White Paper

EMC CENTERA VIRTUAL ARCHIVE
Planning and Configuration Guide

Abstract

This white paper provides best practices for using EMC Centera Virtual Archive in a customer environment. The guide starts with an overview of Virtual Archive and discusses typical use case scenarios, performance considerations, scalability/sizing, and network latency.

May, 2013

EMC²

Table of Contents

Executive summary

EMC Centera® Virtual Archive is a new EMC technology that enables the aggregation of a set of EMC Centera clusters. This set of clusters forms a virtual, tamper-proof archive with the capacity of all clusters aggregated and available to applications. Once connected to the Virtual Archive, EMC Centera functionality will be accessible to applications in a seamless manner. Capacity and processing power can be added beyond the existing boundaries of a single EMC Centera system or a single data center.

EMC Centera Virtual Archive allows customers to federate multiple Centera systems to create a single, highly scalable, distributed digital archive. Most significantly, the EMC Centera Virtual Archive simplifies management of archives at scale and distance, delivering significant operating efficiencies and flexible allocation and reallocation of archive capacity.

Audience

This white paper is intended for technical pre-sales personnel who have a good technical background and understanding of storage technology. Familiarity with the existing EMC Centera components, concepts, and features is essential and recommended.

Terminology

The following table summarizes the commonly used terms used in this guide.

Terms	Definitions
ARM	Advanced Retention Management feature of EMC Centera
BLOB	Denotes the finite bit string that is the customer data. A single C-Clip™ can contain pointers to multiple BLOBs. The content of a BLOB is opaque to CentraStar®. BLOBs also have associated Content Addresses, but they are not exposed to the applications accessing EMC Centera. This is to ensure the path to the data is always through the appropriate pointer (C-Clip)
Centera SDK	Centera Software Development Kit
Centera Viewer (CV)	Management tool used for EMC Centera configuration
Centera Virtual Archive	A collection of federated clusters with Virtual Archive software
Centera Virtual Archive software	The Virtual Archive software that is hosted on one cluster on multiple cubes/nodes
Cluster	A cluster consists of one or more EMC Centera cubes interconnected to present a single storage array
Cube	A single EMC Centera unit containing a minimum of four and a maximum of 32[1] nodes including a set of cube switches
EMC CentraStar	EMC Centera firmware running on the Centera hardware
EOSL	End of Service Life
Federation	A collection of clusters either acting as a Host or Member System that

---

[1] GEN3 or older clusters have 32 nodes in a cube. GEN4 and GEN4-LP clusters have 16 nodes in a cube.

	together form a virtual archive
Federation Source	The Federation in a replication configuration that sends data to a remote site
Federation Target	The Federation in a replication configuration that receives data written to the Federation Source by the application(s)
GEN4-LP	Generation 4 Low Power Centera hardware
Host System	A cluster that executes the Centera Virtual Archive software
Member System	Any cluster in a Federation that is eligible to store and retrieve content and does not host the Virtual Archive software
Replication Source	The primary cluster in a replication configuration that sends data to a remote cluster
Replication Target	The cluster in a replication configuration that receives a copy of the content written to the primary cluster by the application(s)
Restore Source	The cluster where restore is started, from which content is copied to a specified restore target
Restore Target	The cluster that receives a copy of the content from a cluster running restore
LAN	Local Area Network
VIM	Vendor Interface Module
WAN	Wide Area Network

Table 1. Terminology

Virtual Archive

Virtual Archive is designed to provide customers the ability to address increasing storage demands by adding more EMC Centera storage in an effective manner. Virtual Archive addresses the need to expand archives into other data centers and increases the abstraction of the technology to also include software versions and other architectures. The latter is needed to ensure seamless access to a long-term archive and to deal with the unknown (not-yet-known new storage technology).

Virtual Archive is software-based and can be configured to run on any GEN4-LP running EMC CentraStar version 4.0+. Virtual Archive allows for scalability of EMC Centera clusters up to 512 nodes.

This section covers the main components and concepts of Virtual Archive used throughout the guide.

Federation

A Federation is a collection of clusters that together form a Virtual Archive. It is comprised of a single Host System and one or more Member Systems. After the Virtual Archive software is installed on standard EMC Centera GEN4-LP hardware, an EMC Centera service representative will need to perform the process of creating the Federation on the Host System and then adding additional Member Systems.

Virtual Archive software

Starting with CentraStar 4.2.2 the Virtual Archive software is part of the CentraStar software. The CVA routing software takes the responsibility of routing data traffic between the Host and Member Systems. When a request is received from an

application server(s), the Host System decides which system in the Federation will service the request. The request is then routed back via the Host System to the application server(s).

Figure 1 shows the Virtual Archive data flow in a federated environment comprised of a Federation of a Host and a Member System.

