



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/20151105-LabsOutage

[< Incident documentation](#)

Contents [\[hide\]](#)

- 1 [Summary](#)
- 2 [Timeline](#)
- 3 [Conclusions](#)
- 4 [Actionables](#)
 - 4.1 [Buy more virt nodes, with bigger hard drives](#)
 - 4.2 [If we have the space, reduce the overcommit ratio](#)
 - 4.3 [Address the fuzziness of overcommitting in some way to ensure that we don't schedule on full drives](#)
 - 4.4 [Page when a virt node is running out of disk space](#)

Summary

On 2015-11-05 at around 17:10, the VM partition of labvirt1002 filled up. This resulted in many of the hosted instances switching into a 'paused' state and becoming unreachable. One big instance was deleted and a few others migrated off, the affected instances unpaused, and full service restored by 17:45. Other than the disposable instance sacrificed to make space, no data was lost.

A netsplit on freenode slightly delayed response to this incident.

Timeline

- [earlier] Labvirt1002 has more than 100Gb of free space. Ebernardson creates an elastic search instance, estest1003, and the nova scheduler places it on Labvirt1002 because the overprovisioning rules permit it. estest1003 starts to sync with the search cluster and starts gobbling Gbs of disk space.
- [17:03] Ebernardson reports that estest1003 appears to have 'disappeared'
- [17:10] Hashar reports that some CI vms are misbehaving. Coren responds and determines that affected instances are all on labvirt1002.
- [17:15] Andrew joins the response and determines that the drive /var/lib/nova/instances is full on labvirt1002.
- [17:20] Coren begins migrating several instances to labvirt1005.
- [17:35] Free space on labvirt1002 is now at 1%, which means that still-running instances are out of danger.
- [17:37] With ebernardson's permission, andrew deletes the estest1003 instance, freeing up more than 100Gb of space.
- [17:38] <icinga-wm_> RECOVERY - Disk space on labvirt1002 is OK: DISK OK
- [17:40] Andrew gets a list of paused instances, on labcontrol1001:

```
$ sudo su -  
# source ~/novaenv.sh  
# nova list --all-tenants --host labvirt1002
```

- [17:40] Andrew unpauses affected instances, on labvirt1002:

```
$ sudo su -  
# virsh resume <instance-id>
```

- [17:45] All services are restored

Conclusions

Labs uses copy-on-write images so that disk space on a virt host is only consumed when an instance actually uses it. For example, a minimal Jessie image of size m1.small has a drive 'size' of 20Gb but initially only

[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](#)

consumes 751M of space on its host. The large majority of labs instances never fill their allocated drives.

For this reason we support overcommitting of disk space. This is a bit risky, as it's possible that a large number of instances could unexpectedly fill their drives at the same time, thus maxing out the host's drive space. Setting the overcommit ratio, therefore, is a compromise between guaranteed uptime (that would be a ratio of 1:1 allocated-space to permitted space) and efficient use of resources.

As of the week of the 5th, the overcommit ratio was 2.1. I (Andrew) arrived at that number experimentally by verifying that it prevented scheduling on hosts less than 10% available free space. It's all fuzzy, though, since everything is subject to the number and fullness of instances on any given node. Apparently the scheduler thought there was room for just one more instance, which was true except that that one last instance turned out to be one that was just about to grow to 150Gb in size.

Actionables


Buy more virt nodes, with bigger hard drives

✓ **Done** <https://gerrit.wikimedia.org/r/#/c/251954/> -- note that labvirt1010 (and, presumably, future purchases) have nearly 2x the hard drive space as labvirt1002.

If we have the space, reduce the overcommit ratio

✓ **Done** <https://gerrit.wikimedia.org/r/#/c/251954/>

Address the fuzziness of overcommitting in some way to ensure that we don't schedule on full drives

 **In progress** The proposed upstream patch <https://review.openstack.org/#/c/242251/> will allow us to demand a certain amount of free disk space on virt nodes. It still won't guarantee against a perfect storm of instance growth, but it will at least stop the scheduler for making bad choices regardless of overcommit math.

Page when a virt node is running out of disk space

✓ **Done** <https://gerrit.wikimedia.org/r/#/c/251297/> and <https://gerrit.wikimedia.org/r/#/c/251292/>

Category: [Incident documentation](#)

This page was last edited on 9 November 2015, at 16:19.

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

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

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

[Wikitech](#)

