Lawrence Livermore National Laboratory
Integrated Computing and Communications Department

FY06 I/O Integration Blueprint

Reviewed and Released v1.7
UCRL-TR-220607

April 1, 2006

K.Cupps	**BG/L**
K.Fitzgerald	**Lustre deployment, Interconnects, IO Testbed**
M.Gary & J.Shoopman	**Archival Storage**
P.Hamilton	**Purple**
B.Loewe	**Scalable IO**
J.Slavec	**Facility Networks**
D.Smith & D.Fox	**Network Attached Storage**
D.Southard	**Visualization**
D.P.Wiltzius	**Facility Networks, Interconnects, Special Projects**

Other contributors:
John Allen, Ned Bass, Bill Boas, Tom Connell, Doug East, Brent Gorda, Bryan Lawver, Steve Louis, Terri Quinn, Mark Seager

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Purpose

This document provides an understanding of the near and long term computing and I/O resources in the Secure Computing Facility (SCF) and Open Computing Facility (OCF). Requirements for data flows, storage capacities and transfer rates are determined. Recommendations are made for architectures, timeframes for major deliverables, and procurements for the next fiscal year.

Scope

This document will provide an understanding of the resources creating most of the data (e.g., platforms, compute servers, visualization servers, etc) in the timeframe of next year and up to 5 years out. Also considered will be special mandates (e.g., Lustre going to production, etc). *The goal will be to recommend architectures, timeframes for major deliverables, and procurements for this fiscal year for OCF and SCF.* All resources that have significant I/O flows for local and tri-Lab remote computing resources and user communities will be considered: platforms, compute servers, visualization servers, archival storage system, network attached storage, site-wide global file systems, and networks. The general outline of this document is as follows:

1) Plans for next year and discussion of 2-5 year plans regarding major drivers of I/O requirements:
    a) platforms (e.g., Purple, BlueGene/L, Peloton)
    b) compute servers
    c) visualization servers
    d) special strategies, mandates, considerations, projects (e.g., swinging BlueGene/L, ICSI program deployments, etc)
2) Current architecture discussion that includes network topology, network attached storage (e.g., NFS servers), archival storage (e.g., HPSS), visualization servers, site-wide global file system (e.g., Lustre), and WAN connectivity to tri-Lab remote resources, collaborators and users. In this section information will also be provided, to the extent possible, on the past and current IO capacity and bandwidth.
3) IO requirements and analysis. This section will also try to summarize and focus the future architecture, capacities and bandwidths, and the steps to achieve the future state.
4) Identification of I/O issues and recommendations for addressing issues to achieve the architecture, capacities, and transfer rates suggested throughout the document.
5) Plans for FY06 and beyond. This will focus on providing major FY06 deliverables, and procurement plans. Longer term milestones and such may also be re-stated or summarized in this section.

Appendices will be used to provide additional information that is relevant to the IO architecture of our HPC facilities, but not necessarily of primary consideration.

- A: Inventory of major resources in the OCF and SCF. Mostly tables and statistics, no analysis.
- B: SCF and OCF computing model. This is our generalized HPC "usage model."
- C: Link aggregation tutorial. A topic of the day for networks, since our bandwidth requirements are going beyond the commodity market for single links (now at 10GigE max). Others welcomed!
- D: Technology watch. Whatever people want to contribute on new and interesting technology developments relevant to HPC IO architectures, beyond the immediate future.
- E: IB related protocols explained
- F: Lustre deployment strategy versus costs
- G: Lustre deployment strategy evaluation matrix

Executive Summary

The most valuable aspect of the FY06 I/O Blueprint effort is the exchange of information, and the integration discussions leading to design and procurement decisions. The essence of the resultant recommendations and decisions is captured in this document, primarily in sections 4 and 5. Section 4 identifies high level issues and makes recommendations on how to react to or mitigate those issues. This section first addresses strategic issues (in FY06, these are Lustre and Storage), then goes on to the tactical issues. Section 5 summarizes the high level deliverables as stated in the nWBS, and the large procurements.

To establish a framework for the I/O Blueprint effort, section 1 summarized the most significant drivers (e.g., machine deployments, WAN upgrade) for I/O requirements in FY06, and beyond. Section 2 provides a review of the architecture, which in general evolves slowly over the years. Section 3 is an attempt to identify I/O requirements based on historical trends, but applying some analysis considering the I/O requirement drivers summarized in section 1. These sections are supplemented by several Appendices that provide additional detail or support information for parts of the document.

For FY06 the I/O requirements were influenced by the platform deployments of Purple, BlueGene/L and Peloton. Relatively minor I/O requirements were provided by the Green Data Oasis and the DisCom WAN upgrade. Initial discussions involved the network deployments required for archival, Purple, visualization, and Lustre (primarily on OCF), and the funding of the network procurements. The Storage group has their situation in control, which was surprising considering the huge increase in computational capacity deployed in FY06 between Purple and BlueGene/L. The strategy of the Storage group for meeting these exceptional requirements is nicely described in section 4.

Most of the I/O Blueprint meetings were spent discussing the deployment strategy for Lustre. Two extremes were considered: the “distributed” scheme where each platform would be deployed with its own Lustre file system that would meet the throughput requirements for that platform, and the “centralized” scheme where two Lustre file systems would exist but only one would provide the throughput for the largest platform. Both schemes provide global access by mounting all Lustre file systems on all computing resources. The cost savings potential of the centralized scheme was recognized (with the aid of a cost model implemented as an Excel spreadsheet), but were outweighed by reliability and performance concerns given the current state of Lustre. Until Lustre development and/or product availability mitigates these reliability and performance concerns, the recommendation favors a Lustre deployment scheme that results in more Lustre file systems (e.g., “distributed”). Further, costs can be significantly decreased by careful consideration of the details of the each Lustre deployment, such as co-locating the Lustre storage with the platforms and visualization servers to the extent possible.

It is not clear at this time what events in FY07 will precipitate the I/O requirements. One topic that came up this year that may warrant significant consideration in FY07 is scalable I/O. Another may be open interconnect technologies as the challenge to InfiniBand mounted by the Open Fabric (formerly OpenIB) Ethernet vendor (e.g., iWarp, RDMA) community results in a critical mass of low cost 10 Gigabit NICs and large, low cost switch switches. We’ll return next year to provide you with a complete, accurate and unbiased report on these and other exciting developments!

1) Plans for major drivers of I/O requirements

Major computing resource plans for FY05 and beyond are:

1) Purple will be put into Limited Availability (LA) and then General Availability on SCF in FY06.

2) BlueGene/L will be put into LA on OCF, and then GA on SCF in FY06.

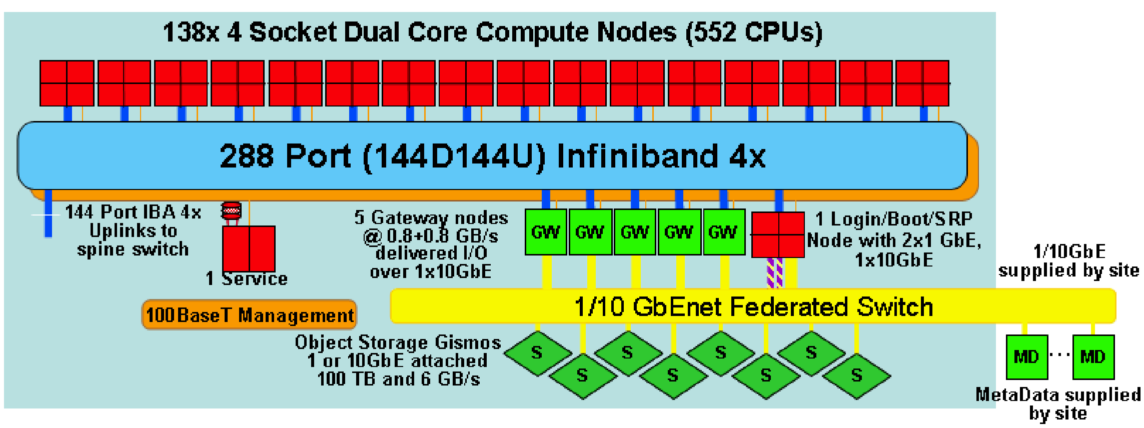
   a) Identify (network, Lustre) hardware needed for BlueGene/L to be sited on SCF.

   b) Anticipate additional issues if BlueGene/L will swing between OCF and SCF.

3) Advanced Architecture, BGP goal of 1 Petaflop (PF). At the end of CY07 a 512 node prototype will be sited using 4 core compute node ASICs, real SMP (cache coherency), 850-900MHz, otherwise same architecture, interconnect/network, etc. Housed in 72 cabinets with 64:1 compute node to IO node ratio for 1152 IO nodes, each 10GigE to Lustre, ideally with same storage capacity ratio as BG/L (or whatever we can afford). Memory is 2-4GB per node likely (possibly 8GB). Production unit expected in end of CY2008.

4) The Tri-Lab Linux Capacity Cluster (TLCC) initiative is a 2 year procurement for capacity computing funded at $45M/year ($90M over 2 years, shared by the tri-Labs as per HQ and the ASC Execs). This initiative was terminated due to lack of funding, but LLNL did leverage this concept for an M&IC funded machine – Peloton.

   The notion is to establish a standard Scalable Unit (SU) for the tri-Labs for capacity computing, where the SU specifies the CPU (AMD Opteron), interconnect (InfiniBand (IB)), gateway nodes, login nodes and object storage (e.g., Lustre), with the 1/10GigE networking provided by the site (see **Figure 1**). The intent of the original TLCC initiative was to provide a competitive cost advantage for such procurements, additional operational stability within each SU generation, and the opportunity for the tri-Labs to share software at



many levels, from BIOS to OS to tools and applications.

Figure 1: Single Unit (SU) Configuration

Multiple SUs can be combined for larger clusters with additional IB switches for a multi-staged interconnect. The SU processor and object storage technologies are expected to be identical within a cluster. Over the 2 year period technologies relevant to the SU are expected to evolve. Faster processors that are backward socket compatible are expected to be available several times over this 2 year period. Faster memory speeds may occur that would require a new motherboard. Towards the end of the 2 year period we expect to see new processors (probably not backward socket compatible) with faster memory and new motherboards.

There are two reference configurations:

a) Single SU reference configuration (90KW) – see ***Figure 1***

   - 5.1TF from 144 35.2GF/s quad socket 2.2 GHz dual core 16GB AMD SMP nodes interconnected with IB 4x (10Gbps) adapters for a total of 2.3TB DDR400 SDRAM at 25.6GB/s (memory

B:F=0.45, BW B:F=0.73). Less than 3 microsecond MPI latency and 2GBs MPI bandwidth (B:F=0.06).

- 102 TB disk (B:F=20) on (8 Lustre OSS') object storage systems interconnected with 1GigE (10GigE for Lustre) to 8 (Lustre) Gateway nodes resulting in 5GB/s parallel IO (B:F=0.001).
- No local disk. Local (non-removable) flash memory or remote boot to RAID5 boot and swap disks. 400 MB/s POSIX serial IO to any file system. One Login node with 1 (jumbo frame) 10GigE, 1 GigE for login traffic (e.g., SSH) and 1 GigE for NFS traffic. From each Login node, 800MB/s to archive over 10GigE and IBA connections.
- Software for build and acceptance RHEL 4.0, CHAOS, SLURM/LCRM, OpenIB, MPICH2, GNU Fortran, C and C++ compiler , Synthetic Work Load (SWL).

b) Multiple SU reference configuration (360KW, 864 sq ft).

(1) 20TF from 576 SMP interconnected with IB 4x adapters, staged IB switches, 9.2TB DDR400 SDRAM at 25.6GB/s (memory B:F=0.45, BW B:F=0.73). Less than 3 microsecond MPI latency and 2GBs MPI bandwidth (B:F=0.06).

(2) 408 TB disk (B:F=20) on (32 Lustre OSS') object storage systems interconnected with 1GigE (10GigE for Lustre) to 32 (Lustre) Gateway nodes resulting in 20GB/s parallel IO (B:F=0.001).

(3) No local disk. Local (non-removable) flash memory or remote boot to RAID5 boot and swap disks. 400 MB/s POSIX serial IO to any file system. Four Login nodes with 1 (jumbo frame) 10GigE, 1 GigE for login traffic (e.g., SSH) and 1 GigE for NFS traffic. From each Login node, 800MB/s to archive over 10GigE and IBA connections.

(4) Software for build and acceptance RHEL 4.0, CHAOS, SLURM/LCRM, OpenIB, MPICH2, GNU Fortran, C and C++ compiler , Synthetic Work Load (SWL).

With M&IC funding LLNL will purchase **Peloton** - 45TF of (approx 9) SUs in early CY2006. Peloton will use Lustre for the object storage with 10Gbps connections to the Object Storage Systems (OSS).

4) Green Data Oasis. This is an M&IC funded proposal - $850K in FY05 - driven by several programs to serve data to an international user community from the Green network (unclassified and unrestricted; outside the LLNL perimeter firewall for the restricted network, but likely will continue to require firewalls). For example, the Earth System Grid climate simulation output is a growing data set that will approach 100TB in a few years. This data set is to be shared by a large international community. Internet access with a large bandwidth capacity shared by many users is desirable.

5) LLNL's collaborative 3 year Scientific Data Management project with the University of California – San Diego (UCSD). This project provides for UCSD researchers doing science runs on LLNL's Thunder machine. UCSD will transfer large amounts (1-100TB) of data daily via the Storage Resource Broker (SRB) data grid protocol from Thunder to UCSD over ESnet for post-processing. It is expected that this project, along with the Green Data Oasis, will result in LLNL putting into production a 10Gbps ESnet network connection to the west coast research networks such as Internet2, and eventually as LLNL's production Internet path also. For the first part of the first year (Q1FY06) the throughput requirement for SRB data grid access from UCSD to data on Thunder is about 35MB/s (3TB/day). UCSD is anxious to realize the full potential of their SRB data grid resources, which are 10's of TB/day (about 1GB/s sustained).

6) The Integrated Cyber Security Initiative (ICSI) is expected to be deployed in collaboration with LC, use LC support staff and, over the long term, co-funded by LC.

7) ASC platform. Nothing on horizon for FY06. Competition for funding is as follows, prioritized:

a) Sandia NM upgrades Red Storm to 100TF.

b) LANL gets $75M funding in FY05 for 150TF machine.

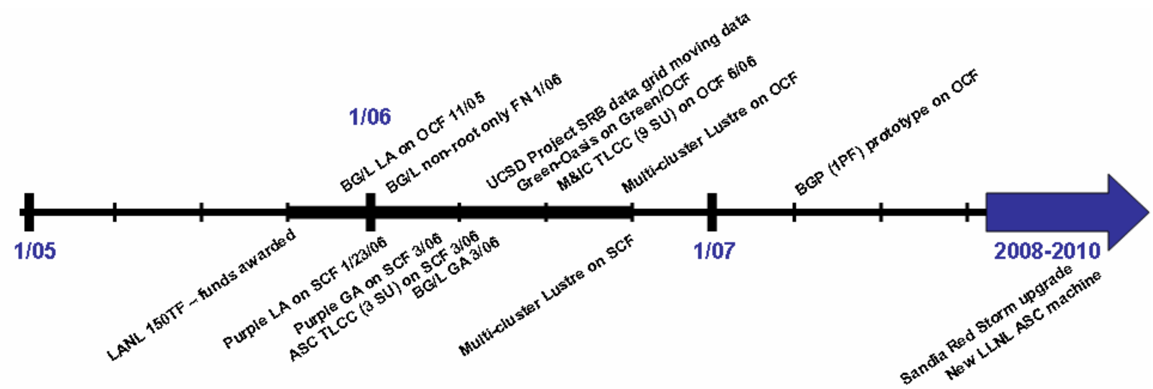
c) LLNL gets funding for machine in FY09-FY10

8) DisCom WAN contract has been re-competed and it appears the new contract will provide for 10Gbps from LLNL to SNL/NM, and an additional link at 10Gbps from LLNL to LANL. This is a 4x increase in link bandwidth, and an extra link (LLNL-LANL) compared to the current DisCom WAN contract.

Deployment Timelines

The timelines in ***Figure 2*** show the deployment schedule for BlueGene/L, Purple and several other major milestones. Milestones for OCF are above the timeline and milestones for SCF are below the timeline.

Figure 2: Estimated deployment timelines as of Q4FY05



2) Architecture

The long-term plan for the core IO Infrastructure is an integrated high performance parallel global file system (i.e., Site Wide Global File System (SWGFS)) and data storage architecture that provides users with fast and uniform access from all the compute resources (see ***Figure 3***). The SWGFS is currently implemented as a set of local Lustre file systems, with plans to go global in FY06 (i.e., inter-cluster Lustre).

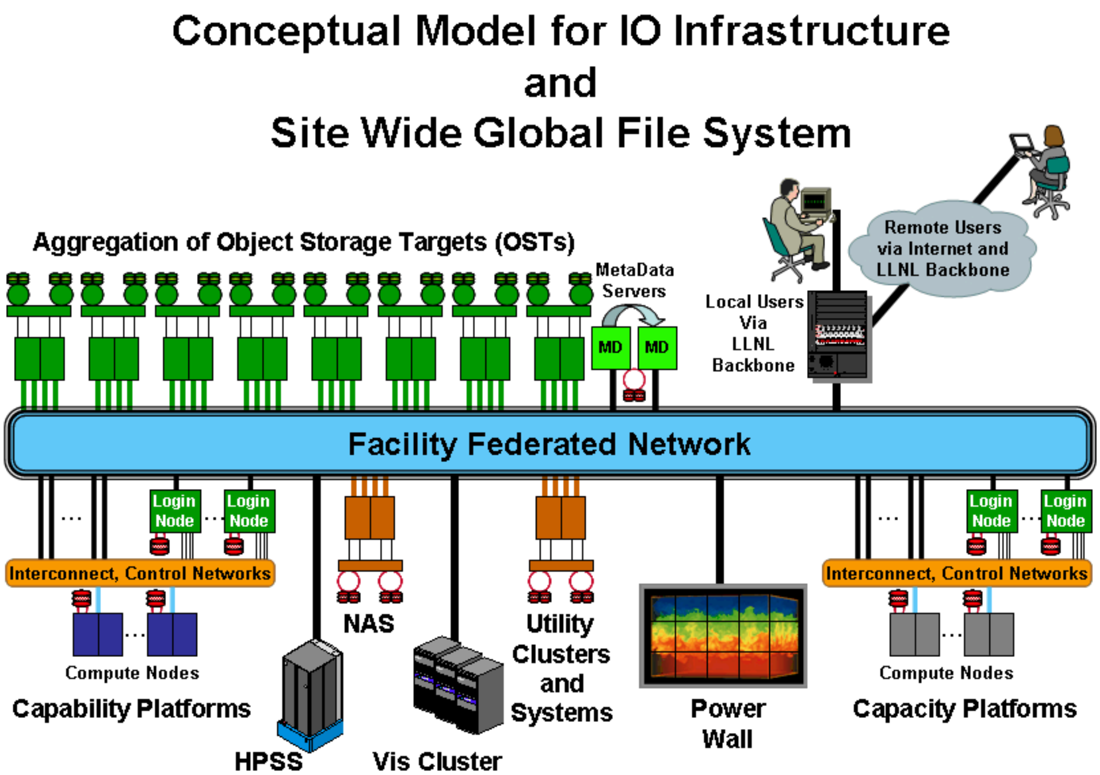


Figure 3: Goal for OCF and SCF facility-wide IO architecture

The architectures of the platform interconnect and the facility federated network infrastructures can utilize a range of different technologies depending on the preferred approach based on performance, cost and availability. For LC in FY05 Quadrics was the most widely deployed platform interconnect, 1 and 10 Gigabit Ethernet for the network infrastructure, and Fiber Channel and Serial ATA as the Storage Area Network.