Note: Applications will always need to connect to the Host System regardless of its location.
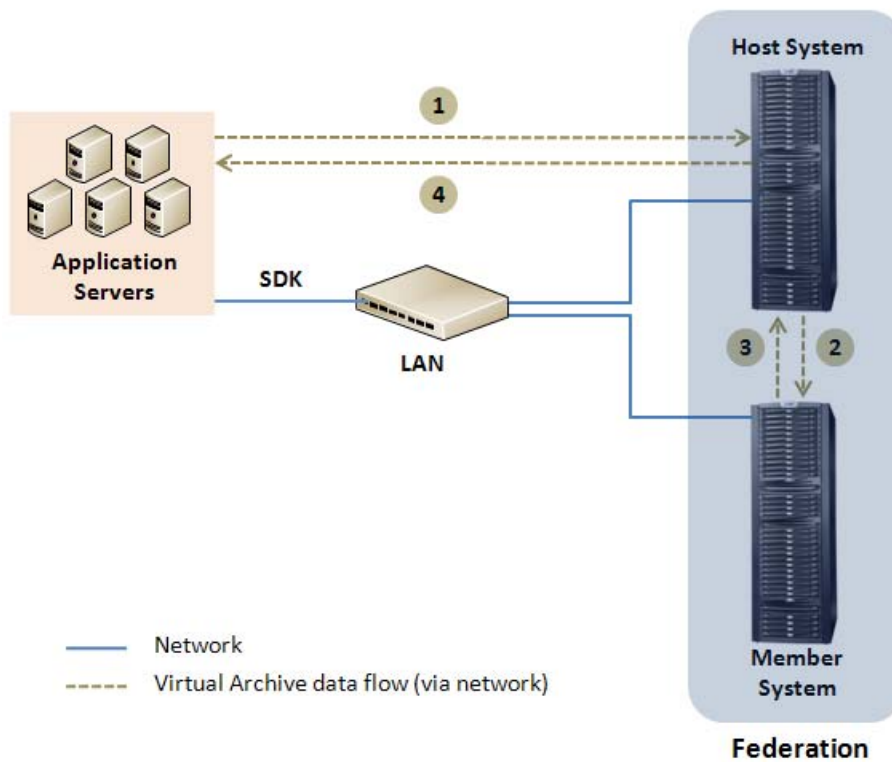
Figure 1. Virtual Archive data flow

Host System

A Host System is the cluster in a Federation on which the Virtual Archive software is configured to run. The routing software gets installed on all nodes in the host systems. The CVA software is active on all nodes that have access and management roles. Applications connect to the Federation through the Host System. Virtual Archive software smartly routes and balances the application data streams across all clusters in the Federation, including itself, based on load distribution, capacity distribution, available capacity, and the nowrite state of the cluster.

With the release of Virtual Archive 1.1, the Host System can be either a new or existing GEN4LP cluster. There is only one Host System supported in a Federation.

Member System

A Member System is any cluster in a Federation that is eligible to store and retrieve content. Member systems must be running CentraStar 4.1 or later.

Campus Federation

A Campus Federation is comprised of data centers that are interconnected and accessible via a LAN. Virtual Archive Host and Member Systems need to be located within certain proximity. In a Campus Federation, latency, packet loss, and bandwidth should all be taken into account.

Virtual Archive Recommended Configurations

This section provides a summary of the recommended best practices that should be considered when deploying Virtual Archive.

Applications

One of the key factors to consider when deploying Virtual Archive is the set of applications to be used in the customer environment. Virtual Archive has been designed to work seamlessly with EMC Centera integrated applications, but some caveats do apply.

SDK and XAM Versions

Virtual Archive supports EMC Centera SDK version 3.1 and later, as well as version 1.0 patch 3 of the EMC Centera SDK for XAM (also known as EMC Centera VIM). This means that all applications to be used with Virtual Archive must be integrated with these or later versions. Note that the vast majority of applications are integrated with EMC Centera SDK version 3.1 or later; however, in some cases, an upgrade to EMC Centera SDK or EMC Centera SDK for XAM (Centera VIM) may be required.

Application Behavior

Although Virtual Archive supports all Centera SDK and XAM functionality, there are a few additional considerations.

- SDK considerations

    EMC Centera Virtual Archive has been designed to take advantage of protocol improvements introduced in EMC Centera SDK 3.2 and Centera SDK for XAM 1.0 patch 3. For this reason, the latest version of Centera SDK and Centera SDK for XAM should be used, when possible.

- Large datasets

    When using EMC Centera SDK 3.1 with Virtual Archive, if the Host System is unstable, for example, access nodes are unstable or bounce, then transactions with datasets greater than 2 GB may not complete successfully. In this scenario BLOBs may end up in different Member Systems and need to be

moved to a single Member System that has the CDF stored. Having a CDF and BLOBs on the same Host or Member System is known as co-location. During co-location of large BLOBs the SDK may time out, thus resulting in a failed transaction. Therefore, in order to use Virtual Archive in such situations it is recommended that the application(s) upgrade to EMC Centera SDK version 3.2 or ensure that the Centera is in a healthy state.

Specific Applications

The following applications have been identified with specific problems areas, if deployed with Virtual Archive.

- Bus-Tech MAS

    The MAS device is integrated with EMC Centera SDK 3.1 and is known to write and extend large datasets. In order to use Virtual Archive with this application it is recommended that the application use EMC Centera SDK version 3.2 or later.

- OpenText LEA

    In order to use Virtual Archive it is recommended that the customer use standard EMC Centera replication for solutions requiring replication and avoid using STORM ISO Device Manager Application replication.

