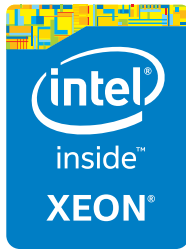# Scale-Out Storage Infrastructure for Apache* Hadoop* Big Data Analytics with Cloudian HyperStore® & Intel®-based Storage Servers

Intel® Xeon® Processor E5-2600 v3 Product Family

Audience and Purpose

For companies looking to build their own cloud storage infrastructure, including enterprise IT organizations and cloud service providers or cloud hosting providers, the decision to use cloud and cloud storage for the delivery of IT services is best made by starting with the knowledge and experience gained from previous work. This white paper gathers into one place the essentials of a scale-out storage reference architecture coupled with a real world example from the Cloudian support organization that is using the Cloudian HyperStore® appliances and the Hortonworks* Hadoop* Data Platform to analyze Big Data logs and troubleshoot customer issues. The reference architecture, based on the Cloudian HyperStore appliances, built on industry standard servers with Intel® Xeon® E5-2600 v3 and Atom™ C2750 series processors, 1GbE and 10GbE Intel® Converged Network Adapters, and Intel® Solid State Drives Data Center Family, creates a multi-node, single data center storage deployment optimized for analytics. The white paper contains details on the topology, hardware, and software-deployed installation and configuration options that should significantly reduce the learning curve for building and operating your first smart data storage platform for analytics.

It should be noted that the creation and operation of a storage cloud requires significant integration and customization to be based on existing IT infrastructure and business requirements. As a result, it is not expected that the reference architecture described in this paper can be used "as-is." For example, adapting to an existing network and identifying management requirements are out of the scope of this paper. Therefore, it is expected that the user of this paper will make adjustments to the design presented in order to meet their specific requirements.

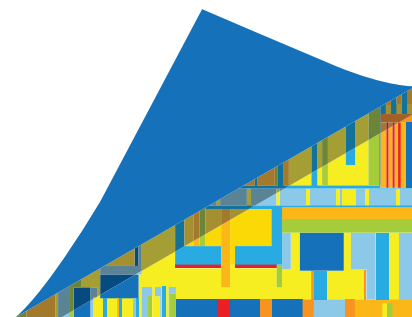This paper also assumes that the reader has basic knowledge of cloud storage infrastructure components and services.

May 2015

Table of Contents

Executive Summary

Cloud storage providers, online applications, and private storage clouds developed by organizations in the media, entertainment, scientific, medical research, and financial services sectors are leading drivers for massive new varieties and capacities of unstructured storage. Applications such as Hadoop are driving a new generation of intelligent data centers with the need for more scalable, durable, and easy to manage storage with very low cost of ownership. To meet these exploding capacity requirements, service providers and enterprises are turning to a software-based scale-out storage infrastructure that combines industry-standard servers and storage components.

A software-based scale-out storage infrastructure gives enterprises the ability to leverage the latest advancements in cost-effective CPU and storage technology. It allows enterprises to keep their environments in lock step with the ever-increasing storage and I/O demands of critical business applications. For scale-out storage architectures, more powerful CPUs lead to greater scale and performance. Intel, for example, typically releases a new generation of CPU every 12-18 months. In addition, disk manufacturers continue to drive innovation into the hard disk drive market space, delivering increased disk drive densities and a lower cost per GB. Cloudian HyperStore software running on the latest off-the-shelf Intel-based servers allows enterprises to take advantage of these technology updates earlier, gaining significant efficiency benefits.

Previously, enterprises looking to deploy storage clouds have had to rely on non-scalable, expensive, proprietary storage, or they've had to invest in configuring and maintaining open-source based storage solutions. Cloudian HyperStore scale-out software brings data protection and scalability improvements to traditional proprietary storage or open-source software-based scale-out storage clouds.

On the other side, many data center administrators want the ease of deploying a turnkey storage offering, a pre-certified storage appliance that takes the guesswork out of configuring the right server and storage combinations. A turnkey system helps to speed up deployment and mitigate the risk of application downtime or performance problems that can occur from misconfigured systems. Rapid deployment and risk mitigation are key for enterprises, where IT staffs are stretched too thin. Also it gives the IT organization a single vendor to support the entire hardware and software stack.

For IT organizations who prefer a turnkey system, Cloudian provides the HyperStore appliance, built on industry-standard and energy-efficient servers using Intel® Xeon® E5-2600 v3 processors, Intel® Ethernet adapters, and Intel® SSDs, pre-loaded with HyperStore software.

This reference architecture defines a scale-out storage solution using HyperStore software running on HyperStore appliances that can be used for several different use cases: File Sync & Share and Remote Office Collaboration, Enterprise Backup and Archiving, Private Cloud Storage, secondary storage for Citrix CloudPlatform* and OpenStack, and Big Data Analytics. This paper focuses on the latter use case: scale-out storage for Big Data analytics.

Introduction

With the popularity of rich media, the proliferation of mobile devices and the digitization of content, there has been an exponential growth in the amount of unstructured data that IT is managing—and this growth is accelerating. This unprecedented growth in unstructured content is simply not sustainable for current NAS/SAN infrastructures. In fact, businesses need to rethink how to manage their whole storage infrastructure. Backups and restores are taking longer. Migrations from older storage systems to new storage systems are labor intensive. Provisioning storage for users is more frequent and time consuming. And the list goes on and on.

Fortunately, Web 2.0 companies such as Facebook*, Google*, and Amazon* faced this same storage challenge a few years back. Cloudian HyperStore software is based on the same principles as these successful companies while addressing enterprises' challenges associated with scalability, ease-of-use and insights generation by built in analytics. Cloudian HyperStore software allows IT to create a very elastic storage infrastructure that can be easily managed and expanded or shrunk based on the demand from applications or end-users. Cloudian HyperStore also provides cost effective Hadoop ready storage. Enterprises can run Hadoop analytics directly on the data in-place at petabyte scale, getting enormous business value and allowing them to drive decision-making and build intelligence into everything they do.

In-place analytics enables enterprises to derive meaningful business intelligence from their data quickly, efficiently and economically. From the financial industry to medical research, there is no shortage of markets that will benefit from turning big data into smart data in pursuit of realizing market- and revenue-shifting insights.

Hadoop, first used by large cloud providers, is designed to allow massive amounts of compute resources to process

very large, unstructured datasets. The larger the data set, the deeper and more accurate the insights. Hadoop does include the Hadoop Distributed File System (HDFS) that employs a Just a Bunch of Disks (JBOD) approach, offering a 'cheap-and-deep' method of storage. However this presents several challenges for traditional enterprises looking to implement it.

1. The first is capacity inefficiency. To protect data, HDFS as a default makes three copies of the datasets.

2. The second challenge is the master node. It contains all the meta-data information for the Hadoop cluster and all that information is typically stored on direct attached storage. Loss of the master node means loss of the cluster and the data.

3. The third challenge is that data must be moved to HDFS over the network to be analyzed, which is a lengthy process for large data sets. Most of the time the data set to be analyzed is created and managed by some other process and stored on a more traditional file or object based storage system.

4. The fourth is scalability and disaster recovery. HDFS can't scale to trillions of data objects and doesn't support cross data center replication.

5. The fifth is the ability to support and be optimized for a wide range of data sizes, from a few kilobytes to tens of megabytes.

6. Lastly the need to provide a REST based API to access and manipulate the data in HDFS.

Adding object storage capabilities and features to the Hadoop infrastructure can address these issues.

The Hadoop plus Cloudian integration enables enterprises to run powerful data analysis, using applications like MapReduce*, Pig*, and Hive* directly on Cloudian HyperStore storage platform. This allows organizations to analyze large data sets without having to move the data. Cloudian's scale-out architecture stores all data in a single namespace, for analyzing data across diverse sources, enabling richer insights to be derived more quickly.

