



Everyday* Lustre

A short survey on Lustre tools

Sven Trautmann
Engineer, Lustre Giraffe Team, Sun Microsystems

Outline

- Motivation
- Setup Tools
- Management Tools
- Monitoring
- Conclusion

Who am I, What am I doing here?

- Engineer in Lustre Giraffe team since October last year
- current product: SunHPC Stack Linux Edition 2.0.1
- We try to provide a complete Linux HPC solution including
 - provisioning tools with BoIB (cobbler, oneSIS)
 - monitoring tools (ganglia, nagios)
 - managing tools (powerman, conman, freeipmi, cfengine, ...)
 - scheduling (slurm, SGE)
 - development tools (Clustertools, Sun Studio, ...)
 - and of course: Lustre server and client packages + OFED stack
- currently at Aachen University, Sandia NL, ANU/BoM, CLUMEQ ...

More giraffe information

- [http://www.sun.com/software/products/
hpcsoftware/](http://www.sun.com/software/products/hpcsoftware/)
- <http://blogs.sun.com/giraffe/>
- linux_hpc_swstack@lists.lustre.org
- irc.freenode.net, Port 6667, #sunhpc
- <http://www.randomgiraffefacts.com>
(not related)

Motivation

- Lustre is a complex system, learning curve is pretty steep
- starts with setting up Lustre and continues with managing and monitoring
- most tasks can be handled by scripts, most sides have their own solution
- there are a few tools out there to make working with Lustre easier
- this talk will discuss some of these tools

Lustre setup tools

The Lustre setup experience

- install Lustre packages (client & server)
- create software RAIDs (optional)
- setup Lustre network(s)
- format all MDT and OST file systems
- configure HA solution of your choice (optional)
- mount all Lustre targets
- try to mount Lustre file system from clients
- fix some problems and try again

Lustre setup tools

- three solutions (sorry if I missed any)
 - `lustre_admin` (Bull) last release in 2007
 - `shine` (CEA)
 - <http://sourceforge.net/projects/lustre-shine/>
 - `lustre_config`
 - comes with Lustre
 - a little outdated, still worked for 1.8.0.1

shine overview

- CLI tools designed to setup and manage the Lustre file system
- implemented in Python
- clustershell Python module is used to communicate with other nodes
- in heavy development, open source
- main developers work for CEA

shine workflow

- create configuration file for every Lustre file system
- `shine install -m example.lmf`
- `shine format -f <fsname>`
- `shine start -f <fsname>`
- `shine mount -f <fsname>`

shine model file example

```
[root@sven00 shine]# cat /etc/shine/models/example.lmf
fs_name: sven
nid_map: nodes=sven0[0-1] nids=sven0[0-1]@tcp0
mount_path: /sven
mgt: node=sven01 dev=/dev/sdb1
mdt: node=sven01 dev=/dev/sdb2
ost: node=sven01 dev=/dev/sdc
client: node=sven00
failover: no
description: my Lustre Filesystem
stripe_size: 1048576
stripe_count: 1
mdt_mkfs_options: -b 4096 -i 4096
mdt_mount_options: acl,user_xattr
ost_mkfs_options: -b 4096
ost_mount_options: extents,mballoc
mount_options: acl,user_xattr
quota: no
mgt_mount_path: /mnt/$fs_name/mgt
mdt_mount_path: /mnt/$fs_name/mdt/$index
ost_mount_path: /mnt/$fs_name/ost/$index
```

shine installation

```
root@sven00 shine]# shine install \
                  -m /etc/shine/models/example.lmf
Using Lustre model file /etc/shine/models/example.lmf
Configuration files for file system sven have been installed
successfully.
```

Lustre targets summary:

- 1 MGT on sven01
- 1 MDT on sven01
- 1 OST on sven01

Use `shine format -f sven' to initialize the file system.

```
[root@sven00 shine]#
```

shine format

```
[root@sven00 shine]# shine format -f sven
Format sven on sven01: are you sure? (y)es/ (N)o: y
Starting format of 3 targets on sven01
[18:54] In progress for 3 target(s) on sven01 ...
[18:54] In progress for 2 target(s) on sven01 ...
Format successful.

FILESYSTEM COMPONENTS STATUS (sven)
+-----+-----+-----+
| type | # | nodes | status |
+-----+-----+-----+
| MGT | 1 | sven01 | offline (1) |
| MDT | 1 | sven01 | offline (1) |
| OST | 1 | sven01 | offline (1) |
+-----+-----+-----+
[root@sven00 shine]#
```

