Understanding Hypervisors

- Load balancing virtualizes systems and resources by mapping a logical address to a physical address. Another fundamental technology for abstraction creates virtual systems out of physical systems.
- If load balancing is like playing a game of hot potato, then virtual machine technologies is akin to playing slice and dice with the potato.
- Given a computer system with a certain set of resources, you can set aside portions of those resources to create a virtual machine.
- The standpoint of applications or users, a virtual machine has all the attributes and characteristics of a physical system but is strictly software that emulates a physical machine.
- A system virtual machine (or a hardware virtual machine) has its own address space in memory, its own processor resource allocation, and its own device I/O using its own virtual device drivers.
- Virtual machines are designed to run only a single application or process and are referred to as process virtual machines.

Understanding Hypervisors

- A virtual machine is a computer that is walled off from the physical computer that the virtual machine is running on
- This makes virtual machine technology very useful for running old versions of operating systems.
- Testing applications in what amounts to a sandbox, or in the case of cloud computing, creating virtual machine instances that can be assigned a workload.
- Virtual machines provide the capability of running multiple machine instances, each with their own operating system.
- VMMs to manage application provisioning, provide for machine instance cloning and replication, allow for graceful system failover, and provide several other desirable features.
- The downside of virtual machine technologies is that having resources indirectly addressed means there is some level of overhead.

Understanding Hypervisors

Virtual machine types :

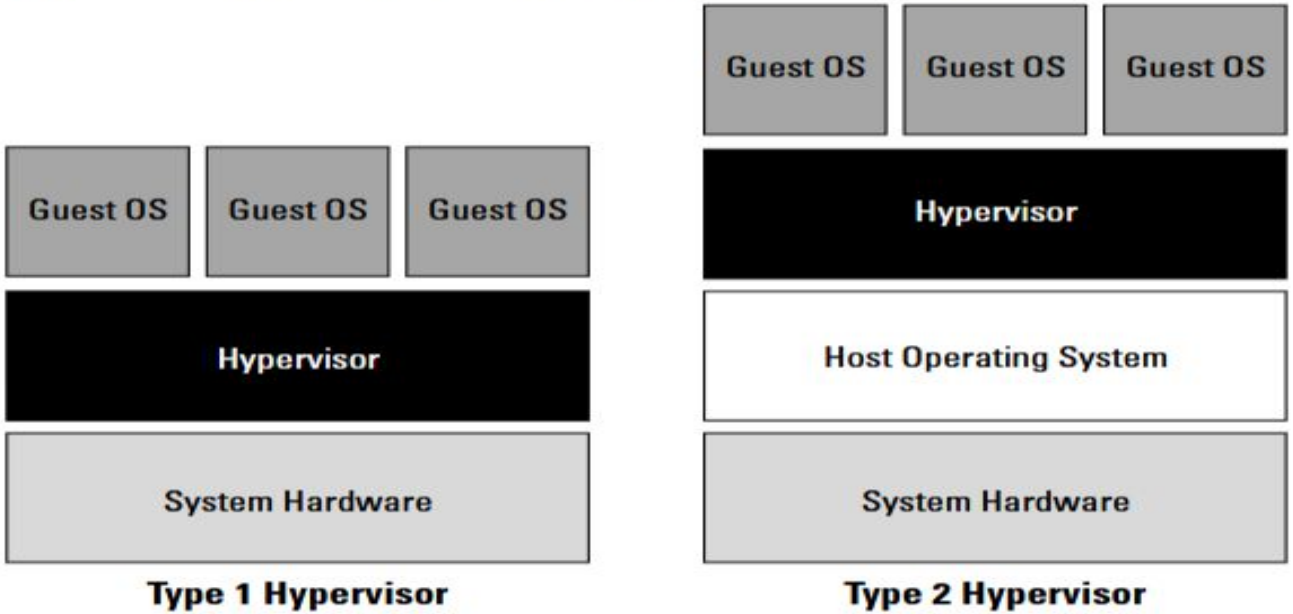
- A low-level program is required to provide system resource access to virtual machines, and this program is referred to as the hypervisor or Virtual Machine Monitor (VMM).
- A hypervisor running on bare metal is a Type 1 VM or native VM.
- Examples of Type 1 Virtual Machine Monitors are LynxSecure, RTS Hypervisor, Oracle VM, Sun xVM Server, VirtualLogix VLX, VMware ESX and ESXi, and Wind River VxWorks, among others.
- The operating system loaded into a virtual machine is referred to as the guest operating system, and there is no constraint on running the same guest on multiple VMs on a physical system.
- Type 1 VMs have no host operating system because they are installed on a bare system.
- An operating system running on a Type 1 VM is a full virtualization because it is a complete simulation of the hardware that it is running on.