- CommVault QiNetix Suite

    This application is integrated with EMC Centera SDK 3.1, and is known to write very large datasets (up to 2 GB). These transactions may not complete successfully if the Federation is under heavy load. It is recommended that the application use EMC Centera SDK version 3.2 or later.

- EMC Centera Universal Access (CUA)

    In order to use CUA in a Virtual Archive environment it is recommended that CUA version 4.1.1 be used. CUA 4.1 or later will use EMC Centera SDK 3.2, which is the recommended SDK version to be used with Virtual Archive. CUA 4.0 and older versions use a version of SDK earlier than 3.2.

    Note: For any recommendations or clarification regarding application usage with Virtual Archive, please contact *CenteraVirtualArchive@emc.com* .

Sizing

Many factors influence sizing and the scenario could vary for different customer environments. Factors influencing the sizing decision may include, but are not limited to, the following:

- Application load

- Total number of access, storage, and management nodes required on a Host System

- Number of open pool connections by the applications

- Number of applications accessing the EMC Centera system.

These factors are discussed in this section.

Number of Access Nodes

The number of access nodes on a Host System should be planned carefully. The cluster that hosts the Virtual Archive software requires at least the same number of nodes with access roles as found on a Member System(s) in the Federation. If an application is designed to open up many pool connections to the EMC Centera (going beyond the recommended best practice of 25 connections per access node) and has a high volume ingest and retrieval rate, then with the introduction of Virtual Archive, the number of access nodes should be increased on the Host System to match the Member System(s).

Table 2 summarizes the access nodes sizing of a GEN4-LP for the Host System based on the number of access nodes that were on the Member System(s).

GEN4-LP Centera	# Nodes	# Access Nodes			
Member System 1	16	2	4	6	8
Member System 2	16	2	4	6	8
Host System	16	4	8	8^2	8^2

Table 2 Host System – Access node sizing

When installing Virtual Archive in EMC Centera storage environments that have multiple applications and/or high ingest rates, for example, consolidating multiple applications to use a single Federation, it is recommended that a Host System with a large number of access nodes be used. This will allow customers to have an equal or better performance experience after they have installed Virtual Archive.

Number of Management Nodes

Virtual Archive software uses management nodes for the purpose of keeping the Federation in sync. Multiple management nodes should be configured on each Host and Member System for redundancy. A minimum of two management nodes are recommended on all systems in the Federation.

---

[2] A maximum of eight access nodes are allowed on a 16-node EMC Centera.

Cluster Size

In order to choose the Virtual Archive Host System cluster size users should look at the following areas:

1. Application load
2. Number of applications
3. Content protection scheme

- Application load

    An important factor to consider when sizing an environment for Virtual Archive is the application load and number of applications. Ingest and retrieval rates could vary from one application to another. Therefore, understanding application load helps assess whether the EMC Centera cluster is being used most effectively.

    If Virtual Archive is used as a means of adding more capacity, then the number of access nodes may need to be increased to get the same or better performance for the existing applications and application load. The same sizing rules as for a single Centera cluster apply in this case.

- Content protection scheme

    EMC Centera offers two protection schemes, Content Protection Parity (CPP) and Content Protection Mirroring (CPM). CPP stores six data fragments and one parity fragment and CPM stores four objects (two BLOBs, two CDFs) on two unique nodes. A minimum of an eight-node configuration is required for CPP and a four-node configuration for CPM.

    As a standard best practice, applications having small objects use CPM and applications having large object sizes use CPP. However, Virtual Archive will not elect a Member System based on file size. It does not enforce any restrictions on the protection scheme selection and it is not required that the Host and Member Systems have the same protection scheme. Best practice recommendations for having a CPP or CPM cluster should be followed according to application recommendation.

Note: More information about capacity and protection schemes is available in the white paper EMC Centera Capacity Reporting for Gen4LP on CentraStar 4.0 - A Detailed Review available on http://www.emc.com and http://powerlink.emc.com.

Network and Latency

As a general networking concept poor network connectivity can increase latency and packet loss. High bandwidth utilization can also affect the application ingest, retrieval and query rates. This section provides best practices recommendations on how to avoid network issues for Federation in a campus environment.

Bandwidth

The network between the clusters in the Federation must accommodate the application write, read, and query requests to the Federation in a campus environment. This means that the network between the clusters in the Federation must support the aggregate bandwidth required by all applications. The best practice for Federations with high application loads is to support twice the aggregate bandwidth required by all applications.

Network Address Translation (NAT)

EMC Centera Virtual Archive does not support the use of NAT between Host and Member Systems that are federated. If NAT is used, all Member Systems must be on the same logical side of the NAT boundary.

Location of application servers and the Host System

EMC Centera is an IP-based device and uses an IP network for connectivity of its data operations. As of CentraStar 4.2.2 the Virtual Archive is installed with the CentaStar software but must be configured to run via the CentraStar CLI. For more information on how to configure the system please reference the CentraStar 4.2.2 user's manual. Applications will connect only to the Host System as it is responsible for making sure that data is routed between systems in the Federation and application servers. Therefore, general network design practices to keep latency low and avoid packet loss and congestion apply.