The IO architecture for the Open Computing Facility (OCF) and Secure Computing Facility (SCF) is nearly identical (see ***Figure 4***), with only minor differences (e.g., WAN connectivity to Internet (OCF) versus SecureNet (SCF), etc). The Facility Federated Network referred to in ***Figure 3***is presently used to connect machines with the local Lustre file systems, presently via gateway nodes (not shown) on each machine. Soon (FY06?) the local Lustre systems will be transitioned to an inter-cluster Lustre file system by cross-mounting the Lustre file systems and coalescing the MetaData servers.

The sections below will provide an overview of the each sub-system of the extended OCF/SCF facilities. This overview will describe the current architecture, quantify the elements of the sub-system, provide an estimate of the aggregate source/sink throughput rates and capacity, raise issues if appropriate, and provide a list of recommended activities (e.g., performance tests, protocol analysis, etc) that would improve the understanding of the sub-system.

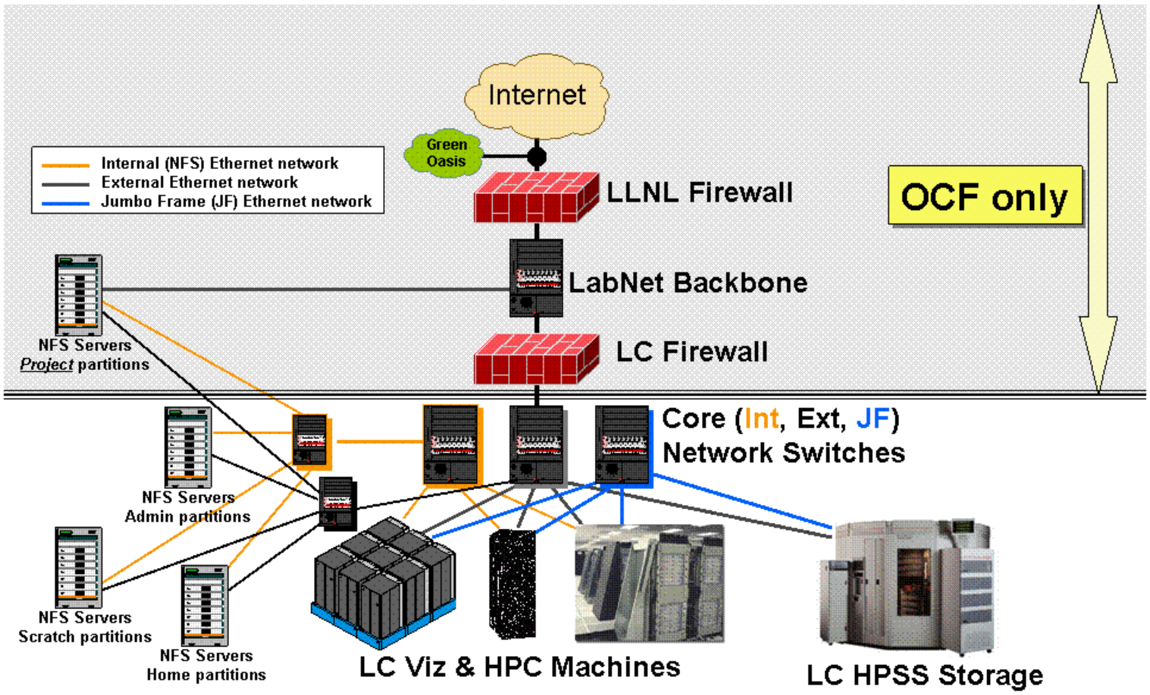


Figure 4: IO architecture overview for FY05

Facility Network

Presently there are three networks in production on the OCF and SCF facilities, as shown in ***Figure 4***: the Internal (or NFS) network for NFS traffic, the External network for inter-facility/desktop traffic, and the Jumbo Frame[1] (JF) network for intra-, tri-Lab facility traffic. Machines or resources on these networks will have one or more network interface cards (NICs) for that network. For example a typical machine login node will have at least 3 NICs, one each for the Internal, External and JF network. The larger machine login nodes have 2, 4 or even 8 JF NICs for reasons explained below. A fourth network is being deployed: the Facility Federated network for Lustre (hence also referred to as the Lustre network). This is network is also configured to support jumbo frames.

Although the OCF and SCF have three or even four networks, these networks do share switches when advantageous (primarily a co-location consideration) to improve switch utilization and costs. Think of this as mostly one physical network with multiple logical networks; that is, virtualization of the networks. Of course, the physical separation between OCF and SCF is always maintained.

External network

The External network provides access from outside the OCF/SCF facility for interactive traffic, some file transfers, etc. Specifically, it provides access to OCF/SCF resources for remote and local users via Open/Closed LabNet and the Internet/SecureNet.

Internal network

The Internal network, used for NFS traffic, does not route addresses outside the OCF/SCF facility (e.g., to desktops or LabNet or Internet/SecureNet) for enhanced security.

[1] Jumbo Frame is a non-standard but widely supported feature of Ethernet where the maximum Ethernet packet payload is increased from 1500 to 9000 bytes. Larger packets decrease the load on the host CPU through fewer interrupts and less protocol processing, hence raising the network throughput (significantly (2-3x) with leading edge networks like 10Gigabit Ethernet in FY05) while lowering CPU utilization compared to standard frames.

Jumbo Frame network
The Jumbo Frame network provides high throughput access between tri-Lab facility resources (e.g., HPSS, visualization, HPC machines, clusters) primarily for file transfers, typically utilizing multiple, parallel streams (TCP sockets) across the 4 JF subnets and if necessary over the DisCom WAN to other tri-Lab facilities. “Jumbo Frame” is a poor name for this network since there are other networks that are configured to support jumbo frame Ethernet packets; perhaps a more appropriate name would be “Data Movement Network” or such. The parallelism of data movement is enhanced by machine nodes that have 1 or more NICs on each of the 4 JF subnets. A machine node (e.g., Purple) with multiple NICs on the same JF subnet will be required to bundle these NICs to one logical channel via the IEEE Link Aggregation Control Protocol (802.3ad). On rare occasions, a JF NIC will be configured with an IP address on 2 or 4 of the JF subnet (i.e., aliases) to establish connectivity to each JF subnet without the cost of additional NICs, but this practice is not recommended due to the resultant performance limitation.

	OCF		SCF		OCF Federated	
Year	2004	2005	2004	2005	2004	2005
Small core chassis	6	10	3	3	0	0
Large core chassis	17	20	11	18	0	17
1GigE ports	1518	2140	732	1012	0	2420
10GigE ports	16	40	20	48	0	50

Table 1: Growth of facility core network components and ports in production

The growth of the network for OCF and SCF, both in the number of chassis and number of 1 and 10 Gigabit Ethernet ports, is shown in ***Table 1***. The Facility Federated (Lustre) switch began deployment in the OCF in FY05. There was also a significant increase in the number of chassis and 1 and 10 Gigabit Ethernet ports in the SCF, and even more so in the OCF. Compared to the SCF, there are more network components on the OCF mostly due to the number of buildings in which LC has OCF resources sited (e.g., B113, B115, B117, B439, B451, B453).

System and NIC Type	B/w (Gbps) per NIC
Ext	0.25
Int (e.g., NFS client)	0.60
1 GigE JF	0.60
10 GigE JF	5.00
Lustre Gwy/Node 1GigE	0.60
Lustre Gwy/Node 10GigE	5.00
1 GigE OSS/OST	0.50
10 GigE OSS/OST	5.00
NFS Server - NetApp	0.40
NFS Server - BlueArc	0.40
NFS Server - Panasas	0.40

Table 2: Per NIC throughput by network and server type

Source/sink throughput estimates are provided throughout this document based on the per NIC throughput in ***Table 2***. These numbers are estimated simplistically on the basis of the throughput of a NIC for each type of network (e.g., external, internal, jumbo frame, etc) and system (e.g., machine, Lustre client, Lustre OSS/OST, NFS server). This is certainly overly-simplistic until the total network throughput capacity of the client/server is factored (e.g., bandwidth per NIC of a machine with one NIC will likely be more than if the machine had 4 NICs). The throughput of transferring data to/from the machine and media should also be factored.

DisCom WAN
The DisCom WAN must complete a competitive RFP for the OC-48 (2.5Gbps) wide-area network link between LLNL and SNL/NM so the replacement WAN is in production before the current WAN contract expires in 1/31/2006. With this RFP the Labs are moving from the current ATM Fastlane encryptors to HAIPE (High

Assurance Internet Protocol Encryption) compliant 1Gbps Taclane (NSA Type 1 IP) encryptors available since mid-CY2005. An IP based WAN is desirable since there are few ATM network components available (we have only one vendor for OC-48 speeds even), ATM product development is stagnant (little R&D funding for these products), and the existing products are challenging to configure to meet our performance requirements. With IP encryptors, we can use mainstream network components for the DisCom WAN.

The new RFP will accommodate a dual path 10Gbps WAN with a link from LLNL to Albuquerque via the west coast (as the current WAN does) and a second, new link from LLNL to Albuquerque or Santa Fe via Denver. LANL has negotiated multiple 10Gbps paths ("Geomax") between LANL and Albuquerque to meet their institutional and ASC (Advanced Simulation and Computing) network requirements. This increase in bandwidth will provide adequate access to the Purple and BG/L machines at LLNL, and when combined with Geomax provides a route diverse, redundant path to LLNL for both Sandia/NM and LANL.

Additional information

- Improve estimates of throughput for NIC and services to ensure they are applicable to the real world, and improve and document method by which these are obtained.
- For UCSD-LLNL project, benchmark throughput from Thunder to UCSD TeraGrid resources. Do benchmarks from Thunder to test machine on LabNet backbone (outside LC firewall), then to test machine (lc-test) outside LLNL perimeter firewall, then to UCSD test machine.
- For UCSD-LLNL project, monitor throughput from Thunder (or test machine network-close to Thunder), and intermediate machines, to UCSD using Netmon and authenticated iperf tool. Make results viewable to UCSD.
- Robust testing of port aggregation algorithms used by Linux and network components to understand and document anomalies, esp. those encountered in HPC facility vs ISP (see footnote **Error! Bookmark not defined.**).
- Acquire 10 Gigabit Ethernet NICs for PCI Express for performance testing, with CX4 GBICs for Cisco's 10 Gigabit Ethernet switch ports.
- How to instrument the network at key points to better understand what we are doing right, or could do better?
- Review network monitoring and performance statistics gathering for IP based DisCom WAN expected in early 2006. Baseline (and document) the latency for all links.

Network Attached Storage (NAS e.g., NFS)

The Network Attached Storage (NAS) service is from a Gigabit Ethernet NIC connected to the Internal (aka private) network. Typically each login node of each machine has one Internal Gigabit Ethernet NIC. All NFS servers are also connected to the Internal network with one or more Gigabit Ethernet connections (typically 1 or 2, but one NFS server has 60 Gigabit Ethernet ports – see Appendix A for details). The NFS servers only use the NFSv3 protocol, all are configured for NFS over UDP or TCP, and none are configured for jumbo frame.

Most NFS systems also have network connections to the External network. This connection will provide NFS services to machines outside the facility, such as the office machines for LC staff.

On the OCF, some NFS servers are also dual-homed to provide NFS service to the OCF resources and to certain projects. Dual-homed NFS servers have an Internal network connection (for NFS on OCF) and, to provide access for the select projects, an External network connection that goes directly to the LabNet backbone since the LC firewall does not allow NFS traffic to pass through (see). This allows projects to share data with their local machines and OCF resources without having to transfer large files (i.e., datasets) in entirety.

Each NFS server is configured with up to four partitions: Admin, Home, Project, and Scratch. These categories have different performance, reliability, availability and feature set requirements.

Admin Space
The Admin partition is allocated from the NFS servers to LC platforms for system management, configuration management, and shared binaries. An example is the /usr/local file system where a cluster has shared applications, libraries and tools. The system space is fully backed up and has snapshots enabled for easy recovery of files.

Home Space
The Home partition is for user home directories (most users have a quota of 16GB), hence has the highest reliability requirement and so the servers are configured as a fail-over cluster. The files in the home directories tend to be small, so the operations-per-second that a server sustains is the primary performance measure.

The home directory space is backed up to tape monthly, and incremental backups are made daily. Full backups are required for disaster recovery, but the restore time defines how long users are without access to their files. This restore time dictates how large a home directory file system is created. The home directory space has a snapshot or point-in-time-copy feature that lets the user retrieve a file from the file system on demand without needing to go to tape. This is very convenient for both the user and the system administration staff. The snapshot schedule is set at twice a day; noon and 7pm, a total of four snapshots are kept on line. The snapshot feature adds five percent overhead to the usable space in a file system.

Project Space
The Project partition is co-funded by projects for their use with LC computing resources and, as explained above, can be configured for access from the user's local computing environment.

Scratch Space
The Scratch partition is a temporary file storage area for users which is not backed up and is subject to purging. Machines have very little dedicated disk space, so the working space for application results and checkpoints is provided via the NFS scratch partition. Users have also found the center wide scratch space to be a convenient way to share files between capacity and capability clusters.

The primary feature of the scratch space is throughput. The I/O access to this space needs to be fast so that user calculations can continue. The total capacity available is important since the space is purged when the scratch space fills up. To prevent a run-away process from filling up this shared resource, a 100GB quota is imposed on every user. The scratch space is not backed up and snapshots are not enabled. It is the responsibility of the users to save to the archive any data in the scratch space that is needed long term.

NAS capacities

Table 3 shows the approximate NAS capacity for OCF and SCF are shown for FY04, and the NAS capacity for FY05. This capacity, and the aggregate bandwidth discussed later, is provided by 18 NFS mounted server systems on OCF and 12 on SCF. On both the OCF and SCF, the largest NFS server is the Panasas NFS server (approx 70TB) and that is for scratch space, the rest of the capacity is provided mostly by Netapp servers but with a few (1-2) Bluearc servers.

	OCF (TB)		SCF (TB)	
	FY04	FY05	FY04	FY05
Admin	4.5	**8.4**	2.3	**4.0**
Home	18.2	18.2	10.3	10.3
Project	13.0	**13.8**	7.0	7.0
Scratch	80.0	**94.0**	80.0	**96.4**

Table 3: NFS storage capacity on OCF and SCF for each of the 4 partitions

NFS over UDP vs TCP; jumbo frame
Historically NFS has been based on UDP/IP. One problem with UDP/IP is the poor throughput experienced by NFS clients in environments that drop packets. Dropped packets are particularly disastrous for NFS since NFS generally uses an 8KB datagram, which are then fragmented within six UDP/IP packets that each fit in a (standard frame) 1500 byte Ethernet packet. If one of the Ethernet packets is dropped, the rest are discarded and need to be re-transmitted. Another problem is that jumbo frame Ethernet can only be deployed with UDP/IP in the most carefully controlled network environment, which in practice prevents the use of jumbo frame and the typically significant improvement in throughput achieved with NFS over jumbo frame.

With NFS running over TCP/IP, jumbo frames can then be enabled on the NFS servers. Machines can be configured for jumbo frame or not and the MTU (aka packet size) discovery that is part of establishing a TCP connection will determine the proper MTU (e.g., standard or jumbo frame). With jumbo frame, an 8KB NFS datagram will now fit in one jumbo frame Ethernet packet avoiding the inefficiencies experienced with dropped packets. Furthermore, if the NFS over TCP/IP connection does not use jumbo frame the NFS datagram is packaged in several Ethernet packets as with the UDP/IP scenario. However, TCP will re-transmit only the dropped data and not the entire NFS datagram as in the case with NFS over UDP/IP.

Additional information

- Note NFS capacity per SCF and OCF and per partition for FY04, to help see growth rate.
- Performance of "typical" NFS server – per interface and per filer system. Establish total throughput capacity per filer, and per interface.
- Configure NFS servers for jumbo frame? Quantify NFS throughput improvement with Jumbo Frame. Impact on Internal network configuration.
- Test performance and tolerance to network errors of NFS over UDP/IP vs TCP/IP, maybe via analysis of information on NFS servers and clients. All NFS clients configured for TCP/IP vs UDP/IP?
- Documentation of end-end performance, relevant features, and security testing with NFSv4.

Scalable I/O

The scalable I/O project (SIOP) provides high-performance parallel file system and I/O library support for all major platforms at LLNL. It performs acceptance testing as part of the procurement process, and collaborates with platform partners, academic researchers, and vendors to address ASC high-performance I/O needs.

The SIOP will accept, benchmark, and support LLNL parallel file systems, developing any testing tools necessary to ensure a stable and high-performance file system. The project continues to provide customer support for parallel file systems, middleware (MPI-IO) and higher-level I/O libraries (HDF5, PnetCDF) and any application I/O issues. With an aim at anticipating future I/O issues, the SIOP will continue to pursue future technologies for file systems with its ASC Alliance contracts (UCSC – Metadata, NWU – Distributed Caching, UMich – NFSv4) and IETF Open group designs for pNFSv4 and POSIX Extensions.

Additional information

- Document of end-end performance, relevant features, and security testing with NFSv4.

Storage (e.g., HPSS)

The HPSS storage system and architecture is largely an independent sub-system of the facility-wide architecture (see ***Figure 4***), requiring adequate network bandwidth for the Internal and External network. Furthermore, the HPSS storage architecture (***Figure 5***) is very similar for both OCF and SCF. In this figure, the three core networks (Internal – orange, External – black, and Jumbo Frame - blue) are depicted. Also shown is the HPSS "migration" jumbo frame network strictly for the HPSS disk and tape "background" migration activity. Finally, the Storage Area Network (SAN) is shown with connections to the IO devices and HPSS movers. The SAN consists of about 60 Brocade Fibre Channel (FC) switches connected to HPSS movers and IO devices by one or more FC Host Bus Adapters (HBA) each with a bandwidth of 2 Gbps (e.g., 250MB/s). Most of the 60 Brocade FC switches are small switches deployed to connect 4-5 tape drives to a mover machine; the rest are larger switches used to share disk controllers with multiple mover machines for transfers of smaller files.

