Contribution ID: **22**   Type: **Oral Presentation**

Unified Account Management for High Performance Computing as a Service with Microservice Architecture

Tuesday, 20 March 2018 11:00 (30 minutes)

In recent years, High Performance Computing (HPC) has developed rapidly in China. From Chinese Academy of Sciences (CAS) level, Scientific Computing Grid (ScGrid), is a general-purpose computing platform started from 2006 in CAS, which provided a problem solving environment for computing users through grid computing and cloud computing technologies. Then ScGrid becomes Supercomputing Cloud, an important port of Chins Science Cloud from 2011. From national level, China National Grid (CNGrid) has integrated massive HPC resources from several national supercomputing centers and other large centers distributed geographically, and been providing efficient computing services for users in diverse disciplines and research areas. During more than 10 years, CNGrid and ScGrid has integrated tens of HPC resources distributed geographically across China, comprising 6 National Supercomputer Centers of Tianjin, Jinan, Changsha, and Shenzhen, Guangzhou, Wuxi, and also dozens of teraflops-scale HPC resources belong to universities and institutes. In total, the computing capability is more than 200PF and the storage capacity is more than 160PB in CNGrid.

As worked in the operation and management center of CNGrid and ScGrid for many years, we notice that users prefer to manage their jobs at different supercomputers and clusters via a global account on different remote clients such as science gateways, desktop applications and even scripts. And they don't like to apply for an account to each supercomputer and login into the supercomputer in specific way.

Therefore, we described Unified Account Management as a Service (UAMS) to access and use all HPC resources via a global account for each user in this paper. We addressed and solved challenges for mapping a global account to many local accounts, and provided unified account registration, management and authentication for different collaborative web gateways, command toolkits and other desktop applications. UAMS was designed in accordance with the core rules of simplicity, compatibility and reusability. In architecture design, we focused on loosely-coupled style to acquire good scalability and update internal modules transparently. In implementation, we applied widely accepted knowledge for the definitions of the RESTful API and divided them into several isolated microservices according to their usages and scenarios. For security, all sensitive data transferred in wide-network is protected by HTTPS with transport layer security outside of CNGrid and secure communication channels provided by OpenSSH inside of CNGrid. In addition, all parameters submitted to RESTful web services are strictly checked in format and variable type.

By providing these frequently important but always challenging capabilities as a service, UAMS allows users to use tens of HPC resources and clients via only an account, and makes it easy for developers to implement clients and services related HPC with advantages of numerous users and single sign-on capability. Based on UAMS, representative clients are introduced and reviewed combined with different authentication schemes. Finally, analysis and test of UAMS shows that it can support authentication in milliseconds level and has good scalability. In future, we plan to implement federated account service that enable a local HPC account similar to a global account to login the national HPC environment, access and use all HPC resources in CNGrid.

Summary

In this paper, UAMS is proposed to provide account management and authentication schemes for accessing and using national HPC environment via a global account of CNGrid. UAMS is implemented in self-contained microservices and solved challenges for mapping a global account to many local accounts, and provides unified account registration, management and authentication for administrators and users. It is creative for users to update the requirements of HPC in the full life-cycle of accounts. Correspondingly, administrators adjust assign available privileges to satisfy the demands.

For developers, a group of RESTful API is implemented in several microservices and provides full life-cycle account services as a whole for developers. Besides, UAMS provides the basic username and password authentication scheme and its variants for developers to integrate the authentication service into their clients.

With UAMS, multiple clients, consisting of command-line toolkits, web gateways and scripts, have integrated authentication service based different authentication schemes.

For performance, the result of test shows that the UAMS can provide account management and authentication services at milliseconds level. For scalability, UAMS could elastically stretch the capacity at the granularity of a microservice to deal with massive requests when the load is heavy and release resource when the load becomes slight.

Primary author: Dr CAO, Rongqiang (Computer Network Information Center, Chinese Academy of Sciences)

Co-authors: Prof. XIAO, Haili (Computer Network Information Center, Chinese Academy of Sciences); Ms LU, Shasha (Computer Network Information Center, Chinese Academy of Sciences); Dr WANG, Xiaoning (Computer Network Information Center, Chinese Academy of Sciences); Prof. CHI, Xuebin (Computer Network Information Center, Chinese Academy of Sciences)

Presenter: Dr CAO, Rongqiang (Computer Network Information Center, Chinese Academy of Sciences)

Session Classification: VRE

Track Classification: Virtual Research Environment (including Middleware, tools, services, workflow, … etc.)