Campus Federation

With Virtual Archive it is recommended that the application servers, Host System, and Member Systems be located within a company campus on the same network. If not, performance issues could arise due to latency and packet loss introduced by additional network hops. Figure 2 shows an example of a Campus Federation whereby all the application servers, the Host System, and Member System are located within the same data center and interconnected via a single LAN switch.
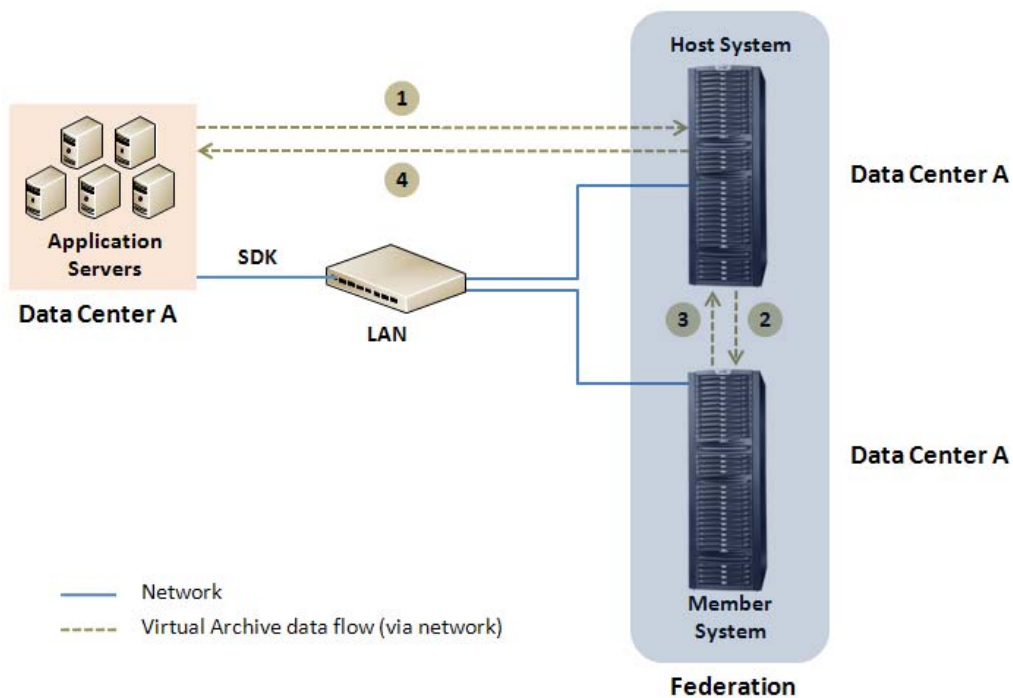
Figure 2. Campus Federation

- Virtual Archive and LAN

    Within a Campus Federation it is possible to have the Host and Member Systems located in different data centers provided there is acceptable LAN connectivity.

    Figure 3 shows the scenario where the application servers and Host System are located in data center A, and the Member System is located in data center B. However, since all components are located in the same location and interconnected via LAN, this configuration meets product specifications.
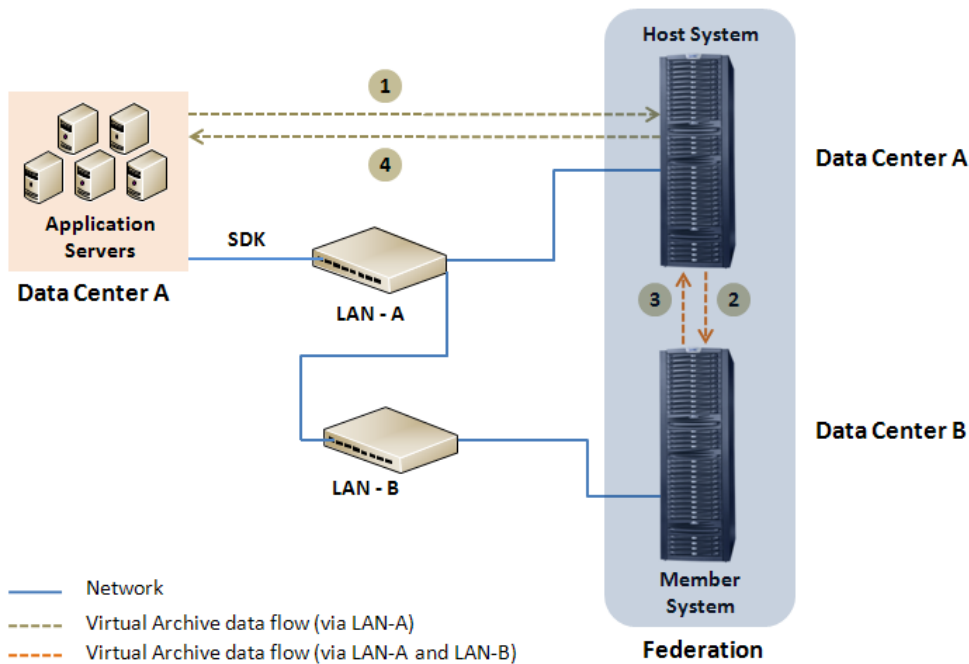
Figure 3. Federation in a multi-LAN environment

Replication and Failover

Virtual Archive allows content to be replicated in uni-directional and bi-directional topologies. Virtual Archive requires a one-to-one mapping of a primary Federation to a replica Federation. Inward star and chain replication topologies are not supported in Virtual Archive environments.

