Introduction to Hadoop

A Dell Technical White Paper

By Joey Jablonski

Dell | Hadoop White Paper Series

The power to do more

Table of Contents

Introduction to Hadoop

Hadoop is a rapidly evolving ecosystem of components for implementing the Google MapReduce algorithms in a scalable fashion on commodity hardware. Hadoop enables users to store and process large volumes of data and analyze it in ways not previously possible with less scalable solutions or standard SQL-based approaches.

As an evolving technology solution, Hadoop design considerations are new to most users and not common knowledge. As part of the Dell | Hadoop solution, Dell has developed a series of best practices and architectural considerations to use when designing and implementing Hadoop solutions.

Hadoop is a highly scalable compute and storage platform. While most users will not initially deploy servers numbered in the hundreds or thousands, Dell recommends following the design principles that drive large, hyper-scale deployments. This ensures that as you start with a small Hadoop environment, you can easily scale that environment without rework to existing servers, software, deployment strategies, and network connectivity.

Hadoop challenges

With all large environments, deployment of the servers and software is an important consideration. Dell provides best practices for the deployment of Hadoop solutions. These best practices are implemented through a set of tools to automate the configuration of the hardware, installation of the operating system (OS), and installation of the Hadoop software stack from Cloudera.

As with many other types of information technology (IT) solutions, change management and systems monitoring are a primary consideration within Hadoop. The IT operations team needs to ensure tools are in place to properly track and implement changes, and notify staff when unexpected events occur within the Hadoop environment.

Hadoop is a constantly growing, complex ecosystem of software and provides no guidance to the best platform for it to run on. The Hadoop community leaves the platform decisions to end users, most of whom do not have a background in hardware or the necessary lab environment to benchmark all possible design solutions. Hadoop is a complex set of software with more than 200 tunable parameters. Each parameter affects others as tuning is completed for a Hadoop environment and will change over time as job structure changes, data layout evolves, and data volume grows.

As data centers have grown and the number of servers under management for a given organization has expanded, users are more conscious of the impact new hardware will have on existing data centers and equipment.

Hadoop node types

Hadoop has a variety of node types within each Hadoop cluster; these include DataNodes, NameNodes, and EdgeNodes. Names of these nodes can vary from site to site, but the functionality is common across the sites. Hadoop's architecture is modular, allowing individual components to be scaled up and down as the needs of the environment change. The base node types for a Hadoop cluster are:

- **NameNode** – The NameNode is the central location for information about the file system deployed in a Hadoop environment. An environment can have one or two NameNodes, configured to provide minimal redundancy between the NameNodes. The NameNode is contacted by clients of the Hadoop Distributed File System (HDFS) to locate information within the file system and provide updates for data they have added, moved, manipulated, or deleted.
- **DataNode** – DataNodes make up the majority of the servers contained in a Hadoop environment. Common Hadoop environments will have more than one DataNode, and oftentimes they will number in the hundreds based on capacity and performance needs. The DataNode serves two functions: It contains a portion of the data in the HDFS and it acts as a compute platform for running jobs, some of which will utilize the local data within the HDFS.
- **EdgeNode** – The EdgeNode is the access point for the external applications, tools, and users that need to utilize the Hadoop environment. The EdgeNode sits between the Hadoop cluster and the corporate network to provide access control, policy enforcement, logging, and gateway services to the Hadoop environment. A typical Hadoop environment will have a minimum of one EdgeNode and more based on performance needs.

