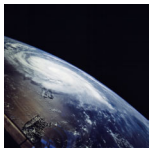
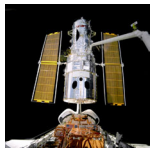
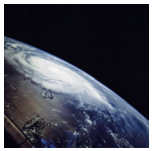
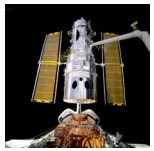




NCCS User Forum

15 May 2008

Agenda


Welcome & Introduction
Phil Webster NCCS

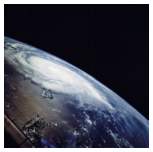
Current System Status
Fred Reitz, Operations Manager
System Issues *Utilization*
Pending Upgrades *Changes*

New Compute Capability at the NCCS
Dan Duffy, Lead Architect
Schedule *Impact of Discover Changes*
Storage Cluster *Architecture* *Quad Core*

User Updates
Sadie Duffy, User Services Lead
One of a Kind Data *SIVO announcements*
Allocation Updates *Changes*
Transition support

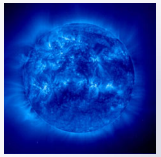
Questions / Comments
Phil Webster

Data-Centric Conceptual Architecture



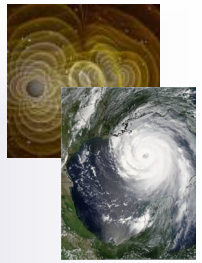
Data Portal & Data Stage

- Developing requirements for follow on system
- Significant increase in storage and compute capability



Collaborative Environments

- Web based tools
- Modeling Guru



Analysis

- Conceptual framework for analysis environment
- Matlab on Discover

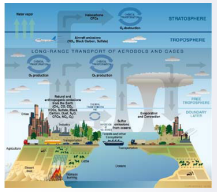
Visualization

- Collaboration with Scientific Visualization Studio to provide tools
- Visualization nodes on Discover

DATA

- Single NCCS home, scratch, and application file system
- Data management initiative

High Speed Networks



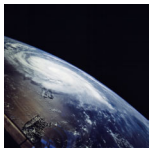
Compute

- Significant upgrade to Discover
- Decommission Explore

Archive

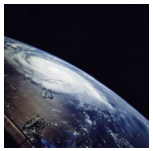
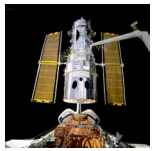
- Increased disk cache for longer file retention
- Increase tape capacity to meet growth requirements
- Upgrade file server (newmintz) and data analysis host (dirac)

NCCS Staff Transitions



- Lead for User Services
  - Sadie Duffy will leave 5/16/08
  - New Lead will be on site mid-June
- Operations Manager
  - Fred Reitz
    - Frederick.Reitz@nasa.gov
    - 301 286-2516

Agenda


Welcome & Introduction
Phil Webster NCCS

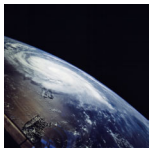
Current System Status
Fred Reitz, Operations Manager
System Issues *Utilization*
Pending Upgrades *Changes*

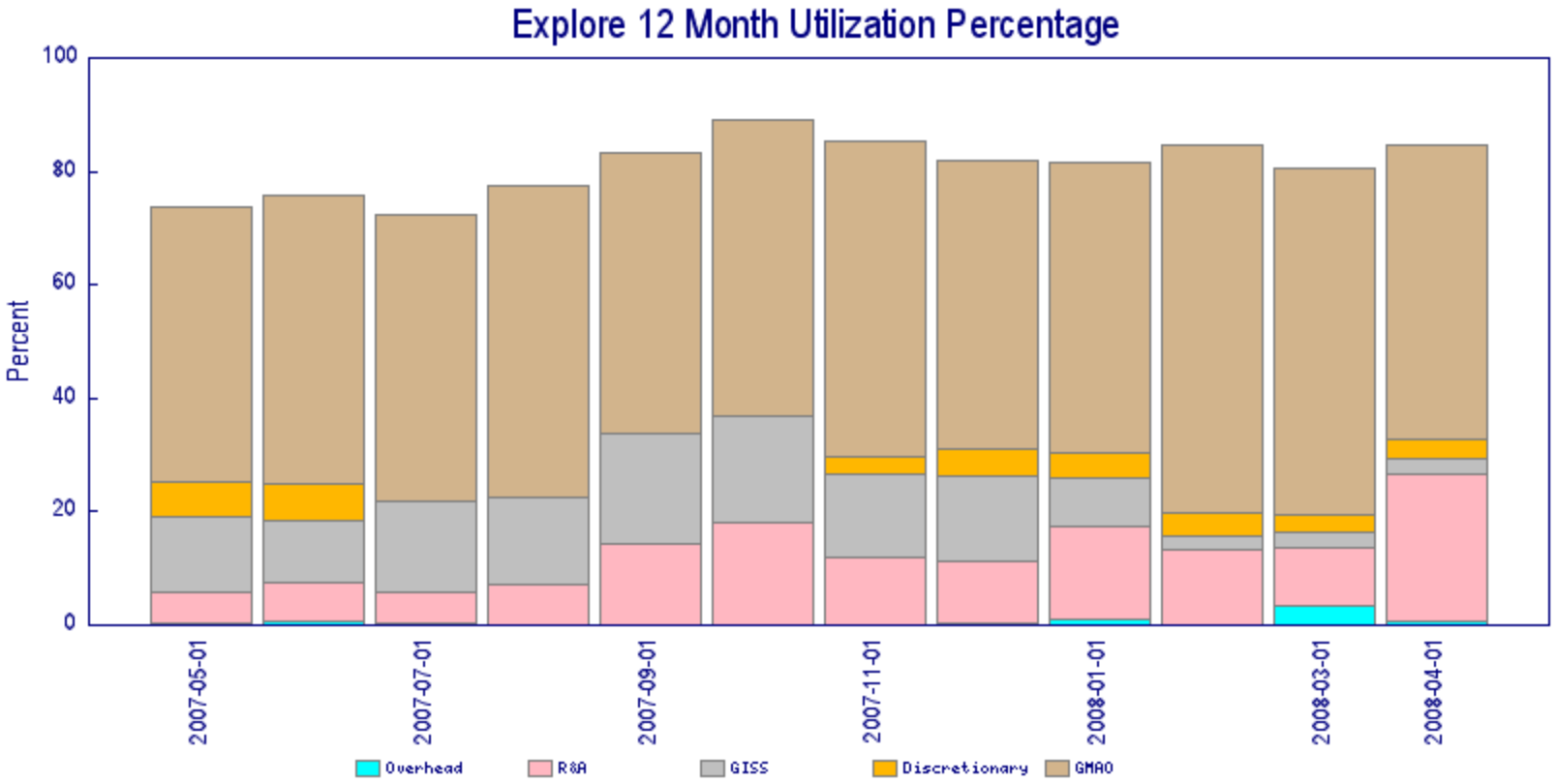
New Compute Capability at the NCCS
Dan Duffy, Lead Architect
Schedule *Impact of Discover Changes*
Storage Cluster *Architecture* *Quad Core*

