Table of Contents

EXECUTIVE SUMMARY 4
THE DATA EXPLOSION 4
A NEW TYPE OF STORAGE 4
A NEW IT CHALLENGE 4
HYPERSTORE OVERVIEW 5
THE MULTI-CLOUD JOURNEY 5
MULTI-CLOUD CHALLENGES 6
HYPERSTORE UNIQUE FEATURES 6
HYPERSTORE WITHIN A MULTI-CLOUD ARCHITECTURE 6
HYPERSTORE MULTI-CLOUD UNIQUE FEATURES 7
HYPERSTORE LIMITLESSLY SCALES ON DEMAND 7
EFFORTLESS DATA MOVEMENT BETWEEN ON-PREM AND THE CLOUD 7
AUTO-TIERING 7
S3 COMPATIBILITY WITH BROAD APPLICATION SUPPORT 7
HYPERSTORE MULTI-CLOUD USE CASES 8
USE CASE SUMMARY 8
CROSS-CLOUD DATA PROTECTION 8
FLEXIBLE CLOUD PROCUREMENT 9
FILE SERVICES 9
CLOUD BURSTING 10
DATA ANALYSIS 10
HYPERSTORE CLOUD DEPLOYMENT EXAMPLES 10
HYBRID CLOUD TIERING EXAMPLE 10
MULTI-CLOUD DEPLOYMENT DEMONSTRATION 15
MULTI-CLOUD DEPLOYMENT NOTES 15
MULTI-CLOUD INFRASTRUCTURE DEPLOYMENT REQUIREMENTS 16
PRIVATE NETWORK IP CONFIGURATION 16
ENCRYPTED VPN DEPLOYMENT NOTES 16
NETWORK SECURITY NOTES 16
TERMINOLGY NOTES 16
AMAZON AWS CONFIGURATION 17
GOOGLE GCP CONFIGURATION 21
MICROSOFT AZURE CONFIGURATION 24
CONFIGURE WINDOWS RRAS ON AMAZON AWS 28
DEPLOY CLOUDIAN HYPERSTORE 33
CONFIGURE HYPERSTORE WITH A DISTRIBUTED EC STORAGE POLICY 34
VIEW SUMMARY FOR DETAILS 35
SIMPLE DASHBOARD 35
SINGLE CLUSTER USAGE & PERFORMANCE VIEW 35
EXECUTIVE SUMMARY
In this technical white paper, we discuss the challenges of moving to public and multi-cloud environments, provide an in-depth view of Cloudian HyperStore and how it can meet the challenges of integrating into private and public cloud hybrid architecture. Next, there will be an overview of HyperStore, an introduction to the multi-cloud journey and a discussion of HyperStore’s unique features that allow for seamless integration into your organization’s overall cloud strategy. After that, we discuss various multi-cloud use cases and deep-dive into HyperStore’s hybrid cloud and multi-cloud capabilities. Finally, we’ll discuss possible multi-cloud architectures and walk through detailed technical deployment examples to show HyperStore’s unique advantages.

THE DATA EXPLOSION
With the popularity of rich media, the proliferation of mobile devices and the digitization of content, there has been and continues to be exponential growth in the quantity of unstructured data that IT is managing. In fact, IDC predicts that all data will grow to 40 zettabytes by 2020, resulting in a 50-fold growth from the beginning of 2010. 90% of this data growth will be unstructured. The growth is not slowing down; in fact, it continues to accelerate in size and scope.

This explosive growth in data and content is simply not sustainable for current NAS and SAN infrastructures. Backups and restores are taking longer. Migrating data from older storage systems to new storage systems are labor intensive and expensive. Provisioning storage for users is more frequent and time consuming.

Not only does all this unstructured data increase the cost of managing the infrastructure, it also impacts the internal consumers of storage. Most IT organizations are faced with a flat to declining storage budget and are forced to manage the ever-increasing storage with the same or reduced IT resources. In short, the costs and complexity of traditional storage systems continue to increase. A radical change in storage infrastructure is needed if enterprise IT is ever going to tame the data explosion.

A NEW TYPE OF STORAGE
Software-defined (SDS) object storage offers an alternative approach to NAS/SAN systems based on expensive proprietary hardware. Cloudian HyperStore® is a scale-out S3 compatible object storage system designed to manage massive amounts of unstructured data. It is an SDS platform which runs on any standard x86 server platform. This dramatically reduces the cost for datacenter storage whilst providing limitless scalability, extreme availability and unprecedented reliability leveraging the 18 to 24-month CPU innovation cycle. Compare this with the typical 3-year product refresh cycle from proprietary storage vendors. In addition, manufacturers continue to drive innovation into the hard disk drive market space, delivering increased disk drive densities and a lower cost per gigabyte (GB). Software-defined object storage architecture allows enterprises to take advantage of these technology updates earlier and meet the explosive storage growth demand. Operational efficiencies are realized on-premise in the data center, within a private cloud and with HyperStore, these efficiencies extend to the public cloud as well.