Understanding Hypervisors

- Some hypervisors are installed over an operating system and are referred to as Type 2 or hosted VM.
- Examples of Type 2 Virtual Machine Monitors are Containers, KVM, Microsoft Hyper V, Parallels Desktop for Mac, Wind River Simics, VMWare Fusion, Virtual Server 2005 R2, Xen, Windows Virtual PC, and VMware Workstation 6.0 and Server, among others.
- This is a very rich product category. Type 2 virtual machines are installed over a host operating system; for Microsoft Hyper-V, that operating system would be Windows Server.
- In the section that follows, the Xen hypervisor (which runs on top of a Linux host OS) is more fully described. Xen is used by AmazoN Web Services to provide Amazon Machine Instances (AMIs). On a Type 2 VM, a software interface is created that emulates the devices with which a system would normally interact.
- This abstraction is meant to place many I/O operations outside the virtual environment, which makes it both programmatically easier and more efficient to execute device.
- I/O than it would be inside a virtual environment. This type of virtualization is sometimes referred to as para virtualization, and it is found in hypervisors such as Microsoft's Hyper-V and Xen. It is the host operating system that is performing the I/O through a para-API.

Understanding Hypervisors

VMware's vSphere cloud computing infrastructure model

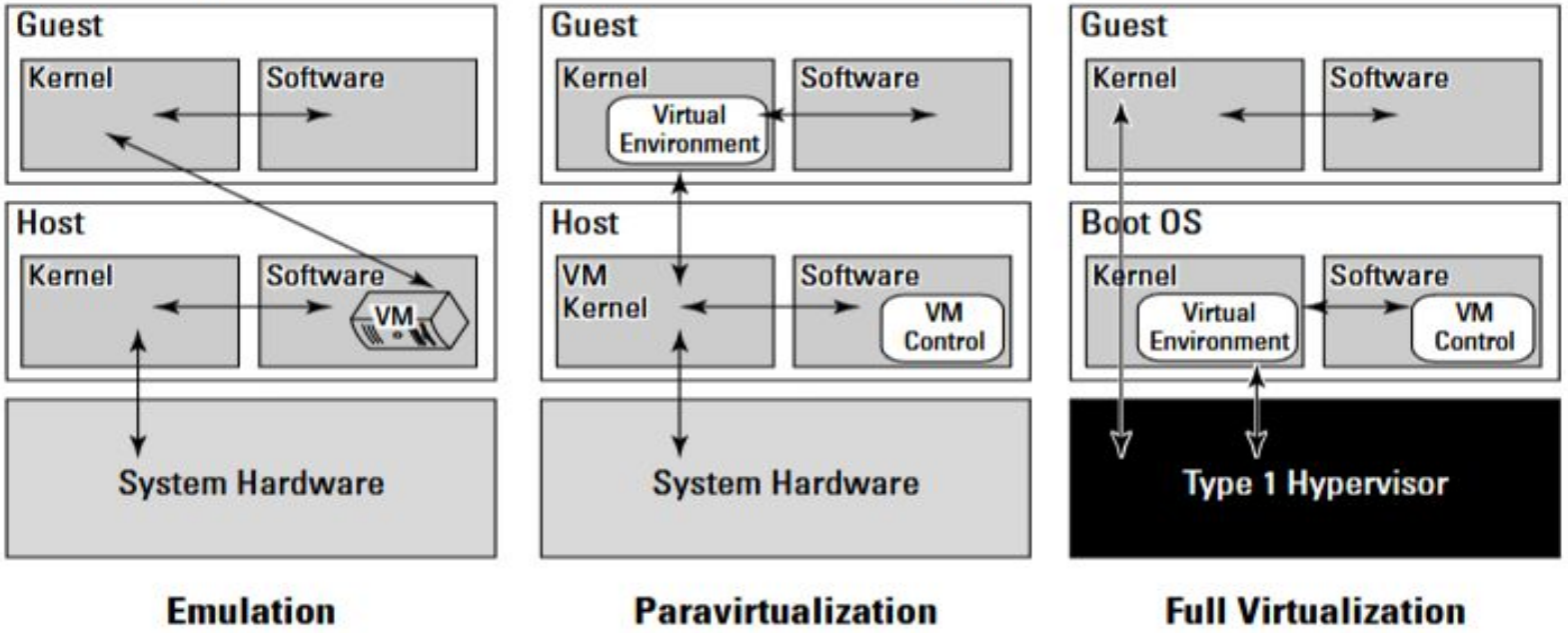


Understanding Hypervisors

- The difference between emulation, para virtualization, and full virtualization. In emulation, the virtual machine simulates hardware, so it can be independent of the underlying system hardware. A guest operating system using emulation does not need to be modified in any way.
- Para virtualization requires that the host operating system provide a virtual machine interface for the guest operating system and that the guest access hardware through that host VM.
