



Cloud.com CloudStack Administration Guide

Version 2.2.0 – 2.2.3

Revised September 8, 2011



Copyright © 2011 Citrix Systems, Inc. All rights reserved. Specifications are subject to change without notice. The Cloud.com logo, Cloud.com, and CloudStack are trademarks or registered trademarks of Citrix Systems, Inc. All other brands or products are trademarks or registered trademarks of their respective holders.

Contents

1	About CloudStack.....	9
1.1	Service Offerings, Disk Offerings, Network Offerings and Templates	9
1.2	Accounts, Users, and Domains	9
1.3	Server Types	9
1.3.1	Management Server.....	10
1.3.2	Hosts	10
1.4	Networking Features and Virtualization.....	10
1.4.1	Direct Attached Networking	11
1.4.2	Virtual Networking.....	11
1.4.3	Combining Virtual Networking and Direct Attached Networking.....	12
1.5	Storage Features and Virtualization	12
1.6	Administrator Controlled Allocation	12
1.7	Guest VM Management	13
1.8	Manageability.....	13
1.9	API and Extensibility	13
1.10	Scalability and Availability	14
2	Selecting Hardware and Software	15
2.1	Hosts.....	15
2.2	Management Servers	15
2.3	Storage	15
2.4	Network.....	15
2.5	Hypervisor Support.....	16
2.6	Guest OS and Software Support.....	16
3	Planning a Deployment.....	17
3.1	Management Server Farm.....	17
3.2	Scaling Concepts.....	17

- 3.2.1 Hosts 17
- 3.2.2 Clusters..... 17
- 3.2.3 Pods..... 18
- 3.2.4 Availability Zones 18
- 3.3 Multi-Site Deployment 19
- 4 Defining Your Service Offering 20
- 5 Understanding Network Types and Network Virtualization 21
 - 5.1 Guest Network 21
 - 5.2 Network Virtualization within One Pod..... 22
 - 5.3 Network Virtualization within One Availability Zone 23
 - 5.4 Network Virtualization 25
 - 5.5 Private Address Allocation 25
 - 5.6 Public Address Allocation 25
 - 5.7 External Network Elements..... 26
 - 5.7.1 Initial Setup 26
 - 5.7.2 Additional Configuration 26
- 6 Network Virtualization Features 27
 - 6.1 Guest Virtual Networks 27
 - 6.2 IP Forwarding and Firewalling 27
 - 6.3 IP Load Balancing..... 27
 - 6.4 Port Forwarding..... 27
 - 6.5 DNS and DHCP 27
 - 6.6 VPN 28
 - 6.6.1 Mac OS X 28
 - 6.6.2 Windows 28
 - 6.7 Working with Additional Networks 28
 - 6.7.1 Network Scope 28

- 6.7.2 Default and Non-default Networks 28
- 6.7.3 Adding an Additional Network 29
- 7 Storage Features and Types 31
 - 7.1 Primary Storage 31
 - 7.1.1 Tags 32
 - 7.1.2 Maintenance Mode 32
 - 7.2 Secondary Storage 33
 - 7.3 Changing the Secondary Storage IP Address 33
 - 7.4 Changing Secondary Storage Servers 33
 - 7.5 Working with Volumes 34
 - 7.5.1 Moving Volumes 34
 - 7.5.2 Resizing Volumes 34
 - 7.5.3 Volume Deletion and Garbage Collection 34
 - 7.6 Working with ISO Images 35
 - 7.7 Working with Blank VMs 35
 - 7.8 Working with Templates 35
 - 7.8.1 The Default Template 35
 - 7.8.2 Creating Templates 36
 - 7.8.3 Uploading Templates 37
 - 7.8.4 Extracting Templates 38
 - 7.8.5 Public Templates 38
 - 7.8.6 Private Templates 38
 - 7.8.7 Deleting Templates 38
 - 7.8.8 Running Sysprep for Windows Templates 38
 - 7.8.9 Importing AMIs 43
 - 7.8.10 Creating an Ubuntu 10.04 LTS Template for XenServer 46
 - 7.8.11 Converting a Hyper-V VM to a Template 47

- 7.8.12 Adding Password Management to Your Templates 48
- 7.9 Working with Snapshots..... 49
 - 7.9.1 Automatic Snapshot Creation and Retention 49
 - 7.9.2 Incremental Snapshots and Backup 50
 - 7.9.3 Volume Status 50
 - 7.9.4 Snapshot Restore 50
 - 7.9.5 Performance Considerations..... 50
- 8 Working with System Virtual Machines 51
 - 8.1 Console Proxy 51
 - 8.1.1 Changing the Console Proxy SSL Certificate and Domain 52
 - 8.2 Virtual Router 52
 - 8.3 Secondary Storage VM 52
- 9 System Reliability and HA 53
 - 9.1 Management Server 53
 - 9.2 Host 53
 - 9.3 Primary Storage Outage and Data Loss 53
 - 9.4 Secondary Storage Outage and Data Loss 53
 - 9.5 HA-Enabled VM (Service Provider and Enterprise Edition) 53
- 10 Management Features..... 55
 - 10.1 Users, Accounts, Administrators, and Domains 55
 - 10.1.1 Root Administrators 55
 - 10.1.2 Domain Administrators..... 55
 - 10.2 Provisioning 55
 - 10.2.1 Register 55
 - 10.3 Changing User and Administrator Passwords 55
 - 10.4 VM Lifecycle Management 56
 - 10.4.1 VM Creation..... 56

- 10.4.2 VM Deletion 56
- 10.4.3 VM Lifecycle..... 56
- 10.4.4 Remote Access..... 57
- 10.5 Changing the Database Configuration..... 57
- 10.6 PV Drivers 57
- 10.7 Administrator Alerts 57
- 10.8 Limits 57
 - 10.8.1 Configuration Limits..... 57
 - 10.8.2 ROOT Domain Limits..... 58
 - 10.8.3 Per-Domain Limits 58
- 11 Working with Hosts..... 60
 - 11.1 Adding Hosts to a Cluster 60
 - 11.1.1 vSphere Host Addition 60
 - 11.1.2 XenServer Host Addition..... 60
 - 11.1.3 KVM Host Addition 60
 - 11.2 Scheduled Maintenance and Maintenance Mode 60
 - 11.3 Removing Hosts..... 61
 - 11.3.1 XenServer and KVM Hosts 61
 - 11.3.2 vSphere Hosts 61
 - 11.4 Re-installing Hosts 62
 - 11.5 Changing Host IP Address..... 62
 - 11.6 Changing Host Password 62
 - 11.7 Host Allocation 62
 - 11.7.1 OS Preferences 62
 - 11.7.2 Over-Provisioning and Service Offering Limits 63
 - 11.8 VLAN Provisioning 63
- 12 Working with Usage..... 64

- 13 User Interface and API 65
 - 13.1 User Interface 65
 - 13.1.1 Admin User Interface 65
 - 13.1.2 End User Interface 65
 - 13.2 API 66
 - 13.2.1 Provisioning and Authentication API 66
 - 13.2.2 Allocators 66
 - 13.2.3 User Data and Meta Data 66
- 14 Tuning 68
 - 14.1 Increase Management Server Maximum Memory 68
 - 14.2 Set Database Buffer Pool Size 68
- 15 Troubleshooting 69
 - 15.1 Event Logs 69
 - 15.1.1 Standard Events 69
 - 15.1.2 Long Running Job Events 69
 - 15.1.3 Event Log Queries 69
 - 15.2 Working with Server Logs 71
 - 15.3 Data Loss on Exported Primary Storage 71
 - 15.4 Maintenance mode not working on vCenter 72
 - 15.5 Unable to deploy VMs from uploaded vSphere template 72
- 16 Appendix A—Time Zones 73

1 About CloudStack

The Cloud.com™ CloudStack™ platform is a complete software suite used to create Infrastructure as a Service (IaaS) clouds. Target customers include service providers and enterprises.

- The Cloud.com CloudStack platform enables service providers to set up an on-demand, elastic cloud computing service that is similar to the Amazon EC2™ service. It enables a utility computing service by allowing service providers to sell self-service virtual machine instances, storage volumes, and networking configurations over the Internet.
- The Cloud.com CloudStack platform enables enterprises to set up an on-premise private cloud for use by their own employees. The current generation of virtualization infrastructure shipped by VMware®, Citrix®, and Microsoft® targets enterprise IT departments who manage virtual machines in the same way as they would manage physical machines. The Cloud.com CloudStack platform, on the other hand, enables self service of virtual machines by users outside of IT departments.

The Cloud.com CloudStack platform includes the Management Server and extensions to industry-standard hypervisor software (E.g. XenServer®, vSphere, KVM) installed on Hosts running in the cloud. The Management Server is deployed on a farm of management servers. The administrator provisions resources (Hosts, storage devices, IP addresses, etc.) into the Management Server and the Management Server manages those resources. The Management server presents web interfaces to end users and administrators that enable them to take actions on some or all of their instances in the IaaS cloud.

1.1 Service Offerings, Disk Offerings, Network Offerings and Templates

The CloudStack platform allows the administrator to define Service Offerings and Disk Offerings. These allow the administrator to define the virtual hardware (CPU speed and count, RAM size, and disk size) that the user can select when creating a new instance.

The Network Offering is defined by the CloudStack. It describes the feature set that is available to end users from the virtual router or external networking devices.

The administrator can also provision templates in the system. Templates are the base OS images that the user can select when creating a new instance. For example, the CloudStack platform includes CentOS as a template. All popular Linux and Windows OS versions are supported.

1.2 Accounts, Users, and Domains

CloudStack platform users are assigned accounts. An account is typically a customer of the service provider or a department in a large organization. Accounts are the unit of isolation in the cloud. Accounts are grouped by domains. Domains usually contain accounts that have some logical relationship to each other and a set of delegated administrators with some authority over the domain and its subdomains. For example, a service provider with several resellers could create a domain for each reseller.

Multiple users can exist in an account. Users are like aliases in the account. Users in the same account are not isolated from each other. Most installations need not surface the notion of users; they just have one user per account.

1.3 Server Types

There are two required types of servers in the CloudStack platform: Management Servers and Hosts.

