



NASA Center for Computational Sciences

NCCS User Forum

6 June 2006



Introducing the User Forum



NASA Center for Computational Sciences

- The NCCS User Forum is a quarterly meeting designed to facilitate *dialogue* with the NCCS users
- Topics will vary and may include:
 - Current NCCS services and systems
 - Suggestions for system utilization
 - Future services or systems
 - Questions and discussion with the user community
- Meeting will be available via remote access
- We are seeking your feedback



Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
- Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
- User Training and Support—Tom Clune
- User Services Updates—Sadie Duffy
- HSM Review and Best Practices—Dan Duffy
- Discussion



HPC Trends



NASA Center for Computational Sciences

- Technology Trends
 - Processor Architectures
 - Communication and Networking
 - System Software
- Architecture Trends
 - Clusters
 - Grids
- Industry Trends
 - Commercial markets drive design and features of HPC products
 - Increasing emphasis on data-intensive computing



NCCS HPC Planning



NASA Center for Computational Sciences

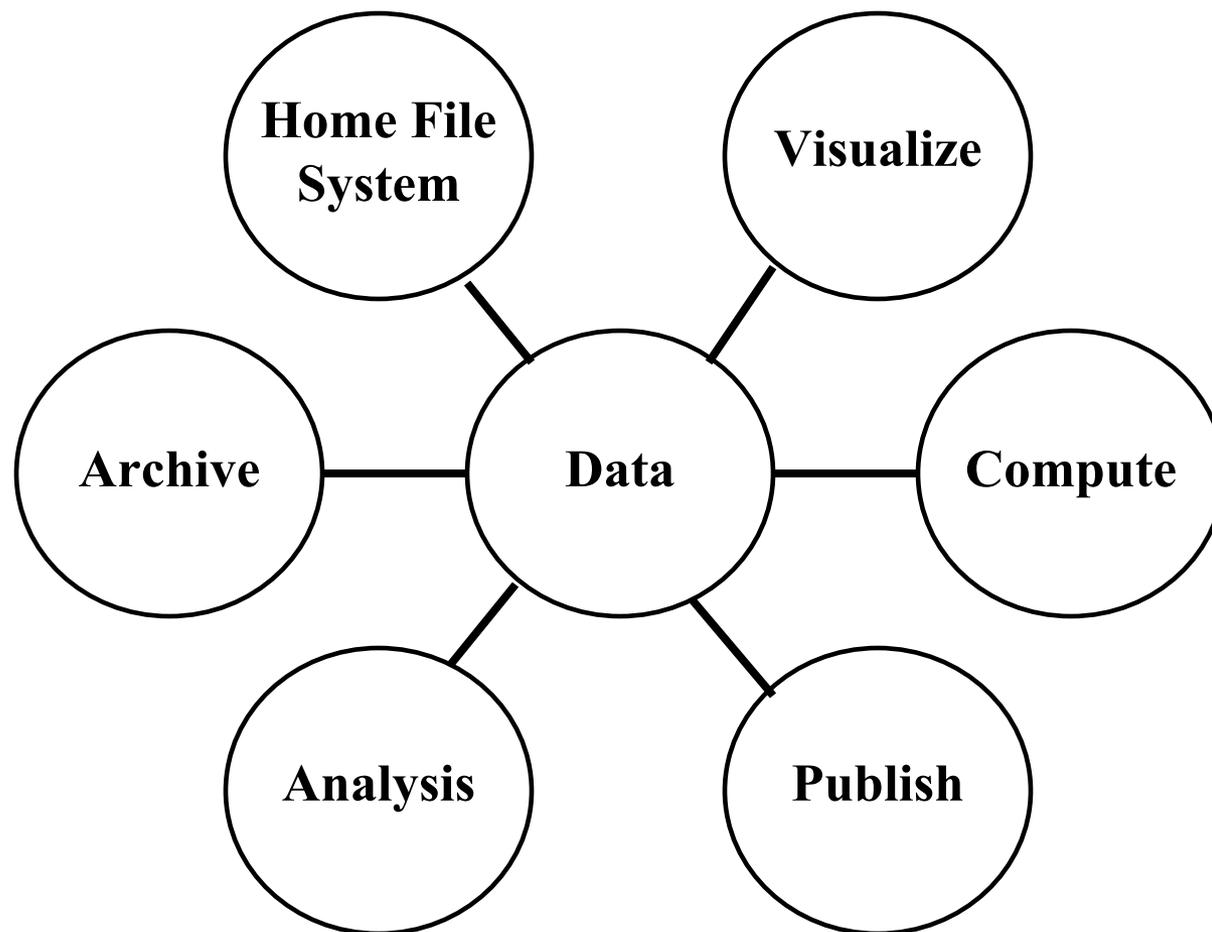
- Improving access to larger amounts of data
 - Greater on-line capacity
 - Faster access
- Establishing systems standardization
 - Common file system across all systems
 - Common operating system for all systems
 - Common user environment – compilers, libraries and tools
- Improving interaction between production and analysis activities
 - Data access from outside NCCS
 - Enhance analysis, visualization, publications opportunities



