



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/20140228-Cirrus

[< Incident documentation](#)

Contents [\[hide\]](#)

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

Summary

From Nik Everett's email to Ops

CirrusSearch flaked out Feb 28 around 19:30 UTC and I brought it back from the dead around 21:25 UTC. During the time it was flaking out searches that used it (mediawiki.org, wikidata.org, ca.wikipedia.org, and everything in Italian) took a long, long time or failed immediately with a message about this being a temporary problem we're working on fixing.

Timeline

We added four new Elasticsearch servers on Rack D (yay) around 18:45 UTC The Elasticsearch cluster started serving simple requests very slowly around 19:30 UTC I was alerted to a search issue on IRC at 20:45 UTC I fixed the offending Elasticsearch servers around 21:25 UTC Query times recovered shortly after that

Explanation

We very carefully installed the same version of Elasticsearch and Java as we use on the other machines then used puppet to configure the Elasticsearch machines to join the cluster. It looks like they only picked up half the configuration provided by puppet (/etc/elasticsearch/elasticsearch.yml but not /etc/defaults/elasticsearch). Unfortunately for us that is the bad half to miss because /etc/default/elasticsearch contains the JVM heap settings.

The servers came online with the default amount of heap which worked fine until Elasticsearch migrated a sufficiently large index to them. At that point the heap filled up and Java does what it does in that case and spun forever trying to free garbage. It pretty much pegged one CPU and rendered the entire application unresponsive. Unfortunately (again) pegging one CPU isn't that weird for Elasticsearch. It'll do that when it is merging. The application normally stays responsive because the rest of the JVM keeps moving along. That doesn't happen when heap is full.

Knocking out one of those machines caused tons of searches to block, presumably waiting for those machine to respond. I'll have to dig around to see if I can find the timeout but we're obviously using the default which in our case is way way way to long. We then filled the pool queue and started rejecting requests to search altogether.

When I found the problem all I had to do was kill -9 the Elasticsearch servers and restart them. -9 is required because JVMs don't catch the regular signal if they are too busy garbage collecting.

Actionables

What we're doing to prevent it from happening again:

- Status: ■ **Done** - We're going to monitor the slow query log and have icinga start complaining if it grows very quickly. We normally get a couple of slow queries per day so this shouldn't be too noisy. We're going to also have to monitor error counts, especially once we get more timeouts.
 - https://bugzilla.wikimedia.org/show_bug.cgi?id=62077
- Status: ■ **Done** - We're going to sprinkle more timeouts all over the place. Certainly in Cirrus while waiting on Elasticsearch and figure out how to tell Elasticsearch what the shard timeouts should be as well.

[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](#)

- https://bugzilla.wikimedia.org/show_bug.cgi?id=62079
- Status: ■ **Done** - We're going to figure out why we only got half the settings. This is complicated because we can't let puppet restart Elasticsearch because Elasticsearch restarts must be done one node at a time.
- Was puppet error coding error. Fixed.

Category: [Incident documentation](#)

This page was last edited on 28 April 2014, at 18:20.

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](#)