shine start & mount

```
[root@svn00 ~]# shine start -f svn
Starting 3 targets on svn01
[21:27] In progress for 1 target(s) on svn01 ...
Start successful.

FILESYSTEM COMPONENTS STATUS (svn)
+---+---+---+---+
| type | # | nodes | status |
+---+---+---+---+
| MGT | 1 | svn01 | online (1) |
| OST | 1 | svn01 | online (1) |
| MDT | 1 | svn01 | online (1) |
+---+---+---+---+
[root@svn00 ~]# shine mount -f svn
Starting svn clients on svn00...
Mount successful on svn00
[root@svn00 ~]#
```

shine comments

- shine and clustershell need to be installed on all nodes, MDS, OSS and clients
- no fail-over (as of shine 0.903)
- no entry in modprobe.conf for other networks than tcp
- only one file system per MGS (?)
- almost no documentation (?)
- no software RAID setup
- shine looks very promising though

lustre_config

- create a csv file describing the Lustre file system
- enable password-less ssh login to all Lustre nodes
- call: `lustre_config <csv_file>`

Lustre csv file

```
f0011,options lnet networks=o2ib0,/root/mds.fs,/mnt/mds,mgs|  
mdt,,,,-device-size=16777216 --param="mdt.group_upcall=/usr/  
sbin/l_getgroups" --stripe-count-hint=4,-i 4095,,  
  
f0012,options lnet networks=o2ib0,/root/oss.fs,/mnt/ost,ost,,  
5.6.9.11@o2ib,0,-device-size=33554432,-i 16357,,  
  
f0013,options lnet networks=o2ib0,/root/oss.fs,/mnt/ost,ost,,  
5.6.9.11@o2ib,1,-device-size=33554432,-i 16357,,  
  
f0014,options lnet networks=o2ib0,/root/oss.fs,/mnt/ost,ost,,  
5.6.9.11@o2ib,2,-device-size=33554432,-i 16357,,  
  
f0015,options lnet networks=o2ib0,/root/oss.fs,/mnt/ost,ost,,  
5.6.9.11@o2ib,3,-device-size=33554432,-i 16357,,
```

lustre_config comments

- comes with Lustre package
- bash script, uses ssh
- currently no new development
- lots of features (try `lustre_config -h`)
 - fail-over, lustre networks ...

Management Tools

What can be managed?

- Lustre parameters can be tuned/changed
 - some parameters need to be changed on all nodes
 - tools like pdsh can help
- other management tasks: deactivating OSTs
- mounting/un-mounting clients
- shine can mount clients and tune Lustre (in theory, not tested yet)

