Contribution ID: **0**   Type: **Plenary reports**

Geographically Distributed Software Defined Storage

Friday, 8 July 2016 09:00 (20 minutes)

The volume of the coming data in HEP is growing. Also growing volume of the data to be hold long time. Actually large volume of data –big data –is distributed around the planet. In other words now there is situation where the data storage does integrate storage resources from many data centers located far from each other. That means the methods, approaches how to organize, manage the globally distributed data storage are required.

For personal needs the distributed storage has several examples like owncloud.org, pydio.com, seafile.com, sparkleshare.org. For enterprise level there are a number of distributed storage systems SWIFT (part of Openstack), CEPH and the like which are mostly object storage.

When distributed storage integrate several data center resources the organization of data links becomes very important issue especially if several parallel data links between data centers are used. The situation on data centers and in data links might be changed each hour. All that means each part of distributed data storage has to be able to rearrange usage of data links and storage servers in each data center. In addition for each customer of distributed storage the different requirements have to be satisfied.

Above topics are planned to be discussed in the proposal of data storage architecture.

Primary author: Mr SHEVEL, andrey (PNPI, ITMO)

Co-authors: Mr KAIRKANOV, Arsen (ITMO);   Mr SADOV, Oleg (ITMO);   Dr KHORUZHNIKOV, Sergey (ITMO);   Dr GRUDININ, Vladimir (ITMO)

Presenter: Mr SHEVEL, andrey (PNPI, ITMO)

Session Classification: Plenary reports

Track Classification: 10. Databases, Distributed Storage systems, Big data Analytics