Page   Discussion   Read   View source   View history   Search Wikitech

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/20140613-Videoscalers

< Incident documentation

Contents [hide]

Summary

Due to a change to the way the job-loop script works, that was not syncronized on the videoscalers, processing of videos stopped effectively in the morning of Friday June 13th 2014. Completing the change on the video scalers puppet manifests resolved the issue and practically restarted the jobs.

Timeline

On Friday June 13th, 10:26 UTC a series of changes to the job-loop bash script that is in charge of running our videoscalers has been merged in the repository (https://gerrit.wikimedia.org/r/#/c/139208/) As this change forgot to remove the extra unused argument from the videoscalers, processing of videos effectively stopped at that point (any change to jobloop triggers a restart of the service in puppet).

Brian Wolff was notified that some uploaded video on commons were not being converted, so he opened an RT ticket (#7693) on Saturday June 14th, 22:17 UTC. I've seen the ticket on Sunday afternoon, and upon investigation (with the notable help of Reedy), we were able to pin the problem to this change. As soon as we've merged https://gerrit.wikimedia.org/r/#/c/139682/ the video scalers instantly began converting videos again. So the incident closed at 14:26 UTC on Sunday June 15th, 2014.

Conclusions

While this is a very honest error in a merge, which could happen again and again, we've got some food for thoughts:

- Job-loop does *not* log or control execution of the jobs it launches. Worse, there is no log of the program running anywhere. The only way to understand if the jobs were running correctly was basically grepping on fluorine for output. The non-working state was completely silent
- In general jobs are poorly monitored and poorly monitorable

So while these kind of errors are pretty normal and could keep happening, we should at least be notified when we do so. This, together with the need for a better job control/logging, make job-loop the weak link here.

Actionables

- Status: ▪ **Declined** - Add monitoring for individual job types on single machines.
- Status: ▪ **on-going** - Make job-loop aware of the status of launched jobs, or rethink and rewrite it completely.
  - https://gerrit.wikimedia.org/r/#/c/144612/ is going live before HHVM (hopefully)

Category: Incident documentation

This page was last edited on 8 July 2014, at 18:26.