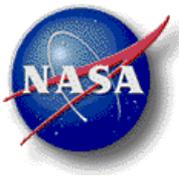


NCCS User Forum

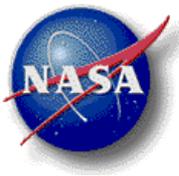
September 14, 2010



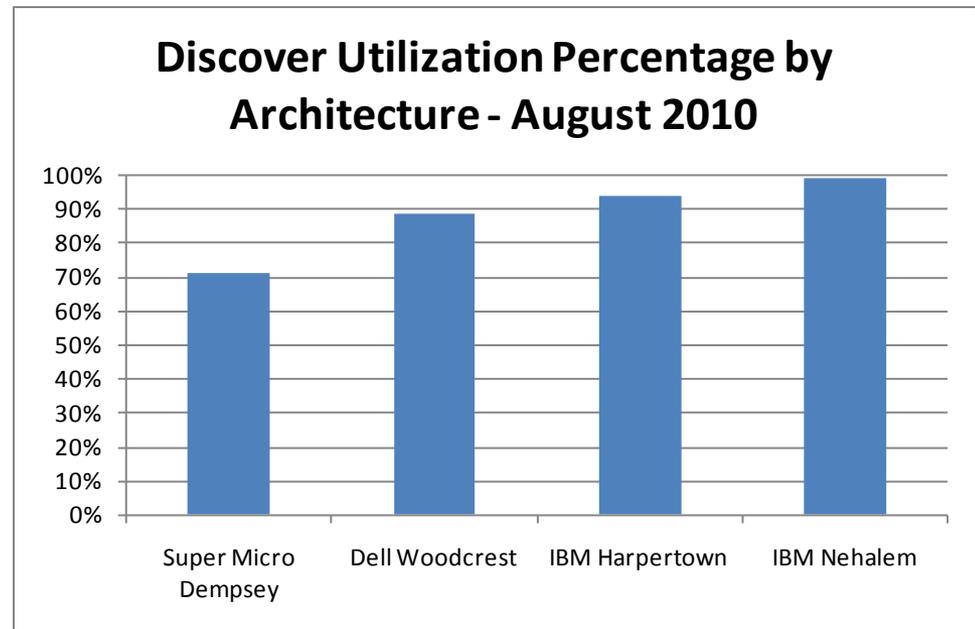
Agenda – September 14, 2010

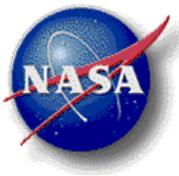


- **Welcome & Introduction (Phil Webster, CISTO Chief)**
- Current System Status (Fred Reitz, NCCS Operations Manager)
- User Services Update (Tyler Simon, NCCS User Services Group)
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- Earth System Grid Data Node Update
- Questions & Wrap-Up (Phil Webster)
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)



- NCCS computational machines are very busy (at capacity) with your science work.
 - Nehalem nodes utilization routinely more than 95% on workdays.
- Help is coming: SCU7 will double the peak capacity of Discover.





Accomplishments



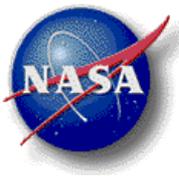
- **Systems**

- Discover SCU7, will double Discover's Peak TFLOPs
- Dirac archive
 - Disk nearly quadrupled
 - Migrate to new Distributed/Parallel Servers in October
- Collaboration with Intel to scale Intel MPI beyond 7,700 cores



- **Services**

- Earth System Grid Data Node for NCCS Users' IPCC Contributions
- iRODS-based Data Management System prototype
- Users' web access to NCCS "Footprints" trouble-ticket system



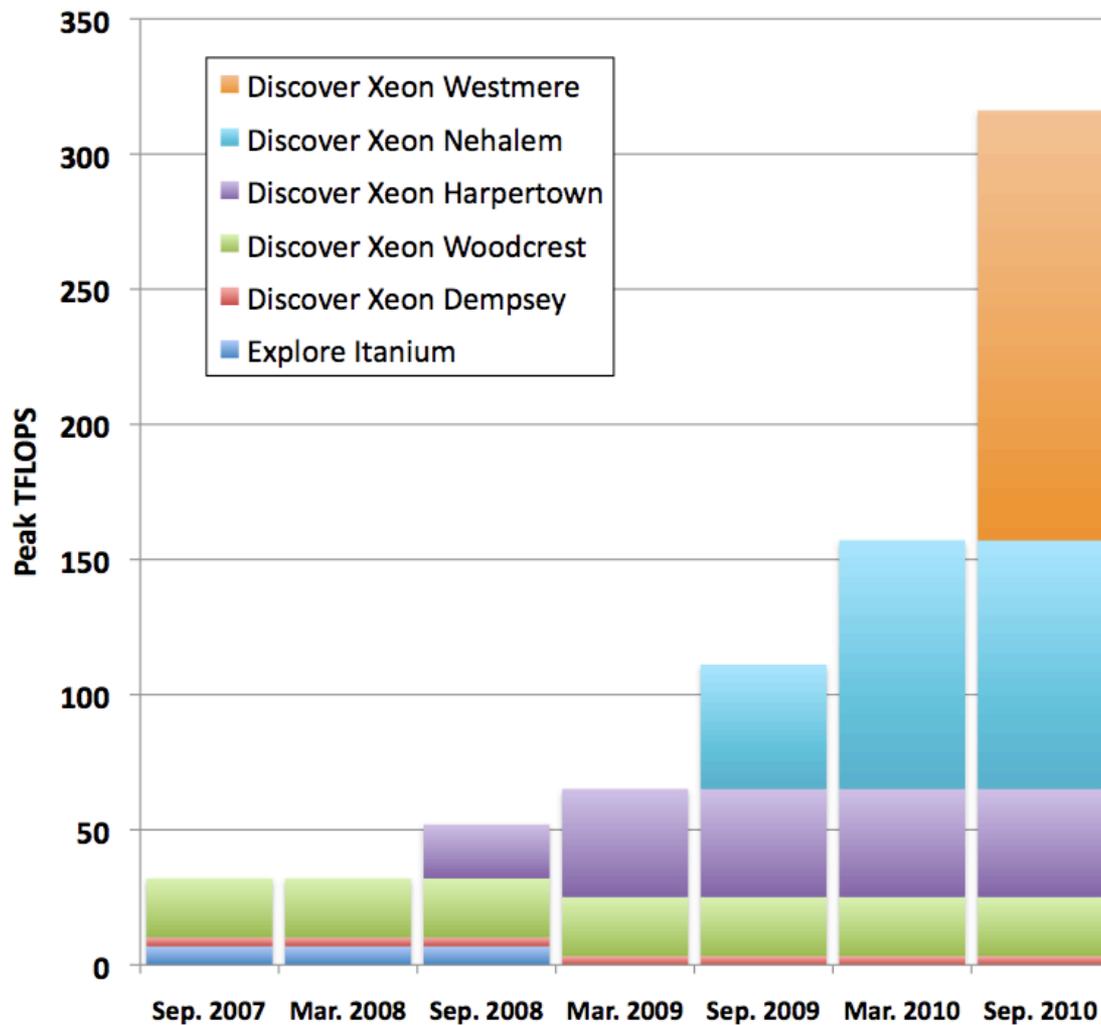
SMD Compute Allocation Period Announcement

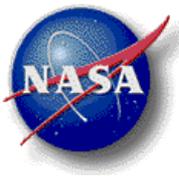


- Call for Computer Allocations for period starting November 1, 2010
 - “The Science Mission Directorate (SMD) will select from requests submitted to the e-Books online system **by September 20** for 1-year allocation awards beginning November 1. Any current projects set to expire on October 31 must have a request in e-Books to be considered for renewal.”
 - <https://ebooks.reisys.com/gsfcc/nccs/submission/index.jsp?solId=27>
- PI applications will specify “core-hours,” as before, but resources will be *awarded* using new System Billing Units (under development with NAS)
 - Allocation award letters will be stated in both units for information purposes.
 - When change begins, all awards will be converted to the new units in the accounting system.
 - Email will be sent to all users before the change to charging via SBUs begins.
- Please email support@hec.nasa.gov with any questions



