



Deep Dive on SimpliVity's OmniStack

A Technical Whitepaper

By Hans De Leenheer and Stephen Foskett

August 2013

Introduction

This paper is an in-depth look at OmniStack, the technology that powers SimpliVity's OmniCube, a 2U data center building block that delivers a broad set of infrastructure functions for the virtualized environment. OmniCube is leading the emerging market for “hyper-converged” infrastructure.

Appliance Sprawl and SimpliVity's Response

Historically, IT environments have used a single server per workload or application. As a result of excessive unused resources in those servers – and with the help of virtualization – datacenters have been simplified and workloads have been consolidated to just the resources necessary.

Shared storage, typically in the form of an enterprise storage array, is a requirement for datacenter virtualization. Other new challenges also brought dedicated hardware resources back into the environment – for example, backup deduplication appliances, WAN optimizers, cloud gateways and flash cache storage appliances. All of these devices add complexity, expense, and management difficulty to the virtual environment.

SimpliVity brings all these functions together in a single 2U offering. OmniCube brings comparable performance, protection, and functionality at 3x lower acquisition cost, power and space savings, and lower overall operational cost. Plus, as an integrated system, OmniCube is easier to manage.

Converged Infrastructure

A few years ago, enterprise IT was buzzing with talk of “infrastructure stacks” and “converged infrastructure”. These solutions sought to bring together compute, storage and networking as a single consolidated solution. But many of the offerings in this space are simply blueprints of different components from multiple manufacturers brought together under one SKU (stock-keeping unit). The only real advantage these solutions offer is that the combination of components has been tested and verified in a lab to be supported by the respective vendors.

SimpliVity took convergence a step further and developed a *hyper-converged infrastructure* solution. This means that all of the infrastructure components are not only in one SKU, they are all integrated in a single physical node – the OmniCube. This hyper-converged solution includes server and storage resources, along with a set of advanced functions for managing and protection the virtual machines, and includes a specialized accelerator card to optimize performance and efficiency of data storage.

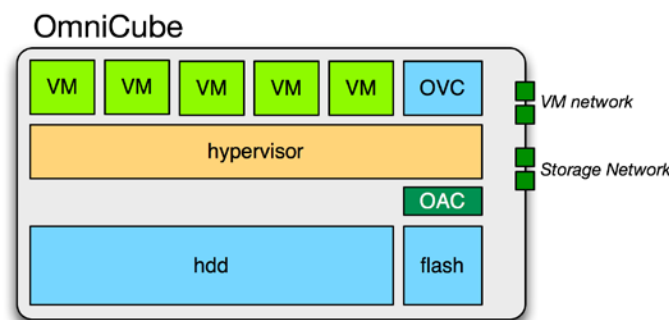
The SimpliVity solution doesn't end with hardware components. A tremendous effort has been made to simplify the initial deployment without adding unnecessary management interfaces. Everything is centrally managed through a single pane of glass in the vCenter interface.

SimpliVity OmniStack

The OmniCube Building Block Architecture

OmniStack is the set of technologies powering the OmniCube, and includes several innovations, which this paper will describe. OmniCube's functionality begins with the OmniCube Virtual Controller (SVC). This is a VM that is deployed with the initial configuration of the OmniCube node. All of the intelligence of the system resides in this controller.

The second important part of the OmniStack is the Data Virtualization Engine (DVE), which performs inline deduplication and compression on the data flow. To enable high performance with no impact to latency, OmniCube contains the OmniCube Accelerator Card (OAC), a purpose built PCIe device that contains a field-programmable gate array (FPGA) as well as several other components to accelerate system processing. The OmniCube Accelerator is directly attached to the SVC and processes all of the writes that enter OmniCube in real time.



Powered by this OmniStack technology, OmniCube is the SimpliVity hyper-converged building block. Two or more of these building blocks (also called 'nodes'), can be combined into a datacenter federation, creating a scalable pool of shared resources.

