

The DAS vs NAS Debate For Apache Hadoop

August 10, 2016 Jagdish Mirani (<https://content.pivotal.io/authors/jagdish-mirani>)



If you want to spin up an **Apache Hadoop**[®] (<http://hadoop.apache.org/>)

cluster, you need to grapple with the question of how to attach your disks. Historically, this decision has favored direct attached storage (DAS). This approach is in keeping with the fundamental Hadoop principle of moving processing to a where the data lives, thereby taking advantage of disk locality to optimize performance. Disk locality is so core to Hadoop that virtually any description of Hadoop starts with this.

The alternative is to use network attached storage (NAS). In contrast to DAS, NAS separates the compute and storage layers so that storage can be shared across a number of servers by shipping data over the network. Historically, this heavy dependence on the network made NAS an order of magnitude slower. Remember, the state of the art was 1GbE networks, and switches were slower and more expensive. I/O requirements for demanding Hadoop-based applications could only be met by DAS.

Today, 10-40 GbE networks are commonplace, with a line of sight toward even higher performing networks in the near future. On the other hand, disk performance has not kept up with improvements in network, memory, and compute resources. For NAS proponents, technology has advanced and shared storage costs have declined, making NAS a cleaner approach, relegating **disk locality to a position of irrelevance** (http://people.eecs.berkeley.edu/~ganesha/disk-irrelevant_hotos2011.pdf). While this may not have tipped the scale 100% in the direction of NAS, the line has been re-drawn such that old assumptions no longer hold.

Hadoop's Original Tenets And The Case For DAS

With shipping large data over the network a big pain point, the Hadoop framework relies on data partitioning and local processing. This approach was believed to be superior because it gave you the best price-performance.

Fundamental performance advantage: Ardent Hadoop advocates still argue that there is no getting around the fact that putting all the server components together gives you a tremendous performance advantage, and workloads with high input/output operations per second are most easily met by DAS. This, in their view, is an indisputable architectural imperative. Additionally, the trade-off between speed and economy can be balanced by choosing either flash storage for performance alongside spinning disks for economy.

Node-level result filtering: DAS advocates also point to another key performance factor—the speed of the result set that is returned by the database. In DAS configurations, each node filters out unwanted records only returning the required/filtered data set. In a shared disk configuration, all the data has to be moved to the processing unit before any filtering can occur. This can increase the amount of data placed on the network by multiple orders of magnitude causing higher latency than a DAS approach.

Avoid encoding overhead: Like DAS, NAS creates copies of data to prevent data loss. To minimize storage overhead, NAS systems use a system of encoding copies so that the storage overhead is minimized. This can reduce storage overhead to only 40%. What happens, however, when you create a copy or read from it? DAS proponents are quick to point out there is a performance hit from translating data to and from its encoded form.

Commodity infrastructure: DAS can serve up a large number of low cost spindles (inexpensive SATA drives or flash) without complex controllers or proprietary software. SATA hard disk drives and SATA controllers are inexpensive commodities. They are mass market products and are priced to offer the best price-performance. The hardware is low-cost enough to do multi-server replication and use hardware to solve for performance and data protection.

Intermediate data storage efficiencies: DAS advocates are quick to neutralize the ‘NAS needs less hardware’ argument. Even in NAS configurations, MapReduce jobs will need to store intermediate data in local storage. **Apache Hive™**

