TECHNICAL WHITE PAPER

Utilizing AWS PrivateLink with Rubrik CDM

Joe Kelly and Amir Khayatbashi
March 2021
RWP-0575

TABLE OF CONTENTS

INTRODUCTION

Welcome to *How It Works: Utilizing Amazon PrivateLink with Rubrik CDM*. The purpose of this document is to aid the reader in familiarizing themselves with the features, architecture, and workflows of Rubrik's CDM capabilities utilizing Amazon PrivateLink. Such information will prove valuable while evaluating, designing, or implementing the technologies described herein.

This solution walkthrough will acquaint users with how to configure Amazon S3 as a Rubrik Cloud Data Management (CDM) platform archival location utilizing Amazon PrivateLink Interface Endpoints. Users will be provided a step-by-step walkthrough of the configuration process and an understanding of the design choices and recommended practices for Rubrik CloudOut with Amazon S3.

AUDIENCE

This guide is for anyone who wants to better understand the usage of CloudOut with Amazon PrivateLink support for S3. This includes architects, engineers, and administrators responsible for AWS infrastructure and data protection operations as well as individuals with a vested interest in security and networking.

OBJECTIVES

The goal of this guide is to provide the reader with a clear and concise point of technical reference regarding architecture and workflows utilized by Rubrik CDM and Amazon PrivateLink. The documented methods within are but two options for utilizing PrivateLink for S3. However, any S3 bound communications can utilize PrivateLink whether that traffic originates from on-premises or from within a VPC itself. This could include cloud compute traffic for CloudOn and Archive Consolidation. This document addresses only CloudOut architectures with AWS.

WHAT IS AWS PRIVATELINK?

AWS PrivateLink provides a means for customers to privately and securely connect their resources, whether on-premise or within AWS, to AWS services using Amazon's private network. Prior to AWS PrivateLink, such services were accessed through public IP addresses, VPC Internet Gateway's, VPC Peering or Customer Proxy's. AWS PrivateLink provides several advantages over these traditional methods, such as:

- Private IP address space for traffic in and out of AWS.
- Vastly simplified and scalable network management. The need for complex networking across multiple VPC's, route table modifications, or whitelisted public IP addresses is no longer necessary.
- Reduced data transfer charges for communications between private VPC resources and AWS services. Primary cost savings realized would be for data egress charges for communications that would typically be over the internet and the absence of the logical network entities to support that communication.

By utilizing AWS PrivateLink, AWS and Rubrik customers can safely, securely and efficiently connect to a CloudOut Archive location over your private network. The following diagram shows a traditional CloudOut Archive architecture utilizing S3's public regional endpoints and will be used as a point of reference going forward.