Reference: http://hortonworks.com/partner/cloudian/
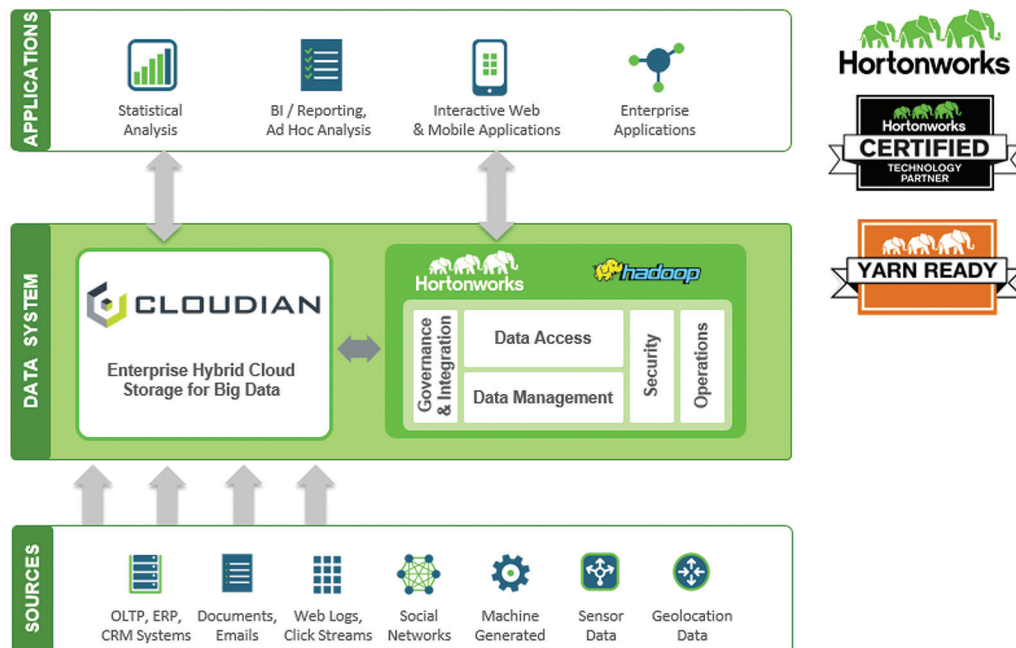
Figure 1: Cloudian Hyperstore And Hortonworks* Architecture

HyperStore Product Overview

Cloudian HyperStore Software

Cloudian HyperStore software delivers a fully Amazon S3* API compliant, multi-tenant, and multi-datacenter hybrid cloud storage solution. Cloud service providers use Cloudian HyperStore software to deploy public clouds and managed private clouds. Enterprises use HyperStore software to deploy private and hybrid clouds.

HyperStore software employs a fully distributed and replicated peer-to-peer architecture with no single point of failure. It easily scales horizontally using commodity hardware so deployments can start with a few servers in a single datacenter and then scale out as usage increases to thousands of servers distributed across multiple datacenters managing hundreds of petabytes of data. Its distributed architecture with automatic replication and recovery services makes it highly resilient to network and node failures without data loss. Similarly, when scaling the storage cluster or performing maintenance, changes in node availability are automatically detected without service interruption.

Features like hybrid cloud streaming, virtual nodes, configurable erasure coding, and data compression and encryption provide highly efficient storage and data management that lets users store and access their data where they want it, when they want it.

Multi–Data Center and Multi–Region Support

Each HyperStore software implementation starts with two or more distributed nodes and then objects are replicated or erasure coded across the available nodes for data durability and availability. Administrators can configure the number of replicas or erasure code strategy required to meet SLA and cost objectives, including the option to replicate copies to other datacenters for geo redundancy. Reads and writes are always performed at the local data center with remote replication performed in the background to avoid latency of remote writes.

HyperStore software supports multiple regions with shared multi-tenant management for administration, users and groups. While groups and users are shared across regions for authentication, bucket names are unique across regions. Data can be placed in specific regions for security, policy, cost or other reasons. Cloudian HyperStore supports thousands of nodes with billions of objects per region and millions of accounts across regions.
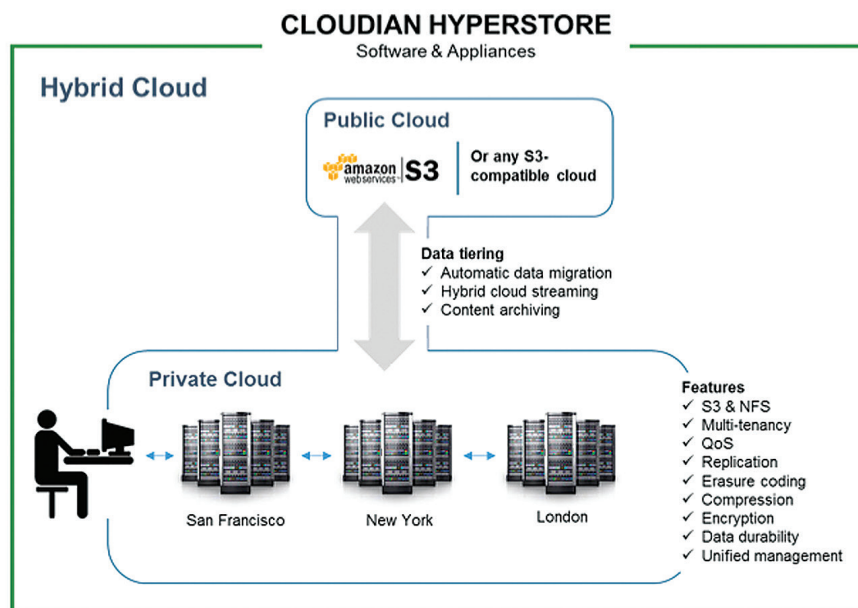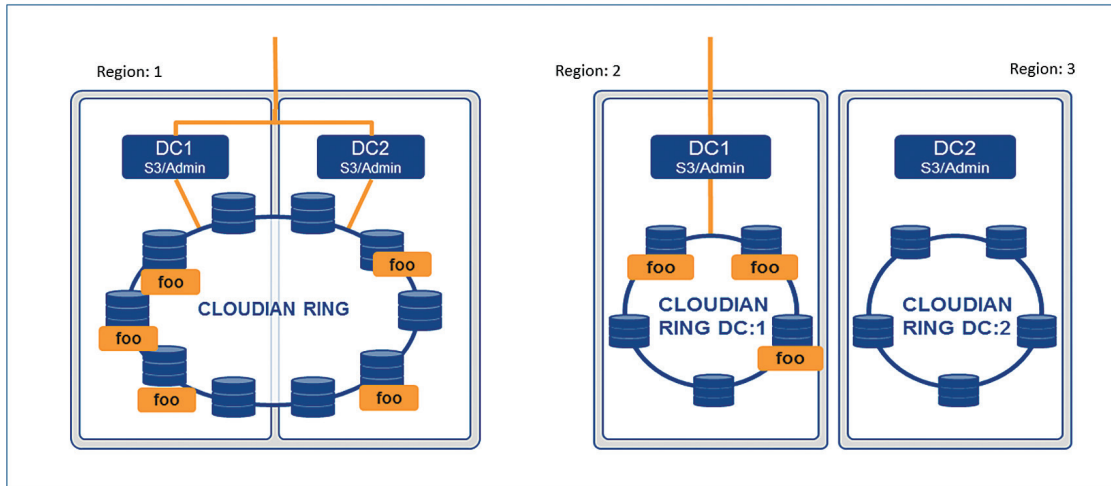

Figure 2: Cloudian Hyperstore Hybrid Cloud

Figure 3: Hyperstore Multi Data Center Topology Example

Configurable Data Consistency

Cloudian HyperStore also provides the ability to configure the level of data consistency. The default consistency requirement for read and write operations is defined as "quorum", meaning that a read or write operation must succeed on a quorum (or set number) of replica copies before a success response is returned to the client application. On the other hand, for data that is considered mission critical, the replication policy can be set to wait until an acknowledgment is received from nodes across multiple data centers before an acknowledgment is sent back to the application. Consistency levels don't just impact availability, but also latency. Latency increases as the consistency level increases and negatively affects performance. During write operations, a higher consistency level requires additional overhead because more nodes must be written to before sending a response back to the client. The same occurs when reading data because more replicas need to be compared to deliver the newest version of the data.

