



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/20180808-Network

[< Incident documentation](#)

Contents [\[hide\]](#)

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

Summary

Topology changes made to improve the redundancy and stability of the switch stack asw2-a-eqiad caused it to drop ~1/3 of the packets transiting through its members for about 1h. This packet drop caused internal services to timeout/retry, exact user facing issues TBD but at least [an increase of 5xx errors](#) [🔗](#).

Timeline (UTC)

17:14 - First topology change made

17:43 - Last topology change made ([T201145#4489225](#))

17:47 - First Icinga alerts, some high API latencies, puppetfails, etc. IRC spam is bad, but no major pages or signs of broader user-facing issues yet.

18:07 - Replaced fpc1-fpc3 link for [T201095](#) (Unaware of the alerts)

18:10 - Started investigating asw2-a-eqiad

18:30 - Disabled fpc1-fpc3 link

18:33 - Minor user-facing disturbances begin showing up as a low-but-unusual rate of 503s

18:42 - 503 rate begins climbing significantly, reaching ~5% of all cache_text request rate at peak (probably roughly all of the misses and passes (e.g. logged-in traffic), only cache hits being served). [Grafana](#) [🔗](#)

18:47 - Disabled fpc2-fpc4 link

18:47 - First Icinga recoveries

18:50 - 503 burst that began at 18:42 comes back to normal near-zero rate.

19:18 - [eqiad front edge depooled in DNS](#) [🔗](#), to stabilize and reduce risk during follow-on investigations fixups (takes 10 minutes for DNS TTLs to expire as this comes into effect)

Conclusions

- Virtual Chassis are black boxes, which makes it more difficult to investigate issues
- Topology changes included cable move, which makes a rollback more difficult
- Our current topologies are unsupported, this outage revealed that any changes, even though toward a more supported configuration can have bad consequences.
- Logging work done in SAL could have reduced the response time
- This event caused a driver issue on new cp1* servers, causing their link to be up on the switch side, but down on the server side

Actionables

- Status: ■ **In progress** - Fix asw2-a-eqiad topology [phab:T201145](#)
- Status: ■ **TODO** - Repool eqiad front edge traffic once eqiad is stable

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](#)

This page was last edited on 19 October 2018, at 12:14.

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](#)

