

# BESTMAN STORAGE CONFIGURATION AT OU SWT2

HORST SEVERINI  
OSG STORAGE FORUM  
SEPTEMBER 22, 2010

## Outline

- Introduction
- US ATLAS Tier 2 Center Setup
- Storage Setup
- BeStMan/checksum/Gridftp Server Selection Setup
- Summary and Outlook

## Introduction

---

- OUHEP involved in computing efforts for both ATLAS and DØ
- Also very active in various Open Science Grid (OSG) activities in recent years
- Working closely with Langston University (LU) and OSU as part of the Oklahoma Center for High Energy Physics (OCHEP)
- Collaborating with LU and UT Arlington as part of the US ATLAS SW Tier 2 Center
- Additionally, using OUHEP desktop cluster and large OU Supercomputing for Education and Research (OSCER) cluster as Tier 3 resources



## US ATLAS Tier 2 Hardware

---

- 54 Node (432 Core) 2.27/2.33 GHz Xeon-64 (E5420/E5345)
- 8 Support Nodes (5 head, 3 storage) 2.27/2.33 GHz Xeon-64
- 2-3 GB RAM per Core
- 200 TB of usable DDN/Lustre storage (264 TB raw)
- Platform PCM 1.2a (RHEL5.3 64 bit)
- OSG 1.2
- tier2-01: head node / gatekeeper
- tier2-02: storage transfer node (SRM/BeStMan/gsiftp)
- tier2-03/4: gsiftp servers
- tier2-05: LFC, Frontier-Squid



## OU Network

- OU connected at 10 Gbps to NLR, I2, and ESnet, via OneNet and GPN
- 10 Gbps connection straight from machine room
- 3-4 Gbps from Tier 2 cluster to BNL
- Tier 3 clusters at 1-2 Gbps



## OUHEP Tier 3 Resources

---

- Desktop Cluster: 46 Node 82 CPU)  $\approx$  2 GHz P4/Xeon, 20 TB storage (additional 80 TB on order)
  - OSG Production site, OSG Integration site, OSG SAM station (DØ)
  - Used for DØ SAMGrid production, ATLAS MC, OSG and SAMGrid integration testing, and local theory calculations
  - BeStMan2 and Gridftp server selection plugin testing



## Lustre Setup

---

- DDN S2A 9900 and DDN S2A 3000
- 240 TB raw (120 2TB drives) and 24 TB raw (48 500GB drives)
- 3 storage servers: 1 mds and 2 oss; R710 dual quad, 36 GB of RAM
- Lustre 1.8.3.ddn: 175 TB data and 12 TB app partitions
- Max performance: 850 MB/s read/write from multiple clients
- gridftp performance: 125 MB/s per stream, 200 MB/s per gridftp server
- So Far: 400 MB/s with two gridftp servers (not maxed out yet)

## BeStMan-Gateway Setup

---

- Version 2.2.1.3.10 (will upgrade to .13 during next Lustre upgrade downtime – or BeStMan2)
- Minor Modifications:
  - MAX\_JAVA\_HEAP=2048 (bestman/sbin/bestman.server)
  - `<parameter name="containerThreads" value="512"/>`  
(bestman/conf/server-config.wsdd)
  - Customized Checksum Calculation (see next page)
- Calculate space usage of tokens every 4h with 'du' script
- <http://www.usatlas.bnl.gov/twiki/bin/view/Admins/BestMan>

## Adler32 Checksum Calculation

---

- Store checksum in extended file system attribute, for speed
- Works for Lustre, GPFS, ext4, ...
- Use two different python scripts:

adler32.py: check for attribute, return it if exists; if not, call

calc\_adler32.py: calculate checksum, store in attribute

- Edit `vdt-app-data/bestman/conf/bestman.rc`:

```
defaultChecksumType=adler32
```

```
showChecksumWhenListingFile=true
```

```
hexChecksumCommand=sudo /home/osg/bestman_apps/adler32/adler32.py
```





## Adler32 Checksum Calculation (cont.)

---

- Edit /etc/sudoers:

```
daemon ALL=(root) NOPASSWD: /home/osg/bestman_apps/adler32/adler32.py
```

- Edit /etc/fstab:

```
10.173.10.15:/data /storage/data lustre _netdev,localflock,user_xattr 0 0
```

- Re-mount file system and restart BeStMan
- Run hourly cron job to pre-calculate checksums of all new files
- Next PanDA Pilot version has hook to trigger calculation immediately, to further reduce load on bestman server
- <https://twiki.grid.iu.edu/bin/view/Storage/BestmanAdler32Checksum>

## BeStMan2

---

- In VDT\_2099 now:  
[http://vdt.cs.wisc.edu/vdt\\_2099\\_cache:Bestman2](http://vdt.cs.wisc.edu/vdt_2099_cache:Bestman2)
- Installs and configures just like BeStMan, except for http port, which is no longer necessary
- Works fine out of the box, just like BeStMan

## Gridftp Server Selection Plugin

---

- Short Java program, MountPoint.java
- Reads servers.txt: PATH=SERVER\_URL
- Compile and archive MountPoint.java
- Add to bestman.rc:
  - protocolSelectionPolicy=class=MountPoint&jarFile=mount.jar&name=gsiftp
  - pluginLib=/opt/bestman/bestman/lib/plugin
- Documentation in the Works

## Summary and Outlook

---

- SWT2 Storage Infrastructure in very good shape
- Good Performance and Stability
- BeStMan-Gateway working well, both on Lustre and Xrootd
- To do
  - Convert OU BeStMan to BeStMan2 as soon as stable version in VDT
  - Grow OU Storage by at least Factor of Two
  - Utilize recently awarded NSF PetaStore MRI grant