User Updates
Sadie Duffy, User Services Lead
One of a Kind Data *SIVO announcements*
Allocation Updates *Changes*
Transition support

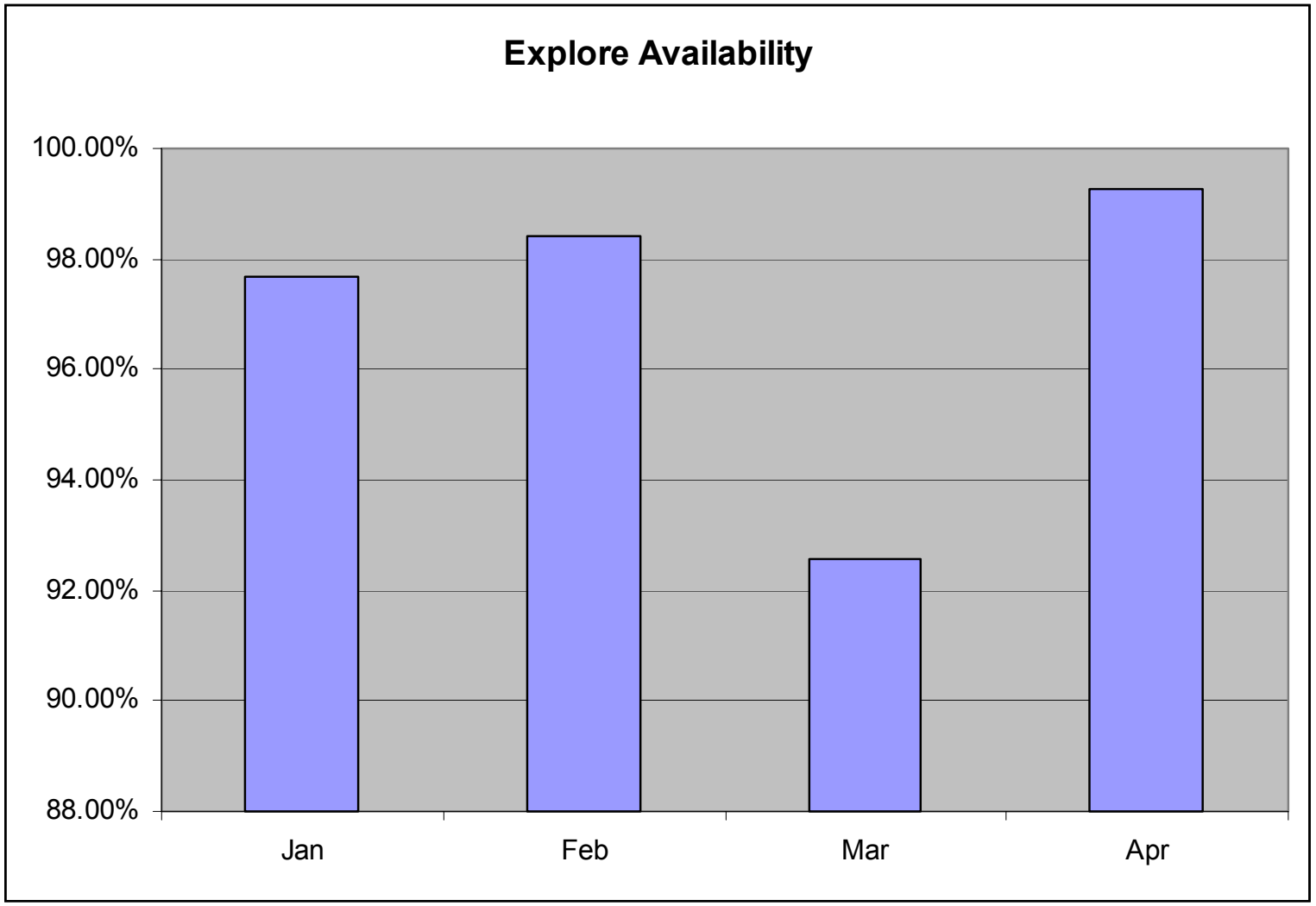
Questions / Comments
Phil Webster

Explore Utilization
Past 12 Months