The following scenario explains the requirements for setting up replication in a federated environment with Virtual Archive.

Replication in a Federated Environment

In order to set up replication in a federated environment an EMC Centera service representative needs to make sure that the Federation Target site has an equal number of clusters. The available capacity for clusters on the Federation Target can be equal to or greater than the clusters on the Federation Source site. There needs to be a one-to-one mapping between the Federation Source and Federation Target site clusters. Figure 4 gives an example of how replication would look in a federated environment.
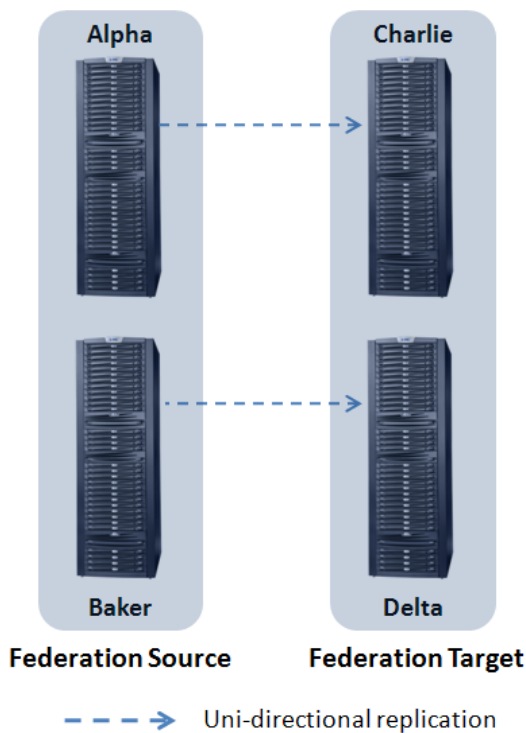
Figure 4. Virtual Archive setup in a uni-directional replicated environment

In Figure 4, replication needs to be set up between Alpha and Charlie (Host Systems at the Federation Source and Federation Target sites, respectively) and between Baker and Delta (Member Systems on the Federation Source and Federation Target sites, respectively).

However, when setting up replication between Alpha and Charlie, an EMC service representative needs to make sure that the correct data port is used. Virtual Archive software is installed and uses the data port 3218; therefore, port 13218 should be used when supplying the Federation Target address for replication shown in

Figure 5:

```
Config# set cluster replication
Replication Enabled? (yes, no) [no]: yes
Replication Address: 10.241.44.19:13218,10.241.44.20:13218
Failover Address [10.241.44.19:13218,10.241.44.20:13218]:
10.241.44.19:13218,10.241.44.20:13218
Replicate Delete? (yes, no) [no]:
Profile Name: replication
Location of .pea file [prompt]: C:\replication.pea
Config#
```

Figure 5. Replication setup on a Host System using CLI

Failover in a Federated Environment

Failover in a federated environment is different from the traditional EMC Centera SDK failover. The failover IP address is set manually on the Host System by using the CLI command `federation set failover`.

Figure 6 shows configuration of a failover address on a Host System.

```
Config# federation set failover
Federation failover enabled (yes, no) [no]:yes
Failover address [not set]: 10.241.44.19:13218,10.241.44.20:13218
Issue the command?
 (yes, no) [no]:yes
Config#
```

Figure 6. Virtual Archive failover setup using CLI

If the Federation Source is inaccessible for any reason, the EMC Centera SDK will fail over to the Federation Target. Failing over to the Federation Target only takes care of reads (writes do not fail over). This is shown in **Error! Reference source not found.**
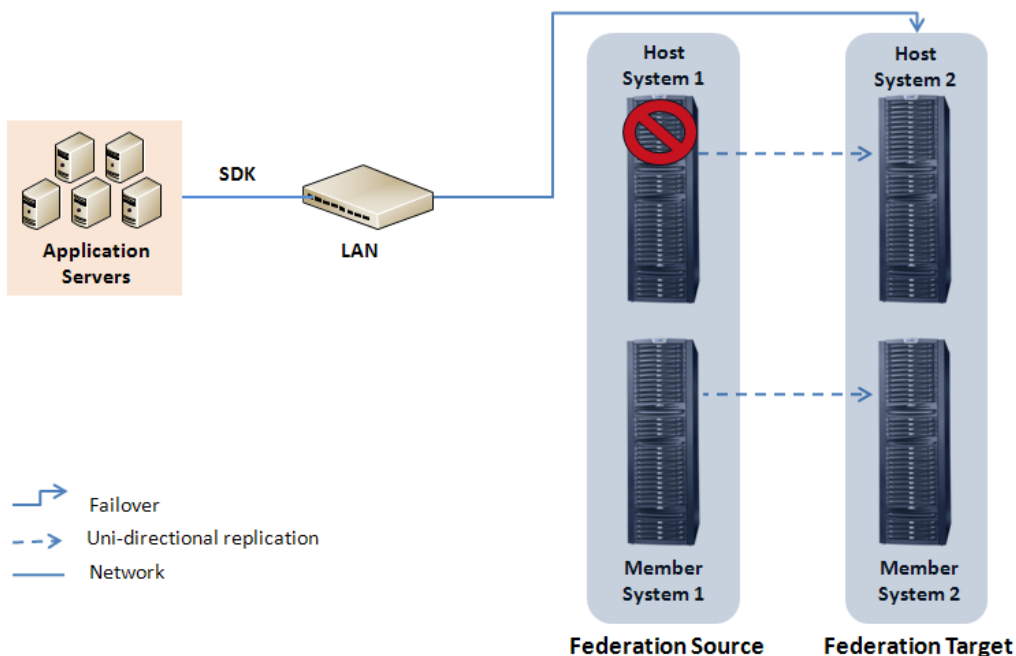