A NEW IT CHALLENGE
Exponential growth of unstructured data isn’t the only challenge. Gartner’s view from The Future of the Data Center in the Cloud Era is that “A multi-cloud strategy will become the common strategy for 70% of enterprises by 2019.” Most organizations are increasingly using a combination of public clouds, infrastructure-as-a-service (IaaS) providers and on-premises infrastructure, called private clouds, to provide IT infrastructure resources and applications. This combination of infrastructure services is becoming increasingly common and is being known as a hybrid or multi-cloud. A hybrid-cloud typically consists of on-premises infrastructure with at least one public cloud provider, while a multi-cloud typically consists of 2 or more public cloud providers.
Why is the adoption rate increasing? Organizations realize that sourcing flexibility, leveraging existing technology relationships and taking advantage of specific feature sets that allow greater efficiencies at scale is critical in today's business environment. Organizations that move to a multi-cloud strategy also take advantage of the accelerated pace of public cloud innovation. Traditional IT infrastructure simply can't keep up because it is inherently inflexible, has long upgrade cycles and is becoming more expensive to maintain.

Cloud innovation allows for nimble application development and faster distribution of IT deliverables with greater efficiency. These organizations may use multi-cloud in a specific geography, vertical market or even for flexibility in their own private cloud. Furthermore, a multi-cloud strategy can meet regulatory requirements such as GDPR, provide greater operational efficiency and can leverage public cloud IaaS provider's integration with traditional IT environments. Thus, moving to a multi-cloud fosters innovation, a competitive edge and ultimately flexibility to business models that traditional IT infrastructure cannot provide. A multi-cloud strategy can also effectively address enterprises' broader business goals, whether that be the use of more price-competitive cloud services or taking advantage of the speed, capacity or features offered by a specific cloud provider. Traditional IT infrastructure is increasingly struggling to meet these challenges too.

HYPERSTORE OVERVIEW

Cloudian HyperStore enables data centers to provide highly cost-effective on-premise unstructured data storage repositories. Cloudian HyperStore is a software platform that can span across the enterprise as a private cloud as well as out into public cloud environments. Cloudian HyperStore is available as stand-alone software or fully integrated with hardware as a Cloudian HyperStore appliance. HyperStore can be deployed pre-installed on-premise, provisioned within a public cloud or any combination. HyperStore easily scales to limitless capacities and offers multi-datacenter cloud storage capabilities.

Cloudian HyperStore is the industry's most compatible S3 storage system. Since S3 is the de-facto public cloud storage protocol standard, HyperStore seamlessly fits into an organization’s IT cloud strategy. HyperStore has fully automated data tiering to all major public clouds, including AWS, Azure and Google Cloud Platform. It fully supports S3 applications and has flexible security options. HyperStore deployment models include fully on-premises storage, distributed off-site storage, storage-as-a-service or even other cloud combinations as illustrated below.

THE MULTI-CLOUD JOURNEY

As you’ve seen, HyperStore is built for cloud environments. Moving an organization to a multi-cloud infrastructure is a journey. Organizations typically start with HyperStore on-premise to take advantage of the limitlessly scalable architecture for data capacity intensive applications, native S3 support, intelligent search and integrated management capabilities as private cloud infrastructure. Later, as organizations begin to explore the benefits of the public cloud, such as replication for DR, flexible capacity expansion and content distribution, a hybrid-cloud infrastructure emerges. For example, a hybrid-cloud can leverage cloud resources for capacity and compute using HyperStore’s tiering capabilities. As these benefits are realized, organizations learn that to take full advantage of nimble application development and faster distribution of data with greater efficiency and availability, they need to evolve beyond a hybrid-cloud and towards a multi-cloud strategy. Cloudian enables the benefits of a multi-cloud by having a single namespace that allows users to manage, protect and find data anywhere through a single framework via a common language and tools.
MULTI-CLOUD CHALLENGES

Once a hybrid or multi-cloud strategy is embraced, there are unique challenges for IT organizations along the way. One example is vendor lock-in. Each vendor typically has various unique APIs that prevent mobility of applications to competitive platforms. Also, public cloud vendors typically have different strengths and so it can be difficult to move applications between on-premise infrastructure and public cloud locations. In addition to those challenges, there is yet another siloed IT infrastructure to manage. Objects and files are not necessarily accessible by the same applications and tools throughout the multi-cloud environment.

HYPERSTORE UNIQUE FEATURES

Cloudian Hyperstore software is uniquely positioned to address multi-cloud challenges. Cloudian HyperStore offers a single namespace, search, multi-tenancy, WORM, hybrid cloud streaming and configurable storage policies with flexible protection levels and redundancy through ISA-L erasure coding, replication factors, data compression and server-side encryption whether on-premise, confined to a public cloud or both. With Cloudian HyperStore, seamless data management is possible allowing users on demand access to their data anywhere and anytime. Built on a robust object storage platform for effortless data sharing, cloud storage providers around the world use Cloudian HyperStore to deploy and manage both public and private clouds, while enterprises can take advantage of HyperStore capabilities on-premise, within hybrid clouds and even within a multi-cloud.