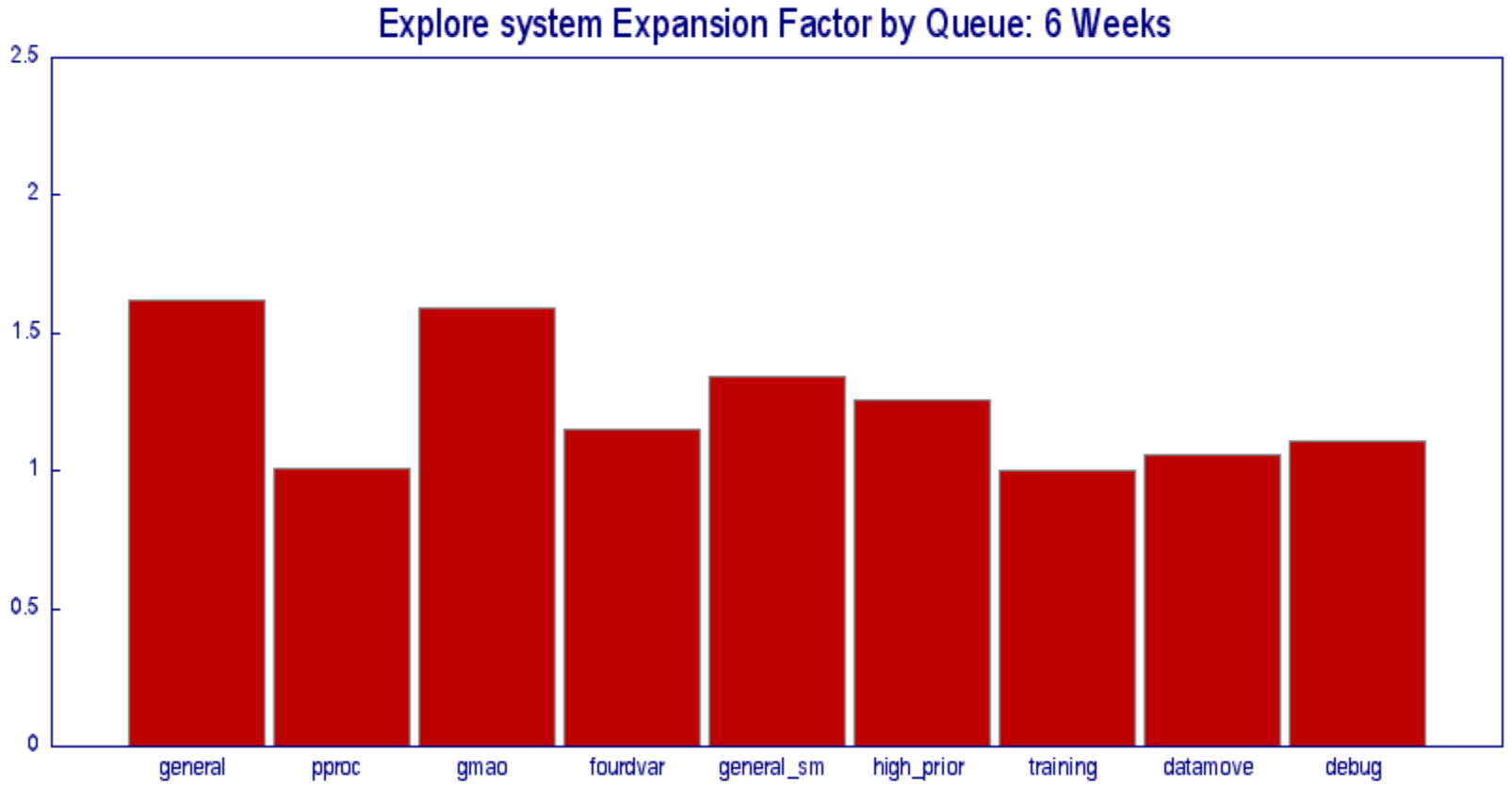
Explore Availability



Explore Queue Expansion Factor

$$\frac{\text{Queue Wait Time + Run Time}}{\text{Run Time}}$$

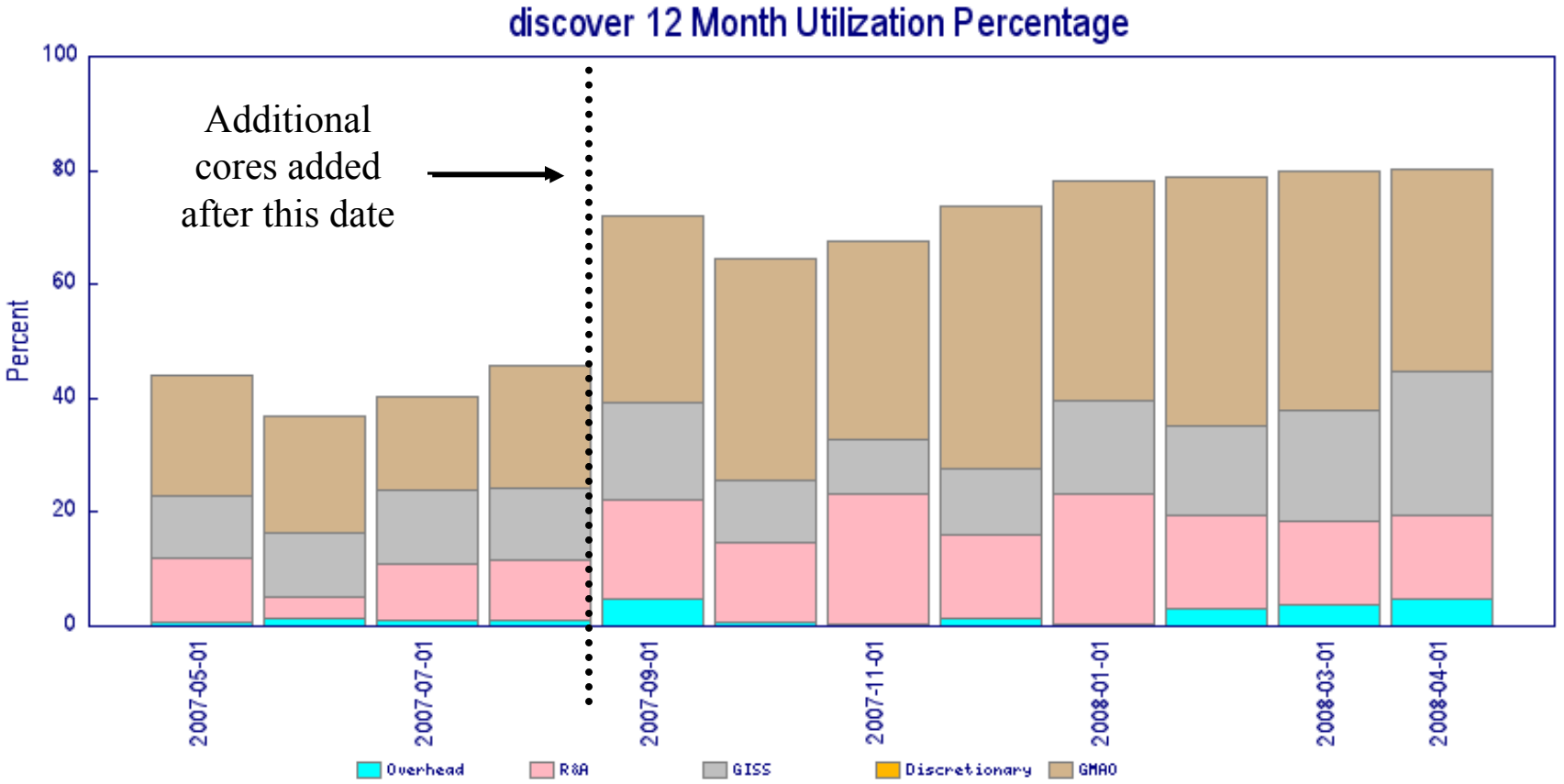
Weighted over all queues for all jobs
(Background and Test queues excluded)



