Choosing the Right Storage for Your Cloud

Version 1.0 – April 2021

Scope

OpenNebula provides a variety of ways for Virtual Machines and containers to access storage. It supports multiple traditional storage models including NAS, SAN, NFS, iSCSI, and Fiber Channel (FC), which allow virtualized applications to access storage resources in the same way as they would on a regular physical machine. It also supports distributed Software-Defined Storage (SDS) models like Ceph, GlusterFS, StorPool, and LinStor, that allow you to create and scale elastic pools of storage and hyperconvergence deployments. Deciding which is the right storage backend for your cloud depends on your performance, scalability, and availability requirements; your existing storage infrastructure; your budget for new hardware, licenses, and support; and your skills and the IT staff you want to dedicate to its operation. This report describes OneStor, a local direct attached storage solution enhanced with caching, replica and snapshotting mechanisms that has been specially designed for OpenNebula cloud infrastructures. OneStor brings significant benefits to any enterprise, with a clear reduction in complexity, resource consumption and operational costs.

What Is OpenNebula?

OpenNebula is a **powerful, but easy to use, open source solution to build and manage Enterprise Clouds** that combines existing virtualization technologies with advanced features for multi-tenancy, automatic provision, and elasticity to offer on-demand virtualized services and applications. It provides a single, feature-rich and flexible platform with **unified management of IT infrastructure and applications** that **avoids vendor lock-in** and **reduces complexity, resource consumption, and operational costs**.

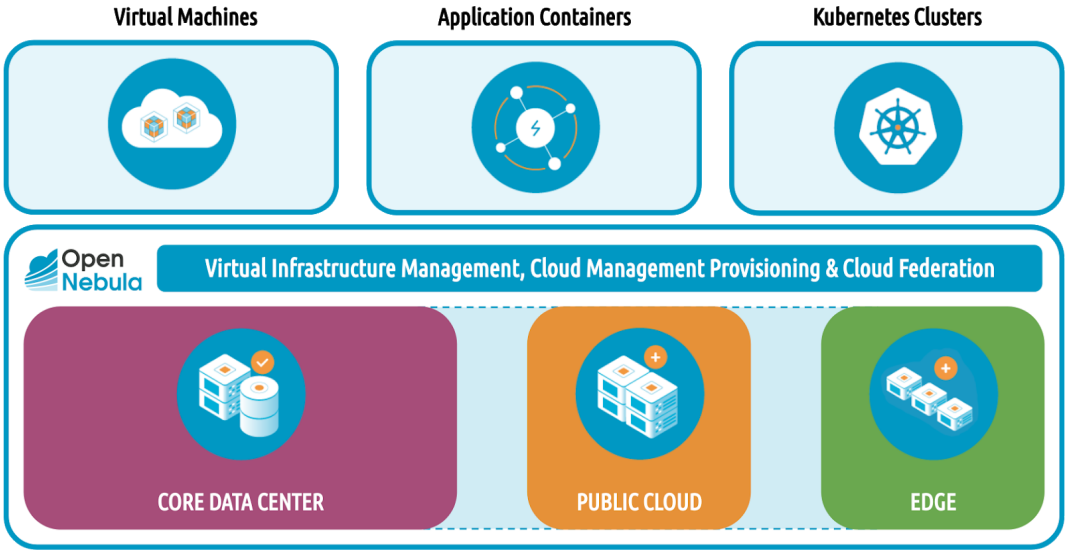

Figure 1. Available solutions for container orchestration with OpenNebula.

OpenNebula manages:

- **Any Application**: Combine containers with Virtual Machine workloads in a common shared environment to offer the best of both worlds: mature virtualization technology and orchestration of application containers.
- **Any Infrastructure**: Unlock the power of a true hybrid and multi-cloud platform by combining edge, public, hosted, and private cloud operations.
- **Any Virtualization**: Integrate multiple types of virtualization technologies to meet your workload needs, from a fully virtualized environment to system containers and serverless deployments.

A standard OpenNebula Cloud Architecture consists of the **Cloud Management Cluster** with the Front-end node(s), and the **Cloud Infrastructure**, made of one or several workload Clusters with the hypervisor nodes and the storage system, which can be located at multiple geographical locations, all interconnected with multiple networks for internal storage and node management, and for private and public guest (VM or container) communication.

