

Fabien Wernli

Some figures

Management Installation Administration

Monitorii Syslog RLS

Conclusion

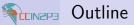
System Management and Monitoring at CC-IN2P3

Do sysadmins dream of IPMI 3.0?

Fabien Wernli

Centre de Calcul IN2P3

Rencontre Texas Instruments 8 Février 2008



Fabien Wernli

Some figures

Management Installation Administration

Monitoring

Syslog RLS Smurf

Conclusion

Some figures

Management 🍋

- Installation
- Administration

Monitoring

- Syslog
- ► RLS
- Smurf

Conclusion



CCIN2P3 Figures

System Management and Monitoring at CC-IN2P3

Eabien Wernli

Some figures

- more than 1102 linux computing nodes (oldest: 4 cores, newest: 8)
- Additional 478 worker nodes = 3824 cores "last week"
- tape library 4 STK9310, 1 SL8500
- \sim 2.8PB disk (DS8k+FAStT+Thumpers,) (XROOTD, DCACHE, HPSS, GPFS)
- Additional 800TB "last week" (IBM DCS9550)
- 600kW cooling
- 600+kW cooling in June 2007 (did we get that?)
- 1Gbit computing nodes, 1/2Gbit Disk servers, 2x2.5Gbit RENATER, 10GBit GEANT, 2x10Gbit CERN+GridKa
- more 10Gbit links to come (GRIF, PIC, ...)



Fabien Wernli

Some figures

Management Installation Administration

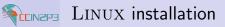
Monitoring

Syslog RLS Smurf

Conclusion

Question		
	now	very soon
comp. nodes	1102	1580
Disk	2.8PB	3.6PB
Sysadmins	6	5 (one went to Korea)

How to cope with so many toys?



Fabien Wernli

Some figures

Management

Installation Administration

Monitoring Syslog BLS

Conclusion

- ► Home-grown installation system : EPIMETHEUS
- Clients: KICKSTART/ANACONDA
- Template-like configuration "à la" QUATTOR
- Post-install scripts
- ► CVS
- We handle upgrades between major OS releases only if absolutely necessary
- ▶ supported platforms: *i386* and *x86_64*
- $\blacktriangleright~$ Scales well: ~ 200 simultaneous clients aren't a problem



Fabien Wernli

Some figures

Management

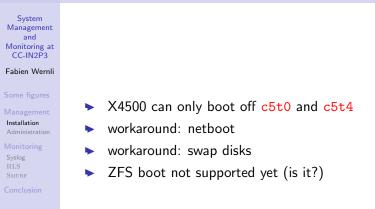
Installation Administration

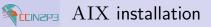
Monitorin Syslog RLS Smurf

Conclusion

- Custom JUMPSTART Installation
- Home-grown template scripts
- post-installation
- We do not handle upgrades between major OS versions
- This can indeed be a PIA and often breaks everything
- Reinstallation is simpler, safer and fast
- supported platforms: x86 and sparc

SOLARIS installation (Thumper)





Fabien Wernli

Some figures

Management

Installation Administration

Monitoring Syslog RLS Smurf

Conclusion

Standard NIM installation

- Postinstall scripts (I guess: ask our local AIX guy)
 - We still do not handle upgrades between major OS versions (still a guess ;))



$CONSERVER \mbox{ operation}$ What to do if there is no ${\rm KVM}$ on 1K nodes



Fabien Wernli

Some figures

Management Installation Administration

Monitoring

Syslog RLS SMURF

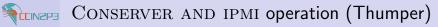
Conclusion

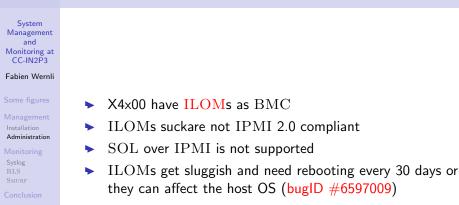
- At least one console server (conserver) per subnet
- Oldest consoles: serial port on serial concentrator, like Equinox ESP-16, perle IOLAN STS, ...
- Various SOL systems do not support IPMI 2.0
 - expect scripts for proprietary systems

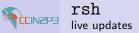
Examples

Supermicro ipmicli, DellPoweredge 2850 solproxy, IBM e326m IPMI 1.5, Sun X4x00 do not support IPMI 2.0 SOL