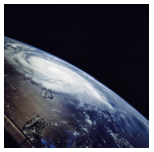
System Being Decommissioned

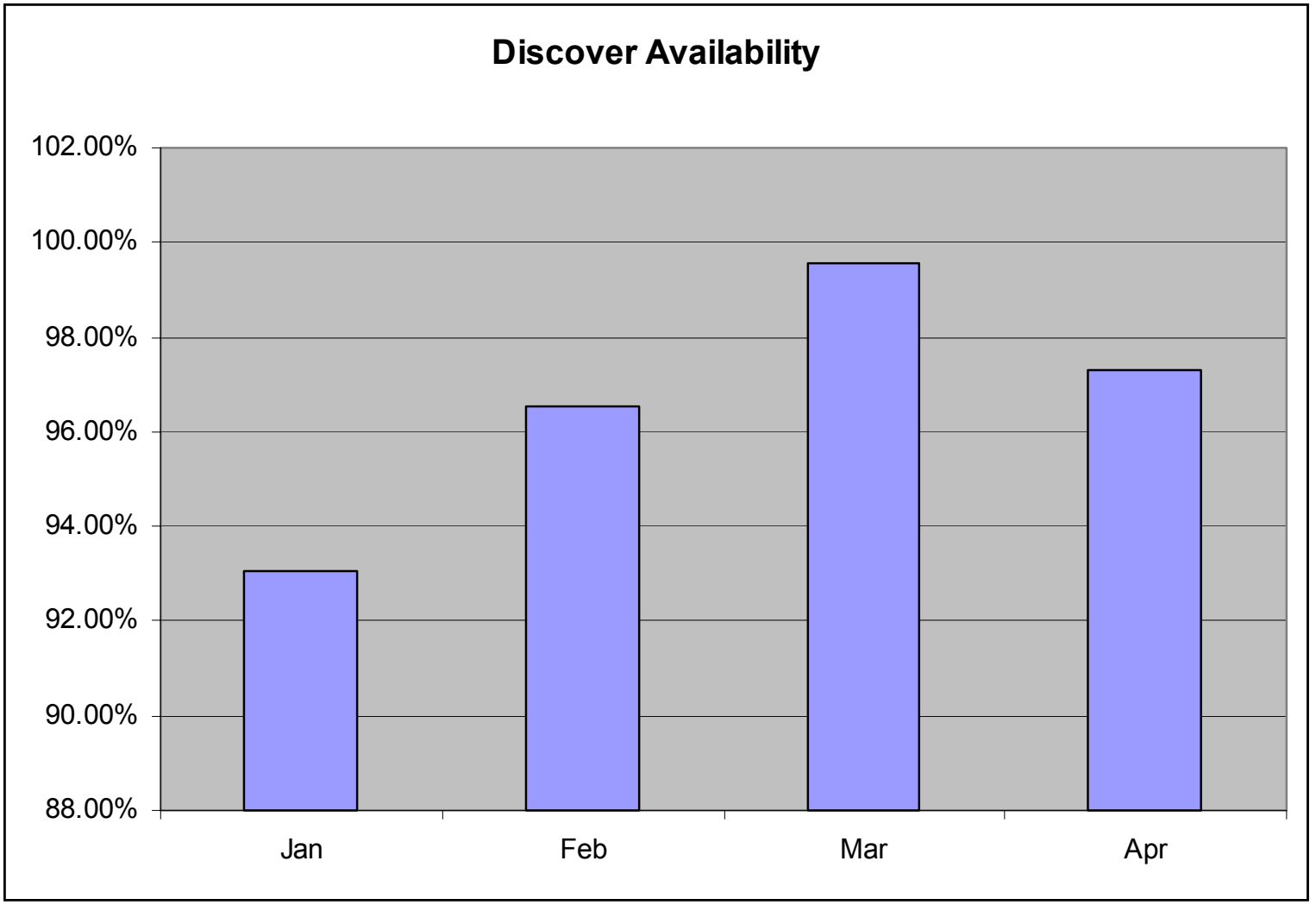
- Leased system
- System will be shut down October 1st, and system must be returned to vendor by mid October
- NCCS is in the process of negotiating the purchase of the disks associated with explore and users will be contacted with details concerning data migration (which can occur after the system leaves in September)
- Please begin porting applications to discover immediately—User Services will have additional details.
- Palm hardware will remain, however there are plans to repurpose the system for a DMF upgrade. (/home and /nobackup filesystems will be available for a short window via dirac)

Discover Utilization
Past 12 Months



Discover Cluster Availability

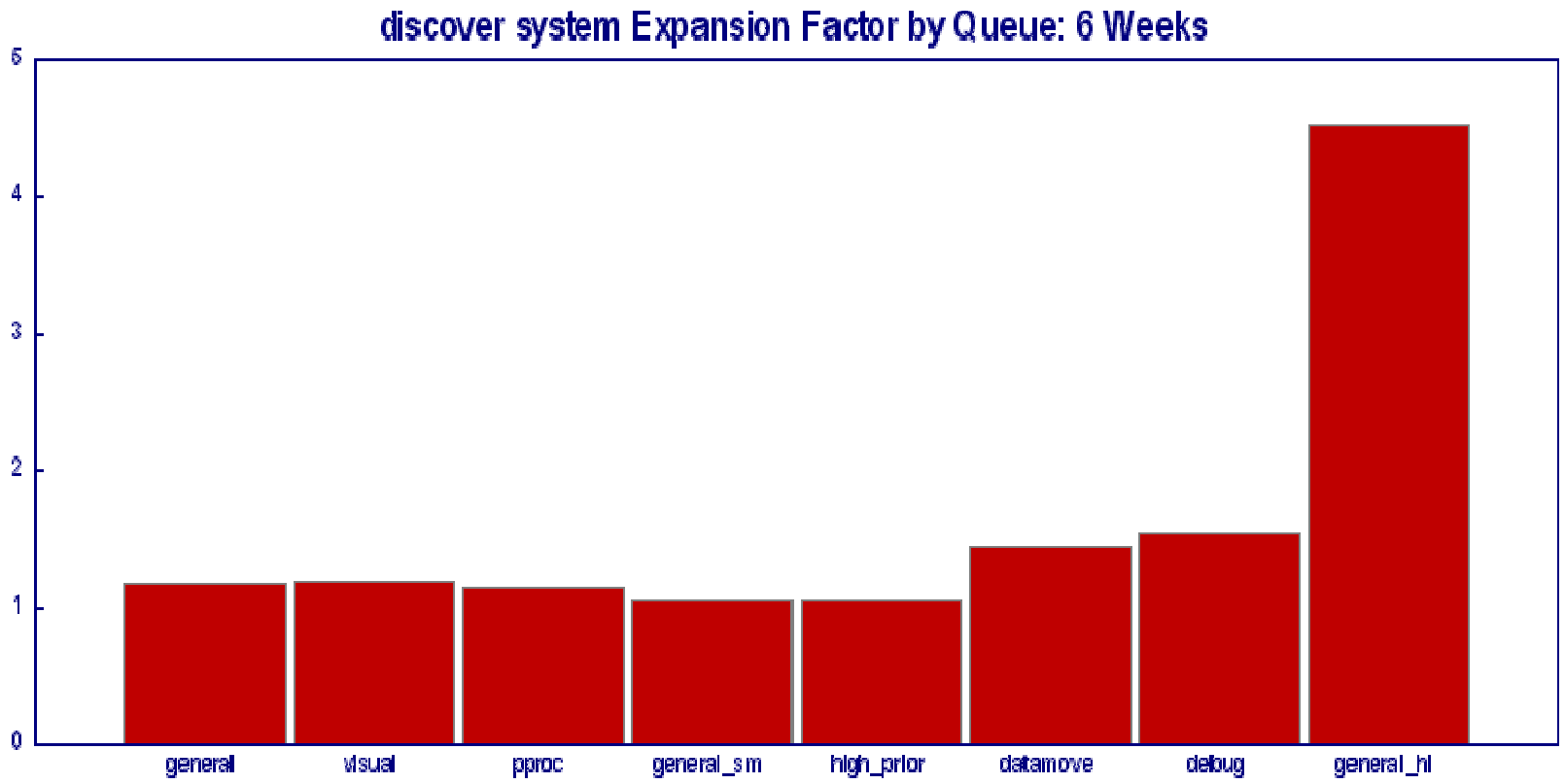


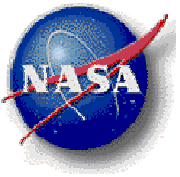


Discover Queue Expansion Factor

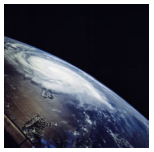
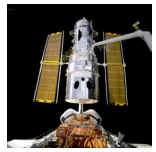
$$\frac{\text{Queue Wait Time + Run Time}}{\text{Run Time}}$$