The OpenNebula Cloud Infrastructure can combine multiple clusters with different configurations and technologies to better meet your needs. In general, there are two types of OpenNebula Clusters:

- **Customized Clusters** (based on either VMware[1] or open source[2] technology), which are typically deployed on-premises to meet specific requirements.
- **Edge Clusters**[3] (based on a combination of open source technologies), which can be deployed on demand both on-premises and on public cloud and edge providers, with a high degree of integration and automation.

One of the key components for performance, scalability, and reliability in your cloud is the storage solution. In VMware cloud deployments, OpenNebula interacts as a consumer of vCenter storage, and as such, it supports all storage backends that can be mounted by ESX, including Fiber Channel SAN, local disks, and NFS shares. In open source cloud environments, OpenNebula supports all the existing major storage solutions and choosing the right storage to meet various performance, scalability, availability and cost requirements can be a challenging and time consuming task. This white paper introduces the storage solutions supported by the OpenNebula software and describes OneStor, the recommended storage solution with a very simple design that avoids vendor lock-in and reduces complexity, resource consumption, and operational costs.

Cloud Storage Solutions

OpenNebula, through its Datastore model, supports the following four main types of storage solutions:

Local Direct Attached Storage

This storage configuration uses a centralized storage area for the VM disk image repository and the local DAS (Direct Attached Storage) area of each host to run VMs. Local Storage solutions are the most cost effective and efficient approach but do not provide any inherent recovery mechanism when a node fails. **OneStor** is a local storage solution enhanced with **caching, replica, backup, and snapshotting mechanisms** that brings recovery and migration enterprise features while maintaining its low-maintenance and low-cost benefits.

[1] https://support.opennebula.pro/hc/en-us/articles/206652953-VMware-Cloud-Reference-Architecture-White-Paper
[2] https://support.opennebula.pro/hc/en-us/articles/204210319-Open-Cloud-Reference-Architecture-White-Paper
[3] https://support.opennebula.pro/hc/en-us/articles/360050302811-Edge-Cloud-Architecture-White-Paper

Network Attached Storage

This storage configuration assumes that your hosts can access and mount shared file volumes located on a NAS (Network Attached Storage) server, typically an **NAS device** or an **NFS server** in small deployments. OpenNebula integration with **GlusterFS** or **Lustre** is performed through the NAS configuration. These shared volumes both store the VM disk images and provide storage area to run VMs.

Storage Area Network

This storage configuration assumes that hosts have access to block-level storage devices (LUNs) exported by an Storage Area Network (SAN) server using a suitable protocol like **iSCSI** or **Fiber Channel**. The Hosts interface the devices through the **LVM** abstraction layer. Virtual Machines run from an LV (logical volume) device instead of plain files. This reduces the overhead of having a filesystem in place and thus it usually increases I/O performance. SAN solutions are more expensive and complex to set up and manage, but provide better performance. They offer the speed of DAS with the sharing, flexibility, and reliability of NAS.

Distributed Systems

This storage configuration assumes a Software-Defined Storage (SDS) system that is distributed on multiple file servers or multiple locations. While they are designed to offer high scalability and availability, the deployment and operation of these storage platforms require experience, greater financial investment, dedicated hardware (sometimes following a hyperconverged model), and a significant amount of human resources. OpenNebula integrates with distributed systems like **Ceph**, **LinStor,** and **StorPool**. Some of these solutions were designed following a hyperconverged approach and others, like Ceph, were designed to build dedicated storage clusters and their deployment in hyperconverged architectures usually results in very poor performance and a balance between storage and compute needs that is difficult to achieve.

Other

OpenNebula also brings a **Raw Device Mapping Image** Datastore that enables raw access to block devices on Nodes for fast VM deployments due to a non-existent transfer operation from the Image Datastore to the System Datastore, and **iSCSI - Libvirt** Datastore that is used to register the already existing iSCSI volume available to the hypervisor Nodes.