HYPERSTORE WITHIN A MULTI-CLOUD ARCHITECTURE

HyperStore enables a single namespace, a fully peer-to-peer architecture and converged access that allows for file and object data to become interchangeable. This architecture makes a single view of the data possible so that organizations can find and manage data anywhere; on-premise, in any public cloud provider or both concurrently using a single API that ties it all together. Thus HyperStore, a multi-cloud controller, provides a unifying element that combines cloud resources into a single management framework. This lets you centralize policy and reduce risk with granular data management. The scale-out architecture of HyperStore is inherently highly-available and allows for parallel access of the data no matter where HyperStore is located. More nodes equate to more performance and greater data durability. Data is always stored in the cloud’s native format, whether that be in Microsoft Azure, Amazon Web Services (AWS) or Google Cloud Platform (GCP) in a single-shared highly-scalable data pool.

![Single Data Pool — Files & Objects](image-url)

Figure 1 – Multi-cloud Deployment Model
HYPERSTORE MULTI-CLOUD UNIQUE FEATURES

As you’ve seen, HyperStore limitlessly scales on demand both on-premise and in the cloud, which easily enables hybrid and multi-cloud architectures. With a single namespace, you can easily move your data between on-premise and the public cloud. This is enabled by HyperStore’s auto-tiering capability that allows data to move on-premise or from another cloud provider immediately, or on a defined daily and/or weekly schedule to a destination of your choice. Amazon has set the cloud storage standard via the S3 API making it the largest object storage environment and with HyperStore as the most compatible S3 storage system, the amount of applications HyperStore can leverage continues to grow. The next few sections dive into the details on each of these unique multi-cloud features.

HYPERSTORE LIMITLESSLY SCALES ON DEMAND

No matter your storage demand, you can seamlessly grow your storage in any dimension as fast or as slow as desired. You can add heterogeneous datacenters and regions and/or add a single node or multiple nodes, all in one operation. You can even add a different number of nodes in each of your datacenters (DC) if your storage demand requires it. For example, you could add five nodes in DC1, eight in DC2 and one in DC3. These DCs could be physical or within a public cloud provider or both! For example, DC1 might be on-premise, DC2 may be within Amazon AWS and DC3 within Microsoft Azure.

 Along with limitless scalability, Cloudian HyperStore offers true enterprise data protection. With replication factors and the ISA-L Erasure Coding, Cloudian HyperStore optimizes storage protection for all data objects. Data protection and durably choices are flexible, enabling efficient storage redundancy to meet your specific business needs. As HyperStore scales, the number of data protection choices only increase either on-premise or in the public cloud or both, providing many options solely based on the needs of an organization.

EFFORTLESS DATA MOVEMENT BETWEEN ON-PREM AND THE CLOUD

Not only does Cloudian HyperStore scale on demand with robust data protection, it simplifies data management. Cloudian HyperStore enables storing and retrieving your data where you want, when you want, using unique features like object streaming and dynamic auto-tiering to public cloud providers. Data can move seamlessly between your on-premises storage and the public cloud regardless of data type and size. For example, it is possible to move data collected on premise, sync the data to the cloud, generate custom meta-data using tools in the cloud, sync the meta-data to local storage and then perform machine learning or intelligent search analysis.

AUTO-TIERING

Cloudian enables seamless integration with on-premise HyperStore cloud storage and the public cloud. HyperStore supports auto-tiering from a local HyperStore bucket to any of several types of destinations systems including S3-compliant systems: Amazon S3, Amazon Glacier, Google Storage Cloud, a HyperStore region or system, or a different S3-compliant system of your choosing. In addition, HyperStore supports auto-tiering to Microsoft Azure and Spectra Logic Black Pearl. Auto-tiering is configurable on a per-bucket basis and can be enabled to happen immediately, or on a defined daily and/or weekly schedule.

S3 COMPATIBILITY WITH BROAD APPLICATION SUPPORT

With Amazon setting the cloud storage standard making it the largest object storage environment, and Amazon S3 API becoming the de facto standard for developers writing storage applications for the cloud infrastructure, it is imperative that every cloud, hybrid storage solution be S3 compliant. Cloudian HyperStore, is the most S3 compliant storage platform. With complete S3 compatibility, Cloudian HyperStore ensures seamless S3 integration with every available AWS/S3 application. Cloudian HyperStore allows unmatched customer choice in deploying applications and storage on-and off-premises. The highly active S3 developer community generates lots of innovative applications in categories including: enterprise secure file sharing; backup, data retention and archiving; NFS/CIFS gateways; and desktop file storage and backup; Cloudian HyperStore uniquely supports them all.
HYPERSTORE MULTI-CLOUD USE CASES

As you’ve seen, Cloudian HyperStore offers a scalable storage system with simplified data management, data protection, high availability, geo distribution and intelligent search capabilities via a unified software stack.

Seamless extension to other clouds is possible because HyperStore’s storage control and access layers are tightly integrated. There are no separate access layers, schedulers or controllers because HyperStore was designed from the beginning to have a simple and scalable architecture. This is unlike other object storage vendors that have bolt-on software and a distinct S3 gateway access layer. These systems lose flexibility, are complicated to scale and can have a single point of failure.

HyperStore instead provides an identical, tightly integrated software stack which can easily grow to include one or more public cloud infrastructures. HyperStore’s architecture provides flexible scalability options. The tight integration and unified management framework enable several multi-cloud use cases; cross-cloud data protection, flexible cloud procurement, file services, cloud bursting and cloud-based data analysis.

