



Toolforge webservices are in the final stages of [migrating to the toolforge.org domain](#).
Please help us clean up older documentation referring to tools.wmflabs.org!

Incident documentation/20150817-ToolLabs-WebgridOutage

[< Incident documentation](#)

Contents [\[hide\]](#)

- [1 Summary](#)
- [2 Timeline](#)
 - [2.1 2015-08-17](#)
 - [2.2 2015-08-18](#)
- [3 Conclusions](#)
- [4 Actionables](#)

Summary

The grid engine exec daemon did not restart correctly on all host after reboots, causing the trusty webgrid queue to be overloaded. This in turn prevented webservice jobs from (re)starting, which caused outages for several tools on Tool Labs.

Timeline

- Earlier: tools-webgrid-lighttpd-1404 and tools-webgrid-lighttpd-1406 are rebooted, and gridengine-exec does not start correctly
- 2015-08-17 16:17: Andrew disables tools-webgrid-lighttpd-1406, marking the start of the downtime

2015-08-17

- 21:10 Valhallasw receives complaints about webservice not starting on a private channel and investigates
- 21:20 notices job is stuck in 'qw' state. Notices this is due to a lack of available vmem, but initially memory, but initially blames this on the 4GB h_vmem requirement of this job
- 21:31 Valhallasw realizes the 4GB is standard for webservice jobs, and calls on Coren and Andrew in #wikimedia-labs
- andrew suggests this is due to a lack of webgrid nodes
- 21:37 Valhallasw goes to sleep.
- 22:13 Andrew builds a new webgrid node
- 22:13 Andrew asks Yuvi for advice on how to add a new tools-webgrid-lighttpd node to the grid; Yuvi is on vacation in Europe and asleep.

2015-08-18

- 01:00 Complaints about webservices not starting on IRC; also on labs-l
- 07:25 Valhallasw looks into the issue again, spends 1.5 on building host and adding it to the grid due to no documentation
- 08:30 during configuration, Valhallasw notices queues on tools-webgrid-lighttpd-1403.eqiad.wmflabs and tools-webgrid-lighttpd-1404.eqiad.wmflabs are also in alarm state.
- 08:36 Valhallasw brings back online tools-webgrid-lighttpd-1403.eqiad.wmflabs and tools-webgrid-lighttpd-1404.eqiad.wmflabs, as well as the new webgrid host tools-webgrid-lighttpd-1411.tools.eqiad.wmflabs, marking the end of the downtime

Total downtime: approx. 16 hours

Conclusions

- exec nodes do not always start the gridengine-exec daemon on boot, causing the node to be unavailable for jobs

[Main page](#)
[Recent changes](#)
[Server admin log \(Prod\)](#)
[Server admin log \(RelEng\)](#)
[Deployments](#)
[SRE/Operations Help](#)
[Incident status](#)

[Cloud VPS & Toolforge](#)

[Cloud VPS documentation](#)

[Toolforge documentation](#)

[Request Cloud VPS project](#)

[Server admin log \(Cloud VPS\)](#)

[Tools](#)

[What links here](#)

[Related changes](#)

[Special pages](#)

[Permanent link](#)

[Page information](#)

[Cite this page](#)

[Print/export](#)

[Create a book](#)

[Download as PDF](#)

[Printable version](#)

- there was no documentation on building a new exec host, which made it difficult for Andrew and Valhallasw to do so
- maintenance (virt reboots) was planned in a time slot where Coren was unavailable, which meant there was no-one with SGE-specific knowledge around to investigate issues
- there is no monitoring for queue status, so we did not notice the queues were offline before issues arose
- there is no monitoring to check there are enough hosts available for current and predicted load

Actionables

- Status: ■ **Done** Make sure gridengine-exec starts on boot ([bug T109728](#))
- Status: ■ **Done** Write documentation on building new exec hosts ([bug T109417](#))
- Status: ■ **Declined** Add monitoring for queue status ([bug T109730](#))
- Status: ■ **Declined** Add monitoring for expected load issues ([bug T109732](#))

Category: [Incident documentation](#)

This page was last edited on 22 August 2015, at 21:14.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. See [Terms of Use](#) for details.

[Privacy policy](#) [About](#)
Wikitech

[Disclaimers](#) [Code of Conduct](#) [Developers](#) [Statistics](#) [Cookie statement](#) [Mobile view](#)