Weighted over all queues for all jobs
(Background and Test queues excluded)



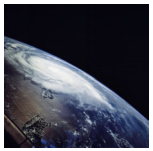
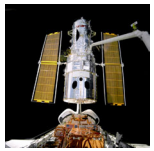


Current Issues
Discover

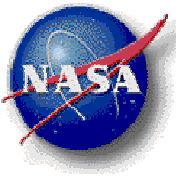




- **Swap and Memory Issues**
  - **Symptom:** Jobs either excessively swap, exhaust nodes of swap, or exhaust nodes of memory.
  - **Outcome:** Job failures and/or filesystem problems
  - **Status:**
    - Most occurrences caught via monitoring
    - System Admins work with individual users when problems occur
    - Overall frequency of problem has been reduced

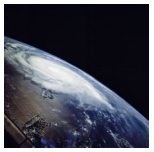
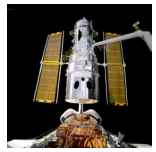
Current Issues
Discover


- **Longer Runtimes Than Expected**
  - **Symptom:** Several jobs that previously ran in 12 hours were not completing.
  - **Outcome:** Job timeouts
  - **Status:**
    - Users reduced number of simulations per job
    - Increased wall time limit for some queues
    - Identified and replaced marginal disk
    - Identified and replaced failed disk
    - Upgraded storage subsystem firmware
    - Moved data to reduce I/O contention
    - Monitoring to identify other contributing factors



Things to Remember Discover

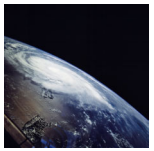
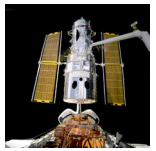




- **NEVER** use "/gpfsm/..." or "/nfs3m/..." or "/archive/g##/..." to reference data on discover. These pathnames may change at any time.

- **ALWAYS** use the following pathnames when accessing data on Discover
  - **$HOME** **for /discover/home/<userid>**
  - **$NOBACKUP** **for /discover/nobackup/<userid>**
    /discover/nobackup/projects/...
  - **$ARCHIVE** **for /archive/u/<userid>**

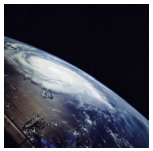
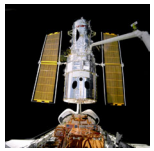
- These pathnames will always point to your data, even if the underlying filesystem or location of the data changes.

Future Enhancements


- **Discover Cluster**
  - Software OS
    - SLES 10 SP1 *Jul 2008*
  - Hardware platform – Sept 2008
  - Storage augmentation
    - Arriving June 2008
    - Filesystems, user $NOBACKUP space, project nobackup space to be moved
    - Data movement to be coordinated with users and projects to minimize impact
- **Data Portal**
  - Hardware platform – Jul/Aug 2008

Agenda


Welcome & Introduction
Phil Webster NCCS

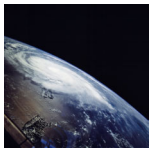
Current System Status
Fred Reitz, Operations Manager
System Issues *Utilization*
Pending Upgrades *Changes*

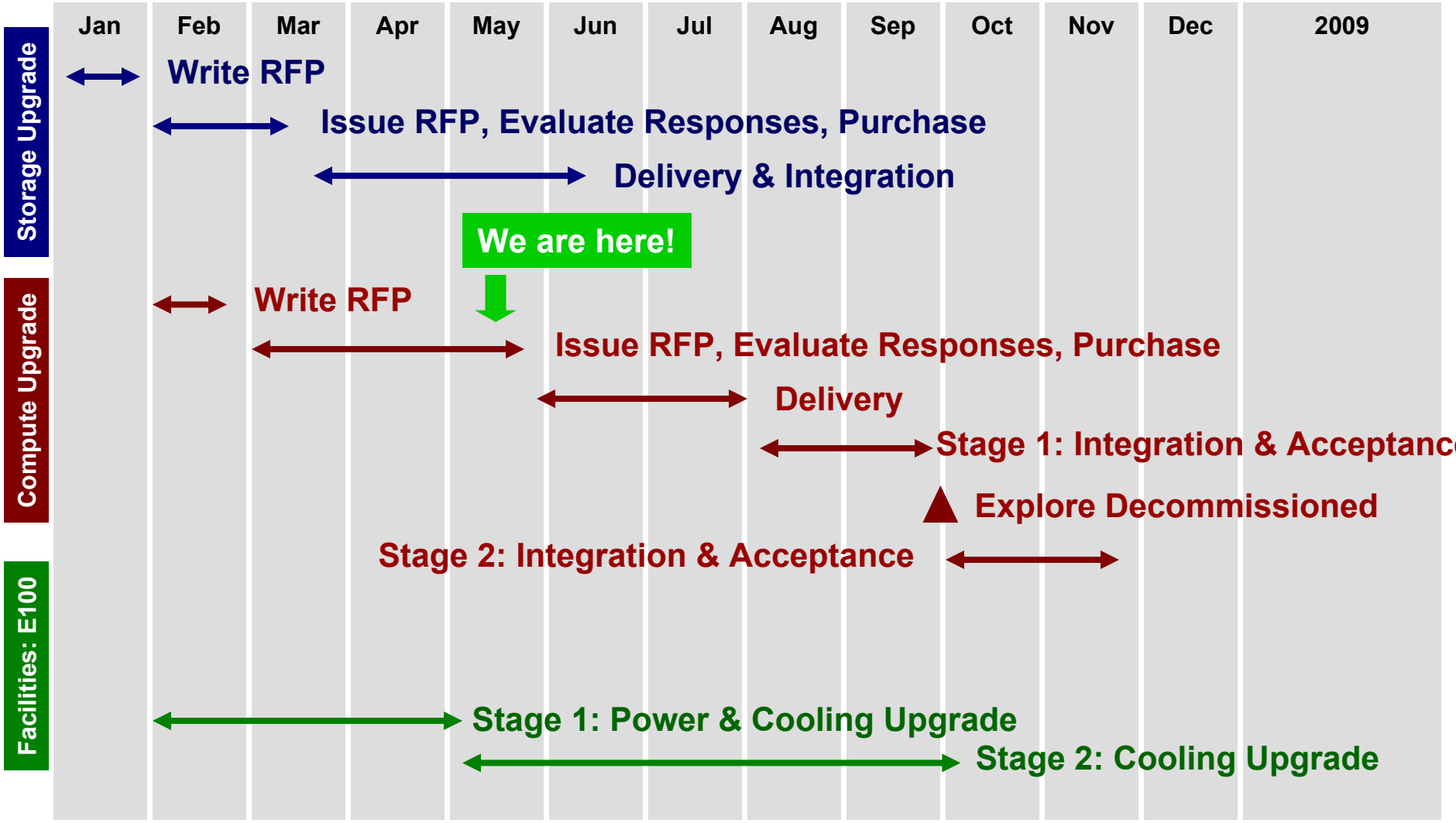
New Compute Capability at the NCCS
Dan Duffy, Lead Architect
Schedule *Impact of Discover Changes*
Storage Cluster *Architecture* *Quad Core*

