Contribution ID: **122**   Type: **Short Talk**

Evolution of the HEPS Jupyter-based remote data analysis System

Thursday, 20 May 2021 11:55 (13 minutes)

High Energy Photon Source (HEPS) has the characteristic of large amount of data, high timeliness, and diverse requirements for scientific data analysis. Generally, researchers need to spend a lot of time in the configuration of the experimental environment. In response to the above problems, we introduce a remote data analysis system for HEPS. The platform provides users a web-based interactive interface with Jupyter, which makes scientists are able to process data analysis anytime and anywhere. Particularly, we discuss the system architecture as well as the key points of this system. A solution of managing and scheduling heterogeneous computing resources (CPU and GPU) is proposed, which adopts Kubernetes to achieve centralized heterogeneous resources management and resource expansion on demand. An improved Kubernetes resource scheduler is discussed. The schedular dispatches resources to upper applications in combination with the cluster status, which can transparently and quickly deployment the data analysis environment for users in seconds and reach the maximum resource utilization. We also introduce an automated deployment solution to improve the work efficiency of developers and help deploy multidisciplinary applications faster and better in the production environment. A unified certification is illustrated to make sure the security of remote data access and data analysis. Finally, we will show the running status of the system.

Primary authors: LIU, Zhibin (Institute of High Energy Physis, CAS; University of Chinese Academy of Sciences); HUANG, Qiulan (Institute of High Energy Physics, CAS); TIAN, Haolai (Institute of High Energy Physics, CAS; Spallation Neutron Source Science Center); HU, Yu (Institute of High Energy Physics, CAS); SHI, Jingyan (Institute of High Energy Physics, CAS); DU, Ran (Institute of High Energy Physics, CAS); HU, Hao (Institute of High Energy Physics, CAS); WANG, Lu (Institute of High Energy Physics, CAS); QI, Fazhi (Institute of High Energy Physics, CAS)

Presenter: LIU, Zhibin (Institute of High Energy Physis, CAS; University of Chinese Academy of Sciences)

Session Classification: Virtualisation

Track Classification: Distributed Computing, Data Management and Facilities