1.3.1 Management Server

The CloudStack Management Server runs in a Tomcat container and requires MySQL for persistence. The MySQL database required by the Management Server may optionally be placed on a separate system from the Management Server itself. This type of server is called a "Database Server". Replication is also supported.

The Management Server:

- Provides the web user interfaces for the administrator and end users.
- Provides the APIs for the CloudStack platform.
- Manages the assignment of guest VMs to particular Hosts.
- Manages the assignment of public and private IP addresses to particular accounts.
- Manages the allocation of storage to guests' virtual disk images.
- Manages snapshots, templates, and ISO images, possibly replicating them across data centers.
- Provides a single point of configuration for the cloud.

1.3.2 Hosts

Hosts are the resource in the cloud that host the guest virtual machines. For example, a Linux KVM-enabled server, a Citrix XenServer server, and an ESXi server are Hosts.

Hosts:

- Provide all the CPU, memory, storage, and networking resources needed to host the virtual machines.
- Interconnect using a high bandwidth TCP/IP network and connect to the Internet.
- May reside in multiple data centers across different geographic locations.
- May have different capacities (E.g. Different CPU speeds, different amounts of RAM, etc.).
- Are high-quality commodity hardware, and are reliable individually, but can fail frequently when a large number of servers are involved.

1.4 Networking Features and Virtualization

The CloudStack platform manages the allocation of private, direct, and public IP addresses. The administrator configures the system with available public, direct, and private IP addresses. There are two primary types of networks that can be created: "Direct Attached" and "Virtual".

The CloudStack refers to a Zone that allows virtual networking and direct attached, tagged networking as having "Advanced" networking. "Basic" networking refers to a Zone that allows only direct attached, untagged networking.

1.4.1 Direct Attached Networking

With Direct Attached networks the guest VMs are directly assigned IP addresses on the local subnet. They access the internet directly, and are not NATed by any components of the CloudStack. Their packets do not travel through a virtual router. Consequently they cannot take advantage of software load balancing, firewalling, and port forwarding features in the CloudStack.

Direct Attached guests may be isolated from other Direct Attached guests or not. With "Direct Tagged", the administrator assigns a specific Zone-wide VLAN ID and IP range to an account. Direct Tagged guests created by that account use that VLAN for guest-guest traffic and are isolated from other accounts' guests. Direct Attached guests receive their IP address from the virtual router, whose function is only for DHCP in this case. Direct Tagged is useful for linking guests with other systems, such as managed servers.

"Direct Untagged" provides isolation through the use of Amazon-style security groups and does not require VLANs. All guests are on the same shared broadcast domain, even if they are from different accounts. Direct Untagged is most useful for private clouds. Direct Untagged is available on all hypervisor types, but only XenServer and KVM nodes can use Security Groups.

1.4.2 Virtual Networking

With Virtual Networking, the guests of an account are allocated their own private, virtual network. An account's virtual network is isolated from the virtual networks of other accounts via a zone-wide VLAN. All guests in this network for an account get a network interface on this VLAN.

There are two types of virtual networking: virtual router based and external router based. The CloudStack provides a virtual router in its installation. This virtual router is capable of providing DNS, DHCP, gateway, NAT, load balancing, and VPN services. External router based virtual networking uses an external network device (e.g. Juniper SRX) to provide gateway and NAT services to the guests. DNS and DHCP continue to be provided by the virtual router with external element based networking.

A deployment that uses Virtual Networking must use either the virtual router or an external router.

Inter-guest traffic travels via a zone-wide VLAN and not through the virtual router. The use of VLANs provides isolation: the guests of different accounts are on different VLANs.

In virtual networking every account is given an initial public IP address. The user may acquire additional public IP addresses. Public IP addresses are routable from the Internet.

1.4.2.1 Virtual Networking with the Virtual Router

Every account is assigned a virtual router. All public IP addresses acquired by the account are assigned to the virtual router. The router is the gateway for guest VM traffic to and from the Internet, and it provides DNS and DHCP services to the guests. It also NATs all Internet traffic. The virtual router's presence in traffic enables the CloudStack platform to present several networking features to the end user. The virtual router can be configured by the user to forward traffic from a public IP to a particular guest VM. The port's traffic can also be load balanced across multiple guest VMs, providing for increased availability of a service behind the public IP address.

1.4.2.2 Virtual Networking with External Routers

The CloudStack is configured to use an external network element for the Zone. Every account is still assigned a virtual router. All public IP addresses acquired by the account are assigned to the external network element. The external router is the gateway for guest VM traffic to and from the Internet. It also NATs all Internet traffic. The virtual router provides DNS and DHCP services to the guests.

Load balancing via external elements is also possible.

1.4.3 Combining Virtual Networking and Direct Attached Networking

A single account may have guest VMs that have virtual networking and guest VMs that have tagged, direct attached networking. In this case there are two virtual routers for the account. One virtual router is responsible for the Zone VLAN used for the guests; the second virtual router is responsible for the tagged, direct attached VLAN assigned to the account.

Basic networking may not be combined with virtual networking or direct tagged networking in the same Zone. A cloud can have one zone with Direct Untagged and a second Zone with some combination of virtual network and direct tagged networking.

1.5 Storage Features and Virtualization

Templates define the base OS image that will be used when a guest is first booted. For example, a template might be 64-bit CentOS 5.3. Every template has a privacy level associated with it. Privacy levels include:

- **Public.** The template is available to all users.
- **Private.** The template is available to only its owner. A user can make his own private templates available to specific users.

Administrators and end users can add templates to the system. Users can see the template owner when viewing the template.

The CloudStack platform defines a volume as a unit of storage available to a guest VM. Volumes are either root disks or data disks. The root disk has "/" in the file system and is usually the boot device. Data disks provide for additional storage (E.g. As "/opt" or "D:"). Every guest VM has a root disk and a data disk. End users can mount multiple data disks to guest VMs. Users choose data disks from the disk offerings created by administrators. The user can create a template from a volume as well; this is the standard procedure for private template creation. Volumes are hypervisor-specific: a volume from one hypervisor type may not be used on a guest of another hypervisor type.

ISO images may be stored in the system and made available with a privacy level similar to templates. ISO images are classified as either bootable or not bootable. A bootable ISO image is one that contains an OS image (E.g. An Ubuntu 10.04 installation CD). The CloudStack platform allows a user to boot a guest VM off of an ISO image. Users can also attach ISO images to guest VMs. For example, this enables installing PV drivers into Windows. ISO images are not hypervisor-specific.

Snapshots may be taken for volumes, including both root and data disks. The administrator places a limit on the number of stored snapshots per user. Users can create new volumes from the snapshot for recovery of particular files and they can create templates from snapshots to boot from a restored disk. Snapshots may be set to occur on a recurring schedule. A completed snapshot is copied from primary storage to secondary storage, where it is stored until deleted or purged by newer snapshots.

The administrator provisions primary and secondary storage in the CloudStack platform. Both primary and secondary storage can be accessible via either iSCSI or NFS. Primary storage stores the guest VM virtual disk image. It is typically located close to the Hosts. Secondary storage stores the templates, ISO images, and snapshot data. There is usually one instance of secondary storage for hundreds of Hosts. The CloudStack platform manages the allocation of guest virtual disks to particular primary storage devices.

1.6 Administrator Controlled Allocation

The CloudStack platform chooses an available Host to create a new guest VM. The chosen Host will always be close to where the guest's virtual disk image is stored. Both vertical and horizontal allocation is allowed. Vertical allocation consumes all the resources of a given Host before allocating any guests on a second Host. This reduces power consumption in the cloud. Horizontal allocation places a guest on each Host in a round-robin fashion. This may yield better performance to the guests in some cases. The CloudStack

platform also allows an element of CPU over-provisioning as configured by the administrator. Over-provisioning allows the administrator to commit more CPU cycles to the allocated guests than are actually available from the hardware.

The CloudStack platform also provides a pluggable interface for adding new allocators. These custom allocators can provide any policy the administrator desires.

1.7 Guest VM Management

The CloudStack platform provides several guest management operations for end users and administrators. VMs may be stopped, started, rebooted, and destroyed.

Guests have a name and group. Guest names and groups are opaque to the CloudStack platform and are available for end users to organize their VMs. Each VM can have three names for use in different contexts. Only two of these names can be controlled by the user:

- Instance name – a unique, immutable ID that is generated by CloudStack and can not be modified by the user. This name conforms to the requirements in IETF RFC 1123.
- Display name – the name displayed in the CloudStack web UI. Can be set by the user. Defaults to instance name.
- Name – host name that the DHCP server assigns to the VM. Can be set by the user. Defaults to instance name.

Guests can be configured to be Highly Available (HA). An HA-enabled guest is monitored by the system. If the system detects that the guest is down, it will attempt to restart the guest, possibly on a different Host.

The CloudStack platform cannot distinguish a guest VM that was shut down by the user (E.g. Via the “shutdown” command in Linux) from a VM that exited unexpectedly. If an HA-enabled guest is shut down inside the VM, the CloudStack platform will restart it. The user must go through the CloudStack UI or API to shut down an HA-enabled guest.

1.8 Manageability

The system provides alerts and events to help with the management of the cloud. Alerts are notices to an administrator, generally delivered by e-mail, notifying the administrator that an error has occurred in the cloud. Alert behavior is configurable.

Events track all of the user and administrator actions in the cloud. For example, every guest VM start creates an associated event. Events are stored in the Management Server’s database.

The CloudStack platform allows administrators to place a Host into maintenance mode. When maintenance mode is activated the node is first removed from the pool of nodes available to receive new guest VMs. Then, the guest VMs currently running on the node are seamlessly migrated to another Host not in maintenance mode. This migration uses live migration technology and does not interrupt the execution of the guest. See "Scheduled Maintenance and Maintenance Mode" on page 60.

Host and guest performance monitoring is available to end users and administrators. This allows the user to monitor their utilization of resources and determine when it is appropriate to choose a more powerful service offering or larger disk.

1.9 API and Extensibility

The CloudStack platform end user and administrator web interfaces are built on the same HTTP query interface that is available for integration. This simple interface enables the creation of command line tools and new user interfaces to suit particular needs.

