cloudera

WHITE PAPER

Hadoop and HDFS:

Storage for Next Generation Data Management

Version: Q414-102

Table of Content

The rise of big data and the recognition of the value it might bring to the enterprise demand that modern data management systems accommodate evolving and variable data formats while providing scalable and accessible, yet cost-efficient, storage.

Introduction

Enterprises are rapidly discovering the power of Apache Hadoop and big data to drive more powerful, data-driven insights and better position their organizations in an increasingly dynamic and competitive economy. Hadoop's innovation in scalability, cost efficiency, and flexibility affords enterprises an unprecedented ability to handle greater volume, greater velocity, and greater variety of data than traditional data management technologies. For these information-driven enterprises, this next step in the evolution of data management is an enterprise data hub.

Powered by Hadoop, an enterprise data hub (EDH) offers a single, unified data management system that combines distributed storage and computation, which can expand indefinitely, store any amount of data, of any type, forever, and make that data available to any user. An enterprise data hub allows for powerful, flexible, and simultaneous analytics of that data within the system by bringing a wide range of processing, computation, and query engines and third-party applications directly to the data, thus reversing the traditional data flow of moving data to computing environments and avoiding the issues surrounding movement, duplication, and singular, specialized data silos. Yet, the era of big data poses important questions around how Hadoop and an enterprise data hub fit with previous investments in data management and storage architectures.

This paper examines the role of the underlying storage substrate of Hadoop and an EDH—the Hadoop Distributed File System (HDFS)—as a key enabler for the data management paradigm shift of bringing the application to the data and the importance of this shift and the capabilities of HDFS for today's enterprise. This paper will also cover the design principles of HDFS and provides an overview of the technical components, as well as the features of HDFS that enterprises depend upon for critical Hadoop environments. Lastly, the paper will consider alternative or mixed storage systems in place of HDFS, as well as clarify common misconceptions about HDFS.

Storage for the Modern Enterprise

The foundation of any data management system is the storage system and its associated components. And the dynamics that embody big data—the oft-repeated volume, variety, and velocity qualifiers of big data—warrant that the modern storage system adapt well beyond archiving and straightforward data retrieval and accommodate ever-changing and variable data formats in addition to explosive growth. An example of the challenge presented by big data is the movement of data itself; moving data at scale in a conventional data management environment, where the storage system is separate from the computing engine, is highly constrained by the network between the two systems. As business users seek to employ more data and more computing resources with these increasing amounts of data, the network limitations will only exacerbate this fundamental problem facing the conventional storage and computing architecture.

Enterprises seeking to realize the full potential of an enterprise data hub as the location for all data and the engine for multiple forms of computing need a single, cohesive system that addresses the following considerations.

Co-Designed Integration

Storage and compute facilities should be tightly coupled in order to collapse multiple workloads, which include batch processing, interactive SQL, interactive search, machine learning, iterative data processing, and low-latency data stores, onto a single set of system resources using single, shared data sets to avoid data movement or synchronization. This design allows enterprise IT teams to service multiple business audiences with the right tools, frameworks, insights, and data sets needed for the specific job, yet consolidate operations and maximize efficiency across these elements.

Flexibility

The data management system should be able to store any type of data, including the bulk of big data—unstructured data—while simultaneously employing multiple types of processing and analytic frameworks against the same data, regardless of structure. Enterprise IT teams can capitalize on this versatility for their business communities as data sources and computing needs expand, merge, and change.

Scalability

The data management system should expand from terabytes (TB) to petabytes (PB) simply by adding servers without needing to redesign the data processing, analysis, networking, or storage systems. For example, a large insurance company easily embraced more data, such as complex weather models, to solve new questions about risk exposure, by only growing their enterprise data hub to house the new data. This capability also extends to computational frameworks so that EDH operators can maximize throughput and achieve linear performance gains without incurring network slowdowns and blockages.

Cost Efficiency

Enterprises should be able to leverage more and richer data by storing known, yet low-valued, historical data or data that might have value (i.e. probable value) in cost effective, yet accessible archives that maintain the native format and full details of the data. Such a storage system greatly enhances exploration and examination and is particularly conducive to discovering trends and anomalies that might otherwise be lost due to normalization or sampling requirements imposed by more traditional approaches. This capability translates to significant opportunities to gain value from data that otherwise is shunted to offline archives (and thus is inaccessible for analysis) or simply deleted due to a lack of a clear, or high, return-on-investment (ROI) for storage.