USE CASE SUMMARY
Cloudian HyperStore can be used in a number of ways to protect, replicate, deliver file services, meet on-demand workloads and analyze your data.

<table>
<thead>
<tr>
<th>USE CASE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross-cloud Data Protection</td>
<td>Replicate data objects from one cloud to another</td>
</tr>
<tr>
<td>Flexible Cloud Procurement</td>
<td>Choose the best cloud to meet your business needs</td>
</tr>
<tr>
<td>File Services</td>
<td>Deliver Windows and/or Linux file services from the cloud</td>
</tr>
<tr>
<td>Cloud Bursting</td>
<td>Migrate on-demand compute workloads from on-premise to the cloud</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>Use cloud compute resources to analyze on-premise data</td>
</tr>
</tbody>
</table>

CROSS-CLOUD DATA PROTECTION
Cross-cloud data protection is a hybrid-cloud or multi-cloud protection scheme. Cross-cloud data protection is possible because of HyperStore’s unified software architecture. With a single namespace, HyperStore can leverage cloud capacity and replication tools provided by the application or within HyperStore itself. With HyperStore, it is easy to manage data across clouds and to seamlessly replicate or migrate across clouds or across regions. HyperStore enables policy-based data replication on a per-bucket basis to maintain data protection through world-class data durability options. This allows an organization to have flexible protection choices to meet specific business needs. HyperStore is completely cloud agnostic and runs in all public cloud environments and uses a common API to access data anytime and anyplace. For example, it is possible to setup a policy-based data replication scheme between two different cloud providers as illustrated below.
FLEXIBLE CLOUD PROCUREMENT

HyperStore’s architecture enables flexible cloud procurement, where you can choose the best cloud provider for your organization’s business needs and applications. A particular cloud provider may be attractive due to cost, features, or location. Since HyperStore is cloud agnostic and has a unified API, it will run in any cloud or within multiple clouds. For example, it is possible to run S3-compatible applications in any cloud architecture. Typical use cases include data protection/archive, content distribution along with intelligent search capabilities and compute-intensive applications.

FILE SERVICES

HyperStore along with the HyperFile NAS controller enables enterprise-wide file access with no plugins required. Hybrid clouds that have a combination of public cloud and on-premise HyperStore enables file service flexibility. HyperStore in the public cloud can be leveraged for capacity and on-premise for fast access, data replication or other applications. Use any public cloud provider that meets your organization’s needs as HyperStore is cloud agnostic. HyperStore can be used for collaboration and/or content distribution applications on-premise and in Microsoft Azure, Google Cloud Platform, or Amazon AWS or any combination of cloud providers you choose.
**CLOUD BURSTING**

Cloud bursting leverages the capacity of the cloud. Run any application, object or file anywhere whether that is on-premise or across clouds. Seamlessly synchronize your data for application portability. Your application may have seasonal demands requiring additional resources. Add compute, storage capacity and performance to meet the demands of your application and then remove from the public cloud when no longer needed. This right-sizes your application needs and reduces the total cost of ownership (TCO) and improves your return on investment (ROI).

**DATA ANALYSIS**

Cloudian HyperStore contains intelligent search and analytics capabilities to extract business value from your data. Leverage cloud resources for cognitive analysis such as data analytics, machine learning and artificial intelligence workloads. You can gather data on-premise and synchronize to the cloud of your choice for analysis, generate custom metadata and synchronize metadata from HyperStore to your local primary storage for high-performing data analysis workloads.

**HYPERSTORE CLOUD DEPLOYMENT EXAMPLES**

In this section, practical deployment examples are discussed. These specific examples shouldn’t be used in production as they are meant to demonstrate basic functionality and provide a foundation for future technical discussions when thinking about deployment scenarios. Bear in mind that every environment is different, so it is important to dive into the specific business needs and security policies for any project under consideration.

**HYBRID CLOUD TIERING EXAMPLE**

Tiering moves object data from a source HyperStore bucket to a cloud provider destination bucket. By default, the metadata stays on HyperStore. Object data can also be restored back to the source bucket, so tiering is one method of cross-cloud data protection. Objects are typically tiered to a public cloud provider. Other possible destinations include Spectra BlackPearl, or even another Cloudian HyperStore system.

This example demonstrates cross-cloud data protection from an on-premise HyperStore bucket to a Microsoft Azure Blob Storage Container using tiering. After the data is tiered, it is stored in the Azure container in the native format. The example assumes that HyperStore is already deployed and configured with a group, user and storage policy. Furthermore, it is assumed that a source bucket created by a user already exists with no
For our example, please note the following considerations:

- Microsoft Azure does not support buckets on HyperStore configured for versioning
- Any data that exists prior to configuring tiering will remain in the source bucket
- Select Bridge Mode to tier objects immediately upon arrival to the HyperStore bucket
- If the Retain Local Copy box is checked the objects will be copied, not moved from the bucket for a fixed number of days. Both the object metadata and data will remain on HyperStore.
- Within the Cloudian Management Console (CMC), the object icon will change to blue after a tiering operation takes place. This shows the meta-data is still present and data has moved or copied to the destination bucket/container.
- Restores of the object data are possible after the data is tiered. There is typically a one-day minimum wait time. During the wait time, the object icon will change to red indicating a restore is scheduled.
- For security purposes, if the bucket’s lifecycle tiering policy is modified, the secret key must be entered again for the changes to take effect

