

0.1 Qubole Data Service

Qubole is a data services company who offers cloud services, data management, and data system security to companies. The BigData Platform in Qubole Data services contains all important elements for companies to process their data, and it is not hard to use. Multiple cloud providers can be chose, the company could select the cloud provider for their own commercial purpose (the most of the cloud providers in Qubole Data Service are open source)[**QDS**].

0.2 CensOS Project

The CensOS Project is a open source project that focuses on developing robust open source ecosystem to personal user, open source community, and companies like Amazon, Google and so on. The CensOS Project provides the both individual users and companies cloud image and powerful cloud developing tools, which enables them to build their own cloud service upon the services offer by the CensOS Project[**centOS**].

0.3 HPCC Systems

HPCC (High Performance computer cluster) systems are open source tool which offers the BigData related services[**HPCC**]. HPCC contains tools that deal with complex data structure and large scale of data amount. It is a powerful open source tool for data analyze, especially for the significant size of data. The functionalities of HPCC such as fast querying to different databases, , data visualization, and data management have good reputation to many users.

0.4 Google Cloud Platform - Big data solutions

Big data solutions is a part of Google Cloud Platform services; it offers special service on data analyzes[**GoogleCP_BDS**]. The users (usually companies) could combine with the cloud services (see more in Google Cloud Platform - Cloud Dataproc) to manage their data and proceed data analyzing machine learning in order to predict business decision. The Big data solutions offers all of the services about data management, efficient data query, and machines intelligence services, which is invaluable for commercial purpose[**GoogleCP_BDS**].

0.5 Google Cloud Platform - Cloud Dataproc

Cloud Dataproc is a part of Google Cloud Platform services. It is a efficient, not sophisticated, and well managed cloud service for companies who running Apache Spark and Apache Hadoop clusters as cloud base. The companies can adopt their cloud system in to the Cloud Dataproc, and use Google Cloud Platform services to manage data, mine useful information from data and so on (see an example in Google Cloud Platform - Big data solutions). It combines with other services in Google Cloud Platform services to generates a complete cloud service for dealing with large amount of data size[**GoogleCP_CD**].

References