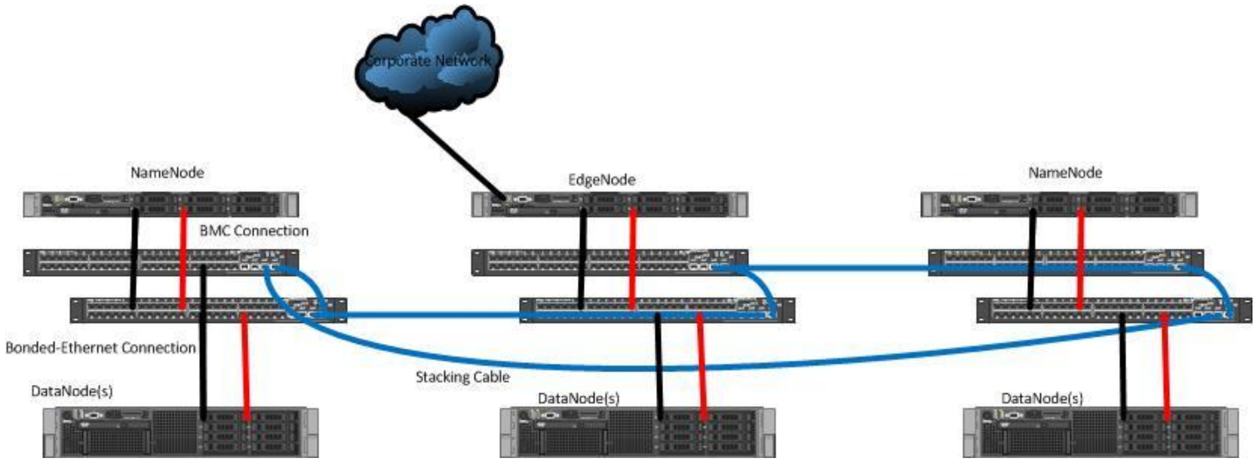

Figure 1. Node types within Hadoop clusters.

Hadoop uses

Hadoop was originally developed to be an open implementation of Google MapReduce and Google File System. As the ecosystem around Hadoop has matured, a variety of tools have been developed to streamline data access, data management, security, and specialized additions for verticals and industries. Despite this large ecosystem, there are several primary uses and workloads for Hadoop that can be outlined as:

- **Compute** – A common use of Hadoop is as a distributed compute platform for analyzing or processing large amounts of data. The compute use is characterized by the need for large numbers of CPUs and large amounts of memory to store in-process data. The Hadoop ecosystem provides the application programming interfaces (APIs) necessary to distribute and track workloads as they are run on large numbers of individual machines.
- **Storage** – One primary component of the Hadoop ecosystem is HDFS—the Hadoop Distributed File System. The HDFS allows users to have a single addressable namespace, spread across many hundreds or thousands of servers, creating a single large file system. HDFS manages the replication of the data on this file system to ensure hardware failures do not lead to data loss. Many users will use this scalable file system as a place to store large amounts of data that is then accessed within jobs run in Hadoop or by external systems.
- **Database** – The Hadoop ecosystem contains components that allow the data within the HDFS to be presented in a SQL-like interface. This allows standard tools to INSERT, SELECT, and UPDATE data within the Hadoop environment, with minimal code changes to existing applications. Users will commonly employ this method for presenting data in a SQL format for easy integration with existing systems and streamlined access by users.

What is Hadoop good for?

When the original MapReduce algorithms were released, and Hadoop was subsequently developed around them, these tools were designed for specific uses. The original use was for managing large data sets that needed to be easily searched. As time has progressed and as the Hadoop ecosystem has evolved, several other specific uses have emerged for Hadoop as a powerful solution.

- **Large Data Sets** – MapReduce paired with HDFS is a successful solution for storing large volumes of unstructured data.
- **Scalable Algorithms** – Any algorithm that can scale to many cores with minimal inter-process communication will be able to exploit the distributed processing capability of Hadoop.
- **Log Management** – Hadoop is commonly used for storage and analysis of large sets of logs from diverse locations. Because of the distributed nature and scalability of Hadoop, it creates a solid platform for managing, manipulating, and analyzing diverse logs from a variety of sources within an organization.
- **Extract-Transform-Load (ETL) Platform** – Many companies today have a variety of data warehouse and diverse relational database management system (RDBMS) platforms in their IT environments. Keeping data up to date and synchronized between these separate platforms can be a struggle. Hadoop enables a single central location

for data to be fed into, then processed by ETL-type jobs and used to update other, separate data warehouse environments.

Not so much?

As with all applications, some actions are not optimal for Hadoop. Because of the Hadoop architecture, some actions will have less improvement than others as the environment is scaled up.

- **Small File Archive** – Because of the Hadoop architecture, it struggles to keep up with a single file system name space if large numbers of small objects and files are being created in the HDFS. Slowness for these operations will occur from two places, most notably the single NameNode getting overwhelmed with large numbers of small I/O requests and the network working to keep up with the large numbers of small packets being sent across the network and processed.
- **High Availability** – A single NameNode becomes a single point of failure and should be planned for in the uptime requirements of the file system. Hadoop utilizes a single NameNode in its default configuration. While a second, passive NameNode can be configured, this must be accounted for in the solution design.

Hadoop architecture and components

Hadoop design

Figure 2 depicts the Dell representation of the Hadoop ecosystem. This model does not include the applications and end-user presentation components, but does enable those to be built in a standard way and scaled as your needs grow and your Hadoop environment is expanded.

