



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/20130508-memcached

[< Incident documentation](#)

Problem

copy/pasted from Mark Bergsma's email

We had an outage today, due to a SFP+ module for a single memcached server (mc1009) going bad in a switch, causing the switch to partially reboot, and causing all application servers to get blocked due to a single down memcached. As usual this caused all kinds of cascading failures on other clusters such as Squid/Varnish. When not overloaded, these clusters would only serve cached pages at that point.

At 13:30 UTC, the EEPROM of the SFP+ module for mc1009 on asw2-a5-eqiad became unreadable over the i2c bus, and apparently this caused the switch to reboot its entire ethernet chassis subsystem. This gave a lot of noise, with the entire rack going offline, and lots of monitoring and log messages. Fortunately, the other servers came back up after a minute or two. I noticed the switch problems after a few minutes, but it took me a while to check and wade through everything and determine that in the end only a single memcached server was still missing. At 13:58 I removed mc1009 from the MediaWiki memcached pool, and two minutes later, all services were back up.

Obviously having all Memcached servers still on a single switch in a single rack is very bad, and I've just restarted our old plans to move half of them (as well as half of Varnish) to another rack in row C.

Also very bad is that the failure of a single memcached server can still knock us over completely. Can we not improve on this, with the new PECL client? Or what's the status of our plans with twemproxy/nutcracker?

Solution

copy/pasted from Asher Feldman's email

As of last night, we've been routing all memcached requests in eqiad via twemproxy. A short while ago, I rebooted mc1009 without any other preparation. The main impact was just log noise (libmemcached can be misleading when logging "SERVER HAS FAILED AND IS DISABLED UNTIL TIMED") but as far as I've been able to tell, there was no impact to site availability or general performance. Graphite showed an increase in external store selects as I'd expect but php request latency (at the average, 50th, and 90th percentiles) was unimpacted, as were memcached get times. Using enwiki as a logged in user seemed completely normal, and I found no user reports of issues during that time frame either. Huge difference vs. the mc1009 failure on May 6 that started this thread! It also helps shave around 12ms of setup time from every request.

The twemproxy deployment is still provisional and needs additional work (mostly around logging and stats collection) but I'm pleased so far and am going to proceed.

Startup Performance

90th percentile Setup.php-memcached time went from 14.8ms to 4ms as a result of the libmemcached upgrade alone, and to 0.9ms with twemproxy.

Category: [Incident documentation](#)

This page was last edited on 18 June 2013, at 16:31.

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

[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](#)