The CloudStack platform pluggable allocation architecture allows the creation of new types of allocators for the selection of storage and Hosts.

1.10 Scalability and Availability

The CloudStack platform has been designed to support tens of thousands of Hosts located in multiple data centers. Administrators define a Pod as the unit of scale. Typically a Pod would be a rack of hardware. Scaling out the cloud becomes the process of adding new Pods and provisioning the added resources with the Management Server.

The CloudStack platform has a number of features to increase the availability of the system. The Management Server itself may be deployed in a farm where the servers are load balanced. MySQL may be configured to use replication to provide for a manual failover in the event of database loss. For the Hosts, the CloudStack platform supports NIC bonding and the use of separate networks for storage as well as iSCSI Multipath.

2 Selecting Hardware and Software

The CloudStack platform has been designed to support a wide variety of hardware for Hosts, storage, and network devices. The following sections describe the requirements and in some cases provide statements on models that have been certified.

2.1 Hosts

For 64-bit x86 machines with processors supporting either AMD-V or Intel VT virtualization, extensions are required.

VMware provides a hardware compatibility list for vSphere at <http://www.vmware.com/resources/compatibility/search.php> for those customers using VMware vSphere as their hypervisor.

Citrix provides a hardware compatibility list for XenServer at <http://hcl.xensource.com/>, for those customers using the Citrix XenServer as their hypervisor.

RedHat provides a hardware compatibility list for RHEL at <https://hardware.redhat.com/>; however, it does not appear possible to do a search to constrain results to hardware that supports KVM.

Each machine should have at minimum 36 GB local disk storage and one or more Gigabit Ethernet (GbE) cards. We recommend 10 Gbps cards for best performance.

The CloudStack platform automatically detects the amount of CPU and memory resources provided by the Hosts.

2.2 Management Servers

The Management Server requires a 64-bit version of Linux. RHEL/CentOS 5.3 and later (including RHEL6) are supported. For the Community Edition, Ubuntu 10.04 and Fedora 14 are supported. The Management Server may be placed on a virtual machine.

A load balancer may be used to load balance traffic from the web and connections from the Hosts.

2.3 Storage

The CloudStack platform is designed to work with all standards-compliant iSCSI and NFS servers that are supported by the underlying hypervisor, including, for example:

- Dell EqualLogic™ for iSCSI
- Network Appliances filers for NFS and iSCSI
- Scale Computing for NFS

2.4 Network

The CloudStack platform is designed to work with all standards-compliant layer-2 and layer-3 network switches, including, for example:

- Cisco™ 3750-E or compatible Gigabit Ethernet switch
- Dell™ 6224 Gigabit Ethernet switch

2.5 Hypervisor Support

VMware vSphere 4.1, Citrix XenServer 5.6, Citrix XenServer 5.6 FP1 and RHEL 6 are supported on the Hosts for the commercial version of the CloudStack. The Community Edition additionally supports Ubuntu 10.04, RHEL/CentOS 5.5, and Fedora 14.

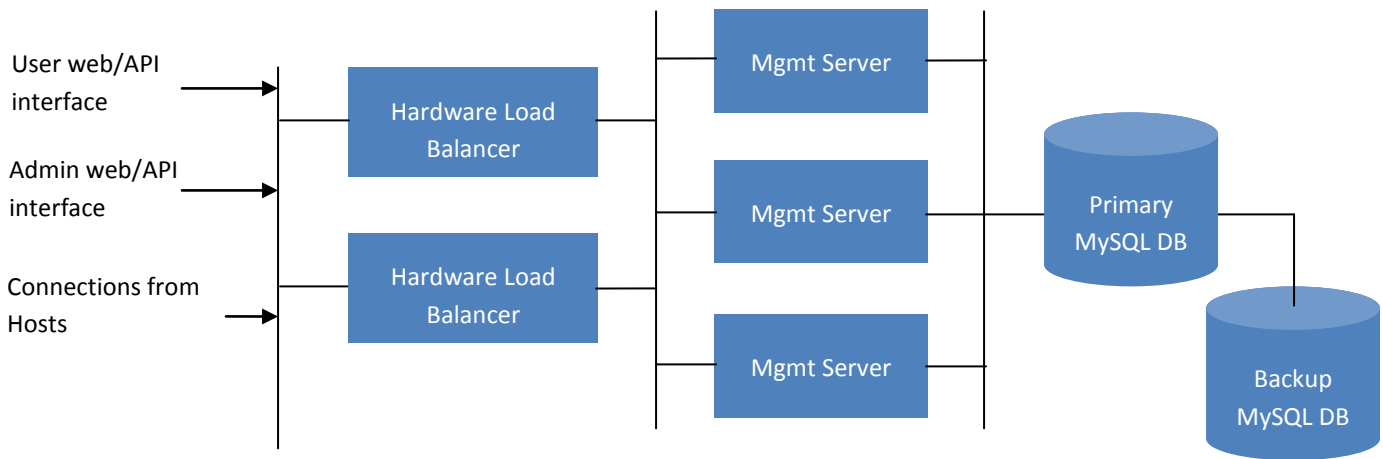
2.6 Guest OS and Software Support

The CloudStack platform works with all operating systems supported by the underlying hypervisor.

3 Planning a Deployment

3.1 Management Server Farm

The CloudStack Management Server is deployed on one or more front-end servers connected to a single MySQL database. Optionally a pair of hardware load balancers distributes requests from the web. A backup management server set may be deployed using MySQL replication at a remote site to add DR capabilities.



The administrator must decide the following.

- Whether or not load balancers will be used
- How many Management Servers will be deployed
- If MySQL replication will be deployed to enable disaster recovery.

3.2 Scaling Concepts

3.2.1 Hosts

Hosts are the basic physical scaling block of the CloudStack platform. Additional Hosts can be added at any time to provide more capacity for guest VMs.

Hosts are not visible to the end user. An end user cannot determine which Host their guest has been assigned to.

3.2.2 Clusters

Clusters are the second level of physical scaling in the CloudStack platform. A Cluster is a collection of Hosts that have access to shared Primary Storage and are of the same hypervisor type. Nodes in the same Cluster can live migrate instances to and from each other. Clusters are not visible to the end user. Size of the cluster is limited by the underlying hypervisor, although the CloudStack recommends less in most cases; see the Best Practices section in the Installation Guide.

Every VMware cluster is managed by a vCenter server. Administrator must register the vCenter server with CloudStack. There may be multiple vCenter servers per zone. Each vCenter server may manage multiple VMware clusters.

Hosts that are in the same Cluster are in the same subnet.

Clusters are still required for deployments that use local storage. There is just one Host per Cluster.

3.2.3 Pods

Pods are the third level of physical scaling in the CloudStack platform.

With shared Primary Storage a Pod is a collection of clusters. It may be exactly one Cluster establishing a 1:1 mapping between Cluster and Pod. Multiple Clusters per Pod are supported; currently CloudStack has been tested with up to two Clusters per Pod.

With local storage a Pod is a collection of Hosts. There are no practical limits to the number of Hosts in a Pod.

The Management Server is used to add and remove Hosts and primary storage from Clusters and Pods.

A Pod is frequently mapped to a single rack with a layer-2 switch. Hosts in the same Pod are in the same subnet.

Pods are not visible to the end user.

3.2.4 Availability Zones

Availability Zones are the fourth level of physical scaling in the CloudStack platform. An Availability Zone is a collection of Pods and secondary storage. An Availability Zone will include one or more layer-3 switches. The Availability Zone implies some form of physical isolation and redundancy (E.g. Separate power supply and network uplink) from other Availability Zones. It does not necessarily mean geographic distribution, and there may be one or more Availability Zones in a data center.

Availability Zones are visible to the end user. They must select an Availability Zone for their guest when started. They may also be required to copy their private templates to additional Availability Zones to enable creation of guest VMs in those zones from their templates.

Availability Zones may be public or private. Public availability zones are visible to all users. This means that any user may create a guest in that Zone. Private Availability Zones are reserved for a specific domain. Only users in that domain or its subdomains may create guests in that Zone.

Hosts in the same Availability Zone are directly accessible to each other without having to go through a firewall. Nodes in different Availability Zones can access each other through statically configured VPN tunnels.

The administrator must decide the following.

- How many Hosts to place in a Pod.
- How many primary storage servers to place in a Pod and total capacity for the storage servers.
- How many Pods to place in an Availability Zone.
- How many Clusters to have per Pod
- How much secondary storage to deploy in an Availability Zone.

3.3 Multi-Site Deployment

The CloudStack platform scales well into multiple sites through the use of Availability Zones. Figure 2 is an example of a multi-site deployment.