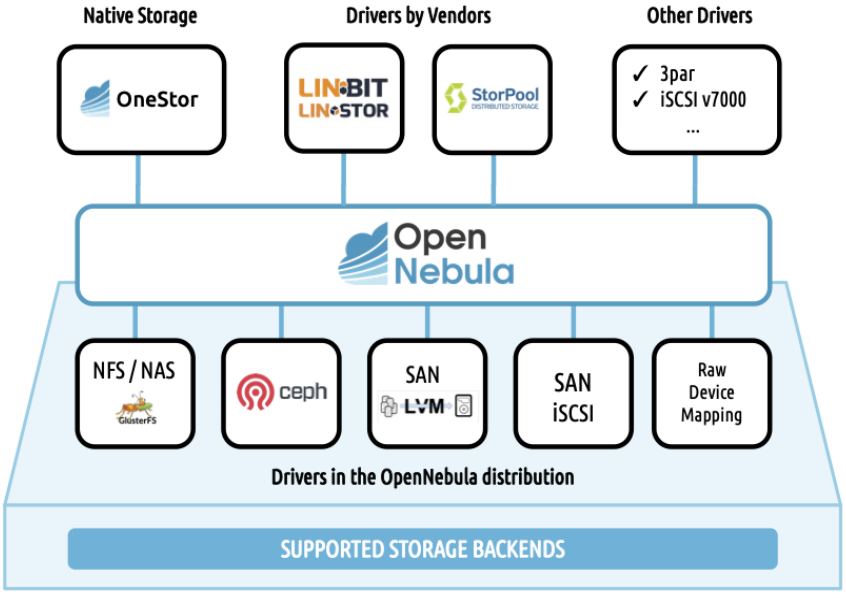

Figure 2. Storage solutions supported by OpenNebula.

OneStor Performance and Availability

OneStor is a storage solution fully supported by OpenNebula Systems that has been developed for the efficient management of disk images in OpenNebula cloud environments. OneStor is one of the multiple storage solutions[4] that can be used to build Customized Clusters and the one that we have selected as the native storage solution for OpenNebula Edge Clusters.[5]

OneStor has been designed to meet the following functional requirements:

- **Access to external (public) and internal (private) marketplaces** that act as global image repositories. Examples include OpenNebula Marketplace or Docker Hub, but also private HTTP repositories or container registries.
- **Minimize image transferring with image caching** when Clusters consist of a large number of nodes or when they have been deployed remotely and are connected over public internet links.
- Install **Replica Servers within each Cluster to increase availability.**
- **Maximize application I/O performance** to provide native physical storage IOPS and latency performance.
- **Simple deployment to reduce the complexity and technology footprint** of the solution.
- Allow **live migration** across hosts within a Cluster to simplify operation and maintenance.
- Provide **persistent storage** for the stateful applications, schedule **recurring snapshots** of a volume, and **schedule recurring backups** to NFS, HTTP or S3-compatible secondary storage without disrupting running workload.

