

Main page
Recent changes
Server admin log (Prod)
Server admin log
(RelEng)
Deployments
SRE/Operations Help

Cloud VPS & Toolforge

Cloud VPS documentation

Incident status

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/20190611-maps

< Incident documentation

document status: in-review

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

While https://phabricator.wikimedia.org/T224395 ☑ is on which led to depooling codfw. Codfw was pooled back and only two nodes were available (see Incident documentation/20190603-maps). It was necessary to reimage all codfw servers to pickup the new partitioning scheme. And when maps2003 was reimaged which was the last to reimage, kartotherian codfw which is pooled had no node to route traffic to hence request timeouts and SREs were paged. Icinga started alerting from 8:21am

Impact

Users who were currently served from codfw were affect. They were affected for about 6min(from logstash)

Detection

lcinga alerts for kartotherian endpoints. Also from pybal, no nodes were pooled at codfw due to

Timeline

All times in UTC.

- 8:18 Reimaging of maps2003 started. OUTAGE BEGINS
- 8:21 Icinga alerts on IRC and pages were sent out
- 8:26 maps200[124] was pooled OUTAGE ENDS

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?

· Outage was at least detected quickly

What went poorly?

Where did we get lucky?

• It was easy to link outage to the ongoing reimage of maps2003

Links to relevant documentation

We should streamline processes for maps to reduce outages. Maybe via a cookbook to ensure we capture all activities and steps before performing any activity on maps.

Actionables

Explicit next steps to prevent this from happening again as much as possible, with Phabricator tasks linked for every step.

- Streamline all maps activities via a script(cookbook) to make sure we don't miss anything?
- Always check all asumptions (never trust any)

Categories: Incident documentation in-reviews | Incident documentation | Maps outages, 2019

This page was last edited on 12 February 2020, at 20:17.

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

Privacy policy About

Disclaimers Code of Conduct Developers Statistics Cookie statement Mobile view

[Powered By MediaWiki

Wikitech