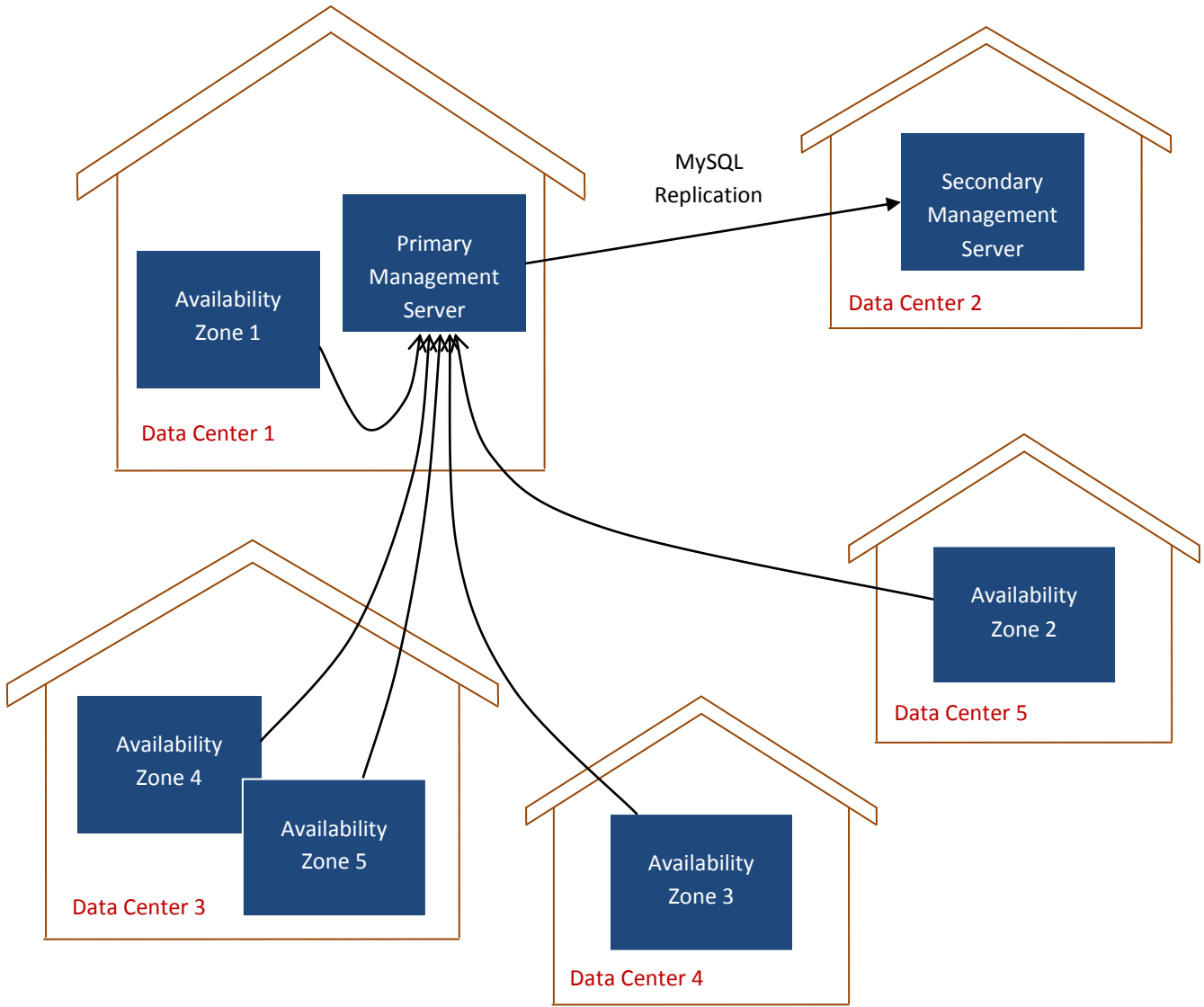


Figure 2 Example of a Multi-Site Deployment

Data Center 1 houses the primary Management Server as well as Availability Zone 1. The MySQL database is replicated in real time to the secondary Management Server installation in Data Center 2.

4 Defining Your Service Offering

The service offering defines the virtual hardware that the end users will be able to choose from. This includes CPU core count and speed, memory, and disk size. Here is an example of a service offering:

A virtual machine instance that is equivalent to a 1 GHz Intel® Core™ 2 CPU, with 1 GB memory at \$0.20/hour. Network traffic metered at \$0.10/GB.

The users expect that a service offering includes the following elements:

- CPU, memory, and network resource guarantees.
- How resources are metered.
- How the resource usage is charged.
- How often the charges are generated.

The CloudStack platform allows the administrator to configure the resource guarantee. It then emits usage records that the administrator can integrate with their billing system.

Service offerings cannot be changed once created.

A service offering that is no longer in use by any virtual machines can be permanently deleted.

A service offering that is still in use can be deleted by the administrator. However it will remain in the database until all the virtual machines referencing it have been deleted. After deletion by the administrator, a service offering will not be available to end users that are creating new instances.

The CloudStack platform separates service offerings into computing service offerings and storage service offerings. The computing service offering specifies:

- Guest CPU
- Guest RAM
- Guest Networking type (virtual or direct)
- Tags on the root disk

The disk offering specifies:

- Disk size (optional). An offering without a disk size will allow users to pick their own.
- Tags on the data disk

5 Understanding Network Types and Network Virtualization

In the CloudStack platform there are several types of networks, some real and some virtual. These include:

- **Guest Network.** The virtual network that the guest virtual machines connect to. It provides the isolation discussed previously.
- **Private Network.** The physical network that carries guest-guest traffic between Hosts when virtual networking is used.
- **Public Network.** The physical network that provides the guests with access to the Internet. This network also carries guest-guest traffic when Direct Attached networking is used.
- **Management Network.** The physical network that provides the link between the Management Servers, hypervisors, and storage devices.
- **Storage Network.** An optional physical network that provides the link between the hypervisors and storage devices.

There need not be a physical separation between these network types. For example, the CloudStack platform can run successfully on a single node installation that has a single NIC. Further, in all cases the private network and the management network are the same network.

Optionally, with the Enterprise and Service Provider Edition, a NIC may be dedicated to the public network. This can be used to isolate the public network traffic from the private network.

Optionally, a NIC may be dedicated to a separate Storage Network. This can be used to isolate storage traffic from other types of traffic. For example, a 1 Gbps NIC could be used for the private network while a 10 Gbps NIC is used for storage access.

See the Install Guide for instructions on procedures for these configurations.

Network virtualization is the process of creating a network for use by guest virtual machines. This virtual network has the characteristics of a LAN from the viewpoint of the guests. The resources used to create the virtual network may come from many sources and may rely on software-based components to a greater extent than is common in physical networks.

5.1 Guest Network

Each account that has a guest with a virtual network Service Offering is assigned a virtual network in its Availability Zone. A guest virtual network can be configured to any private address space. This document uses a Class A network in 10.0.0.0/8 private address space for its examples. The guest virtual network is an overlay network on top of the private network and is managed by the CloudStack platform.

A guest virtual network is valid within only one Availability Zone. Therefore virtual machines in different Availability Zones cannot communicate with each other using their IP addresses in the guest virtual network. Virtual machines in different Availability Zones must communicate with each other by routing through a public IP address.

Figure 1 illustrates a typical guest virtual network setup.

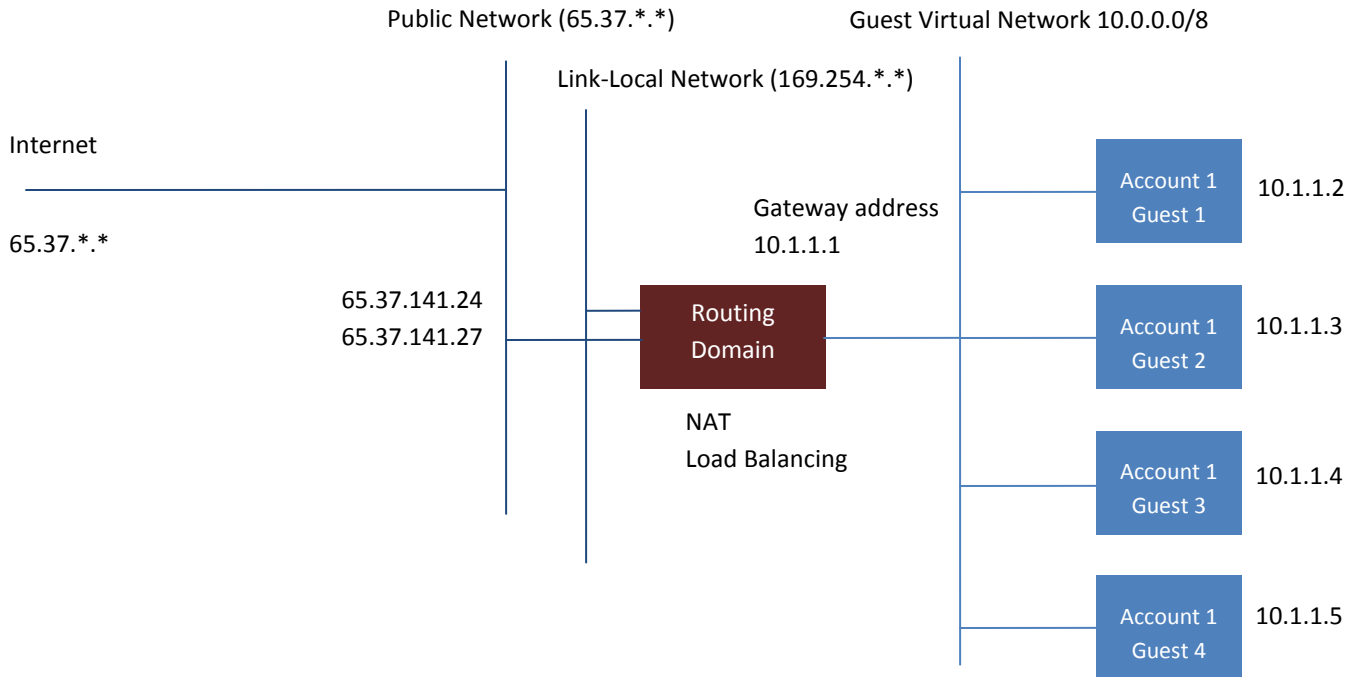


Figure 1 Guest Virtual Network Setup

The Management Server automatically creates a Virtual Router for each guest virtual network. A virtual router is a special virtual machine that runs on the Hosts. Each virtual router has three network interfaces. Its eth0 interface serves as the gateway for the guest virtual network and has the IP address of 10.1.1.1. Its eth1 interface resides on the link-local network and is used by the system to configure the virtual router. Its eth2 interface is assigned a public IP address on the public network.

The virtual router provides DHCP and will automatically assign an IP address for each guest VM in the 10.0.0.0/8 network. The user can manually reconfigure guest VMs to assume different IP addresses.

Source NAT is automatically configured in the virtual router to forward outbound traffic for all guest VMs.

5.2 Network Virtualization within One Pod

Figure 2 illustrates network setup within a single Pod. The Hosts are connected to a Pod-level switch. At a minimum the Hosts should have one physical uplink to each switch. Bonded NICs are supported as well. The Pod-level switch is a pair of redundant gigabit switches with 10 G uplinks.