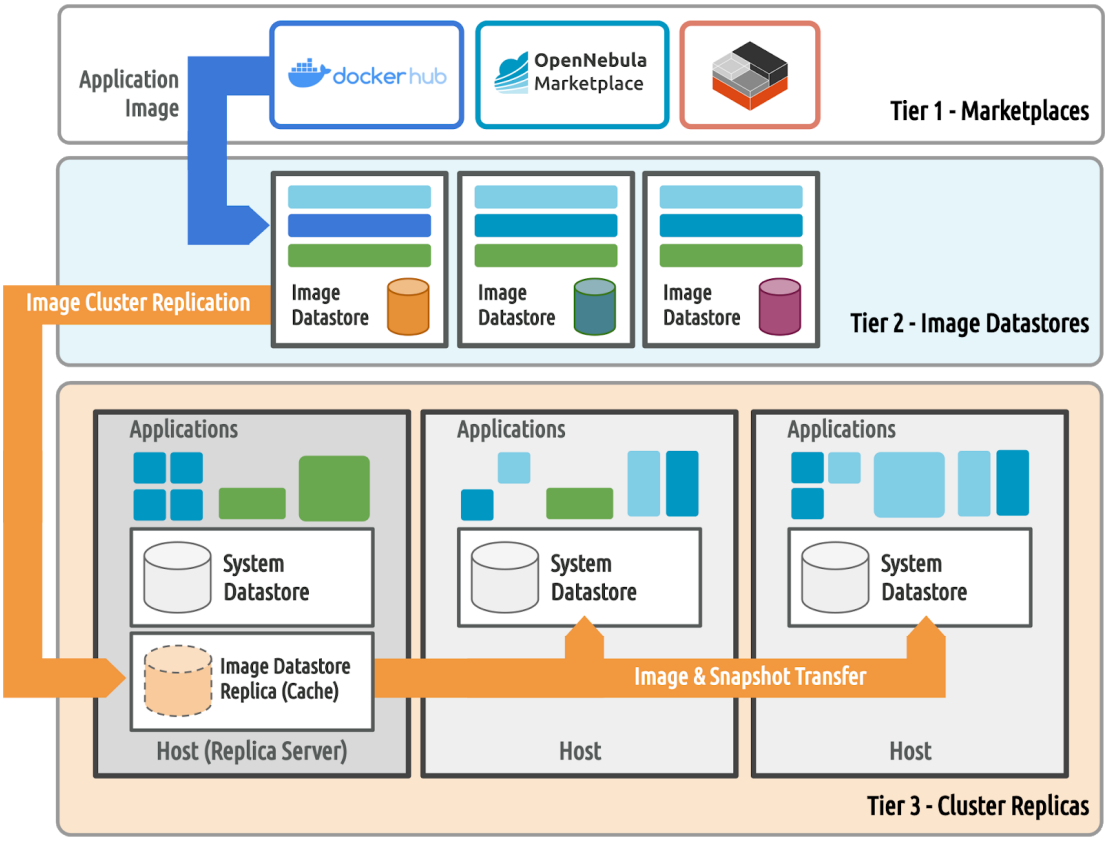

Figure 3. The 3-tier architecture of the OneStor storage system

[4] https://docs.opennebula.io/stable/open_cluster_deployment/storage_setup/index.html
[5] https://docs.opennebula.io/stable/management_and_operations/edge_cluster_management/index.html

Application images are based on files with **qcow2 format** to reduce file transfer and instantiation times, ease backup solutions, and implement advanced features like snapshotting in an efficient way. OneStor combines a **3-tier global architecture** for application image distribution with an enhanced **SSH transfer mode** with **replica caching** and **snapshotting** that greatly improves its **scalability**, **performance,** and **reliability**. **Images can be persistent** and their changes copied back to the Image Datastore after VM shutdown.

3-Tier Architecture for Image Distribution

OneStor implements a 3-tier global architecture for image distribution:

- **Tier 1 - Global Marketplaces**: This tier consists of the remote servers and storage implementing the global application image repositories.
- **Tier 2 - Image Datastores**: This tier consists of the zone Image Datastores in the Front-end and provides the primary image storage location for the OpenNebula instance.
- **Tier 3 - Cluster Replicas**: Application images are cached within a Cluster in dedicated replica hosts to minimize image transferring.

Application Image Caching in Replica Servers

The replica mode caches the images in Replica Servers within each Cluster so they are available close to the hypervisors to reduce the bandwidth requirements to the tier-2 image datastore servers and considerably reduce deployment times. This is especially important in highly distributed deployments where copying images from the tier-2 Front-end to the tier-3 Cluster hypervisors could be very slow.

Snapshots and Fault Tolerance

Availability is a critical aspect of cloud architecture design, mostly when it comes to application data recovery, and integrity. OneStor provides snapshot built-in capabilities to achieve high availability and guarantee that your data and services are always available. Moreover, maintenance work can be performed while the system is operating and host failures can be mitigated thanks to the support for application **live-migration within clusters**.

The 3-Tier Replica Storage Architecture implements an availability system based on **periodic snapshots** that are used to recover from VM and host failures by fencing the node to prevent split-brain conditions (soft, because of local I/O) and automatically restarting the application in another node. Application snapshots are kept within the Cluster (tier-3) to enable fast recovery from the last application checkpoint.

The snapshot operation does not impact the I/O of neighbouring applications because it is based on the QEMU Redirect-on-Write feature, nor the available network bandwidth because it uses a delta-transfer algorithm to reduce the information transferred to the Cluster replica server. Another important aspect to consider is the VM recovery time. Compared to recovering a VM without any snapshot the time is similar, as the base images are already located on the Cluster replica (tier-3) and the only additional overhead is the transfer of the disk snapshot, already available in the Cluster as well.

Availability can be improved within each node by **HW replication** (HD Raid and NIC) and network paths, and within the application by implementing **application-level HA**, when data and application state integrity is required, across multiple clusters.