NCCS Computing Capacity Evolution 2007-2010

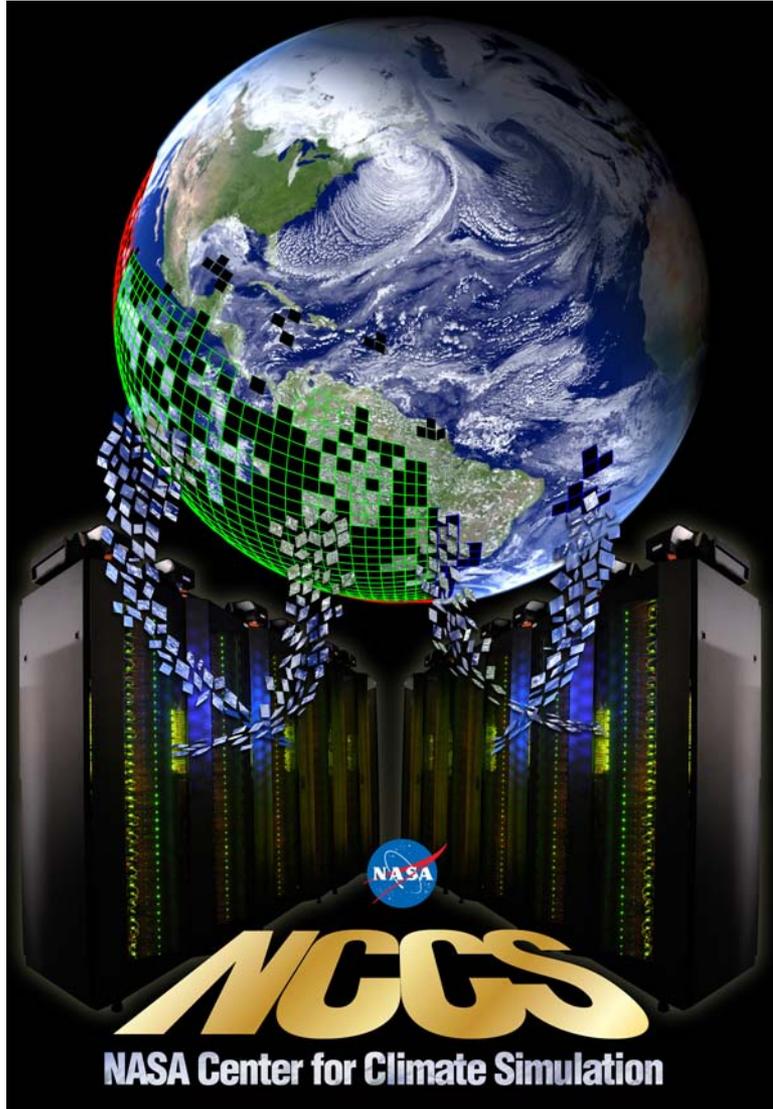




Agenda – September 14, 2010



- Welcome & Introduction (Phil Webster, CISTO Chief)
- **Current System Status (Fred Reitz, NCCS Operations Manager)**
- User Services Update (Tyler Simon, NCCS User Services Group)
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- Earth System Grid Data Node Update
- Questions & Wrap-Up (Phil Webster)
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)

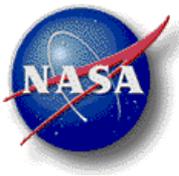


Operations and Maintenance

NASA Center for Climate Simulation
(NCCS) Project

Fred Reitz

September 14, 2010



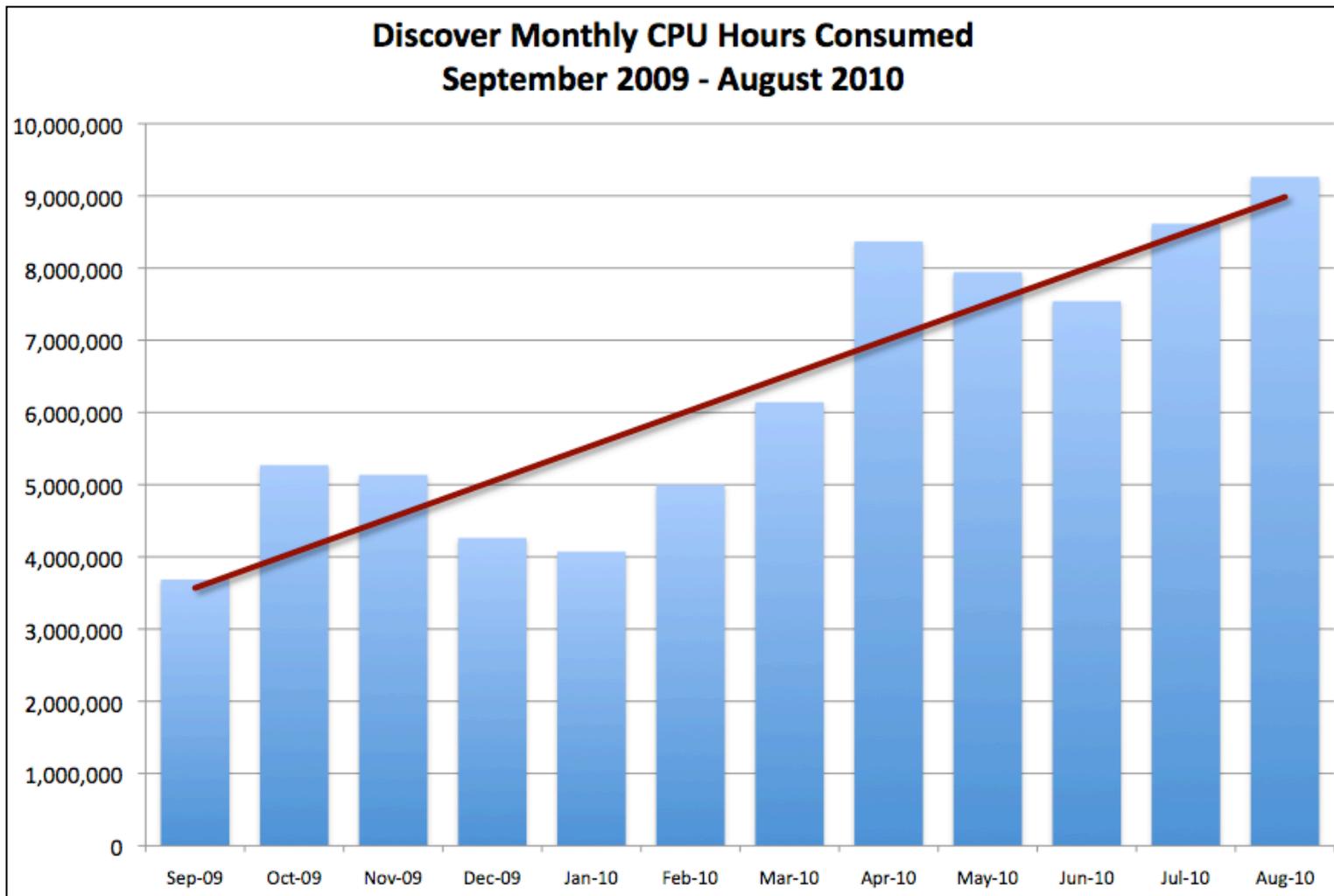
Accomplishments

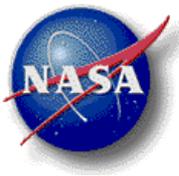


- *Discover*
 - Provided GRIP mission support. Still in mission support mode.
 - SCU7 preparations
 - Installed initial Intel Xeon “Westmere” hardware in Test and Development System (TDS)
 - Received SCU7 computing node hardware (159 TF peak)
 - Took delivery of power and cooling equipment, SCU7 cooling doors
 - Implemented new I/O monitoring tools.
 - Provided resources to Intel MPI Development Team to test scaling at 7000+ cores and investigate performance problems with Intel MPI libraries.
 - Installed two NVIDIA GPGPUs in *Discover* TDS, production *Discover* cluster.
- **Mass Storage**
 - Installed new Distributed DMF hardware in preparation for upcoming archive upgrade.
 - Implemented significant reconfiguration of DMF tape and disk Storage Area Networks in preparation for migration to new Distributed DMF servers and disk array.
- *Dataportal*
 - Created separate GPFS cluster for ESG nodes
 - Installed new *Dataportal* database hardware
 - Installed new *Dataportal* disk array
 - Replaced network switch with newer model



Discover Total CPU Consumption Past 12 Months (CPU Hours)