The server hardware is a 2U standardized x86 server from a tier 1 server manufacturer, which guarantees proven quality and serviceability for the end-user. Inside this server is performance-optimized flash storage, with 4 solid-state drives (SSDs) in a hardware RAID-5 configuration, as well as capacity-optimized storage with 8 or 20 hard disk drives (HDDs) in a hardware RAID-6 configuration. This storage is presented directly to the SVC, making the storage hardware layer completely transparent to the storage controller.

From a network perspective, there are two Gigabit Ethernet (1GbE) ports for the VM network as well as two ten Gigabit Ethernet (10 GbE) ports for the storage network. Additional 1GbE or 10GbE ports can be added in the future.

Installation

OmniCube arrives with VMware ESXi pre-installed. To get started, one must simply install the SimpliVity VMware vCenter plugin and connect the hypervisor to vCenter. It will automatically be detected as a new empty SimpliVity box and a guided configuration process will begin.

A new Virtual Controller will be deployed and two standard vSwitches will be installed - one for the 1GbE NICs (VM network) and one for the 10GbE NICs (storage network). Three logical networks are created as well: management, storage (front-end NFS), and federation (back-end traffic). If the front-end NFS port group is made visible through physical switches, an existing ESXi host in the datacenter could use the OmniStack datastores as well.

In a matter of just a few minutes the first VMs can be deployed. Adding nodes to the stack follows the same simple procedure.

Optimized Data Flow on Write

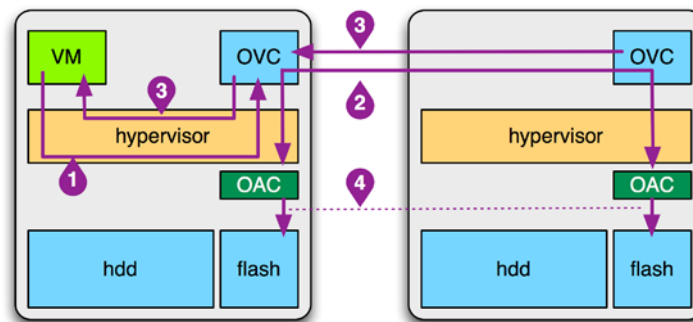
The data flow in the OmniStack is optimized first for low latency and performance (I/O operations per second or IOPS). There are four main steps when new data comes in:

The IO from the VM goes through the hypervisor storage stack to the virtual controller (NFS).

The controller sends the data to both a local and a remote OmniCube Accelerator on another node in the same datacenter.

Once the local and remote Accelerator have received the data it will be acknowledged to the VM.

The data gets processed for compression and deduplication by both OmniCube Accelerators and written to their underlying storage layer by the Virtual Controller



Because a second host already received the data, there is no need to write the data to disk first before acknowledging the IO to the VM. This will tremendously increase IOPS and decrease latency without sacrificing reliability or risking data loss.

Data Optimization

All incoming data is processed in real-time by the OmniCube Accelerator for both deduplication and compression. A significant benefit of real-time processing is that there is no need for uncompressed data to be stored on disk before it can be optimized post-process. Additionally, when there are multiple machines with the same type of data, such as operating system binaries, that data will only have to be written once. This saves storage space and gives the disks more spare time for read actions.

Global Federated Architecture

One of the more substantial benefits of the SimpliVity OmniStack is the data optimized footprint. As the blocks of data are deduplicated, they have a key in the global cache. Whenever a block of new data needs to be written that already exists, a new key is simply added to this table.

This process is exactly the same for copying the data. For example, through the support for VAAI in version 2.1 of the software, cloning a VM takes just a few seconds.

Data Protection Policies

This is also true for data protection, even across multiple sites. For example, when a Windows Server 2012 machine has to be copied over the WAN to another site, only the unique, compressed, optimized blocks of data will be copied. If at least one Windows Server 2012 instance is already running at the other location, then none of the blocks referenced by Windows need to be moved over the WAN. Even then, unique blocks would already be deduplicated and compressed making it bandwidth friendly.

This means that VMs can be protected across the country or throughout the world multiple times a day. The data protection points are always full VM images as they are the composition of all unique data blocks of that machine at any given time. There are no “full plus incremental” protection points necessary for recovery.

