



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/20150331-LabsNFS-Overload

[< Incident documentation](#)

Contents [\[hide\]](#)

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

Summary

Almost all labs instances suffered super high load / kernel stalls, to a point where they were basically dead. rsyncs running as a result of [previous incident](#) were basically eating up all io on labstore1001, causing knock on effects on all instances. Killing the rsyncs restored service, but ultimate cause is unknown since the rsync was running for hours without issue, and was resumed later and has completed in a few hours without issue.

Timeline

```
05:37 Shinken sends email of beta cluster being down
05:43 Shinken sends email of toollabs being down. Several users also report th
is on IRC and labs-l around this time
06:00 Yuvi gets to a computer, starts investigating. Looks like several
instances were 'freezing' frequently, and not responding at all
06:20 Yuvi assumes that the problem is with the underlying Virt* hosts load
since that was the problem earlier, and fixes the one host with load (virt1002)
by turning off a few VMs that were causing load there.. No help. Can not debug
properly since ssh to most hosts I tried fails pretty badly, and am debugging
blind. More vague re-arranging of chairs
07:08 _joe_ manages to ssh in, finds out that instances are stalling due to
labstore issues with messages like
[Tue Mar 31 07:07:16 2015] nfs: server labstore.svc.eqiad.wmnet not
responding, still trying
[Tue Mar 31 07:07:29 2015] nfs: server labstore.svc.eqiad.wmnet OK
07:12: Labstore load abnormally high. 3 rsync processes running on a screen (one
rsync job).
07:16: All rsyncs are killed, things start returning to normal
```

Conclusions

labstore being unavailable / overloaded basically kills all of labs, and is a SPOF. Our NFS setup should be modernized and lots more alerting / monitoring set up.

Actionables

1. Thorough monitoring of our NFS setup <https://phabricator.wikimedia.org/T94606>
2. Spread knowledge of our NFS setup amongst our opsen better (Coren is doing a tech talk soon, I believe)
3. labstore1002 runs jessie, but switchover is not yet tested. Do switchover after a lot of testing, and re-install labstore1001 <https://phabricator.wikimedia.org/T94607> and <https://phabricator.wikimedia.org/T94609>

Category: [Incident documentation](#)

[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](#)

