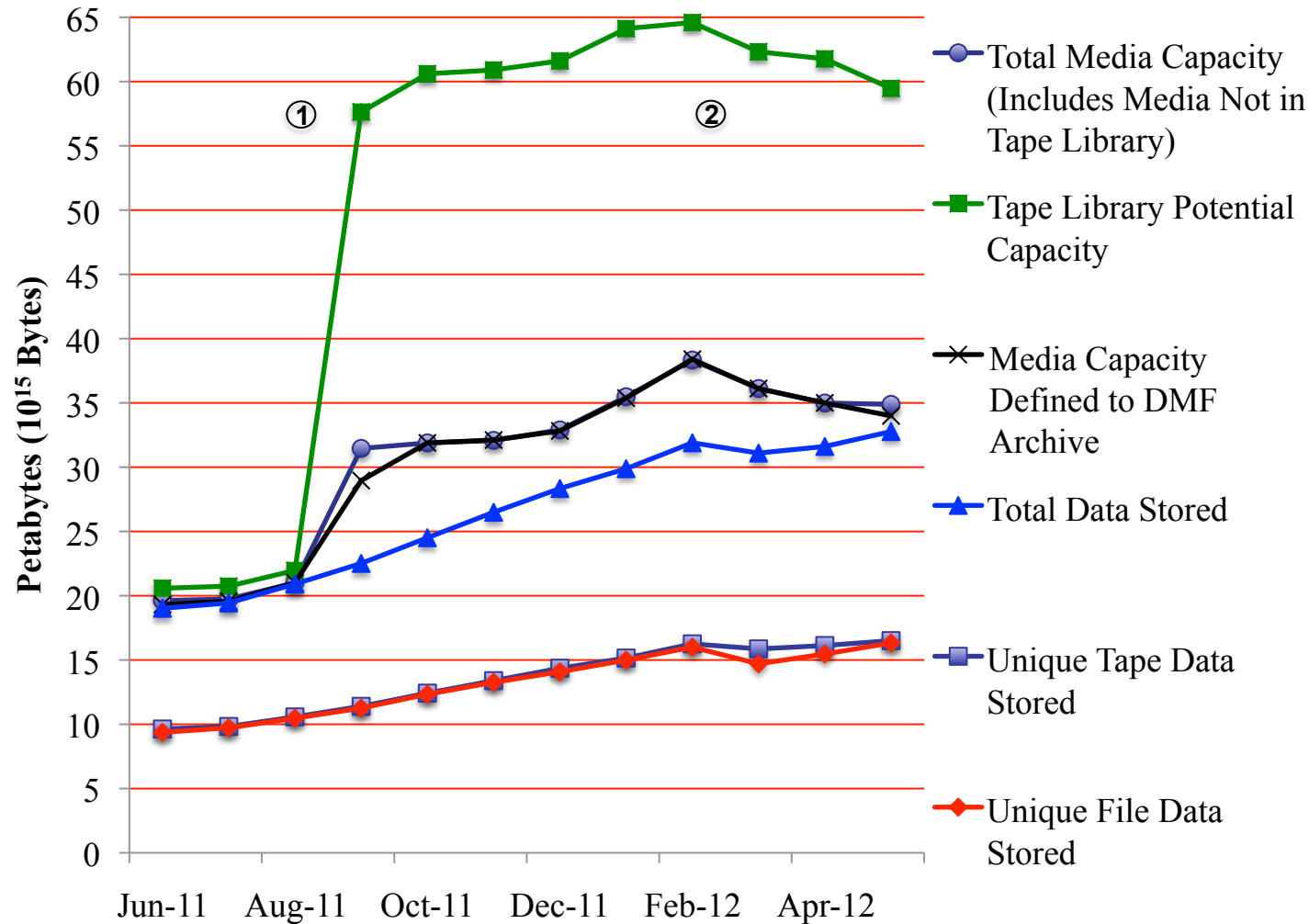
Discover Linux Cluster Downtime
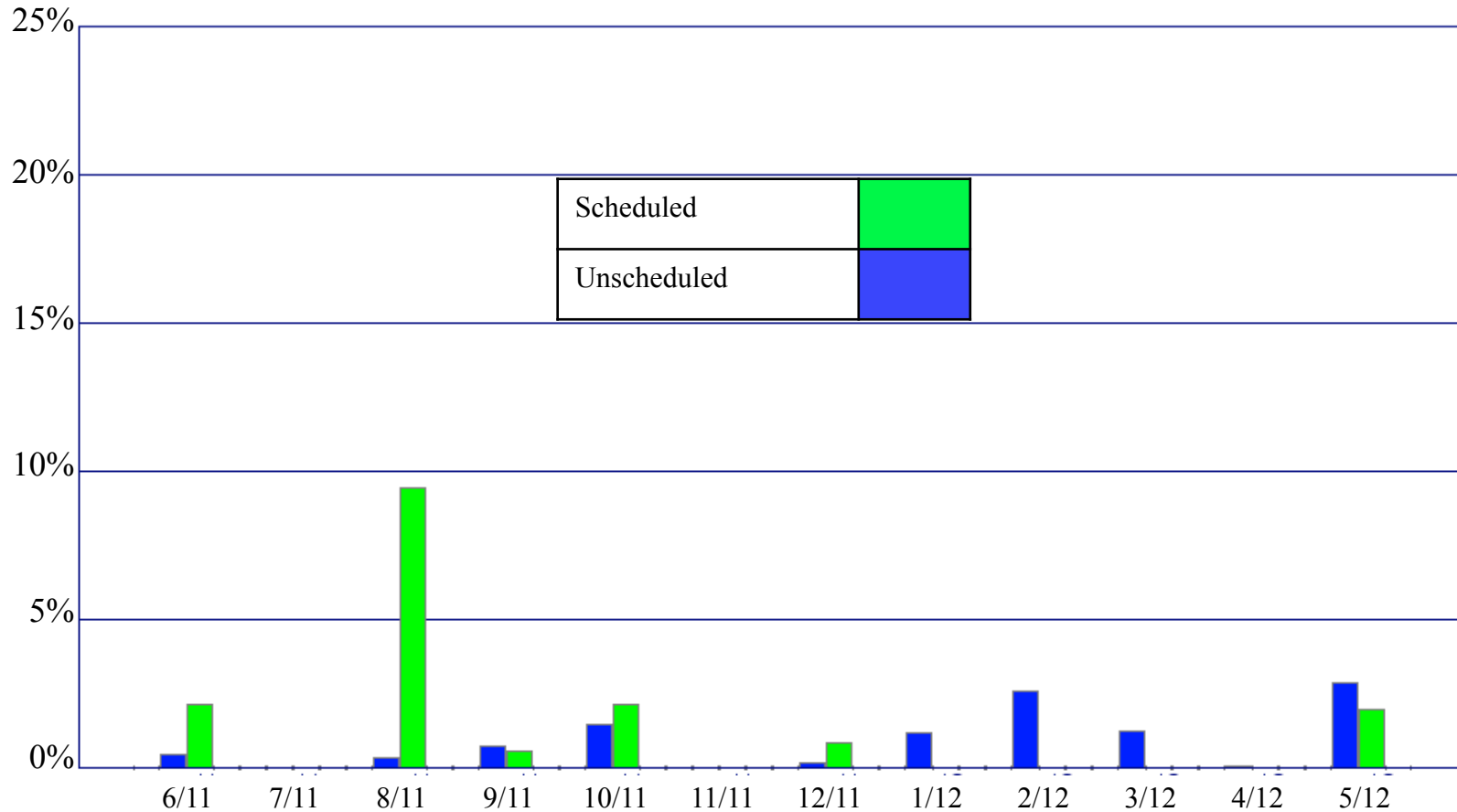
August 2011 scheduled
downtime included a 2-week
outage for Westmere upgrade
to SCUs 3 and 4 (512 nodes).