Conceptual Architecture



NASA Center for Computational Sciences





Data Centric Computing



NASA Center for Computational Sciences

Data central to scientific success and must support all aspects of research.

- Compute intensive – multiple systems
 - Standard HPC
 - Characterized by data creation
- Analysis
 - Requires access to archived data
 - Interactive sessions
- Visualization
 - Specialized hardware and software
 - Interactive sessions and remote display
- Archive
 - Long term retention of valuable user data
- Publication
 - Access to data by non-NCCS users or users remotely located
 - Web-based data services
- Home files systems
 - All users files accessible across entire computing center



Acquisition Strategy



NASA Center for Computational Sciences

- Evaluate prototype
 - Application performance
 - Ease of installation and integration
 - Manageability
- Decision to build-out
 - Revisit architecture
 - Revise procurement



Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
- Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
- User Training and Support—Tom Clune
- User Services Updates—Sadie Duffy
- HSM Review and Best Practices—Dan Duffy
- Discussion



Goals for the Cluster Environment



NASA Center for Computational Sciences

- Minimize the learning curve of the user
- Create a consistent user environment between existing systems and new cluster
 - SUSE Linux O/S
 - Intel Processors and Compilers
 - PBS
 - Modules
 - Home Directories



Commodity AND Stable



NASA Center for Computational Sciences

- Commodity Components
 - Lower cost solution
 - Xeon Dempsey processors
 - Very similar to the DoD TI-06 system being installed at ARL (Aberdeen, Maryland)
- Stable and Highly Maintainable
 - IBM GPFS chosen as the most stable file system for a cluster environment
 - Good support and training through LNXI and all partners
 - On-site parts cache
 - Training the NCCS team to be the first level of support for both hardware and software



Discover Architecture



NASA Center for Computational Sciences

- Base Unit (3.3 TF Peak)
 - 3.2 GHz Intel dual core Dempsey
 - 128 Nodes/256 Sockets (all compute)
 - GPFS & 60.0 TB DDN (raw)
- Interconnect
 - SilverStorm Infiniband
 - Full bisection bandwidth within a single unit
 - Blocking inter-unit connectivity
- Scalable Compute Unit (SCU) (6.6 TF Peak) – representative architecture
 - 3.2 GHz Intel dual core Dempsey
 - 256 Nodes/512 Sockets (all compute)
 - GPFS & ~90 TB DDN (raw)



Software Stack



NASA Center for Computational Sciences

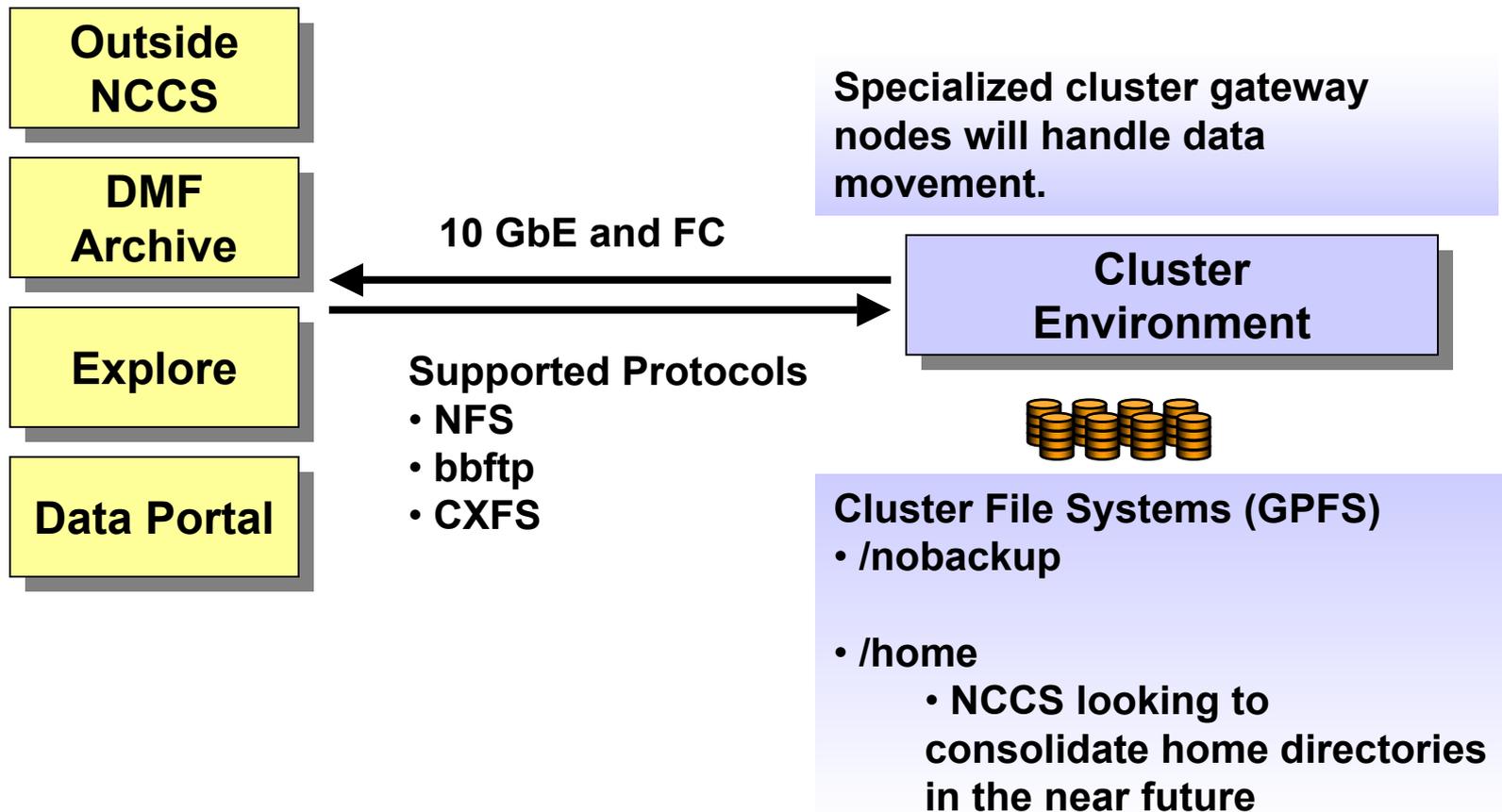
- 3rd party software included with base unit
 - Intel Compilers (5-pack), Vtune, Trace Analyzer, Trace Collector, Math Kernel Library
 - PGI CDK (5-pack)
 - MVAPICH (through SilverStorm)
- NCCS looking into the following additional software packages
 - Intel MPI and Cluster OpenMP
 - IDL for some of the visualization and analysis nodes
 - Matlab for some of the visualization and analysis nodes
 - TotalView (two concurrent runs up to 64P)
 - NAG Fortran
 - Absoft Fortran
 - FPGA software tools
- Variety of open source tools



