



Intel Designs a Big Data Network Based on Cloudera Enterprise* for a Regional Satellite Dish TV Provider

Intel and Cloudera* provide architecture tips and guidelines for deploying a Big Data Hadoop*-based network from scratch.



Why Intel and Cloudera

Intel and Cloudera take the guesswork out of Apache* Hadoop*. Using a unique collaborative approach, we deliver excellent performance, security, and quality distribution, built on open standards. Because we work with hundreds of vendors across the ecosystem, a solution built on Cloudera Enterprise can ensure freedom from lock-in, enabling you to build a robust Big Data solution to meet the needs of your business today and into the future.

- Uniquely aligned product roadmaps for software and hardware to drive innovation faster, providing many industry firsts with Hadoop*.
- Deep partnerships with virtually every provider in the data center, streamlining the process for building Big Data solutions.
- Proven track records of identifying the driving industry standards, so you don't run the risk of stranding yourself on an island.

A regional satellite dish television provider asked Intel to help them redesign their IT infrastructure to accommodate a big data environment. Intel designed their system from the ground up. Before deploying a Hadoop* cluster, you should make certain your environment—including the operating system, network, firewall, and hardware resources—is tuned appropriately to achieve the best performance and avoid common pitfalls encountered during a cluster setup.

In this solution brief, we provide information on common tuning and setup activities required before setting up a Hadoop* cluster.

HDFS disks: Use JBOD, not RAID

For HDFS data disks, we recommend using JBOD rather than RAID, as a JBOD configuration makes the best use of available space, performs better, and is more reliable.

We do not recommend mirrored or striped RAID volumes. By design, HDFS takes care of replicating the data across multiple nodes (default is 3) to make the file system tolerate disk and node failures. Mirroring is overkill, as we will be mirroring an already replicated data.

On RAID stripes, all data is written in a distributed fashion across the disks. As disks age, they tend to develop errors, and eventually reads become slower. If the disks are on RAID, the slower disks could hinder overall read performance. In

a JBOD configuration, however, it is easy to remove slow old disks. For similar reasons, we also do not recommend Linux LVM.

Mount with *noatime* option

With every disk read, the file system records the last access time. Since there may be potentially several terabytes or even petabytes of data stored on the cluster, this could lead to a significant overhead while to record the access time metadata. We recommend disabling this feature when mounting the data disk.

OS disks: Use mirrored/RAID disks

It is good practice to configure your OS disks in RAID1, especially the master nodes. Some of the major components like HDFS NameNode and Yarn Resource Manager can be made highly available (HA), but other components cannot be HA.

Don't store logs, OS on same partition

Most of Hadoop's components store their logs under `/var/log` (configurable). Over extended periods of time, logs tend to use several gigabytes of disk space. If the operating system and logs are on the same disk, the OS could run out of space, which could lead to several issues.

File system

We recommend running a stable file system on the raw disks, such as Ext3 or Ext4.

Hardware profile

The hardware cluster we configured for the Customer (Figure 1) consisted of seven nodes (three masters and four worker nodes).

Operating system

Use a stable, tested operating system that is compatible with the Hadoop* distribution and version you are installing. It is common for Hadoop* distributions to be supported only on certain OS platforms. Providing a stable and supported OS as a platform for the Hadoop* cluster to run on, is crucial.

The most commonly used OSes today are Red Hat Enterprise Linux (6.x, 7.x) and CentOS (6.x, 7.x). Other supported OSes include: Oracle Enterprise Linux (6.x, 7.x), SUSE Linux Enterprise Server (11), Debian (Wheezy, Squeeze), and Ubuntu (Trusty, Precise).

Note: These are all 64-bit operating systems.

Check your Hadoop* distribution's documentation for an accurate list for the version you are setting up.

In addition to running on a supported operating system, we recommend a few configurations to enable Hadoop* to run optimally, including the following.

Use NTP (Network Time Protocol)

The time on all nodes in your cluster needs to be synchronized to prevent the cluster from running into weird errors.

/etc/security/limits.conf

Increase the limits of number of open files and number of processes to the upper limit of 32,768 for Hadoop*-related system users.

HDFS, being a distributed file system, opens a lot of files at the OS level across your cluster (even when you want to access only a single file on the HDFS level) and also spawns

several thousands of processes across the cluster from time to time.

Make these changes for Hadoop*'s system users (HDFS, HBase, MapReduce, etc.).

Disable SELinux

Hadoop* and its ecosystem work in a more complicated manner than traditional software. Many of our clients disable SELinux because constructing policies to enforce it on a Hadoop* cluster is so difficult.

Disable swappiness

Because swapping slows down your Hadoop* cluster, we recommend to turning swapping off altogether. In the `/etc/sysctl.conf` file, disable swap by setting `vm.swappiness=0`. For the latest kernels, set `vm.swappiness=1`.

