



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/20111127-Reddit

[< Incident documentation](#)

On 27th November 2011 at about 21:48 UTC, Nagios started reporting load problem with db1025, a MySQL slave server used by the Fundraising team for reporting purposes. About 7 minutes later, db1025 and the Apache servers started spilling 'socket timeout' error messages. At about the same time, several Technical Ops engineers were paged (by Nagios, Watchmouse, and Reedy). Chats on IRC Operations and Tech channels started reporting site issues.

Mark and Apergos started investigating into it. Subsequently other members of Engineering and TechOps joined in the triage.

A summary of the sequence of events, impact, root cause analysis and recommendations are noted below.

Contents [\[hide\]](#)

- [1 Sequence of Events](#)
- [2 Impact](#)
- [3 Root Cause Analysis](#)
- [4 Recommendations/Lessons Learned](#)

Sequence of Events

UTC time

- 21:48 nagios-wm: PROBLEM - MySQL slave status on db1025 is CRITICAL:
- 21:55 nagios-wm: PROBLEM - DPKG on db1025 is CRITICAL: CHECK_NRPE: Socket timeout after 10 seconds./db/DatabaseMysql.php 'Database::getLag()' fix from [r90644](#)'
- 21:55 nagios-wm: PROBLEM - Apache HTTP on mw44 is CRITICAL: CRITICAL - Socket timeout after 10 seconds
- 21:58 Reedy contacted Apergos and paged Mark

The next 15 or so minutes, nagios continued to report problems with the Apache servers with some intermittent Apache recovery messages.

- 22.00 - 22.25 : Mark, Apergos and Jeff worked on moving db1025 application dependencies to db1008 (master). Mark then brought down db1025, which was showing extremely high load.
- 22.26 : Apache servers started to recover. The initial diagnosis was that db1025 experienced a hardware failure and that failure caused the mediawiki Apache servers to be overwhelmed by requests waiting for that database to recover , thus bringing down the application servers.

There was then a period of calm again. Mark, Jeff and Apergos restarted db1025 and pointed the fundraising reporting database (which was now redirected to db1008) to storage3, another MySQL slave of db1008. Jeff investigated and found evidence of a failed disk on db1025.

- 22:55: Nagios started reporting load problem again on Storage3 db server and some of the Apache servers started having socket timeout problems.
- 22.56 - 23:08 : Storage3 was indeed overloaded and Apergos reverted the application to read from db1008 instead of Storage3. the Apache servers recovered. Domas reported the problem could be with fundraiser stats query.
- 22.10: Mark found lots of queries on SpecialFundraiserStatistics in the db slow log. Investigating into the squid logfiles, Binasher reported the fundraiser stats page (SpecialFundraiserStatistics) got linked to on reddit. Awjr then disabled the fundraiser stat page. The site has been stable since.

Impact

During the incidents, users who were editing or logged-in, experienced the outage. Mobile users were impacted

[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](#)

as well. However anonymous web surfers to WMF sites were mostly not impacted because the squid servers were functioning unless the requested contents were not cached. Watchmouse (status.wikimedia.org) reports seem to confirm that. The total outage lasted about 46 minutes (31 minutes initially, followed by 15 minutes).

Root Cause Analysis

The team noticed the following:

- The ContributionReporting extension could be optimized further. Feedback and recommendations have been sent to the fundraiser engineering team.
- it appeared that the surged traffic came from the reddit community, after a link to the fundraising stats page was posted under Brandon's IAMA on reddit. Since that page wasn't prepared to handle a cache stampede, the surge was the proverbial straw that broke the Apache servers.
- While the ContributionReporting extension made extensive use of memcached, (cache was set for 15 minutes), when the cache expired, the re-caching took a long time and because that page was getting a lot of traffic, the site experienced a 'cache stampede' aka Michael Jackson effect.
- Technically, the DB server (db1025 and later storage3) was brought to its knees by the bad queries, and all apaches were not overloaded, but simply got stuck with all requests waiting for the results of that particular db server.

Recommendations/Lessons Learned

The fundraiser engineering team will be investigating and implementing improvements to the ContributionReporting extension to prevent future [cache stampedes](#). Discussion about possible remedies can be found on [Bugzilla](#) and in various message threads on wikitech-l.

Categories: [Incident documentation](#) | [Fundraising](#)

This page was last edited on 18 May 2018, at 04:38.

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. See [Terms of Use](#) for details.

[Privacy policy](#) [About](#)
[Wikitech](#)

[Disclaimers](#) [Code of Conduct](#) [Developers](#) [Statistics](#) [Cookie statement](#) [Mobile view](#)

