An Implementation of HTCaaS User Web Portal: Easy Start of HTCaaS

Sangwan Kim*, Seungwoo Rho*, Seoyoung Kim*, Jik-Soo Kim*, Soonwook Hwang*
*Korea Institute of Science and Technology Information
E-mail : {sangwan, seungwoo0926, sssyyy77, jiksoo.kim, hwang}@kisti.re.kr

1. Introduction

HTCaaS(High-Throughput Computing as a Service)[1] is a scheduling system for HTC(High Throughput Computing) styled scientific tasks. An HTC computing job consists of many independent tasks which are distributed to multiples computing resources and executed independently on each resource and the results are gathered together. In KISTI, we have been developed HTCaaS and are going to serve HTCaaS as a operation service for scientists. To do that, we also have developed a web portal system for using HTCaaS more easily. Using HTCaaS web portal, users can describe their HTC problem in a web interface, can submit the tasks through HTCaaS middleware, and can monitor the progress of each tasks in the same web portal. In this paper, we introduce some details of the web portal, functionality, how to use it and implementation technology.

2. HTCaaS Overview

HTCaaS is a pilot-job based multi-level scheduling system, in which task execution is performed in two steps. In the first step, a centralized system executes a pilot job (called as "agent" in HTCaaS) in remote resources through a legacy scheduling system, and in the second step, the pilot job executes real tasks in each resources. In HTCaaS the pilot job fetched tasks from a centralized queue system during run-time. Multi-level scheduling is of merit if the execution tasks are very fine-grained (small), which means relatively large scheduling overhead, or if the computing resources are heterogeneous, which means various job submission methods and authentication mechanism among the different computing resources.

3. HTCaaS Architecture

Figure 1 shows the overall architecture of HTCaaS system. Account Manager manages user information and provides integrated authentication and authorization mechanisms to access various computing infrastructures. User Data Manager is responsible for managing user input and output data during the course of job execution. Job Manager mostly performs job life-cycle management, i.e., from the job submission to the completion. Job Manager maintains separate job queues per user, receives a meta-job (written in JSDL) which can be composed of multiple tasks from a user, validates the meta-job, automatically splits the meta-job into multiple tasks, and controls the execution of each task. Monitoring Manager periodically checks job executions and active agents by interacting with DB Manager, and if needed it initiates failure recovery mechanisms for agents (tasks).

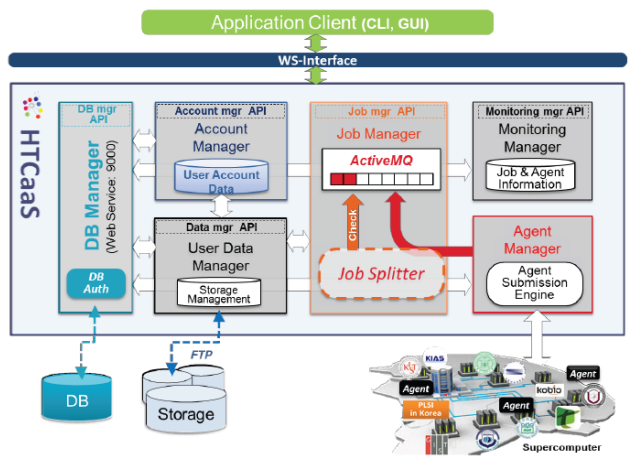

Figure 1. HTCaaS System Architecture

4. HTCaaS Web Portal

We have developed clients for HTCaaS: CLI(command line interface) client, GUI(Java Eclipse RCP) client. Using HTCaaS clients users can submit their HTC jobs to HTCaaS system, monitor the status of running tasks, or get the results of the job. But a demerit of CLI or GUI clients is that users should install a client in their desktop or notebook. We need a web portal client which can be access using a standard web browser without installing or upgrading any program.

HTCaaS web portal is written in java and built using spring framework[2], which is a well-known standard software framework, and runs on top of Apache Tomcat web server.

HTCaaS web portal has the following main features:

- Job script generation : Users can make a job script by providing job name, expected maximum running time of each task, and selecting a template among pre-defined job template list. The total number of tasks of a job is determined by the selected template or parameters from user's input.

- Job submission and progress monitoring : After creating a job script, users can submit the job script and can monitor the progress of submitted jobs. User can monitor how many tasks are waiting, running or done in real-time.

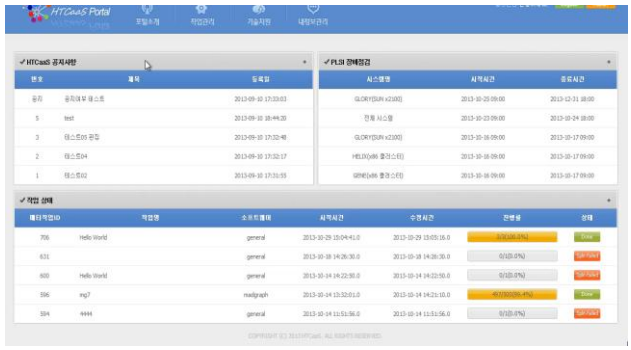

Figure 2. HTCaaS web portal main page

- Tutorial for beginners : Not-practiced users takes much time for understanding how things works and how to submit a job. For these beginning users, web portal system provides a guiding system. By clicking the "Beginner Tutorial" banner in the page, users can start a tutorial mode. In tutorial mode, red guiding messages appears in the page describing what user should do at each step. Tutorial guides user to practice this system submitting a simple but a real job and get the results.

Figure 3. Turial mode for beginners

- Administration : System admin can manage HTCaaS web portal in the admin page, He can manage user account, user group, application software, job template, and services.

5. Conclusion

HTCaaS is a pilot-job based multi-level scheduling system for executing HTC application in distributed multiple resources. Users can access to HTCaaS using various clients, CLI, GUI clients or web portal system. Web portal have merits over CLI or GUI clients, it is easy to use for users and easy to manage for administrator. Using the HTCaaS web portal system, users can make job scripts, submit a job through HTCaaS or monitor the progress. It provides a tutorial mode for beginning users.

KISTI supercomputing center will support HTC-styled computing problem. Especially the PLSI resources distributed in the country are appropriate in HTC computing job. Using HTCaaS and a web portal system, we can provide computing resources effectively for users to run their HTC jobs.

6. References

[1] JS Kim, S Rho, S Kim, S Kim, and S Hwang, "HTCaaS: Leveraging Distributed Supercomputing Infrastructures for Large-Scale Scientific Computing", 6th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers (MTAGS) 2013, ACM, Denver, Colorado USA, 17-22 Nov. 2013

[2] Spring Framework, http://projects.spring.io/spring-framework/