Figure 7. Federation failover scenario

Use Case Considerations

Supported Use Cases

This section provides information on the typical use cases that are supported by Virtual Archive.

Note: To discuss use cases not listed in this section please contact *CenteraVirtualArchive@emc.com*.

Capacity Expansion

This section will discuss the use cases where Virtual Archive can be used as a means of expanding capacity for EMC Centera systems that are at four racks.

- Adding Capacity to an Existing Cluster

    Customers often face the need to increase the storage capacity of their existing EMC Centera clusters, which they do by adding more EMC Centera storage to their existing environment. In most cases they expand capacity by adding additional nodes to their existing clusters, but if there is a need to expand to multiple cabinets they now have the option to use Virtual Archive.

- Steps for adding Virtual Archive in an existing environment

    Capacity can be added to any cluster in the Federation that is at four racks. When using Virtual Archive as a means to federate existing clusters and add more capacity it is recommended that the new cluster runs the Virtual Archive software. Once the Federation is established, additional new clusters or nodes can be added. The steps to add capacity by using Virtual Archive are outlined next:

    1. Configure the new cluster to match the configuration settings of the existing cluster environment. This means that all clusters must run the same compliance mode. If the Host System is running in basic compliance mode and the Member System is running GE than the Host System must have the GE license.

        The Cluster Configuration Sync Tool can be used to determine differences between the Host and Member Systems.

    2. Configure the Virtual Archive host system on the new or existing cluster. Refer to the release notes for the CentraStar version supported by Virtual Archive and the CentraStar user's manual on how to configure the CVA and router software.

Note: Clusters in a Federation do not have to be at the same CentraStar code level.

    3. Create a Federation on the Host System.

    4. Add a new or existing Member System to the Federation. You must use the Cluster Configuration Sync Tool to merge pools and profiles to the Federation.

    5. Update the application connection string so that it now uses the Host System IP.

Figures 8 through 10 describe the process of adding capacity to a new or existing Host System and then adding the existing EMC Centera cluster as a Member System.
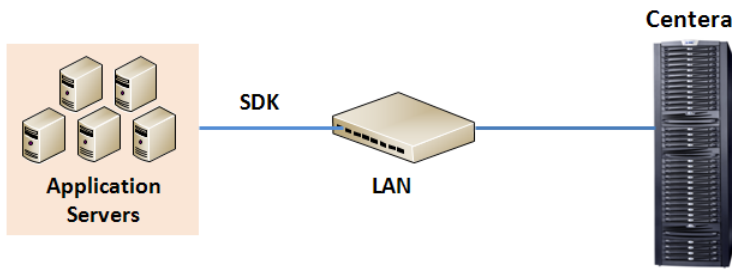

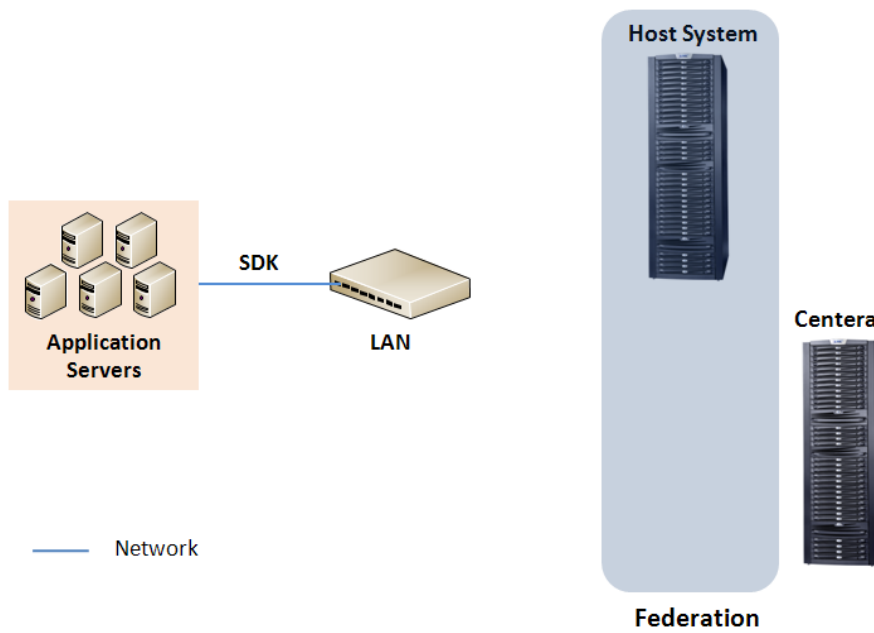Figure 8. Customer environment before adding Virtual Archive