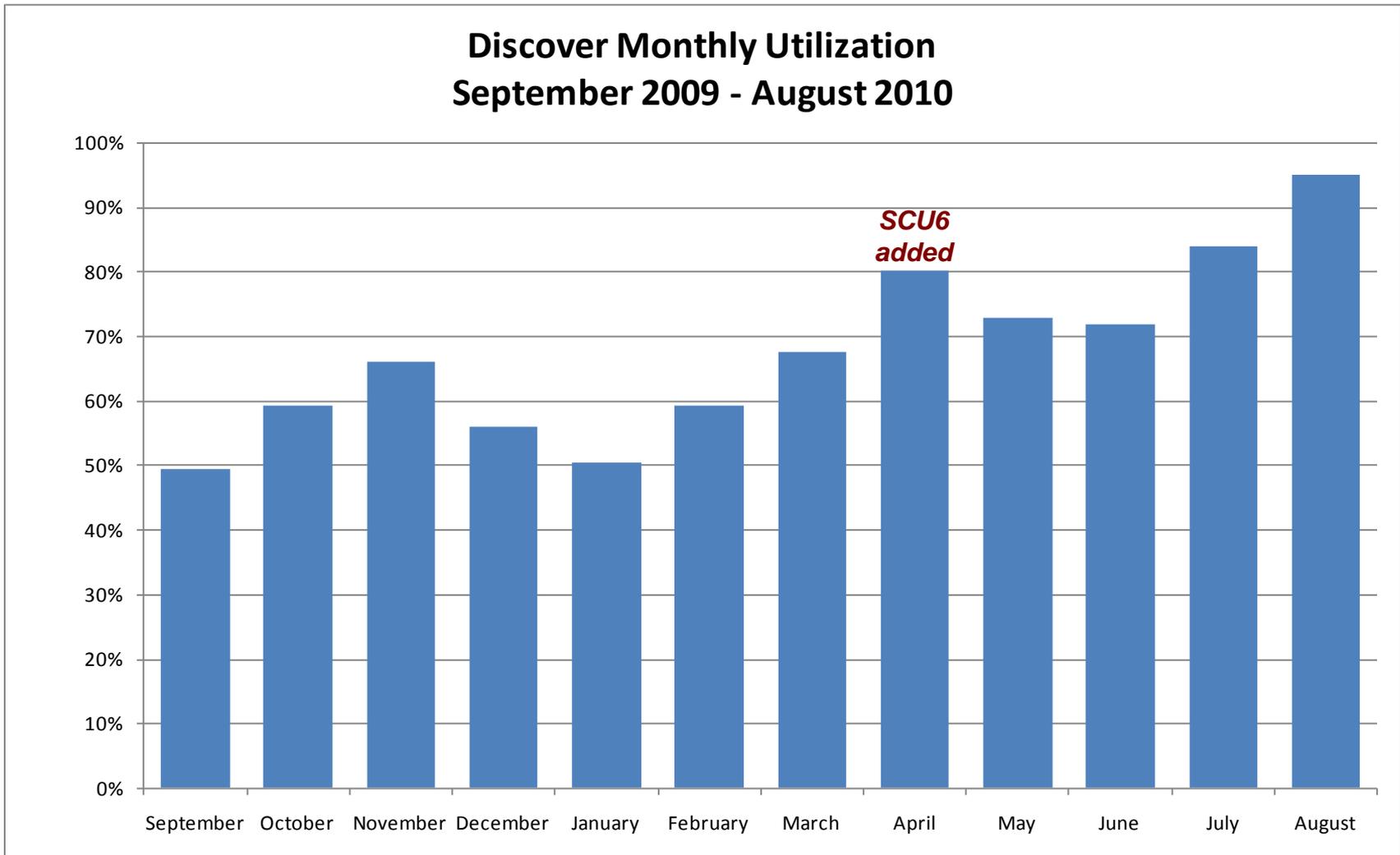


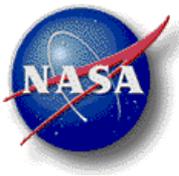


Discover Utilization Past 12 Months (Percentage)



Discover Monthly Utilization September 2009 - August 2010

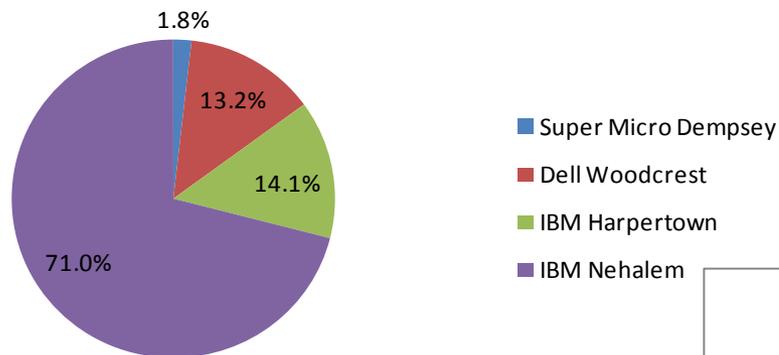




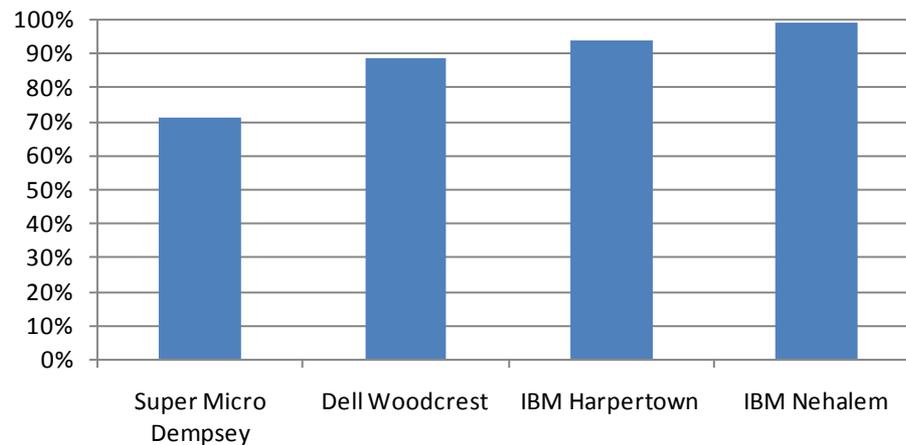
Discover Workload Distribution and Utilization by Architecture – August 2010

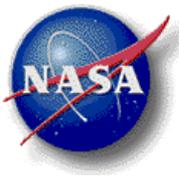


Discover Workload Distribution by Architecture - August 2010

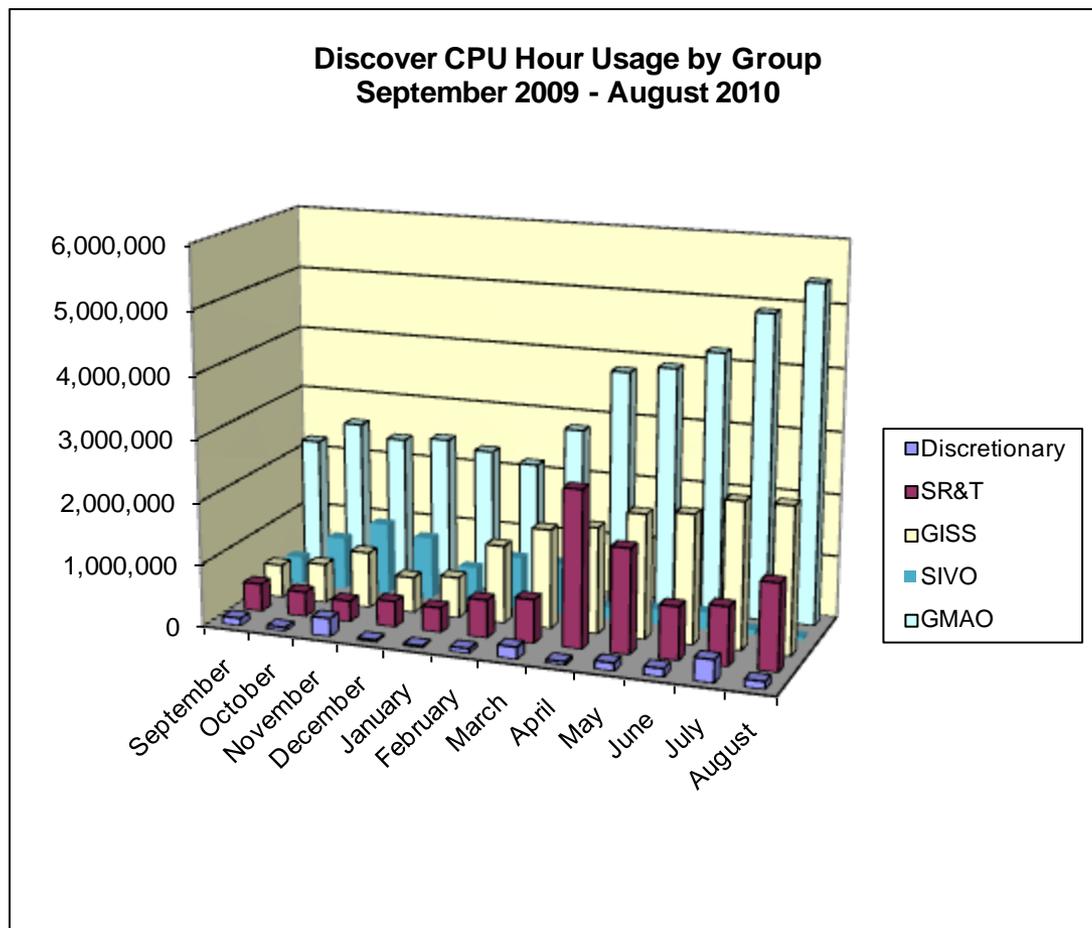


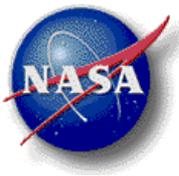
Discover Utilization Percentage by Architecture - August 2010



