
CERN

Batch Processing at CERN

Jérôme Belleman, Ulrich Schwickerath, Paolo Stivanin – IT-PES-PS

(Present of) Batch Processing at CERN

Jérôme Belleman, Ulrich Schwickerath, Paolo Stivanin – IT-PES-PS

Outline

Current Setup

Operational Issues and Recent Incidents

Rebooting/Reinstalling Our Cluster

Need to Improve Workflows

Current Setup

A Platform/IBM LSF 7.0.6 Cluster

- $\approx$ 4 000 nodes
  - SLC5 $\xrightarrow{100\%}$ SLC6
  - Physical $\xrightarrow{92\%}$ Virtual machines
  - Quattor $\xrightarrow{100\%}$ Puppet
- $>$ 65 000 cores
- 400 000 jobs/day
- $\pm$45 000 running jobs

Operational Issues and Recent Incidents

Daily Operations

- Crashed nodes
- Renewing versions
- Various faults

Security Incidents

- Heartbleed
- A nameless one during Xmas
- GHOST
- One involving one of our own in-house scripts

$\rightarrow$ Reboot/Reinstall the cluster

Rebooting/Reinstalling Our Cluster

Warning Users

We have dedicated resources:

- Notify users
- Alleviate capacity loss

Draining Worker Nodes

- Preferably by stage-draining them
- For a node, up to 2 weeks
- In several $\approx 10\%$ chunks
- Manually track, kill stuck jobs

Applying Fixes

Installing RPMs:

- Easy, in principle
- Broken Yum/RPM DBs

Running Puppet:

- Not a problem if nodes run it regularly
- Sometimes, they just don't
- Crashed nodes

Rebooting

- Broadcasting reboot command
- Soft reboot for VMs
- Some nodes get stuck when rebooting
- Some nodes even get stuck when shutting down

Reinstalling Physical Nodes

```
ai-installhost lxbcd0123
```

- Blind operation
- Some nodes sometimes never come back alive

Reinstalling Virtual Nodes

```
ai-bs-vm --flavor large -g batch/share -i 'SLC6 CERN' b6789abcde
```

- Fresh, faulty VMs
- Manual fixing
- Faster to remove and spawn a new one?

Need to Improve Workflows

Making Worker Nodes

```
ai-bs-vm --flavor large -g batch/share -i 'SLC6 CERN' b6789abcde
```

- Keep creating as many VMs as possible
- Make sure they're not faulty
- Park them in spare
- Move them to shared resources
- Remove faulty ones

Making Worker Nodes

```
batchthatch make 100
```

- Keep creating as many VMs as possible
- Make sure they're not faulty
- Park them in spare
- Move them to shared resources
- Remove faulty ones

Making Worker Nodes

```
batchthatch make 100
```

- Keep creating as many VMs as possible
- Make sure they're not faulty
- Park them in spare
- Move them to shared resources
- Remove faulty ones

→ Batch Factory. What about Heat? And Vcycle?

Resetting Worker Nodes

batchnudge

- Listen on GNI message bus
- Look for `no_contact` exceptions
- SSH, ping, count jobs, check console
- Reboot
- Make sure it breathes again

Resetting Worker Nodes

batchnudge

- Listen on GNI message bus
- Look for `no_contact` exceptions
- SSH, ping, count jobs, check console
- Reboot
- Make sure it breathes again

Other exceptions? $\rightarrow$ Batch Factory

Central vs. Distributed Management

A central, robust conductor:

- Using e.g. MCollective/wassh/Parallel SSH
- Some say it doesn't scale

vs.

Responsible worker nodes:

- Which can self-heal
- Serf?

Collaboration

- Avoid interferences
- A central place to track work
- Roger to record states and comments

Delegation

- What if not all can be automated?
- What if our Sys Admin Team could help?
- Rundeck

Conclusion

Outlook

- Some components are already there
- Now glueing them together
- Tools we need now...
- ... and which we'll use with HTCondor

www.cern.ch