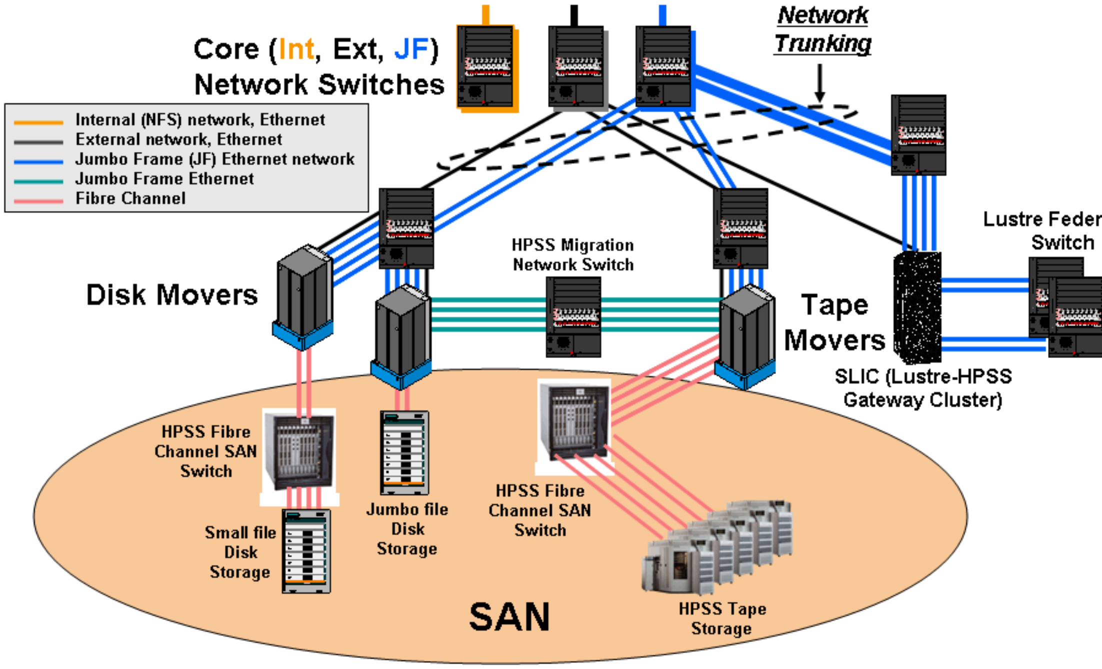


Figure 5: LC HPSS Storage Sub-system Architecture

HPSS User Applications
The users read, write and manage their files on HPSS through three user applications: PNFT, HTAR, and PFTP. HPSS keeps statistics on each of these three tools. PFTP (Parallel FTP) and FTP (serial FTP) exist on LC machines as one client. The selection of serial or parallel, and the selection of network interfaces on the source and sink machines is determined dynamically and dependent on the source and sink machine and considering the network architecture between those two machines.

PTFT and HTAR were written to use parallel streams load balanced using multiple Jumbo Frame NICs to achieve high throughput for data transfers. PNFT was implemented using PFTP as the underlying data movement mechanism, so PNFT also achieves high throughput by parallel streams over multiple network paths when appropriate. On OCF and SCF in FY2005 PFTP was used to move 75% of the files, PNFT for 25% and HTAR for 5% (these are approximate to give an indication of the trend) – but be aware that HTAR stats count as one file an operation on a collection of files (e.g., HPSS jargon is “member files”), so that distribution statistic is inherently skewed against HTAR. Other tools such as Hopper use one or more of these HPSS user applications to do the data movement, HPSS does not keep statistics on the use of higher level data management tools such as Hopper.

Parallel Streams and HPSS Movers
As mentioned above data movement applications use multiple, parallel streams over different physical network for greater throughput. HPSS also employs striping across multiple physical media to achieve higher aggregate throughput. The HPSS movers must coordinate the multiple data streams effectively not only for higher throughput, but also considering availability of data if one or more mover machines become unavailable. For this reason HPSS does striping to multiple media, but the striped media for one application are all connected to one mover machine. Hence multiple data streams over the network are intelligently distributed on a mover machine across multiple mover processes and media drives and controllers to achieve effective end-to-end parallelization.

Storage Lustre Interface Cluster

On OCF, HPSS also provides a gateway, the Storage Lustre Interface Cluster (SLIC, see Appendix A for more details), between HPSS and the Lustre federated network. Users log into the SLIC cluster to FTP files between HPSS and the mounted Lustre file systems. SLIC is a Linux cluster since Lustre is only directly accessible from Linux at this time. Software development efforts are underway in FY05 to provide access to Lustre for HPSS movers, mitigating the need for HPSS-Lustre gateways. Each of the 10 SLIC nodes has four Gigabit Ethernet interfaces for each of the Jumbo Frame and Lustre Federated switch network, and one External Gigabit Ethernet interface – a total of 40 Jumbo Frame Gigabit Ethernet NICs, 40 Gigabit Ethernet NICs for the Lustre federated switch, and 10 Gigabit Ethernet NICs for the External network. The 40 Jumbo Frame NICs on SLIC are connected to an aggregation switch in B543, and trunked to the core Jumbo Frame switch with two 10GigE links. Recent measurements with Thunder (essentially the same hardware and OS as SLIC) and a high-end dual Xeon Intel processor get over 900Mbps for standard frame Ethernet. Hence the SLIC gateway is capable of substantial network throughput.

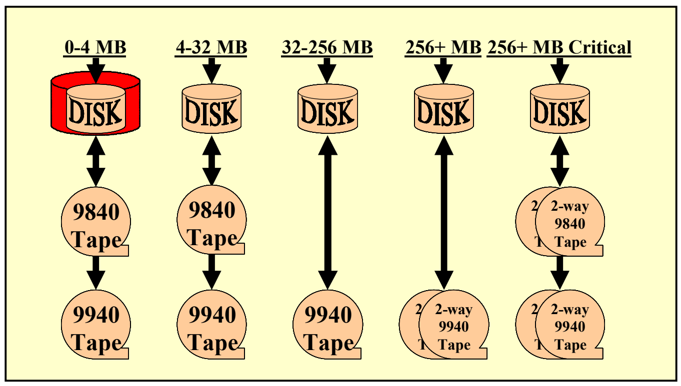


Figure 6: HPSS Storage Hierarchies by Class of Service (COS)

Class of Service (COS) and Storage Hierarchy

HPSS defines 5 classes of service based on the file size: 0-4MB, 4-32MB, 32-356MB, >256MB (256+MB), and 256+MB Critical. ***Figure 6*** illustrates the migration strategy of HPSS based on the COS. For example, the smallest files are migrated from disk to two types of tape, and the largest files are migrated and mirrored to one or, for critical COS files, two types of tape.

Keeping in mind that the y-axis in ***Figure 7*** is logarithmic, notice that the number of files archived in FY05 decreases as the COS file size increases, but the amount archived increases – ignoring the "critical COS." This trend is probably assisted by the fact that the Htar tool concatenates a large number of files (e.g., a sub-directory tree) into one file, and the statistics for writes count this as one (large) file.

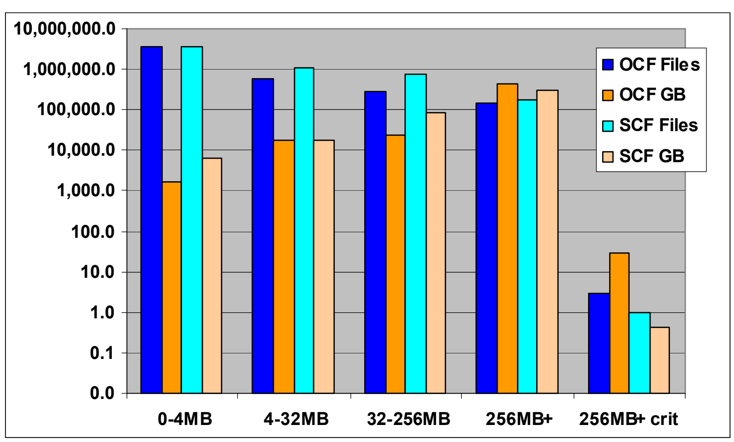


Figure 7: FY05 File and Data Storage by COS

With disk prices dropping and capacity ever-growing, the capacity of the small file disk cache will be increased to be able to store all HPSS small files (under 4MB) online. This strategy has several benefits: it minimizes the problems associated with storing/retrieving small files on tape, and it minimizes stage time for the files which are most-read by users. Aggressive repack will be required to stage millions of existing HPSS small files from tape to the new disk in order to fully realize these benefits. The dual-copy strategy for medium-size files (4-32MB) will remain unchanged as the required disk capacity and cost is prohibitive.

HPSS Movers

Each mover node is connected to the External network (e.g., serial FTP), Jumbo Frame network (e.g., PFTP), and Migration network. Additionally, a mover node is directly connected to disk or tape IO devices, or more typically to the SAN with 1 or more FC HBAs. These connections are shown in ***Figure 5***.

All disk movers are IBM 6Mx machines running AIX. Each disk mover node is designated for one of two groups of Class of Services (COS): COS for <256MB and COS for 256MB+ ("jumbo"). Jumbo disk mover configuration is optimized for throughput. Each have four FC HBAs directly attached to all 4 ports on the DDN RAID controller. The other disk movers (the COSes with files <256MB) are optimized more for capacity than throughput. Each have 8 FC HBAs connected to the SAN to 16 ports of RAID controllers. DDN RAID achieve about 600MB/s for reads and writes.

All tape movers are IBM 6Mx machines running AIX, but there are plans to deploy a large number of Linux tape mover machines in FY06. The tape movers each have 4 FC HBAs, each connected to a FC switch with 5 tape drives – a total of 20 tape drives per tape mover machine. The STK 9840 tapes are capable of 15MB/s compressed (10MB/s before compression). The STK 9940B tapes are capable of 45MB/s compressed (30MB/s before compression).

Understanding that the disk and tape mover nodes provide about 130 (jumbo frame) Gigabit Ethernet connections to the two HPSS switches (shown in ***Figure 5***, above the mover node icons), there is network bandwidth oversubscription in the trunks between the HPSS switches and the core network switches (top row of 3 switches in the figure). Currently, there are two 10GigE trunks connecting the HPSS switch on the left (for the disk movers) and one 10GigE trunk for the switch on the right (for the tape movers). Considering everything, this results in approximately a 6:1 network bandwidth oversubscription between the Core switch and the HPSS movers. This trunking bandwidth may be inadequate, particularly for the tape movers which require a sustained data flow to maintain streaming tapes.

For FY06, the number of mover nodes, peripherals and size and connections to the HPSS disk cache will increase to provide specified increases in throughput and to accommodate the retirement of obsolete hardware. Linux tape movers will also be introduced, along with new tape media: STK Titanium (**under NDA** - 120MB/s uncompressed) and IBM LTO-3 (67MB/s uncompressed). Tape movers will be configured with two Jumbo Frame NICs (aliased to the 4 JF subnets), and two Migration jumbo frame NICs. Tape movers for Titanium will have 2 FC2 (i.e., 2Gbps) HBAs connected to two STK Titanium drives; tape movers for LTO-3 will have 2 FC2 HBAs connected to four IBM LTO-3 tape drives. Since tape drives perform best when streaming, it is important to maintain data throughput to/from the tape drives to keep them spinning. As such, these tape drives are able to adjust their speed (e.g., 3-5 speeds) to best match the source (tape write) or sink (tape read) throughput of the data.

SAN (Storage Area Network)

The SAN uses 2Gbps Fibre Channel (FC) to provide connectivity between HPSS movers and tape and disk controllers. The SAN strategy for disk and tape is different. HPSS movers for the jumbo (256MB+) file COS are directly attached to disks and hence use no SAN. HPSS movers for small through large file COS use a SAN of larger FC switches configured to isolate disk controller ports to specific mover HBAs (via FC zones). If a failure occurs in a mover, disk controller or part of the SAN, the FC switches can be quickly reconfigured to maintain access while suffering only a modest decrease in throughput.

Each HPSS tape mover has a small FC switch that connects the mover machine to 4-5 tape drives. This is a cost effective solution to provide adequate bandwidth through the mover to the respective SAN attached tape drives.

HPSS Metadata

The HPSS metadata storage is the critical directory that maps between the file system hierarchy and the storage device. Being such a critical element of HPSS, the metadata has two independent servers and storage. The metadata storage is mirrored and located in two geographically diverse locations at LLNL for both OCF and SCF: in B115 and in B451/3. A robust SAN is built to provide this diverse connectivity, and also to provide connections to tape drives dedicated to backing up the metadata.

Additional information

- Is the network bandwidth oversubscription between the core and disk movers, and core and tape movers, appropriate? Note that the HPSS disk movers and HPSS tape movers each have their separate HPSS switch. Special consideration for tape movers to ensure the tapes are streaming?
- Benchmark/document mover architecture and throughput:
    - Mover architecture: number of tape mover machines, and disk mover machines; number of disk (or FC SAN) HBAs per mover, number of tape drives per mover, movers that are AIX or Linux
    - Device throughput: streaming tape speed (read/write), max per disk controller IO (read/write)
    - Read/write from mover with tape drive and disk SAN
- Determine performance of HPSS mover node standard frame Ethernet (e.g., External network) and jumbo frame Ethernet, with 1-9 Ethernet interfaces active with the configuration (above) in production. This will validate the per mover peak throughput of the HPSS sub-system, which currently assumes 250Mbps for the External NIC and 600Mbps per jumbo frame NIC.
- Data movement tools should use the jumbo frame (intra-facility data movement network) network for optimal throughput and parallelism: pnft, htar, pftp. Confirm effectiveness of data movement e.g., especially if pnft uses pftp, are all 4 stripes used effectively?

Lustre, Inter-Cluster (SWGFS)

The goal of the Site Wide Global File System (SWGFS, aka Cluster-Wide File System in ***Figure 3***, and presently implemented as Lustre) is to provide a high performance file system that incrementally scales in throughput and capacity. The SWGFS is shared by all or many of the machines on the site (facility). The scalability requirement has some nuances that are best explained with an example. When a capability (i.e., largest) machine is added to the facility the global file system requirements (i.e., throughput and capacity) of that machine will be met by adding resources to the existing SWGFS implementation, rather than providing all global file system resources with the capability machine. In fact, the throughput requirement of the SWGFS is to "only" meet the peak throughput requirement of the largest machine on site (and NOT the aggregate peak throughput of all machines on site, which would be larger). It is assumed that in practice any machine's peak throughput is of relatively short duration, and occurs infrequently. It is also our experience that capacity requirements are typically met once the throughput requirements are met, due to the rapidly increasing capacity of disks. Also, the SWGFS must deal with several generations of storage products since the SWGFS grows incrementally over time using the latest storage products at the time purchased – a substantial challenge facing Lustre. The SWGFS must also have one metadata server (versus a federation of metadata servers and associated storage systems) since the largest capability machine is expected to meet peak throughput requirements via access to the entire SWGFS.

In the past the hardware and software for the cluster parallel file systems were tightly integrated with the cluster itself and purchased as part of a package deal. With the SWGFS model, integration of a new system would require an incremental upgrade in the capacity and throughput of the SWGFS. Hence the cost of providing a global file system for a machine via SWGFS would be much less than if the machine had a dedicated global file system. The lower cost of the global file system for a machine tracks the lower cost of capability and capacity computing platforms that are increasingly utilizing open (i.e., non-proprietary) software and hardware solutions, such as InfiniBand for interconnects.

The architectural model for high performance file systems is for all cluster-wide (global) file systems to be replaced by the SWGFS, a single global scalable parallel file system serving all the platforms. The rate of progress toward the complete SWGFS is determined by the stability of the Lustre file system as it is deployed and the effort needed to meet full security requirements. Currently the SWGFS, Lustre, is deployed as a set of independent systems that are mounted on machines needing access (see ***Table 4*** and ***Table 5***).

Unclassified Lustre OSS (OCF) 256 Gbps Total					
System	Storage Capacity (TB)	Gateway Nodes / OSS	GigE NICs	10GigE NICs	External I/O
Lustre GM1	91.5	64	2		64.0 Gbps
Lustre GM2	91.5	64	2		64.0 Gbps
Lustre GA1	45.0	32	2		32.0 Gbps
Lustre GA2	45.0	32	2		32.0 Gbps
Lustre GT1	192.0	64	2		64.0 Gbps

Table 4: FY05 capacity and configuration for Lustre systems on OCF

Classified Lustre OSS (SCF) 280 Gbps Total					
System	Storage Capacity (TB)	Gateway Nodes / OSS	GigE NICs	10GigE NICs	External I/O
Lustre GL1	152.0	56	2		56.0 Gbps
Lustre GB1	400.0	112	2		112.0 Gbps
Lustre GB2	400.0	112	2		112.0 Gbps

Table 5: FY05 capacity and configuration for Lustre systems on SCF

The change to the SWGFS model may also bring changes to the computing usage model. For example users will no longer be required to explicitly move data between the computing platforms and the visualization servers. There will also be no reason to keep multiple copies of large files to have them conveniently available. The architecture of SWGFS and its connectivity to the archive is still in the planning stages.

To complete the integration of the SWGFS into the SCF and OCF environments, some convenient access method is required for usage from other than the major platforms. This is expected to be done with one or more NFS portals. With the introduction of NFSv4 over the next several years, it is planned that clients which have the throughput capability will able to use parallel extensions to the NFSv4 protocol to get higher performance to files within SWGFS.

Determining requirements for throughput and capacity of the SWGFS takes some understanding of the file system usage patterns and delivered I/O performance. The benchmark programs used to verify the file system performance for acceptance purposes is very different from most applications. Because of the less optimized file access patterns, applications typically see I/O rates of roughly 25% of the peak. Allowance was made for this difference in the guidelines used to specify the required I/O bandwidth when a new system was procured.

Additional information

- Are we doing adequate testing of Lustre client, gateway and OSS/OST to ensure we understand performance?
- Can we monitor OSS performance to ensure appropriate performance during operations?
- Proceed with architecture design reviews to accommodate IB or 10GigE as Lustre interconnect between clients/gateways and servers (OSS/OST).

Visualization

Visualization resources provide an essential and critical part of the high performance computing environment. Furthermore, the visualization resources are often used in real-time by the end-user to post-process computational data from the various computing resources. Visualization resources have evolved in response to technologies, computer architectures, and facility architectures:

- The first generation visualization machines were characterized by SGI machines which provided a multi-processor shared memory machine architecture, very high speed hardware assisted graphics rendering, and high speed local disk IO (~600MB/s per thread). These machines were accessed for visualization by only a very few users at a time, reserving one or several of the machine's frame buffers for RGB ("video") image delivery to the end-user's office by special hardware (e.g., Lightwave Communication RGB switch and extenders). This generation visualization servers set a high bar for user expectations. The disadvantages were: expensive system, graphics engines, and peripherals; difficulty in moving data to machine for visualization processing; maintenance was expensive and required skills that were diminishing as the SGI market presence faded.