Optimized for all data sizes

Built on the Cassandra* NOSQL database, enhanced with file system properties, HyperStore can store vast amounts of unstructured data without object size limitations. Erasure coding enables deep archive efficiency and flexible redundancy, giving you robust data protection without consuming precious disk space. Object replicas are employed for frequently used data. This gives HyperStore improved storage scaling and finer control over data availability.

Storage Node Heterogeneity

Cloudian's vNode technology enables data centers to intermix node types. In other words, storage nodes deployed into a cluster can be of dissimilar size. For example, a 24TB node could be installed alongside a 48TB node and the HyperStore operating system will automatically pool and load balance these resources as they are added to the cluster. This gives businesses the flexibility to add capacity and CPU resources at the desired granularly. It also helps to improve efficiencies as the right resources can be added to the cluster only when needed.

Amazon S3 Data Tiering

HyperStore appliances give you the option of on-premises storage as well as easy access to the AWS cloud. This unique hybrid cloud storage approach means you can choose to leverage AWS infrastructure for long term bulk storage while keeping your most critical data close at hand. With unique features like object streaming and dynamic auto-tiering, data moves seamlessly between your on-premises cloud and Amazon S3, regardless of file size.

Quality of Service (QoS)

QoS and metering are foundational capabilities for implementing a multi-tenant private cloud storage solution. With HyperStore, storage administrators can set a maximum allowable limit on both storage consumption and I/O, based on the user or a group of users, and then charge back those users on a monthly basis, just like a utility. Configurable group and user-level QoS ensure groups and users do not exceed storage quotas or consume bandwidth in a manner that impacts other tenants. A CFO could be assigned a high priority privilege (Platinum Service Level) to financial records while an end-user accessing sync and share data could be given lower priority access (i.e. Silver).

Security

With data security breaches becoming more commonplace, it is essential for businesses to safeguard their data from the prying eyes of data hackers and unauthorized users. HyperStore simplifies the data encryption process by providing transparent key management. HyperStore AES-256 Server-Side Encryption enables enterprises and service providers to easily encrypt data stored at rest. SSL encryption ensures data confidentiality for data in transit (HTTPS). And with S3- compatible object-level ACLs, system administrators can secure buckets and objects with either no access, read-only or read-write permissions for everyone or named users and groups.

Compression

Compression can reduce storage and network consumption by up to 40%, while accelerating data replication speeds and reducing network bandwidth requirements. With less data to store on disk and less data to move over the network, businesses can get more life out of their existing storage and network investments, further improving their ROI and lowering their TCO. HyperStore offers three different types of data compression technology—lz4, snapp and zlib.
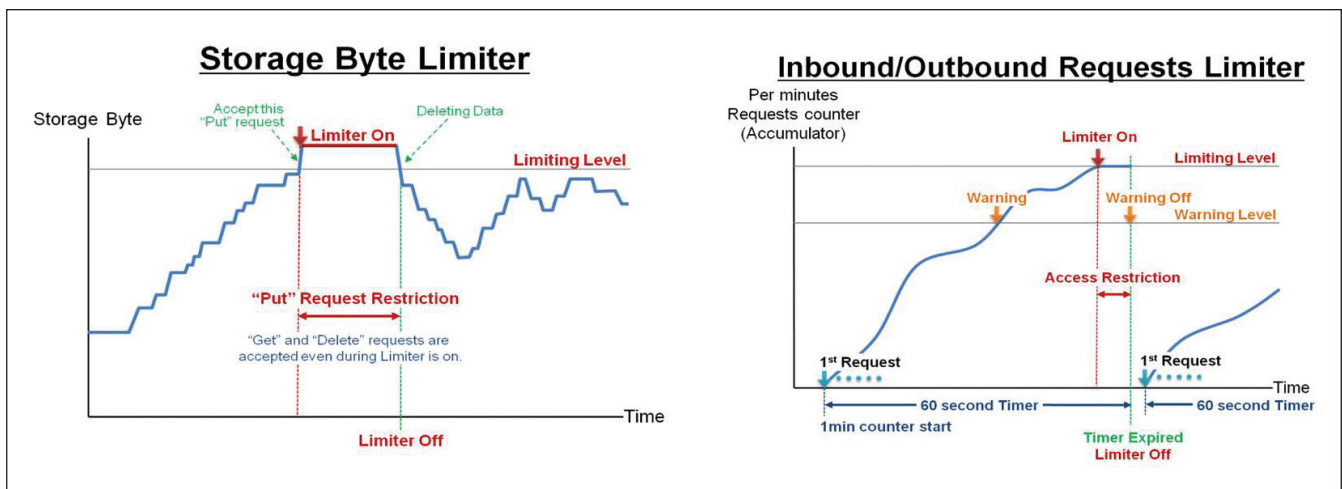

Figure 4: Cloudian Hyperstore QoS

Multi-Tenancy

Multi-tenant management starts with account segmentation where each account is logically segmented within the software. Data for accounts is only accessible by account users and the group administrator. Advanced identity and access management features allow system administrators to provision and manage groups and users, define specific classes of service for groups and users, and configure billing and charge-back policies. Both administrators and users benefit from unique reporting options and account and data management capabilities. Multiple credentials per user are also supported.

GUI and API for Billing, Management and Monitoring

HyperStore provides system monitoring, data management, provisioning and management of users, groups, rating plans, QoS and billing via graphical user interface or RESTful APIs. The graphical user interface is also highly and easily customizable to provide easy integration with existing environments.

Figure 5: Cloudian Hyperstore Multi-Tenancy Admin Interface

Fully S3 Compliant API for the Broadest Application Support

HyperStore software's fully S3 compliant API ensures seamless interoperability of applications developed for Amazon S3-based storage clouds. HyperStore supports many advanced S3 features such as Multi-Part Upload, Object Versioning and S3 compatible ACL, Location Constraint and many more.

Cloudian HyperStore Appliance

Cloudian HyperStore software is also available in fully integrated Intel-powered HyperStore appliances.

Software-based object storage offers an alternative approach to NAS/SAN systems. But some enterprises don't have the time or the people resources to integrate their own solutions. Instead, they prefer the ease of deployment that an appliance offers, along with the cost savings that a scale-out software solution can deliver. The all-in-one appliance approach gives them the confidence to deploy these systems that have been pre-tested and configured to optimally work with HyperStore software.

Intel-powered HyperStore appliances come in three enterprise-focused configurations. All three models provide redundancy and high availability features like hot-pluggable disk drives and dual hot-pluggable power supplies.
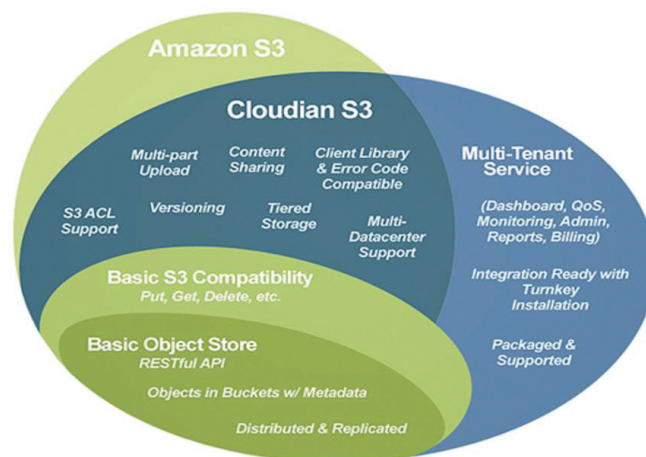

Figure 6: S3 API Coverage Diagram

Entry Level Clouds	Active Archives	Enterprise Performance
HSA1024	**HSA1048**	**HSA2060**
The entry level HSA1024 is a 1U system equipped with 32GB of RAM and a 4xGigE NIC. The HSA1024 delivers 24TB of capacity and is perfect for low throughput workloads like sync and share applications.	The mid-tier ultra-dense HSA1048 is a 1U system equipped with 32GB of RAM and a 4xGigE NIC. The HSA1048 delivers 48TB and is targeted at higher capacity entry point deployments. This configuration is a suitable building block for large multi-media content libraries, medical records, data center backups and other big content applications.	The high performance HSA2060 is a 2U system equipped with 64GB of RAM and two Intel NIC devices -2x1GigE and 2x10GigE. The HSA2060 delivers 60TB and is built for high I/O bandwidth requirements like in media and entertainment, energy, finance, and health care environments. The system intelligently places metadata into memory and SSD to ensure high performance.

