Contribution ID: **381**   Type: **Poster**

A New Data Access Mechanism for HDFS

Tuesday, 11 October 2016 16:30 (15 minutes)

With the era of big data emerging, Hadoop has become de facto standard of big data processing. However, it is still difficult to get High Energy Physics (HEP) applications run efficiently on HDFS platform. There are two reasons to explain. Firstly, Random access to events data is not supported by HDFS platform. Secondly, it is difficult to make HEP applications adequate to Hadoop data processing mode. In order to address this problem, a new read and write mechanism of HDFS is proposed. With this mechanism, data access is done on local filesystem instead of through HDFS streaming interface. For data writing, the first file replica is written to the local DataNode, the rest replicas produced by copy of the first replica stored on other DataNodes. The first replica is written under the Blocks storage directory and calculates data checksum after write completion. For data reading, DataNode Daemon provides the data access interface for local Blocks, and Map tasks can read the file replica directly on local DataNode when running locally. To enable files modified by users, three attributes including permissions, owner and group are imposed on Block objects. Blocks stored on DataNode have the same attributes as the file they belong to. Users can modify Blocks when the Map task running locally, and HDFS is responsible to update the rest replicas later after data access done. To further improve the performance of Hadoop system, two optimization on Hadoop scheduler are conducted. Firstly, a Hadoop task selection strategy is presented based on disk I/O performance. With this strategy, an appropriate Map task is selected according to disk workloads, so that disk balance workload is achieved on DataNodes. Secondly, a complete localization task execution mechanism is implemented for I/O intensive jobs. Test results show that average CPU utilization is improved by 10% with the new task selection strategy, data read and write performance is improved about 10% and 40% separately.

Tertiary Keyword (Optional)

Secondary Keyword (Optional)

Data processing workflows and frameworks/pipelines

Primary Keyword (Mandatory)

Cloud technologies

Primary author: Mr LI, Qiang (INSTITUTE OF HIGH ENERGY PHYSICS)

Co-authors: Prof. SUN, Gongxing (INSTITUTE OF HIGH ENERGY PHYSICS); Mr WEI, Zhanchen (INSTITUTE OF HIGH ENERGY PHYSICS); Mr SUN, Zhenyu (INSTITUTE OF HIGH ENERGY PHYSICS)

Session Classification: Posters A / Break

Track Classification: Track 4: Data Handling