The representation is broken down into the Hadoop use cases from above: Compute, Storage, and Database workloads. Each workload has specific characteristics for operations, deployment, architecture, and management. Dell's solutions are designed to optimize for these workloads and enable you to better understand how and where Hadoop is best deployed.

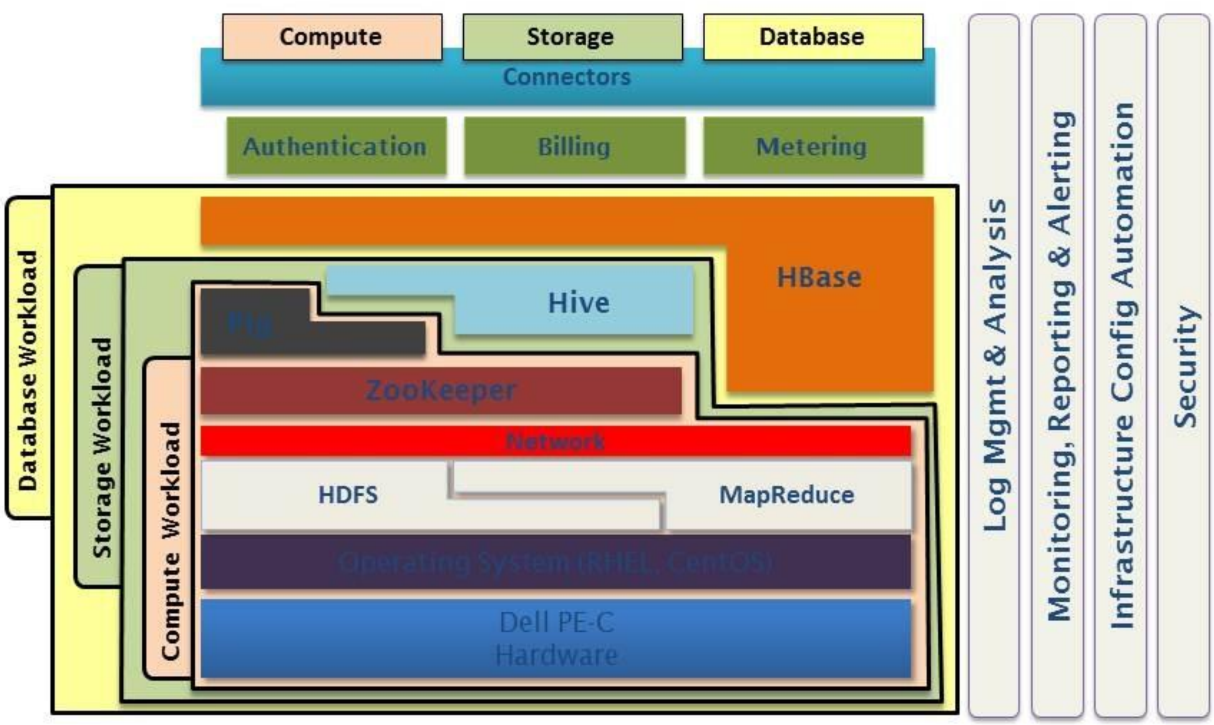

Figure 2. The Dell representation of the Hadoop ecosystem.

Network

Figure 3 depicts a common network architecture for the top-of-rack (ToR) switches within a Hadoop environment. These connect directly to the DataNodes and allow for all inter-node communication within the Hadoop environment. Hadoop networks should utilize inexpensive components that are employed in a way that maximizes performance for DataNode communication. The Hadoop software has many features to ensure availability in the event of hardware failure. This saves the common costs associated with expensive chassis-based switches with redundant power supplies and controllers. Some costs can be saved in the environment, with no impact to performance, by utilizing common ToR switches for all node connectivity. An example of such a switch is the Dell™ PowerConnect™ 6248.

This recommended architecture ensures connectivity in the event of a switch failure by creating link aggregation groups (LAGs) between the DataNodes and two separate switches, eliminating the potential for a single switch failure causing a loss of connectivity.