Elasticity

The data management system needs to handle the increase in user communities and the decrease in service level agreement (SLA) windows and time-to-value requirements that result from globalization, ecommerce and competitive dynamics, and the growing awareness of data as a valuable business asset. Enterprise IT teams need to scale computing and storage without sacrificing delivery speed and agility.

Availability

Enterprise IT administrators require uninterrupted business continuity capabilities in the face of software and hardware failures. As user communities diversify both in terms of size and business objectives, so do the associated business services and data sources, and these twin forces present challenges to enterprise IT teams when meeting SLAs and recovery time objectives (RTO), especially as these services and the data within an EDH become more critical to these audiences.

Self-Healing

The data management system needs to adjust automatically to prolonged machine failures in order to maintain business SLAs at lower administrative total cost of ownership (TCO). IT operators anticipate a degree of unexpected failures within their ecosystem, especially within a distributed data management environment where a combination of multiple servers, services, roles, and physical components are working in concert, yet these events should not impact operations or incur significant overhead or costs.

Security

The data management system must protect data from unauthorized access. Enterprise user communities often seek out new data sets and insights for inclusion for reporting and value creation, and enterprise IT teams need systems that enable the commingling of sensitive and non-sensitive data and data sets to promote efficient and secure sharing, yet still prevent unsanctioned data use due to misguided or malevolent activity.

These key requirements and features have led many IT teams to reexamine their traditional practices of data management and consider alternative approaches and architectures.

The Challenges of Big Data

Enterprise IT teams have made attempts in the past to redesign their traditional data management architectures to meet these new requirements. These architectures often consist of independent datavsystems, where each system houses a purpose-built data silo catering to a particular set of workloads.vThese data silos are then connected by extract-transform-load (ETL) data pipelines that copy (i.e. move)vdata between them. This approach has proven bearable for traditional, pre-planned data sets as ITvoperators are able to forecast the effort of change and resource usage within these boundaries, yetvencounters a number of challenges when faced with the new, larger, and changing data sets representedvby big data.

Scalability

Traditional architectures often face scalability issues due to the replication of data between systems. Eachvsizable increase in data volume requires corresponding sizable adjustments to network and processing invthe ETL pipeline to maintain SLAs for performance and reliability, and these links in the pipeline quickly becomevimpediments to successful operations. This becomes increasingly untenable as systems shift fromvhandling a few TBs toward many PBs of information, let alone the growing number of analytical systemsvthat IT teams must and thus become additional replication targets to satisfy today's information-drivenvbusiness communities.

Locality

The separation of storage from compute is a key challenge for traditional approaches. Even after data is ingested into a storage system, the data processing or analytics that uses the shared storage system must access the data remotely. Again, as data sizes grow from a few TBs toward PBs of information for a particular task, the interconnect between the processing framework and the storage system increasingly becomes the bottleneck. after data is ingested into a storage system, the data processing or analytics that uses the shared

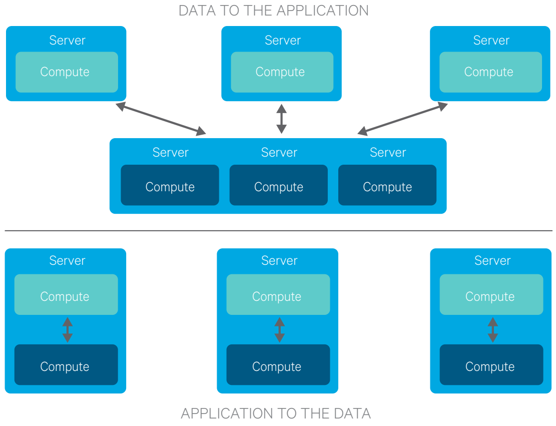

Figure 1: Data Locality

storage system must access the data remotely. Again, as data sizes grow from a few TBs toward PBs of information for a particular task, the interconnect between the processing framework and the storage system increasingly becomes the bottleneck.

Flexibility

The flexibility—rather the lack of flexibility—of schemas in conventional storage systems, notably the relational database, is an impediment to big data adoption for many enterprises. Often the elements of the data pipeline are executed in a traditional database system that requires the developers to predefine the data structures prior to loading data into the structures for subsequent processing. As organizations look to leverage more, changing, and different types of data, the cost of modeling and altering schemas in these systems becomes prohibitively burdensome.