User Updates
Sadie Duffy, User Services Lead
One of a Kind Data *SIVO announcements*
Allocation Updates *Changes*
Transition support

Questions / Comments
Phil Webster

Overall Acquisition Planning Schedule 2008 – 2009





What does this schedule mean to you?
Expect some outages – Please be patient

Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec | 2009

Discover Mods

- GPFS 3.2 Upgrade (RDMA)
- SLES 10 Software Stack Upgrade

Storage Upgrade

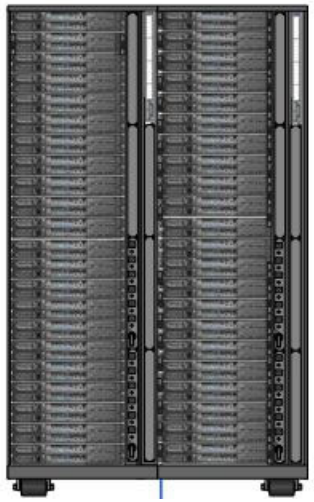
- Additional Storage On-line

Compute Upgrade

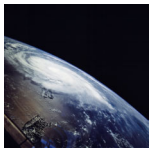
- Stage 1 Compute Capability Available for Users
- Decommission Explore
- Stage 2 Compute Capability Available for Users

Vendor Proposals and Selection

- NCCS received five (5) proposals from four (4) vendors
  - Subsequently narrowed down the field to two vendors and had subsequent negotiations with the final two
- Selected Solution
  - IBM IDataPlex solution: ~40 TF Peak
  - Very similar architecture to Discover
  - Doubled the size of a scalable unit
  - Full non-blocking bisection bandwidth for up to 2,048 nodes
  - Dual-socket, quad-core Intel Xeon nodes
  - Doubled the memory footprint (2 GB/core or 16 GB/node)
  - Lowest risk solution



More Details of the Compute Upgrade



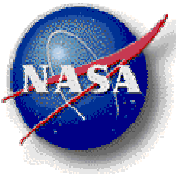
- Two (2) scalable units
  - More traditional 1U pizza box solutions
  - 2,048 cores each; 1:1 blocking within an SCU
  - 2.5 GHz Intel Quad-Core Xeon Harpertown with 1,333 MHz FSB
  - Dual-socket node with 2 GB/core
    - 8 cores per node
    - 16 GB per node
  - Single management solution
- Stage 1: Turn on the equivalent of one scalable unit (~20 TF)
- Stage 2: Turn on the rest of the compute nodes

Storage

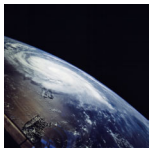
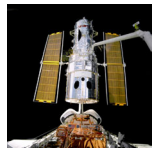
- **Base: 512c, 3.3 TF**
- **SCU1: 1,024c, 10.9TF**
- **SCU2: 1,024c, 10.9TF**
- **SCU3: 2,048c, 20.0TF**
- **SCU4: 2,048c, 20.0TF**

Going Green:

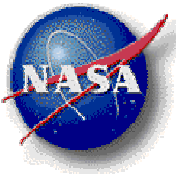
Processor	Power W	Speed GHz	GF/W
Woodcrest	125	2.66	11.7
Harpertown	50	2.5	5.0



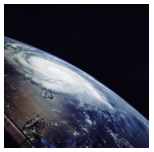
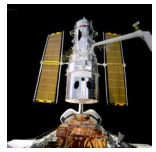
Dual-core to Quad-core
What should I expect?


- Binary compatibility
  - Unless your application is compiled with optimizations specific to the chip (-X options), your binary should work on ALL compute processors within the Discover cluster.
  - You will not have to recompile to use the new nodes.
- Floating point performance
  - Intel has made some significant improvements from the Woodcrest (currently on Discover) to the Harpertown (to be delivered in the upgrade).
  - Floating point performance should at least stay the same (even with a slower clock) and some codes will speed up.

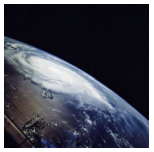


Dual-core to Quad-core
That's not the whole story...

