Operational Report

Suggestions for Decreasing Operational Load within CCF

What is Covered

- **Networking**
- **NIMI**
- **Computer Security**
- **LQCD**
- **FermiGrid**
- **OSG**
- **Storage**

Networking

- 10 - Nameservers Currently Windows - going to Linux
- 2 - DHCP servers - Solaris
- 3 - Hosts for Pinger Support - Unix
- 5 - MRTG hosts - Linux
- 2 - Radius Servers
- 10 - Production Servers - Netflow Collection, Monitoring, Management - Ciena, Big Sister, etc.
- 32 - production

Networking

- 15 test/backup systems
- Maintaining these machines – .5 FTE

Networking

- Already a success story
  - Most issues are troubleshooting and upgrading
  - Installation is already outsourced to contractors
    - Equivalent of 1 FTE
    - Will increase to 1.5 FTE
    - Best for industry standard tasks

Networking

- Implement the Cisco intelligent wireless management system (WISM) that would dynamically tune our wireless network to adjust to individual AP failures & problems. Wireless LAN problems are the number one source of trouble tickets & problem reports. WISM should reduce the level of that effort.
  - Estimated cost ~$75k.
  - 0.5fte * 2 months

Networking

- Complete the WH Fiber-To The-Desktop (FTTD) project. Problems with pillar network devices & connections are still common. Physically implemented network reconfigurations to accommodate reassigning offices to different work groups is also common. WH FTTD should result in a major drop in network problems, and greatly simply office network reconfigurations.
  - Estimated cost ~$165k.
  - 0.75fte * 3 months, plus contractor (part of $165k...)

NIMI

- 4 servers
- 3 instances of PostgreSQL Dbs
- applications
- Chih-Hao Huang (35%), Maxim Grigoriev (5%-10%), Randy Reitz (10%)
- Check on health of Tissue
- Check on health of NIMI collectors (email notification)
- Check/restart services

NIMI

- better monitoring system, including monitoring the usage/performance of the system.
  - 2 FTE months
- reorganization of product structure: make installed product layout match that in the code repository.
  - 1 FTE month

NIMI

- training of the personnel who maintain NIMI
  - database, python, Soap, Zope, OS, and Network
- Only sees operational load increasing in the next several years

Computer Security

- Fermi Scanner Farm
  - 4 nodes, +1 future node
- Inventory
  - 1 node, +1 future
- Central Logger
  - 1 node
- Sniffers
  - 4 nodes, -3 future

Computer Security

- Consolidator, web server...
  - 2 nodes
- KDC's
  - 6 + 3 nodes
- KCA's
  - 2 nodes
- Total – 26 nodes - 3

Computer Security

- OS/Patch management
  - Frank Nagy (20%)
- Day-to-day effort
  - Frank Nagy (20%)
- Application management
  - Randy Reitz (90%), Tim Rupp (100%)

Computer Security

- We need remote access to our machines via the console interface. This isn't straightforward with Linux. Basically, servers should have a BIOS that knows about remote management. Also, we need the terminal server infrastructure. I would like operations to setup, support, maintain terminal servers for all of the CST machines.

Computer Security

- CST has purchased 'white' box hardware. It's cheap. It looks like a server, but managing it from afar is never easy. The recent power outages has demonstrated the superiority of SUN hardware. SUN knows how that build a server box that can be managed fairly easily (lights out) over a serial link. I think Dell is getting this expertise, but Dell still favors Windows. If we spent a little more on the hardware, we can get it back on maintenance.

LQCD

- 648 worker nodes
  - Grow to 1200 by October
  - Will need help for installation
- 1 head node with 15 TB disk
  - Will add another by October
- Dcache and PNFS
  - 3 pool nodes
    - Will add 7 by end of july
  - PNFS server, door node, pool manager node

LQCD

- Infiniband fabric
  - 144 port spine
  - 33 leaf switches
- 1 node for monitoring and controlling fabric
- Fabric will double by October
- 128 port Myrinet fabric
- PBS batch system

LQCD Tasks

- Health of Infiniband and Myrinet fabrics
- Health of all the nodes
- Status of the batch system

LQCD

- Amitoj Singh (50%), Kurt Ruthmansdorfer (100%), Don Holmgren (10%-25%)
- Would benefit from 1 more FTE now
- Will only grow even more over the next 5 years.
- Next major increase in FY08
  - Will need another FTE