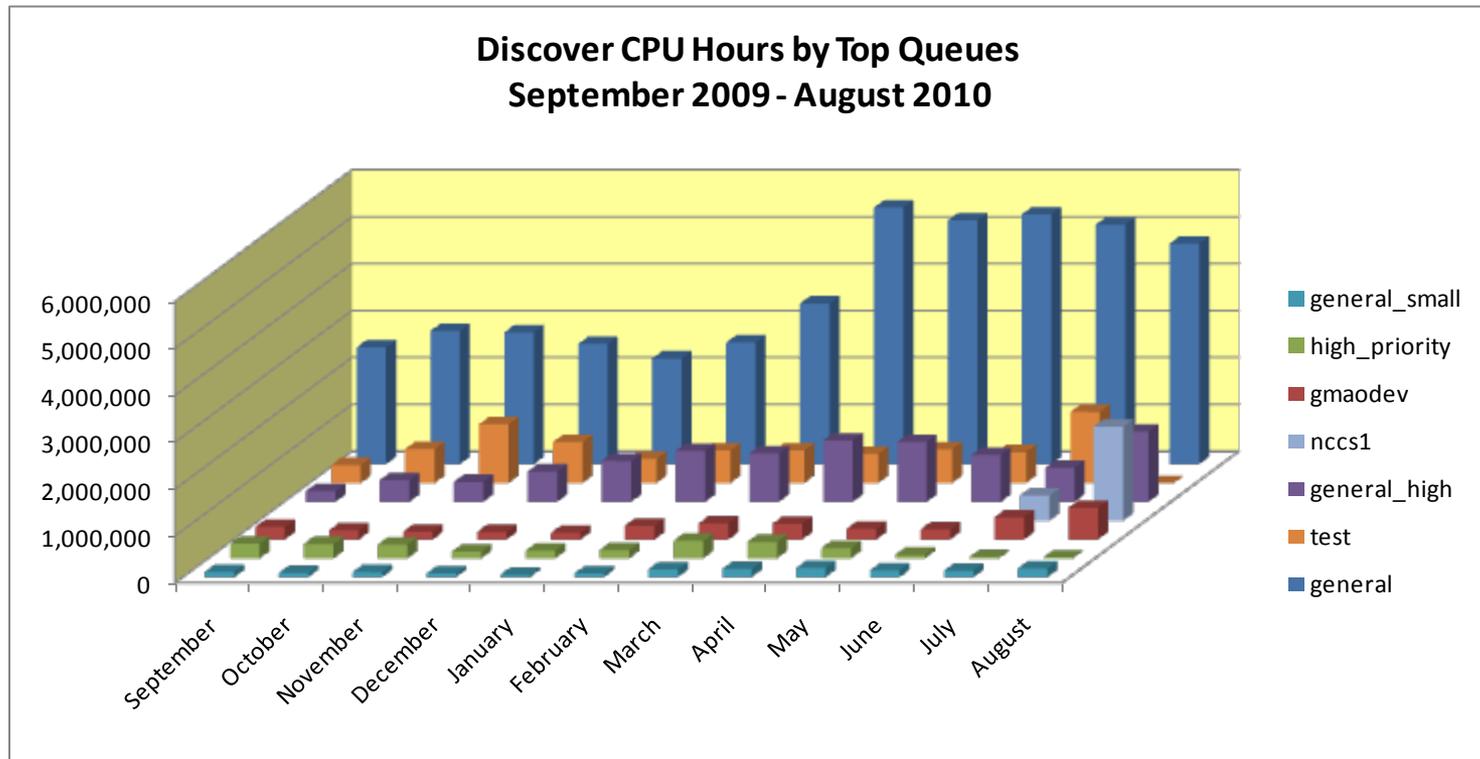


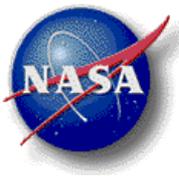
Discover CPU Consumption by Group Past 12 Months





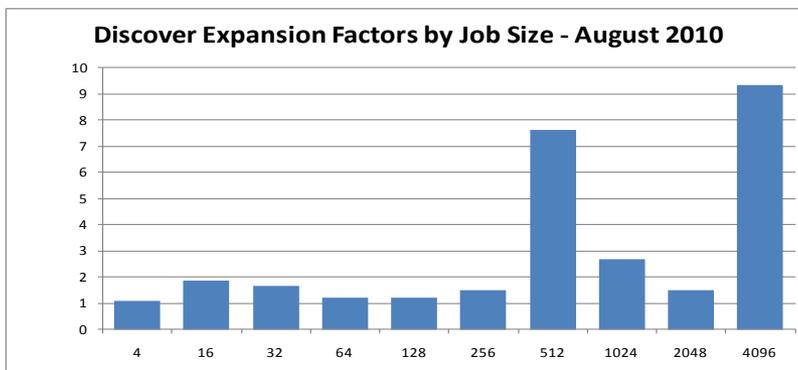
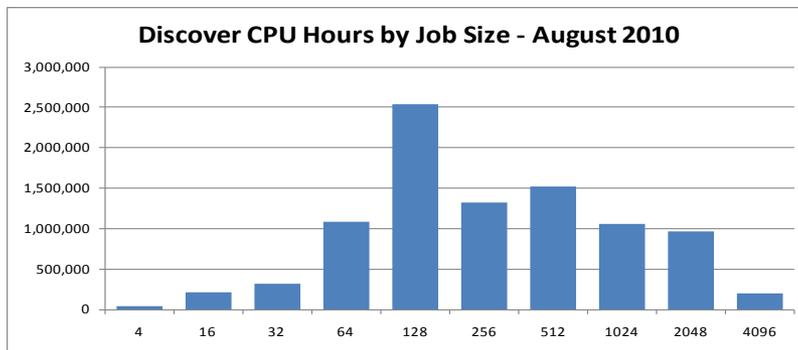
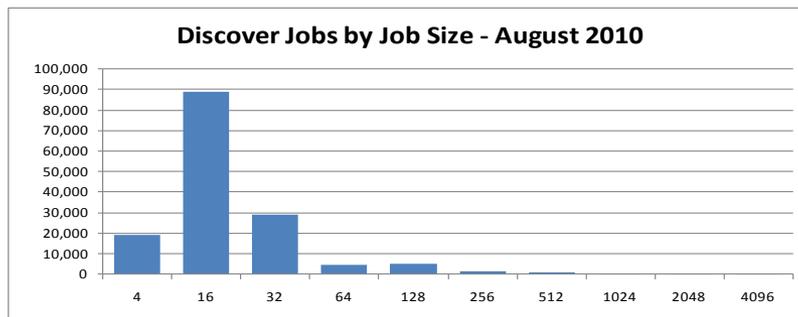
Discover CPU Consumption by Top Queues Past 12 Months





Discover Job Analysis by Job Size

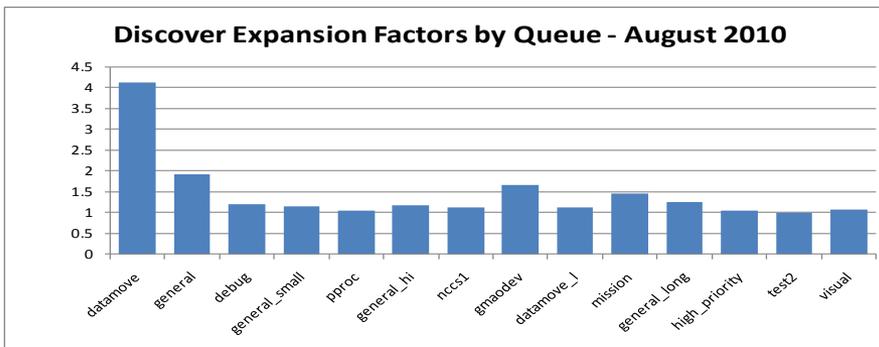
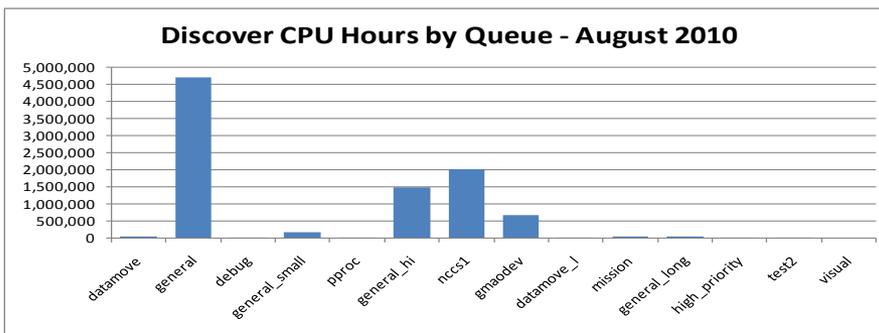
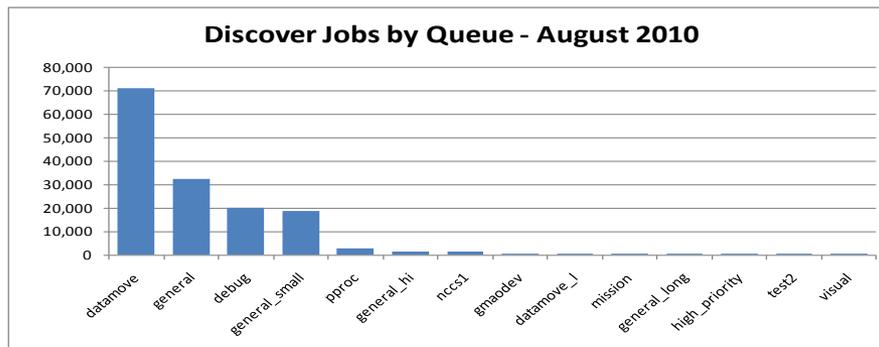
August 2010