To tier objects from a HyperStore S3 bucket to Azure, do the following steps:

1. Log into the Microsoft Azure portal and create storage account if it does not already exist.
2. In Azure create a blob container as a HyperStore bucket tiering target.
3. From Azure, gather the following configuration information for later configuration of tiering on HyperStore and save to a temporary text file.
   a. The Account Name, Account Key and Azure Container Name.
4. On HS login to the Cloudian Management Console (CMC) as Admin and do the following:
   a. Navigate to **Cluster -> Cluster Config -> Configuration Settings** and enable **Auto-Tiering** as shown.

   ![Auto-Tiering Configuration Settings](image)

   b. Click on **Save** at the bottom of the window. It may take a few minutes for the changes to propagate to all of the nodes in the cluster.

   ![Save Button](image)

   c. Log out as admin user

5. From the CMC login page, select the desired Group and User and login. The example uses TestGroup and the username is User1.

6. Click on the **Buckets & Objects** tab to navigate to the bucket configuration page.
7. From the Buckets tab click on Properties -> Lifecycle Policy. The following window appears.

8. Click on + Add New Rule
   a. Enter a Rule Name
   b. Leave Object Prefix blank to tier all files in the bucket
   c. Select ENABLE TIERING to configure the tiering options and schedule

9. The Object Tiering window appears after checking the box ENABLE TIERING above.
   a. Check Tier to Azure under Destination
   b. Enter Azure credentials under the END POINT from text file created earlier. Type in or copy the Account Name, Account Key and Container Name.
   c. Under Schedule, check Current Version since Microsoft Azure does not support versioning. The example shows that only objects 7 days older than After Date will be tiered to the Azure container. Please note the following:
      - Previous Version does not apply and should not be selected
      - Minimum time is tier 1 day after last access time.
   d. In the example, we wish to keep a local copy of the object since we are using the destination Azure container for protection purposes. To keep a local copy, enable Retain Local Copy. Note that a 7-day retention time is set. After aging 7 days, the objects will be deleted from the HyperStore bucket, but the metadata will remain.
   e. In the example Bridge Mode is selected. Any new object that appears in the bucket will be immediately be PUT to the destination Azure container. In this configuration, the HyperStore bucket effectively acts as a passthrough bridge.
   f. Keep all other options as default
   g. Click on Save.
10. The HyperStore tiering configuration is completed. Be sure a green bar appears at the top of the window that appears showing a status of success.

11. When completed successfully, the Lifecycle Policy tab will show the new rule for the bucket.

12. Run a test to verify that tiering to the Azure container is working.
   a. Use the CMC or another S3 tool, such as s3cmd or Cloudberry Explorer, to upload a 10 MB test file to the HyperStore bucket.
   b. Verify the object exists in the bucket
   c. Since Bridge Mode is enabled, the object will tier to the Azure container immediately.
   d. Log in to the Microsoft Azure storage account and click on Refresh. Verify the container contents shows the test file.
MULTI-CLOUD DEPLOYMENT DEMONSTRATION

Here we demonstrate as a proof of concept, how to install and configure a 9 node Cloudian HyperStore cluster in a multi-cloud within Amazon AWS, Microsoft Azure and Google GCP, the three most popular IaaS (Infrastructure as a service) public cloud providers.

The primary goal is to pilot the solution and show that the connectivity works between the nodes which spans all three IaaS providers. A second goal is to demonstrate that a distributed erasure coded (EC) storage policy works throughout the cluster. The diagram below shows the connectivity and private IP address spaces.

MULTI-CLOUD DEPLOYMENT NOTES

In the following sections there will be detailed instructions on how to setup each cloud provider networking configuration so that a single 9-node HyperStore cluster can be deployed that spans all three of the IaaS providers. Once complete, HyperStore needs to be installed on the VM instances and then configured to use a distributed EC storage policy. After confirming a successful deployment, a 10 MB test file will be uploaded to the HyperStore cluster and using the object locator, we will show that the object fragments are distributed across each of the three IaaS providers, Amazon AWS, Microsoft Azure and Google GCP.

Each of these IaaS providers will have their own private LAN along with an externally public-facing IP address. To connect each of these sites using the IaaS public-facing IP addresses, we will use VPN encrypted tunnels. Note this technique requires that traffic go through the internet, so in this demonstration the network connectivity will neither be fast nor secure. Due to these limitations it is NOT RECOMMENDED to use this technique for production. In production, typically there will be at least one on-premise site and dedicated network infrastructure, such as AWS direct connect, for optimal performance and high security.
MULTI-CLOUD INFRASTRUCTURE DEPLOYMENT REQUIREMENTS

Below is a brief list of requirements to deploy a HyperStore cluster that spans each of the three IaaS public cloud providers.

