



Deployment work at CERN: monitoring task

WP4 workshop

Barcelona project conference 5/03

Jan van Eldik CERN IT/FIO



The Cern CC

- ◆ Different functionalities (interactive, batch, diskserver, www...)
- ◆ Different OS (RH6*, RH7*, RHAS, Solaris, Windows)
- ◆ Different users (production- and service-managers, "normal" users...)
- ◆ Different owners (experiments, developers, service-providers)
- ◆ Different sysadmins (service-managers, outsourced company...)
- ◆ Different ideas about monitoring...

this talk focuses on IT/FIO Linux production farms



Cern CC Linux production clusters -I

Name	Type	Nodes	OS
LXPLUS	Interactive	100	RH73, RH61
LXBATCH	Batch	1000	RH73, RH61
Diskserver		200	RH7*, RH6*
Tapeserver		80	RH7*, RH6*
~15 clusters	Mixed...	100	RH73, RH61



Cern CC Linux production clusters -II

- ◆ ~20 **production** clusters, with ~1500 nodes
- ◆ All need exception and/or monitoring
- ◆ The WP4 MSA provides this since June 2002...



Installing and configuring MSA

- ◆ Clusters are described in BIS database
- ◆ CERN-CC-MSA RPM installed (by hand) in /usr/local/MSA
- ◆ 3 sensors: sensorLinuxProc, FioSensor.pl, internal
- ◆ CCConfig + private Perl module (+SUE) generates MSA.cfg
- ◆ Between 80 and 120 metrics per node
- ◆ Sampling intervals between 60 seconds and 1 day
- ◆ Local cache in /var/spool/MSA
- ◆ LSB compliancy: lograte, startup scripts, chkconfig
- ◆ needs a few Perl modules: Proc::ProcessTable, Digest::MD5
- ◆ Installation is AFS independent...



Operator alarms

- ◆ Cern CC operators use (home grown) SURE alarm display
- ◆ MSA sends status information (ie alarms)
- ◆ Dedicated sensor ForSure.pl:
 - Knows desired state values for ~60 metrics
daemons, CPU load, file system occupancies, software versions...
 - Reads actual values from local cache
 - Compares them to builds an "alarm state"
 - Sends this information every 5 minutes (if not: NO_CONTACT alarm)
- ◆ SURE is limited: no historical data, no performance data, no global correlations, does not scale to 10K nodes, frozen



Measurement repositories

- ◆ MSA sends its data to several repositories
 - PVSS
 - Commercial SCADA tool, Cern has a license
 - Toolkit to build GUIs, archiving (FlatFile / Raima), scripting language, ...
 - Not intended for Computer Centre use
 - WP4 fmonservers:
 - ORACLE
 - Handles 500 inserts per second
 - Installed on RH Advanced Server
 - FlatFile archive
 - does not scale (runs out of inodes...)
 - deprecated, but useful for development



What's next

- ◆ IT/FIO/SM: "historical data must reside in Oracle" -> Pvss2Oracle
- ◆ Fabric monitoring solution for LCG1 @ Cern: decision end of May
- ◆ PVSS solution has operator interface, WP4 not yet...
- ◆ Pvss2Oracle not yet fully tested...

to be continued...