This reference architecture is based on the top-end appliance, the HSA2060.

Reference Architecture

The typical Enterprise has data streaming in from all directions. The challenge is to take unstructured data and synthesize it, quantify it, and increase its business value by deriving actionable results. The following reference architecture and use case demonstrate how an enterprise can use Cloudian HyperStore and Apache™ Hadoop® to build an efficient data content store platform for running analytic software like MapReduce, Pig, and Hive.

Apache Hadoop provides a variety of file systems to use when processing data. You can specify which file system to use by the prefix of the URI used to access the data. For example, s3n://mybucket/path references an Amazon or Cloudian HyperStore bucket.

The main advantage of using the s3n file system versus the traditional HDFS file system is that it eliminates the need of extracting data from HyperStore in order to be processed by the Hadoop cluster. This increases analysis speed and eliminates redundant data storage.

Another advantage of the Hadoop s3n file system is that you can access files on Cloudian HyperStore that were written with other tools. Conversely, other tools can access files written by Hadoop.

Moreover, multiple Hadoop clusters can access Cloudian HyperStore storage directly and the processing results can be stored in HyperStore for long-term retention in both replicated or erasure coded format. Also, Cloudian
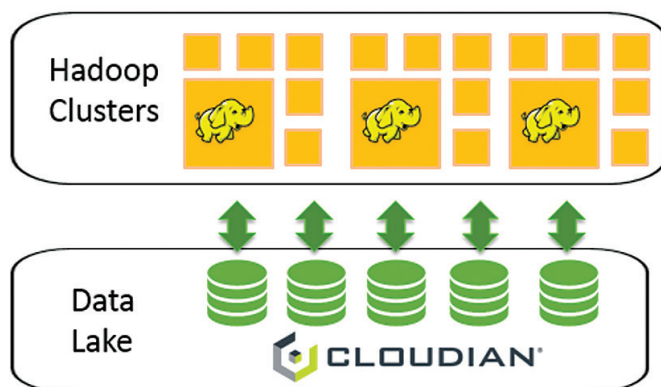
Figure 8: Multiple Hadoop Clusters And Cloudian Hyperstore

HyperStore supports encryption and compression. These storage types help improve data efficiency and security.

Finally, enterprises can increase/decrease computational analysis and storage capacity by scaling up /down HyperStore storage servers and Hadoop compute servers independently of each other. The decoupling of compute nodes and storage nodes allows you to optimize Hadoop clusters for particular workloads and/or shutdown clusters at job completion.

The following table summarizes the benefits of the combined solution.

Figure 7: Hadoop and Hyperstore Diagram

Features & Benefits of the Combined Solution

1. Use multiple HDP clusters to process the same data set.
2. Run optimized HDP clusters for particular workloads.
3. Shut down clusters when data analysis is finished.
4. No redundant storage of data.
5. Scale capacity independently from compute.
6. Keep your data safe with Cloudian HyperStore data at rest encryption.
7. Shrink data footprint by using Cloudian HyperStore compression.
8. Augment Hadoop with life cycle data management.

The Architecture

Analytics can be run against large data sets more easily when Hadoop is co-located with the data being analyzed. Customer data that many organizations are collecting on a regular basis can be mined for myriad insights, such as understanding the capabilities customers most value, or improving capabilities in future products.

For example, Cloudian is using its HyperStore systems running in-place Hadoop analytics as a key diagnostic tool within the customer support function. Customer-owned HyperStore systems in the field send telemetric data to Cloudian support for end-user, operations, and trend analysis. The Cloudian support organization in return feeds actionable information back to the customer. Cloudian calls this feature Smart Support.

Figure 9 illustrates the overall configuration. On the left side of Figure 9, Cloudian HyperStore provides cloud storage functionalities to end-users. On the right side of Figure 9, Cloudian HyperStore is used as a content storage to collect the logs that the Smart Support feature is sending. This design can extend to any device sending logs, data telemetry, statistics, etc. to a Cloudian HyperStore storage cluster where the data is stored and analyzed in place by the Hadoop cluster.
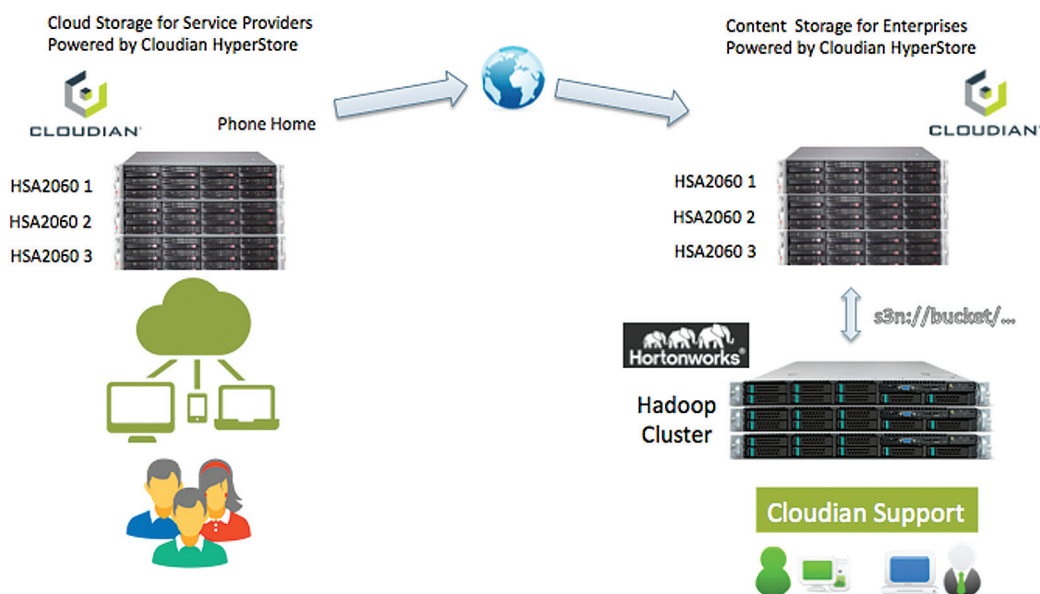
Figure 9: Smartsupport and Hyperstore Hadoop Integration

As shown in Figure 9, three Cloudian HyperStore HSA2060 appliances were deployed to create the content storage and three standard Intel-based servers were used to host the Apache Hadoop software. This portion of the test bed is detailed in Figure 10 - see nodes hyperstore01, hyperstore02, hyperstore03, hdp01, hdp02, and hdp03. The HyperStore appliances are connected via a back-end redundant 10Gb network and a front-end redundant 1Gb network. The Hadoop cluster communicates and transfers data with the HyperStore cluster via the 1Gb network.
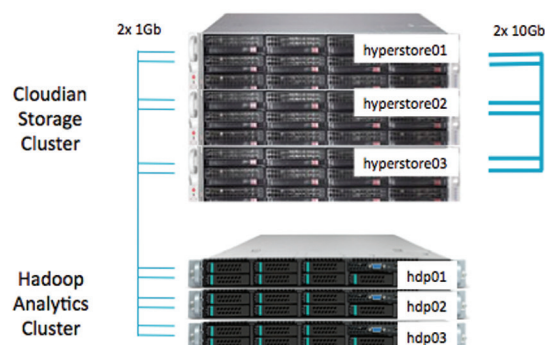
Figure 10: Hyperstore-Hadoop Setup

Technical review

This section describes HyperStore appliances and software configuration. Specifically, it shows the steps to install and configure Hadoop software on Intel nodes and HyperStore software on Intel-powered HyperStore appliances. This "cookbook" style tutorial should help the reader recreate the reference architecture shown in Figure 10: HyperStore-Hadoop setup.