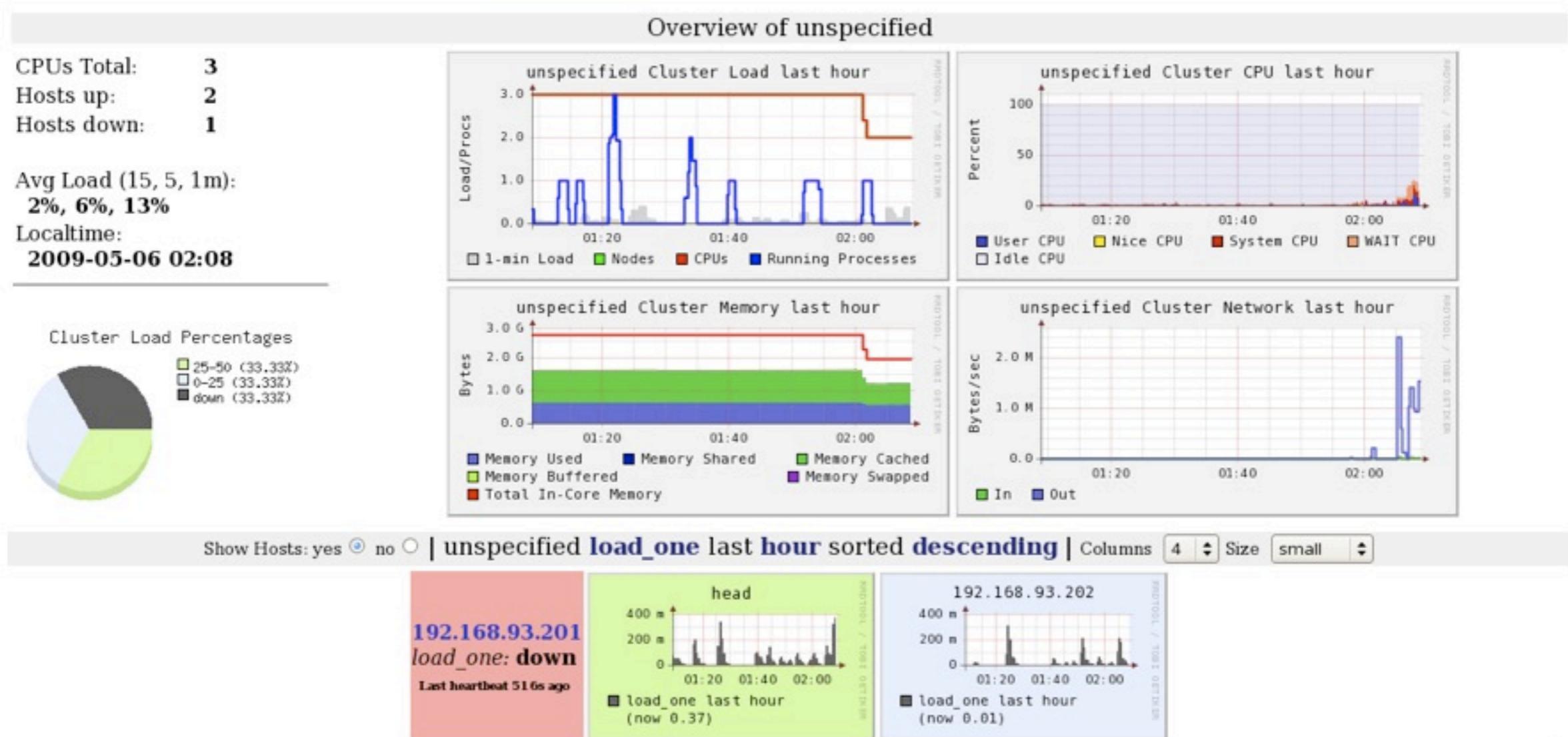
Lustre Monitoring Tools

Monitoring tools

- general monitoring tools
 - ganglia, nagios, cacti ...
- filesystem tools
 - collectl, ...
- Lustre specific tools
 - LMT2

Ganglia, Nagios ...

SunHPC Linux Grid > unspecified > --Choose a Node ▾



Nagios®

General

- Home
- Documentation

Current Status

- Tactical Overview
- Map
- Hosts
- Services

Host Groups

- Summary
 - Grid
- Service Groups**
- Summary
 - Grid
- Problems**
- Services (Unhandled)
 - Hosts (Unhandled)
 - Network Outages

Quick Search:

Reports

- Availability
- Trends
- Alerts
 - History
 - Summary
 - Histogram
- Notifications
- Event Log

System

- Comments
- Downtime
- Process Info
- Performance Info

Current Network Status
Last Updated: Thu May 7 10:20:46 EDT 2009
Updated every 90 seconds
Nagios® 3.1.0 - www.nagios.org
Logged in as nagiosadmin

[View Service Status Detail For All Host Groups](#)
[View Host Status Detail For All Host Groups](#)
[View Status Summary For All Host Groups](#)
[View Status Grid For All Host Groups](#)

Host Status Totals

Up	Down	Unreachable	Pending
1	0	0	20

[All Problems](#) [All Types](#)
0 21

Service Status Totals

Ok	Warning	Unknown	Critical	Pending
40	0	0	8	0

[All Problems](#) [All Types](#)
8 48

Service Overview For All Host Groups**Linux Servers (linux-servers)**

Host	Status	Services	Actions
localhost	UP	8 OK	

SunHPC Servers (sunhpc-servers)

Host	Status	Services	Actions
d10-0	PENDING	2 OK	
d10-1	PENDING	2 OK	
d10-2	PENDING	2 OK	
d10-3	PENDING	2 OK	
d10-4	PENDING	2 OK	
d10-5	PENDING	2 CRITICAL	
d10-6	PENDING	2 OK	
d10-7	PENDING	2 CRITICAL	
d10-8	PENDING	2 CRITICAL	
d10-9	PENDING	2 CRITICAL	

