



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/20190506-zuul

< Incident documentation

## Contents [hide]

- 1 Summary
  - 1.1 Impact
  - 1.2 Detection
- 2 Timeline
  - 2.1 Incident 1
  - 2.2 Incident 2
- 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

CI was unresponsive and was not voting on Gerrit patches. It then had a very slow response time in processing changes. Those are two distinct events having different root causes, but in the end cause a bad user experience.

## Impact

The incident affected Zuul, the CI system. It affected any CI user, mainly SREs, developers and volunteers which were not getting CI jobs running for their patches or, later with a very large delay.

## Detection

An Icinga error got reported but not immediately taken in account by releng

SRE reported errors on #-operations, #-releng, #\_security and directly to Antoine "hashar" Musso

## Timeline

All times in UTC.

### Incident 1

- 09:26 Antoine deploys Zuul config change <https://phabricator.wikimedia.org/T105474>
- 10:00 CI reported processing times skyrocket for many repositories.
- 10:33: second to last operations/puppet patch processed
- 10:35 Giuseppe (and then Arturo) notice CI is down. Attempts to reach anyone in release engineering in IRC fail.
- 11:18: final operations/puppet patch processed before fix
- 11:30(approx): People complain about CI not working in `#wikimedia-operations`
- 11:51: Seeing how everyone's work is blocked, Giuseppe opens <https://phabricator.wikimedia.org/T222605> and sets its severity to Unbreak now! Further attempts to reach to anyone in `#wikimedia-releng` by several people fail.
- 12:11: first automated alert about the issue: `PROBLEM - Work requests waiting in Zuul Gearman server on contint1001 is CRITICAL: CRITICAL: 57.14% of data above the critical threshold [140.0]` <https://grafana.wikimedia.org/dashboard/db/zuul-gearman?panelId=10&fullscreen&orgId=1>
- 12:56: mark attempts to reach Antoine on IRC (but he misses it due to other notifications)
- 14:00 roughly: Antoine notices the problem.
- "14:04" reverted zuul patch <https://phabricator.wikimedia.org/T222605> that broke the gate and submit queue
- 14:08: operations/puppet patches begin to be processed again

### Incident 2

- 14:18 Long chain of patches is send to <https://gerrit.wikimedia.org/r/#/c/mediawiki/tools/phan/SecurityCheckPlugin/>
- 14:18: Zuul queue size explodes. They are merger:merge jobs being a cartesian products of all changes send to the above repo or 3000 merges
- 14:30: despite queue size of 3000+, only 2 jobs running at any time between approx 14:30 and 15:00
- 14:57: number of running jobs returns to 20+
- "16:00" CI still unresponsive slow due to job queue <https://grafana.wikimedia.org/d/000000322/zuul-gearman?from=1557144860227&to=1557155660227&orgId=1>
- 15:30 Zuul gearman server has rougly 2000 merger:merge jobs enqueued
- 16:08 Antoine makes `/srv/zuul/git/mediawiki/tools/phan/SecurityCheckPhanPlugin` unreadable on contint1001 and contint2001 to have zuul-merger fail quickly when attempting to do the merger:merge function

