



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/20171001-labvirt1015

[< Incident documentation](#)

Contents [\[hide\]](#)

- [1 Summary](#)
- [2 Timeline](#)
- [3 Conclusions](#)
- [4 Actionables](#)

Summary

On Friday, 2017-09-29, Andrew migrated a few VMs to the recently-repaired labvirt1015 virt host. Early on Sunday labvirt1015 crashed, causing downtime for all the hosted VMs.

Some time later the problem was discovered, and the affected VMS migrated to a more trustworthy host. Affected instances were:

- puppet-phabricator
- wdqs-deploy
- tools-clushmaster-01
- search-jessie
- gerrit-test3
- phab-01
- integration-slave-jessie-1003
- integration-slave-jessie-1004

Timeline

- Previously: during initial setup labvirt1015 was found to be unexpected halted several times. This issue was investigated and (we thought) fixed; this process is documented in [T171473](#)[🔗](#). During that process, Andrew dntimed the instance for 30 days to silence the alerts caused by both the crashes and the frequent reboots that were part of the debugging and repair process.
- 2017-09-19: Labvirt1015 has a new CPU, has been up for several days and is declared fixed.
- 2017-09-29: Labvirt1015 is still up and seems fine. As part of [T176044](#)[🔗](#) andrew rebuilds it with a known good kernel, and begins evacuating VMs from labvirt1016 to 1015 in order to free up 1016 for a rebuild.
- 2017-10-01 01:21:01: Labvirt1015 halts, taking all hosted VMs with it. This failure does not page or alert because the 30 days of downtime are still in progress.
- 2017-10-01 09:06: A user on IRC (paladox) reports a VM outage. The message is missed because it's in the early hours of Sunday morning for the WMCS team.
- 2017-10-01 19:00: Madhu notices that the host is down and that the downtime is affecting hosted instances.
- 2017-10-01 22:00: WMCS team converses on IRC and Whatsapp, agrees to reboot labvirt1015 and evacuate VMS.
- 2017-10-01 23:00: All VMs save one are migrated to labvirt1017 and resume normal function. The remaining VM is enormous and takes several hours to copy, but completes migration later in the evening.

Conclusions

- Labvirt1015 is not actually fixed, but its failure is intermittent and sneaky and it will be hard to know when we can trust it again.
- The only real mistake here was the silencing of Icinga warnings for labvirt1015, that resulted in a considerable delay before we responded to the outage.

[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](#)

Actionables

- Replace labvirt1015
- Include a review of current icinga acks and downtimes as part of closing any hardware-related task.

Category: [Incident documentation](#)

This page was last edited on 1 October 2017, at 22:52.

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](#)

