

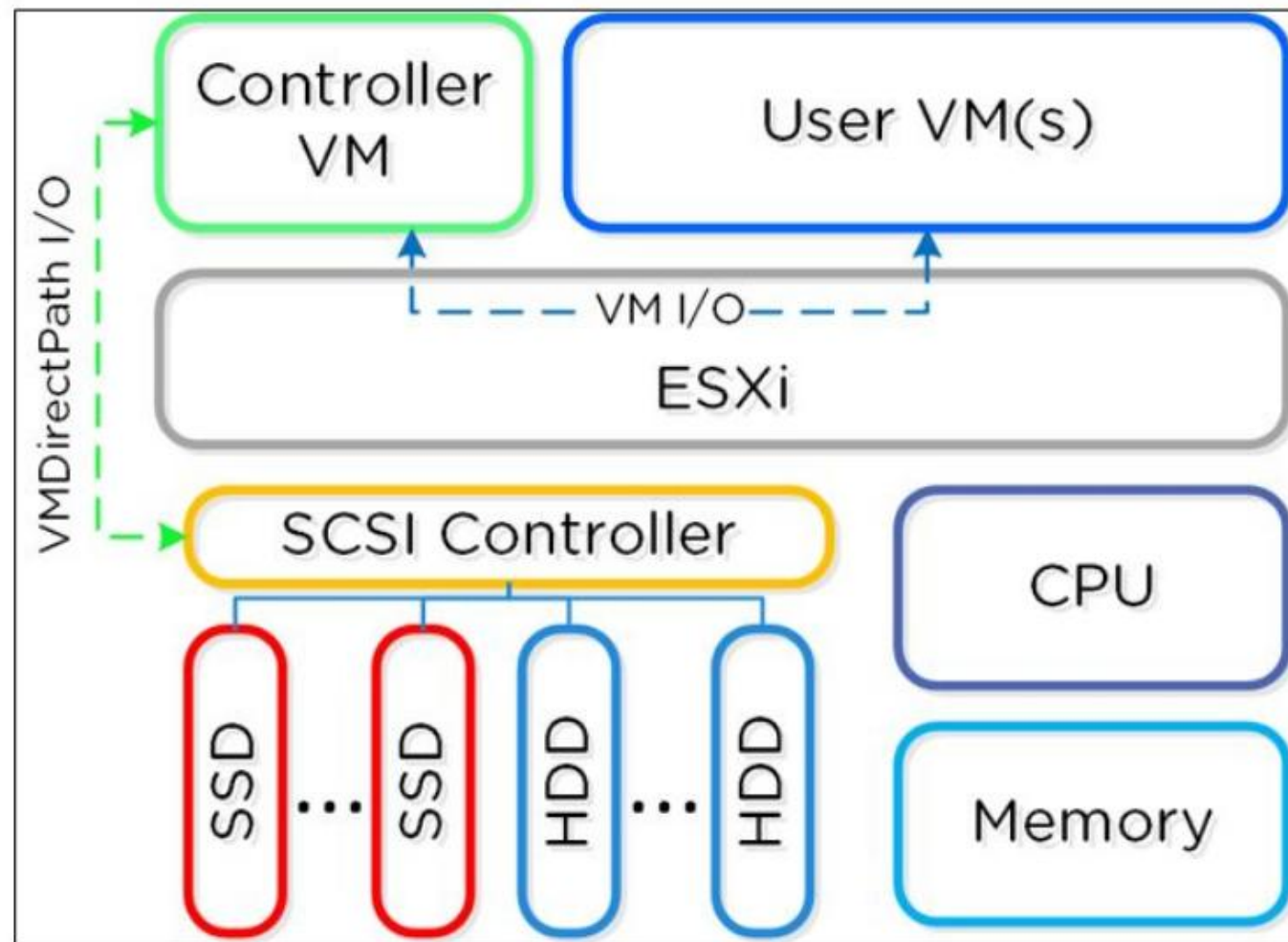
ESXi architecture

Overview of ESXi

Lenovo

ESXi node architecture

In ESXi deployments, the CVM runs as a VM and disks are presented using VMDirectPath I/O. This allows the full PCI controller (and attached devices) to be passed directly through to the CVM and bypass the hypervisor.