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

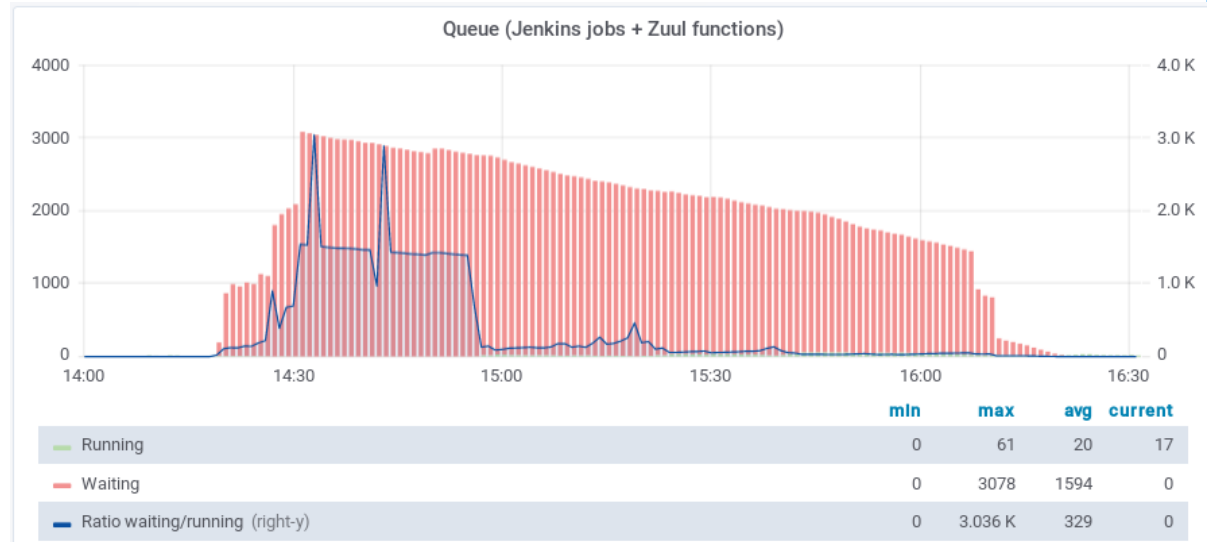
What links here  
Related changes  
Special pages  
Permanent link  
Page information  
Cite this page

Print/export

Create a book  
Download as PDF  
Printable version

- 16:12 Queue drained

View of the Gearman queue from 14:00 UTC to 16:30 UTC. The peak is reached at 14:30 UTC and is slowly drained by the two zuul-merger processes running on contint1001 and contint2001. Around 16:10 UTC, the quick drain is due to purposely causing those jobs to fail fast.



## Conclusions

### What went well?

- In both cases the root cause has been fairly easily to pin point. Respectively a config change and a huge amount of merger:merge function enqueued.

### What went poorly?

- it was difficult to get in touch with those who could fix the issue: IRC pings and an UBN ticket went unseen. (perhaps a phone call would have been better?)
- there is no shared knowledge about the CI stack beside a few people. During European time, its only covered by Antoine.

### Where did we get lucky?

- Root causes were "easy"

## Links to relevant documentation

- The documentation is at [https://www.mediawiki.org/wiki/Continuous\\_integration/Zuul](https://www.mediawiki.org/wiki/Continuous_integration/Zuul) but does not convey any helpful hint regarding this incidents.

## Actionables

- review contact list for release engineering systems (Zuul/Jenkins/Gerrit/Phabricator) with identified persons, chain of escalation etc. List them on the documentation page.
  - TODO: Create a task for this!
  - A specific set of SLOs around what is escalation-worthy and how to escalate to releng might also be helpful.
- [phab:T158054](#) - create and end-to-end job that measures the average time on queue for CI or create an alert that if average time is > X.
- [phab:T222645](#) - Get a couple more zuul-merger process to help drain merge requests T222645
- [phab:T140297](#) - Review upstream patch <https://review.opendev.org/#/c/643703/>
- **Done** Document the workaround to drain the merger queue quickly (make the git dir unwritable on the zuul-mergers)
  - [https://www.mediawiki.org/wiki/Continuous\\_integration/Zuul#Very\\_high\\_queue\\_of\\_merger:merge\\_functions](https://www.mediawiki.org/wiki/Continuous_integration/Zuul#Very_high_queue_of_merger:merge_functions)
- **Done** [phab:T105474](#) - find a proper solution to reject CR+2 patches from test pipelines **and have it properly tested**
  - Probably due to lack of <https://review.opendev.org/#/c/589762/>

Category: Incident documentation

This page was last edited on 13 May 2019, at 10:47.

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