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

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/20190228-eventbus extension and logstash

< Incident documentation

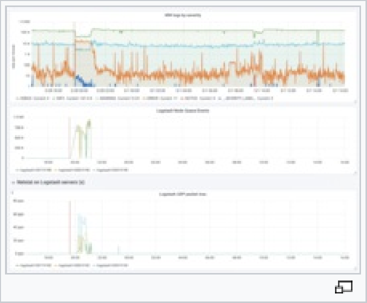
Contents [hide]

Summary

An invalid mediawiki-config deployment caused EventBus extension to be misconfigured. This resulted in a spike in mediawiki errors which caused logstash instances to become overloaded and stop reliably processing logs until logstash persisted queueing was disabled.

See also: https://phabricator.wikimedia.org/T217385

Timeline



- 19:40: https://gerrit.wikimedia.org/r/#/c/operations/mediawiki-config/+/492770/ was deployed as part of Morning SWAT. It was tested on mwdebug1002 before being deployed to the rest of fleet, but the misconfiguration would only manifest on group2 wikis.
- 19:40 group2 wikis can no longer send events to eventlogging-service-eventbus.
- 19:40: mediawiki log rate increases significantly
- 19:40: logstash queued events begin to increase
- 20:10: logstash udp loss begins to rise
- 20:10: Mediawiki .wmf19 is deployed to group2 wikis. EventBus extension configuration now works everywhere.
- 20:15: logstash udp loss alerts fire
- 20:20: troubleshooting -- all associated services in stack appear healthy despite no logs being processed. logstash instances are logging normally with no errors, elasticsearch cluster health green, kafka cluster is in sync and healthy. logstash restarts do not improve the situation
- ~20:30: Hashar notices that error logs are related to EventBus and notifies Ottomata. Hashar and Ottomata revert https://gerrit.wikimedia.org/r/#/c/operations/mediawiki-config/+/492770/ just in case.
- 20:47: logstash persistent queue files are cleared and logstash instances restarted
- 20:47: logstash queued events begin to increase again
- 20:53: logstash udp loss alerts clear
- 21:07: logstash udp loss alerts fire again
- 21:09: logstash persisted queue disabled, instances restarted
- 21:13: logstash udp loss alerts clear
- 21:37: kafka-logging consumers back in sync
- 21:37: end of incident
- 22:29: Ottomata realizes that the failed event data is in the udp_localhost-err topic in the Kafka logging cluster. He consumes the events and replays them to eventlogging-service-eventbus.

Conclusions

- Logstash is not able to cope with a sudden significant increase in log volume, in this case mediawiki errors.
- The logstash persistent queue mechanism becomes unstable under significant load, resulting in broken

logstash instances that remain in a broken state across restarts.

Actionables

- Disable logstash persisted queues (done)
- Improve visibility of logging pipeline health by creating grafana dashboard with closely related metrics side by side (namely kafka consumer lag, mw log volume, logstash loss/queues) (done)
- Consider using Kafka logstash consumer throttling with Kafka quotas: https://kafka.apache.org/documentation/#design_quotas

Category: Incident documentation