



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/20150607-LabsNFS-Outage

[< Incident documentation](#)

## Summary

The Labs NFS server (labstore1001) suffered a 4-hour outage beginning at 01:00 on June 7th. The original cause of the outage is unknown but the duration was greatly increased due to errors on the switch. That issue has been patched over but may need further attention. The original cause of the problem may have been a side-effect of network congestion, or may have also be related to the switch failure; more investigation is needed.

## Timeline

- [01:00] First sign of trouble. Labstore1001 ganglia graphs get ugly, syslog starts saying "peername failed (err 107)"
- [02:00] Alex Monk texts Andrew Bogott with concerns about labs failures. Andrew is buying groceries.
- [02:15] Andrew joins IRC and investigates. Labstore1001 is running 100% IO, high CPU, but low-to-normal network usage. top/iotop only indicate cryptic 'nfsd' and 'kworker' processes as hogging IO.
- [03:00] Ori joins in. Andrew finally looks at dmesg and notes "peername failed (err 107)" which suggests a file descriptor shortage. Andrew tries various ways to restart or flush NFS[1], to no avail, leading to...
- [03:10] Andrew reboots labcontrol1001. When it comes back up, it has no network connectivity at all.
- [03:20] Andrew reboots labcontrol1001 again, and rubs eyes in disbelief.
- [03:20] Reedy graciously joins the hunt.
- [03:30] Icinga starts throwing puppet failures for multiple hosts. Puppet is refreshing geoip for hosts (and, in some cases, failing) Ori concludes that there is a more wide-reaching network issue.
- [03:45] Andrew texts Faidon
- [04:00] Andrew texts Brandon and Alex as well
- [04:05] Andrew gets wise and /calls/ Faidon who appears almost immediately
- [04:10] Faidon logs into the associated switch and discovers a host of errors like

```
Jun  7 04:04:55 asw-c-eqiad fpc1 MRVL-  
L2:mrvl_fdb_mac_entry_rebake(),482:fdb_mac_entry_mc_set() failed(-1)  
Jun  7 04:04:55 asw-c-eqiad fpc1 RT-HAL,rt_entry_topo_handler,4121:  
l2_half_vectors->l2_entry_rebake failed  
Jun  7 04:05:02 asw-c-eqiad fpc1 MRVL-  
L2:mrvl_fdb_mac_entry_mc_set(),1089:Sanity Checks Failed(Invalid Params:-2)
```

- [04:20] Faidon disables ge-3/0/5, ge-3/0/11, ae3 and waits to see if errors stop (they did)
- [04:35] Faidon dismantles labstore1001's bond0, sets up everything over eth0 again. Also reconfigures the switch accordingly (ae3 was deleted, ge-3/0/5 was enabled). A series of non-recurring switch error messages appear. labstore1001's connectivity is restored.
- [04:37] Faidon disables eth0 and tries eth1 / ge-3/0/11 instead, errors appear again.
- [04:42] Faidon re-reverts, disables eth1 but enables eth0. Another similar non-recurring flood of error messages are logged. labstore1001's connectivity is restored again.
- [04:58] Andrew starts NFS on labstore1001.
- [05:11] After 13mins of waiting for /srv/project recovery (suggesting an unclean shutdown), the mount ultimately fails with:

```
Jun  7 05:11:22 labstore1001 kernel: [ 5477.700259] EXT4-fs warning (device dm-  
9): ext4_enable_quotas:5026: Failed to enable quota tracking (type=0, err=-3).  
Please run e2fsck to fix.  
Jun  7 05:11:23 labstore1001 kernel: [ 5477.775181] EXT4-fs (dm-9): mount  
failed
```

[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](#)

- [05:11] Faidon attempts to remount /srv/project without running e2fsck and succeeds.
- [05:16] Faidon stops nfs-kernel-server & manage-nfs-volumes, unmounts everything and re-runs start-nfs.
- [05:18] NFS is made available again, service is restored.

[1] Running nfs-start on labstore1001 was scary because that script says not to do this unless labstore1002 is off and labstore1002 is /not/ off. That message needs to be fixed, or labstore1002 needs to be turned off, or I've done something horrible. My guess is it's the first thing. At any rate I verified that labstore1002 didn't have anything in /srv/projects so I take it to not be attached to the disks, or something.

## Actionables

---

- [Make Labs NFS alerts paging](#)
- [start-nfs script warning message is scary and wrong](#)
- [Make a flowchart for locating and halting misbehaving NFS clients](#)

Category: [Incident documentation](#)

This page was last edited on 8 June 2015, at 18:02.

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](#)