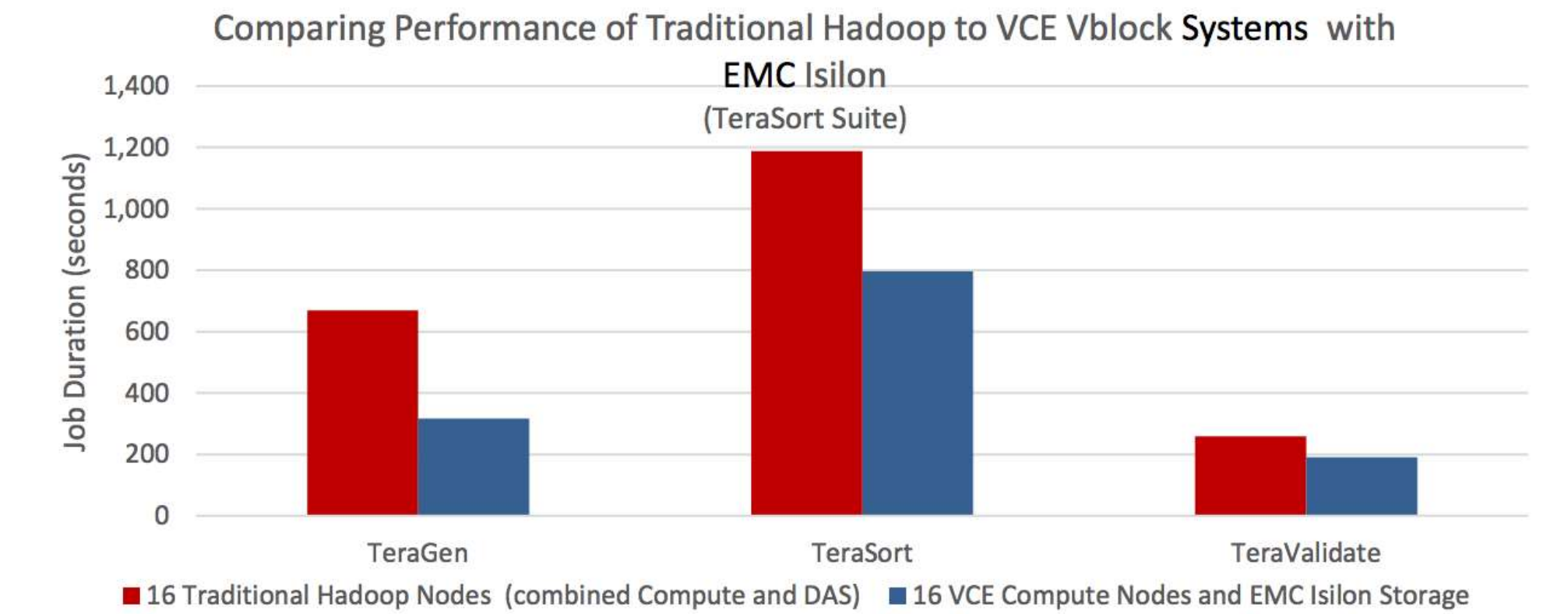
(<https://hive.apache.org/>) and **Apache Pig** (<https://pig.apache.org/>) jobs are translated into MapReduce or **Apache Tez™** (<https://tez.apache.org/>) or **Apache Spark™** (<http://spark.apache.org/>), storing their intermediate data to the local storage as well. So, to be on the safe side, the compute nodes will need to be configured with local storage that equals the amount of raw data you want to store in your cluster anyway. So, based on this argument, NAS will end up needing more storage overall.

The Case For Hadoop With NAS

Initially, the interest in NAS was relegated to a less important role—as archival storage for the primary DAS storage in a Hadoop cluster. However, with advances in I/O bandwidth and switch performance, the tide is turning towards using NAS as the primary storage layer.

Network performance improvements: The crux of the argument against NAS is now under question—can NAS provide comparable performance to DAS? The data suggests that the crossover point has been reached. Network rarely becomes a bottleneck if you have enough 10GbE connectors and the highest performance switches.

The benchmark data bears this out, and there is plenty of it. Here’s an example of a performance benchmark comparison based on three common workloads used for measuring Hadoop performance: TeraGen generates a random data set; TeraSort sorts one terabyte of data produced by TeraGen; TeraValidate validates the sorted output data from TeraSort. The compute nodes with a shared NAS configuration does better than traditional DAS in every test.



Source: Enterprise Storage Group Lab Review: VCE Vblock Systems with EMC Isilon for Enterprise Hadoop

Several other studies corroborate similar performance claims, such as this one:

“While counterintuitive, our experiments prove that using remote storage to make data highly available outperforms local disk HDFS relying on data locality.” **Accenture Technology Labs: Cloud-based Hadoop Deployments: Benefits and Considerations** (https://www.accenture.com/t00010101T000000__w__/jp-ja/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Local/ja-jp/PDF_2/Accenture-Cloud-Based-Hadoop-Deployments-Benefits-and-Considerations.pdf)

See ‘References’ below for links to additional performance data.

Cost of NAS coming down: NAS storage has evolved to use low-cost, readily available hardware. Also, decoupling storage from compute gives you finer grained control over these resources. This will lead to cost savings by preventing you from over-provisioning storage. Also, as noted above, NAS results in less storage for copies of data because copies are encoded to minimize storage overhead. NAS proponents counter the performance hit that results from encoding by pointing out that the performance hit is only encountered when making copies (encoding) or when restoring from a copy (decoding).

Enterprise-grade management: The presence of out-of-the-box enterprise features, like manageability and availability, has typically favored NAS. Tasks like creating dynamic point-in-time snapshots, replication and backups, and ensuring high-availability are known to favor NAS. NAS storage comes with its own storage management utilities that are often already in use, and no specialized storage management software is required. NAS storage is also known to have fewer component failures. Virtualization can neutralize some of these differences by making DAS easier to manage. But again, you can't have the benefits of virtualization and be concerned with managing physical servers—the two are diametrically opposed.

Why Microservices Architecture Favors NAS

Business agility necessitates IT agility, and this consideration favors the use of NAS. The high level of elasticity offered by NAS allows organizations to quickly provision application and storage needs. Storage can be provisioned per application, so that you can choose the appropriate cost/performance characteristics and other features required for each application. Data volumes can grow to petabytes of storage without concern.

The dynamic provisioning of microservices on any hardware makes data locality unattainable. Modern architectures provision microservices as virtual machines or in containers like Docker. Data microservices that store data in NAS can be assigned to individual microservices or containers that need access to the data. Moving workloads will cause a loss of locality unless the system dynamically rebalances to accommodate the movement of the workload, which is impractical and cumbersome for large volumes of data. So, it is difficult to get the performance benefits of disk localization and the portability benefits of virtualization at the same time.

