14th International Conference on

Metabolomics and Proteomics

March 28-29, 2019   Osaka, Japan

iProX: An integrated proteome resource

Tao Chen
Beijing Institute of Life Omics, China

Sharing of research data in public repositories has become best practice in academia. With the accumulation of massive data, network bandwidth and storage requirements are rapidly increasing. The ProteomeXchange (PX) consortium implements a mode of centralized metadata and distributed raw data management, which promotes effective data sharing. To facilitate open access of proteome data worldwide, we have developed the integrated proteome resource iProX as a public platform for collecting and sharing raw data, analysis results and metadata obtained from proteomics experiments. The iProX repository employs a web-based proteome data submission process and open sharing of mass spectrometry-based proteomics datasets. Also, it deploys extensive controlled vocabularies and ontologies to annotate proteomics datasets. Users can use a GUI to provide and access data through a fast Aspera-based transfer tool. iProX is a full member of the PX consortium; all released datasets are freely accessible to the public. iProX is based on a high availability architecture and has been deployed as part of the proteomics infrastructure of China, ensuring long-term and stable resource support. iProX will facilitate worldwide data analysis and sharing of proteomics experiments. Based on a hyper-converged architecture, the application, virtualization and infrastructure layer are implemented for the iProX repository. The infrastructure layer includes hardware resources of computing, storage and network, while the virtualization layer enables a unified and virtualized resource pool to provide the real-time running resources required by iProX. The application layer uses virtual machines to achieve load balance and high availability, including load balance, application, database and datafile servers. The load balance is achieved by using the nginx and keep alived technologies on load balance and application servers, whereas two database servers are real-time synchronized to guarantee the high availability and two file servers are used for high speed data transfer and backup.

Biography

Tao Chen is currently working as an Assistant Professor at Beijing Institute of Life Omics. She has published more than 10 papers in reputed journals of proteomics, bioinformatics and computer science.

taochen1019@163.com

Notes: