Project Columbia Data Management Overview

Alan K. Powers
NASA Ames Research Center
NASA Advanced Supercomputing Division
MS 258-6, Moffett Field CA 94035-1000
Phone: +1-650-604-3991    FAX: +1-650-604-1777
E-mail: powers@nas.nasa.gov

Presented at the THIC Meeting at the Sony Auditorium,
3300 Zanker Rd, San Jose CA 95134-1940
February 28 – March 1, 2006
U.S. and NASA High-end Computing Prior to Columbia

• Japan’s Earth Simulator held the world record for 2 1/2 years
  – Funding-to-Production required ~7 years
  – Build and Integration alone took 4 years
  – Security model: Users required to be onsite with data in hand
• “Leadership” Systems required years of planning/development
  – Livermore National Lab (IBM Blue Gene)
  – Oak Ridge National Lab (Cray X1/X1E)
  – Sandia National Lab (Cray Redstorm)
• NASA HEC funding declining over 10-year period
  – Fell from Top 10 to 59th ranking in Top 500 list
  – Surpassed by systems in several foreign countries
    • (China, Korea, Brazil, Canada)
  – NAS division at risk of losing all funding
Previous Experience - Columbia Could NOT be Built in 120 days

• 2003 - SGI 512-processor SSI at NAS
  – Installation Start to Production-Ready required 75 days
  – First Linux-based (SSI) System of this size to be installed
  – Using this process would take 1500 days

• 2004 - Install 20-512p systems in 120 days
  – 12-A3700 and 8-A3700bx2 (Not yet manufactured)
  – Still had production systems and service users
  – Still had to deinstall 13,000 sq ft of systems
  – Remove all old cables under raise floor (6+ pallets)
  – While upgrading plumbing, power, & cable plant
  – 9 bx2 systems arrived in a 10 day period
  – Still had to add users to new systems during build out
  – Benchmark all 10,240p system - hoping no component fails
Needed Everything Perfect

• Team Developed Highly-compressed Rapid Deployment Process for a 512p node
  – Hardware Roll-in (12-16 Racks), cable up, power up (day 1)
  – Hardware Memory/Processor/Network Detail Diagnostic (day 2)
  – Operating System/Job Scheduler/File System Installation and Security Acceptance (Day 3)
  – System Verification and Validation Testing with Major Application (day 4)
  – Users Given Full Access for Production Runs (day 5)

• The new 5 day streamline process had to be done in parallel to meet 9 new BX2 systems arriving within a 10 day period.
But nothing is perfect (Overcoming obstacles)

- 40% RIF in staff + 0% raises
- 2 day planned power outage - taking 6 days - Major 18” water pipe broke - creating a hole big enough for a car to fit in.
- 10 Gige delayed shipping from Cisco for 30 days
- Power whips could only be added when PDU was off.
- Elevator used to move hardware to 2nd floor broke.
- Fork truck used to move hardware downstairs broke.
- Infiniband rack delayed every day for about 2 weeks.
- Mellonox Infiniband cables 2/3 bad and was discovered after installation. This caused a shortage of cables.
- Initial design used “certified” 19m cables. Redesign to use 17m cables.
- SGI had a list of 800 items that would prevent them from meeting the delivery schedule.
- IB - HBA had design problem impacting benchmarking - at end of install - caused 3 parallel attempts to finish benchmarking.
The Columbia Construction Project
People, Planning, Power, Plumbing and Processors = SCIENCE

Internal Networks
Internode Comm – Infiniband
Data Transfer – Gigabit Ethernet
Hi-Speed – 10 Gigabit Ethernet

Systems – SGI Altiix 3700 and 3700-BX2
Processors – 10,240 Intel Itanium 2
Global Shared Memory – 20 Terabytes

Front-End – SGI Altiix 3700 (64 proc)
Online Storage – 440 Terabytes RAID
Offline Storage – 6 Petabyte STK Silos

3/1/2006
Columbia Nodes, Network, IO Paths

**Front End**
- 24p, 24p, 20p Altix 3700

**Networking**
- InfiniBand Switch (288port)
- InfiniBand Switch (6x32port)
- InfiniBand (4 per 3700, 8 per 3700Bx2))
- 10GigE Switch 32-port
- 10GigE Cards (1 Per 512p)
- GigE Switch 384-port, 96-port
- GigE (up to 12 per 512)
- Altix 3700Bx2 2048 Numalink Kit

**Compute Nodes**
- Altix 3700 12x512p
- Altix 3700Bx2 8x512p
- C1-14,16 - I2 1.5 Ghz 6 MB cache
- C15-17-20 - I2 1.6 Ghz 9 MB cache

**Hyperwall Access**
- 16p Altix 3700 (limited access)

**Storage Area Network**
- Brocade Switch 2x128port

**Storage (440 TB)**
- FC RAID 8x20 TB (8 Racks)
- SATARAID 8x35TB (8 Racks)
World-Wide Recognition of Project Columbia

- International Press Coverage - 300 articles
- Intel full-page ad in Wall Street Journal honoring Columbia
  - October 20th, 2004
- HPCwire Editor’s Choice Award
  - “Most Innovative Implementation”
  - “Best Collaboration of Government and Industry”
- Government Computer News Annual Gala Agency Award (above photo)
  - Prestigious Government Information Systems Organization
  - Black-Tie Dinner and Award Ceremony
  - 132 Project Nominations considered, resulting in 10 awards
  - Attended by over 1000 people representing all branches of the federal government
Columbia Web Status Page