Rack Component	Active Archives
HyperStore Node	SuperMicro* SSG-6027R-E1R12L 2U Chassis 1x Intel® Xeon® E5-2658 V2 8 x 8GB 1Gx64 DDR3-1600 UDIMM, 12 Toshiba Tomcat* 4TB 7200RPM, 3.5in, SATA HDD 2 Intel® S3500 240GB 2.5in SSD 1 Quad ported 1GbE Intel® i350 GbE Controller 1 Dual ported 10GbE Intel® X540 10GbE Controller
Hadoop Node	SuperMicro 5018A-AR12L - 1U Chassis Intel® Atom™ C2750 (8 core) 4 x 8GB 1Gx64 DDR3 UDIMM 4 x Toshiba Tomcat 2TB 7200RPM, 3.5in, SATA HDD 1 Dual ported 10GbE Intel® X540 10GbE Controller
Network Switches	Netgear* M7100 10GbE Ethernet switch 48 x 10Base-T/1000Base-TX/1000Base-T 4 x 1/10 GbE dual-speed SFP+ ports
Software	Cloudian HyperStore 5.1 Hortonworks® Hadoop HDP 2.1

Apache Hadoop Installation

Download HDP from http://hortonworks.com/hdp/downloads/. Use the Automated (Ambari) version as shown in Figure 10. Using Ambari is the recommended way to set up HDP for a production environment. Apache™ Ambari simplifies the provisioning, management, and monitoring of your cluster.
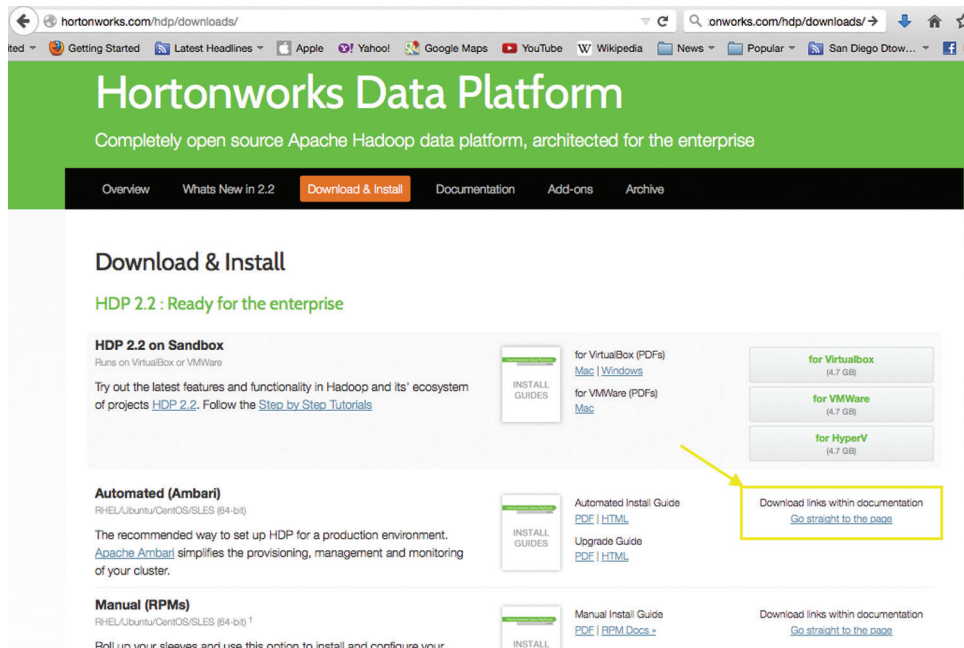

Figure 11: Install Page

1. Log in to any of the Intel servers as root.

2. Download the Ambari repository file to a directory on the Intel server.
   wget -nv http://public-repo-1.hortonworks.com/ambari/centos6/2.x/updates/2.0.0/ambari.repo *-O /etc/yum.repos.d/ ambari.repo*

3. Confirm that the repository is configured by checking the repo list.
   yum repolist

4. Install the Ambari package.
   yum install ambari-server

5. Install the PostgreSQL Ambari database.
   yum install postgresql postgresql-server

6. Setup the Ambari server
   ambari-server setup

7. Start the Ambari Server
   ambari-server start

8. Using a web browser log in to the Ambari server and start the cluster creation wizard
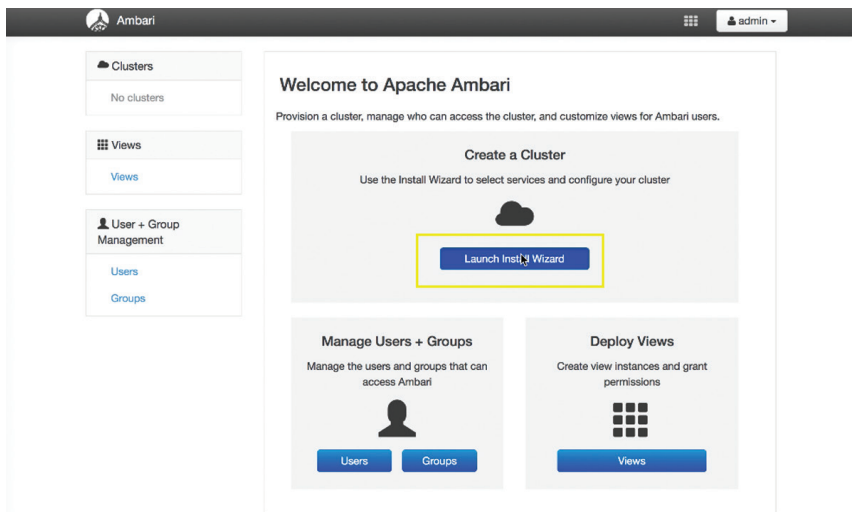
Figure 12: Ambari Page

9. Follow the prompts

Cloudian HyperStore Appliances Configuration

1. Log in to the Cloudian HyperStore appliance hyperstore1 with user "root" and password "password"
2. Launch the Cloudian HyperStore appliance configuration tool *configure_appliance.sh:*
   [root@cloudian-node1]# configure_appliance.sh

On launch, the tool displays a task menu for setting up the host:
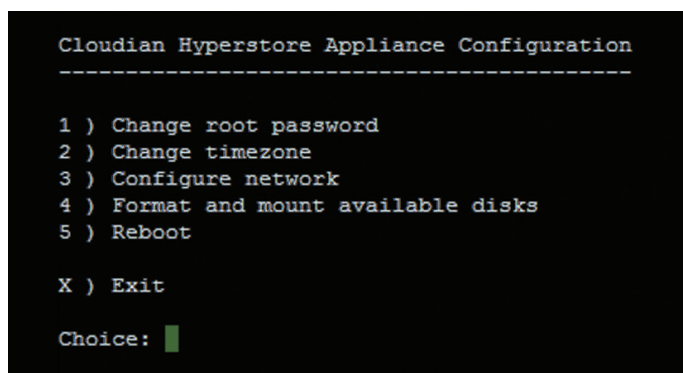
Figure 13: Hyperstore Appliance Configuration Menu

Complete Steps 1 through 4 for each HyperStore appliance - hyperstore1, hyperstore2, and hyperstore3.