- Newer boxes all support IPMI 2.0 SOL: simple ipmitool wrapper
- Most scripts submitted to http://conserver.com/contrib
- ▶ Need to setup a "special" conserver for maintenance Techs

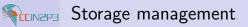


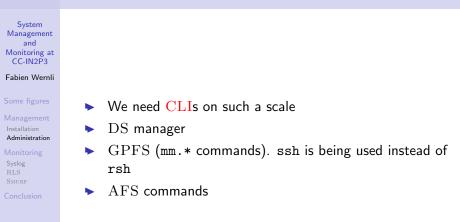




- Fabien Wernli
- Some figures
- Management Installation Administration
- Monitorin Syslog RLS SMURF
- Conclusion

- Massive configuration updates (no config change framework for all 3OS')
- Mostly at jobs sent over via rsh: AOR (at-over-rsh)
- Use hourly crontab used on all (AFS) nodes
- rsh with AFS token no longer supported since OPENAFS 1.2.13+
- > This is being used by some users, and by admins
- Replacement needed: ssh with pubkey auth and proper credentials: need newer OPENSSH than one provided in SL (sshd_config: Option Match appeared in 4.4)







Central syslog server

System Management and Monitoring at CC-IN2P3

Eabien Wernli

- Syslog

syslogd.conf

Example

*.notice

@ccsyslog.in2p3.fr

- syslog agents:
 - kernel messages (klogd)
 - snmptrapd daemons on conservers (trap e.g. V40z or X4x00 alerts)
 - application messages if asked for by service admins
- swatch (http://swatch.sourceforge.net/ triggers alerts for different patterns

Examples

mmfs:(.*), kernel: [a-z]+: status error: . * DriveReady SeekComplete

> Eabien Wernli System Management and Monitoring at CC-IN2P3



Thumper Alerts

System Management and Monitoring at CC-IN2P3

SNMP traps:

Eabien Wernli

Syslog

Examples

set /SP/alert/rules/1 destination=134.158.168.251 level=information

SEL

Examples

ipmitool -Ilanplus -Uroot -Pdontchangeme -H ilom57 sel elist

We still need to add more specific traps and filters (PEF)



$\operatorname{RLS}_{\operatorname{An alert}}$ messaging service

System Management and Monitoring at CC-IN2P3

- Fabien Wernli
- Some figures
- Management Installation Administration
- Monitoring Syslog RLS
- Conclusion

Alert service

Many agents

Examples

- ► Filesystem full (local daemon)
- I/O error (swatch)
- HPSS errors
- Refused connection (tcpd)
- ▶ WEBRLS client with hyperlinks to repair actions
- X11 client with support for acknowledging

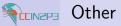


$\begin{array}{l} SMURF\\ \text{System Monitoring using } RRD{}_{\mathrm{TOOL}} \end{array}$

- System Management and Monitoring at CC-IN2P3
- Fabien Wernli
- Some figures
- Management Installation Administration
- Monitoring Syslog RLS SMURF
- Conclusion

- System monitoring of all nodes
- CPU, disk, IO, RAM, ...
 - The Matrix: http://ccsmurf:8080/thematrix/ = overview of activity of all nodes at a given time
 - Uses <u>RRDTOOL</u> http://oss.oetiker.ch/rrdtool/
 - Central server: data pushed by nodes using ssh every 10 min (15s resolution)
- DRRAW (http://web.taranis.org/drraw/
- Could be used to monitor various Thumper specifics (*e.g.* overview of all disks, number of SEL entries, ...)
 Example: monitoring of controller activity

http://ccsmurf:8080/cgi-bin/drraw.cgi?Mode=view&Graph=1163773235.20833

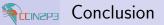


Fabien Wernli

Some figures

- Management Installation Administration
- Monitoring Syslog RLS SMURF
- Conclusion

- mail
- ► SMS
- GPFS statistics
- ► NGOP
- LEMON
- NAGIOS
- ► DCACHE



Fabien Wernli

Some figures

Management Installation Administration

Monitoring

Syslog RLS Smurf

Conclusion

The bottom line

Administrating thousands of systems using a few sysadmins is either cool or lethal