From a time-to-implement perspective, changes to schemas commonly impact the upstream and downstream stages of the processing pipeline, and IT teams often mitigate these far-reaching effects with careful, if not laborious, planning, testing, and deployment cycles. Business users can encounter operational latencies of weeks, if not months, to implement changes to complex data pipelines. This resistance to change affects the ROI of the bytes stored, as business teams must face an equation that balances the value of data against the cost of change - an equation that often pushes business users to avoid introducing less valuable or uncertain valued data into the processing mix due to an unclear rate of return versus the known and high cost of ingestion and inclusion.

Technical Innovation

Enterprise IT operators typically discover that not only are traditional technologies costly for storing massive data volumes, they also inherently pose risks of future proprietary lock-in. Proprietary licensing models place the vendor into an advantageous long-term position as customers must either continue paying the cost to license from that single vendor only or pay the cost to re-architect their solution for a different technology. As data grows in size and usage, IT teams find it increasingly costly to migrate to a different technology both in terms of physically moving the larger data sets as well as the development time.

Ultimately, most enterprise IT teams have concluded that the growth bottlenecks and the shrinking ETL windows in conjunction with the increasing need to analyze a greater variety of data sources mandates a new architecture for processing and analyzing big data.

Data at the Center of the Enterprise

An enterprise data hub, powered by Hadoop and HDFS, is uniquely capable of meeting the challenges of big data by employing a fundamentally different approach to data management than conventional storage and computing architectures. The core tenet of this approach is that Hadoop puts data at the center of the architecture and applications move to the data, which is the reverse of the legacy model of replicating and pipelining data between different computing systems.

Accomplishing this shift in data management requires a comprehensive end-to-end architecture where the HDFS storage system and the Hadoop computing frameworks are tightly co-designed. This strategy provides a number of key capabilities that meet the demands of next generation data management.

Integrated Computing and Storage

HDFS and the Hadoop frameworks are tightly integrated and physically collocated within the same server in a cluster to provide the shortest path between data and computing, which brings accessibility and throughput to any workload with any data within the system. This design thus benefits all workloads that a business user might require for their solution - from batch processing with MapReduce to interactive engines like Apache HBase, Cloudera Impala, and Cloudera Search.

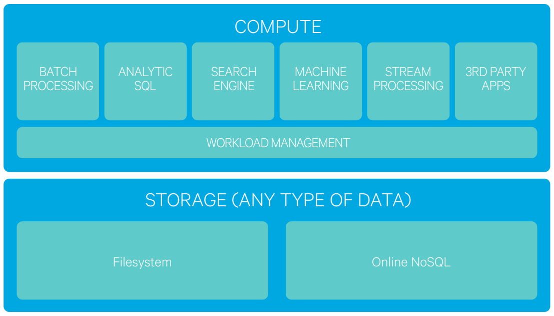

Figure 2: Computing and Storage with Hadoop

Unified and Flexible Storage

The design of HDFS lets IT operators ingest data once into efficient, native, and open formats that allow shared, simultaneous access to that data by all current and future Hadoop processing and analytic frameworks. This design stores data without any pre-defined structure and shifts the projection and interpretation of the data to the moment of processing or analysis by the given framework of choice (i.e. schema-on-read). Business users and processes thus can access any and all data, in its native full fidelity, in whatever framework deemed most appropriate to the task at hand.

This design avoids the issue facing many IT teams when using traditional architectures, where the a priori data structures required for data ingest force the decision of which data to keep and which to discard in order to normalize to these structures, thereby potentially overlooking key information. This design and convergence also avoids the network bottleneck and resource time and effort inherent in legacy architectures that must replicate data between different systems purpose-built for a specific type of data processing or analytics. This flexibility also helps IT teams mitigate the operational impacts of altering schemas within the data processing pipeline, a key benefit as organizations embrace an increasing variety of data types and sources.

Scale and Performance

Hadoop’s collocated compute with distributed, “shared nothing” storage allows IT operators to add servers to a cluster to accommodate data size and usage increases without impacting the day-to-day operations and business SLAs or requiring a redesign of the hardware architecture. This grants the enterprise a straightforward and predictable provisioning model for future capacity and growth.

As mentioned, the Hadoop computing frameworks run distributed across the servers of a cluster and access the data locally and directly on each individual HDFS node. In addition to avoiding the traditional network or storage interconnect bottlenecks, this design serves as the foundation for linear performance by combining the high I/O intrinsic to direct-attached storage (DAS) with HDFS’s distribution and scale-out model -- adding capacity simultaneously adds throughput and computing cycles. For the IT operator, this feature means both capacity and performance characteristics of a Hadoop cluster are direct and predictable, so if a business community requests more demanding SLAs, the IT team can simply add additional servers to the cluster to gain the needed I/O and CPU processing.