SunHPC Service Processors (sunhpc-service-processors)

Host	Status	Services	Actions
d10-0-52	PENDING	2 OK	
d10-1-50	PENDING	2 OK	
d10-2-52	PENDING	2 OK	
d10-3-50	PENDING	2 OK	
d10-4-50	PENDING	2 OK	
d10-5-52	PENDING	2 OK	
d10-6-52	PENDING	2 OK	
d10-7-52	PENDING	2 OK	
d10-8-52	PENDING	2 OK	
d10-9-52	PENDING	2 OK	

Nagios - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://localhost/nagios/ Google

Nagios®

General

- Home
- Documentation

Current Status

- Tactical Overview
- Map
- Hosts
- Services
- Host Groups
 - Summary
 - Grid
- Service Groups
 - Summary
 - Grid
- Problems
 - Services (Unhandled)
 - Hosts (Unhandled)
 - Network Outages

Quick Search:

Reports

- Availability
- Trends
- Alerts
 - History
 - Summary
 - Histogram
- Notifications
- Event Log

System

Current Network Status

Last Updated: Thu May 7 15:01:25 CEST 2009
Updated every 90 seconds
Nagios® 3.1.0 - www.nagios.org
Logged in as nagiosadmin

View History For all hosts
View Notifications For All Hosts
View Host Status Detail For All Hosts

Host Status Totals

Up	Down	Unreachable	Pending
1	0	0	0

All Problems All Types

0	1
---	---

Service Status Totals

Ok	Warning	Unknown	Critical	Pending
6	1	0	1	0

All Problems All Types

2	8
---	---

Service Status Details For All Hosts

Host	Service	Status	Last Check	Duration	Attempt	Status Information
localhost	Current Load	OK	05-07-2009 14:58:42	34d 20h 43m 13s	1/4	OK - load average: 0.08, 0.51, 0.35
	Current Users	OK	05-07-2009 14:59:20	34d 20h 41m 53s	1/4	USERS OK - 2 users currently logged in
	HTTP	CRITICAL	05-07-2009 14:57:57	0d 0h 6m 28s	4/4	Connection refused
	PING	OK	05-07-2009 15:00:35	34d 20h 46m 20s	1/4	PING OK - Packet loss = 0%, RTA = 0.07 ms
	Root Partition	WARNING	05-07-2009 14:58:34	1d 17h 49m 2s	4/4	DISK WARNING - free space: / 2948 MB (1.7% inode=90%):
	SSH	OK	05-07-2009 14:59:12	33d 21h 14m 43s	1/4	SSH OK - OpenSSH_4.3 (protocol 2.0)
	Swap Usage	OK	05-07-2009 14:59:49	34d 20h 44m 28s	1/4	SWAP OK - 100% free (2015 MB out of 2015 MB)
	Total Processes	OK	05-07-2009 15:01:12	34d 20h 43m 4s	1/4	PROCS OK: 38 processes with STATE = RSZDT

8 Matching Service Entries Displayed

Ganglia, Nagios ...

- these tools were not developed for Lustre
- some parameters can be monitored using these tools
- e.g. network load, software RAID state
- most Lustre specific things can be monitored, but not out of the box

collectl

- developed by HP
- shows file system statistics
- latest version supports Lustre
- was in SunHPC Linux stack 1.x, not in 2.0
- will probably be in 2.1 again

collectl example output

```
#<-----CPU-----><-----Disks-----><-----Network----->
#cpu sys inter  ctxsw KBRead  Reads KBWrit Writes KBIn PktIn KBOut PktOut
  1   0 1000    125      0       0     0     0     0     0     0     0     0
  2   2 1010    175     48      2     0     0     0     0     0     0     0
  1   0 1037    129      0       0    717    95     0     0     0     0     0
  1   1 1014    150      0       0     0     0     0     0     0     0     0
  1   1 1002    140      0       0     44     2     0     0     0     0     0
  1   0 1013    155      0       0     0     0     0     0     0     0     0
  0   0 1000    134      0       0     0     0     0     0     0     0     0
  1   0 1012    164      0       0     0     0     0     0     0     0     0
  1   1 998     133      0       0     0     0     0     0     0     0     0
  1   0 1012    141      0       0     0     0     0     0     0     0     0
  1   1 1003    151      0       0     12     2     0     0     0     0     0
  1   0 1011    155      0       0     0     0     0     0     0     0     0
  0   0 996     141      0       0     0     0     0     0     0     0     0
```

