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/20180404-mathoid

< Incident documentation

Summary

The Mathoid service in eqiad experienced an outage due to increased load. An operator mistake in provisioning more capacity extended the total outage time by 10 mins.

Timeline

- 22:26 UTC The SRE team gets paged for a mathoid outage.
- 22:27 UTC Pages about recovery arrive
- 22:36 UTC Paged again about a failure.
- 22:37 UTC Filippo calls Alex, who happens to be awake.
- 22:40 UTC Alex logs in, realizes that extra traffic is being directed to the mathoid service (reason unknown, probably some pages/template edited), and decides to increase the capacity allocated to the cluster
- 22:54 UTC alex issues the command to increase the number of pods from 4 to 16.
- 22:55 UTC mathoid pages again.
- ~23:00 UTC Giuseppe advises that the service IP in the kubernetes Service resource is wrong
- 23:01 UTC alex fixes the above mistake, still no luck
- 23:05 UTC Alex realizes the there is no network policy for the service applied so all traffic is denied
- 23:08 UTC Alex removes the helm release and installs a new one
- 23:09 UTC Mathoid recovers

Conclusions

An operator mistake when increasing the number of pods from 4 to 16 in the invocation caused a wrong service IP in the kubernetes service configuration. We need better tooling to avoid such things in the future. The helm release did not have a network policy applied after a pod increase. The exact cause is yet unknown, more testing is needed.

Actionables

- Create documentation for the most common operations of the Kubernetes cluster and their services, like modifying the cluster size.

Notes

Usually, most requests for Math formulae are handled by RESTBase (as most are asks for renders); Mathoid itself handles only responding to checks for erroneous formulae and renders for new formulae. Because of that, traffic to Mathoid itself is rather limited (compared to actual external traffic). However, it is not uncommon for editors to import large numbers of formulae at once (e.g. by copy/pasting math code from one wiki page to another and then modifying for language/style), causing surges in traffic to the service itself. During the move of the service from SCB to Kubernetes, however, the number of workers able to handle requests was dropped from 96 to 4, which can easily cause the service to stop accepting requests as rendering MathML, SVG and PNG versions of a formula takes a while.

Category: Incident documentation