- The second generation visualization machines were Linux clusters that addressed issues with limited success. The first visualization cluster, Vivid, did not have a cluster file system and was more of a proof of concept, focusing more on the demonstration of visualization tools. Then followed Sphere and Gvis, which tried the PVC cluster file system with little success. Lately the Lustre file system is being used for visualization servers: Sphere share's the MCR Luster system; PVC the BGL Lustre system; and Gvis has a Lustre file system on the classified. However, the throughput and stability, although improving, are a limiting factor in the visualization cluster overall performance. Meanwhile, the visualization tools on clusters nicely complement the visualization tools and capabilities available on the HPC platforms. And CHAOS, Redhat Linux with LC's modifications and management tools, is being run on the visualization clusters. With visualization clusters users run VisIt real time and get the images delivered to their office terminal via the network using commodity components (vs the Lightwave Communication RGB extenders with the SGI).
- The current generation of visualization servers will continue with Linux clusters. Improved performance from Lustre is expected. Until Lustre is deployed "globally" within a facility, a visualization server will likely be associated with a platform and that platform's Lustre system (e.g., Gauss, BG/L and the BG/L Lustre system). Another change is more use of InfiniBand (IB) as the interconnect. Reliance on IB as the interconnect also brings the expectation for IB switches to provide 10Gigabit Ethernet ports for connectivity to the Gigabit Ethernet system area network. Associated with IB and the IB-Ethernet gateway functionality is the reliance on the success of OpenIB. In summary, visualization cluster deployments recently completed or in progress are: Vertex (16 dual Opteron nodes, InfiniBand interconnect), Klein (10 dual Xeon nodes, Elan4 interconnect), and Gauss (256 Opteron nodes, InfiniBand interconnect) that will follow BG/L to classified and share BG/L's Lustre file system.

Interconnect and System Area Networks

Proprietary interconnect solutions will continue to be a critical part of leading edge computers such as IBM's Purple and the BlueGene/L system. However, large Linux clusters at LLNL have been using solutions that are not proprietary to the machine vendor, such as Myrinet, Quadrics and even Ethernet in a few cases.

The current interconnect strategy acknowledges the continuing role of proprietary solutions for the leading edge platforms (e.g., BG/L, Purple). However, an industry standard solution is desirable for cluster platforms. InfiniBand (IB) has been selected by the tri-Lab community as the open interconnect strategy for clusters. In support of this strategy, LLNL is participating in IB collaborations for developing standards. Additionally, LLNL is promoting the delivery of high performance IB software stack through participation in the OpenIB Alliance. See the appendix pertaining to the IB protocols, stacks, and interfaces established and/or under development by industry.

Special Projects

Green Data Oasis

The Green Data Oasis (GDO) is an M&IC funded proposal - $850K in FY05 - driven by several programs to serve data to an international user community from the Green network. The data on the GDO is expected to grow to several Petabytes over a few years, and will be shared with a large international community at very high throughput. Hence Internet access with a large bandwidth capacity shared by many users is desirable.

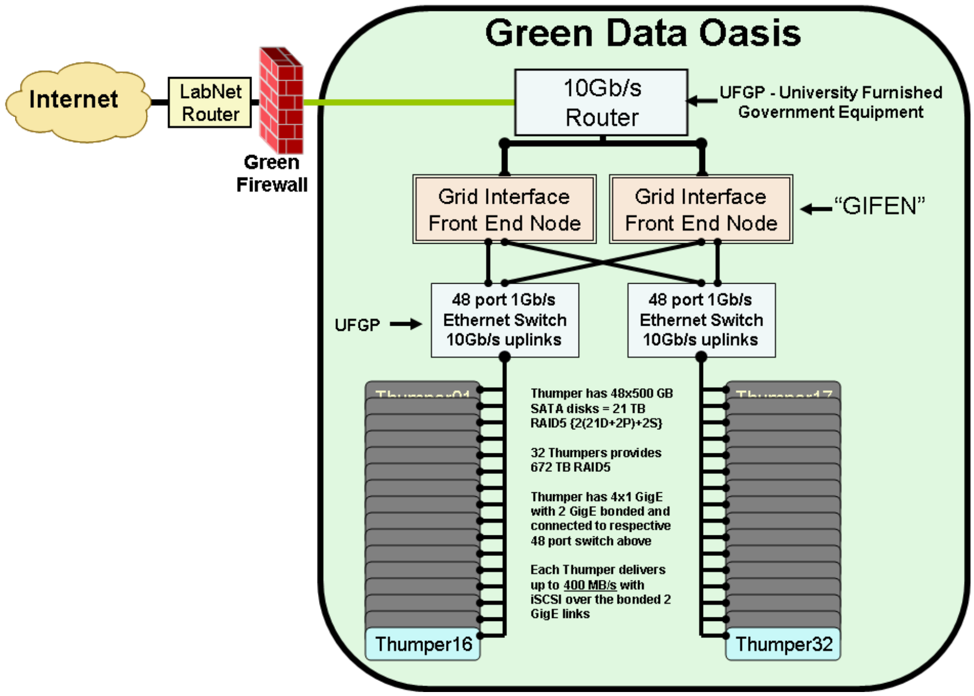


Figure 8: Green Data Oasis architecture

The GDO architecture (see ***Figure 8***) is based on Sun's new Thumper disk storage technology accessed by Sun's ZFS (Zettabyte File System), which provides high performance, scalability, and data integrity[2]. Local and Internet access to the GDO is over 10GigE to two grid interface front end servers (GIFEN) supporting several data movement interfaces including gridFTP, FTP, and HTTP (see ***Figure 9***).

The deployment schedule of the GDO is tempered by the availability of the Thumper product and other hardware. It is a phased deployment to provide some components early for software and architecture development, testing, and validation. By Q2CY2006 it is expected that the GDO will be in production.

Phase 1, started by Q4CY2005, will begin with the delivery of the two GIFEN units from Sun with three 10Gigabit Ethernet NICs. LLNL will provide the network routers and network connectivity to the Green network. These GIFEN will include four 73 GB or larger disks to allow for software development and testing, and architecture validation.

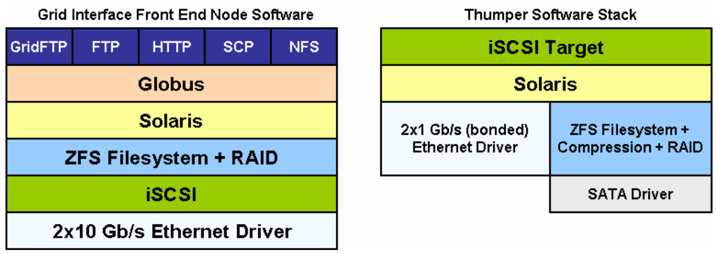
---

[2] ZFS claims:

- **Simple administration**. ZFS automates and consolidates complicated storage administration concepts, reducing administrative overhead by 80 percent.
- **Provable data integrity**. ZFS protects all data with 64-bit checksums that detect and correct silent data corruption.
- **Unlimited scalability**. As the world's first 128-bit file system, ZFS offers 16 billion times the capacity of 32- or 64-bit systems.
- **Blazing performance**. ZFS is based on a transactional object model that removes most of the traditional constraints on the order of issuing I/Os, which results in huge performance gains.

Figure 9: GDO Software architecture

In phase 2 the GIFEN CPU will be upgraded and the PCI-X bus will be replaced by PCI-Express (PCIe).



The throughput for FTP demonstrated between the 10 Gigabit Ethernet connected GIFEN machines will be 1 GB/s for file sizes ranging from 1 GB to 10 TB. The FTP client on one GIFEN will source data from ZFS on RAID5 disks via iSCSI over 10 Gigabit Ethernet to the bonded Gigabit Ethernet links to the disks. The FTP daemon will be on the other GIFEN to a separate ZFS on RAID5 disks via iSCSI over 10 Gigabit Ethernet to the bonded Gigabit Ethernet links to the disks.

Also in phase 2, 32 Thumper storage units will be delivered. Each Thumper unit will deliver at least 6 GB/s streams memory bandwidth, 150 MB/s aggregate TTCP performance, and 300 MB/s aggregate local ZFS file IO rate from the two ZFS file systems on the unit using IOR.

UCSD-LLNL Scientific Data Management project

The goal of this three year University of California – San Diego (UCSD) and LLNL project is to improve the conduct of science through the provision of scientific data management technology that enables the organization, manipulation, and analysis of observational and simulation data. The project is driven by two exemplary scientific applications:

1) Global climate modeling to determine the impact of climate changes on water supply. A Detection and Attribution analysis will be used to answer the question: Can we detect a global warming signal in main hydrological features of the western United States? This will involve making runs of global climate and downscaling models that will be unprecedented in scope.
2) Cosmology simulations of the early universe and the generation of simulated observations for the Large aperture Synoptic Survey Telescope.

UCSD will provide the driving applications for climate modeling and cosmology. The simulation codes will be based on current state-of-the-art algorithms. Regional models will be forced by the output of the global model simulations to improve resolution for assessing impact on water supplies. Adaptive mesh refinement extensions will be used to model galaxies. Both applications will be capable of generating hundreds of terabytes to petabytes of simulation output.

The simulation codes will be executed on compute resources at LLNL (primarily Thunder at the outset), with the output cached on storage systems at LLNL and stored on archives at SDSC. Data grid technology will be deployed between LLNL and SDSC to facilitate the movement of data to UCSD for analysis by researchers. Existing data collections at UCSD will be federated with the simulation output to enable comparisons with observation data, and support collaborations with researchers within the earth science and astrophysics disciplines. SDSC will research, implement, and provide support for data grid technology that improves the ability to organize, share, publish, visualize, and analyze both simulation output and observational data.

Data management research activities will be selected that improve the ability to conduct science research. Example research areas include:

- integration of the SRB data grid with data repositories at LLNL
- possible federation with the Earth Science Grid for access to the PCMDI collection
- development of data movement capabilities for terascale applications
- integration with disk caches for enhanced analysis of simulation output
- integration with visualization systems for display of hundred-terabyte 4D data sets

- integration with feature detection systems for automating analysis of simulations
- integration with dataflow management system for managing interactions with data
- integration with digital libraries for managing descriptions of features in the data
- demonstration of prototype LSST data management pipeline
- development of terascale database applications

Collaborations will be sought that facilitate interactions with other researchers within the scientific disciplines, broaden the scope to national-scale projects, and promote on-going research past the initial three-year period. Opportunities to support larger-scale projects such as the Large Aperture Synoptic Survey Telescope will be pursued.

Integrated Cyber Security Initiative (ICSI)

ICSI is an NNSA funded project to provide uniform identity management solution for NWC's classified environment that meets all of the federal policy and guidelines. ICSI is chartered to build the essential infrastructure (e.g., common hardware and software) at each of the NNSA sites and define the basic protocols for application integration. These services will be provided to the NNSA complex over SecureNet, often in a more limited capacity. ICSI is funded for the 2-3 year deployment of a set of production services. Beyond that, each site is expected to find funding to provide support for the ICSI services – that is, provide long term operational support. Current wisdom is that LC will provide long term operational support for an appropriate set of ICSI services. At this time, LC is expected to provide support for ICSI as prescribed below (ICSI funding may have been negotiated for some of these efforts):

- ICSI will set up an "institutional data aggregation" server (aka "people database", such as badging, clearance and training info for each person) on classified. LC will host the machine, perform system administration, and maintain the appropriate security plan (currently added to LC's security plan). This information is for LLNL use only and will be managed by LLNL's Computer Security Program (CSP) However, ICSI is required to provide a subset of this data over SecureNet to the other ICSI sites within the NNSA.
- LC will provide a regular (e.g., weekly) feed of the SCF (classified) LC accounts to the institutional data aggregation server (see bullet above).
- LC will host the machine, provide system administration support, and maintain the appropriate security plan for the classified Entrust server.
- Yet to be negotiated by ICSI is the expectation that LC will provide (via SAML and/or Kerberos) front-end authentication for LLNL applications on the classified.
- LC will provide help desk support for classified Entrust and future institutional applications on the classified such as IDPR (Integrated Design and Production Release), and "cross complex data exchange." The LC help desk will need to become knowledgeable about these LLNL applications, modify business process to include new and/or enhanced services (for example, modifying user forms to include secure email), assist in cross-site Entrust account management and troubleshooting, etc.
- The ICSI services will need access to SecureNet to provide these services between sites. ICSI will also need support for the Taclane IP encryptor on the unclassified network for ICSI's "Secure Test Environment."

ICSI has negotiated the majority of these requirements with LC, often providing the requested funding to LC for the efforts. We need to keep in mind that ICSI is funded primarily for the deployment of these services and not the long term operational support. Since these are services on the classified, LC will probably be asked to provide the long term operational support of these services while the funding model remains uncertain.

Additional information

- Network and security architecture details of Green Oasis, including plans for Green network upgrade and connectivity plans to ESnet 10GigE BAWAN connection, with peering upgrades to get 10Gbps to UCSD.

- Identification of and schedules for milestones that would impact LC.

3) I/O Throughput and Capacity Requirements and Analysis

The IO architecture ideally would be designed so that in totality, and for each system, there would be a perfect balance between 1) network connectivity and bandwidth capacity (e.g., number and aggregate bandwidth of NICs), 2) IO connectivity and bandwidth, and 3) the throughput capacity (e.g., machine's memory, processing and IO performance for all IO applications). This is probably an impractical if not unachievable goal, so compromises will be made. Over the long term this section should quantify the parameters above for the systems, and lend some discussion and rationale to the decisions reached in the final design and implementation of the IO architecture of the OCF and SCF.

Reminders of activities that are sources of requirements (see section 1 for more discussion):

1) Purple LA/GA on SCF: impact on network, storage, NAS, etc.
2) BG/L to SCF, and/or swings: impact on Lustre, network, storage, etc.
3) BGP in CY07-08.
4) TLCC to SCF (not likely to happen). Peloton SUs to OCF.
5) Green Data Oasis.
6) LLNL-UCSD collaboration on Data Management project.
7) ASC platform plans at other Labs: SNL/NM Red Storm upgrade, LANL funded for new platform (nothing known so far).
8) Higher speed (10Gbps vs current 2.4Gbps), dual path DisCom WAN in early to mid-CY06.

Facility Network

The network requirements will be set in large part by the implementation decisions of the other IO architecture elements: computing resources, storage, NAS, Lustre, and so on. That notwithstanding, some analysis can be done with the information at hand.

The growth in the number of network chassis, 1 GigE ports and 10 GigE ports (see ***Table 1***) is rather significant: about a 40% growth in the number of chassis, 40% growth in 1 GigE ports and 140% in growth of 10 GigE ports – excluding the Federated switch. This raises several concerns. One is the cost of maintenance (which scales with the number of 65xx chassis; other hardware is more complicated), about $350K in FY05. Another is the management of these systems, not only for performance and errors but also for configuration control.

System and NIC Type	B/w (Gbps) per NIC
Ext	0.25
Int (e.g., NFS client)	0.60
1 GigE JF	0.60
10 GigE JF	5.00
Lustre Gwy/Node 1GigE	0.60
Lustre Gwy/Node 10GigE	5.00
1 GigE OSS/OST	0.50
10 GigE OSS/OST	5.00
NFS Server - NetApp	0.40
NFS Server - BlueArc	0.40
NFS Server - Panasas	0.40

Table 6: FY05 per NIC throughput by network and server type

Table 6 (identical to ***Table 2***, but duplicated here for convenience) shows the estimated throughput for the network interface for each type of network and service. These numbers should be re-validated periodically in the IO Testbed using the same hardware, software, and applications that are in production. For this year, they are an approximation based on network benchmark performance testing.

Using the network interface throughput in ***Table 6*** and the number of network interfaces in each system, an approximation of the aggregate throughput for each major service is provided in ***Table 7***. Note that the OCF NFS clients ("Machines – Int") can source and sink about 55 Gbps while the NFS servers can source or sink about 33 Gbps. Assuming that the load generated by the clients is reasonably spread out in time, a 2:1 oversubscription of the OCF server NFS bandwidth seems quite adequate. On SCF, the oversubscription is about 4:1 – quite different from OCF.

	OCF Aggregate Throughput (Gbps)	SCF Aggregate Throughput (Gpbs)
Machines - Ext	25.3	49.8
Machines - Int	55.2	119.4
Machines - Jumbo Frame	25.8	185.8 (160 from Purple)
NFS Servers - Ext	9.0	7.3
NFS Servers - Int	33.2	31.2
Lustre clients	862.8	38.4
Lustre servers (OSS/OSTs)	480.0	56.0
HPSS Movers - Ext	8.3	9.3
HPSS Movers - Jumbo Frame	79.2	88.8
HPSS SLIC - Ext	2.5	N/A
HPSS SLIC - Jumbo Frame	24.0	N/A

Table 7: FY05 aggregate throughput estimates for resources on OCF and SCF

Similarly for Lustre, there is about a 2:1 oversubscription of the servers by the clients. Again, consideration of the actual throughput capacity of the clients and servers should be determined. For HPSS, the movers have the network connectivity to source or sink about 80 Gbps while the machines are at about 26 Gbps – a network connectivity oversubscription of the clients by about 3:1 for both OCF and SCF if Purple is ignored. Similar oversubscription exists between the machine clients and the HPSS SLIC gateway, both for serial FTP (over the External network) and PFTP (over the Jumbo Frame network).

The DisCom WAN will grow in FY06 both in bandwidth (from 2.4Gbps to 10Gbps), and connectivity (to SNL/NM, and added is a link to LANL). Also, Gigabit Ethernet IP Taclane encryptors will be used, replacing the ATM UltraFastlane encryptors currently in use. In 12-18 months we expect 10 Gigabit Ethernet Taclane encryptors to become available – quite late to make all 10 Gbps available over both links in a fully redundant architecture.

Recommendations

Investigate how to manage the network components effectively, considering C&A, manpower, staff skill level, cost of tools, and benefits. Historically network management tools provide few benefits compared to their total cost. With the auditing requirements imposed by laws on industry, new tools may prove worthwhile. Tools should be investigated and specific recommendations drafted (e.g., currently looking at *Opsware* for C&A configuration management, and product *statseeker* for performance and error reports and real-time analysis (likely to replace our HP Performance Insight product).

Gather network requirements from visualization, NAS, HPSS and Lustre. Develop a plan for a balanced network connectivity and throughput deployment for those elements.