LMT 2 overview

- <http://code.google.com/p/lmt/>
- development started at LLNL
- monitors Lustre File System servers
- collects data using the Cerebro monitoring system and stores it in a MySQL database
- GUI and CLI clients display historical and real time data

LMT dependencies

- LMT uses Cerebro to collect data
- Cerebro needs genders and LLNL specific genders extension gendersInl
- genders and gendersInl need pdsh and some Perl modules
- LMT Client's CLI commands and GUI is implemented in Java
- LMT uses MySQL database to store data

Cerebro

- Cerebro is a monitoring tool, sharing roots with Ganglia
- <https://computing.llnl.gov/linux/cerebro.html>
- Cerebro comes with pre-defined metrics e.g. for loadavg and memory usage
- configuration similar to Ganglia's gmond
- uses multi-cast or uni-cast, just like Ganglia

LMT server agent

- provides new Cerebro metrics
- lmt-server-agent package provides Lustre metrics for Cerebro

```
root@sven01 ~]# cerebro-stat -l | grep lmt
lmt_mds
lmt_oss
lmt_ost
[root@sven01 ~]# cerebro-stat -m lmt_oss
sven01: 1.0;sven01;0.400000;18.965605
[root@sven01 ~]# cerebro-stat -m lmt_ost
sven01: 1.0;sven01;sven-OST0000;524232;524288;7906228;8256952;0;0
[root@sven01 ~]#
```

LMT Server Setup I

- LMT server provides binding between cerebro and mysql database
- LMT server can run on any node configured as Cerebro server
- needs to run a MySQL service
- one table for every Lustre file system
- MySQL database may be used by other applications as well

LMT Server Setup II

```
[root@sven00 ]# yum install mysql-server
...
[root@sven00 ]# /etc/init.d/mysqld start
...
[root@sven00 ]# /usr/bin/mysqladmin -u root password 'changeme'
[root@sven00 ]# yum install lmt-server
...
[root@sven00 ]# cat /usr/share/lmt/cron/lmtrc
filesys.1.name=sven
filesys.1.mountname=/sven
filesys.1.dbhost=localhost
filesys.1.dbport=3306
filesys.1.dbuser=lwatchadmin
filesys.1.dbauth=
filesys.1 dbname=filesystem_sven
[root@sven00 ]#
```

create LMT tables

```
filesystem
{
    name          sven
    mountpoint   /sven
}
mds
{
    name          sven01
    uuid          sven-MDT0000
    nid           192.168.93.200@tcp0
    device        /dev/sdb2
}

ost
{
    name  sven01
    uuid  sven-OST0000
    nid   {NODENAME}@tcp0
    device /dev/sdc
    skip  1
}
```

LMT database

```
# create_lmtconfig -a -d -t -f lmt-config.txt > lmtdb.in
# mysql -p mysql <lmtdb.in
```

```
[root@sven00 cerebro]# mysql -p
Enter password:
Welcome to the MySQL monitor. Commands end with ; or \g.
Your MySQL connection id is 2
Server version: 5.0.45 Source distribution

Type 'help;' or '\h' for help. Type '\c' to clear the buffer.

mysql> use filesystem_sven
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A
```

```
Database changed
mysql> select * from OST_DATA;
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| OST_ID | TS_ID | READ_BYTES | WRITE_BYTES | PCT_CPU | KBYTES_FREE | KBYTES_USED | INODES_FREE | INODES_USED |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
|     1 |     1 |       0 |       0 |    NULL |      7906228 |      350724 |      524232 |        56 |
|     1 |     2 |       0 |       0 |    NULL |      7906228 |      350724 |      524232 |        56 |
...
|     1 |   276 |       0 | 1935671296 |    NULL |      7064496 |     1192456 |      524200 |        88 |
|     1 |   277 |       0 | 1935671296 |    NULL |      7064496 |     1192456 |      524200 |        88 |
|     1 |   278 |       0 | 1935671296 |    NULL |      7064496 |     1192456 |      524200 |        88 |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+
278 rows in set (0.03 sec)
```

```
mysql> quit
Bye
[root@sven00 cerebro]#
```

LMT Clients

- All clients are implemented in Java
- just MySQL clients, which could be easily implemented in many languages
- two CLI clients ltop and lstat
- one GUI client lwatch
- extra config file in /usr/share/lmt/etc.

