



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/20190214-labsdb1005

[< Incident documentation](#)

## Contents [\[hide\]](#)

- [1 Summary](#)
- [2 Timeline](#)
- [3 Conclusions](#)
- [4 Links to relevant documentation](#)
- [5 Actionables](#)

Tracked in [Phabricator](#)  
**Task T216208**

## Summary

The [PAWS](#) service was offline due to issues connecting to `tools.db.svc.eqiad.wmflabs`, related to the database being in the limit of usage ( [T215993: tools.db.svc.eqiad.wmflabs hitting it's limit?](#) ).

See also the prior day's incident: [20190213-PAWS](#)

In the course of this several tables crashed. All were able to be repaired except:

`s51290__dpl_p/monthly_list_part` and `s51290__dpl_p/t_s_all_dabs` . Those two were unrecoverable.

## Timeline

- 11:54 - Monitoring paged about PAWS being down (error 504 gateway timeout)
- 11:53 - Bryan, Arturo, Brooke, Giovanni started investigating
- 12:00 - Errors coming from the Jupyter Hub pod indicating operation timeout while contacting some service (unclear at this time)
- 12:01 - Connecting to the database works
- 12:04 - Chicocvenancio jointed troubleshooting
- 12:14 - Sill looking at labsdb1005 but no signs of problems (previous outage was caused by DB so focus is there)
- 12:20 - Verbosity level increased for Jupyter Hub, indicating problems connecting to the database. Database is not yet reporting 'too many connections' but also doesn't respond anymore
- 12:41 - MySQL logs show various "Got timeout reading communication packets" errors
- 12:43 - Brooke tries to get DBA help
- 12:46 - Connections to MySQL using the CLI only hang, no command prompt after authentication (even locally using the sock file directly).
- 12:55 - DBAs haven't replied yet
- 13:00 - Connections attempts are failed with 'too many connections' error
- 13:10 - mysqld thread count is at 1051 threads
- 13:14 - Brooke identifies some tools that might be using too many connections. Works starts on finding out and stopping these tools (jobs, pods, crontabs).
- 13:58 - Brooke increased `max_connections_limit` from 1024 to 1280. MySQL clients are able to connect now.
- 14:08 - Connections are still climbing
- 14:28 - Connections aren't getting released. MariaDB is restarted.
- 14:35 - Restart doesn't work, existing mysqld process doesn't die. Attempting a stop only.
- 14:39 - Connections are accepted again. There are many tables marked as crashed and in need of repair.
- 14:44 - Giovanni runs ``mysqlcheck -A --auto-repair``
- 23:39 - Jaime from the database team killed the repair as it was locking up metadata. He also killed several long-running queries.
- 23:47 - Jaime restarted the database in read only mode.
- 00:11 - Bryan confirmed read only tools can use the database at this time.

[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](#)

- 00:15 - Jaime set the database back in read/write mode after repairs finished
- 01:49 - Users report database is locking up again. Queries never return.
- 03:50 - Brooke spends a lot of time looking for a cause or a reason, but the connections keep climbing because queries aren't finishing. They are over 900. Brooke decides not to restart the db because it doesn't seem to help (just puts off the problem coming back) and in the locked up state it risks more data.
- 06:35 - Manuel from the DBA team limited the per-user connections to 20, so that will prevent connection limits being reached.
- 16:18 - Users are reporting the slow queries and unusable condition has returned. Connection numbers stay lower due to the per user limit.
- 18:45 - Brooke shut down replication to labsdb1004 and stopped the database in hopes of moving things over to a new system.
- 03:14 - cloudvirt1019 is in service, a VM-based database server is running and Brooke and Andrew have started a transfer of the database files from labsdb1004 to try and stand that up on it.
- 21:21 - The team got the new database server running as an up-to-date replica of the failing server. There is still a couple items to transfer over to put it fully in production.

## Conclusions

- Flood of requests from Tools can overwhelm labsdb1005
- At various points in the whole episode, slow connections and client or server-side timeouts occurred. This can be caused by overwhelm by connections, but it can also cause connections to pile up if the root cause is something else. It is possible that issues within the database server or the network between it and clients remain.
- After Jaime from the database team got involved, he noticed that a user was running a CREATE DATABASE IF NOT EXISTS query on every connect (or close to it). This activity was stopped and will be discussed with the user in case that was locking up the metadata leading to the overload mentioned above.
- Risk: MariaDB on Jessie is not maintained (might be lacking new settings that prod uses)

## Links to relevant documentation

## Actionables

- [phab:T216167](#) - Verify checkwiki tool against excessive DB usage
- [phab:T216168](#) - Review labsdb1005 MariaDB configuration against prod standards (evaluate improvements).
- [phab:T216170](#) - Add per-user connection limits to toolsdb
- [phab:T193264](#) - Replace labsdb100[4567] with instances on cloudvirt1019 and cloudvirt1020
- [phab:T216202](#) - Likely unrelated, but single disk failure was also found in the RAID
- [phab:T216749](#) - Reclaim/Decommission labsdb1004.eqiad.wmnet and labsdb1005.eqiad.wmnet as soon as they are ready
- Other actionables mentioned in [Incident documentation/20190213-PAWS](#)
- [phab:T216769](#) - List old/unrebootable hardware, notify users about possible doomed services
- [phab:T216770](#) - Document a process for quick failover to the replica where that seems appropriate (with warning on why that is a semi-nuclear option here at times) when DBAs are not available.

Category: [Incident documentation](#)

This page was last edited on 21 February 2019, at 23:27.

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](#)