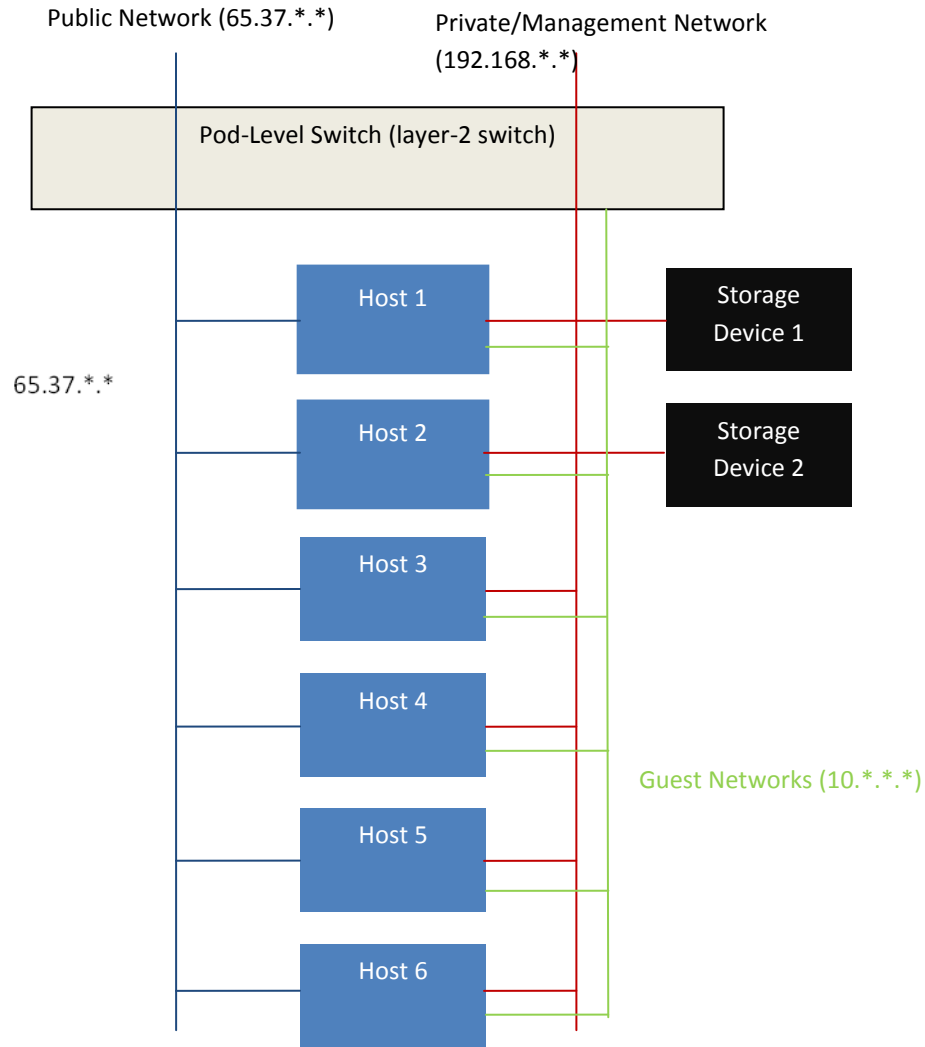


Figure 2 Network Setup within a Single Pod – Logical View

Servers are connected to the private/management and public networks as follows:

- Storage devices are connected to only the private/management network.
- Hosts are connected to both the management network and the public network.
- The Hosts are connected to one or more guest networks.

We recommend the use of multiple physical Ethernet cards to implement each network interface as well as redundant switch fabric in order to maximize throughput and improve reliability.

5.3 Network Virtualization within One Availability Zone

Figure 3 illustrates the network setup within a single Availability Zone.

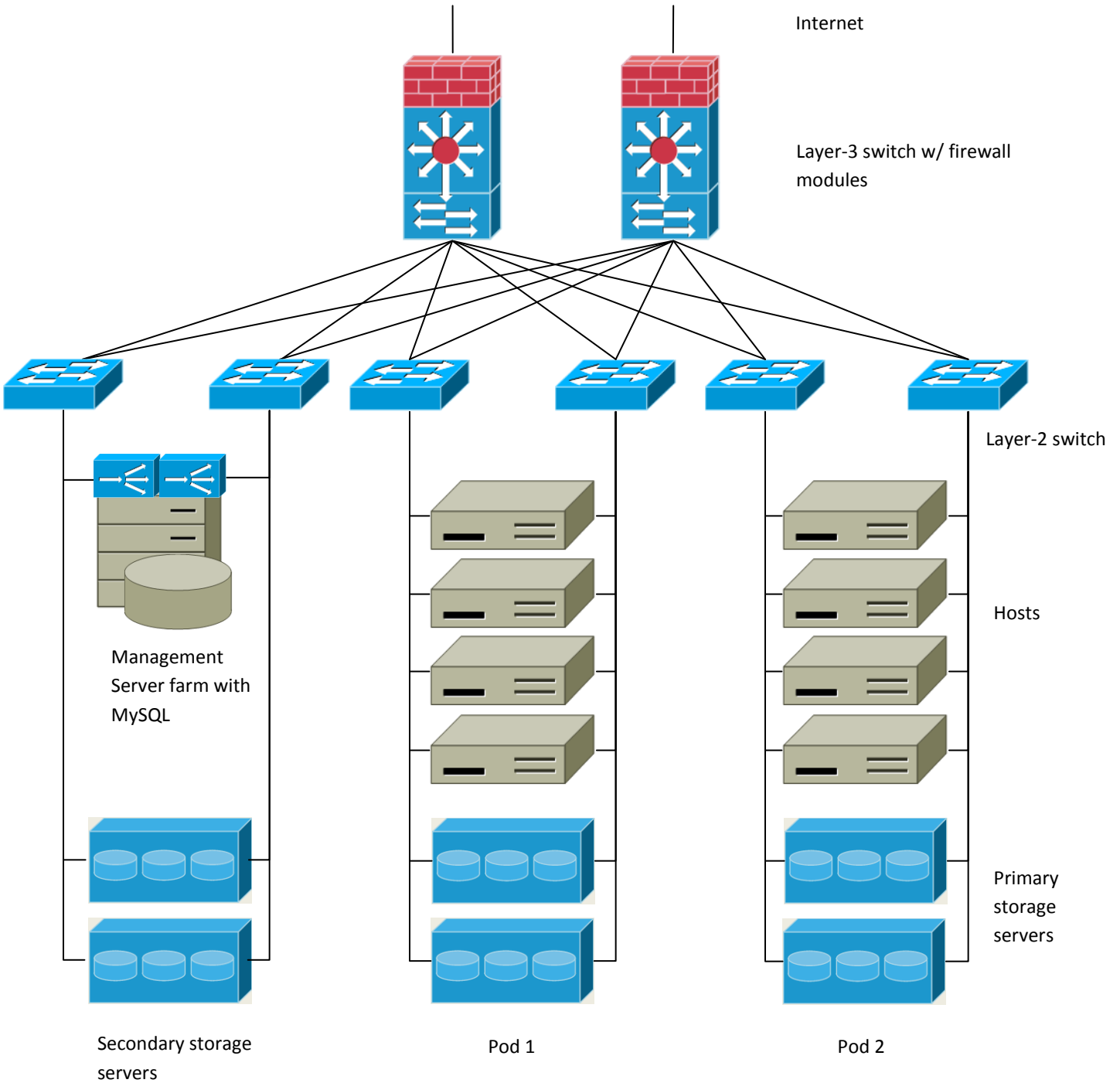


Figure 3 Network Setup within a Single Availability Zone

The private/management network carries traffic in the guest virtual networks.

A firewall for the private/management network operates in the NAT mode. The private/management network typically is assigned IP addresses in the 192.168.0.0/16 Class B private address space. Each Pod is assigned IP addresses in the 192.168.*.0/24 Class C private address space.

Each Availability Zone has its own set of public IP addresses. Public IP addresses from different Availability Zones do not overlap.

The private/management network addresses must be unique across the cloud. You cannot, for example, have a Host in one Availability Zone which has the same private IP address as a Host in another Availability Zone.

5.4 Network Virtualization

Network virtualization uses tagged VLANs to provide isolation between guest virtual networks. With network virtualization there are three types of VLANs that are present within every Availability Zone.

- **The Private/Management VLAN.** This is for the private/management network as defined above.
- **The Public VLANs.** These are for the public network as defined above.
- **The Zone VLANs.** There is one tagged VLAN per guest virtual network with active instances in a Zone..

There are 4094 available VLANs according to the 802.1q standard. The administrator should determine a segmentation of the VLAN namespace that matches their requirements. Here is an example of such a segmentation:

VLAN IDs	Use
< 100	Reserved for administrative purposes
100-199	Public VLANs
200-499	Untagged Private/Management VLANs
500-1999	Zone VLANs
> 2000	Reserved for future use

5.5 Private Address Allocation

The CloudStack platform allocates a private IP address to each system VM that is not a virtual router. Virtual routers only communicate with the hypervisor and use link-local IP addresses. The administrator is responsible for allocating private IP addresses to the hypervisors. The administrator configures the CloudStack platform with the range of IP addresses available for private IP address allocation.

5.6 Public Address Allocation

Each virtual router is assigned at least one public IP addresses. The user may request additional public IP addresses.

The administrator configures the available public IP address pools on a per-Availability Zone basis. Distinct public IP ranges can be added as separate VLANs incrementally. Each public IP range can be used by any Pod inside the same Availability Zone.

5.7 External Network Elements

The CloudStack is capable of replacing the Virtual Router with an external Juniper SRX device and an external F5 load balancer for gateway and load balancing services. This is available only when virtual networking is in use. In this case the VMs in virtual networking and direct tagged networking use the SRX as their gateway.

5.7.1 Initial Setup

When the first VM is created for a new account the CloudStack programs these network elements to work with the VM. This initial setup is performed as part of creating the virtual network for the user. It is done once.

The following objects are created on the firewall:

- A new logical interface to connect to the account's private VLAN. The interface IP is always the first IP of the account's private subnet (e.g. 10.1.1.1).
- A source NAT rule that forwards all outgoing traffic from the account's private VLAN to the public Internet, using the account's public IP address as the source address.
- A firewall filter counter that measures the number of bytes of outgoing traffic for the account.

