



[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](#)

[Create account](#) [Log in](#)

Page [Discussion](#)

[Read](#)

[View source](#)

[View history](#)



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/20200207-mediawiki API down

[< Incident documentation](#)

document status: in-review

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?](#)
 - [3.4 How many people were involved in the remediation?](#)
- [4 Links to relevant documentation](#)
- [5 Actionables](#)

Summary

A bot scraping zhwiki, which we have been monitoring for a while now, started making more expensive requests more aggressively. The bot was concealing itself by using a common User-Agent.

Most requests were similar to:

```
http://zh.wikipedia.org/w/api.php?
action=parse&pageid=2996886&prop=text&wrapoutputclass=wiki-
article&disableeditsection=true&mobileformat=true&mainpage=true&format=json
```

The `wrapoutputclass` url parameter causes a request to bypass [parsercache](#). To make matters worse, the scraper was going through the whole list of French localities on zhwiki, each of which made ample use of some known slow templates, originally seen on occitan wikipedia (euwiki), with the 36k entry table of localities. Each of those requests required 15-60 seconds to parse.

Lastly, while we were investigating, an unscheduled deployment was pushed to production, to fix an [UNB!](#) task. The deployment caused s8 to recive an influx of queries, so it was quickly reverted

[Incident_documentation/20200207-wikidata](#).

Impact

API became almost unresponsive for about 10 minutes and. Application servers were unresponsive for another 10 minutes a little bit after.

Detection

```
14:06:40 <+icinga-wm> PROBLEM - MediaWiki exceptions and fatals per minute on
icinga1001
14:06:41 <+icinga-wm> PROBLEM - High average POST latency for mw requests on
api_appserver in eqiad on icinga1001 is CRITICAL:

14:09:42 <+icinga-wm> PROBLEM - High average GET latency for mw requests on
api_appserver in eqiad on icinga1001 is CRITICAL:
14:09:52 <+icinga-wm> PROBLEM - Apache HTTP on mw1290 is CRITICAL: CRITICAL

14:17:07 <+icinga-wm> RECOVERY - Nginx local proxy to apache on mw1283 is OK:
HTTP OK:
```

Timeline

All times in UTC.

- 14:06 **OUTAGE #1 BEGINS**

We start parsing API logs, where we establish that it the zhwiki bot we have been monitoring, is making very expensive requests. The requests were both bypassing parsercache and included some infamous templates. It is using a very common UA, one that is used by real users as well, so blocking would be not be easy.

- 14:17 **OUTAGE #1 ENDS**
- 14:28 Amir contacts the community [A technical issue with articles of French communes](#)
- 14:46 Amit Emptified the templates [https://zh.wikipedia.org/wiki/Special:用户贡献/Amir_Sarabadani_\(WMDE\)](https://zh.wikipedia.org/wiki/Special:用户贡献/Amir_Sarabadani_(WMDE))



Conclusions

Templates issues are hard to debug.

What went well?

We already were aware of the bot being active in zhwiki as well as its activity. It was the first thing we looked, and it easily stood out in the logs.

What went poorly?

It is hard to pinpoint when an issue is due to a template as well as which template it is. Also, this bot was using a common UA, making it a bit complicated for us to simply block it.

Where did we get lucky?

We had similar issues with euwiki with the same templates, so they were on our radar. We were also lucky that the bot slowed down rather quickly. Also that Amir was online and knew what to do.

How many people were involved in the remediation?

- 4 SREs + 2 software engineers

Links to relevant documentation

None

Actionables

- Emptyify the French Commune Data templates and contact the community (already done)

Categories: [Incident documentation in-reviews](#) | [Incident documentation](#)

This page was last edited on 25 March 2020, at 19:12.

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](#)