3. Choose one of your HyperStore appliances and copy the Cloudian license file into the installation staging directory/*root/CloudianPackages.*

4. Change into the staging directory */root/CloudianPackages,* then run the following command:

    root# ./CloudianHyperStore-5.1.bin <license-file-name>

5. In the staging directory, create a file named survey.csv and in it enter one line for each HyperStore host in the following format:

    <region-name>,<hostname>,<ip4-address>, <datacenter-name>,<rack-name>

    region1,hyperstore1,192.168.2.1,DC1,RAC1

    region1,hyperstore2,192.168.2.1,DC1,RAC1

    region1,hyperstore3,192.168.2.1,DC1,RAC1

6. In the staging directory, launch the HyperStore software installer.

    root# ./cloudianInstall.sh -s survey.csv

7. *From the installer menu, complete tasks #1 through #7 in sequence.*

Figure 14: Hyperstore Software Configuration Menu

8. *During task #2 you will need to provide information about your desired HyperStore deployment such as the S3 service domain and S3 data replication strategy.*
9. Once the HyperStore system is installed and running, you can use the Cloudian Management Console (CMC) to create a group and a user for HDP data. Point your browser to http://<CMC_host_IP_address>:8888/Cloudian
10. You will get a certificate warning. Follow the prompts to add an exception for the certificate and accept it. You should then see the CMC's login screen:
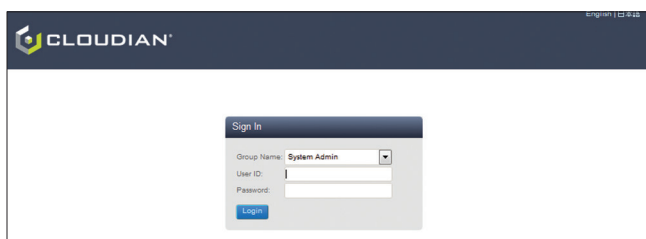
Figure 15: CMC Login Screen

11. Log in with the system administrator user ID *admin* and default password *public*.
12. Select the **Admin** tab
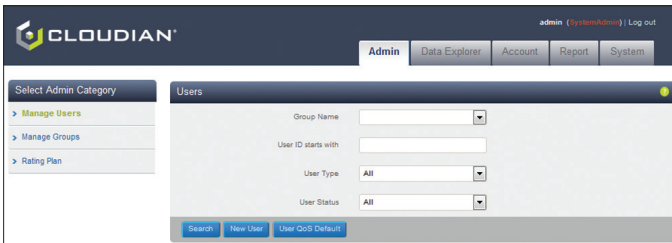
Figure 16: CMC Admin Screen

13. Select **Manage Groups → New Group** and create a new user group **e.g. hdpgroup**
14. Select **Manage Users → New User** and create a new regular user **e.g. hdp**, assigned to the group **hdpgroup**.
15. Log out of the CMC and then log back in as the new HDP user that you created. The CMC UI displays with the **Data Explorer** tab selected. When you log in as a regular user, fewer tabs are available in the CMC UI.
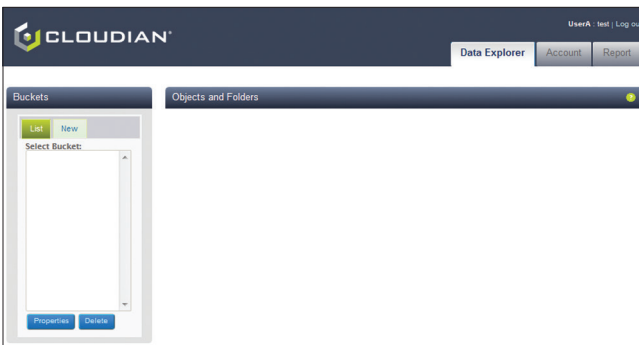
Figure 17: CMC Bucket Creation Screen

Cloudian HyperStore Software Configuration

1. Create a new storage bucket for the user **e.g. hdp**
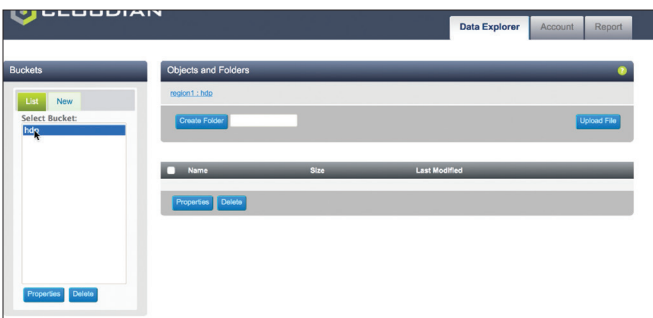
Figure 18: CMC HDP Bucket

2.   Navigate to **Account > Security** Credentials to get the access and secret keys
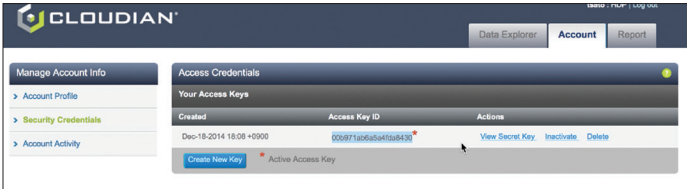
Figure 19: CMC User Credentials Screen

In this example the keys are:

Access Key: 00b971ab6a5a4fda8430

Secret Key: QPsrxQ5jRUeBqh6xo92E672mhBUFLeGGbaCrMkEc

HDP Setup to connect to Cloudian HyperStore

1.   If DNS is not setup in your network, add an entry to the */etc/hosts* file to resolve the Cloudian HyperStore endpoint e.g. s3.cloudianhyperstorestorage.com

> *# echo "54.174.123.34 hdp.s3.cloudianhyperstorestorage.com s3.cloudianhyperstorestorage.com" >> /etc/hosts*

2.   Create a configuration file to access Cloudian HyperStore Storage from the Hadoop cluster

> *# printf "s3service.https-only=false\n*
>
> *s3service.s3-endpoint=s3.cloudian.com\n*
>
> *s3service.s3-endpoint-http-port=80\n" > /etc/hadoop/conf/jets3t.properties*

3.   Add custom properties to the core-site.xml file

> *fs.s3n.awsAccessKeyId: Access Key* [00b971ab6a5a4fda843000b971ab6a5a4fda8430]
>
> *fs.s3n.awsSecretAccessKey: Secret Key* [QPsrxQ5jRUeBqh6xo92E672mhBUFLeGGbaCrMkEc]
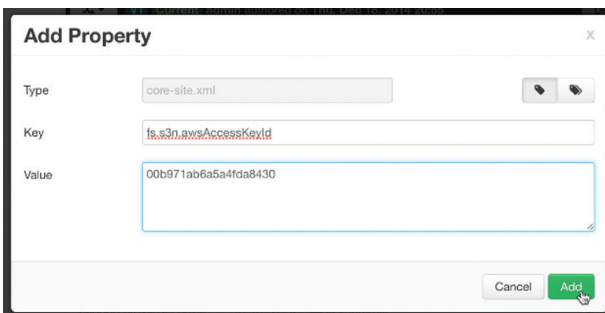
Figure 20: HDP Hyperstore Credentials Setup Screen

Cloudian Big Data Analytics

4. Follow the restart suggestions by Ambari
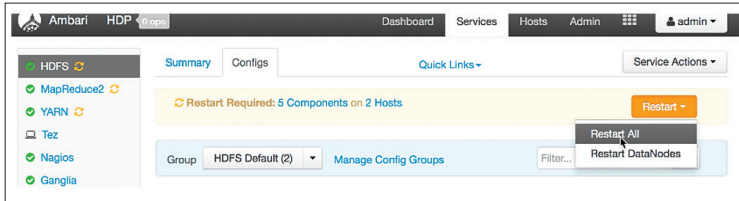
Figure 21: Ambari Final Setup Screen

5. Verify that Hadoop can connect to Cloudian HyperStore storage and list the files present in the bucket
       # hdfs dfs -ls s3n://hdp/

Performance Considerations

This section describes the recommended system configuration settings to maximize performance for the Hadoop analytics use case. These settings can improve performance by 3X or more.