ltop and lstat

```
[root@sven00 ~]# ltop
2009-07-26 00:05:40 --- sven: OST Report

OST Name          Read (MB/s)    Write (MB/s)
sven-OST0000      0.00          0.00

[root@sven00 ~]# lstat

----- 2009-07-26 00:06:25 -----

[OST Name]      [Read MB/s]    [Write MB/s]
sven-OST0000    0.00          0.00

----- 2009-07-26 00:06:30 -----

[OST Name]      [Read MB/s]    [Write MB/s]
sven-OST0000    0.00          0.00

----- 2009-07-26 00:06:35 -----

[OST Name]      [Read MB/s]    [Write MB/s]
sven-OST0000    0.00          0.00
```

lwatch

LWatch-lustre

File Configure lustre

lustre-MDT0000 2009-07-22 19:09:55.0

%CPU	%KB		%Inodes		
0.40	3.10	0.00			
Operation	Samples	Sample /Sec	Avg Value	Std Dev	Units
add_conn	0	0.00	0.00	0.00	reqs
adjust_ioms	0	0.00	0.00	0.00	reqs
attach	0	0.00	0.00	0.00	reqs
bw	0	0.00	0.00	0.00	reqs
bw_async	0	0.00	0.00	0.00	reqs
cancel	0	0.00	0.00	0.00	reqs
cancel_unused	0	0.00	0.00	0.00	reqs
change_cbdata	0	0.00	0.00	0.00	reqs
checkmd	0	0.00	0.00	0.00	reqs
cleanup	0	0.00	0.00	0.00	reqs
close	0	0.00	0.00	0.00	reqs
commitnw	0	0.00	0.00	0.00	reqs
connect	0	0.00	0.00	0.00	reqs
copy	0	0.00	0.00	0.00	reqs
create	0	0.00	0.00	0.00	reqs
del_conn	0	0.00	0.00	0.00	reqs
destroy	0	0.00	0.00	0.00	reqs
destroy_export	0	0.00	0.00	0.00	reqs
detach	0	0.00	0.00	0.00	reqs
disconnect	0	0.00	0.00	0.00	reqs
enqueue	0	0.00	0.00	0.00	reqs
extent_calc	0	0.00	0.00	0.00	reqs
get_info	0	0.00	0.00	0.00	reqs
getattr	0	0.00	0.00	0.00	reqs
getattr_async	0	0.00	0.00	0.00	reqs
getxattr	0	0.00	0.00	0.00	reqs
health_check	0	0.00	0.00	0.00	reqs
import_event	0	0.00	0.00	0.00	reqs
init_export	0	0.00	0.00	0.00	reqs
iocontrol	0	0.00	0.00	0.00	reqs
iterate	0	0.00	0.00	0.00	reqs
join_jru	0	0.00	0.00	0.00	reqs
link	0	0.00	0.00	0.00	reqs
ilog_finish	0	0.00	0.00	0.00	reqs
ilog_init	0	0.00	0.00	0.00	reqs
match	0	0.00	0.00	0.00	reqs
merge_hb	0	0.00	0.00	0.00	reqs
migrate	0	0.00	0.00	0.00	reqs
mkdir	0	0.00	0.00	0.00	reqs
mkcoll	0	0.00	0.00	0.00	reqs