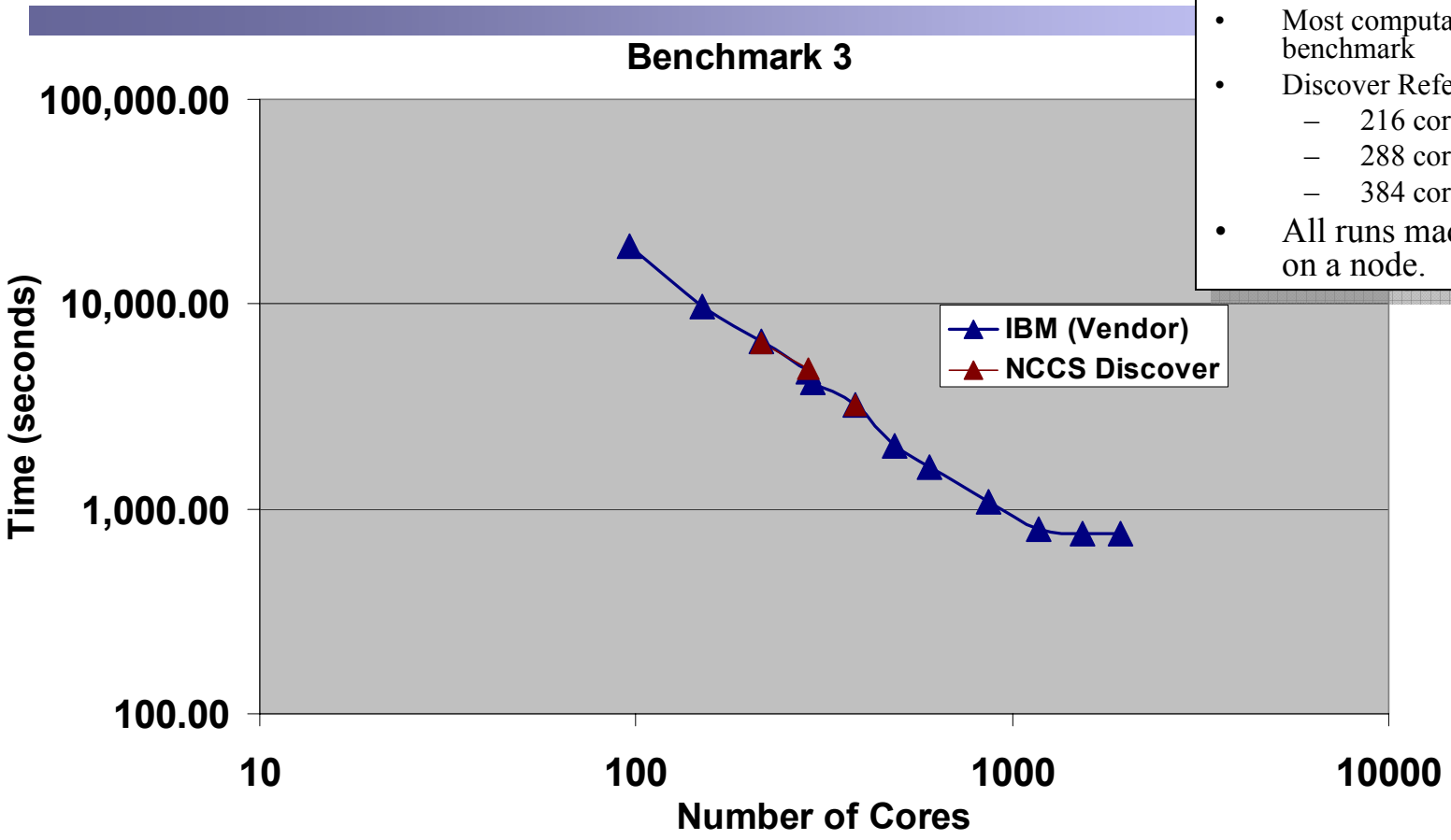




- Memory to processor bandwidth performance
  - The speed of the front side bus will not increase (1,333 MHz).
  - For the quad cores, more cores on a single chip will share the front side bus.
  - The codes where multiple processes on a single chip end up contending for memory at the same time will be affected.
  - Very application dependent.
- PBS
  - How will you select the nodes?
  - Still working that out – details to be released as soon as we can.

Cubed Sphere Finite Volume Dynamic Core Benchmark

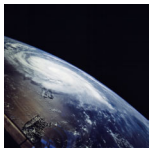


- Non-hydrostatic, 10 KM resolution
- Most computationally intensive benchmark
- Discover Reference Timings
  - 216 cores (6x6) – 6,466.0 s
  - 288 cores (6x8) – 4,879.3 s
  - 384 cores (8x8) – 3,200.1 s
- All runs made using ALL cores on a node.



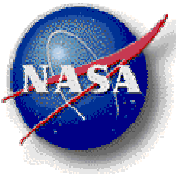
Discover and the new system upgrade should run at the same level of performance using all the cores on a node.

Software Stack

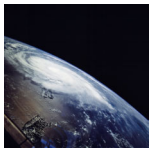
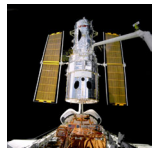


Item	Current Version	Existing Discover Units	IBM Upgrade
Cluster Manager	Clusterworx 3.4	Clusterworx Advanced	XCAT 2.0
OS	SLES9 SP3	SLES10 SP1	SLES10 SP1
MPI	Scali 5.4	Scali 5.6 Intel MPI	Scali 5.6 Intel MPI OpenMPI 1.2.5
Infiniband	Qlogic	OFED 1.3 MPI Latencies: 3 to 5 microseconds	OFED 1.3 MPI Latencies: 1 to 2 microseconds
Compilers	Multiple Versions	Multiple Versions	Multiple Versions
PBS Scheduler	8.0	8.0	8.0
IBM GPFS File System	3.1.15	3.2	3.2

User environment will look virtually identical when running across the different types of nodes.

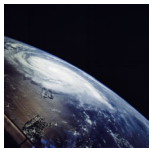
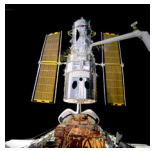


Storage Upgrade


- IBM was selected for the storage upgrade as well.
- The NCCS is going to stay with IBM GPFS for now
  - May consider alternative file systems in the future, such as Lustre
- Storage Upgrade
  - Additional DDN S2A9550
  - Additional 240 TB RAW of storage capacity
  - Low risk upgrade to both the capacity and the throughput
- NCCS is currently migrating file systems around to reduce contention on disks
  - Ultimate goal is to have a very low impact upgrade to the storage environment while increasing both performance and capacity

Agenda


Welcome & Introduction
Phil Webster NCCS

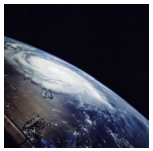
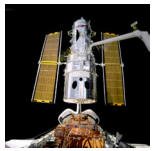
Current System Status
Fred Reitz, Operations Manager
System Issues *Utilization*
Pending Upgrades *Changes*

New Compute Capability at the NCCS
Dan Duffy, Lead Architect
Schedule *Impact of Discover Changes*
Storage Cluster *Architecture* *Quad Core*

User Updates
Sadie Duffy, User Services Lead
One of a Kind Data *SIVO announcements*
Allocation Updates *Changes*
Transition support *NAMS*

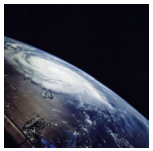
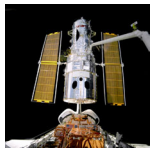
Questions / Comments
Phil Webster

Transition to Discover

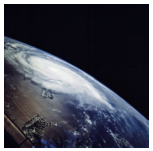
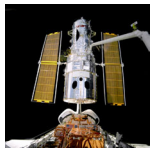

- From Explore to Discover
  - Transition Coordinators assigned to migrating teams…they will be your personal advocate during this transition (you can still get help through User Services)
  - Any team who does not already have an allocation on discover will be granted one (good until the November 1st allocation period)
    - Users from these teams will be granted access to discover
    - Your coordinator will let you know when this occurs
  - Disks containing /explore/nobackup will be retained, and will allow for time for data migration (help is available)
- From Discover to its upgrade
  - Help for utilization of quad cores is available through User Services
    - Code Porting
    - Code Optimization
- To the Discover of the future
  - We will be working with you and soliciting your input

Announcements


- SIVO Survey- The Software Integration and Visualization Office (610.3) is planning a series of lectures and hands-on training classes on high-end computing and various related topics customized for Goddard science applications. SIVO hopes to offer a subset of these topics as a condensed two week school later in this calendar year. Participants can learn a wide variety of new skills including Fortran 2003, debuggers, visualization software, and quality software development practices.
  - SIVO would like to know which topics would be of greatest interest to the community. In addition, SIVO would like your assistance in determining appropriate dates to offer these classes. Please fill out and submit the short on-line survey at the link: http://sivo.gsfc.nasa.gov/school/
- Unique Data
  - If you are storing the only copy of irreproducible data at the NCCS, you NEED to let us know!
  - Send an email to support@nccs.nasa.gov
- Updating systems status on NCCS website
  - Health check of filesystem loads, user required daemons, system load statistics and qstat
  - Targeted for the end of May 2008
- Direct access to ticketing system coming soon
  - Users will be able to log in directly to open tickets, review tickets and search NCCS knowledge base
  - Targeted for the end of May 2008

