

Storage performance

Christine Leroy
c.leroy@cea.fr

Plan

- Storage type at GRIF
- Trick to improve performance
- Monitoring the storage

Storage type at GRIF

- GRIF site is composed of 6 sub-sites with 6 storage Elements
 - Alls sub-sites using DPM glite 3.2 under SL5
 - Different technologies for SEdisks (SAN, DAS)
 - File system used: xfs, ext4, ext3
 - Network interface: 1Gb aggregation (up to 5), 10Gb/s

Trick to improve performamnce

- DPM Specific:

- Splitting the filesystems of one SE disk across the different DPM pools
- Tuning mysql:

```
set-variable=max_connections=300
innodb_buffer_pool_size = 2048M
innodb_additional_mem_pool_size=128M
innodb_flush_method=O_DIRECT
log-slow-queries
```

- RAID and Filesystem configuration to respect the 200mb/s/TB
- network interface to respect the 200mb/s/TB (agregation of 1Gb/s or 10Gb/s)
- Network parameter tuning via sysctl (<http://lcg.in2p3.fr/wiki/index.php/TCP-Tuning>):


```
net.ipv4.tcp_rmem
net.ipv4.tcp_wmem
net.core.rmem_default
net.core.wmem_default
net.core.rmem_max
net.core.wmem_max
net.ipv4.tcp_dsack
net.ipv4.tcp_sack
net.ipv4.tcp_timestamps
net.core.netdev_max_backlog
```

Monitoring the storage

- Monitoring functionality: Nagios local + Nagios Regional
- Monitoring performance: testing the GRIDPP monitoring tool (http://www.gridpp.ac.uk/wiki/DPM_Monitoring)
- Monitoring the usage (accounting): We are extracting (using perl program) data from cns_db for user/group accounting, from dpm_db for token states, and we are parsing the result of dpm-qryconf for the pool state. All this data is then registered in our accounting DB to be graphed with BIRT, (like our local Computing accounting) : Work in progress

Monitoring

Nagios Regional



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

Current Network Status
 Last Updated: Wed Nov 10 15:09:03 CET 2010
 Updated every 601 seconds
 Nagios® Core™ 3.2.1 - www.nagios.org
 Logged in as /O=GRID-FR/C=FR/O=CEA/OU=IRFU/CN=Christine Leroy

[View History For This Host](#)
[View Notifications For This Host](#)
[View Service 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	Pen
9	0	0	0	0

All Problems	All Types
0	9


Service Status Details For Host 'node12.datagrid.cea.fr'

Host ↑↓	Service ↑↓	Status ↑	Last Check ↑↓	Duration ↑	Attempt ↑	Status Information
node12.datagrid.cea.fr	hr.srce.SRM2-CertLifetime	OK	11-10-2010 14:28:25	0d 8h 40m 38s	1/2	CERT LIFETIME OK - Certificate will expire in 233.02 days (J 2011 GMT)
	org.sam.SRM-All-/ops/Role=lcqadmin	OK	11-10-2010 14:13:18	0d 9h 55m 45s	1/4	OK: success.
	org.sam.SRM-Del-/ops/Role=lcqadmin	OK	11-10-2010 14:14:07	0d 9h 55m 10s	1/4	OK: File was deleted from SRM.
	org.sam.SRM-Get-/ops/Role=lcqadmin	OK	11-10-2010 14:14:04	0d 9h 55m 12s	1/4	OK: File was copied from SRM. Diff successful.
	org.sam.SRM-GetSURLs-/ops/Role=lcqadmin	OK	11-10-2010 14:13:23	2d 16h 55m 48s	1/4	OK: Got SRM endpoint(s) and Storage Path(s) from BDII
	org.sam.SRM-GetTURLs-/ops/Role=lcqadmin	OK	11-10-2010 14:13:51	0d 9h 55m 20s	1/4	OK: protocols OK-[rfio, gsiftp]
	org.sam.SRM-Ls-/ops/Role=lcqadmin	OK	11-10-2010 14:13:35	0d 9h 55m 33s	1/4	OK: listing [/dpm/datagrid.cea.fr/home/ops/testfile-put-12893c8846ff444ee.txt]-ok:
	org.sam.SRM-LsDir-/ops/Role=lcqadmin	OK	11-10-2010 14:13:24	0d 9h 55m 43s	1/4	OK: Storage Path[/dpm/datagrid.cea.fr/home/ops]-ok:
	org.sam.SRM-Put-/ops/Role=lcqadmin	OK	11-10-2010 14:13:34	0d 9h 55m 34s	1/4	OK: File was copied to SRM.