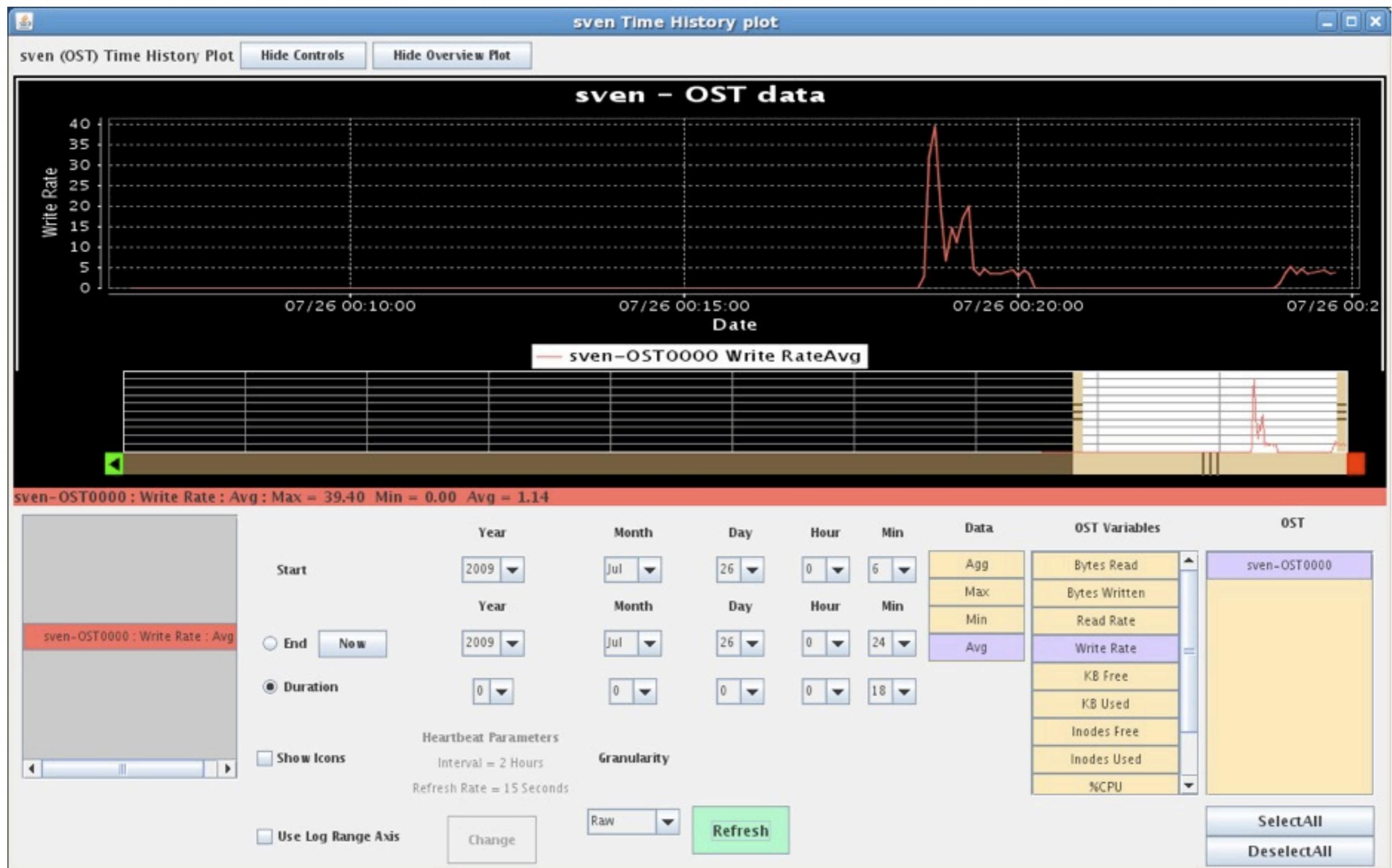
OST 2009-07-22 19:09:55.0

Dst Name	Read Rate	Write Rate	%CPU	%KB	%Inodes
lustre-OST0000	0.00	0.00	****	2.00	0.01
lustre-OST0001	0.00	0.00	****	0.66	0.00
lustre-OST0002	0.00	0.00	****	2.00	0.01
lustre-OST0003	0.00	0.00	****	0.66	0.00
lustre-OST0004	0.00	0.00	****	2.00	0.01
lustre-OST0005	0.00	0.00	****	0.66	0.00
lustre-OST0006	0.00	0.00	****	2.00	0.01
lustre-OST0007	0.00	0.00	****	0.66	0.00
AGGREGATE	0.00	0.00	*****	*****	*****
MAXIMUM	0.00	0.00	****	2.00	0.01
MINIMUM	0.00	0.00	****	0.66	0.00
AVERAGE	0.00	0.00	0.00	1.33	0.00