Re-design network to accommodate new DisCom WAN connectivity and bandwidth using new 1GigE IP Taclane encryptors. Compromises are expected in redundancy and the bandwidth actually made available over the DisCom WAN for classified traffic. Track availability of 10 Gigabit Ethernet Taclane encryptors.

Network Attached Storage

For the purpose of analysis of NAS requirements we will *exclude systems with global parallel file systems or Lustre file systems*. Doing this, the OCF has 3TF and SCF has 17TF of capability and capacity computing.

Conventional wisdom is to provide 0.0005 Bytes/s per FLOP/s <u>each</u> for productive and defensive I/O. This is a total of 8 Gbps per TF. After excluding the appropriate systems (above), the NAS throughput requirement is 24

Gbps for OCF and 136 Gbps for SCF. ***Table 7*** indicates that the throughput that can be sourced or sinked from the aggregate of all Internal interfaces on all machines (i.e., all NAS clients) is 55 Gbps for OCF and 120 Gbps for SCF, which exceeds or matches reasonably well the requirements above. The estimates in this table also suggest that the current bandwidth of all NAS servers is high by about 50% on OCF and low by 4x on SCF.

	OCF (TB)			SCF (TB)		
	FY04	FY05	FY06	FY04	FY05	FY06
Admin	6.4	8.4		3.1	4.0	
Home	18.2	18.2		10.3	10.3	
Project	13.0	13.8		7.0	7.0	
Scratch	80.0	94.0	60.0	80.0	96.4	60.0

Table 8: NFS storage capacity on OCF and SCF for each of the 4 partitions

Conventional wisdom is to provide at least 20 bytes of globally addressable disk space per FLOP/s or 20GB per TF. Again *excluding systems with Lustre or locally global file systems*, the NAS must cover for 3TF on OCF and 3 TF on SCF. These generate NAS capacity requirements in FY05 of 60 TB for OCF and 60 TB for SCF. The current NAS capacity shown in ***Table 8*** meets the requirements for OCF and SCF.

Recommendations
The machines to be deployed in FY05 – BlueGene/L, Purple, and Peloton –include sufficient globally (within that cluster) addressable disk space at sufficient throughput.

Scalable I/O

See next section, Issues and Recommendations.

Storage

Note in ***Figure 10*** that while HPSS reads increased modestly since 2000, the HPSS writes are increasing at a high rate: for SCF 33% per year since 2001 and for OCF 200% per year since 2002. This difference in writes (and presumably storage capacity since file deletions are in the noise) between OCF and SCF may be due to the greater growth in computers and/or users, or difference in "usage model" and/or applications being run.

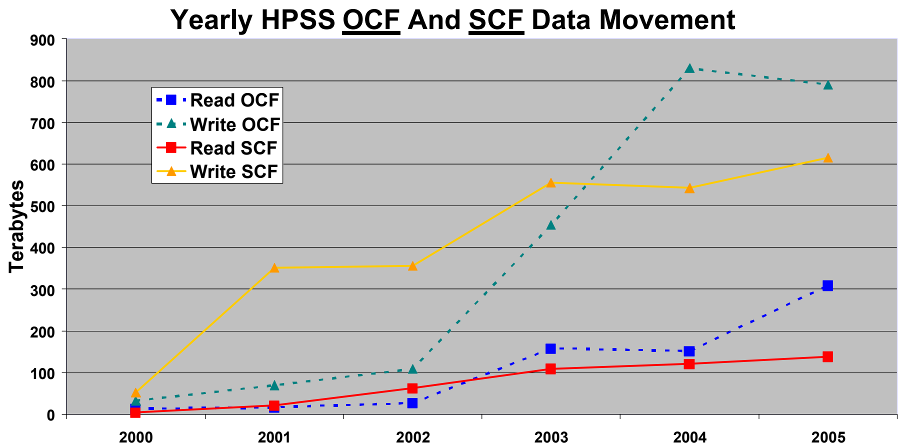


Figure 10: Yearly data read and written to OCF and SCF HPSS

Data storage capacity used for OCF and SCF doubles about every year after 2000 (***Figure 11***, note the vertical grids are spaced every 2 years). This could be tracking the computing capacity in each facility, throughput to storage media, or data movement/management tools. Regardless, an appropriate HPSS storage capacity should be planned and funded.

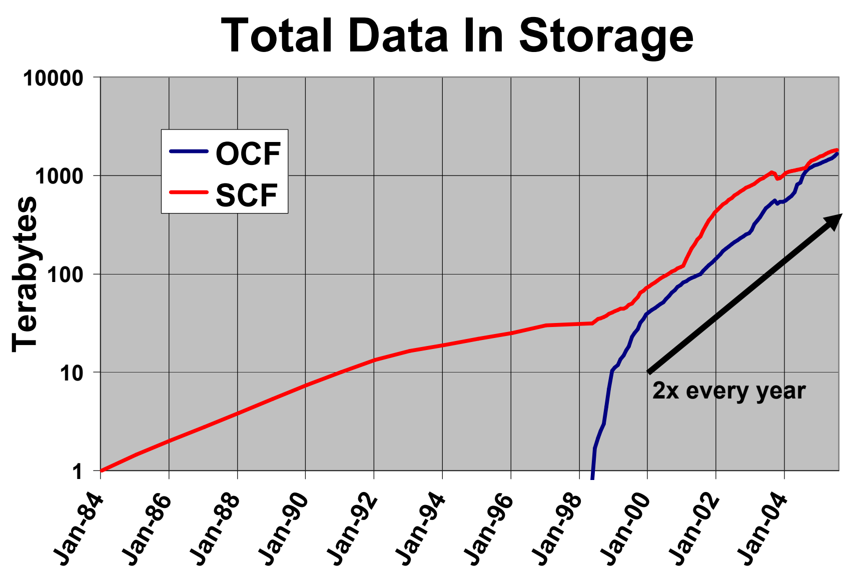


Figure 11: Total data in OCF and SCF storage

The tables below (***Table 9***and ***Table 10***) provide some information on the storage capacity and network connectivity for the HPSS system and movers. This ignores the actual network design, which aggregates the HPSS mover network interfaces provides a "trunking" bandwidth to the facility's core network (see ***Figure 5***). Using the network throughput noted in ***Table 6*** above, we see that in FY05 the aggregate HPSS throughput from all movers and gateways for clients via the Jumbo Frame network (e.g., parallel FTP, DisCom WAN) is about 80 Gbps. HPSS provides about 8 Gbps for the clients via the mover's interfaces to the External network (e.g., serial FTP, etc). In FY05 HPSS set and met a "capability" throughput goal of 20 Gbps (2.5 GB/s).

OCF HPSS Storage Characteristics			
Parameter	**FY04**	**FY05**	**FY06**
Slot Capacity	5.3 PB	11.2 PB	
Tape Capacity Available	3.9 PB	4.9 PB	8.6 PB
Tape Capacity Used		1.7 PB	
Disk Cache	65 TB	75 TB	
Archive Nodes	33	33	33
Jumbo Frame GigE NICs	132	132	
JF Aggregate Peak b/w (Gbps)	12.3 Gbps	79.2 Gbps	
Migration Jumbo Frame GigE NICs	132	132	
Migration JF Peak b/w (Gbps)	12.3 Gbps	79.2 Gbps	
External GigE NICs	33	33	
External Aggregate Peak b/w (Gbps)	8.3 Gbps	8.3 Gbps	

Table 9: HPSS Storage Characteristics for OCF

SCF HPSS Storage Characteristics			
Parameter	**FY04**	**FY05**	**FY06**
Slot Capacity	4.4 PB	12.2 PB	
Tape Capacity Available	3.4 PB	4.4 PB	8.5 PB
Tape Capacity Used		1.8 PB	
Disk Cache	70 TB	75 TB	
Archive Nodes	36	37	37
Jumbo Frame GigE NICs	144	148	
JF Aggregate Peak b/w (Gbps)	21.0 Gbps	88.8 Gbps	
Migration Jumbo Frame GigE NICs	144	148	
Migration JF Peak b/w (Gbps)	21.0 Gbps	88.8 Gbps	
External GigE NICs	36	37	
External Aggregate Peak b/w (Gbps)	9.0 Gbps	9.3 Gbps	

Table 10: HPSS Storage Characteristics for SCF

A strategy should be developed and articulated by the HPSS team regarding throughput goal. One characterization of a strategy, adopted from platforms and Lustre (see Lustre section in Section 2 above) is capability vs capacity. In this context, the capability strategy would be for the goal to meet the requirements for the largest single user/application. The capacity strategy would be to meet the peak aggregate throughput (applying some statistical distribution) of all users/applications. This strategy would then be used to determine the throughput goal to which the network is designed. Specifically the “trunking” between the HPSS aggregating switches and the core network should meet the throughput requirements for HPSS. Since the large HPSS throughput in the tables above is driven less by throughput requirements and more by the design of HPSS (e.g., accommodating COS, striping, storage capacity, etc), the trunking bandwidth provided by the network from HPSS to the facility core network is expected to be less than the bandwidths in the tables above. For example, in FY05 the HPSS movers were connected to the Jumbo Frame network with 80 Gbps throughput, but the throughput goal for HPSS was ¼ that – 20 Gbps (2.5 GB/s). Hence the network trunking from the core to HPSS should be about 20Gbps – which it is in FY05 for both OCF and SCF.

Year	Read OCF		Write OCF		Read SCF		Write SCF	
	TB Per Year	Mbps (ave)	TB Per Year	Mbps (ave)	TB Per Year	Mbps (ave)	TB Per Year	Mbps (ave)
1998	3	1	9	2	1	0	10	3
1999	21	5	26	7	3	1	33	8
2000	14	4	33	8	5	1	52	13
2001	17	4	70	18	21	5	352	89
2002	27	7	109	28	63	16	356	90
2003	158	40	454	115	110	28	556	141
2004	150	38	824	209	121	31	542	137
2005	309	78	791	201	139	35	615	156

Table 11: Trend of data movement by HPSS on OCF and SCF

Over the past 5 years or so the growth of storage capacity has grown on average a factor of 2 per year (see ***Figure 11***). However, this growth has in fact occurred in jumps – see ***Table 11*** e.g., year 2003 on OCF writes and year 2001 on SCF writes for example. This has likely occurred in response to a new computing resource becoming available (e.g., White on SCF in 2001), but should be verified.

Another point of interest is that on both OCF and SCF the HPSS storage has over 20 Gbps bandwidth to the facility's core network switch but (obviously) much less bandwidth would be required to move that data on average (e.g., 800 Mbps in 2005 for OCF writes). Similar to all resources in a high performance computing facility, the IO design is heavily influenced by the peak utilization, which was demonstrated in FY05 to be 2.5GB/s (20Gbps) for Purple using the same tools available to users. Empirically (in this example), assuming the peak throughput is indeed adequate for users, the peak to which the IO architecture is designed is about 25 times the average (over a year/month/week) sustained throughput – a potentially interesting design parameter, perhaps could be used to correlate between increase in capacity and (peak) throughput.

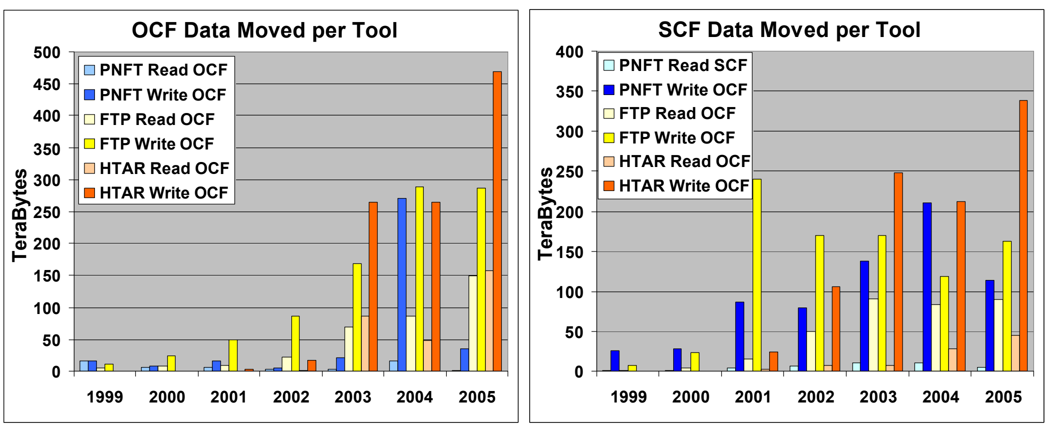


Figure 12: Data movement tools, data moved

Data movement tools have been designed, written and re-written to address many user requirements. Several tools provide a useful user interface while actually using other tools to perform the data movement. Examples are HOPPER (which uses almost all data movement tools, as appropriate for the situation) and PNFT (aka NFT). In ***Figure 12*** the amount of data moved by various tools on both OCF and SCF are shown. The numbers for PFTP (which is used by PNFT for data movement) were adjusted for PNFT. HTAR logs as one file an operation that actually moves a number of files (e.g., a directory of 1,000's of files) so the figures showing the number of files (see ***Figure 13***) miss that information.

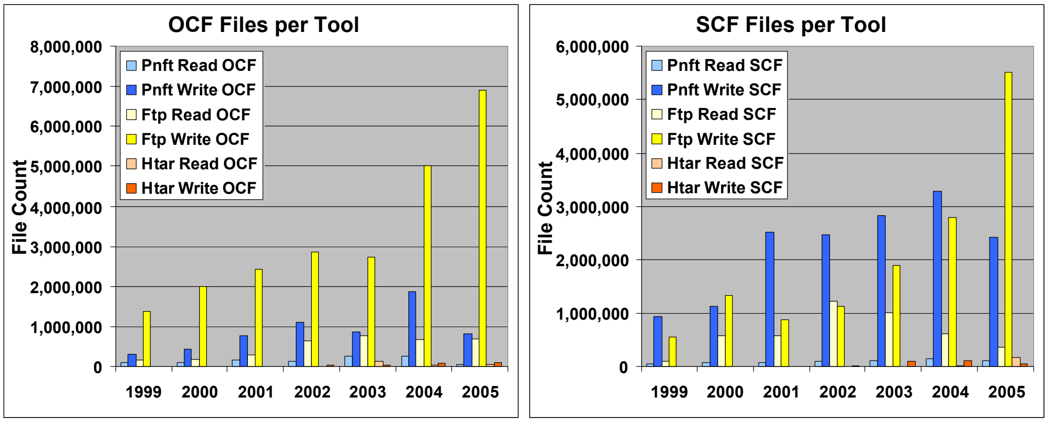


Figure 13: Data movement tools, number of files

Recommendations
Plan to provide adequate storage capacity in FY06 to meet the rapid growth of storage. Verify how capacity used increases significantly e.g., when a new computing resource becomes available. Consider how to change user behavior, tools, and facility design to curb this exponential growth e.g., consider how Lustre (a SWGFS available to every computing resource in the facility) should change user behavior regarding file archiving, consider a re-compute strategy, re-consider quotas, etc.

Develop and articulate a strategy for providing throughput between HPSS and the computing resources in the facility, via the Jumbo Frame network (and hence including remote users on their facility platforms via the DisCom WAN), the external network (and hence remote users on their desktop via the lower bandwidth SecureNet), and via the Lustre network. Design for peak throughput as appropriate (funding etc considered).

There are preliminary discussions for establishing (and possibly developing) a tri-Lab data movement tool. This could provide an opportunity to consolidate the features in the existing tools. Also, achieving good performance over the new 10Gbps DisCom WAN should be a goal for this tool.

Pursue plans for using the DisCom WAN for archiving classified and unclassified data. This may influence the local design of the DisCom WAN.

Lustre, Inter-cluster (Site Wide Global File System)

I/O requirements (i.e., throughput and capacity) of the Site Wide Global File System (SWGFS, aka Cluster-Wide File System in ***Figure 3***, and presently implemented as Lustre) are that of the single, largest machine. The throughput requirement of the SWGFS is to only meet the peak throughput requirement of the largest machine on site (and NOT the aggregate peak throughput of all machines on site). It is assumed that in practice any machine's peak throughput is of relatively short duration, and occurs infrequently. It is also our experience that file system capacity requirements are typically met once the throughput requirements are met, due to the rapidly increasing capacity of disks. Another outcome of this requirements strategy is that the SWGFS must have one metadata server (versus a federation of metadata servers and associated storage systems) since the largest capability machine is expected to meet peak throughput requirements by accessing the entire SWGFS.

Recommendations
Due to the lack of experience and usage statistics with Lustre in production operation, improved guidelines for Lustre I/O requirements are lacking at this time.

Visualization

See visualization Blueprint.

4) Issues, Analysis and Recommendations

Meta-issues for FY06 and into FY07

Lustre implementation end-goal and strategy

Lustre, the implementation of LC's site-wide global parallel file system, was intended to meet several requirements for LC's classified and unclassified high performance computing (HPC) facilities. The first is to provide a file system that can be shared by all computing resources in LC's HPC facilities. The second is to meet the high IO throughput requirements of file systems for the HPC computing resources (specifically the platforms and visualization resources). The third is to minimize costs of the computing infrastructure that supports the platforms.

Potential Issue	Centralized strategy	Distributed strategy
1) Reliability / availability		Pro • Multiple cross mounted FS improves FS fault tolerancy. • Multiple FS reduces impact of component failures or downtime.
2) Cost a) Either deployment strategy requires severe constraints on deployments to realize cost benefits.	**Update platform every 1 (one) year, sum of costs 2006-2013**	
	$166.5M	$193.1M
	Update platform every 2 (two) years, sum of costs 2006-2013	
	$120.5M	$123.1M
	Pro • Generally less costly, cost differential decreases as time between platform deployments increase	Pro • Automated administrative tasks more likely with homogenous hardware. • HPSS archival implementation of smaller peak bandwidth for Lustre file system backup.
3) Performance		Pro • Each platform will have its own file system providing required throughput, less competition for resources. • More Metadata servers, higher aggregate transaction rate. • Homogenous hardware improves parallel performance.
4) User impact	Pro • Few (2) FS on which users will have files. • Fewer paths to files for applications and users.	Pro • FS primary use would be by local platform, mitigating performance and reliability issues. • Improved strategies for Lustre s/w and h/w upgrades. • Fewer users impacted when one FS has problems or down for upgrade/repair. • Reduced purge time. • Changes in applications and scripts to use multiple FS are considered minor and acceptable.

Table 12: Summarized comparison of Lustre deployment strategies

Two deployment strategies

Appendices F and G compare in some detail two deployment strategies of Lustre. The **Centralized strategy** is to create two file systems (FS) and locate the storage in two locations (B439 and B451). As new platforms are deployed, the Lustre storage is alternatively upgraded in the two locations to meet the Lustre storage bandwidth and capacity requirements of the largest machine. The **Distributed strategy** is to co-locate the storage and platform ideally in the same computer room, and purchase the Lustre storage required for that platform. This strategy will result in one Lustre FS per platform that can be cross-mounted to allow access by all computing resources.

