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

# Incident documentation/20191211-MachineVision+cpjobqueue

< Incident documentation

**document status**: final

Contents [hide]



## Summary

Between 2019-12-11 and 2019-12-17, the job queue was blocked by image annotation request jobs that were being enqueued by the MachineVision extension. These jobs were using the release timestamp feature of the job specification interface that is not well supported by the Kafka job queue; release timestamps are implemented as blocking waits. These waits ended up blocking and causing a sizable backlog of a variety of jobs that are in the main pool of jobs not handled in job-specific Kafka topics. Jobs continued to be processed, but very slowly and with severe delays. No jobs appear to have been lost.

Phabricator task: https://phabricator.wikimedia.org/T240518

### Impact

This delayed the execution of a wide variety of tasks that rely on the job queue, including but not limited to:

- Global renames
- Deleting translatable pages
- Echo notifications (https://phabricator.wikimedia.org/T240800)
- File uploads (https://phabricator.wikimedia.org/T240698)
- Recent changes processing
- MassMessages
- Pageview data publication (https://phabricator.wikimedia.org/T240803)

### Detection

The issue was first reported by user 1997kB in https://phabricator.wikimedia.org/T240518. That issue specifically concerned delayed global renames. Several other delays in specific wiki functionality were reported over the next few days. The issue was not detected by any automated alerts.

## 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.**

- 2019-12-11 20:39: **OUTAGE BEGINS:** The group restriction for MachineVision functionality is lifted, enabling it for all users. fetchGoogleCloudVisionAnnotations jobs are enqueued, with 48h delay implemented via release timestamp, for most bitmap images newly uploaded to Commons.
- 2019-12-12 00:05: User 1997kB reports delayed global rename execution in https://phabricator.wikimedia.org/T240518
- 2019-12-12 through 2019-12-15: Additional reports of functionality blocked by job queue delays
- 2019-12-16 06:24: Giuseppe identifies fetchGoogleCloudVisionAnnotations jobs failing with 500s
- 2019-12-16 16:12: Holger alerts Michael H. to a possible problem with fetchGoogleCloudVisionAnnotations jobs
- 2019-12-16 17:46: Michael H. disables new fetchGoogleCloudVisionAnnotation jobs from being enqueued
- 2019-12-16 19:17: Marko disables the job queue from consuming jobs from the fetchGoogleCloudVisionAnnotations topic
- 2019-12-16 19:31 & 19:54: Marko temporarily increases the job queue processing concurrency level to help process the backlog
- 2019-12-17 01:25: **OUTAGE ENDS:** All jobs have recovered to their baseline execution times
- 2019-12-18: Petr identifies the specific issue with Kafka job queue release timestamp support

## 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?

- 1997kB noticed the increasing job queue backlog only a few hours after it started, and included it in the Phab task regarding global rename processing.

### What went poorly?

- The root cause of the delay went unidentified for nearly a week before resolution.

### Where did we get lucky?

- Marko was around to assist with remediation while Petr was out.

### How many people were involved in the remediation?

At least the following individuals were involved:

- 2 SREs (Giuseppe, Jaime)
- 4 software engineers (Marko, Holger, Petr, Michael)
- Martin Urbanec and several other volunteers

## Links to relevant documentation

- Kafka Job Queue

## 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.

- Add alert(s) for unusual job processing backlog increases ( phab:T242721)
- Document the danger of the release timestamp feature in code and on-wiki ( phab:T242722)
- Kafka job queue should improve its handling of unknown new jobs ( phab:T242726)

Category: Incident documentation