Cost Efficiency

HDFS is designed with built-in software scalability, elasticity, flexibility, and availability and thus can run reliably on industry-standard hardware. This design offers the IT administrator options for optimizing hardware expenditures, and when combined with Hadoop's open source economics, costs for a Hadoop-powered enterprise data hub often are an order of magnitude less the cost of traditional architectures. More importantly, building on the foundation of open source avoids the pricing pitfalls of proprietary storage lock-in and the cost of migration. With an open source storage platform, IT teams are not beholden to any single vendor or even required to have a Hadoop vendor at all.

Durability and Security

HDFS provides the backbone for uninterrupted business continuity through built-in software high-availability, snapshots, and data replication facilities. For example, IT administrators can use applications like Cloudera Manager to gain end-to-end execution, management, and monitoring of multi-datacenter replication. HDFS is also designed for fault-tolerance and is self-healing to combat "bit rot" and the sundry outages that can occur at the disk, server, and rack level. For example, HDFS automatically re-replicates data blocks in the event of a prolonged failure to ensure sufficient durability and accessibility in order to maintain business SLAs. Lastly, HDFS (as part of an enterprise data hub) has built in encryption and key management. It also provides industry standard Kerberos authentication and UNIX-style file permissions, and has direct integration with well-known, proven systems like Microsoft's Active Directory. Combined with unified access control from Apache Sentry and data governance from Cloudera Navigator, IT administrators can have a common security model across their network and data management architectures

All of these capabilities ultimately stem from the co-engineering of HDFS with the Hadoop computational frameworks. This single design principle is what uniquely enables Hadoop operators to bring their applications to the data and meet the critical scalability, elasticity, flexibility, availability, and cost efficiency requirements of an enterprise data hub.

The Internals of HDFS

From a technical perspective, HDFS achieve its scale-out capabilities through the interaction of two main components: NameNodes, which handle the management and storage of the file system metadata, such as access permissions, modification time, and data location (i.e. data blocks) and DataNodes, which store and retrieve the data itself on local, direct-attached disks.

The NameNode

In the simplest of terms, when client applications or Hadoop frameworks request access to a file in HDFS, it is the NameNode service that responds with the DataNode locations for the individual data blocks that constitute the requested file. Correspondingly, the NameNode handles activities like new file creation and the locality assignments or metadata changes such as file system hierarchies and file renaming. In order to scale to the thousands of servers deployed in large Hadoop clusters, the NameNode does not participate in the actual read and write flows; the individual DataNodes interact directly with the requestor for all read and write activities.

The HDFS NameNode service is typically deployed on multiple nodes and configured for high availability to ensure uninterrupted business continuity in the event of NameNode software or hardware failures. An individual NameNode has been proven in production to scale to sizes and concurrency greater than any other publicly known commercial file system – companies such as Yahoo and Facebook store over a hundred PBs of data, representing hundreds of millions of files, across thousands of DataNodes. A 10,000 node HDFS cluster with a single NameNode server can expect to handle up to 100,000 concurrent readers and 10,000 concurrent writers. For IT operators who require the even more capacity, HDFS federation can horizontally scale the NameNode infrastructure.

The DataNode

The DataNode service is the real workhorse of HDFS as this service handles all the storage and retrieval of Hadoop data. Once the NameNode retrieves which DataNode or DataNodes contain a given file's data blocks, all the actual data communications stream directly between the client application or framework and the DataNode housing that particular data block. Native Hadoop frameworks capitalize on this direct communication by distributing their workloads (e.g. jobs, queries) across the Hadoop cluster to their collocated computing services so to read the data sets (i.e. the data blocks) locally from each collocated DataNode service. This localized read and computation is the critical element that enables Hadoop to bring its ecosystem of applications directly to the local HDFS data, which in turn avoids the networking bottlenecks that stymie linear scalability in traditional architectures.

Replication

HDFS automatically and transparently replicates data blocks across DataNodes to provide two key capabilities: read scalability and built-in fault tolerance. For the former, HDFS uses data block replicas to parallelize concurrent file access and thus maximize throughput and latency. In this form, data that is more readily available via replication is subject to less resource contention, which is especially important in a multitenant, multi-workload enterprise data hub or a mission-critical application.