LQCD

- Would like more of an expert system to handle all of the monitoring.
  - Current notification is by email
  - Very little automatic restarts etc.
  - Use nannies to handle some problems
  - Need a more generalized framework
  - Project identified to work on this (SCIDac2)
    - Vanderbilt - 1.083 FTE for 4 years
    - Fermilab - .25 FTE for 4 years (Jimn Kowalkowski, Amitoj Singh, Kurt Ruthmansdorfer)

LQCD

- Physicists that run need an automated work flow
  - Connects batch systems, mass storage, and status and monitoring to optimize usage of the cluster
  - Have sets of jobs (1000) for an analysis campaign that must finish in order to correlate all of the output of the sets.
  - Now is adhoc.
  - Also identified in SCIDac 2
    - IIT – 1.083 FTE
    - Fermilab - .5 FTE (Jim Kowalkowski, Jim Simone)

LQCD

- Power outages are an operational problem
  - ½ of LQCD equipment relies on old HVAC equipment that puts operations at risk of an extended outage caused by chiller failure
  - Would benefit from chiller replacement or conversion
  - Would benefit from a small UPS for critical server nodes (dCache...)

FermiGrid Effort

- Steve Timm (50%), Keith Chadwick (90%), Dan Yocum (75%)
- 50%-75% of effort total is operational effort
- A CP4 arriving after paperwork is done
- Need at least 1 more FTE
- Was in budget request for last year

FermiGrid Tasks

- Maintain existing systems
  - fngp-osg
- Deploy new systems – fcdfosg1/2
- d0cabosg2 coming
  - Hardware, OS, VDT configuration etc...
- VDT
  - Testing pre-release candidates of VDT
  - Upgrades - rolling upgrades across fleet

FermiGrid Tasks

- FermiGrid common services infrastructure
  - Fermigrid1/2/3/4 (production)
  - Fermigrid0, fgtest1/2/3/4/5 (test and development)
- Maintaining metrics collection infrastructure

FermiGrid Use

- Run II
  - Operational support will decrease when systems are shut down, many years out
  - Data still needs to be analyzed
- CMS
  - Operational support will be increasing
  - Will likely want 24x7 support of core FermiGrid services
- Neutrino Experiments
  - Mostly Minos and MiniBoone

FermiGrid Use

- General Fermilab
  - ILC, ... ????
- Only see expansion of operational duties (for all experiments) for the next 7 years.
- See no ways to reduce operational effort

OSG Tasks

- Privilege Project
  - GUMS, PRIMA, WS-PRIMA, gPLAZMA
  - Gabriele Garzoglio (35%), Vikram Andem (50%), Ted Hesselroth (50%)
- VO
  - VOMRS, VOMS
  - Tanya Levshina (50%), Parag Mhashilkar (10%)

OSG Tasks

- ReSS
  - Information Gatherer, CONDOR MMS
  - Gabriele Garzoglio (40%), Tanya Levshina (40%), Parag Mhashilkar (10%)
- Risk Assessment
  - Don Petravick (??%), Vikram Andem (50%)

OSG

- Improve privilege infrastructure in the areas of robustness, monitoring, software validation, usability/error reporting. This should reduce the occurrence of problems and empower the local administrator to perform better diagnostics.
  - Effort not identified for many items

Privilege WBS

- 1. Improve robustness of GUMS
- 1.1. Fix GUMS memory management problems (3 FTE weeks) (Starting)
- 1.2. Improve GUMS configuration management (3 FTE weeks) (Not started)
- 1.2.1. implement configuration validation
- 1.2.2. investigate/eliminate side-effects of hot-swapping configuration
- 1.2.3. implement configuration display interface
- 1.2.4. implement configuration GUI
- 1.2.5. improve debugging capabilities (date fields in database?)
- 1.2.6. add query interface for "last changed date"
- 1.3. Investigate redundant servers configuration (3 FTE days) (Not started)

