Technology Insight Paper

Hadoop from Sandbox to Production

Using IBM Spectrum Scale to enhance and harden Hadoop's storage environment

By John Webster

December, 2015

Evaluator Group

Enabling you to make the best technology decisions

Evaluator Group

What "In Production" Means for Hadoop

Reaching production status is now increasingly important for organizations counting on Hadoop to enable them to deliver a new class of analytics capabilities based on Big Data technologies. The reason for sensitivity to this issue is simple. If an enterprise can't transform its interest in Hadoop from shiny new toy to at least mission-sensitive if not mission-critical application status, then why bother playing with it any longer.

In the case of Hadoop, "in production" means different things to different enterprise users. In early releases, the fact that Hadoop's Name Node represented a single point of failure was a known problem that could be dealt with by it early adopters in the webscale community (Google, Facebook, Twitter, Yahoo, etc.). Not so by others in financial services and telecommunications for example who need to see a fix before proceeding further. Other criteria for moving Hadoop from test bed to production status include:

- Readiness to support business-critical applications and their dependent applications
- Data consistency and integrity
- Performance and efficiency at scale
- Conformance to security standards and regulatory requirements
- Availability of qualified IT staff to manage the platform

The Evaluator Group believes that storage has a significant role to play in moving Hadoop from the playground to the enterprise data center. Four of the five criteria above can be addressed with storage-based functions and processes.

Apache Hadoop creates its own storage environment, instantiated the Hadoop Distributed file System (HDFS) features and functions, and the Apache open source community works continuously to improve HDFS. However, with multiple community projects attempting to create and build-in the needed functionality, users are often kept waiting for production-ready functionality to be delivered in future releases.

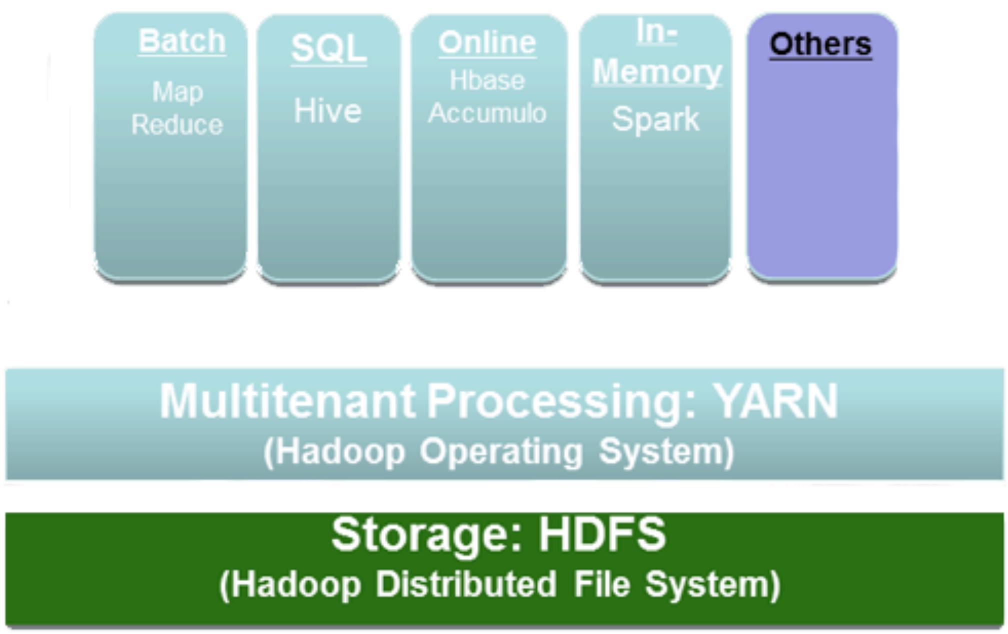
Seeking an alternative to waiting, IT administrators have begun to integrate a data-center grade storage system with Hadoop—one that comes with the required data consistency, integrity, security and governance features built-in. Here we look at one such Hadoop storage deployment option—IBM's

Spectrum Scale as a scale-out, hybrid (solid state plus mechanical disk) storage system integrated with a Hadoop cluster.

Can the Chasm be Crossed?

Hadoop originated from the Internet data centers of Google and Yahoo!. As such, it was designed to deliver high analytic performance at large scale and at low cost. However, there is a "chasm" between large Internet data centers and enterprise data centers defined by differences in management style, spending priorities, compliance and risk-avoidance profiles. Hadoop's HDFS storage environment was not originally designed for long-term data persistence. The assumption was that data would be loaded into a distributed cluster for MapReduce batch processing jobs and then unloaded. The process would be repeated for successive jobs.

Now, enterprise users not only want to run successive MapReduce jobs, they also want to build multiple applications for multiple types of analytics users on top of HDFS (see graphic below). These include OLTP (Hbase) and real time analytics (Storm and Spark). To do this, data needs to be persisted, protected and secured for multiple user groups.



