Calculate Your Total Cost of Ownership of Apache™ Hadoop®

The power of big data solutions continues to grow significantly every year, and the cost of collecting, managing, and storing data is also increasing. As a result, organizations are rethinking their enterprise architecture in order to find ways to reduce cost and increase the flexibility of their data management/storage, processing solution.

As Hadoop deployments grow within an organization, the architectural differences between Hadoop distributions begin to show dramatic cost differences across both capital and operational expenses. These differences can save companies 20-50% in terms of total cost of ownership (TCO).

Less Hardware = Lower TCO. In a recent survey of 52 MapR customers with experience with Apache Hadoop, Cloudera or Hortonworks, 31% of surveyed IT organizations who had experience with another Hadoop distribution selected the MapR Distribution including Hadoop because of lower hardware requirements.*

A Practical TCO Comparison. The example below illustrates a cost comparison for a 500 TB cluster between two vendors’ Hadoop distributions based on a customer-validated TCO model. The TCO for MapR in the example is $3.2 million over three years, compared to another Hadoop distribution at $4.7 million for the same period. MapR provides a 32% savings.


Cost comparison for a 500 TB cluster between two commercial Hadoop distributions
MapR Architecture Significantly Lowers Hadoop TCO in Four Key Areas

With the MapR Distribution including Hadoop, you get the benefits of all the open source innovation with an enterprise-grade data platform that is as reliable and performant as best-in-class DBMS or NAS systems. The MapR architecture dramatically lowers TCO in four key areas:

1. Higher Performance = More Throughput with Less Hardware

MapR is known for record-setting speed. A MapR customer broke the official MinuteSort record by sorting 1.65 TB with a 298-node cluster. That is 1/7th the hardware of the previous record using 2,200 nodes. This shows you can expect smaller data center footprint for the same performance with other Hadoop distributions.

2. MapR No-NameNode Architecture Reduces Hardware Required

MapR provides a fully-distributed data platform which distributes and replicates the NameNode metadata across the cluster. This architecture provides the standard Hadoop NameNode functionality, but without the separate NameNode server. With no NameNode, there are no practical limits to the number of files that can be used in Hadoop. Even more, MapR gives you high availability (HA) without special-purpose NameNode hardware and standby servers required by other distributions. Both aspects lead to better reliability and less administration for Hadoop with a significantly smaller data center footprint.

3. Automatic File Compression

MapR applies compression automatically to files in the cluster. Compression is 2-3x depending on file types and compression settings, which reduces storage requirements, as well as uses less bandwidth on the network, resulting in improved performance.

4. True Multi-Tenancy for Hadoop

Support for sophisticated workload management results in a reduction of disparate systems. MapR allows multiple tenants and workloads in a single deployment by combining YARN with volumes, job and data placement control, and label-based scheduling. With this unique combination, you can perform fine-grained resource management not only over compute, but all the way down to the node. As an example, one media customer was able to consolidate eight single-tenant HBase clusters down to an operationally efficient single cluster with MapR.

MapR delivers on the promise of Hadoop with a proven, enterprise-grade platform that supports a broad set of mission-critical and real-time production uses. MapR brings unprecedented dependability, ease-of-use and world-record speed to Hadoop, NoSQL, database and streaming applications in one unified distribution for Hadoop. MapR is used by more than 500 customers across financial services, government, healthcare, manufacturing, media, retail and telecommunications as well as by leading Global 2000 and Web 2.0 companies. Investors include Google Capital, Lightspeed Venture Partners, Mayfield Fund, NEA, Qualcomm Ventures and Redpoint Ventures.

About MapR

Read more on Hadoop TCO
mapr.com/blog/get-real-hadoop-lower-your-tco