- An account for each IaaS cloud provider
- Interconnect 3 different private subnets on 3 different public clouds using public IP addresses.
- Note the difference in VPN secure protocol support between providers.
  - AWS supports only IKEv1. Azure supports IKEv2. GCP - supports both security methods
- Security and firewall considerations
- Implementing High Availability infrastructure within each of the IaaS providers
- Low latency WAN connections
- Routing table configuration on each IaaS provider
- Virtual VPN appliances as needed. AWS requires a Windows based RRAS EC2 instance for example.
- HyperStore VMs each running RHEL or CentOS 7.x with specific minimum sizing requirements

PRIVATE NETWORK IP CONFIGURATION

For the demonstration HyperStore cluster, we used the following subnets for the private networks.

- 10.240.0.0/24 – Amazon AWS
- 10.230.1.0/24 – Microsoft Azure
- 10.230.2.0/24 – Google GCS Network

Encrypted VPN Deployment Notes

- Microsoft Azure requires a Virtual Private Gateway for the encrypted VPN tunnel which takes about 45 minutes to deploy.
- Amazon AWS requires an elastic IP address and a Windows 2012 R2 installed on an EC2 instance configured with RRAS for the encrypted VPN tunnel to Microsoft Azure.
- Google GCP requires no special changes or virtual appliances since it supports both IKEv1 and v2
- Keep in mind that for VPN encryption Amazon AWS supports only IKEv1. Microsoft Azure supports IKEv2, while Google GCP supports both encryption methods.

Network Security Notes

Note that on each of the IaaS providers AWS, Azure and GCP - the security groups and/or firewalls are not configured for robust security. The security group is configured to allow for all traffic on all ports for 0.0.0.0/0. In production, specific ports should be opened for outbound and inbound connections.

TERMINOLGY NOTES

- Amazon AWS - VPC, Subnet, Route Table, Elastic IP, Security Group, Customer Gateway, Virtual Private Gateway, VPN, EC2 VM
- Google GCP - Project, VPC, External IP Addresses, Firewall rules, VPN, Compute Engine VM Instances
AMAZON AWS CONFIGURATION

Below are the steps to configure the AWS networking infrastructure to support HyperStore connectivity to Google GCP using a Virtual Private Gateway

1) Login to Amazon AWS account and create a VPC. Make sure an elastic IP address is configured for external public access.

![Create VPC](image)

2) Create a subnet in the VPC. This will be the AWS private network IP address range.

![Create Subnet](image)

3) Create a new Route Table

![Create Route Table](image)
4) Create a Customer Gateway for Google GCP private network connectivity. Note the public facing IP address is 35.225.187.109

5) Create a Virtual Private Gateway to Google GCP
6) Create VPN to GCP
7) Download the VPN configuration. Make sure to download and save the shared IKEv1 shared encryption key. This configuration will have endpoints to use and IPSec shared key. Choose vendor agnostic format when downloading the VPN configuration.
**GOOGLE GCP CONFIGURATION**

Create and configure a Google GCP to accept the secure VPN connection from Amazon AWS.

1) Login to Google GCP and create or select a project. Create a new VPC Network.

![Google Cloud Platform](image)

2) Create a VPN connection to Amazon AWS with the following steps

   a) Network for VPN should be configured as multicloud-vpc
   b) Region for VPN should be the same as the region for VPC
   c) IP address of VPN is the GCP public IP address we reserved before.
   d) Under tunnel - Remote Peer Address is the IP of AWS VPN connection.
   e) Pick IKEv1 as IKE security method
   f) Enter the shared key you downloaded from AWS
   g) Pick Route based routing and enter the AWS subnet.
### Google Cloud Platform - Interconnect

**Create a VPN connection**

A virtual private network lets you securely connect your Google Compute Engine resources to your own private network. Google VPN uses IKEv1 or IKEv2 to establish the IPSec connectivity. [Learn more](#).

**Google Compute Engine VPN gateway**

- **Name**: multicloud-vpn
- **Description**: (Optional)
- **Network**: multicloud-vpc
- **Region**: us-central1
- **IP address**:  (Optional)

**Tunnels**

You can have multiple tunnels to a single Peer VPN gateway.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description (Optional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>multicloud-vpn-tunnel-1</td>
<td></td>
</tr>
</tbody>
</table>

- **Remote peer IP address**: 52.45.53.239
- **IKE version**: IKEv1
- **Shared secret**: mysharedkey
- **Routing options**: Dynamic (BGP), Route-based, Policy-based
- **Remote network IP ranges**: Enter multiple IP address ranges (in CIDR notation) by pressing Enter after each one. 10.240.0.0/24

[Add tunnel](#)
3) After successful secure VPN tunnel handshake, you should see the following messages in Google GCP and Amazon AWS.
MICROSOFT AZURE CONFIGURATION
This section discusses how to configure Microsoft Azure networking infrastructure to support HyperStore encrypted VPN connectivity to Amazon AWS and Google GCP private networks.

1) Login to the Microsoft Azure portal and create a Virtual network.
2) Create a Local Network Gateway to connect to the Google GCS public facing IP address
3) Next create a second Local Network Gateway to connect to Amazon AWS
4) Create a Virtual Network Gateway used for the secure VPN tunnel.
5) Create a VPN connection from Microsoft Azure to Google GCP. Microsoft Azure and GCP support IKEv2, so you can pick a shared key and add a tunnel on Google GCP. Use the same shared key to create a secure VPN connection on Microsoft Azure. After adding the VPN connection, you should see the following screens:

6) This completes the Microsoft Azure to Google GCS VPN network configuration.