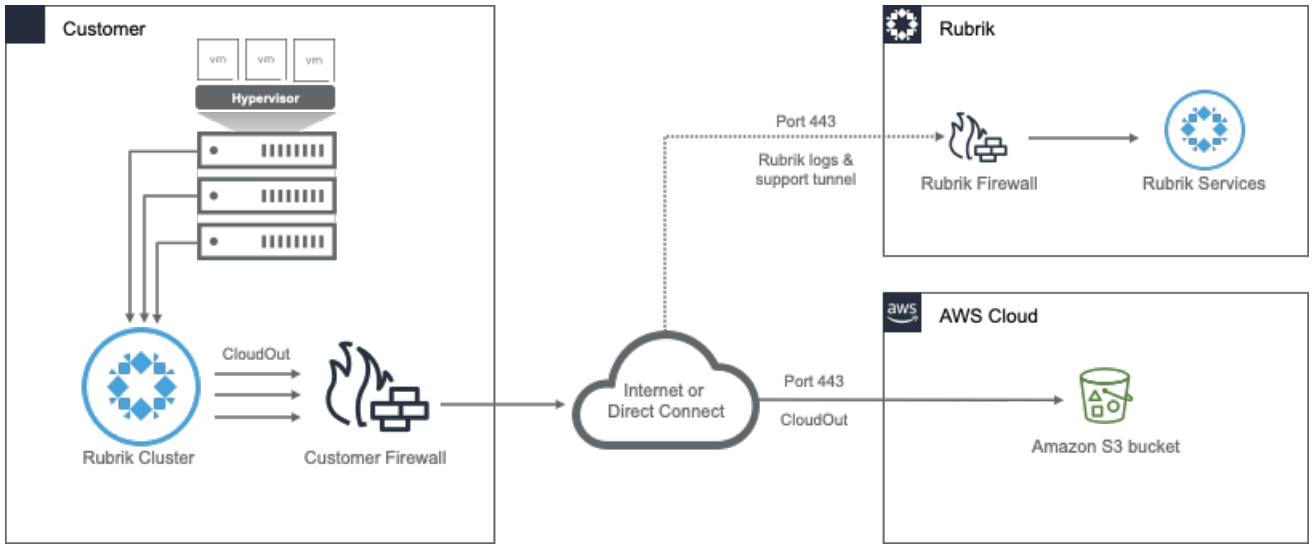

Figure 1 - High Level Architecture: CloudOut for AWS (no PrivateLink)

For information on PrivateLink please refer to the following: https://aws.amazon.com/privatelink

OPTION 1: UTILIZING LOCAL HOST ENTRIES FOR CDM

For those customers, looking for a simplified approach to utilizing Amazon PrivateLink support for S3, CDM local host mappings can be implemented. Not without its constraints, host mappings (like local windows hosts files) are static in nature and typically require manual efforts to accommodate changes in your environment. In addition, in the unlikely event that an Amazon Availability Zone goes offline, multiple host mappings can be utilized pointing to the same bucket name with different interface endpoint IP's. Host entries provide a very simple solution in circumstances where external DNS cannot or won't be utilized for S3 bucket resolution.

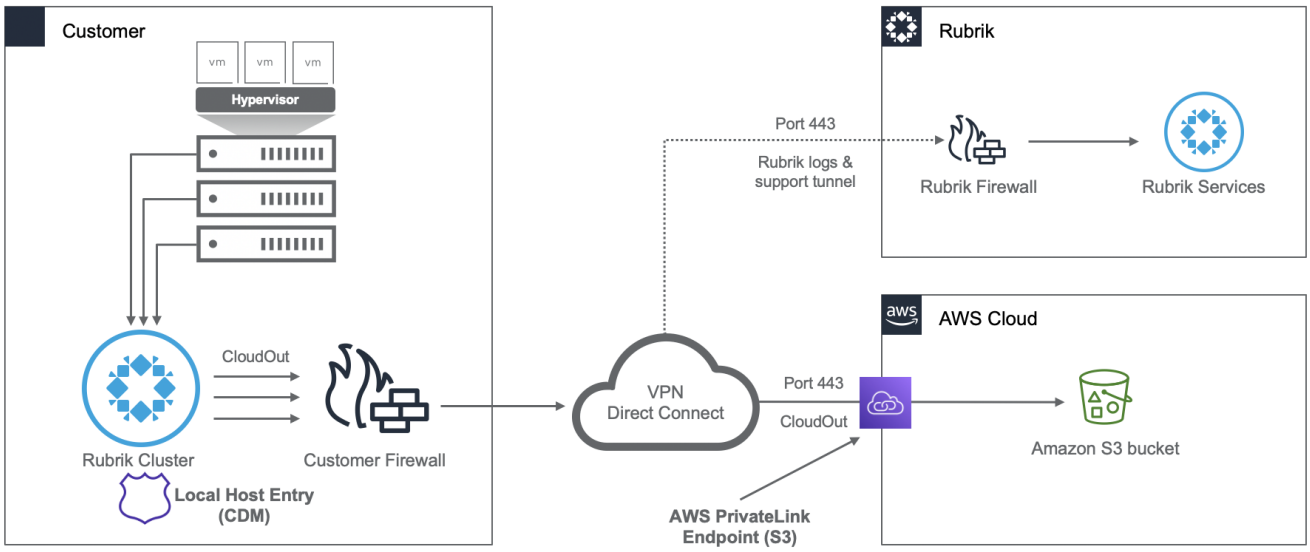

Figure 2 - High Level Architecture: CloudOut for AWS (PrivateLink Enabled)

At a high level, the following workflow is used when configuring CloudOut with AWS PrivateLink support for Amazon S3.

1. Complete CloudOut steps within CDM prior to moving to step 2.

a. For CloudOut configuration prerequisites and steps, please refer to the Rubrik CDM User Guide on the Rubrik Support Portal.