Data Movement Into and Out of Cluster



NASA Center for Computational Sciences



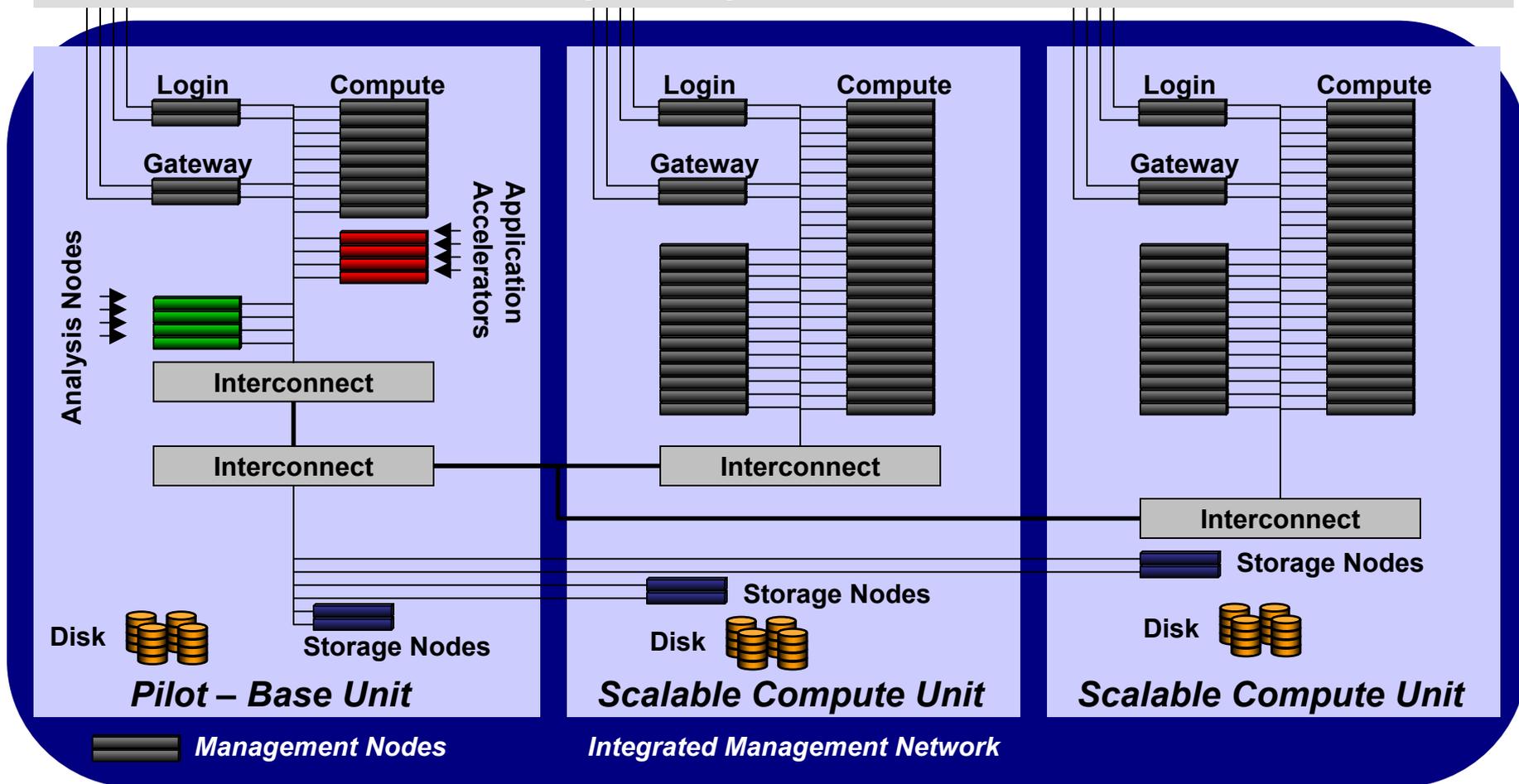


How It All Fits Together



NASA Center for Computational Sciences

GigE or 10 GigE NCCS LAN Switch

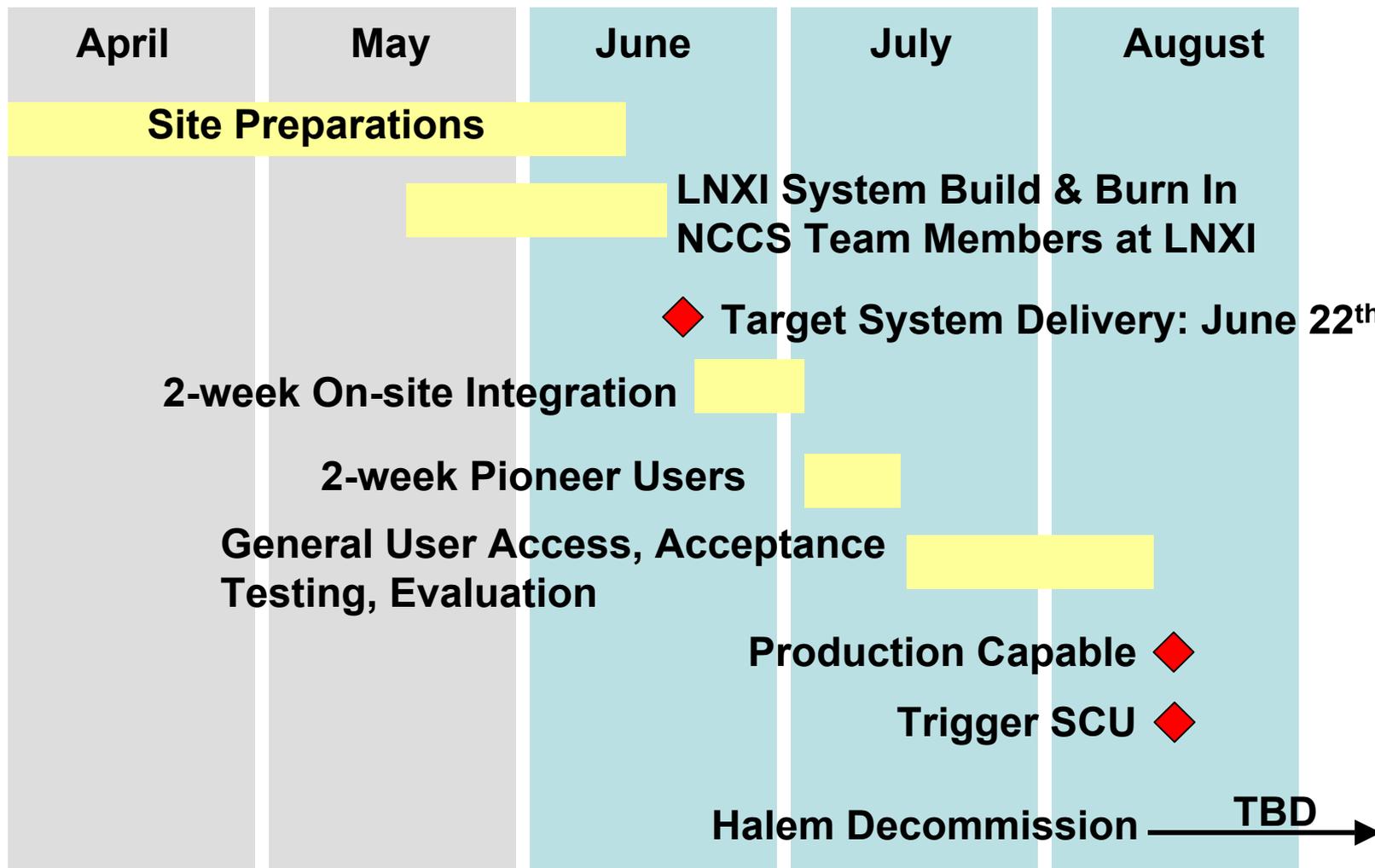




High Level Schedule: Integration



NASA Center for Computational Sciences





Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
- Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
- User Training and Support—Tom Clune
- User Services Updates—Sadie Duffy
- HSM Review and Best Practices—Dan Duffy
- Discussion



SIVO - Code 610.3



NASA Center for Computational Sciences

- Software Integration and Visualization Office (SIVO)
 - Office headed by Mike Seablom
 - Multifaceted support for NASA's high-end computing and scientific software
- SIVO Components
 - Advanced Software Technology Group (T. Clune)
 - Level II help-desk application support
 - Training, Porting, Optimization
 - Scientific Visualization Studio (H. Mitchell)
 - Education and Outreach (D. Herring)



Training for Discover



NASA Center for Computational Sciences

- User Training Classes
 - Commence early July 2006
 - Repeated on demand
- Format
 - Separate sessions on various topics
 - Skip classes that are less relevant
 - Avoid “killing” entire day
 - Wireless technology - (bring a laptop?)
 - Hands-on classes, where appropriate/possible
 - Specialized user training opportunities
- Sample topics
 - General Info - basic architecture, access, filesystems, etc.
 - Detailed Architecture - the nitty-gritty
 - Batch Queues
 - MPI optimization
 - **Hands-on porting sessions**
 - Debuggers and Profilers



