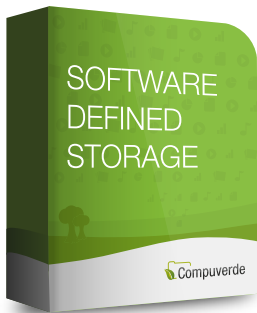


Best practice: Metro Cluster

Metro Cluster is an enterprise-grade, absolute high-availability solution that is designed to tolerate the loss of an entire datacenter, perfect for mission-critical applications that demand zero downtime.



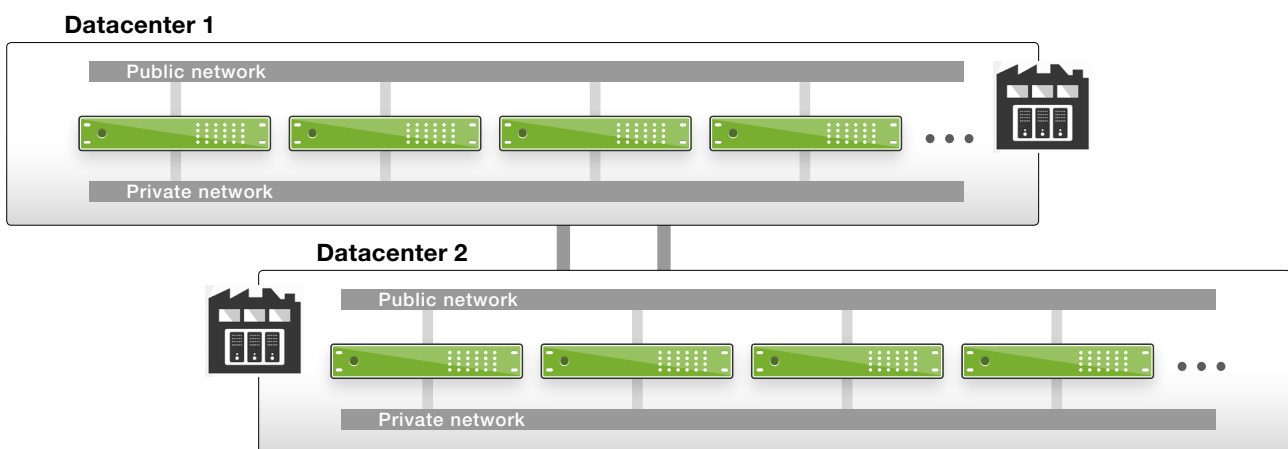
The Compuverde Metro Cluster is a simple, multi-site configuration that offers complete redundancy to support mission-critical data and applications. Data is synchronously mirrored and distributed to two independent physical locations. The Metro configuration can be utilized with both Compuverde vNAS and Hyper-converged setups.

- ✓ Active-active nodes
- ✓ No central orchestration
- ✓ No single point of failure
- ✓ Sub 5 ms between locations

Any Workload

Any Application

Any Protocol



Private network

For synchronization and communication between the nodes. Use high-speed connections for low latency and increased performance. Optionally, there can be separate networks for management and antivirus.

Connection between the sub-clusters should be low latency < 5 ms and with minimum the same bandwidth as the private network.

All the nodes on both the primary and secondary datacenter are available to the clients through the public network.

Management Tool

Designed to help perform administrative tasks, monitor the system, change configuration features, access logs and roll out new firmware updates. Connected to the private network, it is not part of the storage cluster or node-to-node communication.

One single file system spans both datacenters and is available to applications and users inside and to the outside through Ethernet. The storage keeps itself internally synchronized, horizontally throughout the cluster and vertically down to the physical storage so that all nodes will have the same consistent view of the unified file system for all reads, writes and updates.

The two datacenters are classified as primary and secondary. Even though applications and users may experience higher performance when connected to the primary datacenter compared to the secondary, they can connect to any node on any of the two datacenters. Other than this, all nodes share the same tasks and responsibility. There is no complicated

central orchestration, special purpose hardware or potential bottlenecks.

Data is mirrored to the two locations and erasure coding is applied to further ensure that each sub-cluster is protected against failure to nodes.

Should you lose an entire location, data is secured and service will be provided through the other location. If the primary location is lost, the remaining cluster will set itself into a safe state to prevent split-brain conditions. The storage administrator can then decide to demote the failing and promote the other datacenter to resume normal functionality.

2017-11-22

Setup

1. Install the software on each node
2. Create the storage cluster
3. Activate Metro

Everything after step 1 is done through setup wizards and GUI inside the Compuverde management tool.

Scaling

Scaling has to be simple to accommodate any future demands. In a Compuverde Metro Cluster, scaling is simply done by installing new nodes and adding them to your primary and secondary location. You will see a linear increase both in capacity, performance and IOPS as new nodes are added. With Auto Rebalancing, data is automatically moved to the new nodes to keep the cluster in balance. Please ensure that the storage size on both the primary and secondary location stays equal.

Network

The cluster is interconnected through a private, high speed, low latency local area network. This network deals with synchronization between the nodes while the public network delivers access for clients through any of the major protocols: SMB, NFS, iSCSI, Amazon S3 and OpenStack Swift.

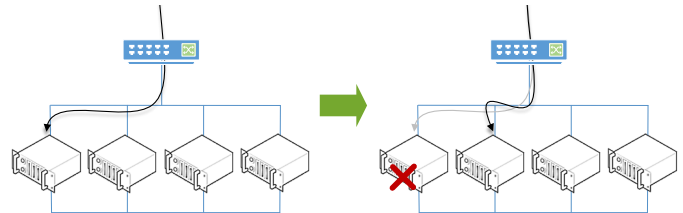
Make sure that IGMP is properly configured for all switches that are affected by private network traffic. All the switches should have IGMP Snooping enabled and only one switch in the group has both IGMP Querier and IGMP Snooping enabled.

Erasure Coding

In addition to being mirrored to two locations, data is erasure coded and distributed within the sub-cluster for additional protection. This ensures that each datacenter can tolerate the loss of one or more disks or nodes (depending on the chosen file encoding). The self-healing process on each non-affected node automatically kicks in to replicate the missing data simultaneously.

Virtual IP

Compuverde's Virtual IP makes your storage cluster highly available. The feature provides a failover mechanism for nodes, ensuring that all nodes appear online at all times even in the case that one goes down. You can read more about this and other Compuverde features in the Technical Overview or Management Manual.



Requirements

- A minimum of eight nodes are recommended, four for each location
- Start with equal capacity on both locations
- Low latency (<5 ms) on the connection between the two locations
- Metro feature is enabled in your license

Recommended specifications each node

CPU	x64 (4+ cores)
RAM	32+ GB
Boot disk	60+ GB
Cache disk	100+ GB SSD / NVMe
Storage disk	SAS / SATA disks
Network	10 Gigabit or better

The Compuverde management tool requires a Windows client (64-bit Windows 7 or later, 100 MB for installing the Management software) on the same subnet as the storage cluster's private interface.