2. Create PrivateLink Interface Endpoint for S3 for AWS accounts where CloudOut Archive Bucket exists.

    a. Search for "Endpoints" within the AWS Management Console

    b. Select "Create Endpoint"

        i. Service Category - AWS Services

        ii. Service Name - `com.amazonaws.<REGION>.s3` (substitute <REGION> for your specific region).

            1. Owner - `Amazon`

            2. Type - `Interface`

    c. Choose the VPC in which you want to create the endpoint.

    d. Choose at a minimum two subnets to mitigate single points of failure for the endpoint itself.

    e. Choose or create a security group that will allow communication between the endpoint network interface and the resources on-premises and or within the VPC to the service.

    f. Policy - `Full Access`.

        i. This policy is an IAM resource policy and is designed to provide another layer of protection for the endpoint service itself. This policy can be modified at any time to reflect the security posture of your organization.

    g. Choose `Create Endpoint`

    h. Once status says `Available`, document the following:

        i. Select the endpoint name and review the "Details" tab.

            1. Document the DNS names for the regional endpoint, i.e. `<vpceID>.s3.REGION.vpce.amazonaws.com`.

                a. Where <vpceID> is the ID of the VPC endpoint ID.

        ii. Select the "Subnets" tab, please document the following:

            1. IPv4 Addresses - these represent the private interface endpoint IP addresses across the availability zones selected during the endpoint creation process.

            2. Network Interface ID - ENI (Elastic Network Interface) of interface endpoint.

                a. *Note - this is helpful from a data flow and troubleshooting perspective as VPC Flow Logs are supported for ENI's.*

3. For CDM version 5.1 and above, the admin CLI can be utilized.
Note - Please contact Rubrik support for versions prior to 5.1.

    a. SSH to Rubrik cluster, login with admin account.

    b. Run command: `network configure_ip_hosts_mapping`

    c. Enter 1 to set entry.

```
VRECBBEE2A3DA >> network configure_ip_hosts_mapping
================================================
Configure IP and hostname mappings in /etc/hosts
================================================

Options: 1) Set entry
         2) Delete entry
         3) Show current entries
         4) Return to main menu
         0) Show options for this command

Please select your option: 1
```

Figure 3 - Host Mapping Command from CDM

d. Type the IP address of the private IP address(s) for the PrivateLink Interface Endpoint for the region where the archive S3 bucket will exist.

```
Set IP host mapping
Type IP address: 10.[redacted]
Hostname: examplebucketname.s3.us-east-2.amazonaws.com
Alias (Optional):
IP host mapping set

Updated entries (not commited yet):
{
    10.[redacted]: [
        examplebucketname.s3.us-east-2.amazonaws.com
    ]
}
Confirm, overwrite /etc/hosts with above entries?
Type "yes" to continue: yes
Commited
```

Figure 4 - Host Mapping Command from CDM (Option 1)

e. Type the FQDN of the bucket to be utilized for the archive location, i.e. "`<bucketname>.s3.<REGION>.amazonaws.com`.

   i. *Note - replace <bucketname> with your bucket name.*

   ii. *Note - replace <REGION>with your region of choice.*

   iii. *Note - a single Availability Zone will be utilized for each host entry. However, multiple host entries can be used pointing to the same S3 bucket name. CDM will randomly utilize one of the host mappings, specifically until that endpoint is no longer available. At that point, communications will failover to another endpoint within the mapping list. Once the originating endpoint is restored, communications will fail back to the endpoint in question.*

f. At the (Optional) 'Alias prompt hit <Enter>.

g. Type `yes` to confirm.

h. Type **4** and hit enter to return to the main menu.

i. Run: `network ping <bucketname>.s3.<REGION>.amazonaws.com`, to verify the correct resolved private IP address. Please note that interface endpoints do not return a reply for ICMP traffic.

   i. Note - Substitute <REGION> for the region location of the archive bucket in question, i.e. us-east-2

4. At this point, CDM is ready to add the archive target
5. Within CDM, add archive location (click on gear icon). The following information is required:
    a. Archive Type - AWS
    b. Region - Archive Bucket Region
    c. Storage Class - Standard
    d. AWS Access Key from user in Step 3
    e. AWS Secret Key from user in Step 3
    f. Bucket Name
    g. Archival Location Name - Autofilled by default
    h. Encryption Key - KMS (recommended)
    i. KMS Master Key ID - Required from Step 4
6. Enable the archival location for specific SLA Domain(s) (assign to resources).
    a. Enable Instant Archive to start archiving immediately.