Periodic Backups

OneStor supports **periodic image backups** to a private Marketplace based on NFS, HTTP or S3-compatible secondary storage.

Application Image Deployment Performance

Disk images are transferred between the Image and System Datastores in each host's local storage to maximize I/O performance. The expected deployment times depend on the interconnection links between tiers[6] and disk sizes, achieving times of a few seconds for cached images.

Application I/O Performance

The applications run from the direct attached storage of the hosts to maximize the available I/O performance delivered to the applications. The I/O performance is close to that of the native host and it is only impacted by the virtualization layer.

Hardware Requirements

OneStor has been implemented with lightweight technology components that already exist in the Linux operating system to accommodate any deployment model both on physical and virtual resources, increase the reliability of the storage backend, and translate into modest hardware requirements, like SATA SSDs and 10 gigabit networks. Moreover, its deployment follows a hyperconverged approach that does not require dedicated servers to implement a distributed storage system. This reduces the complexity of the solution, enabling the use of the local storage area of the Cluster's hosts.

Cloud Storage Comparison

Choosing the best storage option for your business really comes down to deciding which criteria are most important. Once you know which factors are essential to your cloud operations, you can find the storage option to fit your particular needs.

- **Performance**: What is the expected performance of running applications and the potential impact of the underlying storage solution?
- **Scalability**: How does the storage solution scale with the cloud infrastructure?
- **Availability**: What would happen if you lost files? Can you recover workload running when a node fails? From what state? What is the time for recovery? What would downtime do to your business?
- **Hardware Needs**: How much do you have to spend?
- **Maintenance**: Do you have dedicated IT staff to manage your system?
- **Enterprise Support**: Can I get 24/7 support for the storage? At what cost?

	OneStor	NAS/SAN	Distributed Systems
Application I/O Performance	Direct attached storage with optimal performance	Dedicated, high-performance storage system	CPU, network, and disk contention that may diminish performance
Scalability	Medium-sized clusters and any number of clusters	Medium-sized clusters and any number of clusters	Could support very large-scale clusters
Availability	Periodic snapshots in replica hosts	Provided by the data storage device	Replication across hosts
Hardware Needs (Capx)	No need for additional hardware	Data storage device	Additional dedicated hardware and network consumption

[6] https://support.opennebula.pro/hc/en-us/articles/360050302811-Edge-Cloud-Architecture-White-Paper

Maintenance (OpEx)	Minimal	Medium	Complex
Enterprise Support (OpEx)	OpenNebula E2E support (included in OpenNebula subscription)	By third-party vendor (additional cost)	By third-party vendor (additional cost)

Our recommendation is to adopt OneStor unless you are planning to build very large-scale clusters or you are supporting complex, mission-critical applications.

Summary

In this report, we have described OneStor storage configuration and illustrated a comparison between OneStor and the rest of the storage solutions supported by OpenNebula. Choosing the right cloud storage platform is crucial. A wrong decision may result in severe delays, long-term vendor lock-in situations, and an increased TCO. We have demonstrated that OneStor can bring **significant benefits** for any enterprise with a clear **reduction in complexity, resource consumption, and operational costs**. You can enjoy a single vendor experience since OpenNebula Systems offers Enterprise support for the complete software stack through its OpenNebula **Software Subscription** and offers managed cloud services through a new OpenNebula **Managed Subscription** so your team can forget about infrastructure and focus on business workloads.

Ready for a Test Drive?

You can evaluate OpenNebula and build a cloud in just a few minutes by using **miniONE**,[7] our deployment tool for quickly installing an OpenNebula Front-end inside a Virtual Machine or a physical host, which you can then use to easily add remote Edge Clusters based on KVM, LXC or Firecracker.

[7] https://minione.opennebula.io

LET US HELP YOU DESIGN, BUILD, AND OPERATE YOUR CLOUD

CONSULTING & ENGINEERING

Our experts will help you design, integrate, build, and operate an OpenNebula cloud infrastructure

OPENNEBULA SUBSCRIPTION

Get access to our Enterprise Edition and to our support and exclusive services for Corporate Users

MANAGED SERVICES

Our team of experts can fully manage and administer your OpenNebula cloud for you

Sign up for updates at OpenNebula.io/getupdated

Rev1.0_20210426