Comparison of the two strategies, summarized in ***Table 12***, raise many concerns related to the impact on users resulting from the operational support of Lustre with only a few file systems. In general, the operational criticality of keeping all file systems working reliably, with high availability and expected performance is vastly greater with only a few file systems. Meanwhile the proliferation of file systems in the Distributed strategy is characterized as a not too serious inconvenience to the user community, especially since the expectation is that most applications will use only the file system local to the platform on which it is running.

Aside from fewer file system name spaces, the other main advantage of the Centralized strategy is cost. This is discussed in considerable detail in Appendix F. Briefly, the cost savings in this strategy arises from the re-use of the storage in one location to meet the Lustre bandwidth and capacity requirements of the largest platform. The cost advantage of the Centralized strategy is highest when new, larger platforms arrive more frequently since the value of the reused disk depreciates over time, and hence is of higher value when less time elapses. Other factors also influence the costs rather dramatically.

Recommendation

After considering cost and technical factors, the recommendation is to maintain a distributed file system (with global visibility) strategy and progress toward a more centralized file system strategy as various Lustre performance and reliability issues are mitigated.

Archive challenges surrounding BGL and Purple for FY06

The addition of Purple and BGL, two enormous compute engines, brings into question the ability of the archive to handle the data volume and bandwidth that these resources will require. Will the archive be able to handle the load given declining budgets and Data Storage Group (DSG) manpower?

Planning Strategy

The LC has placed a strong emphasis on advance planning; identifying archive requirements based on computational resources as they become evident. This type of planning helps the center to identify opportunities to make early investments in storage and network technologies that will meet these demanding requirements. This proactive approach affords the LC adequate time to pre-deploy large-scale archive hardware and grow the archive capacity and bandwidth gradually, amortizing the cost and work of expanding the archive. The strategy not only includes building early to meet future requirements, but also verifying that the necessary infrastructure (including network) actually meets the specifications through demonstration of DOE L2 milestones (e.g. the FY05 Purple L2 milestone).

Although the LC has always placed an emphasis on advance planning, storage technology is highly dynamic making it extremely challenging to architect and procure hardware ideally suited for meeting the archive requirements before the technologies actually exist. In recent years, the archive procurement strategy has become more agile than in the past, enabling a flexible approach to hardware selection. Just-in-time hardware procurements prevent the archive from being caught in the trap of purchasing older, slower technologies that soon become sub-optimal in terms of speed and capacity.

The LC's strategy has been, whenever possible, to aggressively procure equipment in advance of the arrival of compute infrastructure. At the same time the LC has been using a just-in-time procurement strategy for media and for those items that shouldn't or can't be purchased well in advance of production.

Purple and BG/L Archive Preparation

From day-one, LC management recognized the need for an investment in infrastructure commensurate with supercomputer procurements. FY04, '05 and '06 archive budgets were well funded. Many of these dollars were used to purchase and deploy infrastructure aimed at Purple and BG/L requirements. New STK SL8500 robotics were purchased, capable of serving our needs for at least a decade. STK Titanium drives were procured, quadrupling drive bandwidth and more than doubling cartridge capacities. Combined, these technologies represent over 19.5PB worth of slot capacity in the SCF alone. Lower cost LTO tape technology was purchased and aimed at providing more cost effective capacity for second copies on tape. New, enterprise-class Brocade FibreChannel Switches were installed and both Linux and high-end IBM server technologies were purchased and installed as HPSS movers. HPSS scalability and longevity improvements were designed, developed and deployed in direct support of an order-of-magnitude increase in user load.

While we were actively deploying these hardware and software technologies, BG/L and Purple arrived. Archive, networking and platform personnel were challenged with an FY05 L2 milestone to prove that an adequate I/O infrastructure could be fielded to support Purple. The milestone was successfully met by demonstrating a 2.5GB/s data rate from Purple to the archive. This proof of concept verified that both our software and hardware infrastructure preparation was on-track.

Initial production use of Purple, and especially BG/L, on the unclassified side put an extreme load on the archive. Write and read rate records were set, almost doubling previous highs. In response, DSG personnel rushed LTO tape technology into production to help absorb this load. In the end, the network and archive infrastructure on the OCF proved capable of sustaining production BG/L and Purple loads.

Remaining Challenges

The fact that initial BG/L and Purple production load was handled in OCF does not mean that there are no concerns surrounding these platforms and their file systems. Of immediate concern is the ability of the SCF infrastructure to mimic OCF. Shortly after the systems swing DSG will have to rapidly deploy LTO technology and supporting robotics in the SCF. Sustained production will require similarly rapid deployment of Titanium tape drives in SCF.

The greatest risk to being able to bring these technologies to bear in the SCF on time is manpower. The DSG archive team has lost almost 4 FTEs worth of manpower without replacement, this while deploying and maintaining more and more hardware, software and user data. We no longer are able to proactively deal with problems, but rather are reactive. In the case of fielding new technology in the SCF, we will unfortunately have to be reactive.

BG/L and Purple have shown that users will continue to generate millions of small files. While many will be aggregated using HTAR, a large portion will not. HPSS Release 7.1 will be focused on small file performance (among other scalability issues). While this work is at least 18 months away from being fielded in production, we will be in significant need of these capabilities by the time they are ready.

Electrical power is another challenge. Computing and processing resources hold-sway in LC floor space and power decisions, not infrastructure. Available power is running out, or has run out in B451 and B453 archive areas (B451 is of greatest concern). The lack of locally available power may well force equipment moves (because of the requirement to site movers near associated devices). This is expensive and is especially taxing on DSG manpower.

BG/L swing scenarios are a concern, but are likely minor. Any swing scenario will need to include provisions for a SLIC front-end cluster for both SCF and OCF. OCF archive deployments over the last two years (robotic, tape, disk, FibreChannel…) have removed most archive concerns surrounding swing scenario. Just-in-time media procurements will allow for intelligent allocation of archive dollars during swing periods.

How users actually make use of BG/L, Purple and associated file systems raise many unanswered questions including:

- Will the usage model for Purple and/or BG/L be different from the expected usage model? As computational resources grow, users might find the need to re-design their codes and in so-doing find that they also need to archive either larger amounts of data, or a larger number of small files.
- Users have recently began to consider the archive as a dynamic, random-access storage space, reading more and more of their data back than in the past. What effect will increased read demands have on overall throughput to the archive? Should dual-copy file-size boundaries be reconsidered to increase reliability? Will users be satisfied with the lower read-rate requirements that have been agreed upon to date, or will they demand faster access to stored data?
- Is Lustre going to be successful as a single, global file system, or will islands of file system storage continue to be the model? Developing an HPSS-to-Lustre interface (and inter-connecting hundreds of HPSS servers to Lustre) will be considerably different depending on which file system model is used.
- Will the archive usage model change when users have near-line access to storage from the Lustre file system? Users might choose to archive even more of their data if they find it's easier and more seamless to do so directly from Lustre.

- Will the planned deployment of a tri-lab user interface spark increased usage from offsite computing resources? Will the new tri-lab interface provide a higher peak bandwidth to the archive, making higher capacity needs a reality?

These questions are ongoing topics of discussion in the LC. As BG/L and Purple ramp up production in the SCF, the answers to many of these questions will provide the impetus for significant thrust areas for archive and other infrastructure teams in the LC.

Tactical Issues for FY06

Lustre, Inter-cluster Issues

In FY05 Lustre served as the major file system resource for four major compute platforms providing a total of around 50TF of production compute resource. But throughout most of FY05 the majority of our Lustre development and testing resource was devoted to installing, testing, and resolving issues related to the BlueGene/L system. We progressed from a point where Lustre was barely functional and running at a few gigabytes per second to our current state where we provide around 900TB of storage running at near theoretical disk speeds with a very high degree of reliability. Unfortunately the cost of focusing the majority of our Lustre resources on BGL has been that many of our other high priority projects have been ignored.

Issue A: Lack of Lustre statistics regarding RAS, MTBF

There is no mechanism in place to acquire metrics for RAS (Reliability, Availability, Serviceability) or MTBF (Mean Time Between Failure) of a Lustre file system. There remain questions of what to measure to get these metrics, such as component uptime or availability of files or availability of data blocks/objects. These metrics are complicated by inherent fault tolerancy built into Lustre, such as dual access disks from multiple controllers and OSS machines.

FY06 Action/Recommendation

Further study is needed to understand how to define these metrics to be most useful in operations and prioritization of development, optimization, and troubleshooting efforts.

Issue B: Design and implement an OST failover strategy

A mechanism is lacking for detecting an OST problem and dynamically and, to the user transparently, configure around this failure. The high level design has been worked out, but resources are lacking to address this shortcoming at a higher priority. This development effort would be done by LLNL since it is not in the scope of CFS.

FY06 Action/Recommendation

The high level design is to develop a task to test and verify the health of the OSS servers and components, by probes and queries to the MetaData Server, and then determine the action to take. This may include updating the resource tables in the Lustre clients to move load onto other OSSs, for example. Presently the OSSs are working reliably so there is no urgency. However, when we do start having problems with the OSS machines the impact on Lustre may be severe.

Issue C: Develop, test and deploy InfiniBand connected Lustre storage

InfiniBand is an interconnect technology that is also suitable as a low cost, high bandwidth, reliable storage and system area network technology. However, some InfiniBand features needed for this environment are either immature (e.g., Subnet manager) or lacking (e.g., routers) at this time. Meanwhile the next generation machines, including BGP, continue to have Ethernet integrated into their basic components, with software to drive Ethernet. This is making it more difficult to consider technologies other than Ethernet for the system area network. It is comparatively simple to meet our requirements with a homogenous InfiniBand communication implementation, but is much more complicated even at the middleware level to fully utilize InfiniBand's features in a heterogeneous communication fabric. Presently running TCP/IP over InfiniBand provides unacceptable throughput and/or latency, but would otherwise work in a heterogeneous fabric.

FY06 Action/Recommendation

We continue to meet with InfiniBand vendors to express our requirements for routers and Subnet Managers that are full featured. Resources are also committed to OpenIB to promote the features we require in the InfiniBand Open Source development efforts.

Issue D: Lustre bandwidth issues
The goal in FY06 is 32GB/s sustained Lustre IO throughput. The current disks are limited to 32GB/s writing and less than 30GB/s reading, so this goal is unlikely to be realized.

FY06 Action/Recommendation
After considerable testing and optimization, the throughput limitation of the disks used for Lustre was acknowledged. As a result, this throughput issue has been set aside and the limited Lustre resources have focused their attention on some of the other Lustre issues.

Issue E: Lustre global access performance issue
Consider the practical Lustre environment where a few or (especially in the case of a site-wide global file system) many machines access Lustre simultaneously. Performance will suffer in all metrics for file system: IO speed, file creation, etc with more than one machine using Lustre (e.g., Gauss and BGL). Further, some tests give different performance depending on the Lustre capacity that is used (e.g., nearly full versus nearly empty).

This scenario has been tested in a scaled down scenario with only BGL and Gauss accessing Lustre simultaneously. In this case 65,000 Lustre clients perform a read or write operation on Lustre while 256 Lustre clients, one for each Gauss node, try to access Lustre for "real-time" interactive visualization. The result is that Gauss was effectively locked out of Lustre and ineffective as a visualization resource.

FY06 Action/Recommendation
Studies continue for this issue so there is no recommendation yet. However, one observation is that this issue would be exacerbated by a site-wide global file system as it is currently being promoted, since it would have to be shared by all computational resources in the facility.

Issue F: Metadata speed for Lustre
Lustre operations that involve the MetaData Server (MDS), such as file lookup (e.g., ls command), file creation, file open are very slow. Results of early testing suggests that the problem is not in the hardware, but rather in the locking mechanisms of the MDS – either too many locks, costly locks, or long waits on the locks. This result came from tests that used ramdisk and the latest Lustre MDS software. The use of the ramdisk removes any performance issue with disk access or throughput speeds.

FY06 Action/Recommendation
Further study is recommended.

BlueGene/L Issues
This section discusses BlueGene/L issues to possibly be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: In early CY06 BGL will move to SCF
BGL will move to the classified facility by the end of February 2006. The Lustre file system, federated network switch, HPSS archive Lustre gateway, and visualization server will be moved with it. This leaves a substantial void on the unclassified facility.

FY06 Action/Recommendation
Peloton, and associated Lustre storage and servers, will be arriving on the unclassified facility in CY2006. Funding has been allocated for an adequate network backbone for Lustre (as well as the networking for Peloton) that will be used to interconnect Lustre resources elsewhere on the unclassified facility (e.g., MCR, Thunder). The visualization cluster Prism is targeted for the unclassified facility to provide a visualization capability for Peloton (visualization clusters PVC and Sphere already exist for MCR).

Issue B: Impact of additional 4 (u)BGL frames – network, etc
This was mostly a network connectivity issue, questioning how network hardware would be provided in the location of the new uBGL frames.

FY06 Action/Recommendation
A network chassis with open slots was located in the area of the deployed uBGL frames, and additional linecards were obtained to make the required 128 Gigabit Ethernet connections for uBGL.

Issue C: Swinging BGL between classified and unclassified

The requirement for making BGL available 20% of the time on the unclassified facility is again on the table. Issues related to this requirement include provisioning for an unclassified visualization resource, an adequate Lustre file system (and associated network), and a Lustre gateway to the HPSS archival system. This requirement was investigated a while back, resulting is a material cost of several $100K, but it is felt that this is old information. ICCD management must be advised of the costs and impact on the center for providing this capability.

FY06 Action/Recommendation

Kim Cupps has the task of providing a cost estimate of this BGL swing capability by March 2006.

Purple Issues

This section discusses Purple issues to possibly be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Strategy for Visualization on Purple

The contract for the ASC Purple system did not include a provision for purchase of some number of nodes for visualization use as was the case with the ASC White system. A strategy needs to be developed to provide visualization capabilities for the users of the ASC Purple system.

FY06 Action/Recommendation

Similarly to what was done with White, on Purple there will be a 64 node SLURM pool for visualization. This will give visualization applications immediate and high speed access to application results on the Purple GPFS.

Issue B: How does the Purple/GPFS architecture fit into SWGFS model?

The SWGFS model makes an assumption that the major platforms have access to SWGFS. The recent visualization engines have taken advantage of this and have not purchased much, if any, hardware to support a local file system. The issue is: how can a system like Purple successfully utilize these new visualization engines given the SWGFS model.

FY06 Action/Recommendation

Lustre gateways were not developed for the AIX system, hence Purple will not be able to access Lustre directly. Data between Purple GPFS and Lustre can still be moved via the HPSS SLIC gateway using HPSS data movement tools. However, the biggest driver for Lustre access from Purple is to give separate visualization resources high speed access to application data on Purple. The recommendation, and in fact course of action taken, is to establish a SLURM visualization pool for Purple as its visualization resource. This recommendation and action removes the compelling requirement for Purple to be part of the SWGFS (Lustre).

Issue C: (Purple) GPFS file system architecture

Purple was delivered with nearly 2 PB of disk storage for its GPFS file system. A common practice on previous systems has been to create two file systems instead of one very large one. There is slightly more management overhead with multiple file systems. However, file system availability is higher because any issue with the storage subsystem will not be a single point of failure. The GPFS architecture for its metadata servers does not easily lend itself to easily creating multiple file systems. There could also be a performance impact, such that the bandwidth delivered to any one file system is not the target specified in the SOW. Therefore, the issue is: should Purple have one large file system or multiple file systems and if the latter is chosen, what is the metadata server architecture.

FY06 Action/Recommendation

White currently has 2 GPFS (file systems) specifically to provide greater availability. Since White became GA in 2001 there have been 2 reconstructions of a GPFS file system and 3 extended down times for data recovery. Anecdotal evidence is that the users have waited for the GPFS to come back online rather than change scripts and such to use the alternate GPFS.

Although leaning towards a recommendation of providing two GPFS file systems on Purple, further investigation will be done to better understand the impact of that decision. The GPFS metadata servers have been designed and optimized for one instantiation of GPFS (i.e., one file system). After Pu and Purple have been merged, performance testing will be done on GPFS as a single file system, and configured as 2 file systems. The resulting comparison of performance will be considered to decide whether there will be 1 or 2 GPFS file systems on Purple.

Issue D: (Purple) system usage model

The usage model for Purple (and other ASC platforms) is in the hands of an ASC tri-lab committee. It is entirely possible that remote tri-lab users (e.g. LANL) will be granted a significant allocation on Purple. Depending on the remote users' usage patterns, some amount of data will be sent back to the users' local site. We need to be prepared for this possibility (i.e. testing functionality and network bandwidth).

FY06 Action/Recommendation

After the Pu/Purple merge, test the functionality of tri-Lab access (i.e., authentication); this currently works with Pu so problems are not expected. Do performance testing between Purple and resources at the other two sites (e.g., Q at LANL and Red Storm at Sandia/NM), and fix performance problems to the satisfaction of the tri-Labs, thereby establishing the benchmark throughput level for data movement and remote visualization. Add Purple to the Netmon network performance monitoring to help monitor the network throughput between Purple and the other sites. Jean Shuler is working on the "Purple User Guide", which will include the Purple usage policy. This user guide should make recommendations that take into account the results of the aforementioned testing, for "best practices" for using local versus remote storage and doing remote visualization.

Peloton Issues

This section discusses Peloton issues to possibly be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Unclassified Lustre

When BGL moves to the classified most of the OCF Lustre storage and networking will go with it. This will leave a void for re-creating a federation of Lustre file systems on OCF. In particular, when Peloton is deployed with 800 TB of disk space for its Lustre file system, we will need a design and network equipment to create a federation of Lustre file systems on OCF, including Peloton's.

FY06 Action/Recommendation

Several Lustre designs have been proposed for the OCF after BGL has moved. The decision was to proceed with the design that keeps the machine(s), including Peloton, and the storage for its Lustre file system co-located. This design requires a minimal investment for networking with which to interconnect the remaining OCF Lustre file systems. This OCF Lustre "core" switch is a much smaller version of the BGL Lustre Federated Switch due to the smaller number of IO nodes and Lustre OSS connections.

Interconnects and System Area Networks Issues

This section discusses interconnect and system area network issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Support for open source OpenIB code

The Open Source OpenIB code provides software for InfiniBand (IB). Software includes the drivers and protocol stacks for the Host Bus Adapters (aka NICs, HCAs), and larger products such as the IB Subnet Manager. The issue is with support for the code. Multiple vendors will be providing IB products (e.g., HBAs, switches, etc), customers will want to use OpenIB source, and will expect it to work. When it doesn't, who will respond to problems encountered by customers that appear or are in fact problems with OpenIB codes?

FY06 Action/Recommendation

This is an outstanding issue. When discussed with Bill Boas in late 2005 he did recognize that this was an issue and had some thoughts on how to address it, but nothing in the works yet.

