Contribution ID: **2** Type: **not specified**

Auto-scaling in the cloud: Intelligent HTCondor resource management

Tuesday 21 September 2021 15:00 (30 minutes)

HTCondor is an effective tool to rank and match execute resources against a set of jobs with explicit resource requirements. In the cloud, a subtly different challenge is presented: how to rank execute resource configurations that will be automatically created to run idle jobs (auto-scaled on-demand).

We describe recent work by the HTCondor team and Google Cloud to provide built-in support for commonly desired patterns in cloud auto-scaling. For example, a job can require co-location of execute resources with data stored as Google Cloud Storage objects. Alternatively, a group of jobs might seek to expand into as many cloud regions as possible in search of cost savings or to minimize the wall-clock time of a particular workflow.

Desired slot length

20

Speaker release

Yes

Authors: Dr THOMSON, Ross (Google); DOWNES, Tom (Google)

Presenters: Dr THOMSON, Ross (Google); DOWNES, Tom (Google)

Session Classification: Workshop session

Track Classification: HTCondor user presentations