



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/20171018-wdqs

< [Incident documentation](#)

Contents [hide]

- [1 Summary](#)
- [2 Timeline](#)
- [3 Conclusions](#)
- [4 Actionables](#)

Summary

Unusual load on Wikidata Query Service caused wdqs1003 to stop responding. It was not depooled automatically and had to be restarted manually to fix the issue.

Timeline

- 2017-10-17 21:00 UTC: increased occurrences of queries being throttled on WDQS, all seem to come from the same client: <https://logstash.wikimedia.org/goto/4385319887463b960ca6aeabcb75aba9>
- 2017-10-18 06:04 UTC: icinga alert are raised for wdqs1003, user experience errors, with OutOfMemoryError messages
- 2017-10-18 06:30 UTC: investigation in the logs on wdqs1003 confirm the OutOfMemoryError
- 2017-10-18 06:48 UTC: wdqs1003 is restarted, service is restored

Conclusions

Lot of things went wrong here. In no particular order:

- Throttling should have prevented this higher than usual activity. The throttling we apply at this point only start to do accounting when a user runs a request which takes more than 10 seconds to complete. This matches the previous abuse we have seen, where a single user was running highly expensive queries. This is not aggressive enough for this case, where a client runs many cheap requests in parallel.
- wdqs1003 should have been depooled automatically when failing (the check done by pybal is only checking the nginx in front of blazegraph, not blazegraph itself)
- The JVM should have restarted automatically on OutOfMemory error, but OutOfMemoryError was happening in the creation of native thread, which is a known JVM bug: <https://bugs.openjdk.java.net/browse/JDK-8155004>
- Blazegraph is [catching](#) [OutOfMemoryError](#) in multiple places, which while allowed is usually [not a good idea](#) (note that this probably does not actually affect the outcome of this particular issue)
- As a result of WDQS instability, Recommendation API has also been throwing alerts [bug T178445](#). While it is expected for Recommendation API to fail when one of its dependencies fail, it shows the fragility of the system. It also exposed that the retry mechanism of those checks were putting more load on the system at a time where it was already in a bad state.

Side note: WDQS is by design a fragile service. We allow arbitrary users to run arbitrary SPARQL queries. This is similar to allow any user direct access to our production MySQL databases. This is fine as long as the expectations are managed. It is definitely wrong when services from which we expect stable response time and availability start depending on it. Addressing this is well out of the scope of an incident report, but at some point, splitting WDQS in a public, do whatever you want service and a different more controlled service will make sense. This would be similar to what we do with MySQL, splitting it between a production service and labs replicas, with different expectations.

Side note: We are currently experimenting and tuning GC on WDQS ([bug T175919](#)). While this is necessary, this would probably not have helped in this case, where load just climbed to saturate the service.

[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](#)

Actionables

- fix the pybal check for WDQS [384938](#)
- tune WDQS throttling to be a bit more aggressive [bug T178491](#)
- start the discussion about splitting / duplicating the WDQS clusters [bug T178492](#)
- try to contact the offending client and help him comply with our throttling requirement

Category: [Incident documentation](#)

This page was last edited on 18 October 2017, at 13:22.

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](#)