OPTION 1: DESIGN CONSTRAINTS

With any design decision comes constraints, or rather imposed limitations that affect the overall design. These non-functional requirements can change over time and can therefore influence your functional requirements. For a customer looking to quickly utilize PrivateLink for S3, local host mappings is a quick and easy way to privatize your archive communication channel to AWS. However, utilizing this configuration does have its limits listed below:

- An IP host mapping is a static entry within CDM and therefore would require updating and additions if other archive locations are utilized.

OPTION 2: UTILIZING DNS WITH ROUTE 53 RESOLVERS

For those customers, looking for a more robust approach to utilizing Amazon PrivateLink support for S3, DNS with Route 53 resolvers can be utilized. The following services will need to be implemented from both a customer's premise and AWS environment in addition to the typical CloudOut requirements.

- On premise DNS with Conditional Forwarders.
    - i.e. Domain - `s3.<region>.amazonaws.com`
- AWS Route 53 resolvers (Inbound endpoints) in the region where the archive bucket exists.
    - At a minimum two private subnets will be required across two Availability Zones.
- AWS Route 53 private hosted zone.
    - Region specific
- AWS PrivateLink Interface Endpoint in the region where the archive bucket exists.
    - At a minimum two private subnets will be required across two Availability Zones.

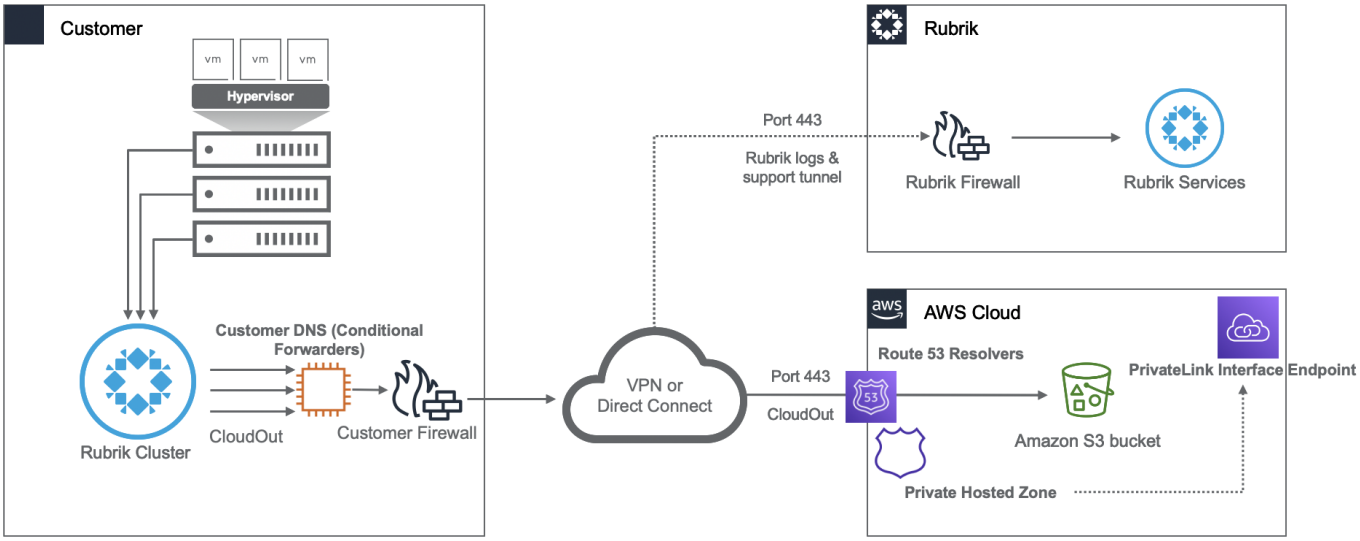

Figure 5 - High Level Architecture: CloudOut for AWS (Route53 Resolvers)

At a high level, the following workflow is used when configuring CloudOut with AWS PrivateLink support for Amazon S3 (using DNS and Route 53 Resolvers).