• http://www.nas.nasa.gov/cgi-bin/nas/status
Moving Data to/from Columbia

- **Transfer Data Policy**
  - Small amounts of data sent to Columbia, but with 2 factor authentication via SFEs up to 10MB/s
  - Columbia nodes push data to external hosts at rates of 1-300 MB/s
  - Data can transferred between Columbia/HEC nodes (30-300 MB/s)
  - Columbia stages data to/from trusted storage area using single factor authentication
  - Soon to have automated way (IA, batch and cron jobs) to transfer data directly to/from Columbia nodes using a new “temporary” ssh key to move data at 200+MB/s
Users Use of Compute Nodes Changing

• From
  – Each 512 node has a ~10 TB XFS local filesystem.
  – Users need to move data to use a different node.
  – At end of batch job - network copy (slow) to archive system

• To
  – Set of nodes (8+) has several ~15 TB Cluster filesystems and ~10 TB XFS filesystem.
  – Users can use any node in the domain without moving their data
  – Compute nodes can be moved between domains
  – Soon - End of batch job - SAN copy to archive system
DLM Upgrade Needed

- 1.9 PB (lou) + .6 PB(susan) / 10 PB capacity
- Maximum daily transferred (lou) - 12 TB
- Average daily transferred - 6 TB
- 85% writes - 15% reads to/from tape
- Files less than 1MB is 90% Files and 2% space.
- Total DLM # Files - 44 Million
- Total DLM RAID archive space - 20 TB & 10 TB
- Vis/pre/post processing
- B4 Columbia - 6/04 - 30 TB/Month
- Columbia - 12/05 - 190 TB/month, 6+x growth
Columbia’s Future DLM Design Goals

• Have a stable, reliable archive system.
• Backup of system data on a separate host from user data.
• Be able to test new OS and DMF/TMF on separate host from production archive system and share transports.
• Move data near peak speeds to and from Columbia hosts and WAN to/from Archive fs & transports (Balance IO).
• For high priority users have IA support for pre/post processing (visualization) of very large datasets (> 1TB).
• Use existing O3k hardware (from chapman) + licenses.
• Upgrade to new host hardware 1-1.5 years after new configuration is in production.
• About 20 days of disk cache (20* daily R/W average)

Able to create a temporary scratch FC RAID filesystems.
Columbia’s Future DLM Requirements

- Support storing 11 TB/day (133.5 MB/s) with 44 TB/day (534.0 MB/s) peak per data copy.
- Support retrieving 2 TB/day (24.3 MB/s) with 4 TB/day (48.5 MB/s) peak per data copy.
- Support 220 TB DMF RAID cache, 50 TB RAID cache, and 150 TB MAID(?) with IO requirements of 8.00 GB/s
- Support 2-10 Gbits/s (2.5 Gbytes/s) across the network, filesystems and tape transports (balance IO).
- Support 6 gige interfaces (.75 GB/s).
- Support 66% growth in IO bandwidth and capacity/yr
- Backups of filesystems in less than 4 hours (2-3 Million inodes/filesystem).

Support compilers and visualization tools
Future DLM Systems, July 2006

- Lou’s - 96p/48, **24p, 24p** GB Origin 3400 (Users)
  - IO requirements need ~12 Cbricks or 48p
  - Use about 48p with CPUsets for IA pre/post processing (5 high priority vis projects)
  - ~270 TB FC RAID
  - 20 (8, 6, 6) STK T10000 and 6000 Media
  - Archive system be a CXFS client in each domain
- Susan’s 28p/14 GB Origin 3400 (Backups)
  - 25 TB FC RAID
  - 6 STK 9940B + 10 more after Lou’s upgrade
Future DLM Will Need Visualization Capabilities

- Time consuming and problematic to move huge amounts of data.
- Desktops, workstations are not designed to visualize or analyze huge amount of data effectively.
- WAN and some LAN not designed to move huge amounts of data.
- Paradigm Shift: from a DLM only to DLM+Vis
- I/A access to a DLM can add value to the data.
Transferring 5 TB of Data; 10x-80x Improvement

- **WAN, 80% Users**: 20.23 days
- **LAN, 20% Users**: 8.09 days
- **SAN, 100% Users**: 4.05 days

### MB/s

- 3 days: 3 MB/s
- 7.5 days: 8.09 MB/s
- 15 days: 4.05 MB/s
- 30 days: 2.02 MB/s
- 45 days: 1.35 MB/s
- 60 days: 1.01 MB/s
- 120 days: 0.51 MB/s
- 240 days: 0.25 MB/s

**3/1/2006**
Archive System Helped Enable ECCO to Visualize Ocean Model

- Turnaround time improved by 50x
- Users not able to remotely digest the amount of data being generated.
- Needed to have online ~ 4TB of data at once.
- I/A visualized data directly on archive system, avoid transferring data to remote/other systems. Huge time saver.
- Wanted More! Later added 10TB to compare 2 different models
Future NAS Evaluations and Upgrades

• 12/05 Evaluate Panasas - 80 TB (40 Gige) & IB RAID
• 1Q/06 Deploy 3 CXFS domains, 2Q/06 again with FC4
• 2Q/06 Evaluate Lustre P1- 50 TB (6 Gbytes/s)
• 1Q/06 Acquire 500 TB RAID for Compute
• 2Q/06 PBSpro - Domains (multinode)
• 2Q/06 Upgrade DMF IO systems
• 3Q/06 Upgrade WAN to 10Gbit to Multi-NASA sites
• TBD Upgrade Columbia
Questions/Comments