Getting EDG Going at Oxford

Ian Stokes-Rees

In The Beginning...

- Oct/Nov 2002 Carmine Cioffi and Pete Gronbech, with help from Steve Traylen, installed EDG v1.2.x
- Small site to experiment with EDG CE, SE, and WN:
  - 20 Gig storage
  - handful of PII or PIII sub 1 GHz Pentium WNs
  - LCFG Controlled
- Objective:
  - Become familiar with EDG software
  - Participate in GridPP with active site
  - Learn about LCFG configuration/control
  - Platform for experimenting with Grid software

Upgrade to EDG v1.4.x

- Dec/Jan 2002/3
- Took much longer than expected (two weeks of almost full-time effort)
- Discovered sensitivity of LCFG configuration parameters (site-cfg.h!)
- All EDG v1.4.x series upgrades and published bug fixes were easy
- Once delicate house of cards configured and running, and no one breathed too hard, it ran reasonably well, except:
  - WNs ran out of HD space for applications
  - LCFG server ran out of space for RPMs
  - CE ran out of space for users (lots of large jobs and left-overs from jobs)

Extending EDG v1.4.x

Since everything was running reasonably smoothly, I decided to start "experimenting"

Nagios

Even with only 7 machines, monitoring needed.

- OK, WARN, CRITICAL, FAIL, UNKNOWN service and host states
- hierarchical alerts and monitoring
- web and command line interfaces
- supports Windows and Linux clients
- easy to install, somewhat complicated to configure, very easy to maintain

RRDtool

Nagios didn't provide good historical tracking or performance metrics, so turned to RRDtool to graph.

- Nagios monitoring channels could be fed into RRDtool DBs
- simple command line interface to file based DBs
- only holds data for period T into the past (from $t_now - T tot_now$)
- aggregates "older" data so DB never grows above fixed size
- generates MRTG style graphs
- simple built-in CGI interface to display graphs of DB statistics

A few more small comments on Monitoring

Ganglia

Uses RRDtool for cluster monitoring. Don't think it has Nagios features. Still want to try it out...

GridICE

Uses Nagios for cluster monitoring. Don't think it has RRDtool features. Documentation difficult to come by ...

Maybe someone will unify, simplify, and rationalise RRDtool and Nagios into a single package, configured by LCFGng, and taking monitoring channels from R-GMA.

Security and RH6.2

- Widespread compromise of RH6.2 systems in PP community, including the OxGrid CE
- EDG activity had died down, seemed like good time to experiment with upgrade to EDG v1.4.x on RH7.3 system
- Commenced early April in attempt to get EDG v1.4.11 running from LCFGng server on all RH7.3 CE, SE, and WNs

Famous Last Words ...

```
Date: Mon, 7 Apr 2003 09:39:41 +0100 (BST)
From: Ian Stokes-Rees <stokes@physics.ox.ac.uk>
To: Nadia Lajili <nlajili@in2p3.fr>
Subject: Re: GRIS unreachable

Due to recent security issues we are likely to
attempt to upgrade our entire site to RH7.3
based system.  It may not be available for a
few days in the meantime.

Ian.
```

91 Days Later

- The Oxford EDG site is now running RH7.3
- Actually only "down" for 4 weeks (mid April to mid May)
- But ... only running Globus and PBS successfully
- Still not talking to RB
- Still trying to get v1.4.x RPMs to install under RH7.3
- Whole point of excercise is to avoid recompiling or hand hacking components in — want stadard(ish) RPMs installed from LCFGng

Experience

- Lack of documentation for SysAdmins who are not intimately involved with EDG software makes installation and configuration difficult. Numerous "gotchas".
- LCFGng easier to use, but still not especially well documented
  - Too many things happen "auto-magically"
  - Web interface OK, but would like Plain Old Logfiles (TM)
  - Don't like C Preprocessor format – too many opportunities for errors
  - Really should be XML with XInclude and XSDL
  - "But it does end up as XML!" you say — Does it?
    `<jobManager_RECORD cfg:name=""fork">`
  - "Just ignore those error or warning messages" is not good enough — some of them are significant

Bugzilla Issues

- Developers need to encourage Bugzilla use and use it properly
- Need more problems being reported to Bugzilla
- Developers must accept bug reports in terms which users are prepared to supply
- Example — nginstallroot:
  - nginstallroot problem took me days to resolve – turned out to be well known bug which hadn't been fixed
  - May 13: I report bug with best information I can find
  - same day, I get response saying I've installed it incorrectly. "Delete it and try again." Bug closed as INVALID.
  - May 20: I have now struggled on for days to try and sort this out and re-report the bug, with some different details
  - May 21: Bug is marked as duplicate of March 14 bug report,

and then CLOSED with WORKSFORME

- In fact, bug still exists, and definition of "WORKSFORME" is if you happen to have read bug #875 and applied hack only documented there, then it should "WORK"
- I point this out, and then developer agrees to actually fix it.

Developers are too close to the code and do not accept that things which are obvious and routine for them are not for the typical Sys Admin

Final Comments

- Far too many dependencies between RPMs — some possibly unnecessary — "Just use –nodep or –force" not a good way forward
- Probably too many RPMs (100s, just for EDG specific s/w)
- Difficult to get a picture of what is going on in order to debug effectively
- Is there value in maintaining EDG v1.4.x series s/w longer?
- Definitely need more documentation, FAQs, and tutorials for software
- Architecture and site requirements to fit in to EDG not particularly clear
- LCFGng makes debugging individual nodes difficult — changes are randomly erased (i.e. they may or may not be erased, and

they may or may not be erased before a reboot)

- Need some good installation walk-throughs for Sys Admins who are not regularly involved with EDG
- TB-Support and GridPP community has been excellent resource
- Concern about required homogenaity EDG systems and inflexibility in configuration — doesn't sound like "Heterogeneous federated computing resources" to me

Bonus: RB Weirdness

I have witnessed the RB do some strange things, but this one is at the top of the list:

```
Fri May 23 11:13:56 2003: gm03.hep.ph.ic.ac.uk (RB) accepts job
Fri May 23 11:13:56 2003: gppui04.gridpp.rl.ac.uk transfers job to RB

... four days pass ...

Tue May 27 14:21:37 2003: RB matches job to
tuber5.phy.bris.ac.uk:2119//jobmanager-pbs-tbq
Tue May 27 14:21:41 2003: JSS refuses jobs at gm03 with "condor command
failed" msg
Tue May 27 14:21:41 2003: RB accepts job back from JSS
Tue May 27 14:21:58 2003: RB logs JobPending state with reason
"Resubmitting"

... six days pass ...
```

```
Mon Jun  2 11:15:26 2003: RB matches job to
bottom.phy.bris.ac.uk:2119//jobmanager-pbs-gridq
Mon Jun  2 11:16:22 2003: JSS refuses job as before, RB accepts job back,
logs JobPending

... two hours pass ...

Mon Jun  2 13:33:42 2003: RB matches job to
tuber5.phy.bris.ac.uk:2119//jobmanager-pbs-bseq
Mon Jun  2 13:34:33 2003: JSS refuses job as before, RB accepts job back,
logs JobPending

... 90 minutes pass ...

Mon Jun  2 15:06:45 2003: RB matches job to
testbed008.cnaf.infn.it:2119//jobmanager-pbs-short
Mon Jun  2 15:32:12 2003: JSS refuses job as before (but takes 25 minutes to
decide)
```

```
Mon Jun  2 15:32:30 2003: RB holds job in JobPending state with reason
"Resubmitting"

... one day passes ...
```

and there are no more log entries. Its current status is ”Waiting”, with StatusReason ”Resubmitting”.