

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

Page Discussion

Read View source

View history

Search Wikitech

Q

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/20181015-eqsin-power

< Incident documentation

Contents [hide]

- 1 Summary
- 2 Timeline
- 3 Conclusions
- 4 Actionables

Summary

The eqsin PoP went offline for approximately 2h10, between approximately 14:15 and 16:25 UTC on 2018-10-14. Of those, 11 minutes were service impacting for the PoP's service region (APAC) for the majority of users, with the long tail of non-conforming resolvers being unable to reach our sites for longer periods of time.

Root cause was a combination of scheduled power maintenance by the facility and lapses in our processes.

Timeline

- Sep 07 08:17 Facility notifies us of a "SERVICE IMPACTING Scheduled Customer Outage Level" with power feed A window.
- Sep 07 08:20 Facility notifies us of a "SERVICE IMPACTING Scheduled Customer Outage Level" with power feed B window.
- Sep 10 10:06 Facility notifies us of power feed B window being rescheduled.
- Oct 12 13:01 Facility sends a "maintenance listed will commence in one hour" notification for power feed A.
- Oct 12 14:16 Power feed A is lost in egsin, and is noticed by monitoring.
- Oct 13 00:00 Maintenance window is over, facility sends a "completed" notice.
- Oct 13 03:07 Power feed A is (seemingly) back up, so the site is repooled ☑. Unknown at the time was that multiple alerts were still outstanding at that point, and power had not been recovered in the top-half of rack 0603, which included all networking equipment.
- Oct 14 13:00 Facility sends a "maintenance listed will commence in one hour" notification for power feed B.
- Oct 14 14:15 Power feed B is lost in eqsin, and with it the PoP goes dark. Users in APAC are unable to access our sites. Monitoring notices and pages, multiple respondents react.
- Oct 14 14:16 Site is depooled ₽.
- Oct 14 14:26 DNS TTLs expire for the majority of recursors, site back up for most of APAC users via ulsfo.
- Oct 14 14:38 Facility's service deks is being contacted to report a double power failure.
- Oct 14 16:02 Service desk responds with a work order; will "be contacted by a technician in 30 to 45 minutes".
- Oct 14 16:25 Site is back up, facility technician informs us that they reset a feed A in-rack breaker that was tripped.
- Oct 14 19:24 Power feed B is restored
- Oct 15 00:00 Maintenance window is over, facility sends a "completed" notice.

Conclusions

- Manual triaging of maintenance alerts is error-prone.
- Repool should always be preceded by a monitoring health-check where all alerts have been cleared or can be reasonably explained.
- eqsin does not have router or rack redundancy, so a power feed + PDU or PSU failure can take the site down.
- eqsin's PDUs are not monitored by neither the facility nor us.

Actionables

NOTE: Please add the #wikimedia-incident

Phabricator project to these follow-up tasks and move them to the "follow-up/actionable" column.

Category: Incident documentation

This page was last edited on 15 October 2018, at 15:21.

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Use for

Privacy policy About Wikitech

Disclaimers Code of Conduct Developers Statistics Cookie statement Mobile view