Because HDFS was not originally designed to persist data in ways that would support multiple user-facing applications, it suffers from the following deficiencies in an enterprise data center setting:

Inefficient and inadequate data protection and disaster recovery capabilities

HDFS relies on the creation of replicated data copies (usually three) at ingest to recover from disk failures, data loss scenarios, loss of connectivity and related outages. While this process does allow a cluster to tolerate disk failure and replacement without an outage, it slows data ingest operations, negatively impacts time to information, and still doesn't totally cover data loss scenarios that include data corruption. It also makes for very inefficient use of storage media—a critical concern when users wish to persist data in the cluster for up to seven years as may be required for regulatory compliance reasons. The Apache Hadoop development community is looking at implementing erasure coding for data protection and recovery in Hadoop as an alternative, but this will likely require a replacement of HDFS with a new storage environment.

The ability to replicate data synchronously between Hadoop clusters does not currently exist in HDFS. Synchronous replication can be a critical requirement for supporting production-level disaster recovery operations. While asynchronous replication is supported, it is open to the creation of file inconsistencies across local/remote cluster replicas over time.

Inability to dis-aggregate storage resources from compute resources

HDFS binds compute and storage together to minimize the "distance" between processing and data for performance at scale. However, this results in some unintended consequences for when HDFS is used as a long-term persistent storage environment. To add storage capacity in the form of data nodes, an administrator has to add processing and networking resources as well, whether or not they are needed. And remember that 1 TB of usable storage equates to 3 TB after the copies are made.

This tight binding of compute and storage also limits an administrator's ability to apply automated storage tiering to take advantage of hybrid SSD/rotating disk architectures. It is well understood by administrators that the price of high-performance flash storage media is dropping on an almost daily basis. They could leverage flash to make up for and even improve upon any loss in performance resulting from compute/storage disaggregation and gain a much more efficient way to save less frequently accessed data within the cluster for years.

Data in/out processes can take longer than the actual query process

One of the major advantages of using Hadoop for analytics applications lies in its ability to run queries against very large volumes of unstructured data. For that reason, Hadoop is often positioned as a "Big Data Lake." The idea is to copy data from active data stores move the copies to the data lake. This process can be time consuming and network resource-intensive depending of the amount of data. But perhaps more critically from the standpoint of Hadoop in production, it can lead to data inconsistencies causing application users to question whether or not they are querying a "single source of the truth."

One way to solve this problem is to run multiple applications producing the data on the same storage system, eliminating the need to create, track, and move data copies over a network. This is somewhat anticipated in the above diagram which shows OLTP applications running in Hbase using the same data store (HDFS) as MapReduce analytics applications. Using an alternative, multipurpose storage environment that offers many use cases simultaneously is another that has the further advantage of not requiring modification of the transactional data architecture on which an enterprise may be dependent. Creating a shared cluster-external storage environment also allows IT administrators to disaggregate compute with storage so that the two can be scaled separately and offers the further advantage of not requiring modification of the OLTP architecture on which an enterprise may be dependent.