- An operating system running as a guest on a para virtualization system must be ported to work with the host interface. Finally, in a full virtualization scheme, the VM is installed as a Type 1 Hypervisor directly onto the hardware.
- All operating systems in full virtualization communicate directly with the VM hypervisor, so guest operating systems do not require any modification.
- Guest operating systems in full virtualization systems are generally faster than other virtualization schemes.
- The Virtual Machine Interface (VMI) open standard (http://vmi.ncsa.uiuc.edu/) that VMware has proposed is an example of a para virtualization API.

Understanding Hypervisors

Emulation, paravirtualization, and full virtualization types



Understanding Hypervisors

- Applications running inside an application virtual machine are generally slow, but these programs are very popular because they provide portability, offer rich programming languages, come with many advanced features, and allow platform independence for their programs.
- Cloud computing applications provide process virtual machine applications, this type of abstraction isn't really suitable for building a large or high-performing cloud network, with one exception.
- The exception is the process VMs that enable a class of parallel cluster computing applications.
- These applications are high-performance systems where the virtual machine is operating one process per cluster node, and the system maintains the necessary intra-application communications over the network interconnect.
- Examples of this type of system are the Parallel Virtual Machine (PVM; see http://www.csm.ornl.gov/pvm/pvm_home.html)
- The Message Passing Interface (MPI; see http://www.mpi-forum.org/). Some people do not consider these application VMs to be true virtual machines, noting that these applications can still access the host operating system services on the specific system on which they are running.

