Background

At IHEP, computing resources are contributed by different experiments including BES, JUNO, DYW, HXMT, etc. The resources were divided into different partitions to satisfy the dedicated experiment data processing requirements. Utilizing Torque&Maui, IHEP had maintained a local cluster with 50 queues serving for above 10 experiments for more than 10 years.

Problems

The separated resource partitions led to imbalance resource load. As shown in the two following diagrams, BES resource partition is quite busy on some time points, even with lots of jobs in queue; Oppositely, DYW resources stay idle for a long time. However, sometimes the situations are contrary.

Basic Method

After migrating resources from Torque&Maui to HTCondor in 2016, job scheduling efficiency has been improved a lot. But resource usage can not increase more due to the separated resources among all experiments. To aim at breaking resource isolation, an efficient sharing strategy was presented to improve the overall resource usage. This strategy consists of two core components: sharing policy and Central Controller. Sharing policy dynamically defines the sharing quota for each experiment group. Central Controller manages the sharing information which is published to worker nodes automatically.

Conclusion

With sharing strategy, overall resource usage of IHEP computing cluster has dramatically increased from around 50% to around 90%. The total wall-time without sharing strategy in 2016 is 40,645,124 CPU hours, while it’s 73,341,585 CPU hours with sharing strategy in 2017, increasing by 80.44%. The results indicate sharing strategy is efficient and integrally promotes experiment data processing.