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/20180214-labvirt1008-failure

< Incident documentation

Contents [hide]

Contents [hide]

1 Summary
2 Timeline
3 Conclusions
4 Actionables

Summary

Labvirt1008 seems to have overheated and gone down. This effects tenants as well as virtual VPS infrastructure

Timeline

- 7:20 UTC labvirt1008 rebooted
- 8:00 UTC moritzm logged a task about it
- There was no paging or alerting. This is a problem.
- 11:00 UTC Chase woke up and started investigating the extent of the outage, and looking for impact on Toolforge especially
- 12:58 UTC Chase sent an email about impact to cloud-announce https://lists.wikimedia.org/pipermail/cloud-announce/2018-February/000023.html with a list of affected instances.

Conclusions

We know that our instance storage is local and ephemeral. We should ensure that is documented for tenants in easy to find places, and re-ensure that our mechanism that keep critical redundant components spread across labvirts are working. In our world though a single hypervisor is a special snowflake and I believe we should have been paged on this outage, but seem not to have been. It was my understanding that a full instance storage partition should have paged if nothing else, and in this case the failure of that check.

Actionables

- Coordinate with DC OPS to deal with overheating phab:T187292
- Look at moving tenant instances to another labvirt (we should have a standing spare)
- Investigate what should have paged and why it did not (and fix it)

Category: Incident documentation

Privacy policy   About Wikitech   Disclaimers   Code of Conduct   Developers   Statistics   Cookie statement   Mobile view