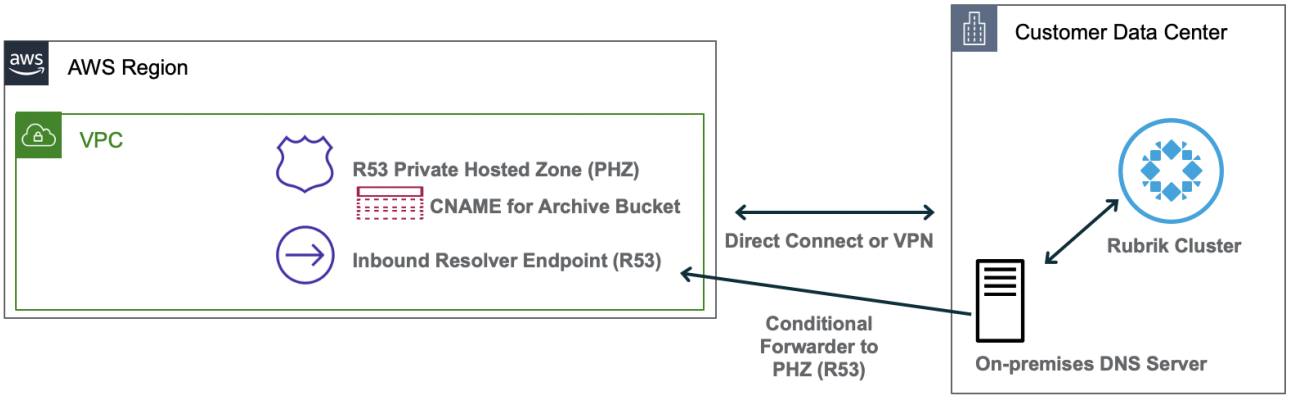

Figure 6 - High Level Architecture: Hybrid DNS Architecture

1. Complete CloudOut steps within CDM prior to moving to step 2.
    a. For CloudOut configuration prerequisites and steps, please refer to the Rubrik CDM User Guide on the Rubrik Support Portal.
2. Create PrivateLink Interface Endpoint for S3 for AWS accounts where CloudOut Archive Bucket exists.
    a. Search for `Endpoints` within the AWS Management Console
    b. Select `Create Endpoint`
        i. Service Category - `AWS Services`
        ii. Service Name - `com.amazonaws.<REGION>.s3` (substitute <REGION> for your specific region).
            1. Owner - `Amazon`

2. Type - `Interface`

c. Choose the VPC in which you want to create the endpoint.

d. Choose at a minimum two subnets to mitigate single points of failure for the endpoint itself.

e. Choose or create a security group that will allow communication between the endpoint network interface and the resources on-premises and or within the VPC to the service.

f. Policy - `Full Access`.

i. This policy is an IAM resource policy and is designed to provide another layer of protection for the endpoint service itself. This policy can be modified at any time to reflect the security posture of your organization.

g. Choose `Create Endpoint`

h. Once status says `Available`, document the following:

i. Select the endpoint name and review the "Details" tab.

1. Document the DNS names for the regional endpoint, i.e. `<vpceID>.s3.<REGION>.vpce.amazonaws.com`

a. Where <vpceID> is the VPC Endpoint ID.

b. Where <REGION> is the region of the VPC Endpoint.

ii. Select the `Subnets` tab, and document the following:

1. IPv4 Addresses - these represent the private interface endpoint IP addresses across the availability zones selected during the endpoint creation process.

2. Network Interface ID - ENI (Elastic Network Interface) of interface endpoint.

a. *Note - this is helpful from a data flow and troubleshooting perspective as VPC Flow Logs are supported for ENI's.*

3. Create an Inbound Endpoint within Route 53 Resolver.

a. Provide Endpoint Name.

b. Select VPC to which all inbound DNS queries will flow.

c. Provide two IP addresses across at a minimum two Availability Zones.

i. Subnets utilized should be private.

ii. Use an IP address that is automatically selected or one you specify.