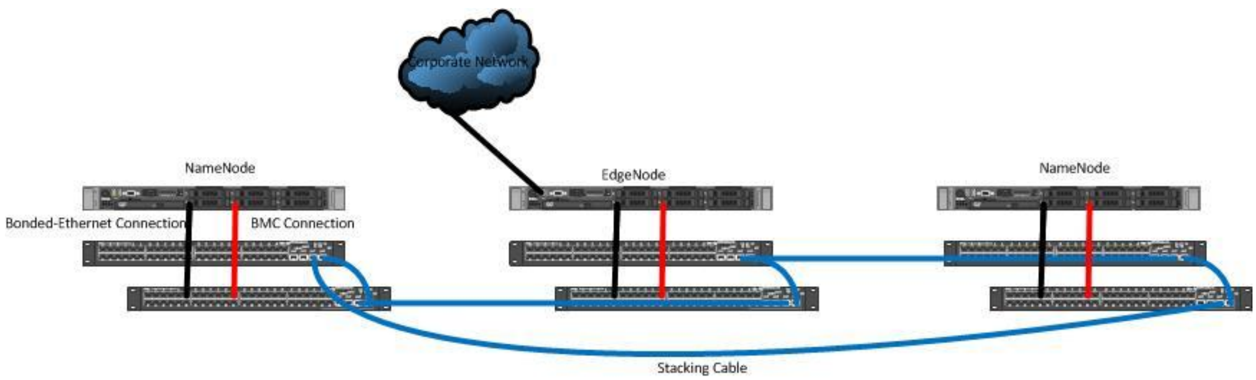

Figure 3. A common network architecture for the top-of-rack switches within a Hadoop environment.

Hadoop is a highly scalable software platform that requires a network designed with the same scalability in mind. To ensure maximum scalability, Dell recommends a network architecture that allows users to start with small Hadoop configurations and grow those over time by adding components, without requiring a rework of the existing environment. Figure 4 depicts a common EoR switch architecture for Hadoop to allow for maximum bandwidth between nodes, while scaling to sizes of environments commonly seen with Hadoop.

Redundancy for the network is most commonly implemented at the end-of-row (EoR) tier of the network. Dell recommends putting two switches in a redundant configuration at the EoR tier and connecting them redundantly to all ToR switches. This will ensure maximum bandwidth between all switches as the environment grows, as well as availability from a single switch failure anywhere in the fabric.

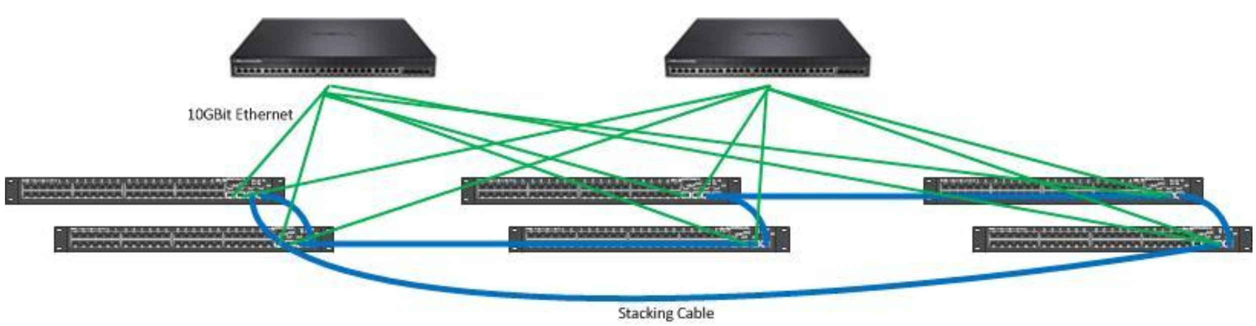

Figure 4. A common ToR switch architecture for Hadoop.

Hadoop performance benchmarks

As with all distributed transaction systems, performance is a primary concern as the environment scales. One method for tracking performance is through the use of common benchmarks for comparing your environment to other similar environments, as well as comparing your environment to itself before and after changes. Within the Hadoop software ecosystem, there are several benchmark tools included that can be used for these comparisons.

Teragen

Teragen is a utility included with Hadoop for use when creating data sets that will be used by Terasort. Teragen utilizes the parallel framework within Hadoop to quickly create large data sets that can be manipulated. The time to create a given data set is an important point when tracking performance of a Hadoop environment.

```
# bin/hadoop jar hadoop-*-examples.jar teragen 10000000000 in-dir
```

Terasort

Terasort is a compute-intensive operation that utilizes the Teragen output as the Terasort input. Terasort will read the data created by Teragen into the system's physical memory and then sort it and write it back out to the HDFS. Terasort will exercise all portions of the Hadoop environment during these operations.

```
# bin/hadoop jar hadoop-*-examples.jar terasort in-dir out-dir
```

Teravalidate

Teravalidate is used to ensure the data produced by Terasort is accurate. It will run across the Terasort output data and verify all data is properly sorted, with no errors produced, and let the user know the status of the results.

```
# bin/hadoop jar hadoop-*-examples.jar teravalidate out-dir report-dir
```

Dell | Hadoop best practices

Dell documents all the facility requirements, including space, weight, power, cooling, and cabling for Dell's defined reference architectures and provides this information to customers prior to any system purchase. This information allows you to make informed decisions regarding the best placement of your Hadoop solutions and the ongoing operational costs associated with them.