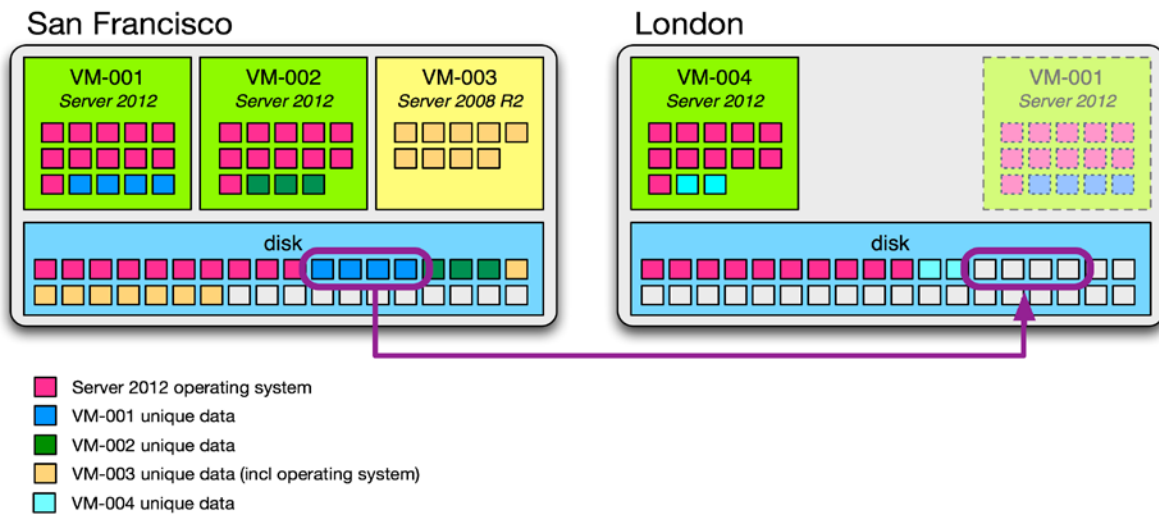
Setting a Protection Policy

A data protection policy consists of multiple rules. A single rule holds the protection schedule, the retention period, the datacenter to copy to (local or remote), and the application consistency requirement for a given machine.

When application consistency is checked, VMware tools will be used for quiescing the application and the VM before taking the VM storage snapshot. If the VM is running Microsoft Windows, Volume Shadow Copy Services (VSS) will be used, while Linux VMs will look for pre-freeze and post-thaw scripts.

The protection policies are per VM. When looking at the VM's SimpliVity tab it will show the active protection policy. A default policy is also set at the datastore level to prevent new VMs running unprotected if the administrator forgets to enable a specific policy for that VM.

In the following illustration, VM-001 resides in San Francisco but could easily be started in London by simply copying the 4 blocks of unique data.



Economically-Efficient Retention

It is important to retain multiple copies of data for reliability and redundancy. Having those copies near production for fast recovery is equally important. Therefore, it is best-practice to have multiple retention policies in a data protection strategy.

A classic data retention policy would include 15 daily recovery points, 4 weekly, 12 monthly and 5, 10 or forever yearly recovery points. But is all of that redundant data storage really necessary? What if long-term copies could be stored more cheaply than short-term recovery points?

SimpliVity has thought about this and has the capability to use Amazon EC2 as one of the datacenter locations for backup data. This significantly reduces storage cost and complexity for long-term retention without impacting the retention schedule and data availability.

Models

SimpliVity now offers three OmniCube models:

1. The CN-2000 (NEW) for Remote Office/Branch Office (ROBO) locations
2. The original CN-3000 as the enterprise model
3. The CN-5000 (NEW) for maximum performance infrastructures.

These models differ in terms of CPU power, memory (RAM) and disk configurations.

OmniCube™ Family of Hardware Platforms



CN-2000 (NEW)

The new CN-2000 is targeted at remote office / branch offices (ROBO) or small and mid-size enterprise (SME) environments. With a usable capacity between 5 to 10 TB depending on the level of optimization of the workload, this is suitable for a complete localized setup. Datacenter components like a Domain Controller and a DNS server, a file server, a small mail server and a document management platform like Microsoft SharePoint all fit perfectly in this small hyper-converged platform. The only difference between the models is their size, which makes this a very useful startup environment that can scale with the company.