**CONFIGURE WINDOWS RRAS ON AMAZON AWS**

This section discusses how to create and configure a secure VPN connection between Amazon AWS and Microsoft Azure. Since Azure doesn’t support IKEv1 and AWS doesn’t support IKEv2, we had to install a Windows 2012 VM on AWS multicloud-vpc and configure and enable Microsoft RRAS on it. This VM is used to establish a secure VPN tunnel from the private Microsoft Azure network to the private Amazon AWS network.

*Note: Another method is to use RHEL7.x and the strongSwan secure VPN application. We don’t cover the detailed instructions here.*

1) Login to Amazon AWS account.

2) Launch a Windows 2012 VM as an EC2 instance in Amazon AWS on the same private network created in an earlier step.

3) Associate the Elastic IP address that was reserved in the first steps with this Windows 2012 VM that will run Windows RRAS.

4) Disable Source/Destination check on this IP. This is an important step as we couldn’t establish the secure VPN tunnel if enabled.

5) Finally update Route Tables in Amazon AWS to route to 10.230.1.0/24 and 10.230.2.0/24 which are the private networks for Microsoft Azure and Google GCS.
4) Enable and configure RRAS on Windows 2012. See the following screenshots for details on setup. Please ensure the security tab configures the secure VPN using IKEv2 to connect the Amazon AWS network with Microsoft Azure.
5) Once this is done, you should see something similar on Microsoft Azure.

6) Finally, update the Route tables in Amazon AWS to route traffic to 10.230.1.0/24 and 10.230.2.0/24
DEPLOY CLOUDIAN HYPERSTORE

Now the private networks and VPN encrypted tunnels are setup and configured, the next step is to deploy the HyperStore software. Detailed instructions for deployment were followed in the Cloudian HyperStore Installation Guide. The following notes provide a summary of what was done for the nine-node cluster in the multi-cloud demonstration. Each IaaS provider is configured as a separate data center (DC) in HyperStore.

1) Bare minimum VM specs for HyperStore are as follows:
   a) 4 vCPUs and 16 GB of RAM for each.
   b) 30 GB boot disk.
   c) Two additional 500 GB disks per VM for data
   d) RHEL 7.x as the base OS.

2) Spin up three VM’s in Microsoft Azure, Amazon AWS and Google GCP. This gives a total of 9 VMs in the HyperStore cluster. Each of these sites are configured as a separate DC in the HyperStore survey file.

3) Once the VMs are deployed test each of the private networks and secure VPN connections from each HyperStore VM in Amazon AWS, Microsoft Azure and Google GCP using the ping command. Ping works because the network security groups are complete open to all network traffic.

4) Test root access to each of the VMs. There were no issues with EC2 VMs in Amazon AWS or Microsoft Azure VMs by default. Configuring VM’s on Google GCP was difficult because root access is disabled by default. We had to modify /etc/ssh/sshd_config to allow root access for all of the 9 nodes.

5) Download the HyperStore 7.x binary and install using the cloudianInstall script.

6) We used the first AWS VM as the HyperStore Puppet Master. All the remaining nodes are deployed from there. It is critical to ensure that both network connectivity and root access are working properly first. See step 4 above.

7) We used 3 DCs for the HyperStore configuration, one each for AWS, Azure and GCP.

8) We did not have a DNS server so DNSMASQ option used for deployment.

9) From the cloudianInstall script make sure to select R to run the Pre-Install Checks on all HyperStore nodes to verify all the pre-reqs are met before deploying HyperStore.
   a) When the pre-install check completes, correct any errors and consider recommendations
   b) Since deploying on VMs, ignore the VM warnings and continue

10) When the HyperStore deployment is complete, run the S3 Validation tests to ensure the HyperStore cluster is online.
    a) Select option 2 from the cloudianInstall.sh and then select option d. Groups, users and buckets will be created and then deleted as part of the validation tests. Make sure the tests pass before continuing. These tests will verify the HyperStore cluster is working properly and is operational.
CONFIGURE HYPERSTORE WITH A DISTRIBUTED EC STORAGE POLICY

With the HyperStore cluster now online, we will create a storage policy to show that objects in S3 buckets will be spanned across all IaaS providers and on all nodes. A distributed EC storage policy will accomplish this objective. With this policy enabled on a bucket, we will show that object fragments will exist on all 9 nodes in the cluster and span the nodes in Amazon AWS, Microsoft Azure and Google GCP.

1) Log into the HyperStore CMC web UI using the default admin user to create a test group, user, bucket and a storage policy called “distEC” with the EC across datacenters option selected.
2) Create a new bucket for the test user and apply the storage policy.
3) Upload a 10 MB file to the bucket using the CMC.
4) Observe that the object fragments exist on all nodes. Navigate to the object locator to view the fragments and locations.
   a) Type in a bucket name
   b) Type in an object name
   c) Leave the Version field blank
5) Click on Find to show the object fragments and locations. See example below for sample output.

