



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/20190529-NFS-keystone

[< Incident documentation](#)

Contents [\[hide\]](#)

- 1 [Summary](#)
 - 1.1 [Impact](#)
 - 1.2 [Detection](#)
- 2 [Timeline](#)
- 3 [Conclusions](#)
 - 3.1 [What went well?](#)
 - 3.2 [What went poorly?](#)
 - 3.3 [Where did we get lucky?](#)
- 4 [Links to relevant documentation](#)
- 5 [Actionables](#)

Summary

What happened?

- We are trying to upgrade cloudcontrol1003 from Debian Jessie to Debian Stretch
- That involves reallocating control plane workload to cloudcontrol1004 which is already Debian Stretch
- The reallocation caused issues in NFS, affecting all NFS attached OpenStack projects

Impact

- Toolforge was down.
- Other Openstack projects using NFS had troubles using NFS.

Detection

Mixed detection mostly at the same time:

- human reports
- icinga alerts

Timeline

All times in UTC.

- 11:51 UTC Andrew merged [512954](#) (Make cloudcontrol1004 the primary keystone host)
- 11:54 UTC labstore1005 nfs-exportd fails to get project list from keystone
- 11:55 UTC labstore1004 nfs-exportd fails to get project list from keystone
- 11:57 UTC Arturo downtimes cloudservices1003.wikimedia.org because we plan to rebuild as stretch
- 11:57 UTC Arturo detected stashbot on IRC not responding
- 11:57 UTC icinga toolschecker alert on IRC about toolforge cron
- 11:59 UTC icinga toolschecker page: stale file handle for toolforge project NFS
- 12:08 UTC Arturo detects that nfs-exportd is using a hardcoded keystone server (cloudcontrol1003.wikimedia.org)
- 12:18 UTC Andrew merged [513097](#) (nfs-exportd: use cloudcontrol1004 endpoint for now)
- 12:25 UTC Krenair reports Toolforge is back apparently
- 12:29 UTC stashbot joins the #wikimedia-cloud-admin IRC channel, indicating that Toolforge is indeed good
- 12:34 UTC Brooke identifies some code in nfs-exportd that can be improved to handle a situation in which keystone returns 401 (not authed)
- 12:35 UTC stashbot joins the #wikimedia-cloud-admin IRC channel again
- 12:37 UTC icinga page for cloudcontrol1004: keystone_novaobserver_delete_tokens, may be a delayed

[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](#)

message for whatever reason

- 12:39 UTC Andrew decides to downtime a bunch of stuff in icinga while on operations
- 12:45 UTC Toolforge seems to be working correctly
- 12:49 UTC Brooke merges [513105](#) (nfs-exportd: if auth errors happen, do not proceed)
- 13:00 UTC we consider the incident done, all systems seems to be working and we understand the issue.

Conclusions

What weaknesses did we learn about and how can we address them?

The following sub-sections should have a couple brief bullet points each.

What went well?

- Alerts went out relatively quickly
- All the OpsEng in the WMCS team immediately were available to handle the outage. Worked together as a team pretty well.

What went poorly?

- The outage was the result of an upgrade operation that we were unable to reproduce in a testing/staging environment
- Since some of the storage environment is in flux at the moment, there were justified concerns that documentation would be outdated

Where did we get lucky?

- incident occurred when the most people were online to assist

Links to relevant documentation

Where is the documentation that someone responding to this alert should have (runbook, plus supporting docs). If that documentation does not exist, there should be an action item to create it.

Actionables

- [513105](#) -- bad branch in nfs-exportd code ✓ **Done**
- [513128](#) -- nfs-exportd: get essential openstack information from yaml files ✓ **Done**
- Improve nfs-exportd to consult keystone about the project list to recognize deleted projects
- hiera cleanup: novacontroller vs keystone_host
- hardcoding keystone endpoint. Many scripts/configs are harcoding keystone endpoint (cloudcontrol1003.wikimedia.org)
- use service FQDN everywhere rather than a concrete server (cloudcontrol: decide on FQDN for service endpoints [task T223902](#))
- revisit the design of the nfs-exportd code: we could remove any ability to drop exports in the code, but that's a security and practical concern because we reuse IP addresses.
- possibly introduce a maintenance flag for openstack services to watch for in hiera or a puppet-controlled file such as novaobserver.yaml
- We don't know why keystone @ cloudcontrol1003 refused to auth queries when we started failovering the service to cloudcontrol1004.
- Add unit tests to nfs-exportd.py -- we can absolutely check for how it behaves in countless failure states easily if it has tests like gridconfigurator does

Category: [Incident documentation](#)

This page was last edited on 30 May 2019, at 21:42.

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