Organize Data by Key

How you organize files inside the bucket makes a significant difference for performance. Make sure you can quickly list and fetch subsets of your data without scanning the whole bucket. Inside object storage, directories do not exist, there are only buckets and keys. To create the illusion of a directory structure, the Hadoop S3 File System prepends the directory pathname to the name of the file to create an S3 key. To improve performance, name your keys hierarchically by putting the most common things you filter by on the left side of your key.

Store Large Files

Combine your data into large files versus multiple small files. This minimizes time spent listing files in your bucket. Less files means less list calls and TCP connections to read your data.

Align Hadoop data chunks with HyperStore data chunks

In Hadoop the default chunk size is 64MB while in Cloudian HyperStore, the default chunk size is 5MB. We measured a 3X improvement in data streaming performance by aligning the chunk sizes. You can do this by editing mts.properties file:

  #cassandra.fs.max_chunk_size = 10485760 <== 10MB

  cassandra.fs.max_chunk_size = 67108864   <== 64MB

Avoid Underscores in bucket names

Many tools convert bucket names to hostnames but underscores are not allowed by DNS. Use dashes instead of underscores in your bucket names.

- **Good Bucket Name:** my-awesome-bucket-of-cats
- **Bad Bucket Name**: my_soon_to_break_bucket_of_cats

Performance Analysis

This section reports the performance results we obtained during our lab testing.

Test #1 Chunk alignment optimization effect on data streaming performance

As described in the Performance Consideration section, the default Hadoop chunk size is 64MB, while in Cloudian HyperStore it is set to 5MB. In this test we measured the time it took to transfer data from Cloudian HyperStore to a Hadoop Pig process using the default value of 5MB and the Hadoop optimized value of 64MB for five different file sizes: 16MB, 32MB, 64MB, 128MB and 256MB.

The tests were performed by storing the two separate data sets in two separate HyperStore buckets, one where the data was stored in 5MB chunks (default size) and one where the data was stored in 64MB chunks (optimized size). We then used a simple Hadoop Pig job to perform a read operation of the entire data set. The tests were executed 3 times, of which we computed and recorded the average values. As you can see from the graph and table below, the optimization resulted in a faster data transfer for all the file sizes tested, achieving an improvement between 1.45X and 3.11X.
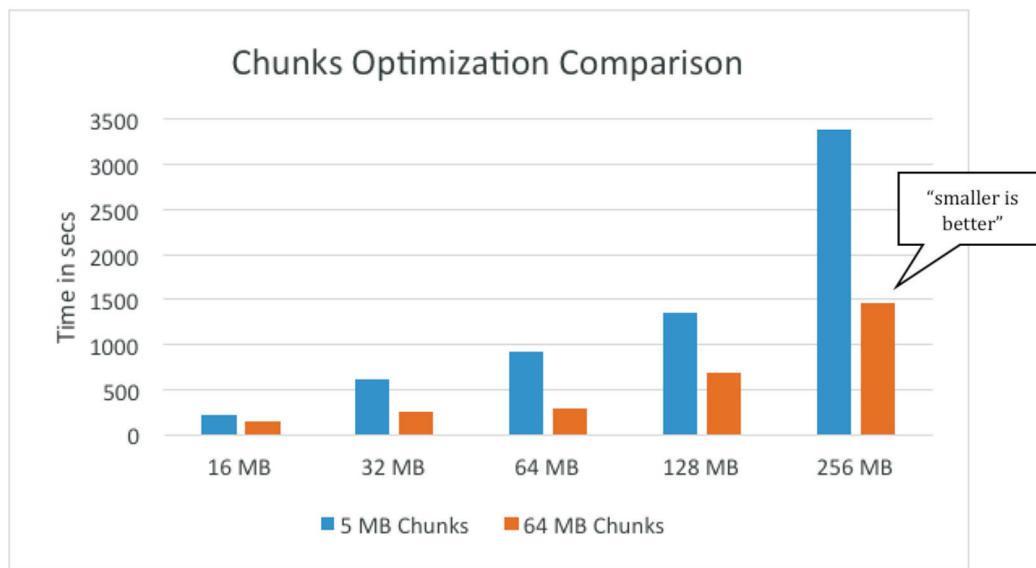
File Sizes	5 MB Chunks	64 MB Chunks	Improvement
16 MB	3 min 35 sec	2 min 28 sec	**1.45X**
32 MB	10 min 20 sec	4 min 12 sec	**2.46X**
64 MB	15 min19 sec	4 min 55 sec	**3.11X**
128 MB	22 min 37 sec	11 min 21 sec	**2X**
256 MB	56 min 30 sec	24 min 12 sec	**2.33X**

Test #2 End-to-end data analysis, HDFS vs. Cloudian HyperStore

In this test we measured the time it took to perform a complete end-to-end data analysis, comparing different data set sizes when using just HDFS and when using Cloudian HyperStore storage via the Hadoop S3n connector. The Hadoop S3n connector allows Hadoop to directly access data stored in Cloudian HyperStore. The tables and graphs below show that in both cases the time it takes for the Pig job to analyze the data does not change, no matter if the data is stored in HDFS or streamed directly from Cloudian HyperStore. The graphs and tables also demonstrate that a

remote device implementing a RESTful API to upload data to a central object storage repository is always faster than a traditional device using a Linux copy command to copy the data. Moreover, in the Cloudian HyperStore case, the intermediate step of copying data from the Hadoop server to HDFS is completely avoided, gaining further execution speed.

In conclusion, the graph below plots the total execution time for the 3 data sets and shows that for all the cases the Cloudian HyperStore configuration performed faster than simply using HDFS alone.

Dataset Sizes	32 MB	64 MB	128 MB
Data copy operation from remote device to Hadoop server	50 sec	105 sec	209 sec
Hadoop Server to HDFS	19 sec	31 sec	66 sec
Pig job duration	49 sec	55 sec	55 sec
Total Time	**118**	**191**	**330**
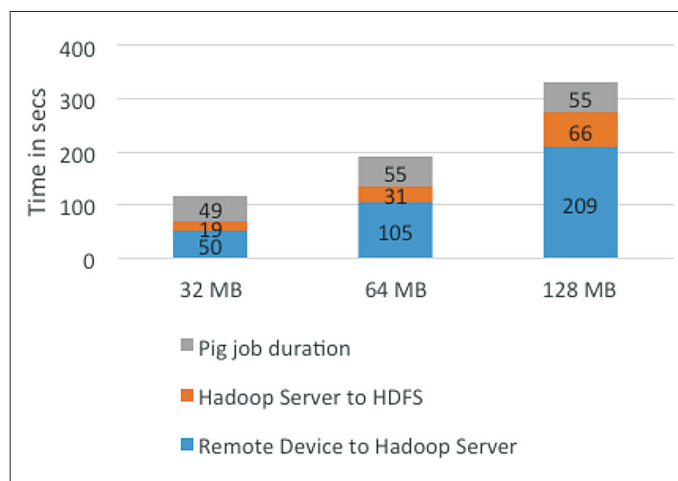
Dataset Sizes	32 MB	64 MB	128 MB
Data upload operation from remote device to HyperStore	29 sec	100 sec	192 sec
Pig job duration	56 sec	56 sec	56 sec
Total Time	**85**	**156**	**248**

Total Duration Comparison

"smaller is better"