Other Resources for Discover



NASA Center for Computational Sciences

- Technical documentation
 - ASTG and the NCCS will develop on-line technical information
- Technical Online Forums
 - Leverage experience of peers
 - Searchable database of related issues and their resolutions
 - ASTG will send out a general announcement as forums are created



Additional Training



NASA Center for Computational Sciences

- ASTG will provide training to maximize NASA's investment in software and science
- What other classes/presentations would be valuable? .
 - Tuning and Analysis (already planned)
 - MPI, OpenMP, Co-Array Fortran?
 - CVS, Subversion?
 - How to build a better Makefile?
 - Best Software Practices?
 - F90/95, F2003, C++?
 - Future directions in HPC?



Other SIVO Services



NASA Center for Computational Sciences

- ASTG exists to support users and developers of high-end scientific software
- Consultation is available for any of the following activities:
 - Optimization (serial, algorithmic, I/O, etc)
 - Parallel implementations
 - OpenMP
 - MPI
 - Use of ESMF
 - Reengineering and/or design/implementation
 - Benchmarking



Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
- Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
- User Training and Support—Tom Clune
- User Services Updates—Sadie Duffy
- HSM Review and Best Practices—Dan Duffy
- Discussion



Cluster Access



NASA Center for Computational Sciences

- User accounts will be added during System Acceptance phase, starting mid July 2006
 - Systematic addition of all Halem users (contact User Services for passwords)
 - Explore users must request cluster access (email User Services for account)
- Open user access on the cluster to facilitate Halem code migration
- Halem to be decommissioned on December 31, 2006
- Cluster will be a resource in the FY07 SMD allocation process



User Services



NASA Center for Computational Sciences

- We are your contact point for **all** service requests
- Help desk available 8am to 8pm EST
 - 301-286-9120 or support@nccs.nasa.gov
- Weekly User Teleconference
 - Discuss current issues, provide updates and answer questions
 - 1:30pm Tuesday (starting June 13th)
 - 866-903-3877 passcode 6684167
- User Portals available
 - Update contact information
 - Metrics on systems used
 - Request to be added to additional groups
- For additional support services contact
 - Sadie Duffy, User Services Lead, 301-286-2954



Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
- Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
- User Training and Support—Tom Clune
- User Services Updates—Sadie Duffy
- HSM Review and Best Practices—Dan Duffy
- Discussion