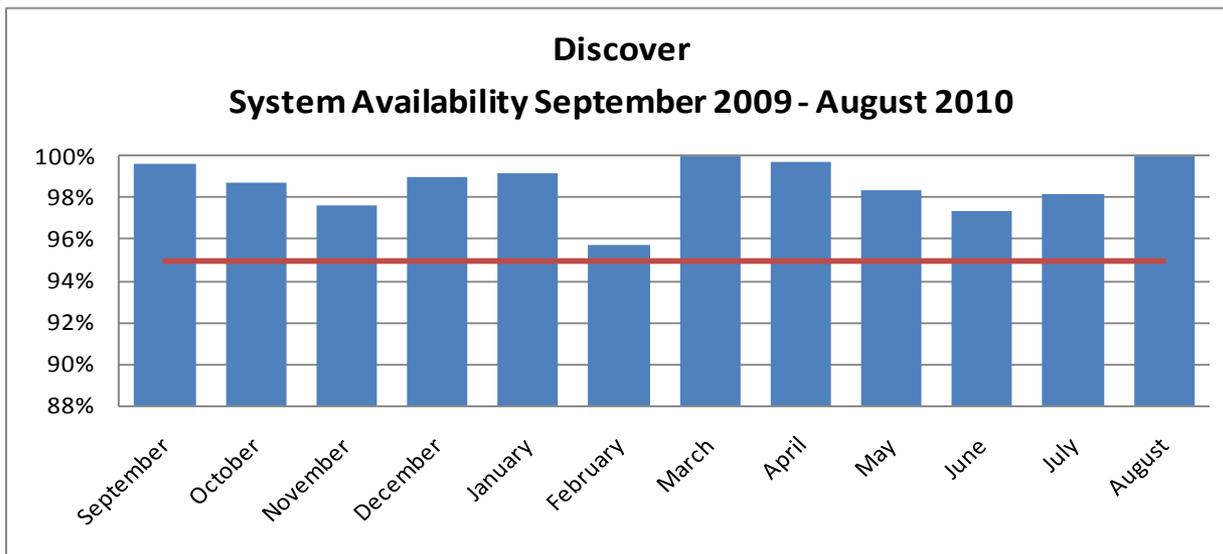
Discover Job Analysis by Queue

August 2010





Discover Availability Past 12 Months



Scheduled Maintenance: June-August

26 June – 8.5 hours

Corrected DDN5 cabling error

21 July – 12 hours

GPFS, OFED, disk controller firmware upgrades

Unscheduled Outages: June-August

4 June – 3 hours, 35 minutes

Internal SCU5 network flapping

10 June – 7 hours, 57 minutes

borgmg management server hang

25 June – 1 hour, 49 minutes

DDN5 disk controller cabling issue

21 July – 50 minutes

Extended maintenance window

23 July – 2 hours, 15 minutes

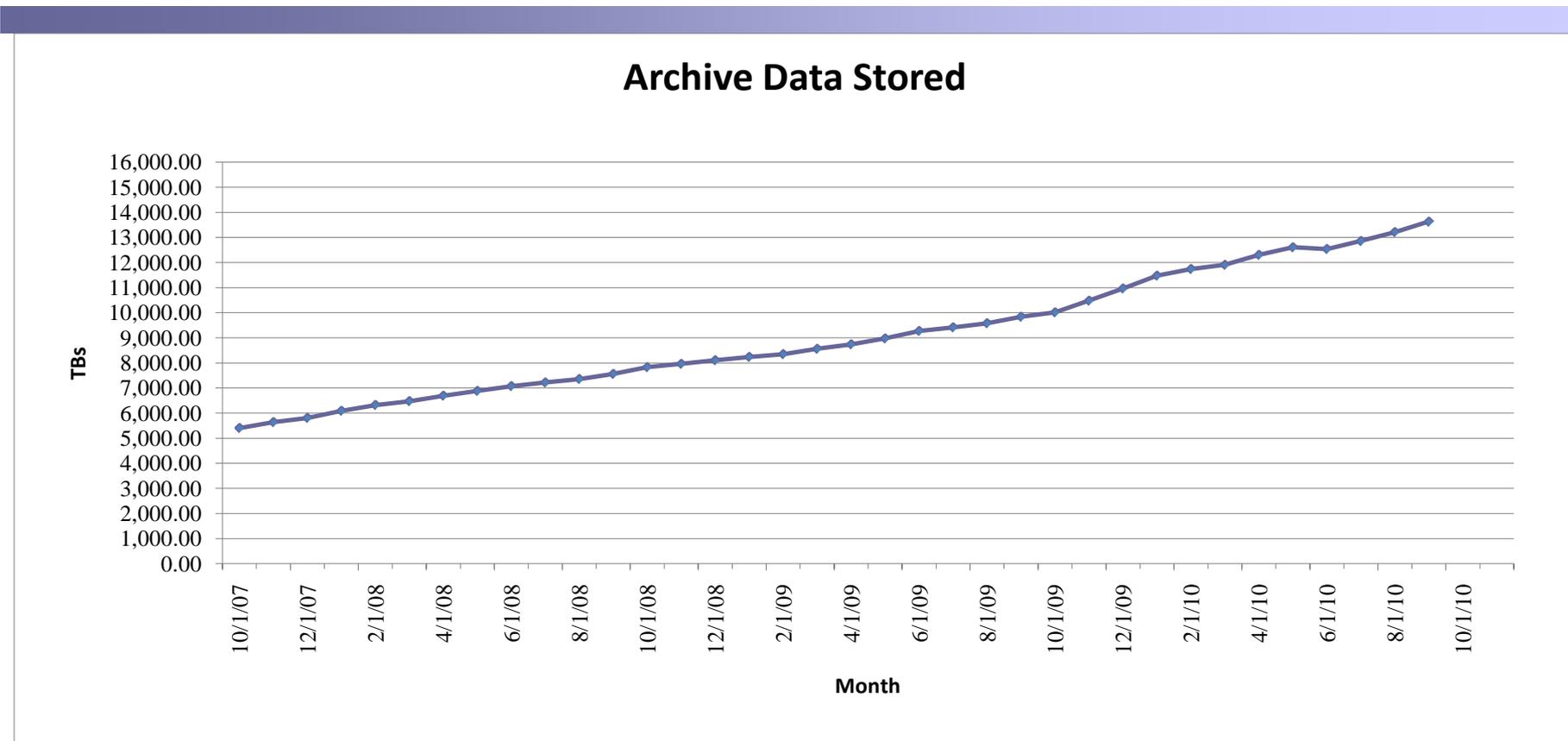
Internal SCU5 network flapping

26 August – 20 minutes

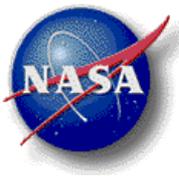
Internal SCU5 network flapping



Archive Data Stored



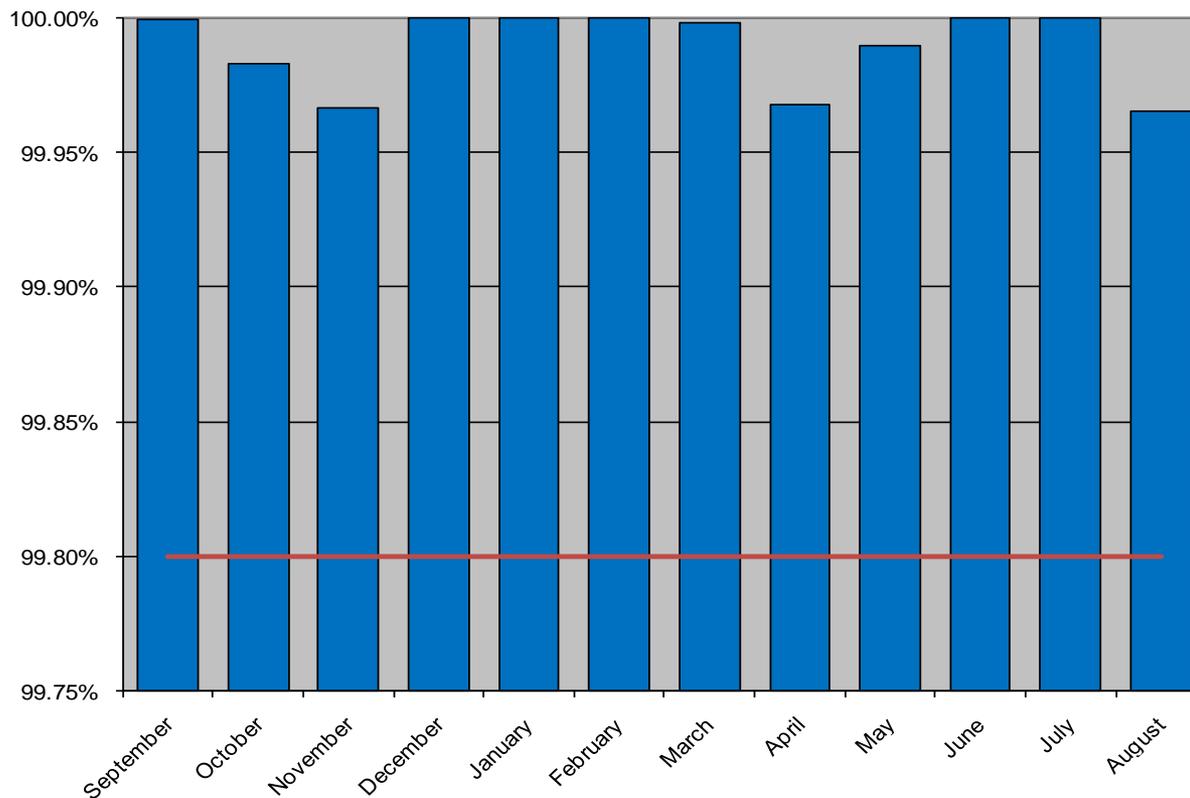
- 4 PBs added over last 12 months, versus 4 PBs in 20 months previously (10/1/07 to 6/1/09)
- NCCS pays an SGI license based on data stored (15 PBs at this time)
- Please remove any data you know you do not need

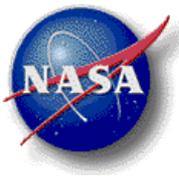


NCCS Network Availability Past 12 months



**NCCS-Managed Network Availability
September 2009 - August 2010**





Issues



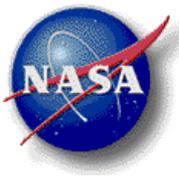
1. User jobs exhibiting high file open/close activity adversely affect other users' work
 - GPFS designed for larger files, streaming I/O
 - NCCS is working with IBM, monitoring system
 - Please work with NCCS, SIVO if your application exhibits high file open/close activity or uses many small files
2. Jobs take longer to begin execution
 - SCU7 will help
3. Users running nodes out of memory can cause GPFS hangs
 - Improved monitoring has reduced GPFS hangs
 - Please work with NCCS, SIVO if your application runs nodes out of memory



