

Scientific data management at HEPS

Wednesday, March 24, 2021 1:00 PM (20 minutes)

The High Energy Photon Source (HEPS) in Beijing is the first national light source of high-energy synchrotron radiation in China, and will be one of the world's brightest fourth-generation synchrotron radiation facilities. Doubtless data are of crucial importance for the scientific discoveries made in the experiments at HEPS. According to the estimated data rates, we predict 30 PB raw experimental data will be produced per month from 14 beamlines at the first stage of HEPS, and the data volume will be even greater after over 90 beamlines are available at the second stage in the near future.

Therefore, successful data management is critically important for the present and future scientific productivity of HEPS. The data management system is responsible for automating the organization, transfer, storage, distribution of the data collected at HEPS. First of all, the main features of the scientific data acquired from all the beamlines and the possible problems exists in data managing and data sharing is explained in this paper. Second, the architecture and data flow of the HEPS data management system are described from the perspective of facility users and IT. Furthermore, key techniques implemented in this system are introduced. Finally, the progress and the effect of the data management system deployed as a pilot test at BSRF are given.

Primary author: Ms HU, Hao (Institute of High Energy Physics)

Presenter: Ms HU, Hao (Institute of High Energy Physics)

Session Classification: Data Management & Big Data Session

Track Classification: Data Management & Big Data