OSS 2009-07-22 19:09:55.0

Dss Name	Read Rate	Write Rate	%CPU	%Space Used	%Inodes Used
f0074	0.00	0.00	0.00	0.66	0.00
f0075	0.00	0.00	0.10	0.66	0.00
f0076	0.00	0.00	0.00	0.66	0.00
f0077	0.00	0.00	0.00	0.66	0.00
AGGREGATE	0.00	0.00	*****	*****	*****
MAXIMUM	0.00	0.00	0.10	0.66	0.00
MINIMUM	0.00	0.00	0.00	0.66	0.00
AVERAGE	0.00	0.00	0.02	0.66	0.00

lwatch





LWatch-lustre

File Configure example

example-MDT0000 2009-07-27 05:20:25.0

%CPU		%KB		%Inodes	
0.10		0.74		0.00	
Operation	Samples	Sample /Sec	Avg Value	Std Dev	Units
add_conn	0	0.00	0.00	0.00	reqs
adjust_kms	0	0.00	0.00	0.00	reqs
attach	0	0.00	0.00	0.00	reqs
bnw	0	0.00	0.00	0.00	reqs
bnw_asymc	0	0.00	0.00	0.00	reqs
cancel	0	0.00	0.00	0.00	reqs
cancel_unused	0	0.00	0.00	0.00	reqs
change_cbdata	0	0.00	0.00	0.00	reqs
checkmd	0	0.00	0.00	0.00	reqs
cleanup	0	0.00	0.00	0.00	reqs
close	0	0.00	0.00	0.00	reqs
commitnw	0	0.00	0.00	0.00	reqs
connect	0	0.00	0.00	0.00	reqs
copy	0	0.00	0.00	0.00	reqs
create	0	0.00	0.00	0.00	reqs
del_conn	0	0.00	0.00	0.00	reqs
destroy	0	0.00	0.00	0.00	reqs
destroy_export	0	0.00	0.00	0.00	reqs
detach	0	0.00	0.00	0.00	reqs
disconnect	0	0.00	0.00	0.00	reqs
enqueue	0	0.00	0.00	0.00	reqs
extent_calc	0	0.00	0.00	0.00	reqs
get_info	0	0.00	0.00	0.00	reqs
getattr	0	0.00	0.00	0.00	reqs
getattr_async	0	0.00	0.00	0.00	reqs
getxattr	0	0.00	0.00	0.00	reqs
health_check	0	0.00	0.00	0.00	reqs
import_event	0	0.00	0.00	0.00	reqs
init_export	0	0.00	0.00	0.00	reqs
ioccontrol	0	0.00	0.00	0.00	reqs
iterate	0	0.00	0.00	0.00	reqs
join_lru	0	0.00	0.00	0.00	reqs
link	0	0.00	0.00	0.00	reqs
llog_finish	0	0.00	0.00	0.00	reqs
llog_init	0	0.00	0.00	0.00	reqs
match	0	0.00	0.00	0.00	reqs
merge_lvb	0	0.00	0.00	0.00	reqs
migrate	0	0.00	0.00	0.00	reqs
mkdir	0	0.00	0.00	0.00	reqs

OST		2009-07-27 05:20:30.0			
Dst Name	Read Rate	Write Rate	%CPU	%KB	%Inodes
example-OST0000	0.00	0.00	****	8.46	0.00
example-OST0001	0.00	0.00	****	8.46	0.00
example-OST0002	0.00	0.00	****	16.91	0.00
example-OST0003	0.00	0.00	****	8.46	0.00
AGGREGATE	0.00	0.00	****	****	****
MAXIMUM	0.00	0.00	****	16.91	0.00
MINIMUM	0.00	0.00	****	8.46	0.00
AVERAGE	0.00	0.00	0.00	10.57	0.00

OSS		2009-07-27 05:20:30.0			
Oss Name	Read Rate	Write Rate	%CPU	%Space Used	%Inodes Used
10076	0.00	0.00	0.00	8.46	0.00
10077	0.00	0.00	0.00	8.46	0.00
10078	0.00	0.00	0.00	16.91	0.00
10079	0.00	0.00	0.00	8.46	0.00
AGGREGATE	0.00	0.00	****	****	****
MAXIMUM	0.00	0.00	0.00	16.91	0.00
MINIMUM	0.00	0.00	0.00	8.46	0.00
AVERAGE	0.00	0.00	0.00	10.57	0.00

What's missing?

- there are a few tools out there which make working with Lustre easier
- problem: no complete solution, combining all the tasks
- shine looks promising, but: not all features are implemented yet
- LMT 2 looks good, but: setup is hard and relies on a lot of LLNL packages

Wish list

- Lustre management/monitoring solution
 - setting up Lustre with fail over
 - plug-ins for different HA solutions
 - monitor errors on all nodes
 - Lustre self-healing and/or provide hints on how to fix problems
 - integrated monitoring solution
 - ...



Thank you! Questions? Comments?
svtr@sun.com