Contribution ID: **72** Type: **Oral Presentation**

Building Infrastructure for Big Earth Data and Cloud Services

Thursday, 4 April 2019 15:00 (30 minutes)

Big Earth Data, as a new type of strategic resource for all nations, brings impetus to Earth Science, and will be a new essential tool to understand the world. To improve the capability of effectively collecting, storing, managing and analyzing Big Earth Data, and to create groundbreaking discoveries, it is imperative to build a Cloud Service Platform and to develop innovative technologies and methodologies for Big Earth Data research. Chinese Academy of Sciences has accumulated vast amount of long-term multi-source scientific data resources and the data acquisition abilities from space to land. Based on these advantages, we propose to construct world leading infrastructure for Big Earth Data and Cloud Service. The infrastructure will support scientific discovery and decision making by integrating and composing a multidisciplinary fusing Big Earth Data Repository, and providing data products and capabilities of computing, storing and analyzing for CASEarth, and.

The program consists of four main parts: (1) The Big Earth Data Infrastructure: Building an extensible Big Earth Data platform to provide advanced computing and storage capabilities by deploying specific computing and storage clusters which converge high-performance computing, high-throughput computing and large-scale data storage. It also aggregates existing computing facilities including China National Grid, China Science & Technology Cloud. (2) The Big Earth Data Repository: Constructing a Big Earth Data Repository to provide data capability by collecting multi-source data of many domains like biology, ecology, environtology, and ingesting data products produced by Big Earth Data research projects. (3) Big Earth Data System Software: Developing systems, including storage management system, computing and processing system, data mining system, grid data engine to provide technical support for Big Earth Data research. (4) A Cloud Service Portal: Developing a Cloud Service Portal to provide unified public services of big data storing, computing, processing, analyzing, and application integrating for users from many research areas.

According to our plan, the infrastructure consisting of specific computing and storage clusters will be deployed before 2019, providing 2PF computing capability and 50PB storage capability. Besides, plenty of services and applications will be deployed on the infrastructure to improve the capabilities of data collecting, storing, managing and analyzing for Earth Science in CAS.

Primary author: CHI, Xuebin (Computer Network Information Center, Chinese Academy of Sciences)

Presenter: CHI, Xuebin (Computer Network Information Center, Chinese Academy of Sciences)

Session Classification: Data Management & Big Data

Track Classification: Data Management & Big Data