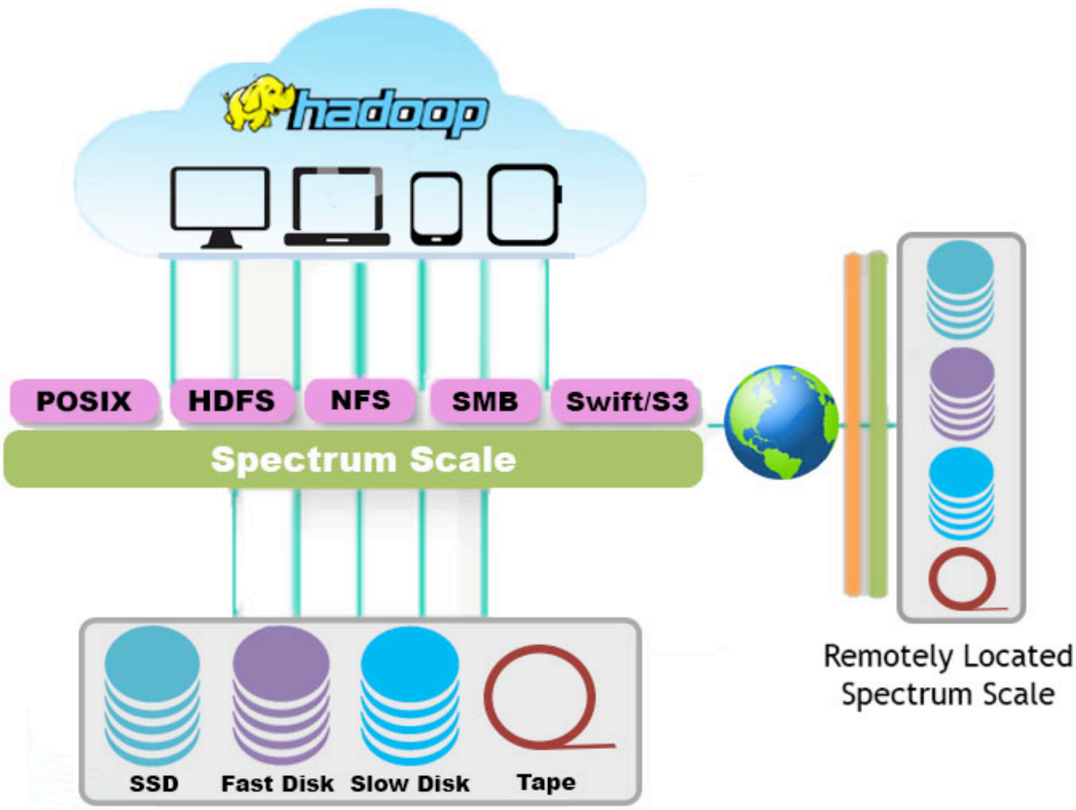
It's complicated

To solve some of the production inhibiting shortcomings of HDFS, the community often creates add-on projects to address deficiencies. RAFT for example can be used to recover from failures without re-computation. DistCp can be used for periodic synchronization of clusters across WAN distances but requires manual processes to reconcile differences when inconsistencies occur, as they will over time. Falcon addresses data lifecycle and management and Ranger centralizes security administration. However, all of them have to be learned and managed as separate entities. Each has a lifecycle of its own that requires tracking, updating and administering. The HDFS environment precludes running many of the common file commands, like copy, which increases the learning curve and opens up the system to human error. Enterprise Hadoop administrators will naturally gravitate to simplicity in this regard. Using a storage environment that has already built-in features, which these add-on projects try to address individually, simplifies management and reduces opportunities for error.

IBM Spectrum Scale

As mentioned earlier, the storage environment plays a significant role in determining whether or not Hadoop can be moved from proof-of-concept to production, and that integrating a data-center grade storage system with Hadoop can remove road-blocks. One such system is IBM's Spectrum Scale

IBM Spectrum Scale is a scalable (to multi-PB range), high performance storage system that can be integrated natively with a Hadoop cluster in a way that replaces HDFS without requiring modifications to Hadoop applications.



Spectrum Scale's major attributes as a persistent storage environment for production-level Hadoop data repositories include:

Unified storage environment—support for both file and object-based data storage with Swift and S3 support

Simplified GUI management for the storage environment that includes automated resource provisioning and storage system performance monitoring

Single global name space to support small to large-scale Hadoop deployments within a single Spectrum Scale environment

Automated storage tiering using solid state flash storage (for performance) and multi-terabyte mechanical disk (for inexpensive capacity) with automated, policy-driven data movement between storage tiers. Tape is also supported as an additional archival storage tier.

Solid state drives can also be used as a high performance storage space for metadata (directories, inodes) and for read and write caching. Read and write caching that is consistent across storage cluster nodes also can be enabled to enhance performance for data intensive applications.

Data access methods include POSIX, NFS, SMB, S3, and Swift to maximize compatibility with existing applications.

Non-disruptive operation is supported for active file management and file placement optimization at the global namespace level without requiring a system outage. Rolling upgrades without a system outage are also supported.

Policy-driven data compression (i.e. which files, when and how controlled by the system administrator) can be implemented on a per file basis for an approximately 2x improvement in storage efficiency and reduced processing load on Hadoop cluster nodes. Compression also makes more efficient use of cache.

Storage-based security features include data at rest encryption as an option and secure erase as well as LDAP/AD for authentication. Authentication and authorization via Active Directory or other LDAP source is also supported.

Snapshots at the file system or file set level and backup to an external storage target (backup appliance and/or tape) are also supported.

Synchronous and asynchronous data replication at LAN, MAN and WAN distances with transactional consistency

IBM zSystem support for users looking to integrate IBM zSystem mainframe data with Hadoop

Spectrum scale is offered to customers via three deployment models:

- Software-only for deployment on industry standard servers
- Elastic Storage Server—a fully integrated hardware/software storage system
- As a cloud-based service

Evaluator Group Assessment:

Hadoop visibility is increasing within the enterprise. Pressure builds from business groups that want to do new kinds of data analytics—like sentiment analysis for example—and have heard that their competitors are doing it on Hadoop. As a result, it is now common to find Hadoop in "shadow IT" where data is being used in ways that would compromise security and data

governance policies and even fail an outside regulatory compliance audit. For this and other reasons, enterprise IT is now being asked to take these Hadoop instances under their wing.

To do this, enterprise IT must feel comfortable in managing Hadoop as a platform for production analytics applications. Enterprise-grade storage platforms that are tailored to replace the HDFS storage environment can address many of their requirements. Here we have shown that IBM's Spectrum Scale can be used to simplify, protect, and manage Hadoop data resources in ways that are familiar to enterprise IT and conform to their existing policies and practices for managing critical and secondary applications that support the day-to-day functioning of the organization.

About Evaluator Group

*Evaluator Group Inc. is dedicated to helping **IT professionals** and vendors create and implement strategies that make the most of the value of their storage and digital information. Evaluator Group services deliver **in-depth, unbiased analysis** on storage architectures, infrastructures and management for IT professionals. Since 1997 Evaluator Group has provided services for thousands of end users and vendor professionals through product and market evaluations, competitive analysis and **education**. **www.evaluatorgroup.com** Follow us on Twitter @evaluator_group*