Understanding Hypervisors

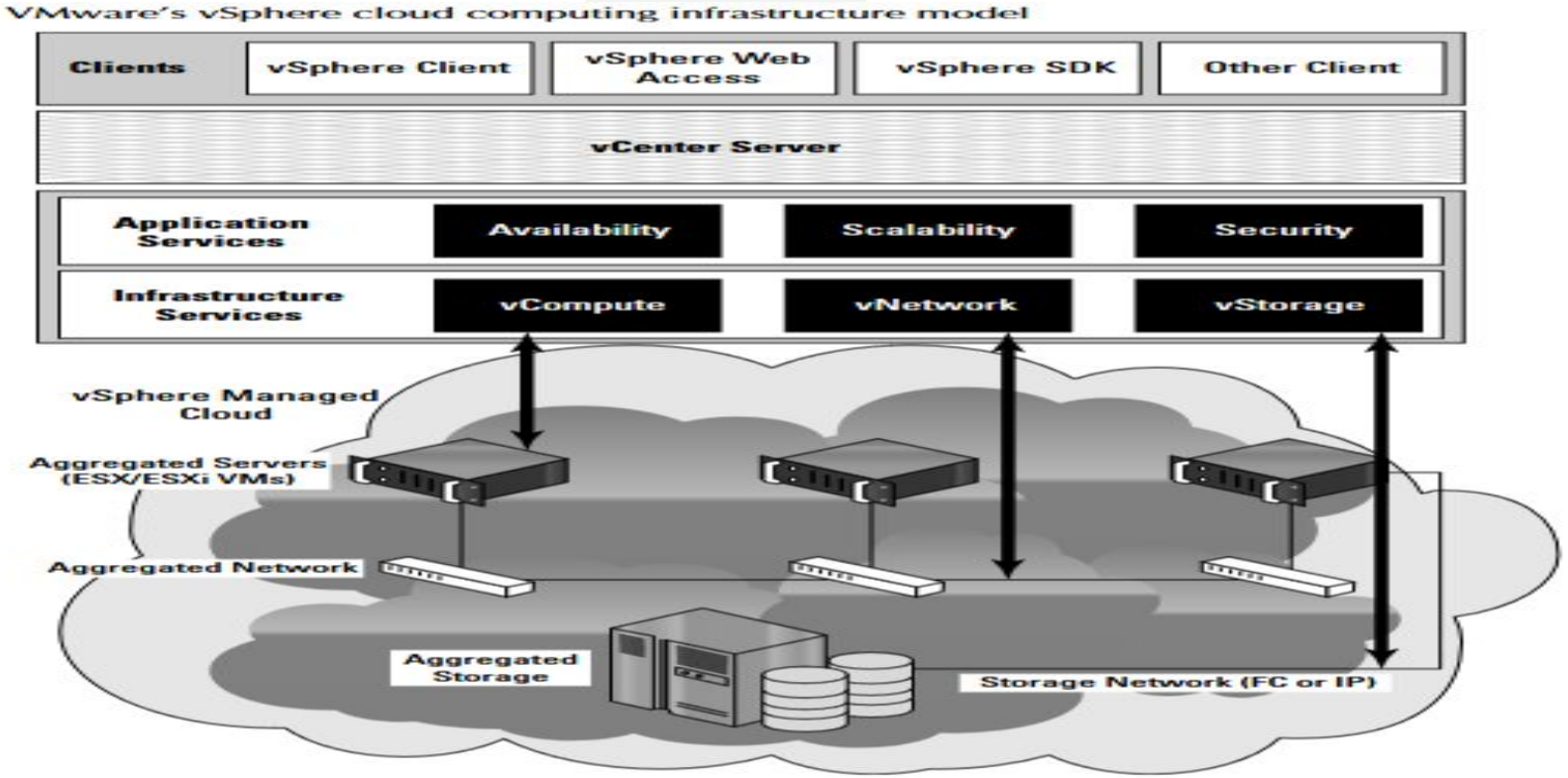
- Some operating systems such as Sun Solaris and IBM AIX 6.1 support a feature known as operating system virtualization.
- This type of virtualization creates virtual servers at the operating system or kernel level. Each virtual server is running in its own virtual environment (VE) as a virtual private server (VPS).
- Different operating systems use different names to describe these machine instances, each of which can support its own guest OS.
- However, unlike true virtual machines, VPS must all be running the same OS and the same version of that OS.
- Sun Solaris 10 uses VPS to create what is called Solaris Zones. With IBM AIX, the VPS is called a System Workload Partition (WPAR).
- This type of virtualization allows for a dense collection of virtual machines with relatively low overhead.