Privilege WBS

- 2. Improve health status reporting for key servers (1 FTE week) (Almost Done)
- 2.1. Better Gatekeeper / Prima error reporting for authorization failures (effort TBD)
- 2.2. VOMS/GUMS health monitors
-
- 3. Improve software validation (8 FTE weeks) (Not started)
- 3.1. Improve validation of basic functionalities (framework available in VDT)
- 3.2. Implement validation of software dependencies
- 3.2.1. Validate PRIMA / GUMS stability across multiple architectures
- 3.2.2. Validate GUMS / VOMS interaction across multiple release versions
- 3.2.3. Validate GUMS / MonLISA interaction
- 3.2.4. Validate GUMS / Tomcat - MySQL stability across multiple release versions

Privilege WBS

- 4. Improve GUMS usability
- 4.1. Improve pool account management (1 FTE week) (Not started)
- 4.1.1. **Improve error logging**
- 4.1.1.1. Add error when authorization is denied because pool accounts are exhausted
- 4.1.2. Improve management interface
- 4.1.2.1. Add interfaces to query used and available pool accounts.
- 4.2. **Implement history log querying interface (2 FTE week) (Not started)**

OSG Support

- Invest more time in training CERN personnel that support VOMRS.
  - 2 FTE weeks at CERN
- For SAM-Grid operations, train people from the experiment in deployment practices and in typical support cases. This training program is under way.
  - Few hours/week for 6 months

Storage Tasks

- Storage Administration
  - 21 server computers
  - Greater than 100 mover computers
  - Greater than 100 tape drives
  - 1 ADIC tape library
  - 1 STK SL8500
  - 6 STK 9310s
  - 1 computer, 12 tape drives (needfile)
  - 11 + 2 nodes in general dCache
  - 67 + 12 nodes in CDF dCache

Storage Tasks

- Enstore, dCache, and SRM applications
- Storage Administration
  - 3.5 FTE of 4 administrators and 1 group leader

Storage Tasks

- Lower Storage applications
  - About .5 FTE of 2 FTEs
- Upper Storage applications and support (dCache, SRM)
  - Greater than 1.5 FTE of 4.5 total FTEs (includes 2 project managers)

Storage

- The operations load grows with the MSS growth, and the MSS is growing rapidly
  - Installs
    - Movers
    - Tape drives
    - Tape libraries
  - Tape investigations
  - Mover Investigations

Storage

- Run II
  - Will need more storage space in future
  - Will be analyzing data long past end of data taking
- CMS
  - Ramping up in next year
  - Will need more space in coming years

Storage

- Screwdriverless Enstore (no self-maintenance, minimal downtime upgrades)
- Electronic workflow
- Improvements in mover exception handling
  - these are by far the majority of day-to-day work
- Improvements in monitoring (organization of hodgepodge of alarms, emails)

Storage

- Regularly executed downtimes

Screwdriverless Enstore

- Transition Hardware and OS requisitioning, install, burn-in, support to CSS
  - Have started to meet on doing this
  - Will come up with a plan
  - Need to improve homogeneity of the system
    - Replace servers
    - Need spares
  - Support perhaps by D1

Screwdriverless Enstore

- Software
  - Dynamically move services to new machines
  - Decrease the number of servers
  - Improve documentation

Storage Software

- Dcache and SRM Operations: A relatively high load due to:
  - rapid development to support increasing service scale and new features.
  - less attention to administrative and monitoring interfaces

Storage Software

- Theme of effort to address this
  - simplify all commonplace administrative tasks
  - extend monitoring to simplify admin tasks, capture expert knowledge
  - define and track metrics to describe service quality from customer view
  - automate development processes like testing to improve service quality

Storage Software

- Admin toolkit to perform major tasks with a single script invocation.
- Create and support a "dcache-server" workgroup in Fermi Sci Linux to simplify installation of new pool nodes, OS upgrades, etc.
- Add metrics and organizational capabilities that describe customer's view of storage rather than the developer's view of the storage system. Maximally leverage the existing billing database.

Storage Software

- Monitoring package simplifying task of creating new plots, tapping into experiment DH experience and work to fill out suite.
- Deploy exist build/test framework, populate with features tests
- Evaluation, co-development, and deployment of Chimera to replace PNFS. Service specific monitoring of Chimera is a prerequisite.

Storage Software

- Extending the monitoring to provide compact summaries of all client problems associated with a site to simplify troubleshooting and move towards more pro-active service and support.
- All of above work will take .5-1 FTE

Summary

- Increase in operational load seen in next few years
  - CMS
  - Run II – no decrease
- Workflow tools desired by many
  - AlphaFlow?
  - Who supports this?
  - Division direction?