Issue B: Scalability

Problems have been observed with performance and latency of early IB implementations in a large cluster environment. Robust testing of large (100s of nodes) clusters with IB interconnect has not been done yet, so there is low confidence that most problem areas have even been identified yet. Other issues related to scalability are included in this section.

FY06 Action/Recommendation
LC staffer Ira Weiny has been tasked with becoming the IB expert at LC. To better address all IB issues we need some local expertise. The recommendation is to do more testing of larger IB clusters regarding performance, interoperability, management, support, and features to better understand issues we'll encounter when deploying larger IB clusters.

Issue C: Open source versus proprietary
Several IB vendors are developing products and software outside the scope of Open Source specifically to differentiate themselves from competitors. Arguments have been made that if vendors collaborate in developing Open Source products, the marketplace for the products will grow and many benefits, including profits, for all will follow.

FY06 Action/Recommendation
Vendors are listening to the arguments for Open Source, but it is not clear that they will execute accordingly. However, most vendors agree to be compatible with OpenIB products – a modest compromise to what is being requested of them.

Issue D: Congestion control
IB products currently have static "routing" tables. In a more complicated IB deployment e.g., multiple switches with oversubscribed trunking links, this can result in congestion. Emerging products will help system administrators identify congestion but otherwise not help in preventing it. IB vendors promise that products further out will employ more sophisticated, dynamic re-routing of traffic to detect and avoid congestion. Such mechanisms are already available in proprietary (non-IB) products from Quadrics, for example.

FY06 Action/Recommendation
The availability of these advanced IB features is often provided by the IB chip designers and manufacturers – of which there are only 2. We need to meet with the IB vendors regularly to get these issues on their product roadmap, and keep them there.

Issue E: Fault tolerance
The IB routing information today is discovered by the IB Subnet Manager upon power-up, and installed in the IB fabric elements (e.g., switches and HCAs). If trunking links or other parts of the IB fabric fail to function or have excessive errors, system managers need to learn of these failures. Ideally, the IB fabric will also dynamically reconfigure itself to accommodate those problems and continue to function until the problems are fixed.

FY06 Action/Recommendation
The availability of these advanced IB features is often provided by the IB chip designers and manufacturers – of which there are only 2. We need to meet with the IB vendors regularly to get these issues on their product roadmap, and keep them there.

Issue F: Technology & product support for large fabrics
Although IB as a technology has been around for many years, many features remain unimplemented or inadequately implemented. Building and supporting large IB fabrics, especially fabrics that encompass interconnects and system area networks (e.g., the Lustre Federated Switch), require features such as sophisticated management tools; dynamic reconfiguration to address congestion, errors and component failures; products to support routing and bridging; and protocols, specifications and products to support enhanced features (such as RDMA) across different communication technologies e.g., IB and Ethernet.

FY06 Action/Recommendation
We need to meet with the IB vendors regularly to get these issues on their product roadmap, and keep them there. Also, having knowledgeable staff with considerable experience with IB will help ensure our suggestions and comments are relevant to these vendors.

Network Issues

This section discusses network issues to possibly be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Upgraded DisCom WAN design, IP encryptors, throughput

The DisCom WAN has been re-competed with Qwest as the winner. LANL's request for redundant bandwidth to SNL and LLNL platforms influenced the final proposal. When this new DisCom WAN is deployed by mid-CY2006 the bandwidth will be increased to 10Gbps from the current 2.5Gbps, a northern redundant path will be added, and the protocol will be IP over Ethernet versus the current IP over ATM (all of this at a lower cost than our current contract).

Figure 14: Qwest DisCom WAN routes

A 10Gbps WAN will provide considerable bandwidth especially given the past usage of the DisCom WAN. It is unknown how the WAN usage will increase given the substantial new computational resources at LLNL, Purple and BlueGene/L. Also, the different IO architecture of BlueGene/L may drastically change the DisCom WAN usage. These factors combined with LANL's lack of a comparable platform resulted in their request for additional WAN bandwidth and increased protection from single points of failure.

The additional DisCom WAN route will mitigate some single points of failure present in the current architecture by adding a "north" route to the present "south" route through southern California that exists today (see ***Figure 14***). The tri-Labs are trying to establish redundancy also in the local route, again trying to remove single points of failure. For example, the south route goes from Sunnyvale to LLNL. Qwest initially proposes to bring the north route into Sunnyvale also. We are asking them to select another local route for the north route, such as Oakland, so the common point would then be LLNL rather than Sunnyvale to LLNL. Similarly, both routes meet in Albuquerque. Rather, we'd like the north route to terminate at Santa Fe and from there go to LANL, removing Albuquerque as the single point of failure.

Moving to IP/Ethernet from IP/ATM will allow us to retire some legacy network hardware, but will also require us to move to IP encryptors which currently are limited to 1GigE. The current strategy is to bring up no more than 4 1GigE encryptors with the 10GigE DisCom WAN, and in mid-CY2006 determine more accurately the availability of 10GigE IP encryptors.

The additional route and increased bandwidth of the DisCom WAN has raised other issues with the tri-Lab network groups. We are seeking other uses of the DisCom WAN, such as for Disaster Recovery archival between sites. We are also concerned that current data movement tools, and the source/sink resources at the tri-Labs, will be unable to utilize a 10Gbps link. Also under consideration is sharing the DisCom WAN for unclassified (but still ASC) uses between the tri-Labs. With the redundant path, dynamic routing protocols (versus the static routing that is currently used) are being considered. A disadvantage to some of these changes is that, if not configured and maintained properly, these changes could destabilize the DisCom WAN. Conversely, these efforts are trying to more fully utilize the DisCom WAN.

FY06 Action/Recommendation
The 10Gbps DisCom WAN local interface will be 10GigE. This demark will be located in B453. The LLNL owned network equipment and associated NSA Type 1 IP encryptors are also located in B453. Additional encryptors – either more 1GigE encryptors or 10GigE encryptors – will be purchased to make the 10Gbps WAN bandwidth available to the LLNL classified facility. By the end of February 2006 the tri-Labs will determine if they want to deploy two 1GigE IP encryptors and then upgrade to newly available and acquired 10GigE IP encryptors at some later date, or deploy four 1GigE IP encryptors (which means most sites will have to procure additional 1GigE IP encryptors). So initially, the 10Gbps DisCom WAN will offer 2 or 4 Gbps initially, with up to 8 or 10GigE by late 2006 or early 2007, depending on IP encryptor availability, requirements, and usage.

The legacy ATM based DisCom WAN equipment and encryptors in B115 will be moved to Qwest (from the AT&T WAN) by early February 2006. The 10Gbps WAN from Qwest will be available by the end of May. The 1GigE IP encryptors will be tested on the 10Gbps WAN for up to one month, and then put into production. This is when the DisCom WAN traffic is moved from the ATM WAN to the production 10Gbps WAN, and the legacy ATM hardware is retired.

Issue C: Merge Jumbo Frame network & Federated Switch (Lustre network)
As noted elsewhere in this document on both the OCF and SCF we have 4 logically separate networks: the external network (access to/from the Internet/SecureNet), internal network (no external access, primarily for NFS traffic), jumbo frame network (4 subnets to facility parallel data movement tools for transferring large data sets quickly), and the Lustre network (Federated Switch, with the SLIC gateway to the jumbo frame network). The assumption is that by merging the jumbo frame network and Lustre network that we will lower costs and/or improve throughput.

FY06 Action/Recommendation
This is not a high priority issue, although it periodically comes up in facility design discussions. No compelling reason for making this change has been identified yet, so the recommendation is keep the jumbo frame network and Lustre as it is today - logically separate.

Issue D: Improving network operations to meet IO requirements
Several factors have converged to raise the priority of deploying tools to help with network operations.

- Budgets are now tighter and the network budget is no exception. In previous years there was sufficient budget to provide the requested networking for platforms, visualization, NAS, and storage. Tools are needed to better understand the actual usage and apply network resources where needed.
- Security and auditing functionality is becoming increasingly important for all IT operations.
- Effective and/or subsidized network operations and management tools have recently become available. Network management tools have been notoriously complex, expensive and yet ineffective.

FY06 Action/Recommendation
Network monitoring tool Statseeker was recently purchased for deployment on the OCF and LabNet, and soon also on the SCF and the DNT local network. This is a tool that is easily configured and used by network support staff without substantial, prolonged training typically required of other tools. To goal is to provide network reports and information on specific areas of the network (e.g., storage, visualization, Purple, BG/L, etc) that will show or indicate average and peak utilization and errors over periods of a day, week, month and year.

Another tool under consideration is Opsware. This tool will provide configuration control, flag configurations or IOS versions that may have security risks or that deviate from best practices, perform inventory, and other auditing and network management functions. This tool was selected by the institution for the desktops and servers, and was selected by the LabNet network team lead (Jason King) for the network C&A (Certification and Accreditation) effort. Budget has been earmarked for this purchase, and approved.

Network Attached Storage Issues

This section discusses Network Attached Storage services issues to be addressed in FY06. Each issue is described and then a recommended action is presented.

Issue A: Growth of NAS bandwidth and capacity appropriate?

There is currently not an easy way to determine if our NAS systems that support the various areas (Admin, Home, Project and Scratch spaces) are scaled appropriately with regards to space, performance (I/O bandwidth, response time), and uptime.

FY06 Action/Recommendation

It is our recommendation that we implement a metrics gathering system to track space, performance, and uptime information over a period of time. With metrics data we will be able to determine trends. In addition to the gathering of the metrics data, baseline metrics for space, performance and uptime metrics should be determined and set. Once baselines have been determined and established we are recommending periodic testing to determine if we are meeting the baselines.

The type of metrics gathering environment needs to be determined but due to budget constraints it is highly likely that LC staff will be used to create scripts to gather the appropriate statistics. In the future, a commercial software product may be investigated.

At least annually we will review our metrics against our baselines and meet with LC customers to review NAS requirements such as uptime, space and performance. Base on the metrics and LC customer input we can then make appropriate recommendations regarding the scaling of our NAS environments (OCF and SCF).

Issue B: Replacement for DFS

Existing applications and services should stop relying on DFS by July 2006 since DFS is losing support. The use of DFS by the tri-Lab user community has diminished substantially over the years, but a few critical applications remain. Aside from some code groups (this requirement is not clearly understood), the DFS based Web service for SimTracker provides the most critical requirement for DFS. A replacement for DFS is needed to meet these needs.

FY06 Action/Recommendation

Tri-Lab consensus is the NFSv4 will replace DFS and fully meet the sharing and security requirements for the current applications of DFS. Currently lacking in the NFSv4 product availability is support for Linux. The Linux NFSv4 support is a development effort at the University of Michigan. Additionally, the Linux changes will need to be incorporated in CHAOS. Beyond that, IBM AIX NFSv4 product is available in AIX 5.3, but not earlier versions of AIX. So Purple can support NFSv4 but not White, since White will not be upgraded to AIX 5.3. Similarly, Sun has a good NFSv4 product. Network Appliance also supports NFSv4. NFSv4 implicitly supports NFSv3, the computing environment to evolve and grow with no special consideration as to NFSv4 support.

Scalable I/O Issues

This section discusses Scalable I/O issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Purple GPFS, BGL Lustre support on SCF

Data integrity, performance (both data and metadata bandwidth), stability, and functionality for Purple GPFS and BGL Lustre parallel file systems require testing and support. Several tools exist currently for this, but the tools' capabilities need to evolve as I/O issues are discovered. Some problems are discovered with these tools prior to users encountering them, but some problems are only found and reported by users. In addition to the support for these parallel file systems with regular testing, customer support is vital for making sure the application codes are making good use of the file system. This covers higher-level I/O library usage, application general I/O design, and file system-specific interaction with customer codes.

FY06 Action/Recommendation

Continued efforts on addressing any data integrity problems, in particular developing and enhancing tools to rigorously test the file system, are necessary for both Purple GPFS and BGL Lustre. IOR has been used successfully for stress, integrity, and performance testing, but further enhancements to the code for testing would be useful. A roadmap for IOR development has been established, and new functionality is being added. Further, additional tools may become necessary for testing and replicating user errors. I/O kernels -- I/O-only portions of real applications -- would be very helpful in understanding user issues with I/O. Such kernels could replicate problems more quickly and be more easily handed to vendors as test case examples. Continued customer support for user applications on these file systems is imperative: a high-performance parallel file system used awkwardly by applications does not offer much gain.

- Provide necessary enhancements of existing parallel file system testing codes, develop additional tools as needed.
- Isolate and make available any I/O kernels; this may be synthetic tests using testing tools or, preferably, actual code from real applications.
- Continue to support code teams using Purple GPFS and BGL Lustre.

Issue B: I/O library tuning on Lustre

Middle- and higher-level I/O libraries that are used on Lustre may need tuning to work effectively with the underlying file system and with the users codes. These libraries include MPI I/O, Parallel NetCDF, and HDF5. There has not been a concerted effort to study the performance characteristics and tuning possibilities for these libraries on Lustre.

FY06 Action/Recommendation

Probably the opportunity to run a study on the higher-level I/O libraries on BGL Lustre is not here yet. While it would have been good to complete such a study and tuning on the open side, the effort on Lustre has rightly been to improve performance, stability, integrity, and functionality of the parallel file system. But while the I/O libraries appear to pose any performance problems currently, studying and tuning of these libraries would likely pay off in performance improvements in the future.

- Monitor I/O library issues to determine when the user's I/O issues are less frequent at the file system level and more often with the library level. Pursue tuning these libraries when it is clear that it offers the highest return on effort.

Issue C: ASC Alliances

We currently have three Tri-Lab ASC Level 3 Academic Alliances for I/O research: UC Santa Cruz's research into scalable metadata, Northwestern University's (NWU) implementation of distributed caching, and University of Michigan's reference implementation of NFSv4. Each of these research contracts is due to complete at the end of FY06, and no money is available to continue these research activities with these universities.

FY06 Action/Recommendation

The work at UCSC shows promise with a distributed, load-balancing metadata server (MDS) approach. This has been gaining interest with IBM, and is hopefully seen as a possibility for multiple metadata servers for Lustre. Sage Weil has been leading this research and will be visiting in February to talk about the design of the distributed MDS. NWU's effort on distributed caching for MPI I/O has been working its way into MPICH's ROMIO source. Likely the fruit of this labor is paying off in this MPI I/O parallel library. University of Michigan's implementation of NFSv4 is something that needs to be adopted by more vendors before we embrace a full DFS replacement with NFSv4 (John Allen says that NFSv4 will replace DFS for current critical applications by mid-CY2006; see NAS issue B). The same is true with parallel NFS and its possibility as universal parallel file system interface to GPFS, Lustre, etc. With little money to continue funding these research efforts, the best recommendation is to see if the seeds sown blossom into products offered by vendors.

- Continue to support Sage Weil's work at UCSC if possible. His funding (distinct from the Level 3 Alliance money) comes from the LLNL Computation Directorate office and may be from a different pool of money than ASC.
- Keep an eye on Sage Weil and other graduating students for possible LLNL employment, again with evaporating dollars.

Issue D: POSIX Extensions standardization
There is an effort by the national labs (LLNL, LANL, SNL, ANL), academia (CMU, NWU, UMinnesota), and industry (Panasas, IBM) to improve POSIX performance for parallel I/O by extending I/O calls. Designed more for serial access than parallel, POSIX has shortcomings for parallel I/O that need fixing. While POSIX is not likely to be replaced, extensions to the POSIX standard for parallel I/O are possible. Some of the extensions would relax expensive coherence and metadata semantics, change data movement from streams of bytes to distributed vectors of bytes, allow group locking and file descriptors, inform storage systems of access patterns, and provide improved access control lists (ACLs) for security.

FY06 Action/Recommendation
It seems this is beginning to get a life of its own as more parties get drawn in and vendors in particular begin to show an interest in the possibilities for improving parallel I/O performance. The proposal for these extensions has been developed, and an open group has been established in the process of getting this into the POSIX standard.

- Maintain involvement to drive these extensions into a standardization that vendors provide.

Issue E: Supporting the Aurora strategic plan's initiative of Predictive Knowledge Systems
In support of Aurora, it is important to address the Predictive Knowledge Systems initiative, and likely this will need involvement from the I/O and Storage efforts. This initiative aims at extracting and analyzing information from enormous amounts of data. The I/O and storage's involvement would be, presumably, to address the data movement necessary for such an undertaking. As yet, however, it is unclear the specifics.

FY06 Action/Recommendation
The first issue, then, would be understanding the nature of the initiative and the issues involved. With more information, some questions would need to be pursued:

- Can we even move as much and as quickly as would be needed yet for this initiative? There may need to be initial research into new I/O models. This may be multi-Terabyte datasets in multi-Petabyte databases that would need to be mined.
- Computational analysis of data is likely viewed as the bottleneck, with I/O feeding the analysis a lesser issue. Is this even true? Is the larger problem looking at the data or moving the data to look at?
- There may be real-time or post-processing aspects that need consideration. For example, for post-processing existing data, can mining happen at the file system cluster level rather than pulling data to compute nodes? For real-time data, can mining happen in distributed caches as data is streamed in from multiple sources?
- To grind through that much data, which is perhaps collected from different sources, data format/conversion may be an issue. Is this being considered and addressed?
- Visualization involvement would likely be necessary for this activity. How would this happen, and is this something the Visualization team can pursue?
- Additional thoughts from Mike Zika on data mining/analysis in his recent talk at an ASC PI meeting include:
    - Mining of petascale datasets would require a new approach to data management. Some thoughts include a transparent lossless data compression (as an I/O layer?) or perhaps remote mining in storage for data extraction as there is not really enough room to keep multi-TB datasets on disk for very long.
    - Additional visualization/analysis clusters would be needed to connect to GPFS. Purple nodes on GPFS are too valuable to be used for this, particularly as much of the time they are idle. Instead, a separate, dedicated cluster to access GPS on Purple, much like the approach to using a Viz cluster on Lustre might prove useful.
- General recommendation: Investigate possible involvement in this initiative and where I/O and storage could contribute.

Storage Issues

This section discusses archival storage issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: HPSS for Green Data Oasis

The current implementation plan for the Green Data Oasis (GDO – see section x) does not allocate resources for an archival data storage system. With ~800 TB of file system storage capacity in the GDO, the authors believe that GDO users will quickly fill the file system (possibly within months), and then discover that capacity limitations and/or file system reliability issues demand that an archival data storage solution be deployed. Unfortunately, no such solution has been planned for or funded.

FY06 Action/Recommendation

Prepare a strawman proposal involving the investment in a modest set of hardware and the leveraging of the free site license of HPSS to provide a robust, scalable, high performance data archive for the GDO. Present this proposal as an option to LC management seeking guidance on whether to proceed. The proposal would allow GDO users to leverage local archival storage expertise and would provide a possible platform to be leveraged by the Aurora Predictive Knowledge System Project.

