Facilitating HPC job debugging through job scripts archival

Andy Georges

2 February 2020

FOSDEM 2020 - HPC, Big Data & Data Science devroom

About

- I am an HPC sysadmin at Ghent University
- Only doing user support very occasionally
- When something is sent my way
- But . . . I am responsible for logging things
- And for the scheduler

Motivation

- HPC clusters run a gazillion jobs over their lifetime
- These jobs sit in the queue after submission
- For a while . . .
- Some jobs die unexpectedly
- Then the user wants to know why
- Probably to avoid it happens again
- And because it cannot be their fault, obviously

The key problem

Figure out what was running in the job under which environment

Surely we can ask the user to provide the job script

- They no longer have it
- They may have changed it (and not under version control) to be used in another job
- They may not recall which version was submitted
- They may claim to know exactly what was submitted and provide you with the wrong script
- In all of the above they would have been acting in good faith

The user is not the only actor

- The scheduler may have changed the script
- Or its settings, like the requested cores, memory, . . .
- Through a submit filter
- But . . . it does keep a copy
- Or does it?

Surely the scheduler can provide the required information when we ask it

- The script is saved
- In the spool directory
- Once the job is queued
- Until it crashes

Should we patch the scheduler?

- Yes, but no, but yes, but no, but maybe, but no
- If the scheduler is FOSS
  - Write a patch
  - To save the *exact* job script in a secondary location
  - Forget about it, to avoid deletion upon job completion
- Maintain said patch forever
  - Unless you can get it upstream
  - But why should it be accepted?
    - Saving a duplicate copy is not the scheduler's task
    - It makes for more work to be done on each job submission
  - You may need to adjust, test, . . . in the next release

Complications

- Your site may be running multiple schedulers
  - Depending on the vendor
- You may need to pay just to get a duplicate copy of the job scripts
- And other sites might too (hey, it’s free money)
- So even if your current scheduler is FOSS and got patched, the next one may be different

Takeaway

The scheduler may not be the best place to obtain job script backups

Enter SArchive

- FOSS (duh), written in Rust
- Separates the front end (finding job scripts for the scheduler) from the back end (archival of said job scripts)
- Started out as a tool for Slurm, but also supports Torque
- Should be trivial to add support for schedulers that also drop job scripts in a spool directory, e.g. Univa Grid Engine, LSF, PBS Pro, . . .

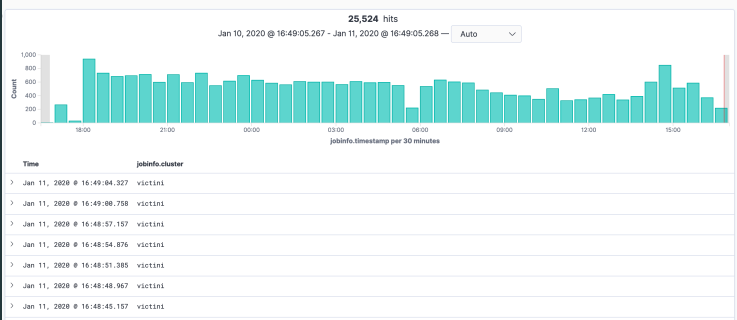
What it does

- Monitor the spool director(y)(ies)
- Upon receiving a desired change notification tell the . . .
- . . . scheduler-savvy front end code to pick up the data as it knows how to
- The resulting job information is pushed onto a FIFO queue for further processing
    - To allow fast processing of data as jobs can be entering the system suddenly in large quantities
- The back-end takes the items out of the FIFO queue and archives the information

Supported back-ends

- Saving to a file hierarchy with YYYY[MM[DD]] sub-directories
- Sending a JSON structure with the job script information to Elasticsearch
- Producing a JSON structure with the job script information to Kafka
- Note: I only implemented the features that we need/use for ES/Kafka (which is fairly limited)

24 hours of job scripts injected into ES through Kafka (6 Ghent University clusters)



25,524 hits

Jan 10, 2020 @ 16:49:05.267 - Jan 11, 2020 @ 16:49:05.268 — Auto

Time	jobinfo.cluster
Jan 11, 2020 @ 16:49:04.327	victini
Jan 11, 2020 @ 16:49:00.758	victini
Jan 11, 2020 @ 16:48:57.157	victini
Jan 11, 2020 @ 16:48:54.876	victini
Jan 11, 2020 @ 16:48:51.385	victini
Jan 11, 2020 @ 16:48:48.967	victini
Jan 11, 2020 @ 16:48:45.157	victini

Resources

- https://github.com/itkovian/sarchive
- https://crates.io/crates/sarchive (may be behind master, depends on dependencies)
- Fork it, add to it and open a PR :)
- Or open an issue if you want or need a feature