Understanding Hypervisors

VMware vSphere :

- VMware vSphere is a management infrastructure framework that virtualizes system, storage, and networking hardware to create cloud computing infrastructures.
- vSphere is the branding for a set of management tools and a set of products previously labeled VMware Infrastructure.
- vSphere provides a set of services that applications can use to access cloud resources, including these:
- VMware vCompute: A service that aggregates servers into an assignable pool
- VMware vStorage: A service that aggregates storage resources into an assignable pool
- VMware vNetwork: A service that creates and manages virtual network interfaces
- Application services: Such as HA (High Availability) and Fault Tolerance
- vCenter Server: A provisioning, management, and monitoring console for VMware cloud infrastructures

Understanding Hypervisors



Understanding Hypervisors

- A vSphere cloud is a pure infrastructure play. The virtualization layer that abstracts processing, memory, and storage uses the VMware ESX or ESXi virtualization server.
- ESX is a Type 1 hypervisor; it installs over bare metal (a clean system) using a Linux kernel to boot and installs the vmkernel hypervisor (virtualization kernel and support files).
- When the system is rebooted, the vmkernel loads first, and then the Linux kernel becomes the first guest operating system to run as a virtual machine on the system and contains the service console.
- VMware is a very highly developed infrastructure and the current leader in this industry. A number of important add-on products are available for cloud computing applications.

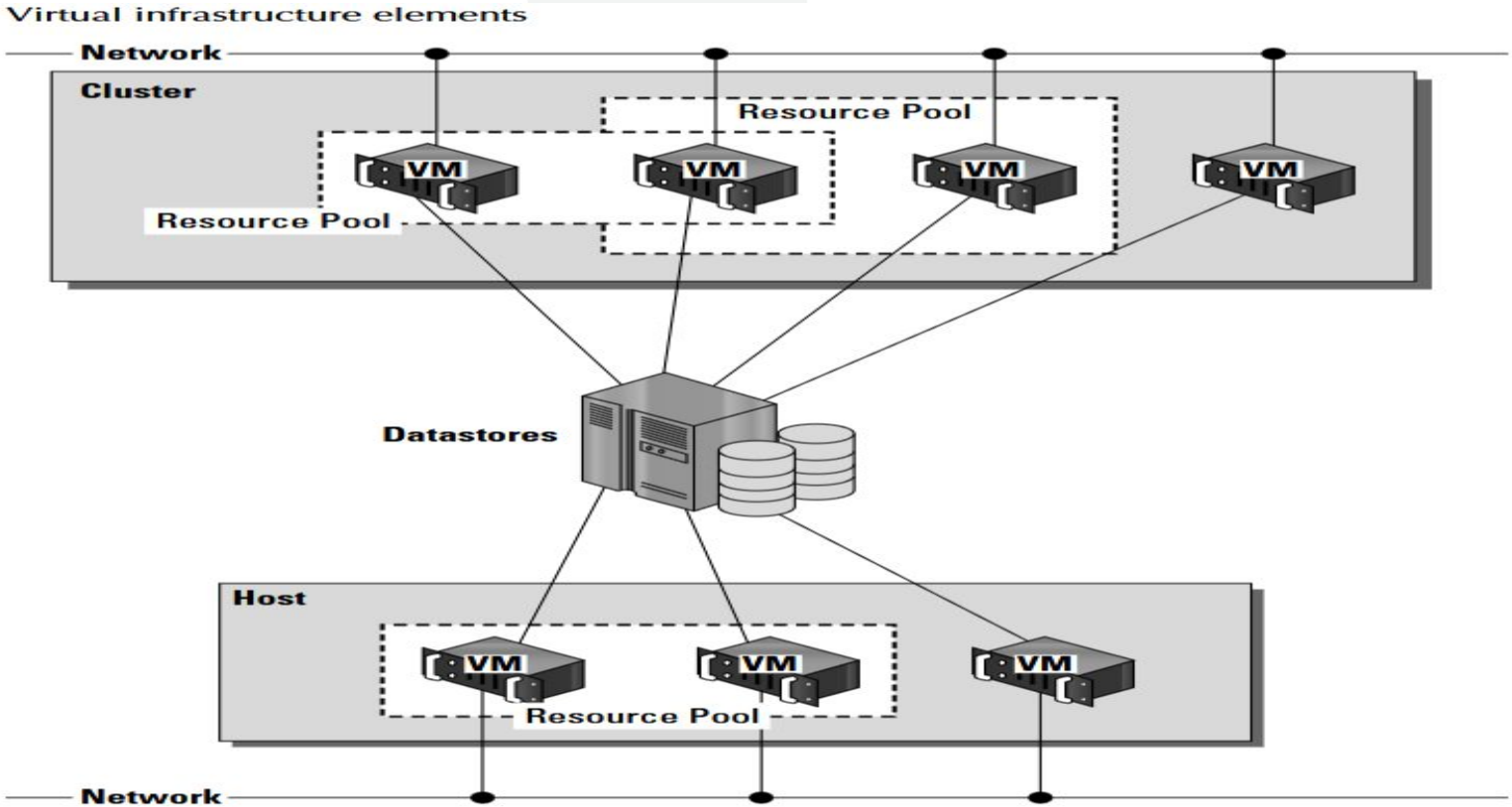
These are among the more notable products:

- **Virtual Machine File System (VMFS):** A high-performance cluster file system for an ESX/ESXi cluster.
- **VMotion:** A service that allows for the migration of a virtual machine from one physical server to another physical server while the virtual server runs continuously and without any interruption of ongoing transactions.

Understanding Hypervisors

- **Storage VMotion:** A product that can migrate files from one datastore to another data store while the virtual machine that uses the datastore continues to run.
- **Virtual SMP:** A feature that allows a virtual machine to run on two or more physical processors at the same time.
- **Distributed Resource Scheduler (DRS):** A system for provisioning virtual machines and load balancing processing resources dynamically across the different physical systems that are in use.
- A part of the DRS called the distributed power management (DPM) module can manage the power consumption of systems.
- **vNetwork Distributed Switch (DVS):** A capability to maintain a network runtime state for virtual machines as they are migrated from one physical system to another.
- **DVS also monitors network** connections, provides firewall services, and enables the use of third party switches such as the Cisco Nexus 1000V to manage virtual networks.

Understanding Hypervisors



Understanding Hypervisors

- Storage resources can be either Direct Attached Storage (DAS) of a server using SCSI, SAS, or SATA connections, Fibre Channel disk arrays/SANs, iSCSI disk arrays/SANs, or Network Attached Storage (NAS) disk arrays.
- Although the lines drawn between the data store and different VMs indicate a direct connection, with the exception of DAS, the other storage types are shared storage solutions.
- Storage virtualization is most commonly achieved through a mapping mechanism where a logical storage address is translated into a physical storage address. Block-based storage such as those used.
- In SANs use a feature called a Logical Unit Identifier (LUN) with specific addresses stored in the form of an offset called the Logical Block Address (LBA). The address space mapping then maps the address of the logical or virtual disk (vdisk) to the logical unit on a storage controller.
- Storage virtualization may be done in software or in hardware, and it allows requests for virtualized storage to be redirected as needed.
- Network virtualization abstracts networking hardware and software into a virtual network that can be managed. A virtual network can create virtual network interfaces (VNICs) or virtual LANs (VLANS) and can be managed by a hypervisor, operating system, or external management console.

Understanding Machine Imaging

- The abstractions that cloud computing needs can be achieved through redirection and virtualization. A third mechanism is commonly used to provide system portability, instantiate applications, and provision and deploy systems in the cloud.

- This third mechanism is through storing the state of a systems using a system image. A system image makes a copy or a clone of the entire computer system inside a single container such as a file.

- The system imaging program is used to make this image and can be used later to restore a system image. Some imaging programs can take snapshots of systems, and most allow you to view the files contained in the image and do partial restores.

- A prominent example of a system image and how it can be used in cloud computing architectures is the Amazon Machine Image (AMI) used by Amazon Web Services to store copies of a virtual machine.

- When you subscribe to AWS, you can choose to use one of its hundreds of canned AMIs or to create a custom system and capture that system's image to an AMI.
- An AMI can be for public use under a free distribution license, for pay-per-use with operating systems such as Windows, or shared by an EC2 user with other users who are given the privilege of access.