Networking

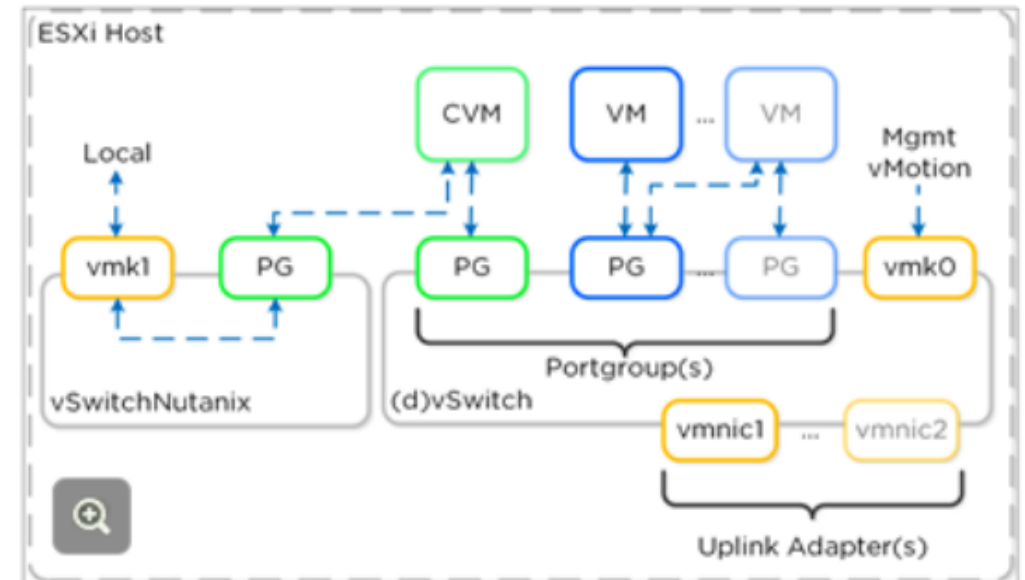
Each ESXi host has a local vSwitch which is used for intra-host communication between the Nutanix CVM and the host. For external communication with VMs, a standard vSwitch (default) or dvSwitch is leveraged.

The local vSwitch (vSwitchNutanix) is for local communication between the Nutanix CVM and the ESXi host. The host has a vmkernel interface on this vSwitch (vmk1 - 192.168.5.1), and the CVM has an interface bound to a port group on this internal switch (svm-iscsi-pg - 192.168.5.2). This is the primary storage communication path.

The external vSwitch can be a standard vSwitch or a dvSwitch. This will host the external interfaces for the ESXi host and CVM as well as the port groups leveraged by VMs on the host. The external vmkernel interface is leveraged for host management, vMotion, and so on. The external CVM interface is used for communication to other Nutanix CVMs. Assuming the VLANs are enabled on the trunk, as many port groups as required can be created.

The diagram shows a conceptual overview of the virtual switch architecture.

Click to enlarge the image.



How ESXi works on Nutanix

The Nutanix platform supports the VMware APIs for Array Integration (VAAI), which allows the hypervisor to offload certain tasks to the array. This is much more efficient as the hypervisor doesn't need to be the "man in the middle." Nutanix currently supports the VAAI primitives for NAS, including the "full file clone," "fast file clone," and "reserve space" primitives. For more information about the various primitives, refer to the following article:

<http://cormachogan.com/2012/11/08/vaa-comparison-block-versus-nas/>.

For both the full and fast file clones, a DSF "fast clone" is carried out. This means creating a writable snapshot (using re-direct on write) for each clone that is created. Each of these clones has its own block map, meaning that chain depth isn't a concern. The following information shows whether or not VAAI is used in specific scenarios:

- Clone VM with Snapshot → VAAI will NOT be used
- Clone VM without Snapshot which is Powered Off → VAAI WILL be used
- Clone VM to a different Datastore/Container → VAAI will NOT be used
- Clone VM which is Powered On → VAAI will NOT be used

These scenarios apply to VMware View:

- View Full Clone (Template with Snapshot) → VAAI will NOT be used
- View Full Clone (Template w/o Snapshot) → VAAI WILL be used
- View Linked Clone (VCAI) → VAAI WILL be used

Users can validate that VAAI operations are taking place by using the "NFS Adapter" Activity Traces page.

Useful ESXi administration links

- vSphere Administration Guide for Acropolis
<https://portal.nutanix.com/#/page/docs/details?targetId=vSphere-Admin6-AOS-511:vSphere-Admin6-AOS-v511>
- Collecting diagnostic information for ESX/ESXi hosts and vCenter Server using the vSphere Web Client <https://kb.vmware.com/s/article/2032892>
- vSphere networking <https://portal.nutanix.com/#/page/docs/details?targetId=vSphere-Admin6-AOS-v511:vsp-vsphere-networking-c.html>
- vSphere networking components
<https://portal.nutanix.com/#/page/docs/details?targetId=vSphere-Admin6-FLEX-AOS-v511:vsp-vsphere-networking-components-r.html>