Read the documentation of the operating version you are running to identify the appropriate way to disable swap.

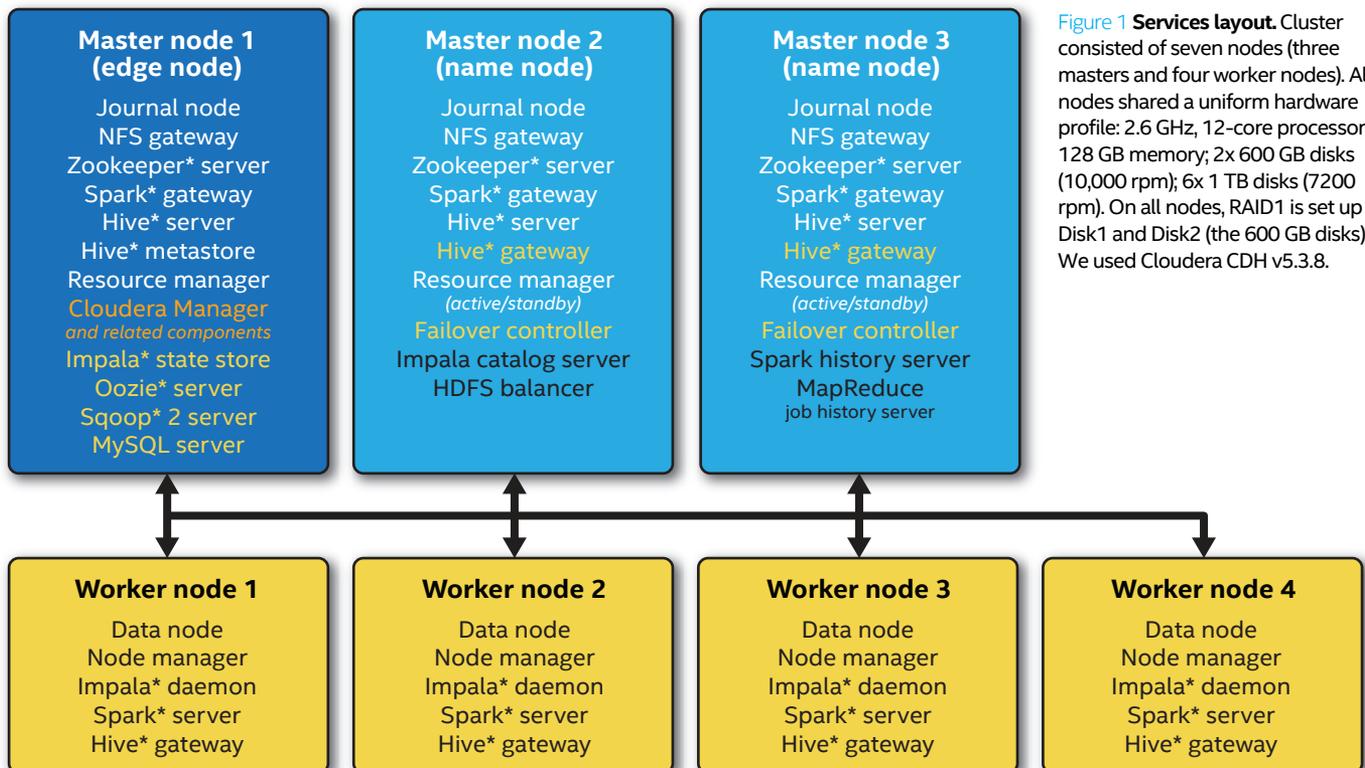


Figure 1 Services layout. Cluster consisted of seven nodes (three masters and four worker nodes). All nodes shared a uniform hardware profile: 2.6 GHz, 12-core processor; 128 GB memory; 2x 600 GB disks (10,000 rpm); 6x 1 TB disks (7200 rpm). On all nodes, RAID1 is set up on Disk1 and Disk2 (the 600 GB disks). We used Cloudera CDH v5.3.8.

Firewalls

Either disable your firewalls or explicitly enable the ports that each component in the Hadoop* ecosystem will use. This could very well be scores or even hundreds of ports, depending on what components you are running.

Disable transparent huge pages (THP)¹

On the operating system level, memory is managed in blocks known as pages. A page is 4096 bytes. 1 MB of memory is equal to 256 pages; 1 GB of memory is equal to 256,000 pages, etc. CPUs have a built-in memory management unit that contains a list of these pages, with each page referenced through a page table entry. Huge pages are blocks of memory that come in 2 MB and 1 GB sizes. The page tables used by the 2 MB pages are suitable for managing multiple gigabytes of memory, whereas the page tables of 1 GB pages are best for scaling to terabytes of memory.