9 Matching Service Entries Displayed

Monitoring

Nagios Local (head node DPM + SEDisk DPM)



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
- Scheduling Queue

Current Network Status
 Last Updated: Wed Nov 10 15:18:12 CET 2010
 Updated every 90 seconds
 Nagios® 3.1.0 - www.nagios.org
 Logged in as *grif*

[View History For This Host](#)
[View Notifications For This Host](#)
[View Service 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
22	0	0	1	0
All Problems		All Types		
1		23		

**Service Status Details For Host
'node12.datagrid.cea.fr'**

Host ↑↓	Service ↑↓	Status ↑	Last Check ↑↓	Duration ↑↓	Attempt ↑	Status Information
node12.datagrid.cea.fr	NCM logfiles	CRITICAL	11-10-2010 14:50:33	0d 4h 27m 37s	1/1	ERROR! Last 5 lines of logfile /var/log/ncm-cdispd.log match regexp 'ncm-ncd finished.' [1-9]-[09]*\$!
	SMART health	OK	11-10-2010 14:27:02	34d 3h 18m 25s	1/2	OK: no SMART errors detected
	SPMA logfiles	OK	11-10-2010 14:52:01	34d 3h 16m 43s	1/1	OK: regexp 'ERROR.*SPMA.*finished' not found in last 1 lines of /var/log/spma.log
	System checks	OK	11-10-2010 14:26:21	19d 20h 6m 19s	1/2	OK: no problem detected.
	[Security Check] hidden processes	OK	11-10-2010 14:26:14	33d 15h 43m 55s	1/3	OK. No hidden proc found, no phalanx2 known trace found.
	cciss drives status	OK	11-10-2010 14:26:20	33d 21h 42m 25s	1/3	OK - nothing alarming found on cciss drives
	dpm daemon	OK	11-10-2010 15:09:41	2d 9h 44m 50s	1/1	TCP OK - 0.002 second response time on port 5015
	eth0.net stats	OK	11-10-2010 15:16:15	34d 3h 9m 0s	1/3	NETIO OK - eth0: RX=220657441016, TX=111889151631
	firewal	OK	11-10-2010 15:08:04	0d 14h 10m 34s	1/2	OK
	host certificate	OK	11-10-2010 14:26:49	34d 3h 6m 0s	1/3	OK - cert expires on Jul 1 13:56:38 2011 GMT
	hung ncm-cdispd processes	OK	11-10-2010 15:05:56	0d 14h 13m 18s	1/2	CMD: /bin/ps axwo 'stat uid pid ppid vsz rss pcpu etime comm args'
	hung nrpe processes	OK	11-10-2010 14:53:37	0d 14h 24m 33s	1/3	PROCS OK: 3 processes with command name 'nrpe'
	hung processes	OK	11-10-2010 15:01:06	34d 3h 1m 30s	1/2	There are 0 issues
	load average	OK	11-10-2010 15:16:31	0d 13h 52m 21s	1/3	OK - load average: 2.33, 2.10, 1.92
	memory usage	OK	11-10-2010 15:16:05	0d 5h 37m 46s	1/3	OK: Memory Usage: (W> 100, C> 100): 34% - Total: 7985 MB, used: 2755 MB, free: 5230 MB
	nrpe daemon	OK	11-10-2010 15:16:43	0d 14h 13m 50s	1/2	NRPE v2.12
	quattor daemons	OK	11-10-2010 15:16:42	0d 14h 13m 30s	1/2	OK
	read only filesystems	OK	11-10-2010 15:15:43	13d 6h 5m 37s	1/2	OK. No problem detected.
	root free space	OK	11-10-2010 14:38:36	34d 3h 22m 55s	1/3	DISK OK - free space: / 6824 MB (89% inode=98%):
	srmv2.2	OK	11-10-2010 15:10:37	0d 9h 48m 4s	1/1	PROCS OK: 1 process with command name 'srmv2.2'
	srmv2.2 daemon	OK	11-10-2010 15:12:26	0d 10h 16m 15s	1/1	TCP OK - 0.000 second response time on port 8446
	tmp partition freespace	OK	11-10-2010 15:16:25	0d 14h 12m 45s	1/3	DISK OK - free space: /tmp 9416 MB (98% inode=99%):
	var partition freespace	OK	11-10-2010 15:15:43	13d 6h 3m 30s	1/3	DISK OK - free space: /var 42844 MB (46% inode=99%):

Local SE accounting