Issue B: Storage Media Protection

Today both classified and unclassified tape media require layered processes surrounding media tracking, labeling and security processes in order to ensure that sensitive data is never compromised. Enterprise tape vendors are actively developing device-level encryption technology which would encrypt data at rest on tape and disk devices. This technology is attractive and could prove extremely effective in the goal to provide ever-increasing levels of security for our nation's decades of investment in classified/unclassified archived data. Done properly, a lost tape would no longer be an issue and defective tapes could be sent to vendors for data recovery (something that is not possible today).

FY06 Action/Recommendation

In FY06, the DSG will keep abreast of this encryption technology as it matures. The DSG will also present the technology concept as a forward-looking option to the Computation OISSO and the Cyber Security Program (CSP) to identify opportunities and challenges (i.e. Key Management) that this technology might hold for LLNL. Assuming the Computation OISSO and CSP agree that at-rest data encryption could possibly be put into operation at LLNL, the DSG will then prepare a cost/benefit analysis for abandoning non-encrypting technologies. This analysis should consider optional levels of staged deployment: SCF new data alone; SCF new and legacy data; all data – SCF and OCF. Dependent upon cost and other factors, non-encrypting technologies from the SCF could be re-deployed in the OCF, making an initial investment in encryption for the SCF more affordable.

Issue C: Meeting tape throughput requirements: new drive, RAIT

There is concern that the STK Titanium (T10K) high-speed tape drive will be capable of outperforming the networks that are attached to the mover nodes that supply the tape drives. At 120 MB/sec, a single 1GbE link will be saturated, and will not be able to keep the T10K running at full speed. While link aggregation might seem to be an obvious option, there is skepticism as to whether or not it provides concurrent bandwidth. Some internal research has shown that although multiple links are active, the data is actually switching between links, using only one link worth of bandwidth at any given instant in time. Another option is to invest in 10GbE networks for the T10K mover nodes. This would be extremely costly and technically challenging, as 10GbE Network Interface Cards (NICs) and ports are expensive.

FY06 Action/Recommendation

Test bandwidth to STK T10K hosted on Linux tape mover nodes using single 1GbE networks and also using link aggregation. Report findings, and suggest network architecture that will balance performance with cost. Also keep abreast of software-based RAIT offerings in the public and private domain.

Issue D: Dual copy with geographically diverse silos

With the physical move of equipment to B451 and B453 complete, the opportunity to enable geographically remote dual-copies of data now exists. In FY06, the DSG will integrate the STK SL8500 libraries located in the B451/B453 areas, which are destined to house the first copy of Small- and Medium-sized files on LTO-3 media with the secondary copy being in silos on 9940B media, located in the B115N/S areas. The SL8500s will also be home to the secondary copy of Mission Critical data, on

STK T10K media, with the primary copy being on 9940B media in the silos. However, this will only address new data written into the archives, and does not account for previously written dual-copied "legacy" data that currently exist only in the silos. In order to provide optimal availability to data, legacy dual-copied data should be geographically separated, along with new data. The issue will be in determining how best to move these legacy data from the silos to the new SL8500s.

FY06 Action/Recommendation

Begin aggressive repack of the primary copy of legacy Small- Medium-sized data currently on 9840 media onto LTO-3. This should begin as soon as LTO-3 proves to be a viable option and enough new media capacity exists to store legacy data in addition to new data being ingested. The secondary copy of Small-Medium will continue to exist on 9940B media in the silos, and therefore will not need to be repacked. Both copies of legacy Mission Critical data will need to be repacked eventually, but initially, only the secondary copy (currently on 9940B media) will need to be repacked to move it from the silos to the T10K media in the SL8500s. This will also free 9940B media that is currently in high demand.

Issue E: Disaster recovery

With the initial goal of reducing Classified Removable Media (CREM) that has in the past been shipped to the Nevada Test Site (NTS) [typically SCF mission-critical customer/project data, home file system backups, HPSS metadata backups, etc.], the DSG is co-developing a Disaster Recovery (DR) solution in conjunction with the Storage Group at LANL. The current plan is for LLNL DR data to be dual-copied into the LLNL HPSS, and into the LANL HPSS [a separate name space], and likewise for LANL DR data. Currently, there are technical issues related to maintaining authentication during the upload, firewall challenges on the unclassified side, as well as performance issues when transferring large amounts of data over the DisCom WAN. The configuration of HPSS name spaces also must be agreed upon and implemented, such that each site's data will be kept isolated from other user data at the opposite site [this is to be implemented using HPSS File Families, which keep like data together on separate media].

FY06 Action/Recommendation

In FY06, the DR teams at LLNL and LANL will work together to resolve the technical and performance issues, agree upon HPSS configurations, and bring the DR solution into being in the SCF.

Issue F: Lustre interconnect with HPSS

An in-depth investigation into the technical merits of utilizing Liblustre as the enabling software to provide interconnect between AIX HPSS movers and Linux-based Lustre file systems proved futile late in the last quarter of FY05. There still exists the desire to provide a cost-efficient, high-speed interconnect between the Lustre file systems and HPSS. In FY06, the DSG will need to explore other options, beginning with Linux-based tape movers, and eventually Linux-based disk movers, as well. There are many technical issues that will need to be addressed with this approach.

FY06 Action/Recommendation

The HPSS tape mover will first need to be tested extensively under CHAOS and Linux in order to provide the optimal platform for Lustre interoperability. The DSG will also need to develop transfer-specific code for the various data transfer interfaces to provide transparent movement of data between Lustre and the HPSS name space without the user being required to use special local file semantics [lfget, lfput, etc.]. By the end of the year, the DSG will also need to perform a cost/benefit analysis weighing the option of deploying new Linux platforms for the disk movers, versus attempting to install Linux on existing IBM disk mover hardware to make best use of an already large hardware investment.

Issue G: Single data movement interface for Tri-Labs

In late August, 2005, ASC Execs challenged storage groups from all three Labs (Los Alamos, Livermore, and Sandia) to explore the technical and cost challenges of deploying a common data movement interface that can be used across compute platforms in the Tri-Lab complex within a two-year timeframe [initially, on TLCC resources]. The ASC Execs believe that users would benefit greatly from a common interface, and they also believe such an approach could provide cost benefits. The storage groups have met several times on the subject, and the teams have prepared a recommendation.

FY06 Action/Recommendation

In the coming year, each storage group will present the recommendation to their local management for feedback. Afterwards, the three storage groups will meet again to review the feedback from their local management and incorporate it into the presentation before delivering the presentation to ASC. Depending on the outcome of that meeting, user discussions, development and implementation plans may begin as early as FY06.

Issue H: Tape drive species delays and possible media shortfalls

Due to technical delays that Sun Microsystems has encountered in delivering a production-ready Titanium tape drive, and recent unanticipated increases in the amount of data being archived, the tape drive integration schedule and architectural design that the DSG had planned for FY06 might need to be modified in order to avoid a media capacity shortfall.

FY06 Action/Recommendation

Develop contingency plans for handling present data ingestion rates using only tape drive species that are currently available.

Visualization Issues

This section discusses visualization issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Appropriate throughput between Lustre and visualization resources

The LLNL clustered visualization needs fall into two major categories: The requirements for driving large, collaborative environments, and the requirements for data reduction and visualization of terascale datasets. Over the last year, we have refined our cluster architectures and no longer use Lustre for driving collaborative environments, save as a last result. Lustre, in its current implementation, does not provide sufficient single-threaded throughput, does not provide reserved bandwidth, and is not reliable enough for use in high-profile collaboration. Those needs are now met by small cache file systems on the clusters that commonly drive Power Wall environments.

For reduction and visualization of terascale datasets, the visualization clusters rely on Lustre. This provides an added advantage in that users have zero-copy access to the simulation data in situ. This enables quick “compute-visualize” workflows, and provides a future path to computational monitoring and computational steering.

Compared with the compute clusters like MCR or Thunder, visualization systems rely more heavily on file system IO for post-processing and visualization jobs. Furthermore, because these clusters are often used interactively, with the cluster reducing and rendering imagery to the display of a remote user, the IO needs are more closely aligned with the notion of capability rather than capacity. Either the IO infrastructure is capable of accessing data with sufficient speed to provide interactive performance, or it is not.

For the viz clusters, necessary Lustre throughput is mainly determined by three factors:

1) The size of the datasets being processed
2) The size of the Lustre store being serviced
3) Job throughput requirements

All the factors should be considered under my first rule of visualization: Data that has not been seen by the visualization system cannot be rendered or visualized. The first factor deals with both individual timesteps and the overall dump size. So if a job has 27 billion zones, each with three degrees of freedom represented as doubles, getting the first image requires the transfer of 648GB of data from Lustre to the viz platform. For “capability runs” the upper bound on this is probably the aggregate RAM of the compute cluster creating the image – e.g., BGL has 3TB of RAM so a full system run might produce timestep dumps approaching that size.

The second factor is the upper bound on the amount of data that is produced. More importantly, it effects how the third factor is considered when sizing viz resources. As compute jobs complete, the researchers turn to the viz resources to gain insight into the results, and to storage to preserve important runs for future study. Because viz and compute share the same storage target, visualization and post-processing of data requires keeping those datasets in the same Lustre store that is used for ongoing computation runs. This reduces the resources available for computing until the older data has been visualized and purged.

The same factors also play into milepost and deliverables scheduling. Because visualization and post-processing are the last steps in many experiments, they inherit any delays caused by earlier issues. As a result, the viz systems are often tasked with quick turnaround on experiments that were completed just prior to important deadlines and reviews. Again, the size of the Lustre store represents the upper bound.

With the current Lustre architecture, there are limited options to balance these requirements when architecting a cluster. Using a gateway none model to route Lustre traffic has proven not to scale. Either the cluster would have a large percentage of nodes dedicated to gateway service (increasing cost), or would be designed with a gateway-to-render node ratio like that used for compute clusters. That ratio results in a visualization system with insufficient connectivity.

The other option, which is currently deployed on gauss, is to give every node a direct connection to Lustre. This provides better bandwidth (and thus improved performance), but is not tunable and does compromise theoretical single-thread performance. Every node gets one network connection worth of bandwidth, no more or less. The throughput bottleneck for Lustre operations is this connection, not the interconnect fabric (which isn't even used for Lustre under this model).

Looking at FY06 and beyond, we hope to see a third option: “bridging” the Lustre traffic into the switch fabric on the cluster. This could be accomplished by using the multi-protocol features present in IB switches to trunk Lustre over Ethernet directly into the switch, or by moving to IB transport for Lustre itself. In either case, we would have the ability to tune the amount of Lustre bandwidth provided to a cluster, rather than for example getting one Gigabit Ethernet per node as our only option.

FY06 Action/Recommendation

Visualization plans the following deployments or re-deployments in FY06:

- Gauss will move to the SCF, and have its 256 Lustre Gigabit Ethernet connections upgraded to full bandwidth (meaning moving from the current 12 10GigE uplinks on from its Lustre switches to a full 24 10GigE uplinks.
- A new OCF cluster, with the working name “son of gauss” will be deployed on the OCF. This cluster will be between 50% and 100% of gauss' size (128-256 nodes) and requires a similar amount of Lustre bandwidth. A bridged Lustre configuration is under consideration.
- A new SCF cluster to do drive the B132 powerwall. Small size (16 nodes). Lustre configuration unknown.
- Continue to work OpenIB issues on Gauss and Vertex.
- Investigate bridged Ethernet/IB solutions.

Issue B: Network connection outside facility for graphics, data

In addition to requiring significant bandwidth between the Lustre file system and the visualization system, visualization applications must also deliver imagery to user desktops. This can be done in multiple ways: move the data (with FTP and such), move the geometry dataset (e.g., triangles), move the pixels, or by extending the RGB signals from the frame buffers to the desktop (e.g., Lightwave). All options but the last (RGB extenders) result in an additional load on visualization cluster login nodes and the network. There are also security/privacy issues to get network traffic from LC resources to local area networks in DNT’s A and B Divisions. As datasets increase in size, the necessary level of interactivity also increases, placing further demands on the network between the user desktop and the cluster.

FY06 Action/Recommendation

The RGB extenders (from Lightwave) are particularly expensive for visualization clusters (compared to monolithic visualization resources such as SGI), unreliable, and scales poorly. Moving images or data over the network faces the security/privacy issues, which also tend to impact the performance since the security issues are mitigated by tunneling X-Windows and GLX data through an encrypted SSH session. This remains an unresolved issue which should be addressed soon to open the way for network-based remote visualization solutions in FY07.

There have been demonstrated technical successes for network-based remote visualization. At considerable cost, a few scientists were provided isolated (for security) network connections on the Jumbo Frame network. The network throughput and interactivity allowed the visualization applications to work to the satisfaction of the scientists. This solution does not scale due to cost, and would probably also raise security concerns if too widely deployed. But it does provide a proof of principle.

Issue C: Appropriate versions of MPI

On previous generation Quadrics clusters, the MPI implementation was provided by vendors and based on MPICH. As we transition to IB interconnects, we have several MPI implementations to choose from, including MVAPICH and OpenMPI. Because some visualization technologies require MPI-2 features not implemented in MPICH and MVAPICH, we will likely provide OpenMPI on visualization clusters as an alternative to MVAPICH.

FY06 Action/Recommendation

The recommendation is to simply provide OpenMPI and MVAPICH on visualization clusters. As a result, all MPI features needed by users will be available on visualization clusters.

I/O Testbed Issues

Test environments have always played a key role in developing, tuning, and debugging software capabilities. Sophisticated test configurations are also required to evaluate emerging hardware and software technologies. As our production systems became more complex and modular in nature, the amount of hardware required to perform meaningful testing increased to the point that we found it advisable to build a robust test infrastructure designed to duplicate our production capabilities as closely as possible. The I/O Testbed also increases production system uptimes by minimizing stand-alone time required to deploy new capabilities. The I/O Testbed has been designed to be a center-wide resource which can be configured to perform a wide range of testing at short notice. Other sections of the I/O Blueprint elaborate upon projects which will be performed using I/O Testbed resources. The following basic areas of research have been planned for FY06.

- **Networking/Interconnects**
  - InfiniBand testing.
  - Evaluate PCI-express (HBA's, HCA's, and switching hardware).
  - Experiment with larger MTU's for Lustre transfers.
  - Evaluate current Ethernet channel bonding technologies.
- **Systems**
  - Continue to support systems that represent our actual production resources.
  - Evaluate blade technologies (high density, integrated power & console management).
  - Track IPMI capabilities and board management issues.
  - Evaluate new chipsets and FSB speeds as they appear.
  - Experiment with diskless compute nodes.
  - Setup a prototype "next generation" system configuration (at this point it looks like the system will be some type of Opteron or Nocona based blade technology using a PCI-express interface to an InfiniBand interconnect).
- **Software**
  - Produce a standard benchmark suite that we can use to evaluate new compute platforms.
  - Track the Open IB software effort.
  - Evaluate OS bypass software alternatives.
  - Develop a Chaos regression test suite.
- **Storage**
  - RAIT development – two software solutions are being explored.
  - Evaluate global file system alternatives.
  - Look for lower cost storage options.
  - Evaluate Linux based movers for HPSS.
  - Continue to facilitate Lustre development.
- **Visualization**
  - Evaluate supporting the specialized chipsets required for graphics.
  - Evaluate the behavior of various PCI-express graphics cards.

Driven by the issues with Peloton, and the Lustre and InfiniBand issues, the IO Testbed will focus in FY06 on the following:

- Acquiring and benchmarking 10GigE PCIe NICs, which are just becoming available Q3 of FY06.
- Gaining a better understanding of InfiniBand, and working with OpenIB and vendors to:
    - Ensure the IB product suite is complete enough for a large scale cluster and,
    - Ensure IB products will provide what we need in the System Area Network (e.g., replacing 10GigE)
- Benchmark the multiple offerings from network and interconnect vendors for low cost 10GigE interconnect solutions (ironically, directly competing in IB's market). Several of these products are under very tight non-disclosure agreements (NDA).

Special Projects - Green Data Oasis Issues

This section discusses Green Data Oasis (GDO) issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Multiplicity of data movement tools

The list of data movement tools and interfaces ranges from NFS to GridFTP. Issues include support and security.

FY06 Action/Recommendation

This issue has been overshadowed by the other issues related to GDO, and will probably be greatly influenced by cyber security considerations. No recommendation at this time.

Issue B: Security of the GDO data

There is acute interest by some people at LLNL for providing more bandwidth on the unrestricted network to the Green Data Oasis (GDO). Presently the unrestricted network is 100Mbps locally and 1Gbps to the ESnet WAN – vastly undersubscribed when compared to the end-to-end 1-10Gbps commonly found at universities and such that are interested in the GDO. There are network hardware procurements, security procurements, possibly development efforts in cyber-security, and facility upgrades needed to realize the 10Gbps available from ESnet since mid-2005 (at no additional cost to LLNL).

Above and beyond the cyber-security issues facing the 10Gbps ESnet WAN upgrade, several potential issues exist with security of the GDO data. Unless the data is reviewed and released for release to the world, some security mechanisms will need to be put in place for the data (e.g., authentication and authorization). This would be in addition to the usual network security (e.g., firewalls, IDS, etc).

FY06 Action/Recommendation

Jason King has identified several cyber-security products that would probably fulfill the requirements of LLNL's Cyber Security Office (CSO). CSO should be directly involved in any ESnet WAN upgrade effort to avoid miscommunication and to ensure their requirements are met. Jason does have equipment available to upgrade the unrestricted network from 100Mbps to 1Gbps, but that upgrade is being held up by lack of power in B256. Further power and security issues will arise to upgrade the ESnet connection and any portion of the unrestricted network to 10Gbps (i.e., 10GigE). Hence power and cooling in B256, as well as security issues for 10GigE, remain outstanding issues.

Brian Carnes is leading the GDO effort in ICCD. Recently the programs in the GDO effort have agreed that they will get all data released as "public domain." A security plan is needed for the GDO. Pam Hamilton's team will do the security plan, and be responsible for deploying and supporting the equipment in the GDO.

Special Projects - UCSD-LLNL Scientific Data Management Issues

This section discusses UCSD-LLNL Scientific Data Management issues to be addressed in FY06. Each issue is described and then recommended actions are presented.

Issue A: Network application throughput monitoring

Several data movement interfaces are expected for this project, such as FTP and SRB. No throughput monitoring is presently available, mostly due to lack of resources from UCSD. Without such resources, it is improbable that the desired throughput will be maintained, or even achieved.

After considerable network testing and with our prompting for an automated network performance monitoring capability the UCSD/SDSC network staff recommended that SDSC and LLNL set up an authenticated IPERF (network performance measurement tool) server. LLNL completed their installation by the fall of 2005. However SDSC continues to be pre-occupied with other higher priority tasks.

FY06 Action/Recommendation

Continue to work with Jim McGraw to ensure this issue is addressed.