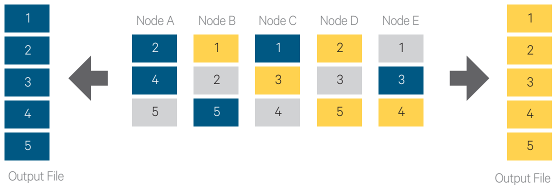

Figure 3: Read Scalability and Data Block Access

Data block redundancy also enables HDFS to stay highly available in the face of multiple software or hardware failures. At scale, the law of large numbers requires HDFS to withstand failures of any single component, including the less probable failures with racks and switches. HDFS's block placement takes rack locality into account to ensure uninterrupted availability of data across the cluster. For widespread or catastrophic failures such as an entire data center or cluster outage, IT operators can employ full-fledged disaster recovery (DR) system tools, like Cloudera Manager, to provide standard recovery capabilities necessary for multi-datacenter operations.

Moving data at scale is a time-consuming effort, even under optimal conditions. Recent achievements in WAN speeds, for example, can transfer 1 TB in roughly 15 minutes, yet a 100 TB data set – not uncommon big data analysis – is transferred in roughly 24 hours.

Traditional Storage with Hadoop

IT operators new to an enterprise data hub and Hadoop often ask whether they should extend their existing traditional enterprise storage systems as their Hadoop storage layer. The answer to this question follows the same theme that motivates the core design principles of Hadoop: consider the best methods for bringing the application to the data. The separation of storage from compute impacts scalability, locality, and elasticity as all data access must occur remotely via the physical interconnect between the storage system and the Hadoop frameworks accessing the data.

For data that already exists in an enterprise storage system, some IT teams may find the effort of migrating this data to Hadoop greater than the disadvantages associated with independent storage and Hadoop systems. An independent storage model is most successful for workloads that are more computationally intensive and the constraints of data volume or rates and network bandwidth are secondary or limited; the data payload costs are minimal. Independent storage is least successful for workloads like interactive SQL analytics, such as Cloudera Impala, or running a NoSQL operational store, like Apache HBase, due to the high rate of data interchange between storage and computing employed by these particular capabilities.

Enterprise IT architects should also consider the management of independent storage and Hadoop systems. Maintaining an incumbent storage system enables IT teams to leverage existing tools and operational experience. Yet, the separation of Hadoop and storage can impose additional operational burdens on these teams as they must now work to coordinate two separate management, auditing, security, backup, and disaster recovery systems.

Misconceptions of HDFS

With all the excitement and interest surrounding the power and application of Hadoop and big data, enterprise IT teams should not be surprised to encounter confusion about the reality of HDFS. While Hadoop and HDFS are now mature and proven foundations for next generation data management, a few common misconceptions persist.

Myth #1: Hadoop and HDFS Are Not Enterprise Ready

The deliberate design of HDFS— distributed, collocated, fault-tolerant, highly available— uniquely enables Hadoop to deliver the critical features needed for modern, enterprise-ready data architectures like an enterprise data hub.

Massive scale and elasticity - Many leading enterprises have deployed Hadoop in production, storing hundreds of PBs and running clusters of thousands of nodes. And all computing frameworks of Hadoop run collocated with this distributed data, thus enabling the direct local reads that achieve the scalability and elasticity needed to match growing data volumes and increased computational demands simply by adding servers to existing clusters, without changes to the architecture.

Native computational frameworks - The same data is simultaneously available for multiple native processing and analytic frameworks from batch processing to interactive analytics, which include a rapidly increasing ecosystem of third party applications.

Unified management - IT operators can easily deploy, manage, and monitor the entire Hadoop system —from the file system (HDFS) to the computational frameworks—with a single management application like Cloudera Manager. IT operators can also control and monitor the dynamic allocation of cluster resources—from compute and memory to file system—with Cloudera Manager, which lets administrators tune cluster usage and capabilities to best serve the business needs and its SLAs.

Unified security - Hadoop security provides strong authentication between all services through direct Kerberos integration with Active Directory and unified authorization capabilities across components through Apache Sentry. Additionally, HDFS has transparent encryption and integration with Navigator Key Trustee, Cloudera's key management solution.

Unified auditing - Data management tools like Cloudera Navigator grant IT and audit teams a single, coordinated system for end-to-end data auditing, lineage, and discovery across Hadoop capabilities, from processing to storage.

Unified point-in-time recovery - HDFS and other Hadoop data applications, like HBase, support snapshots so IT administrators can capture data at a point in time and rollback to previous states if needed. With tools like Cloudera Manager, administrators gain comprehensive management features for snapshots including the coordination of not only the data within HDFS but also the associated metadata, thus enabling IT teams to execute rapid and complete recovery processes.