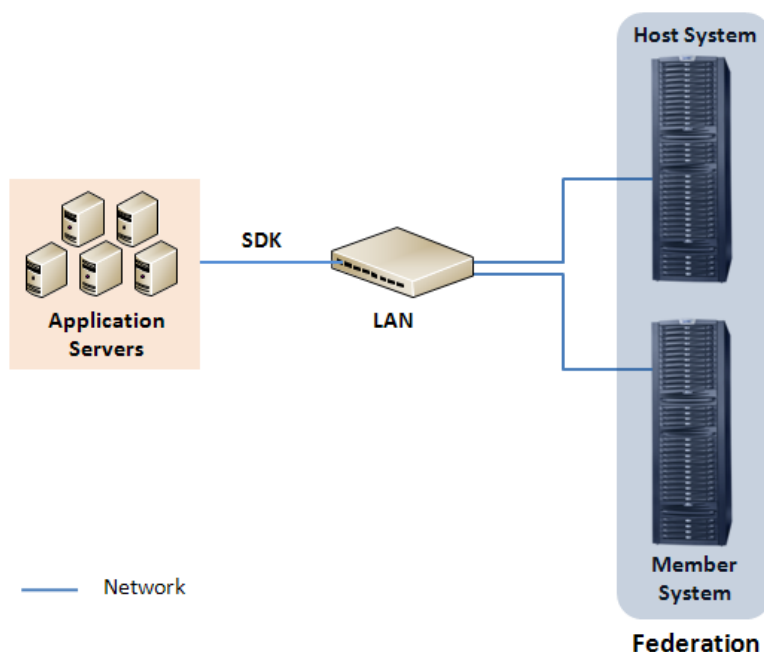Figure 9. Adding a Host System and creating a Federation

Figure 10. Existing EMC Centera cluster added to a Federation as a Member System

- Adding capacity by using Virtual Archive in an existing environment (replicated setup)

  In the scenario where a customer is adding capacity using Virtual Archive and has an existing replicated setup either in uni- or bi-directional replication, a Federation needs to be created at the primary and replica sites.

- Steps for adding Virtual Archive in a replicated environment

  The following steps need to be followed in order to add capacity to the primary and replica sites by using Virtual Archive:

  1. Configure the new or existing cluster at the primary site to match the configuration settings of the Member Systems, using the Cluster Configuration Sync Tool.

  2. Install Virtual Archive software on the new or existing cluster to be the Host System at the primary site or the Federation Source.

  3. Create a Federation on the Host System and add new or existing clusters, up to three, to the Federation.

     You must use the Cluster Configuration Sync Tool to add Member Systems to the Federation.

  4. Perform steps 1 through 3 for the target site or Federation Target.

5. Establish replication between Host Systems at the Federation Source and Federation Target using the CLI command `set cluster replication`. Use port 13218 when setting up replication between the Host System at the Federation Source and Federation Target sites and provide a failover address. This is shown in

6. Figure 11.

```
Config# set cluster replication
Replication Enabled? (yes, no) [no]: yes
Replication Address: 10.241.44.19:13218,10.241.44.20:13218
Failover Address [10.241.44.19:13218,10.241.44.20:13218]:
10.241.44.19:13218,10.241.44.20:13218
Replicate Delete? (yes, no) [no]:
Profile Name: replication
Location of .pea file [prompt]: C:\replication.pea
Config#
```

Figure 11. Replication setup between Host Systems at Federation Source and Target

7. Repeat step 5 for all the Member Systems.

8. Update the application connect string so that it now uses the Host System IP at the Federation Source.

- Federating two existing or new clusters

    Virtual Archive supports federating two to four existing clusters. Using the Cluster Configuration Sync Tool you can synchronize the existing clusters to match cluster configuration settings. These settings include:

    - Pool and profile settings

    - Compliance

    - Retention

   Steps to federate existing clusters include the following:

    1. Install Virtual Archive on one of the existing clusters that will become the Host System.

    2. Create a Federation on the Host System.

    3. Add the second existing cluster as a Member System.

- Clip Invariant support

    Clip Invariant is the process by which Virtual Archive ensures that it keeps only one instance of a clip found on more than one Member System in a Federation. A clip can exist on more than one cluster due to any of the following conditions:

    - The existing clusters in the Federation at one time were a replication pair.

- Restore operation was performed on any of the clusters in the Federation and that resulted in having the same clip existing on more than one cluster.

- Data was manually copied by EMC Service or a customer to a different cluster and that resulted in having the same clip exist on more than one cluster.

