

# Lightning Talk: SRM at Nebraska

Brian Bockelman  
OSG Storage '09

# Talk Focus

- Not our Hadoop install (will be covered later), but our SRM configuration:
  - Operations at Nebraska.
  - BestMan stress tests.
  - BestMan issues.

# SEs and SRM Endpoints

Hadoop (300TB)

dcache07.unl.edu

srm.unl.edu

red-srm1.unl.edu

dCache (100TB)

srm.unl.edu

Panasas (150TB)

ff-srm.unl.edu

4 BestMan servers running (3 in Hadoop, 1 on Panasas)  
1 dCache/SRM server

# BestMan SRM Servers

- Hadoop:
  - 1 for D0 (they caused service issues)
  - 1 for CMS production stageout & RSV/SAM
  - 1 for CMS PhEDEx transfers.
- Panasas: On separate cluster from T2 in Omaha.

More servers,  
same # of admins =  
less time per  
server.

The increase in # of SRM endpoints necessitates increasing automation.

- Must be no harder than Apache HTTP.

# Hardware

- We're getting rid of our "storage cluster".
- "storage" and "compute" hardware are now the same thing - one cluster, less mgmt
- Dedicating a node to BestMan is pointless; BestMan instances are virtualized.
- 5 GridFTP servers for HDFS; 1 with 10Gbps card, 4 with 1Gbps cards.

# Operations

- Hadoop:
  - We use syslog to centralize all ERROR-level logs (this level is manageable).
  - HDFS resilience means HDD failures are low priority.
  - For the most part, Hadoop core requires little maintenance.
  - Datanodes do not get overloaded.

Hadoop is actually fairly hard to kill - it mostly survived bad network switch issues earlier this month.

# Operations

- BestMan:
  - Still hand-installed and configured :(
  - Nagios checks are basic:
    - Make sure the process is running.
    - Make sure the node is alive
  - RSV tests are sufficient for application-level liveness tests.

# Operations

- GridFTP:
  - Never had a gridftp server where the host was alive but gridftp non-functional.
  - Watch for zombie transfers.
  - Only monitor host-level stats
  - Can kickstart gridftp server - come up with everything except host certificates.



# BestMan Scalability

- We've worked closely with A. Sim for scalability.
- We have successfully been able to surpass 50Hz for SRM operations for a single server.
- We can do this separately for each server.
- We now have a simple toolkit for generating SRM load.

# BestMan Usability

- BestMan can not only push through a high rate, but also respond quickly.
- In our testing, responds 10x faster than dCache/SRM using lcg-ls (.3s versus 3s).
  - Both take ~ 3s with srmls.
- (lcgutils + BestMan) makes SRM easier to interact with - now works at approximately “human scale”.

# BestMan Successes

- Over the last year, we've gotten:
  - Patches for any compatibility issues.
  - New plug-in infrastructure for GridFTP selection.
  - Drastic performance improvements.
  - Java API for filesystem operations (currently not used :( )

# Big Issues

- Configuration in general!!!!
  - Dynamic and centralized configuration is not possible.
  - Need to remove/hide flexibility.
  - Prefer no more than 3 options.
  - Editing sudoers is a mess (setuid script?)
- Configuration of gridftp servers
  - Default to dynamic configuration in file?
- Need re-configure without restart!

# Other Issues

- We'd really like to build from source.
- This would allow us to fast-track even more updates, as well as make our own contributions.

# Future Work

- GridFTP servers are more “appliance-like” with their new RPMs (good).
- Would like to do the same thing with SRM.
- Believe the future is in virtualized instances that are automatically load-balanced.
- One endpoint, but a variable amount of resources!
- Investigating use of LVS.

# Future Work

- Themes: Stability, low-maintenance.
- BestMan as a stable service.
  - Transparent interventions.
- Continue to work on decreasing maintenance.
- We want fully-automated installs, **native packaging**.