Porting Applications

- Cloud computing applications have the ability to run on virtual systems and for these systems to be moved as needed to respond to demand.
- Systems (VMs running applications), storage, and network assets can all be virtualized and have sufficient flexibility to give acceptable distributed WAN application performance.
- Developers who write software to run in the cloud will undoubtedly want the ability to port their applications from one cloud vendor to another, but that is a much more difficult proposition.
- Cloud computing is a relatively new area of technology, and the major vendors have technologies that don't interoperate with one another.

The Simple Cloud API :

- If you build an application on a platform such as Microsoft Azure, porting that application to Amazon Web Services or GoogleApps may be difficult, if not impossible.
- In an effort to create an interoperability standard, Zend Technologies has started an open source initiative to create a common application program interface that will allow applications to be portable.

Porting Applications

- The initiative is called the Simple API for Cloud Application Services (http://www.simplecloud.org/), and the effort has drawn interest from several major cloud computing companies.
- Among the founding supporters are IBM, Microsoft, Nivanix, Rackspace, and GoGrid. Simple Cloud API has as its goal a set of common interfaces for:
- **File Storage Services:** Currently Amazon S3, Windows Azure Blob Storage, Nirvanix and Local storage is supported by the Storage API.
- **Document Storage Services**: Amazon SimpleDB and Windows Azure Table Storage are currently supported. Local document storage is planned.
- **Simple Queue Services**: Amazon SQS, Windows Azure Queue Storage, and Local queue services are supported. Zend intends to add the interface to their open source PHP Framework (http://www.frame work.zend.com) as the Zend_Cloud framework component.
- Vendors such as Microsoft and IBM are supplying adapters that will use part of the Simple Cloud API for their cloud application services.

AppZero Virtual Application Appliance

- Applications that run in datacenters are captive to the operating systems and hardware platforms that they run on. Many datacenters are a veritable Noah's Ark of computing.
- Moving an application from one platform to another isn't nearly as simple as moving a machine image from one system to another.
- The situation is further complicated by the fact that applications are tightly coupled with the operating systems on which they run. An application running on Windows, for example, isn't isolated from other applications.
- When the application loads, it often loads or uses different Dynamic Link Libraries (DLL), and it is through the sharing or modification of DLLs that Windows applications get themselves in trouble.
- Further modifications include modifying the registry during installation. These factors make it difficult to port applications from one platform to another without lots of careful work.
- If you are a Platform as a Service (PaaS) application developer, you are packaging a complete software stack that includes not only your application, but the operating system and application logic and rules as well. Vendor lock-in for you application is assured.

AppZero Virtual Application Appliance

- The ability to run an application from whatever platform you want is not one of the characteristics of cloud computing, but you can imagine that it is a very attractive proposition.
- While the Simple Cloud API is useful for applications written in PHP, other methods may be needed to make applications easily portable.
- One company working on this problem is AppZero (http://www. appzero.com/), and its solution is called the Virtual Application Appliance (VAA).
- The AppZero solution creates a virtual application appliance as an architectural layer between the Windows or the UNIX operating system and applications.
- The virtualization layer serves as the mediator for file I/O, memory I/O, and application calls and response to DLLs, which has the effect of sandboxing the application.
- The running application in AppZero changes none of the registry entries or any of the files on the Windows Server.

AppZero Virtual Application Appliance

- VAA creates a container that encapsulates the application and all the application's dependencies within a set of files; it is essentially an Application Image for a specific OS.
- Dependencies include DLL, service settings, necessary configuration files, registry entries, and machine and network settings.
- This container forms an installable server-side application stack that can be run after installation, but has no impact on the underlying operating system.
- VAAs are created using the AppZero Creator wizard, managed with the AppZero Admin tool, and may be installed using the AppZero Director, which creates a VAA runtime application.
- If desired, an application called AppZero Dissolve removes the VAA virtualization layer from the encapsulated application and installs that application directly into the operating system.
- AppZero envisages using VAAs to create what it calls a stateless cloud. In a stateless cloud, the application's state information is stored on a network share where it is available to run on different cloud systems as needed.