The following objects are created on the load balancer:

- A new VLAN that matches the account's provisioned Zone VLAN.
- A self IP for the VLAN. This is always the second IP of the account's private subnet (e.g. 10.1.1.2).

5.7.2 Additional Configuration

Additional user actions (e.g. setting a port forward) will cause further programming of the firewall and load balancer. A user may request additional public IP addresses and forward traffic received at these IPs to specific VMs. This is accomplished by enabling static NAT for a public IP address, assigning the IP to a VM, and specifying a set of protocols and port ranges to open. When a static NAT rule is created, CloudStack programs the Zone's external firewall with the following objects:

- A static NAT rule that maps the public IP address to the private IP address of a VM.
- A security policy that allows traffic within the set of protocols and port ranges that are specified.
- A firewall filter counter that measures the number of bytes of incoming traffic to the public IP.

A user may also create load balancing rules that balance traffic received at a public IP to one or more VMs. When a user creates a rule, specifies an algorithm, and assigns the rule to a set of VMs, CloudStack programs the zone's external load balancer with the following objects:

- A virtual server that listens for traffic at the specified public IP and measures the number of incoming and outgoing bytes.
- A load balancing pool that redirects traffic to VMs with the specified algorithm.
- Pool members and nodes for each VM that the rule is assigned to.

The number of incoming and outgoing bytes through source NAT, static NAT, and load balancing rules is measured and saved on each external element. This data is collected on a regular basis and stored in the CloudStack database.

6 Network Virtualization Features

The CloudStack platform provides network virtualization. Network virtualization allows the guests to communicate with each other using shared infrastructure with the security and user perception that the guests have a private LAN.

The virtual router is the linchpin of the networking features. The Management Server programs the Virtual Router over the management network. The Virtual Router is then able to implement the following features for the guest network.

6.1 Guest Virtual Networks

The IP ranges of the guest virtual networks are set on a per-account basis by the user. This allows the users to configure their network in a fashion that will enable VPN linking between their guest network and their clients.

6.2 IP Forwarding and Firewalling

By default all incoming traffic to the public IP address is rejected. All outgoing traffic from the guests is translated via NAT to the public IP address and is allowed. Users may enable port forwarding for specific public IP addresses, public port to guest IP addresses and guest port.

6.3 IP Load Balancing

The user may choose to associate the same public IP for multiple guests. The system implements a TCP-level load balancer with the following policies.

- Round-robin
- Least connection
- Source IP

This is similar to port forwarding but the destination may be multiple IP addresses.

6.4 Port Forwarding

A port forward service is a set of port forwarding rules that define a policy. A port forward service is then applied to one or more guest VMs. The guest VM then has its inbound network access managed according to the policy defined by the port forwarding service.

A guest VM can be in any number of port forward services.

Port forward services can be defined but have no members.

6.5 DNS and DHCP

The Virtual Router provides DNS and DHCP services to the guests. It proxies DNS requests to the DNS server configured on the Availability Zone.

6.6 VPN

The CloudStack provides a L2TP-based VPN service to guest virtual networks. Clients native to Windows and Mac OS X may be used to connect to the guest network. The user is responsible for creating and managing users for their VPN. The CloudStack does not allow its authentication database to be used for this purpose.

Users need to make sure that not all traffic goes through the VPN. That is, the route installed by the VPN should be only for the guest network and not for all traffic.

6.6.1 Mac OS X

In Mac OS X, in Connect Menu -> Options the user should make sure that the checkbox "Send all traffic over VPN connection" is not checked.

6.6.2 Windows

The procedure to effect this varies by Windows version. Generally the user will need to edit the VPN properties and make sure that the default route is not the VPN.

6.7 Working with Additional Networks

In Advanced Networking Zones the CloudStack allows additional networks to be provisioned. Additional networks are always Direct Tagged networks. These networks are associated with a VLAN and subnet parameters such as Gateway, Netmask, and IP Range. The IP range is the set of IPs that the CloudStack will manage and assign to guests in this network. The IP range can be either public IP addresses or RFC 1918 addresses.

Additional networks may be added at any time to the CloudStack after the initial installation.

A VM's networks are defined at VM creation time. A VM cannot add or remove networks after it has been created, although the user can go into the guest and remove the IP address from the NIC on a particular network.

6.7.1 Network Scope

Additional networks can either be available to all accounts or be assigned to a specific account. Networks that are available to all accounts are Zone-wide. Any user with access to the Zone can create a VM with access to that network. These Zone-wide networks provide little or no isolation between guests.

Networks that are assigned to a specific account provide strong isolation.

6.7.2 Default and Non-default Networks

Networks can be Default networks or non-default networks. VMs have exactly one Default network. The virtual router's DHCP reply will set the guest's default gateway as that for the default network.

Multiple non-default networks may be added to a guest in addition to the single, required default network.

The administrator controls which networks are available as the default network. The CloudStack has a default virtual network offering that determines the user availability of a virtual network. The virtual network may be specified as required, optional, or unavailable. If required, the user must choose the virtual network as the default network for the guest. If optional, the user may choose the virtual network as the default network, but may also choose a direct tagged network as the default network. If

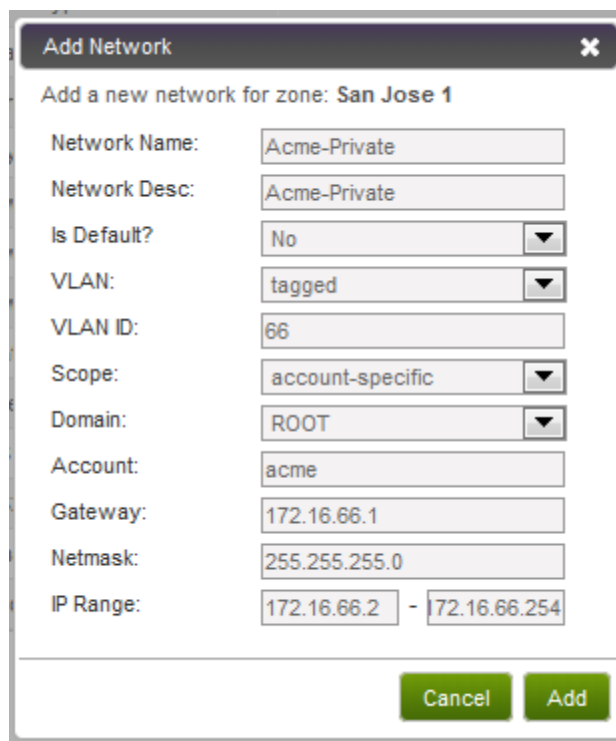
unavailable then the user will not have the virtual network available to them, and they must choose a direct tagged network as their default network.

To set the availability of the virtual network you should edit the "DefaultVirtualizedNetwork" Network Offering. Choose Configuration -> Network Offering -> DefaultVirtualizedNetwork. Then choose edit in the Actions drop down. Set the availability field to your choice.

The additional direct tagged networks you choose to add must be classified as either default or non-default.

6.7.3 Adding an Additional Network

To add a new network to the CloudStack, go to System -> Physical Resources -> { target zone } -> Network. Then click on "Add Network" in the top menu bar. You will see this dialog:



This dialog requires the following information:

- **Network Name.** The name of the network. This will be user-visible.
- **Network Desc:** The description of the network. This will be user-visible.
- **Is Default?:** Choose yes if it is a default network, choose no if not.
- **VLAN:** This is always tagged.
- **VLAN ID:** The VLAN tag for this network.
- **Scope:** Choose account-specific if you would like to limit a single account to accessing this VLAN. Choose zone-wide if all accounts with access to the Zone should be able to access the network.
- **Domain/Account:** Enter the domain and account name for the account, if account-specific was chosen for scope.
- **Gateway:** Enter the gateway for this network.

- **Netmask:** Enter the netmask for this network.
- **IP Range:** Enter the first and last IP addresses that define a range that the CloudStack can assign to guests.

With this information entered click "Add". The network will now be available to newly created guests.



7 Storage Features and Types

The CloudStack platform defines two types of storage: primary and secondary. Primary storage can be accessed by either iSCSI or NFS. Additionally, direct attached storage may be used for primary storage. Secondary storage is always accessed using NFS.

A site’s policies and administrative preferences combined with the advantages and disadvantages of each access protocol should be considered in deciding between NFS, iSCSI, and direct attached storage for primary storage.

In contrast to some other cloud offerings, there is no ephemeral storage in the CloudStack platform. All volumes on all nodes are persistent.

7.1 Primary Storage

Primary Storage is used for storing the guest VM root disks as well as additional data disk volumes. Shared primary storage (all types except local storage) is registered with the Cluster of Hosts. Root volumes are created automatically when a virtual machine is created. Root volumes are deleted when the VM is destroyed. Data volumes can be created and dynamically attached to VMs. Data volumes are not deleted when VMs are destroyed.

The speed of primary storage will impact guest performance. If possible administrators should choose smaller, higher RPM drives for primary storage.

Primary Storage can be added at any time via the administrator UI. Administrators should monitor the capacity of primary storage devices and add additional primary storage as needed. Please see the installation guide for instructions to add primary storage to a cluster.

Administrators add primary storage to the system by creating a CloudStack storage pool. Each storage pool is associated with a cluster. The following table discusses storage options and parameters for different hypervisors.

	VMware vSphere	Citrix XenServer	KVM
Format for Disks, Templates, and Snapshots	VMDK	VHD	QCOW2
iSCSI support	VMFS	Clustered LVM	Yes, via Shared Mountpoint
Fiber Channel support	VMFS	Yes, via Existing SR	Yes, via Shared Mountpoint
NFS support	Y	Y	Y
Local storage support	Y	Y	N
Storage over-provisioning	NFS and iSCSI	NFS	NFS

XenServer uses a clustered LVM system to store VM images on iSCSI and Fiber Channel volumes and does not support over-provisioning in the hypervisor. The storage server itself, however, can support thin-provisioning. As a result the CloudStack can still support storage over-provisioning by running on thin-provisioned storage volumes.

vSphere uses VMFS to store VM images on iSCSI and Fiber Channel volumes and does not support over-provisioning in the hypervisor. The storage server itself, however, can support thin-provisioning. As a result the CloudStack can still support storage over-provisioning by running on thin-provisioned storage volumes.