Changes to Account Processing

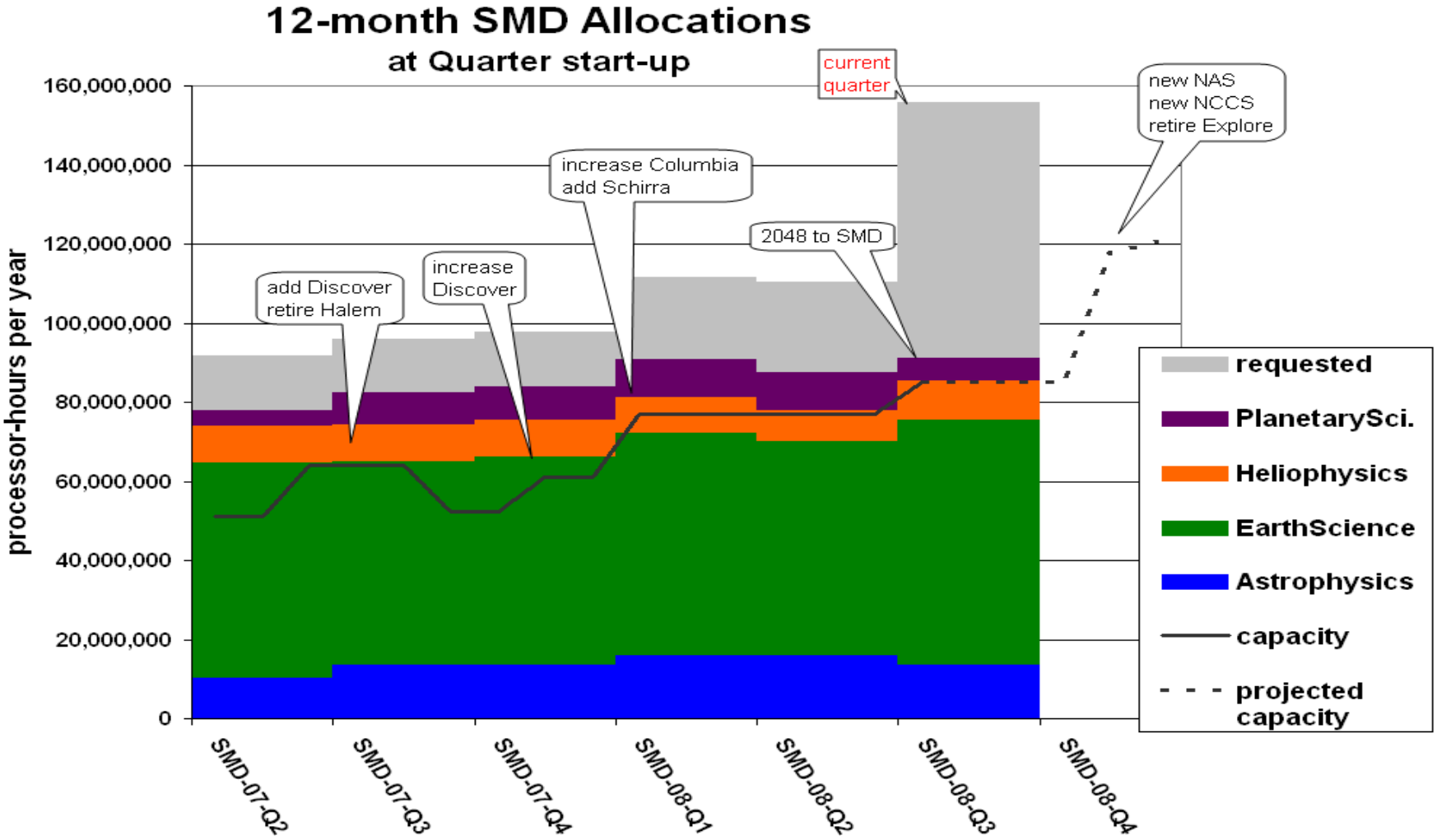




- NAMS (NASA Account Management System)
  - Establish a central agency "clearinghouse" for user accounts for various NASA resources
  - We don't have a lot of information—yet
  - All new users of any NASA IT resource, both local and remote, will have to go through the NAC-I process
  - Deadline for NCCS to migrate to NAMS by end of this fiscal year
  - We will be contacting users who must change their username to utilize NAMS services
- Foreign National Access
  - The "bad" news—all Foreign Nationals will have to go through the e-quip process with a full NAC-I
  - The "good" news—Once integrated with NAMS, when they are processed at one NASA center, they are good at all of them

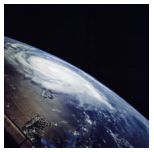
Allocations

- 135 requests were submitted in e-Books, including 5 added late, requesting a total of more than 82 M processor-hours.
- SMD capacity available for allocation across all resources was 53 M processor-hours.
- HQ SMD Science Managers considered the requests and allocated over 41M processor hours.
- Because allocation requests for Columbia totaled 70 M processor-hours but only 30 M processor-hours could be allocated May 1, we plan to allocate 28 M additional processor-hours after the NAS expansion in late summer (no additional PI action will be needed).
- If a project allocation runs low, the PI should email a request for additional hours to: support@HEC.nasa.gov.

SMD Allocations through 08-Q3

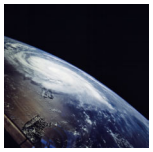


The Modeling Guru



- New "knowledge base" to support scientific modeling within NASA
  - Commercial package, customized by SIVO's ASTG, hosted by NCCS
  - Moderated discussions/forums
  - Document repository
  - Questions and support
- Goal is to leverage and share community expertise
  - Augmentation for level 2 support provided by SIVO to the NCCS
  - Topics/Communities include
    - HPC systems
    - Programming languages (e.g. SIVO F2003 Series)
    - Models: GEOS-5, GMI, modelE, etc
- Access: https://modelingguru.nasa.gov
  - Site currently in beta mode
  - Most categories publicly visible
  - Posting requires login
    - All NCCS users have login by default
    - Anyone with relevant interest can request an ID

Agenda



Welcome & Introduction
Phil Webster NCCS

Current System Status
Fred Reitz, Operations Manager
System Issues *Utilization*
Pending Upgrades *Changes*

New Compute Capability at the NCCS
Dan Duffy, Lead Architect
Schedule *Impact of Discover Changes*
Storage Cluster *Architecture* *Quad Core*

User Updates
Sadie Duffy, User Services Lead
One of a Kind Data *SIVO announcements*
Allocation Updates *Changes*
Transition support

Questions / Comments
Phil Webster

- Questions?
- Comments?