The Dell |Cloudera Hadoop solution addresses the challenges discussed in this white paper in a variety of ways to ensure streamlined deployment of low-risk solutions:

- The Dell |Hadoop solution includes the tools necessary to manage the environment's configuration from a single location and monitor the environment for unexpected changes.
- As part of the Dell | Hadoop solution, Dell provides reference architectures and associated benchmarks to streamline the Hadoop design process and minimize the risk to performance bottlenecks.
- This information allows you to make informed decisions regarding the best placement of your Hadoop solution and the ongoing operational costs associated with the solution.
- Dell provides results from lab tests and recommendations for Hadoop tunable parameters for your use in streamlining the time necessary to take a Hadoop installation from bare hardware to a production, operational environment.

Dell has worked with Cloudera to design, test, and support an integrated solution of hardware and software for implementing the Hadoop ecosystem. This solution has been engineered and validated to work together and provide known performance parameters and deployment methods. Dell recommends that you utilize known hardware and software solutions when deploying Hadoop to ensure low-risk deployments and minimal compatibility issues. Dell's solution ensures that you get maximum performance with minimal testing prior to purchase and minimal risk with solution deployment.

Supported solutions

Dell recommends that you purchase and maintain support for the entire ecosystem of your Hadoop solution. Today's solutions are complex combinations of components that require upgrades as new software becomes available and assistance when staff is working on new parts of the solutions. The Dell | Cloudera Hadoop solution provides a full line of services, including deployment, hardware support, and software support, so you always have a primary contact for assistance to ensure maximum availability and stability of your Hadoop environment.

About the author

Joey Jablonski is a principal solution architect with Dell's Data Center Solutions team. Joey works to define and implement Dell's solutions for big data, including solutions based on Apache Hadoop. Joey has spent more than 10 years working in high performance computing, with an emphasis on interconnects, including Infiniband and parallel file systems. Joey has led technical solution design and implementation at Sun Microsystems and Hewlett-Packard, as well as consulted for customers, including Sandia National Laboratories, BP, ExxonMobil, E*Trade, Juelich Supercomputing Centre, and Clumeq.

Special thanks

The author extends special thanks to:

- Aurelian Dumitru, Senior Cloud Solutions Architect, Dell
- Rebecca Brenton, Manager, Cloud Software Alliances, Dell
- Scott Jensen, Director, Cloud Solutions Software Engineering, Dell

About Dell Next Generation Computing Solutions

When cloud computing is the core of your business and its efficiency and vitality underpin your success, the Dell Next Generation Computing Solutions are Dell's response to your unique needs. We understand your challenges—from compute and power density to global scaling and environmental impact. Dell has the knowledge and expertise to tune your company's "factory" for maximum performance and efficiency.

Dell Next Generation Computing Solutions provide operational models backed by unique product solutions to meet the needs of companies at all stages of their lifecycles. Solutions are designed to meet the needs of small startups while allowing scalability as your company grows.

Deployment and support are tailored to your unique operational requirements. Dell Cloud Computing Solutions can help you minimize the tangible operating costs that have hyperscale impact on your business results.

References

Google MapReduce
http://labs.google.com/papers/mapreduce.html

Google File System
http://labs.google.com/papers/gfs.html

Apache Hadoop
http://hadoop.apache.org/

Cloudera Hadoop
http://www.cloudera.com/products-services/enterprise/

Teragen
http://hadoop.apache.org/common/docs/r0.20.2/api/org/apache/hadoop/examples/terasort/TeraGen.html

Terasort
http://hadoop.apache.org/common/docs/r0.20.2/api/org/apache/hadoop/examples/terasort/TeraSort.html

Teravalidate
http://hadoop.apache.org/common/docs/r0.20.2/api/org/apache/hadoop/examples/terasort/TeraValidate.html

Hadoop Distributed File System
http://hadoop.apache.org/hdfs/

To learn more

To learn more about Dell cloud solutions, contact your Dell representative or visit:

www.dell.com/cloud