Future Enhancements



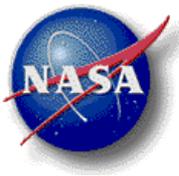
- Data Exploration Theater performance, functionality
- *Discover*
 - PBS v10
 - SLES11
 - SCU7
- *Dataportal*
 - iRODS
 - Additional storage
 - New database servers
- Mass Storage
 - Platform upgrade
 - Additional storage



Archive (*Dirac*) Upgrades



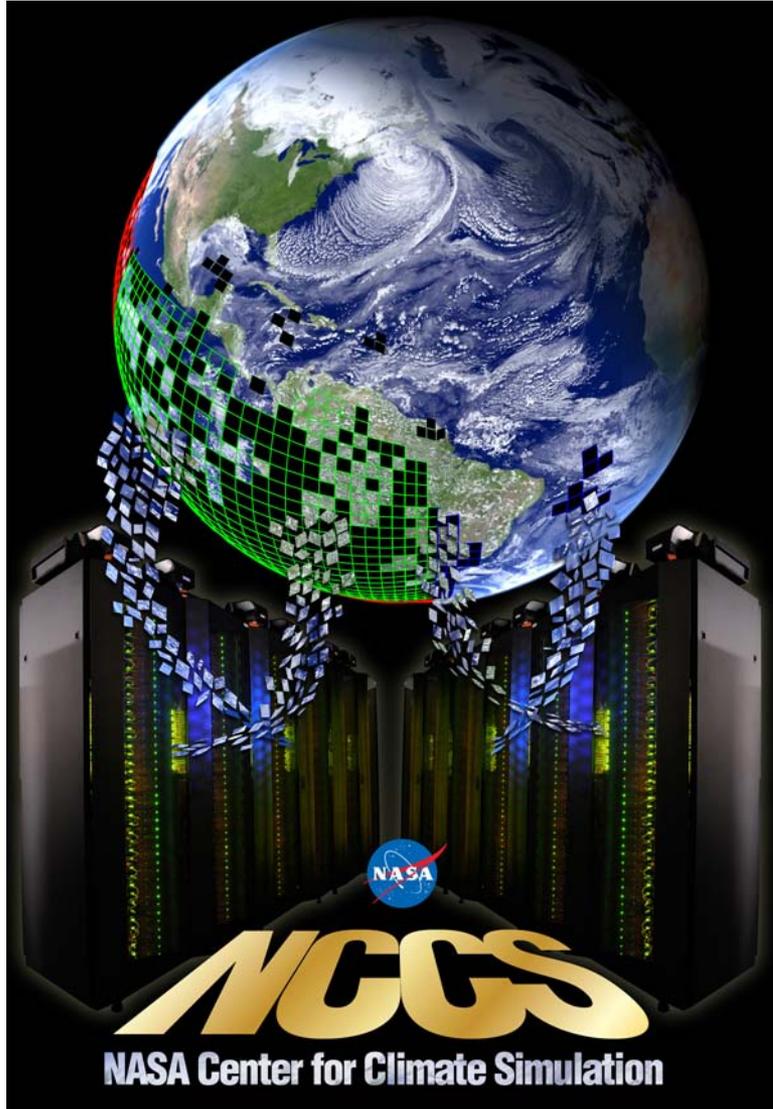
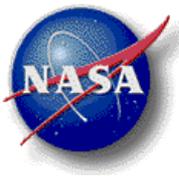
- *Dirac* currently on 5 year old system
- Upgrades
 - New DMF server to replace Bx2
 - Parallel Data Movers for tape
 - Two Login nodes for interactive access to archive data
 - Two NFS nodes for improved NFS access to archive data from *Discover*
 - All systems commodity Xeon based
 - Upgrade to primary disk cache to 600 TB RAW - **Complete**
 - Upgrade to disk cache for DMF databases to improve serviceability and data access
- What does this mean to users?
 - Much higher performing archive
 - Outage for moving the DMF databases between systems (please be patient)
- Timeline
 - Currently scheduled for October 12 and 13



Agenda – September 14, 2010



- Welcome & Introduction (Phil Webster, CISTO Chief)
- Current System Status (Fred Reitz, NCCS Operations Manager)
- **User Services Update (Tyler Simon, NCCS User Services Group)**
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- Earth System Grid Data Node Update
- Questions & Wrap-Up (Phil Webster)
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)



User Services

**NASA Center for Climate Simulation
(NCCS) Project**

Tyler Simon

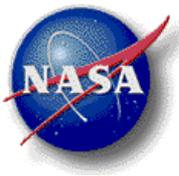
September 14, 2010



NCCS User Services



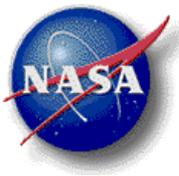
- Online Ticketing
- Compiler/MPI Baselines
- Performance concerns
 - I/O Compute/Archive
 - Job throughput
 - Architectural behavior (demp, wood, harp, neha)
- Tools, benchmarking



Agenda – September 14, 2010



- Welcome & Introduction (Phil Webster, CISTO Chief)
- Current System Status (Fred Reitz, NCCS Operations Manager)
- User Services Update (Tyler Simon, NCCS User Services Group)
- **SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)**
- Earth System Grid Data Node Update
- Questions & Wrap-Up (Phil Webster)
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)



10 Years of NCCS Highlights



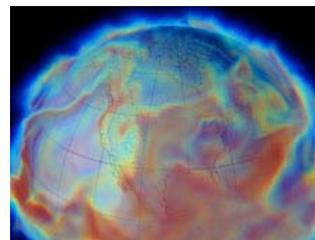
Halem – HP/Alpha Cluster
18th fastest computer in the world



Large SGI Shared Memory Systems at GSFC



Discover – First production IB commodity cluster in NASA HEC



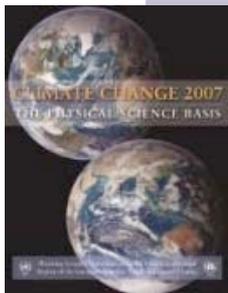
Introduction of Dali analysis environment, interactive access to large data sets



Climate Data Theater
New tool for scientific analysis and visualization

2000

Support for the IPCC AR4 modeling runs by GISS – Nobel Prize



10 GbE NCCS LAN
10 GbE between HEC centers



Introduction of the Data Portal System and Services
WMS, MAP, Cloud Library



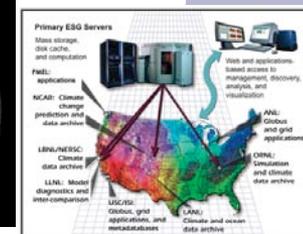
2005

Discover upgrades supports 3 KM global resolution model – highest resolution ever run



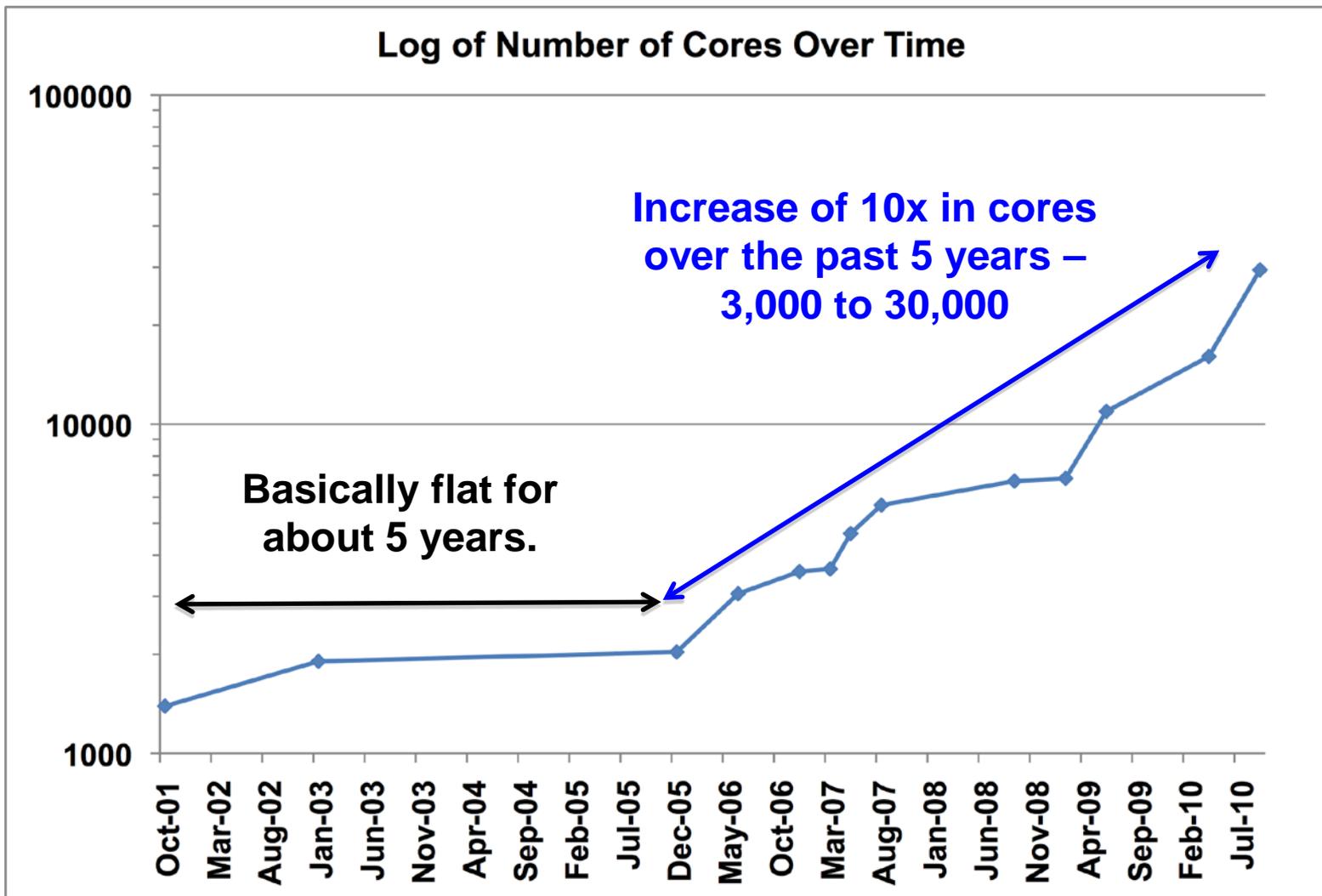
2010

Advancement of Data Services with the Data Management System (DMS) and the Earth Systems Grid (ESG)