In the scenario where a clip resides on more than one Host or Member System in the Federation, upon receiving any data read or update request for the particular clip, Virtual Archive reads the Write Time Stamp (WTS) of all occurrences of the clip. It will then update the clip having the lowest delta between the cluster time and the WTS. Note that all other occurrences of the clip are marked invalid by setting their WTS to a sentinel value. The sentinel value is set at the file system level and clips with WTS marked to a sentinel value will not be read on any future requests. Figure 12 shows the process flow for Clip Invariant scenario.
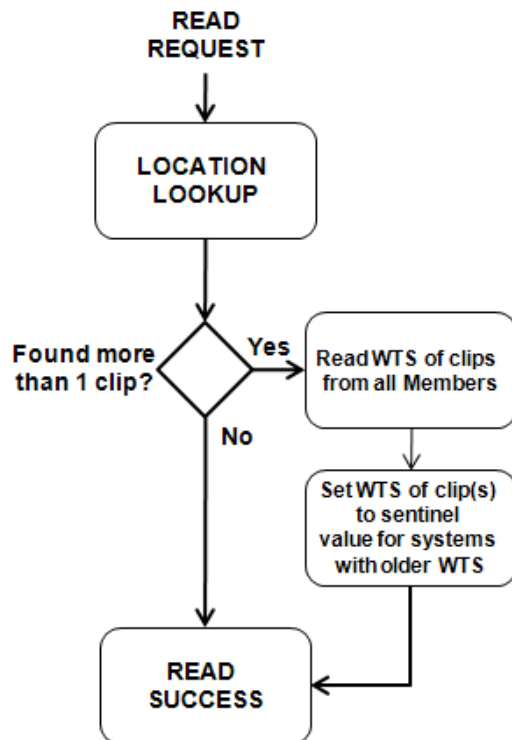
WTS = Write Time Stamp

Figure 12. Clip Invariant process flow

Use Cases Not Supported for a Virtual Archive

The following use cases are not supported.

Virtual Archive in a WAN setup

Virtual Archive software is not suited for any clusters that are geographically dispersed; that are in different cities, states, countries; and that require a WAN-to-WAN connection. Going over the WAN can add huge network latency and packet loss and could potentially result in a degraded performance. It is therefore not a supported use case to have Virtual Archive in an environment where the Host System, Member Systems, and application servers are dispersed across different cities, states, or countries connected by a WAN.

The next figure shows an example of a configuration that is not supported and should be avoided.
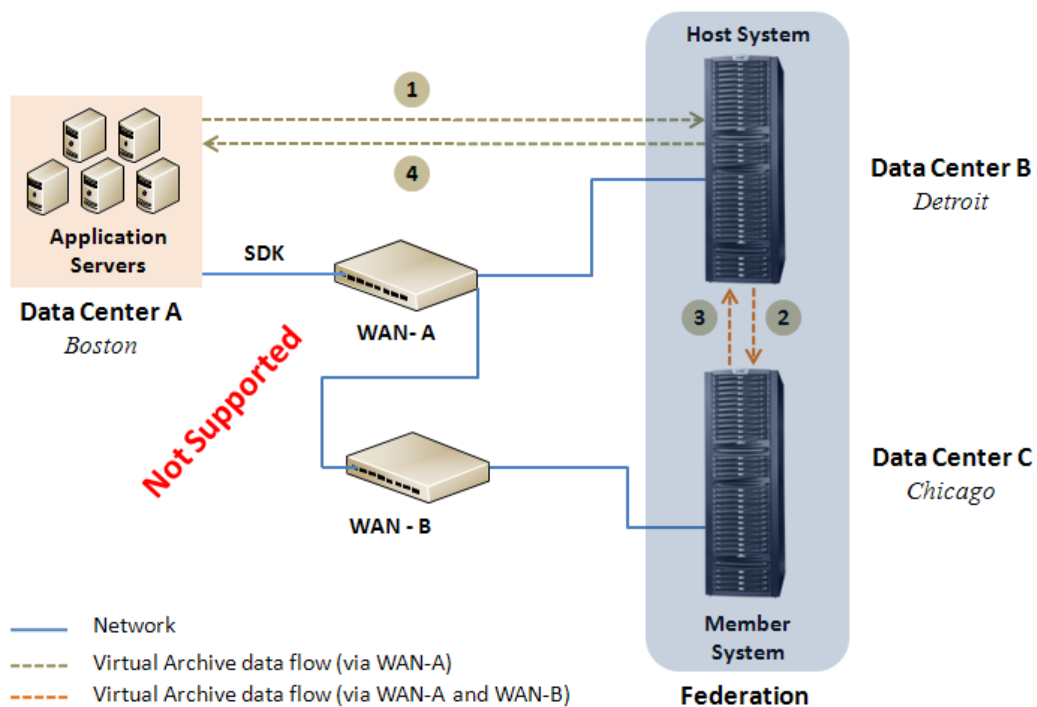

Figure 13. Virtual Archive connectivity via WAN is not supported

Direct connectivity to a Member System in a federated environment

In a federated environment, the application servers should always connect to the Host System. The Host System is the only cluster that has knowledge about the Federation and therefore is responsible for the routing of data transactions. If application servers connect directly to the Member Systems for data operations the application will not be able to read data that was routed by the Host System to Member Systems.

Conclusion

This configuration guide is intended to provide recommended configurations for EMC Centera Virtual Archive. It is expected that after reading this document the reader will have a better

understanding of where this technology can be deployed in customers' existing environments and the various factors that need to be considered.