

VCondor: a dynamic cloud scheduler with HTCondor for multi-queue in cloud based environment

As a new approach to manage computing resource, virtualization technology is more and more widely applied in the high-energy physics field. A virtual computing cluster based on Openstack was built at IHEP, using HTCondor as the job queue management system. In a traditional static cluster, a fixed number of virtual machines are pre-allocated to the job queue of different experiments. However this method cannot be well adapted to the volatility of computing resource requirements.

To solve this problem, VCondor: an elastic computing resource management system in cloud based environment has been designed. This system performs unified management of virtual computing nodes on the basis of job queue in HTCondor, and based on dual resource thresholds as well as the quota service (VMquota). A VM will be created automatically when a job is waiting to run. It will be destroyed when the job is finished and there is no more job in HTCondor queue. The job queue is checked in a period of time such as 10 minutes, so a VM will continue to run if there are new jobs in the period of time. The system is consisted of four loosely-coupled components, including job status monitoring, computing node management, load balance system and the daemon. Job status monitoring system communicates with HTCondor by command lines or APIs to get the current status of each job queue. Computing node management component communicates with Openstack to launch or destroy virtual machines. After a VM is created, it will be added to the resource pool of corresponding experiment group. Then the VM can get a job to run. After the job finishes, the virtual machine will be shutdown. When the VM shutdown in Openstack, it will be removed from the resource pool. Meanwhile, the computing node management system provides an interface to query virtual resources usage. Load balance system provides an interface to get the information of available virtual resources for each experiment from VM Quota. The VM Quota tells load balance system how many virtual machines one experiment can use and reserve them for a period of time such as 30 minutes. The daemon component asks load balance system to decide the number of available virtual resources. It also communicates with job status monitoring system to get the number of queued jobs. Finally, it calls computing node management system to launch or destroy a few of virtual computing nodes.

The practical run shows virtual computing resource dynamically expanded or shrunk while computing requirements change. Additionally, the CPU utilization ratio of computing resource was significantly increased when compared with traditional resource management. VCondor also has good performance when there are multiple condor schedulers and multiple job queues.

Primary author: Mr LI, Haibo (Chinese)

Co-authors: Mr CHENG, Yaodong (IHEP, CAS); Mr CHENG, Zhenjing (Institute of High Energy Physics, Chinese Academy of Sciences)

Presenter: Mr LI, Haibo (Chinese)

Track Classification: Infrastructure Clouds and Virtualisation