

Using Slurm Rest API with Cluster Cockpit

Katrin Nusser, NHR@FAU

Our setup

- Currently in production use on two clusters (fritz & alex) with Slurm 21.08 and 22.05
- Consists of two scripts run every few minutes by cron job
- Python script gathers Slurm job data
 - Two different endpoints of Slurm Rest API used:
 - `/slurm/v0.0.3X/jobs` → equivalent to `scontrol show job`
 - `/slurmdb/v0.0.3X/jobs` → equivalent to `sacct`
 - Mapping of core/GPU IDs in case of node sharing
 - Write payload for started and stopped jobs to file
- Separate (bash) script sends job information to CC Rest API

Pros & Cons

Solution is independent of Slurm infrastructure:

- Problems in monitoring setup have no impact on batch processing
- Can be run as unprivileged user (only Slurm operator role necessary)
- Can be run on separate server if Slurm Rest API is exposed, no direct integration in Slurm necessary
- Not sensitive to Slurm updates (older Rest API versions usually still available)
- Scripting language can be freely chosen

- No direct triggers for started or finished jobs; very short or failed jobs might not be captured
- Changes during runtime of job (e.g. walltime extension) currently not captured

Things to be aware of

- Slurm API endpoint behaves similar to `scontrol` → information on finished jobs only available for a few minutes
- For shared nodes: information on GPU/core IDs is not saved in Slurm DB
→ writing of temporary files and use of two separate scripts ensures that no job information is lost, even in case of longer unavailability of CC
- As with everything Slurm-related: largely undocumented, bugs and unexpected behavior possible