1. *Note - an inbound endpoint will be required for each region that utilizes PrivateLink for S3 access.*

4. Document all resolver IP addresses once status is operational.

5. Create a new conditional forwarder within your organization's DNS server.

a. DNS Domain name - `s3.<REGION>.amazonaws.com`

i. *Note - a conditional forwarder will be required for each region you plan to archive to.*

6. Add Route 53 resolver IP addresses as conditional forwarder master servers.

a. If the master servers are Microsoft DNS servers, check `Store this conditional forwarder in Active Directory, and replicate it as follows: Choose all servers in Domain or Forest`.

7. Validate that an internal DNS server IP address has been added to Rubrik CDM under `Network Configuration/ Network Settings`

8. Add a Route 53 private hosted zone

    a. i.e. Domain - `s3.<REGION>.amazon.aws.com`

        i. Where <REGION> is the name of the S3 region to use.

    b. Record Name

        i. Record name - `<BUCKETNAME>.s3.<REGION>.amazonaws.com"`

            1. Replace <BUCKETNAME> and <REGION> with your specific information

        ii. Type - `CNAME`

        iii. Routing - `Simple`

            1. Note - This routing in particular will randomly route between the underlying endpoint IP addresses when the regional DNS name is utilized as the CNAME value.

        iv. Value/Route traffic to - Regional DNS name for interface endpoint, i.e. `<vpceID>.s3.<REGION>.vpce.amazonaws.com`.

            1. Where <vpceID> is the VPC Endpoint ID

            2. Where <REGION> is the region that the VPC Endpoint is in.

9. Within CDM, add archive location (click on gear icon). The following information is required:

    a. Archive Type - AWS

    b. Region - Archive Bucket Region

    c. Storage Class - Standard

    d. AWS Access Key from user in Step 3

    e. AWS Secret Key from user in Step 3

    f. Bucket Name

    g. Archival Location Name - Autofilled by default

    h. Encryption Key - KMS (recommended)

    i. KMS Master Key ID - Required from Step 4

    j. Once Complete, enable archival location for specific SLA Domains (assign to resources). Enable Instant Archive to start archiving immediately.

        i. *Note - Interface Endpoints do support VPC Flow Logs which can be used in tandem with AWS CloudWatch Log Groups for traffic flow validation and troubleshooting purposes.*

OPTION 2: DESIGN CONSTRAINTS

With any design decision comes constraints, or rather imposed limitations that affect the overall design. These non-functional requirements can change over time and can therefore influence your functional requirements. For a customer looking to utilize a more robust solution, extending to DNS is the most viable solution. This could take on many forms that reach far beyond this document. The documented methods within are but two options for utilizing PrivateLink for S3. The following items could be potential blockers for this solution:

- With robustness sometimes comes complexity. We are injecting additional services that may or may not be utilized in your environment and therefore may incur additional costs.
- Considering we are making use of conditional forwarders within DNS implies we are sending all S3 traffic to `s3.<REGION>.amazonaws.com` which will flow across the PrivateLink interface endpoint. This may or may not be desired for all S3 regional traffic in your environment.
- If additional buckets in additional regions are required, then this architecture would need to be repeated in all corresponding regions.

SUMMARY

This concludes the How it Works guide on Utilizing Amazon PrivateLink with CDM. This architecture allows Rubrik's customers to secure network traffic from on-premises workloads to Amazon S3, greatly minimizing threat exposures experienced over the public internet. Utilizing AWS Direct Connect or VPN, customers can use private IP connectivity to connect to S3 via the Amazon private network, all while simplifying network management and reducing costs. This frictionless architecture allows our mutual customers to rapidly exploit the best of breed data management solution backed by Rubrik and AWS.

VERSION HISTORY

Version	Date	Summary of Changes
1.0	March 2021	Initial Release

rubrik

Global HQ
1001 Page Mill Rd., Building 2
Palo Alto, CA 94304
United States

1-844-4RUBRIK
inquiries@rubrik.com
www.rubrik.com

Rubrik, the Multi-Cloud Data Control™ Company, enables enterprises to maximize value from data that is increasingly fragmented across data centers and clouds. Rubrik delivers a single, policy-driven platform for data recovery, governance, compliance, and cloud mobility. For more information, visit www.rubrik.com and follow @rubrikInc on Twitter. © 2021 Rubrik. Rubrik is a registered trademark of Rubrik, Inc. Other marks may be trademarks of their respective owners.

20210329_v1