KVM supports "Shared Mountpoint" storage. A shared mountpoint is a file system path local to each server in a given Cluster. The path must be the same across all Hosts in the cluster, for example /mnt/primary1. This shared mountpoint is assumed to be a clustered filesystem such as OCFS2. In this case the CloudStack does not attempt to mount or unmount the storage as is done with NFS. The CloudStack requires that the administrator insure that the storage is available.

With NFS storage the CloudStack manages the overprovisioning. In this case the global configuration parameter `storage.overprovisioning.factor` controls the degree of overprovisioning. This is independent of hypervisor type.

Local storage is an option for primary storage for vSphere and XenServer. When the local disk option is enabled, a local disk storage pool is automatically created on each host. To use local storage for the System Virtual Machines (E.g. the Virtual Router), set `system.vm.use.local.storage` to true in global configuration.

The CloudStack supports multiple Primary Storage pools in a Cluster. For example, you could provision 2 NFS servers in primary storage. Or you could provision 1 iSCSI LUN initially and then add a second iSCSI LUN when the first approaches capacity.

7.1.1 Tags

Storage may be "tagged". A tag is a text string attribute associated with Primary Storage, a Disk Offering, or a Service Offering. Tags allow administrators to provide additional information about the storage. For example, that is a "SSD" or it is "slow". Tags are not interpreted by the CloudStack. They are matched against tags placed on service and disk offerings. CloudStack requires all tags on service and disk offerings to exist on the primary storage before it allocates root or data disks on the primary storage. Service and disk offering tags are used to identify the requirements of the storage that those offerings have. For example, the high end service offering may require "fast" for its root disk volume.

The interaction between tags, allocation, and volume copying across Clusters and Pods can be complex. To simplify the situation, use the same set of tags on the Primary Storage for all Clusters in a Pod. Even if different devices are used to present those tags, the set of exposed tags can be the same.

7.1.2 Maintenance Mode

Primary storage may be placed into maintenance mode. This is useful, for example, to replace faulty RAM in a storage device. Maintenance mode for a storage device will first stop any new guests from being provisioned on the storage device. Then it will stop all guests that have any volume on that storage device. When all such guests are stopped the storage device is in maintenance mode and may be shut down. When the storage device is online again you may cancel maintenance mode for the device. The CloudStack will bring the device back online and attempt to start all guests that were running at the time of the entry into maintenance mode.

7.2 Secondary Storage

Secondary Storage is used for storing templates, saved snapshots of guest VMs, and ISO images. The secondary storage has a high read:write ratio and is expected to consist of larger drives with lower IOPS than the primary store. The secondary storage device must be located in the same Availability Zone as the guest VMs it serves. The secondary storage device must serve NFS.

There must be exactly one secondary storage device per Availability Zone.

The Secondary Storage server must be available to all Hosts in the Zone. Additionally, the Secondary Storage VM mounts and writes to Secondary Storage.

Submissions to secondary storage go through the Secondary Storage VM. The Secondary Storage VM can retrieve templates and ISO images from URLs using a variety of protocols.

7.3 Changing the Secondary Storage IP Address

You can change the Secondary Storage IP address after it has been provisioned. After changing the IP address on the host, login to your management server and execute:

```
# mysql
mysql> use cloud;

mysql> select id from host where type = 'SecondaryStorage';

mysql> update host_details set value = 'nfs://192.168.160.20/export/mike-ss1'
where host_id = # and name = 'orig.url';

    - replace # with the id of the secondary server

mysql> update host set name = 'nfs://192.168.160.20/export/mike-ss1' where type
= 'SecondaryStorage';

mysql> update host set url = 'nfs://192.168.160.20/export/mike-ss1' where type
= 'SecondaryStorage';

mysql> update host set guid = 'nfs://192.168.160.20/export/mike-ss1' where type
= 'SecondaryStorage';
```

In the above example, change the URL to use the appropriate IP address and path for your server. Then login to the cloud console UI and stop and start (not reboot) the Secondary Storage VM for that Zone.

7.4 Changing Secondary Storage Servers

You can change the Secondary Storage NFS mount. Perform the following steps to do so:

1. Stop all running Management Servers
2. Wait 30 minutes. This allows any writes to secondary storage to complete.
3. Copy all files from the old secondary storage mount to the new.

4. Use the procedure above to change the IP address for Secondary Storage if required.
5. Start the Management Server(s).

7.5 Working with Volumes

A volume provides storage to a guest VM. The volume can provide for a root disk or an additional data disk. The CloudStack platform supports additional volumes for guest VMs.

Volumes are created for a specific hypervisor type. A volume that has been attached to guest using one hypervisor type (e.g, XenServer) may not be attached to a guest that is using another hypervisor type (e.g. vSphere, KVM). This is because the different hypervisors use different disk image formats.

The creation of data disk volumes is deferred until they are first attached to a guest. The creation of a data disk volume will create an entry in the CloudStack database, but no physical operation is performed on any storage device. This optimization allows the CloudStack to provision the volume nearest to the guest that will use it when the first attachment is made.

7.5.1 Moving Volumes

Volumes may be detached from a guest and attached to another guest. If the two guests are in the same Cluster, this is straightforward. If the two guests are in different Clusters, the volume will first be copied to secondary storage and then copied from secondary storage to the destination cluster. For a large volume, this may take several minutes.

7.5.2 Resizing Volumes

CloudStack does not provide the ability to resize root disks or data disks; the disk size is fixed based on the template used to create the VM. However, the tool [VHD Resizer](http://vmtoolkit.com/files/folders/converters/entry87.aspx) (<http://vmtoolkit.com/files/folders/converters/entry87.aspx>), while not officially supported by Cloud.com or Citrix, might provide a workaround. To increase disk size with VHD Resizer:

1. Get the VHD from the secondary storage.
2. Import it into VHD Resizer.
3. Resize the VHD.
4. Upload the new VHD.
5. Create a new VM.
6. On a Linux guest, extend the file system to reflect the new disk size. Manually resize your partitions using the operating system's utilities. For example, use `resize2fs` or use LVM utilities to add partition space.
7. Take a snapshot, then create a new template from that snapshot.

For more information, see [How to Resize a Provisioning Server 5 Virtual Disk](http://support.citrix.com/article/CTX118608) at the Citrix Knowledge Center (<http://support.citrix.com/article/CTX118608>).

7.5.3 Volume Deletion and Garbage Collection

The deletion of a volume does not delete the snapshots that have been created from the volume.

When a VM is destroyed, data disk volumes that are attached to the VM are not deleted.

Volumes are permanently destroyed using a garbage collection process. The global configuration variables `expunge.delay` and `expunge.interval` determine when the physical deletion of volumes will occur.

- `expunge.delay`: determines how old the volume must be before it is destroyed, in seconds
- `expunge.interval`: determines how often to run the garbage collection check

Administrators should adjust these values depending on site policies around data retention.

7.6 Working with ISO Images

The CloudStack platform supports ISOs and their attachment to guest VMs. An ISO is a read-only file that has an ISO/CD-ROM style file system. Users can upload their own ISOs and mount them on their guest VMs.

ISOs are uploaded based on a URL. HTTP is the supported protocol. Once the ISO is available via HTTP specify an upload URL such as `http://my.web.server/filename.iso`.

ISOs may be public or private, like templates.

ISOs are not hypervisor-specific. That is, a guest on vSphere can mount the exact same image that a guest on KVM can mount.

7.7 Working with Blank VMs

Users can create blank virtual machines. A blank virtual machine is a virtual machine without an OS template. Users can attach an ISO file and install the OS from the CD/DVD-ROM.

7.8 Working with Templates

A template is a virtual disk image that can be used to instantiate a new virtual machine. Templates may be of a variety of operating systems as described later in this document. The administrator and the template creator (the end user) can set different levels of access control on templates.

When templates are deleted, the VMs instantiated from them will continue to run. However, new VMs cannot be created based on the deleted template.

Templates are hypervisor specific. That is, a given template will be associated with a particular hypervisor type. The hypervisor type must be specified on template upload.

7.8.1 The Default Template

The CloudStack platform includes a CentOS template. This template is downloaded by the Secondary Storage VM after the primary and secondary storage are configured. You can use this template in your production deployment or you can delete it and use custom templates.

The root password for the default template is "password".

A default template is provided for each of XenServer, KVM, and vSphere. The templates that are downloaded depend on the hypervisor type that is available in your cloud. Each template is approximately 2.5 GB physical size.

The default template includes the standard iptables rules, which will block most access to the template excluding ssh.

```
# iptables --list
Chain INPUT (policy ACCEPT)
target     prot opt source                destination
RH-Firewall-1-INPUT  all  --  anywhere              anywhere

Chain FORWARD (policy ACCEPT)
target     prot opt source                destination
RH-Firewall-1-INPUT  all  --  anywhere              anywhere

Chain OUTPUT (policy ACCEPT)
target     prot opt source                destination

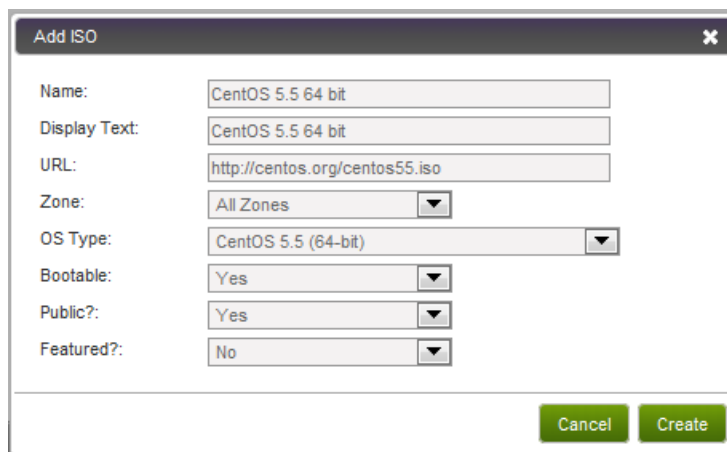
Chain RH-Firewall-1-INPUT (2 references)
target     prot opt source                destination
ACCEPT     all  --  anywhere              anywhere
ACCEPT     icmp --  anywhere              anywhere    icmp any
ACCEPT     esp  --  anywhere              anywhere
ACCEPT     ah   --  anywhere              anywhere
ACCEPT     udp  --  anywhere              224.0.0.251    udp dpt:mdns
ACCEPT     udp  --  anywhere              anywhere       udp dpt:ipp
ACCEPT     tcp  --  anywhere              anywhere       tcp dpt:ipp
ACCEPT     all  --  anywhere              anywhere       state RELATED,ESTABLISHED
ACCEPT     tcp  --  anywhere              anywhere       state NEW tcp dpt:ssh
REJECT     all  --  anywhere              anywhere       reject-with icmp-host-prohibited
```

7.8.2 Creating Templates

Templates can be created from either volumes or ISO images. The procedure to create a template from a volume is available in the web UI.