Built-in availability - The entire Hadoop stack is built to withstand software, node, and rack failures, and this fault tolerance is coordinated from the file system through to the end-client tools and configurable with unified system management applications like Cloudera Manager.

Integration options - HDFS provides flexibility integration options for existing infrastructure, including the ability to mount HDFS via NFS for quick point access. Hadoop also provides native, highly scalable integration frameworks like Apache Flume for streaming and log data sources and Apache Sqoop for relational data sources.

Hadoop and HDFS are the foundations for a wide variety of processing and analytics applications. Cloudera is committed to offering flexibility and choice to our customers in order to meet the needs of their organization. By bringing together a diverse partner ecosystem of software vendors that build directly on and within these foundations, Cloudera is helping customers bring Hadoop and their EDH to more enterprise users and business applications. Cloudera continues to be the industry standard for next-generation enterprise data management and analytics.

Learn more about our software vendors and their applications at the Cloudera Connect partner program. http://www.cloudera.com/content/cloudera/en/partners.html

Myth #2: Hadoop Is Not Secure

With Cloudera's enterprise data hub, Hadoop (including HDFS) features comprehensive security. It supports industry standard Kerberos authentication with direct integration to Active Directory and automated workflows to secure clusters - all through Cloudera Manager. Additionally, Apache Sentry is an integrated component and provides unified authorization and role-based access control across multiple components. Finally, Cloudera Navigator provides enterprise-grade encryption, key management, and unified auditing and lineage for automated data governance.

Myth #3: Separate Data and Compute Does Not Affect Performance

While separating the storage and computing systems allows IT administrators to scale independently storage capacity or computational capacity, this approach is ultimately subject to the quality and performance of the interconnect. Any latency, degradation, or threshold in the network immediately and directly impacts the scale, elasticity, and locality needed by an enterprise data hub to handle growing big data volumes and usage and undermines the computing advantages gained by collocation and moving the application to the data.

Myth #4: HDFS Replication Is Wasteful and Redundant

HDFS leverages data block redundancy for both read scalability and fault tolerance. To increase performance on hotly contested files, IT operators can optionally increase the replication count to provide more accessibility to the data, thus reducing contention and improving overall performance. While traditional solutions like RAID provide fault-tolerance and accessibility, they trade-off data locality and come with traditional storage price tags that are typically 10x or more than the cost of the entire Hadoop stack.

Myth #5: HDFS and the NameNode Has A Single Point of Failure

HDFS is designed with built-in data availability to sustain multiple software or hardware failures, from disks to racks and network switches. Since the release of CDH 4.1 in 2012 and the release of Hadoop 2, the NameNode has had native high availability to eliminate all single points of failure in the system.

Myth #6: HDFS Cannot Handle Small or Many Numbers of Files

HDFS has been proven in production to scale to hundreds of millions of files with a single NameNode. Furthermore, HDFS supports horizontal scaling of the NameNode via federation so that IT operators that need to execute at the largest extremes can confidently support unlimited numbers of files.

Enterprise architects should consider the HDFS best practice of storing files in sizes larger than a single data block in order to maximize the processing efficiency these files. Hadoop computing frameworks gain the advantages of locality by having their distributed local processes collocated with each data block of a file. If the data blocks are too small, the cost of acquiring and processing each data block constituting the file will outweigh the gains achieved with data locality.

Conclusion

The drivers behind the shift in next generation data management—bringing the application to the data and making data the center of the enterprise—positions Hadoop as the natural centerpiece of an enterprise data hub. With an enterprise data hub built with Hadoop, data-driven organizations have a single location to store, process, and analyze all their data, regardless of its type, volume, rate, and retention, using a variety of methods for gaining value from the data, from batch processing to interactive SQL and search. This approach is only possible with a holistic Hadoop architecture and its fully integrated storage and compute frameworks. This key design of Hadoop—collocated storage and compute—lets enterprises meet each data processing need while achieve scale, elasticity, durability, security, and governance demanded by today's big data solutions.

About Cloudera

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data, an enterprise data hub built on Apache® Hadoop™. Cloudera offers enterprises one place to store, process 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. Only Cloudera offers everything needed on a journey to an enterprise data hub, including software for business critical data challenges such as storage, access, management, analysis, security and search. As the leading educator of Hadoop professionals, Cloudera has trained over 40,000 individuals worldwide. Over 800 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. www.cloudera.com.

cloudera.com

1-888-789-1488 or 1-650-362-0488
Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA