Cloudease - Pseudodistributed Distribution Of Hadoop Hive Pi

[DOWNLOAD HERE](#)

CloudEase - PseudoDistributed Distribution of Hadoop Hive Pig HBase Mahout Cascading & Zookeeper Hadoop is a software platform that allows processing of BigData (TeraBytes/PetaBytes) of data. The Hadoop Core contains the basica Map Reduce system and HDFS a distributed file system. While HDFS stores peta bytes of data that need to be processed by the hadoop cluster. The Map Reduce paradigm allows a problem to be broken into thousands/millions of small tasks (Map) and to be processed over the cluster and the results aggregated back (Reduce) into a consistent usable result. Hive and Pig are used for Analytics and DataWarehousing over the Hadoop Cluster. HBase is an internet scale non-relational database system that runs over hadoop. Mahout is for internet scale machine learning systems. Cascading is for Data Processing Workflows. And ZooKeeper is a centralized service for maintaining configuration information, naming, providing distributed synchronization, and providing group services. Setting up Hadoop and managing it is a very cumbersome task that requires a dedicated Admin team. The VirtualBox Cloud Ease virtual machine is a pseudo distributed installation of Hadoop and other tools mentioned above. Everything is installed on a single virtual machine though it behaves like a cluster. It is perfect to kickstart Hadoop development and can be run using VirtualBox on Windows, Linux or Mac host operating systems. It is not intended for a production cluster deployment as is. If you need to run it over VMWare or any other virtualization platform you can export a Open Virtualization Format (OVF) VM Image using Virtual Box and use that. CloudEase is the only distribution which comes with the above 7 tools pre-installed and pre-configured ready to develop with. Specs: The CloudEase VM is built using Ubuntu 64-bit server OS with Jdk 1.6, ant and the tools mentioned above. The VM has a 400GB expandable HDD. Terms: We do not provide technical support. The VM is provided as is without any liabilities or warranties. It is licensed for use by one Named developer.

Similar manuals:

[Hadoop: The Definitive Guide By Tom White](#)

[CloudEase - PseudoDistributed Distribution Of Hadoop Hive Pi](#)

[Pro Hadoop](#)