

Hadoop Storage: DAS vs. Shared



by George Crump, Lead Analyst

Hadoop is a software solution that was developed to solve the challenge of doing a very rapid analysis of vast, often disparate data sets. Also known as big data, the results of these analytics, especially when produced quickly, can significantly improve an organization's ability to solve problems, create new products and to cure diseases. One of the key tenets of Hadoop is to bring the compute to the storage instead of the storage to the compute. The fundamental belief is that the network in-between compute and storage is too slow, impacting time to results.

The Legacy Hadoop Architecture

Hadoop was originally designed to leverage multiple compute nodes. Each of these nodes has its direct attached storage (DAS) which has to be large enough to store the data of any assigned analytics job. One of the nodes in a Hadoop cluster, called the master node, is in charge of coordinating all these nodes. It also maintains a large amount of meta-data so that it understands which data is where on the cluster and what node is the most available. The master node assigns the analysis (Map function) of an analytics request to a node within its cluster that has the appropriate data on its internal storage. When the job is complete, the result (The reduce function) is sent back to the master node and given to the requisition user.

The concept behind this architecture is it can be assembled relatively inexpensively, with standard off-the-shelf hard disk and flash drives and no expensive networking requirements. But is DAS a cheaper alternative after calculating the actual total costs of its use?

The True Cost of DAS

The first challenge that a Hadoop architecture using DAS has to overcome is data protection. How is data going to be protected when it is locally attached? If the node fails or its storage fails, then until repaired, all the data on that node is no longer accessible. To overcome this problem, Hadoop uses a simple replication technology, typically making two extra copies.

Multiple copies give Hadoop DAS two benefits. First, the data is now protected on more than one node, and second, jobs needing access to that data can be started on any one of these nodes.

While some will make a big issue out of replication's 3X capacity requirement, given the cost per GB and density of an 8TB drive this is not much of a concern. What is questionable though, is the replication process itself. Hadoop data sets are large, and this means massive replication across a network connection that is likely being used for other tasks, namely internode communication and the passing of query results. To maintain data confidence means that this data needs to be replicated very quickly at very high speeds. Finally, if there is a node failure, the two other nodes are used to recreate quickly the data that was on the fallen node on another node.

All of these issues may lead a Hadoop environment to implement a more expensive, better quality, dedicated network for replication jobs. In other words, the cost advantage of a "cheap network" is quickly lost as the Hadoop environment matures.

Instead, this same investment could be made on a storage network so that a shared storage system could provide shared data access. A shared storage system would handle data protection on its own, leaving the compute power of the Hadoop nodes dedicated for processing. Also, EVERY node in the cluster would now have access to the data set, meaning the master node could assign jobs to any node in the cluster.

Another added cost to leveraging DAS is the extra expense required to protect the Master node. The master node holds all the meta-data for the group, so losing this node or its storage could be a time-consuming recovery effort and no analytics jobs could be run until this system is back online.

In DAS environments, this means ongoing copying, often manually, the data on this node to another standby node, or implementing a more expensive clustering capability. Shared storage addresses this by storing the master node's data on the same shared storage device as the rest of the environment. Storing that data on the shared storage system provides two benefits. First, that data is automatically replicated and protected. Second, if the master fails, any node in the cluster can access the meta-data and become the new master.

The Data Lake Challenge

Another significant challenge for Hadoop solutions that count on DAS storage is getting data into the environment in the first place. A Hadoop cluster runs the Hadoop Distributed File System as a default. The data that this cluster processes will often come from remote sensors or devices that usually transmit data via NFS or even CIFS. The inability to natively store data into the Hadoop cluster forces many data centers to create a data lake that serves as the central repository. This data is then manually copied to the Hadoop cluster when needed. Multiple systems reduce data efficiency, and once again require a robust network to move data rapidly to the Hadoop cluster.

A shared Hadoop storage system should support multiple protocols like CIFS, NFS, Object, and S3. Multi-protocol support allows it to serve as both a receiver of data from various sources as well as the shared repository for the Hadoop cluster. A single shared system results in reduced storage costs and greater efficiency. Netflix, for example, has evangelized this approach by using S3 to form its "cloud storage" data lake.

Requirements for Hadoop Shared Storage

The justification for using Hadoop in a shared storage environment is compelling, but that storage system should meet certain requirements. One of those is that it be Hadoop compatible. Compatibility can be achieved via native HDFS support or by S3, which Hadoop supports natively. Also, it should provide multi-protocol access beyond S3, including CIFS, NFS, Object and even block access via iSCSI.

The storage system should also be able to balance the demands of performance for analysis and cost effective storage for the data lake. Achieving this balance means typically a mixture of high-speed storage devices, potentially flash, and high capacity storage devices, like 8TB hard disk drives. It also means that the system should support data efficiency techniques like deduplication and compression. Ideally, the system would also support incremental scalability – meaning as new flash and higher capacity drives emerge, then they can simply be added as new nodes to the system. The goal is to ride the economics curve as closely as possible.

Finally, the system should be software focused, leveraging commodity hardware either supplied by the vendor or the customer. The architecture should be similar to the Hadoop cluster; it should scale-out so that it can meet the performance and capacity demands of both the analytic environment and the data lake.

Conclusion

The cost advantages of using DAS in a Hadoop environment quickly erode as that environment moves out of the test phase and into production. As the environment scales and the data center begins to count on it, the extra expense required to maintain performance, reliability and scalability often end up being more expensive than the shared storage alternative.

Shared storage, once the investment in a high-speed network is complete, has many compelling advantages for both Hadoop and the rest of the data center. The multi-faceted capabilities of a scale-out distributed storage software more than offset the initial investment in a high-performance network.

Sponsored By Hedvig

About Hedvig

[Hedvig](#) is a software defined storage system that takes a distributed systems approach to solving storage challenges. The software leverages commodity hardware to create a scale-out storage architecture that provides block, file and object storage services and complete enterprise storage capabilities. You can learn more about Hedvig in [our briefing note](#) or by visiting their [web site directly](#).