



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/20191204-MediaWiki

[< Incident documentation](#)

document status: draft

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?](#)
 - 3.4 [How many people were involved in the remediation?](#)
- 4 [Links to relevant documentation](#)
- 5 [Actionables](#)

Summary

*This is a short (<= 1 paragraph) of what happened. While keeping it short, try to avoid assuming deep knowledge of the systems involved, and also try to differentiate between proximate causes and root causes. **Please ensure to remove private information.***

Roll out of wmf.8 to group1 broke the world.

Impact

Who was affected and how? For user-facing outages, estimate: How many queries were lost? How many users were affected, and which populations (editors only? all readers? just particular geographies?) were affected? etc. Be as specific as you can, and do not assume the reader already has a good idea of what your service is and who uses it.

Detection

Initial indicators of the issue were picked up in logstash and via logspam-watch on mwlog1001. A large number of Icinga alerts followed.

It seems likely that the primary issue was obscured during the initial deploy by a focus on Parsoid errors.

If automated, please add the relevant alerts to this section. Did the appropriate alert(s) fire? Was the alert volume manageable? Did they point to the problem with as much accuracy as possible?

Timeline

This is a step by step outline of what happened to cause the incident and how it was remedied. Include the lead-up to the incident, as well as any epilogue, and clearly indicate when the user-visible outage began and ended.

All times in UTC.

- 20:12 brennen: Train wmf.8 roll fowards from group0 to group1 as well (try 1) [\[1\]](#)
- 20:12 Large amounts of logspam noticed, especially from Parsoid/PHP, and Icinga issues many alerts.
- 20:28 brennen: Train wmf.8 rolled back to just group0 [\[2\]](#)

[Fixes to exclude Parsoid/PHP]

- 23:30 brennen: Train wmf.8 roll fowards from group0 to group1 as well (try 2) [\[3\]](#)
- 23:30 **OUTAGE BEGINS**
- 23:30 Large spike in database errors in logstash ([T239877](#)), shortly thereafter large amounts of Icinga

[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](#)

alerts go off.

- 23:30+ Production group1 **and** group2 wikis become noticeably sluggish, eventually stopping working entirely.
- 23:35 brennen: Attempted train wmf.8 roll back thwarted by canary failures [4]
- 23:38 brennen: Train wmf.8 rolled back to just group0, again [5]
- 23:38 **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?

- for example: automated monitoring detected the incident, outage was root-caused quickly, etc

What went poorly?

- for example: documentation on the affected service was unhelpful, communication difficulties, etc

Where did we get lucky?

- for example: user's error report was exceptionally detailed, incident occurred when the most people were online to assist, etc

How many people were involved in the remediation?

- for example: 2 SREs and 1 software engineer troubleshooting the issue plus 1 incident commander

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.

- [Heterogeneous_deployment/Train_deploys#Rollback](#)

Actionables

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

NOTE: Please add the [#wikimedia-incident](#) Phabricator project to these follow-up tasks and move them to the "follow-up/actionable" column.

- To do #1 (TODO: Create task)
- To do #2 (TODO: Create task)
- ...

Categories: [Incident documentation](#) | [Incident documentation drafts](#)

This page was last edited on 28 April 2020, at 19:54.

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