Legend: ■ Total Time with HDFS only ■ Total Time HyperStore

	32 MB	64 MB	128 MB
Total Time with HDFS only	118	191	330
Total Time HyperStore	85	156	248
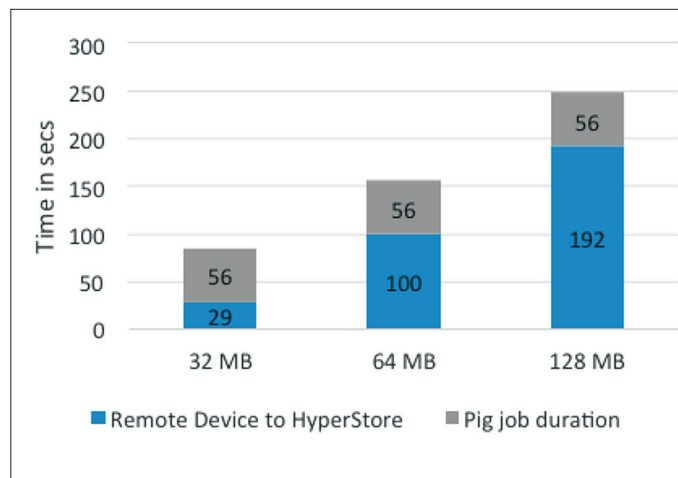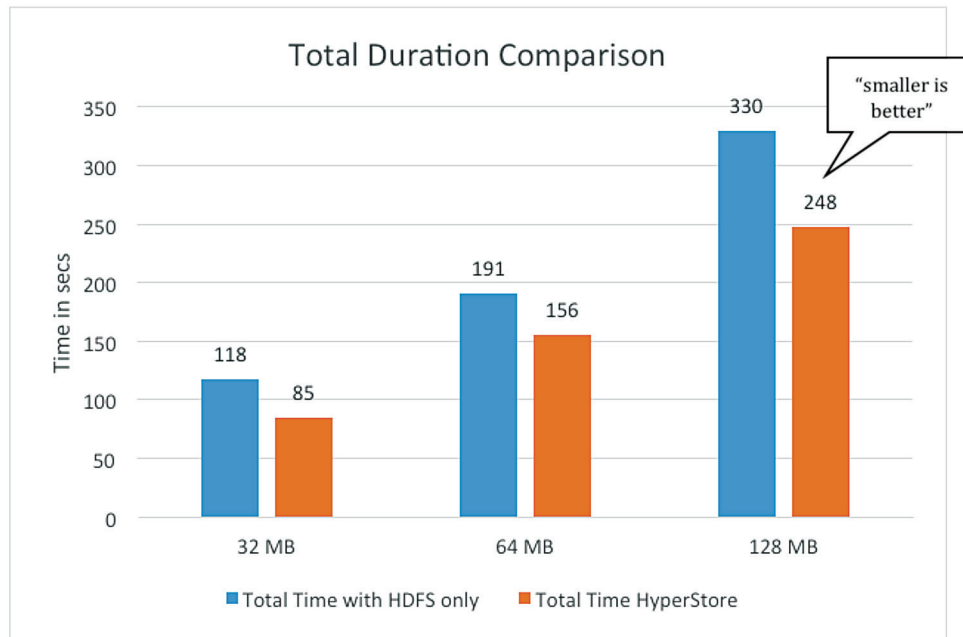
Cloudian Customer Support Issue Use Case

As described in the Reference Architecture chapter, Cloudian is using its HyperStore systems running in-place Hadoop analytics as a key diagnostic tool within the customer support function. Customer-owned HyperStore systems in the field send telemetric data to Cloudian support for end-user, operations, and trend analysis.

As an example, let's use the case where a particular end user is experiencing slower than usual uploads to the Cloudian HyperStore storage cloud. By using the Cloudian Smart Support feature that automatically reports the HyperStore health status back to the Cloudian support organization, Cloudian support engineers get access to log files containing access time, object owner, type of request (e.g. GET, PUT), object size, etc. for every Cloudian node in the cluster.

Figure 22 shows the simple Hadoop Pig job, which loads the data from Cloudian HyperStore via s3n file system. The job aggregates the information contained in the log based on time, object owner, bucket, type and status code. The results are saved back to Cloudian HyperStore.

```
1
2  /*
3   * Simple
4   */
5  A = LOAD 's3n://hdp/phonehome/s3' USING PigStorage('|') AS (
6      timestamp:chararray,
7      address:chararray,
8      owner:chararray,
9      type:chararray,
10     bucket:chararray,
11     user:chararray,
12     reqHeader:int,
13     reqBody:int,
14     respHeader:int,
15     respBody:int,
16     total:int,
17     duration:int,
18     objectKey:chararray,
19     statusCode:chararray,
20     reqId:chararray);
21 B = FOREACH A GENERATE SUBSTRING(timestamp, 0, 13) as hour, *;
22 C = GROUP B BY (hour, owner, bucket, type, statusCode);
23 D = FOREACH C GENERATE flatten(group),
24     SUM(B.reqHeader),
25     SUM(B.reqBody),
26     SUM(B.respHeader),
27     SUM(B.respBody),
28     SUM(B.total),
29     SUM(B.duration),
30     COUNT(B.address),
31     AVG(B.duration),
32     MIN(B.duration),
33     MAX(B.duration);
34 STORE D INTO 's3n://hdp/phonehome_20141210' USING PigStorage(',');
```

Figure 22: Hadoop Pig Script

Figure 23 shows how the output is later imported into Microsoft Excel and analyzed using Pivot Graphs. The Pivot Graph quickly reveals the root cause of the issue: an abnormal increase in the number of delete bucket requests.
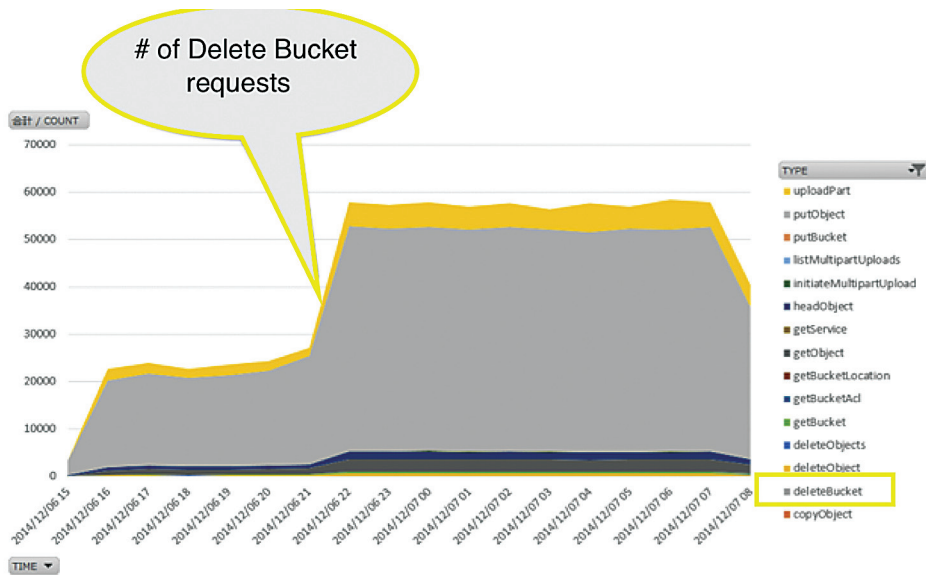
Figure 24: Data Analysis Visualization

Conclusion

The need for a scalable and analytics-ready storage system while controlling cost, optimizing availability, and increasing disaster recoverability can be a challenge given today's IT budgets. To address this, many enterprises are turning to scale-out storage architectures like Cloudian HyperStore. By decoupling storage functionality from the underlying hardware and taking advantage of industry standard servers, Cloudian HyperStore can offer scalability, performance, and data availability far beyond what traditional NAS/SAN arrays that use expensive custom defined hardware can offer.

This paper describes a scale-out storage reference architecture that excels at delivering data analytics by combining Cloudian HyperStore and Apache Hadoop advanced software features with energy-efficient Intel-based servers. The combined solution enables enterprises to run powerful data analysis tools like MapReduce and Pig directly on Cloudian HyperStore storage, reducing CAPEX and OPEX.

Cloudian HyperStore software deploys via ISO image on industry standard servers based on Intel® Xeon® E5-2600 v3 and Atom™ C2700 series processors, 1Gb and 10Gb Intel® Network Adapters, and Intel® Solid State Drives Data Center Family, resulting in a massively scalable and highly available storage environment. A turnkey appliance version is also available.

This reference architecture and deployment steps demonstrate the ease of setup, configuration and usability of the solution.

For more information about Cloudian product offerings, visit http://www.cloudian.com

Disclaimers

(intel®)