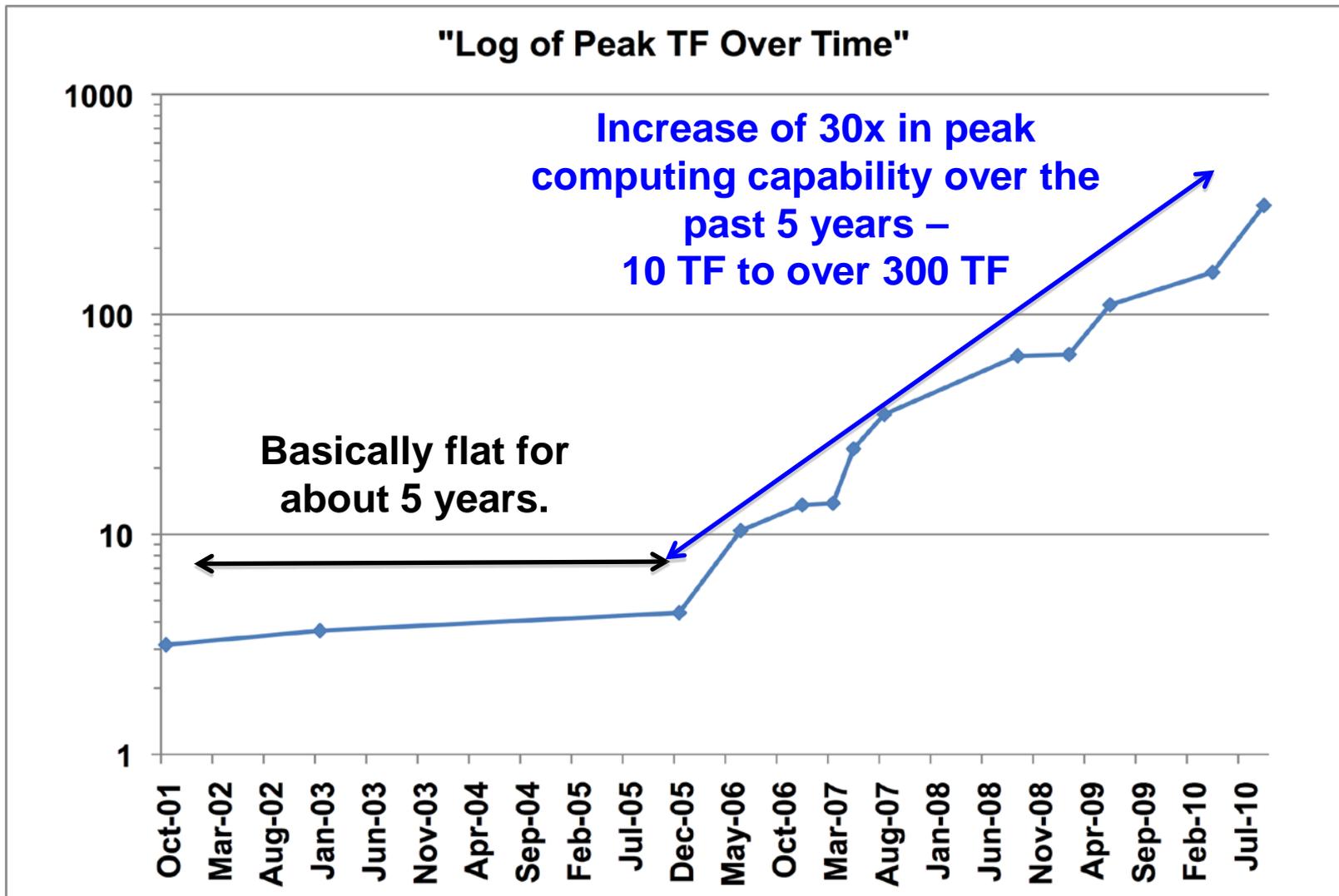


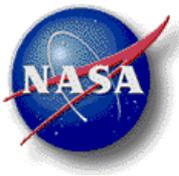
Growth of Computing Cores Over the Last 10 Years





Growth of Peak Computing Capability Over the Last 10 Years





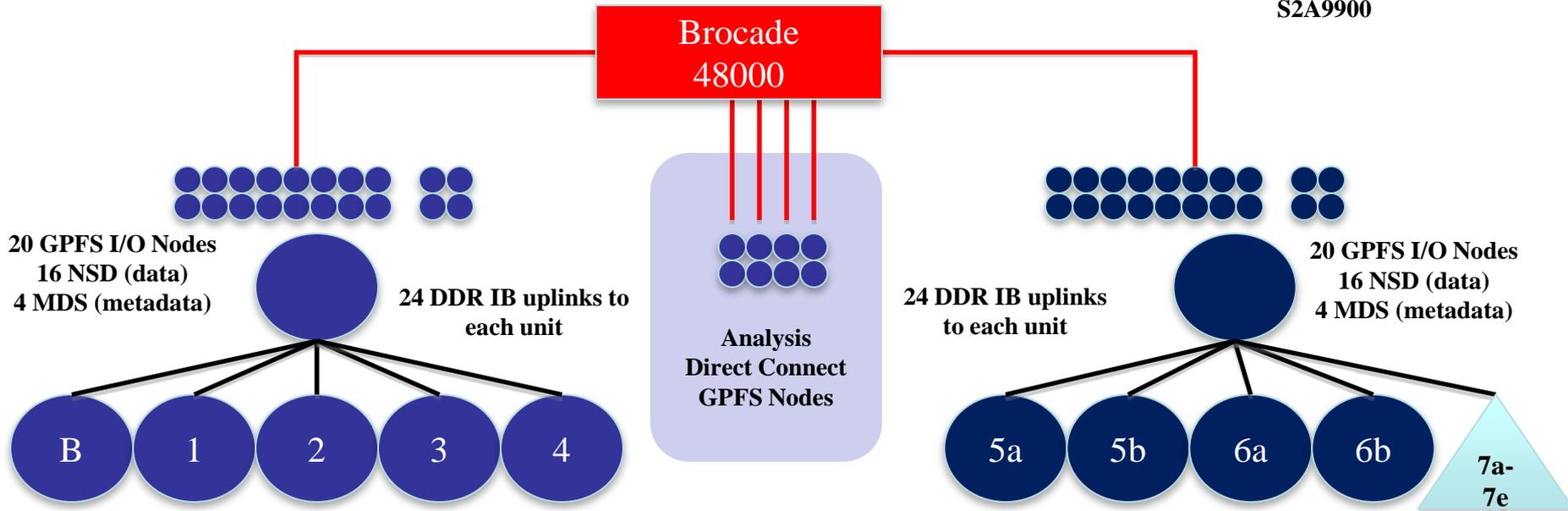
Representative IB and GPFS Architecture



Metadata File Systems:
IBM/Engenio
DS4700



Data File Systems:
Data Direct Networks
S2A9500
S2A9550
S2A9900



- **Base Unit: 512 Dempsey (3.2 GHz)**
- **SCU1: 1,024 Woodcrest (2.66 GHz)**
- **SCU2: 1,024 Woodcrest (2.66 GHz)**
- **SCU3: 2,048 Harpertown (2.5 GHz)**
- **SCU4: 2,048 Harpertown (2.5 GHz)**



Each circle represents a 288-port DDR IB Switch



The triangle represents a 2-to-1 QDR IB Switch fabric

- **SCU5: 4,096 Nehalem (2.8 GHz)**
- **SCU6: 4,096 Nehalem (2.8 GHz)**
- **SCU7: 14,400 Westmere (2.8 GHz)**



FY10 SCU7 Discover Cluster Augmentation



- SCU7 will augment Discover computing nodes with Intel Xeon Westmere processors within the existing SCU5 and SCU6 InfiniBand communication fabric.
 - Dual-socket, hex-core processors.
 - 12 MB L3 cache, 50% larger than Nehalem's.
 - 24 GB of RAM.
 - Quad-Data Rate (QDR) Infiniband
- Initial performance of the GEOS-5 Cubed-Sphere benchmark showed a slight improvement on Westmere processors compared to Nehalem.
 - Test compared use of 4 cores per socket on both Nehalem and Westmere.
- System has been delivered.



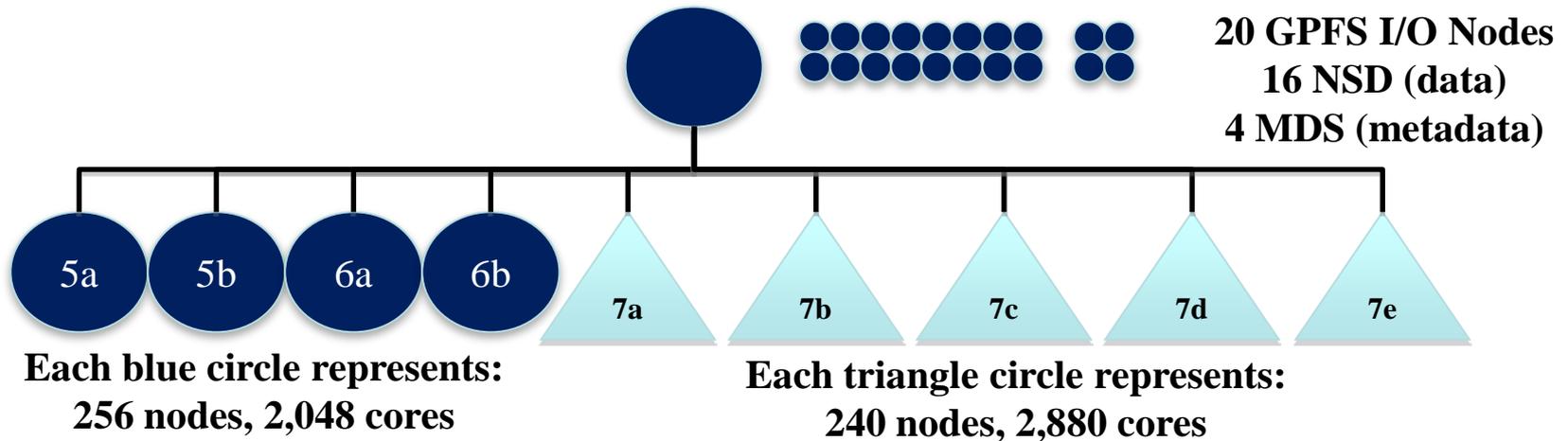
Intel Xeon 5600 processor series: "Westmere"

Xeon Processor	Speed	Cores/Socket	L3 Cache Size
Nehalem	2.8 GHz	4	8 MB
Westmere	2.8 GHz	6	12 MB

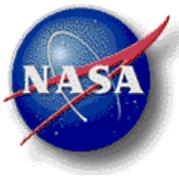
Discover's SCU7 will augment the computing capacity available for high resolution GEOS-5 Cubed Sphere runs and for GISS and GMAO simulations in support of the IPCC 5th Reassessment.



More SCU7 Details and Timeline



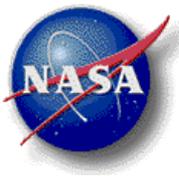
- SCU7 is made up of 5 units of 240 nodes
 - Full QDR bi-section bandwidth (32 Gbps) within those 240 nodes (2,880 cores)
 - 24 DDR (320 Gbps total) uplinks between each SCU7 unit and the top level switch
- To be available for general use by early December.



Agenda – September 14, 2010



- Welcome & Introduction (Phil Webster, CISTO Chief)
- Current System Status (Fred Reitz, NCCS Operations Manager)
- User Services Update (Tyler Simon, NCCS User Services Group)
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- **Earth System Grid Data Node Update**
- Questions & Wrap-Up (Phil Webster)
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)



Earth System Grid (ESG) Data Node for IPCC AR5

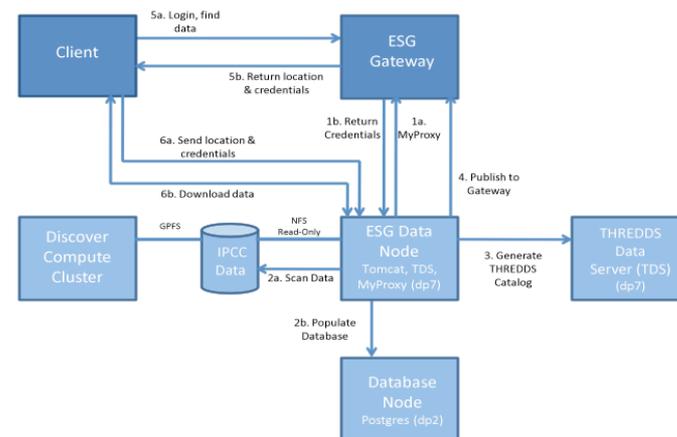


- Earth System Grid purpose: provide data, information, models, analysis, and visualization tools and computational capabilities for management and analysis of climate resources.
- NCCS's initial ESG Data Node is being set up to serve IPCC AR5 climate data (GISS and GMAO contributions).
- ESG allows the user to search and subset all the IPCC AR5 data by CMIP5 variable name.
- GMAO and GISS contributions will be available via the NCCS Data Portal: esg.nccs.nasa.gov
 - Access is via the PCMDI ESG Gateway at the CMIP5 link under: www.earthsystemgrid.org or at PCMDI link (below)
- NCCS's data architecture for the ESG Data Node:
 - AR5 data are exported read-only from their source location to the Data Node.
 - Thus scores of TB are not required to be copied & replicated within the NCCS environment, reducing disk and I/O resource requirements.
- Status: published initial GMAO and GISS sample data to Lawrence Livermore National Laboratories' Program for Climate Model Diagnosis and Intercomparison (PCMDI) ESG Gateway pcmdi3.llnl.gov/esgcat



Earth System Grid ipcc

INTERGOVERNMENTAL PANEL ON climate change

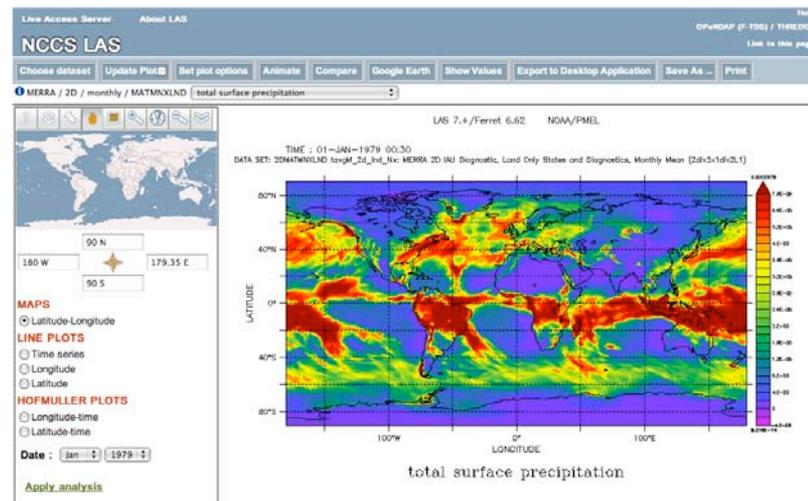


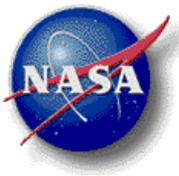


Earth System Grid and NCCS Data Portal: Live Access Server (LAS) for Data Analysis



- NCCS staff has also provided early access to Live Access Server (LAS) analysis services for AR5 and other climate data on the NCCS Data Portal.
 - LAS is to be part of ESG Data Node distribution, but is not yet included in the current distribution.
 - LAS is OPeNDAP-aware and allows users to plot, animate, subset and compare geo-referenced scientific (e.g. climate) data from distributed locations.
- The Live Access Server (LAS) is available under esg.nccs.nasa.gov and, as proof of concept, provides access to the much of the MERRA data stored at the GES DISC as well as some YOTC data.
- Data under LAS is organized by project, 2D, 3D for easy navigation.
- Future plans include offering THREDDS to provide download, OPeNDAP and NetCDF subsetting.

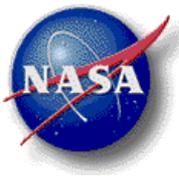




Agenda – September 14, 2010



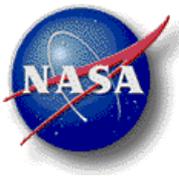
- Welcome & Introduction (Phil Webster, CISTO Chief)
- Current System Status (Fred Reitz, NCCS Operations Manager)
- User Services Update (Tyler Simon, NCCS User Services Group)
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- Earth System Grid Data Node Update
- **Questions & Wrap-Up (Phil Webster)**
- Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)



Questions and Wrap-Up



- Contact information:
 - NCCS User Services:
 - support@nccs.nasa.gov
 - 301-286-9120
 - www.nccs.nasa.gov
 - Lynn Parnell, HPC Lead:
 - Lynn.Parnell@nasa.gov
 - Phil Webster, CISTO Chief:
 - Phil.Webster@nasa.gov



Agenda – September 14, 2010



- Welcome & Introduction (Phil Webster, CISTO Chief)
- Current System Status (Fred Reitz, NCCS Operations Manager)
- User Services Update (Tyler Simon, NCCS User Services Group)
- SCU7 and Other NCCS Systems Updates (Dan Duffy, NCCS Lead Architect)
- Earth System Grid Data Node Update
- Questions & Wrap-Up (Phil Webster)
- **Breakout Session: Footprints Tutorial (Max Guillen, NCCS User Services)**