A wide range of data-intensive applications such as marketing analytics, image processing, machine

learning, andWeb crawling use Apache Hadoop, an open-source, Java-based software system.6 Hadoop

supports distributed applications handling extremely large volumes of data. Many members of the

community contributed to the development and optimization of Hadoop and several related Apache

projects such as Hive and HBase.

Hadoop is used by many organizations from industry, government, and research; the long list of

Hadoop users includes major IT companies such as Apple, IBM, HP, Microsoft, Yahoo!, and Amazon;

media companies such as The New York Times and Fox; social networks, including Twitter, Facebook,

and LinkedIn; and government agencies, such as the U.S. Federal Reserve. In 2011, the Facebook

Hadoop cluster had a capacity of 30 PB.

A Hadoop system has two components, a MapReduce engine and a database (see Figure 8.8). The

database could be the Hadoop File System (HDFS), Amazon S3, or CloudStore, an implementation of

the Google File System discussed in Section 8.5. HDFS is a distributed file system written in Java; it is

portable, but it cannot be directly mounted on an existing operating system. HDFS is not fully POSIX

compliant, but it is highly performant.

The Hadoop engine on the master of a multinode cluster consists of a job tracker and a task tracker,

whereas the engine on a slave has only a task tracker. The job tracker receives a MapReduce job from a client and dispatches the work to the task trackers running on the nodes of a cluster. To

increase efficiency, the job tracker attempts to dispatch the tasks to available slaves closest to the

place where it stored the task data. The task tracker supervises the execution of the work allocated to the

node. Several scheduling algorithms have been implemented in Hadoop engines, including Facebook’s

fair scheduler and Yahoo!’s capacity scheduler; see Section 6.8 for a discussion of cloud scheduling

algorithms.

HDFS replicates data on multiple nodes.The default is three replicas; a large dataset is distributed over

many nodes. The name node running on the master manages the data distribution and data replication and

communicates with data nodes running on all cluster nodes; it shares with the job tracker information

about data placement to minimize communication between the nodes on which data is located and the

ones where it is needed. Although HDFS can be used for applications other than those based on the

MapReduce model, its performance for such applications is not at par with the ones for which it was

originally designed.