Hadoop does not live in a vacuum, and enterprises need to consider their entire data architecture when making these choices. There are substantial manageability and efficiency benefits from using the same data storage approach to Hadoop that you use for the rest of your data constellation. In fact, microservices will need access to both Hadoop and non-Hadoop data, so uniformity in the data layer can be extremely beneficial. The same set of skills that data center administrators use for managing their other applications can also be leveraged for managing their Hadoop deployments.

Deciding Whether DAS Or NAS Is Right For Your Hadoop Cluster

Technology advancements are being separately pursued for each approach, and both will continue to improve. For locally connected storage, companies like **Nutanix (<http://www.nutanix.com/>)** are promoting the notion of hyperconverged architectures, claiming to squeeze performance out of the system by even more tightly coupling the resources. On the other side, proponents of hyperscale architectures are promoting the emergence of new economics around shared storage based on using low-cost, off-the-shelf computing equipment. The debate is far from over and the approaches are diverging. Still, there are some lines that can be drawn based on workload requirements.

When to use DAS for Hadoop: DAS is still a better fit for a specific set of use cases. If your workload is highly predictable, and maps to how the data is partitioned, then DAS will consistently give you better performance. For instance, if all your tables are partitioned on the same key and map to how you query your data, then all your joins can occur locally and you will get optimal performance. Similarly, if you have a small number of relatively flat tables, so that table joins are limited, then DAS can yield superior performance.

When to use NAS for Hadoop: Deviation from this will cause data to be re-partitioned and shipped across the interconnect, which erodes the benefits of locality. Unless these factors are known in advance and not likely to change, a DAS approach may be too limiting in the long run. With the focus on business and IT agility today, this has to be carefully considered.

EMC has taken the NAS approach for Hadoop to a higher level by tailoring their NAS Isilon storage for Hadoop. Isilon’s in-place data analytics approach allows organizations to eliminate the time and resources required to replicate big data into a separate Hadoop infrastructure. Data can be accessed using standard HDFS protocols without moving data into HDFS. The additional benefits of managing a single version of the data source include minimized data file security risks and better control of the data for data governance. Combined with Pivotal’s HDB Hadoop native SQL database, the solution can give you excellent SQL performance. A complete discussion of this combined solution can be found **here** (<http://www.emc.com/collateral/white-papers/h15168-emc-solution-agile-robust-analytics-hadoop-dl-pivotal.pdf>).

Learn More: Performance Studies

- **Enterprise Storage Group Lab Review: VCE Vblock Systems with EMC Isilon for Enterprise Hadoop**
(<http://www.vce.com/asset/documents/esg-lab-review-vblock-with-isilon-for-hadoop.pdf>)
- **Virtualized Hadoop Performance with VMware vSphere® 6 on High Performance Servers**
(<http://www.vmware.com/files/pdf/techpaper/Virtualized-Hadoop-Performance-with-VMware-vSphere6.pdf>)
- **Accenture Technology Labs: Cloud based Hadoop Deployments**
(https://www.accenture.com/t00010101T000000__w__/jp-ja/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Local/ja-jp/PDF_2/Accenture-Cloud-Based-Hadoop-Deployments-Benefits-and-Considerations.pdf)
- **IDC lab validation brief: EMC Isilon Scale-out Data Lake Foundation** (<https://www.emc.com/collateral/analyst-reports/isd707-ar-idc-isilon-scale-out-datalakefoundation.pdf>)
- **Adobe Case Study – Virtualizing Hadoop in Large Scale Infrastructure** (<https://community.emc.com/docs/DOC-41473>)
- **Virtualizing Hadoop in Large-Scale Infrastructures**
(http://www.snia.org/sites/default/files/JustinMurray_VirtualizingHadoopWorkloads.pdf)
- **Accenture – Where to deploy your Hadoop clusters** (https://www.qubole.com/blog/big-data/accenture-hadoop-comparison-study/?nabe=5695374637924352:0&utm_referrer=https%3A%2F%2Fwww.google.com%2F)

About the Author

Jagdish Mirani is an enterprise software executive with extensive experience in Product Management and Product Marketing. Currently he is the Principal Product Marketing Manager for Pivotal’s in-memory data grid product called GemFire. Prior to Pivotal, Jagdish spent 10 years at Oracle in their Data Warehousing and Business Intelligence groups. More recently, Jag was at AgilOne, a startup in the predictive marketing cloud space. Prior to AgilOne, Jag was at Business Objects (now part of SAP), Actuate (now part o OpenText), and NetSuite (now part of Oracle). Jag holds a B.S. in Electrical Engineering and Computer Science from Santa Clara University and an MBA from the U.C. Berkeley Haas School of Business.



More Content by Jagdish Mirani (<https://content.pivotal.io/authors/jagdish-mirani>)