NASA Center for Climate Simulation

NCCS Mass Storage

Legend:
- Total Media Capacity (Includes Media Not in Tape Library)
- Tape Library Potential Capacity
- Media Capacity Defined to DMF Archive
- Total Data Stored
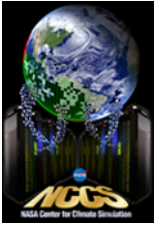- Unique Tape Data Stored
- Unique File Data Stored

① NCCS acquired an additional tape library and higher capacity tape drives and media to accommodate Mass Storage growth from GISS and GMAO CMIP5 simulations and GMAO's increased GEOS-5 resolution.
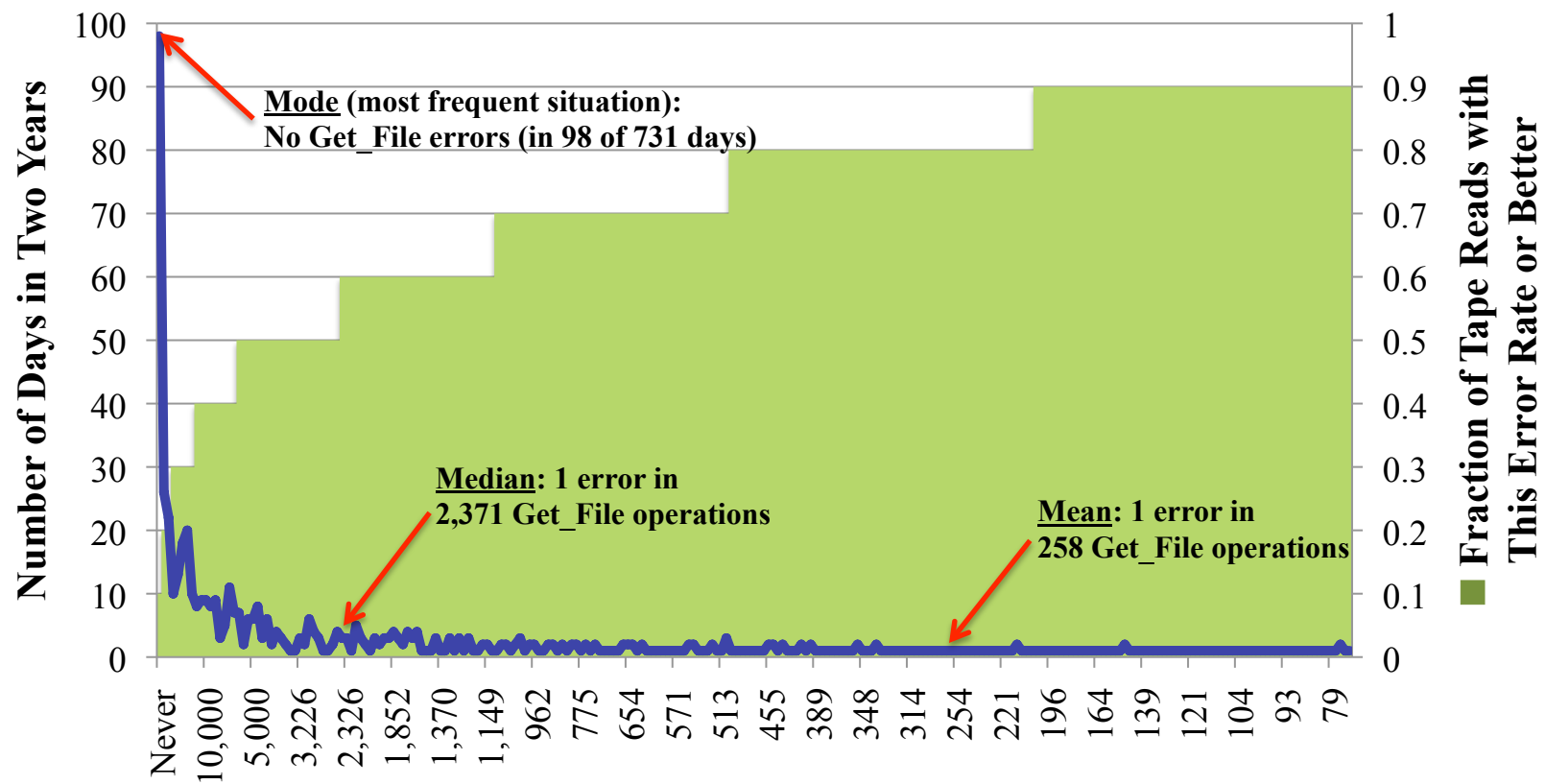
② "Tape Library Potential Capacity" decreased beginning in March 2012 because, at NCCS request, users began deleting unneeded data, and that tape capacity had not all been reclaimed so that new data could be written to the tapes.
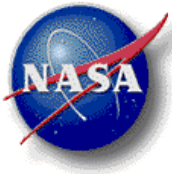
NCCS Observed "Get_File" Tape Operation Error Rates, Temporary and Permanent, May 2010 – May 2012

NCCS Archive Daily Tape Read Error Rates
May 2010 - May 2012

Mode (most frequent situation):
No Get_File errors (in 98 of 731 days)

Median: 1 error in
2,371 Get_File operations

Mean: 1 error in
258 Get_File operations

One Error in How Many Tape "Get_File" Operations?

Summary: Single Tape Copy Default for Archive Starting May 31, 2012

- Default becomes single tape copy of NCCS archive files:
  - …for all newly created files *and*
  - …all *existing* archive files.

- You can request a second tape copy of your **critical** archive files *at any time**:
  - **dmtag –t 2 <archive_file_name>**
  - *Second tape copy will be made within a few hours to a few days.

- See the following for more details:
  - https://www.nccs.nasa.gov/news.html#dmtag
  - http://www.nccs.nasa.gov/primer/data.html#secondcopy
  - https://www.nccs.nasa.gov/images/DMF-dtag.pdf

- Please contact support@nccs.nasa.gov (301-286-9120) if you have questions or concerns.