Technical Specifications:

- Single socket 6 cores
- 128 GB DRAM
- 5-10 TB usable capacity: 4 x 100 GB SSDs + 8 x 1 TB HDDs
- 2x 1GbE and/or 2x 10GbE
- OmniCube Accelerator v2.0

CN-3000 – 2nd generation

First launched in August 2012, the CN-3000 is the SimpliVity flagship. It is the base building block for a mid-sized or larger environment. Where the first generation CN-3000 had 12 cores

total, the new CN-3000 offers 16 or 24 CPU cores. Customers can also now specify various SSD configurations, adding capacity for flash cache. With the big 3 TB HDDs, total system capacity can reach 40 TB or more after data optimization. Of course the CN-3000 2nd generation has the new v2 OmniCube Accelerator as well.

Technical specifications:

- Dual socket 8 or 12 core processors (16-24 total cores)
- From 128 GB to 768 GB RAM
- 4x 200 GB / 400 GB / 800 GB SSDs (max 3.2 TB raw SSD capacity)
- 8x 3 TB NL SAS
- 2x 1GbE and or 2x 10GbE
- OmniCube Accelerator v2.0
- Option to add additional 10GbE or 1GbE ports

CN-5000 (NEW)

While the CN-2000 has been added for smaller environments, now there is also the higher-density CN-5000 with more power in the same footprint. These nodes have higher CPU configurations and bigger RAM options. But the biggest difference is that they come with 20 10,000 rpm disks for even better I/O performance.

Technical specifications:

- Dual socket 10 or 12 core processors (20-24 total cores)
- From 384 GB to 786 GB RAM
- 4x 400 GB or 800 GB SSDs (max 3.2 TB raw SSD capacity)
- 20x 900 GB 10k SAS
- 2x 1GbE and or 2x 10GbE
- OmniCube Accelerator v2.0
- Option to add additional 10GbE or 1GbE ports

Use Cases

SimpliVity's OmniStack is designed to address a wide range of environments and use cases, and the flexibility of OmniCube models allow end users to select the right solution for their requirements. Below are two common use cases that would benefit from the OmniStack technology.

The Retail Store

Many enterprises with retail locations initially deploy centralized servers and data management. They deploy a single datacenter and use web applications for local stores running on thin clients at the remote site. Although this might look cheaper from an infrastructure model, the costs can dramatically increase when stores are cut off from the Internet.

Therefore, a preferred model for retailers is to deploy a small local server rack in each location. But this has proven very expensive and difficult to manage, with many different infrastructure components residing in a back room or office.

With a single SimpliVity CN-2000 unit per store, cost and management is reduced and capability and flexibility improves. Each store would only have a single device, including server, storage, and data protection. Each is linked to the whole OmniStack infrastructure yet each can run independently when connectivity goes down.

Mergers and Acquisitions

When one company acquires another, or when two companies merge, there is always the problem of duplicate infrastructure and the management costs associated with this. Chances are very high that the hardware is not the same and that the CIO will have to decide which hardware to keep and which to remove.

By bringing in SimpliVity OmniCube at the remote locations, VMs can be moved non-disruptively through storage- or enhanced-vMotion. Once all the machines are moved to the new OmniCube only the unique data of those machines needs to be transferred to the existing headquarters. With a single reboot of the VMs, the whole existing infrastructure has been migrated seamlessly to headquarters.

Summary

SimpliVity's OmniStack is a novel hyper-converged enterprise IT infrastructure architecture designed for virtualized environments and distributed enterprises. Consisting of two or more OmniCube nodes, an OmniStack-powered infrastructure incorporates server, storage, and network resources, along with a broad set of core infrastructure functions, in a compact physical footprint. Each OmniCube includes advanced SSD caching and data optimization features as well as VM-centric management integrated with VMware vSphere for performance and ease of use. SimpliVity's advanced federated architecture enables distributed companies to leverage deduplication and compression for bandwidth efficient remote replication, data migrations and data sharing.