Because transparent huge page compaction interacts poorly with Hadoop* and reduces the performance by a large factor, we recommend disabling THP.

Use a supported and stable RDBMS

Many components use an external RDBMS to store metadata and data. For example, Hive* stores the metadata about tables on an external RDBMS; Oozie* stores job workflow history, schedules, etc. on a DBMS.

The most commonly used databases are MySQL 5.x and Oracle 11gR2 or 12c. Other supported databases include Derby and PostgreSQL 9.x (not recommended).

Check the documentation of your Hadoop* distribution to ensure that you are installing a supported operating system.

Use DNS for name resolution

We recommend performing name lookups via a DNS server rather than from the `/etc/hosts` file. This makes the cluster less error-prone.

Note: It is important to note that a Hadoop* cluster can put a significant load on the DNS server. Make sure your DNS server can handle this load.

Hostname: Use FQDN

Make sure the hostname of each of the nodes is set to the fully qualified domain name (FQDN). While enabling security features like Kerberos or TLS, this is crucial.

Keep OS and kernel up to date

Making sure your operating system and kernel are always the latest ensures that you will be taking advantage of the latest bug fixes and performance improvements on the operating system.

Hadoop* components layout

Define the architecture of the Hadoop* cluster and its ecosystem components. The following is based on a real-life architecture on a pre-production cluster we deployed for a major direct-to-home (DTH) provider in the Asia Pacific region.

Choose the right hardware for your workload.² It is important to note that memory is becoming cheaper and many new Hadoop* ecosystem components like Apache Spark* and Apache Impala* make use of the extra memory to process your data faster.

Summary

This document is not a step-by-step instruction manual for deploying Hadoop*, but we hope it proves helpful to those considering such an infrastructure migration.

Let us help your business too.

Spotlight on Cloudera

Cloudera is revolutionizing enterprise data management by offering a unified platform for Big Data, an enterprise data hub built on Apache* Hadoop*. Cloudera offers enterprises one place to store, access, process, secure, and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data.

Cloudera's open source Big Data platform is the most widely adopted in the world, and Cloudera is the most prolific contributor to the open source Hadoop ecosystem. As the leading educator of Hadoop professionals, Cloudera has trained over 40,000 individuals worldwide. Over 1,900 partners and a seasoned professional services team help deliver greater time to value. Finally, only Cloudera provides proactive and predictive support to run an enterprise data hub with confidence. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production.

For more information, visit www.cloudera.com.

cloudera[®]

1. this section stolen from Red Hat webpages.

2. For a reference, check out this Cloudera article: <http://blog.cloudera.com/blog/2015/01/how-to-deploy-apache-hadoop-clusters-like-a-boss>.

Meeting your needs

We look forward to meeting with you to define your requirements and meet your objectives.

- **Accelerate time to value:** Achieve real-time cost savings, respond to market trends, and drive innovation.
- **Secure Big Data:** Deploy a sustainable Big Data program that doesn't put your organization, or you, at risk.
- **Maintain control:** Work with a partner who educates your team so you become self-sufficient.
- **Increase business potential:** Create and execute a plan that helps you adapt now, and in the future.

Contact us

Contact your sales rep or e-mail us at: Hadoop-services@intel.com.

Intel.com/bigdata/services

Hadoop sizing guide

		Cluster size		
		Small	Medium	Large
CPU		Intel® Xeon® Processor E5 v3		
Storage (TB)		<72 TB	72 to 570 TB	>570 TB
Node count	Master	2 to 3	4 to 7	≥8
	Slaves	<12	12 to 95	≥ 96
Memory (GB)	Master	64 GB	128 GB	≥256 GB
	Slaves	48 GB	96 GB	≥128 GB
Network		1 Gbps	10 Gbps	10 Gbps

Hardware configuration is highly dependent on workload. A high storage density cluster may be configured with a 4 TB JBOD hard disk, while a compute intensive cluster may be configured with a higher memory configuration.

References

- <http://blog.cloudera.com/blog/2013/08/how-to-select-the-right-hardware-for-your-new-hadoop-cluster>
- <http://blog.cloudera.com/blog/2015/01/how-to-deploy-apache-hadoop-clusters-like-a-boss>
- https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/6/html/Performance_Tuning_Guide/s-memory-transhuge.html

cloudera®



The results cited in this document are based on research and testing conducted by Intel and Intel's customer and are for informational purposes only.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit <http://www.intel.com/performance>.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at www.intel.com.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725 or by visiting Intel's website at <http://www.intel.com/design/literature.htm>.

Intel and the Intel logo are trademarks of Intel Corporation in the United States and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2015 Intel Corporation. All rights reserved.

♻️ Please Recycle

333539-001