![Object Locator Image]

- NODE IP ADDRESS: List of nodes where object fragments are stored.
- FILE PATH: Path to the fragment.
- LAST MODIFIED: Time the fragment was modified.
- VERSION: Version of the fragment.
- DIGEST: Hash of the fragment.

Example output:

- NODE 1: Fragment exists on node 1.
- NODE 2: Fragment exists on node 2.
- NODE 3: Fragment exists on node 3.
- NODE 4: Fragment exists on node 4.
- NODE 5: Fragment exists on node 5.
- NODE 6: Fragment exists on node 6.
- NODE 7: Fragment exists on node 7.
- NODE 8: Fragment exists on node 8.
- NODE 9: Fragment exists on node 9.
VIEW SUMMARY FOR DETAILS
What follows are sample screenshots and basic descriptions of what will be seen in the CMC manager.

SIMPLE DASHBOARD
The CMC dashboard provides a high-level view of the status of your Cloudian HyperStore object storage service. If you have multiple service regions, there is a separate dashboard view for each region.

SINGLE CLUSTER USAGE & PERFORMANCE VIEW
View cluster usage graphs that cover the past 30 days of activity.

CAPACITY EXPLORER
With the CMC’s Capacity Explorer page, you can view your available S3 object data storage capacity by region, by data center, and by node. First (if you have a multi-region system) choose a region tab at the top of the page. Then, in the graphical display:

- The inner circle represents the service region as a whole
- The middle circle has one segment for each data center in the region
- The **outer circle** has one segment for each node in each data center.

The circle segments are color-coded as follows:

- **Green** indicates that free space is 30% or more of total space for that region, data center, or node. (Slightly different shades of green are used merely to differentiate the concentric circles from each other. Green has the same meaning regardless of the shade of green.)
- **Orange** indicates that free space is between 10% and 29% of total space for that region, data center, or node.
- **Red** indicates that free space is less than 10% of total space for that region, data center, or node.

**VIEW USER AND TENANT USAGE**

In the CMC’s “Usage by Users & Group” page you can generate service usage reports for individual users, for user groups, and for the system as a whole.
STORAGE POLICIES

Central to Cloudian’s data protection are its storage policies. These policies protect data that ensure data durability and high availability to users. The Cloudian HyperStore system lets you pre-configure one or more storage policies and are applied on a per-bucket basis. Users when they create a new storage bucket can then choose which pre-configured storage policy to use to protect the data. For each storage policy that you create, choose from either of two data protection methods: replication or erasure coding.
CLUSTER CONFIGURATION & MONITORING

MULTI-DATACENTER & REGION VIEW

The datacenters page displays a panel for each datacenter in your Cloudian HyperStore system. For each datacenter, each Cloudian HyperStore node in the datacenter is represented by a cube.

VIEW NODE STATUS

For each datacenter, each Cloudian HyperStore node in the datacenter is represented by a cube. Clicking on the cube, allows you to see view node activity and manage node services.
SIMPLE CLUSTER SETTINGS
The Cluster Information page displays static information about your Cloudian HyperStore system such as software version, service hosts and license info. The configuration page allows for globally modifying and updating configurations files, requiring no service restart.

<table>
<thead>
<tr>
<th>CLUSTER INFORMATION</th>
<th>CONFIGURATION SETTINGS</th>
<th>AUTO TIERING CREDENTIALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ SMTP/Email Settings for Alerts/Notifications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ SNMP Trap Destination Settings</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ Usage Tracking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ Quality of Service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ Auto Tiering</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enable AutoTiering</td>
<td>Enabled</td>
<td>Edit</td>
</tr>
<tr>
<td>Default Tiering URL</td>
<td><a href="https://us-ga1.1.com/virtualmachine.com/GO21">https://us-ga1.1.com/virtualmachine.com/GO21</a></td>
<td>Edit</td>
</tr>
<tr>
<td>▶ S3 Request Restrictions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>▶ Auto Replication Schedule</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| SAVE |

SIMPLE NOTIFICATIONS & ALERTS
The Cloudian HyperStore system comes with a set of pre-configured notification rules. The pre-configured notification rules are listed in the “Rules” section of the “Notification Rules” page.
CONCLUSION

Cloudian HyperStore software makes it easy to build fully-featured, Amazon S3-compliant cloud storage, on-premises in a private cloud, a hybrid cloud or in multi-cloud for cross cloud data protection and migration, cloud bursting and/or hybrid cloud data analysis. HyperStore is available as either stand-alone software, or as Cloudian HyperStore appliances. Either way, Cloudian HyperStore storage software ensures unlimited scale that can extend from on-premise to hybrid clouds or even the multi-cloud. Features such as a single namespace, fully automated data tiering and bucket versioning allows for secure support of all S3 applications on-premise or public cloud or both. We illustrated with detailed instructions on how to deploy HyperStore in a hybrid cloud and multi-cloud.

Cloudian HyperStore provides highly efficient storage and seamless data management that lets users store and access their data where they want it, when they want it. Built on a robust object storage platform for effortless data sharing, cloud service providers around the world use Cloudian HyperStore to deploy and manage both public and private clouds, while enterprises rely on it to maintain their private and hybrid clouds which can extend to a multi-cloud environment.