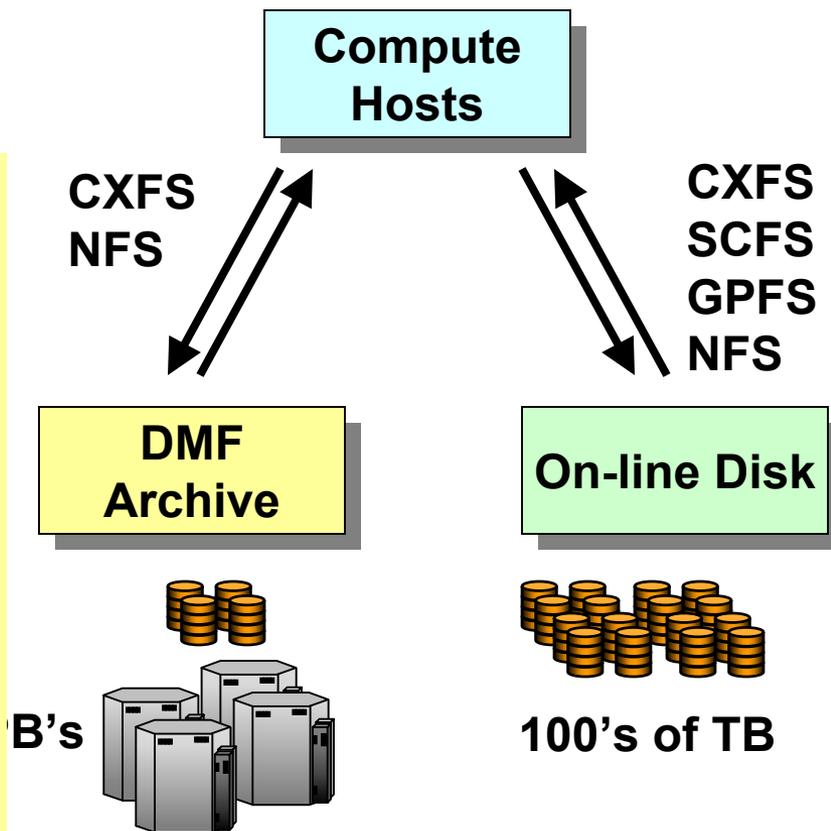


Tiered Storage Environment



NASA Center for Computational Sciences

- Compute hosts have various access bandwidths to the DMF Archive
 - Fast access via CXFS (GB/s aggregate)
 - Slow access via NFS (MB/s aggregate)



- “Permanent” Files
 - Final Products
 - Years
- Write dominated
- Minimize reads from tape
- Slow to fast access

- “Itinerant” files
 - Minutes to days to months
- Examples
 - /nobackup
 - /lscr
 - /dataportal
 - /datastage
 - /home
- Fast access (typically)
- Read and write often
- Not backed up (generally)
- User Managed



Recommended Usage Patterns



NASA Center for Computational Sciences

- On-line Disk Storage Environment
 - Fast reliable disk to support a high turn over rate of data
 - Scratch files for batch jobs, restart files, etc.
 - Data needed for post processing and analysis
 - Not backed up, with the exception of /home
 - User or project managed
- DMF Archive
 - Store final products
 - Write dominated
 - Minimize reads, especially from tape
 - Low turn over rate for files



Future Archive and Disk Modifications



NASA Center for Computational Sciences

- Larger /nobackup file systems
 - Create project oriented /nobackup file systems
 - On the order of 150 TB for Explore
 - Target 30 to 60 day data hit rate
- Increase primary DMF disk cache
 - On the order of 100 TB for newmintz
 - Target 30 to 60 day data hit rate from disk for recalls from DMF
 - Assumes that the /nobackup file systems lessen the need to read from tape
- Upgrade SAN infrastructure to 4 Gb
 - Double the SAN peak bandwidth
- New tape drives and media
 - Faster and larger tapes (500 GB tapes at 120 MB/sec)



Agenda



NASA Center for Computational Sciences

- Introduction and Technical Direction for the NCCS—
Dr. Phil Webster
 - Linux Networx Cluster (Discover) Overview—Dan
Duffy
 - *Explore, Discover and Understand*
 - User Training and Support—Tom Clune
 - User Services Updates—Sadie Duffy
 - HSM Review and Best Practices—Dan Duffy
- Discussion



- Backups



Benchmarking Strategy



NASA Center for Computational Sciences

- Use HPC Industry standard benchmarks
 - NAS Parallel Benchmark
 - HPCC (HPC Challenge)
 - Low level chip benchmarks?
 - Switch benchmarks?
- NCCS developed I/O benchmark
- User applications were not used
 - Effort to prepare benchmarks
 - Selection of a few representative applications
 - Port code to multiple platforms
 - Develop test cases and validation tests
 - Effort by vendors
 - Additional porting
 - Code optimization
 - Lack of representative systems to run benchmarks
 - Results are not representative
 - Systems are not in actual operations
 - I/O is often ignored
 - Code changes do not reflect changes a user would make
 - Applications change during this process



Archive Capability & Capacity



NASA Center for Computational Sciences

- Current data creation rate
 - ~2.2 TB/day (single copy) write
 - ~1.2 TB/day read
- Increasing archive requirements
 - MAP 06 starting in July 2006
 - Base Unit (equal to Halem, ~0.7TB/day) starting in July 2006
 - MERRA starting in Aug 2006
 - Additional scalable units to take the place of Halem and grow to substantially larger peak processing (starting in Oct 2006)
- Total Capacity = 3.74 PB
 - Existing: 2.74 PB
 - New: 2,000 x 500 GB = 1.0 PB
- If projections are on target, additional disk will be needed by Jan 2007
 - NCCS will constantly monitor tape holdings and adjust the target date for additional tape media