



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/20190613-wdqs

[< Incident documentation](#)

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?](#)
- 4 [Actionables](#)

Summary

From June 13 ~15:10UTC to ~15:50 UTC [🔗](#) the public WDQS endpoint in eqiad was overloaded by a bot to the point where it was not serving user queries. There is no reason to think that this bot was malicious. To mitigate this, the python-requests user agent is temporarily banned from accessing WDQS, consistent with our [user agent policy](#) [🔗](#).

Impact

The WDQS public endpoint in eqiad was unavailable [from ~15:25 to ~15:45 UTC](#) [🔗](#).

The python-requests user agent is still being banned, we are waiting to implement a more gentle solution before removing this ban.

The internal WDQS endpoint was not impacted.

Detection

Problem was detected by the Icinga LVS probe.

Timeline

All times in UTC.

- 15:10: load starts to increase on the public wdqs eqiad cluster
- 15:31: Icinga LVS alert for wdqs.svc.eqiad.wmnet

Conclusions

- identifying and throttling bots is a hard problem
- we need to take more drastic action to protect the stability of the service (aggressively throttle generic user agents)

What went well?

- problem was detected automatically in a timely manner
- good collaboration and clear communication between

What went poorly?

- while we do have logic to throttle abusive bots, this throttling was not sufficient to protect the service
- we are still banning python-requests as a user agent, which affects a number of bots

Where did we get lucky?

[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](#)

- This happened during SRE offsite, when most SRE are in the same timezone. Luckily this wasn't when all of them were sleeping!

Actionables

- We know that our throttling of bot is far from perfect (this is a hard problem). Some idea are already being discussed on [phab:T219477](#)
- ban logs are too verbose on disk (mitigated by <https://gerrit.wikimedia.org/r/c/operations/puppet/+516837>)
- our user agent policy discourages the use of generic user agents, we should start to be more aggressive in throttling generic UA or ban requests with no UA (<https://gerrit.wikimedia.org/r/c/wikidata/query/rdf/+516834>)
- our own clients should follow our policy, for example our icinga checks should have a meaningful UA (<https://gerrit.wikimedia.org/r/c/operations/puppet/+517032>)
- remove the ban on the python-requests UA once we have a more gently throttling solution in place

Category: [Incident documentation](#)

This page was last edited on 21 June 2019, at 19:26.

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](#)