The procedure to create a template from an ISO image is as follows:

1. Log into the UI as either an end user or administrator.
2. Click on the Templates tab, and go to the ISO section.
3. Click Add ISO in the top menu bar.



The screenshot shows a dialog box titled "Add ISO" with the following fields and values:

- Name: CentOS 5.5 64 bit
- Display Text: CentOS 5.5 64 bit
- URL: http://centos.org/centos55.iso
- Zone: All Zones
- OS Type: CentOS 5.5 (64-bit)
- Bootable: Yes
- Public?: Yes
- Featured?: No

Buttons: Cancel, Create

You must provide:

- **Name.** Short name for the ISO image. (E.g. Ubuntu 9.10)
 - **Display Text.** Description of the ISO image. (E.g. Ubuntu 9.10 Desktop i386 32 bit)
 - **URL.** The URL that hosts the ISO image. The Management Server must be able to access this location via HTTP. If needed you can place the ISO image directly on the Management Server.
 - **Zone:** The Zone for the ISO to be available in.
 - **OS Type:** Select the OS type of the ISO image. Choose other if the OS Type of the ISO is not listed or if the ISO is not bootable.
 - **Bootable:** Whether or not a guest could boot off this ISO image. For example, a CentOS ISO is bootable, a Microsoft Office ISO is not bootable.
 - **Public:** Choose Yes if this ISO should be available to other users.
 - **Featured:** Choose Yes if you would like this ISO to be more prominent for users to select. Only administrators may make ISOs featured.
4. Click Create.

The Management Server will download the ISO. Depending on the size of the ISO, this may take a long time. The ISO status column will display Ready once it has been successfully downloaded into the secondary storage. Clicking Refresh updates the download percentage.

Important: Do not continue to the next step until the ISO has finished downloading.

5. Go to the Instance section, My Instances and click on Add Instance. In the wizard, click ISO Boot and select the ISO file. A VM will be created and booted from that ISO file.
6. Make any desired configuration changes on the running VM and then stop it.
7. Once stopped, in the Instances section, find the VM. Select it and click on the Volumes tab. Find the root disk volume and in the Actions drop down menu select "Take Snapshot". This will create a snapshot from the volume that will be the basis for the template. Wait for the "Taking Snapshot" procedure to complete before proceeding to step 8.
8. Find the snapshot you have taken in the Storage -> Snapshots section. Select it and in the Actions dropdown for it choose "Create Template".

The new template will be visible in the Templates section when the template creation process has been completed. The template is then available when creating a new VM.

7.8.3 Uploading Templates

Templates are uploaded based on a URL. HTTP is the supported access protocol. The Management Server will download the file from the specified URL, such as <http://my.web.server/filename.vhd.gz>.

The operating system type should be provided when uploading a template. This helps the CloudStack platform and hypervisor perform certain operations and make assumptions that improve the performance of the guest. If the operating system present on the template is not available you should select Other.

Note: Generally you should not choose an older version of the OS that you have. For example, choosing CentOS 5.3 to support a CentOS 5.4 image will in general not work. In those cases you should choose Other.

“Password Enabled” refers to whether or not your template has the CloudStack platform password change script installed. This was discussed previously.

Templates are frequently large files. You can optionally gzip them to decrease upload times.

If you are uploading a template that was created using vSphere Client, be sure the OVA file does not contain an ISO. If it does, the deployment of VMs from the template will fail.

7.8.4 Extracting Templates

End users and Administrators may extract templates from the CloudStack. Navigate to the template in the UI and choose the Download function from the Actions menu.

7.8.5 Public Templates

Public templates are available to all users in all accounts. All users can create virtual machines from these templates.

When a user publishes a template as “public”, the template is available to all users in all domains.

7.8.6 Private Templates

Private templates are only available to the user who created them. By default an uploaded template is private.

Users can create virtual machines from their collection of private templates the same way they create virtual machines from public templates.

7.8.7 Deleting Templates

Templates may be deleted. In general, when a template spans multiple Zones, only the copy that is selected for deletion will be deleted; the same template in other Zones will not be deleted. The provided CentOS template is an exception to this. If the provided CentOS template is deleted it will be deleted from all Zones.

7.8.8 Running Sysprep for Windows Templates

Windows templates must be prepared before they can be provisioned on multiple machines. You first need to upload your Windows ISO and create a VM Instance with this ISO. After you have created your VM with Windows installed, follow these next steps to run sysprep on your VM. Sysprep allows you to create a generic Windows template and avoid any possible SID conflicts.

7.8.8.1 Sysprep for Windows Server 2008 R2

For Windows 2008 R2, you run Windows System Image Manager to create a custom sysprep response XML file. Windows System Image Manager is installed as part of the Windows Automated Installation Kit (AIK). Windows AIK can be downloaded from the Microsoft Download Center at the following location:

<http://www.microsoft.com/downloads/details.aspx?FamilyID=94bb6e34-d890-4932-81a5-5b50c657de08&DisplayLang=en>

Use the following steps to run sysprep for Windows 2008 R2.¹

1. Download and install the Windows AIK.

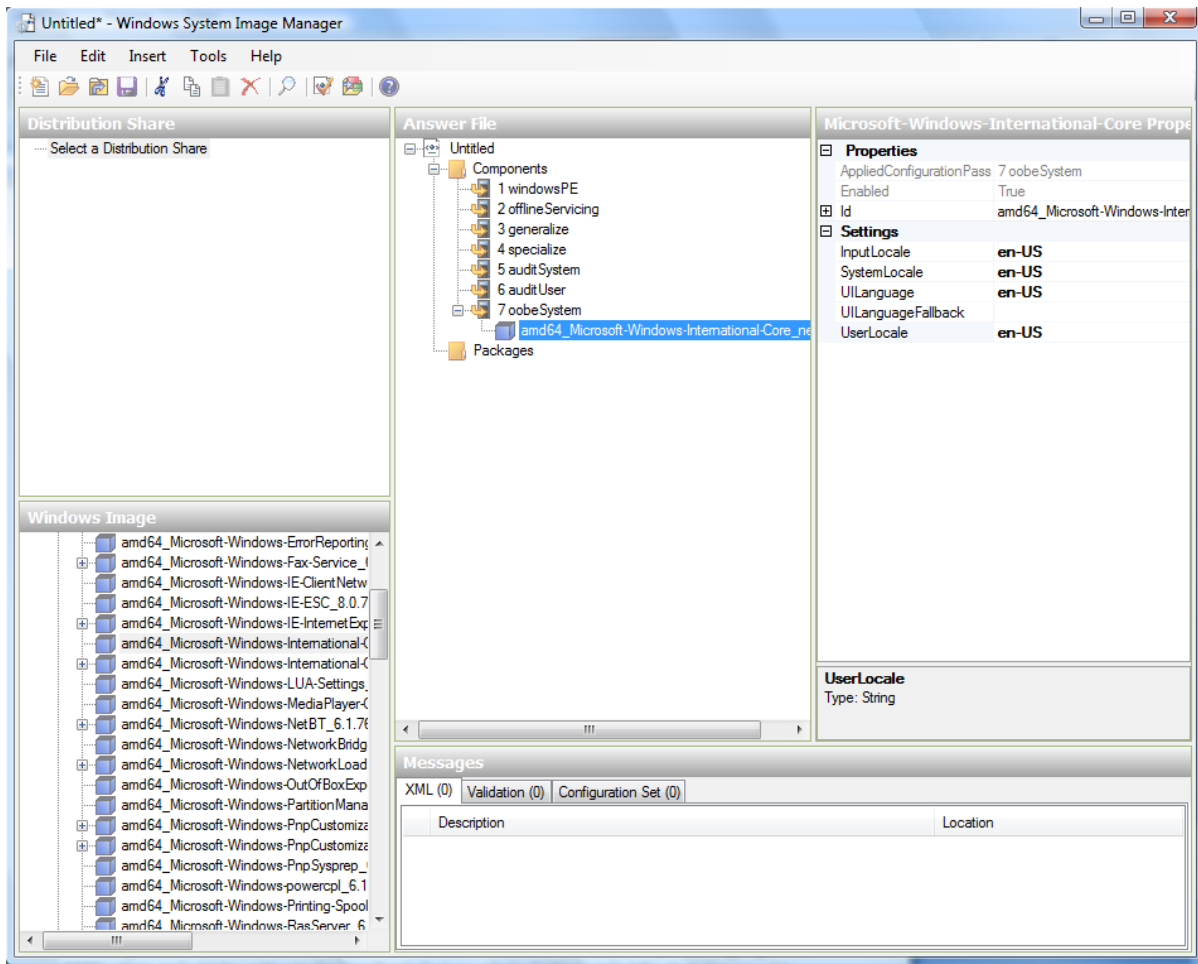
Note: Windows AIK should not be installed on the Windows 2008 R2 VM you just created. Windows AIK should not be part of the template you create. It is only used to create the sysprep answer file.

2. Copy the install.wim file in the \sources directory of the Windows 2008 R2 installation DVD to the hard disk. This is a very large file and may take a long time to copy. Windows AIK requires the WIM file to be writable.
3. Start the Windows System Image Manager, which is part of the Windows AIK.
4. In the Windows Image pane, right click “Select a Windows image or catalog file” to load the install.wim file you just copied.
5. Select the Windows 2008 R2 Edition.

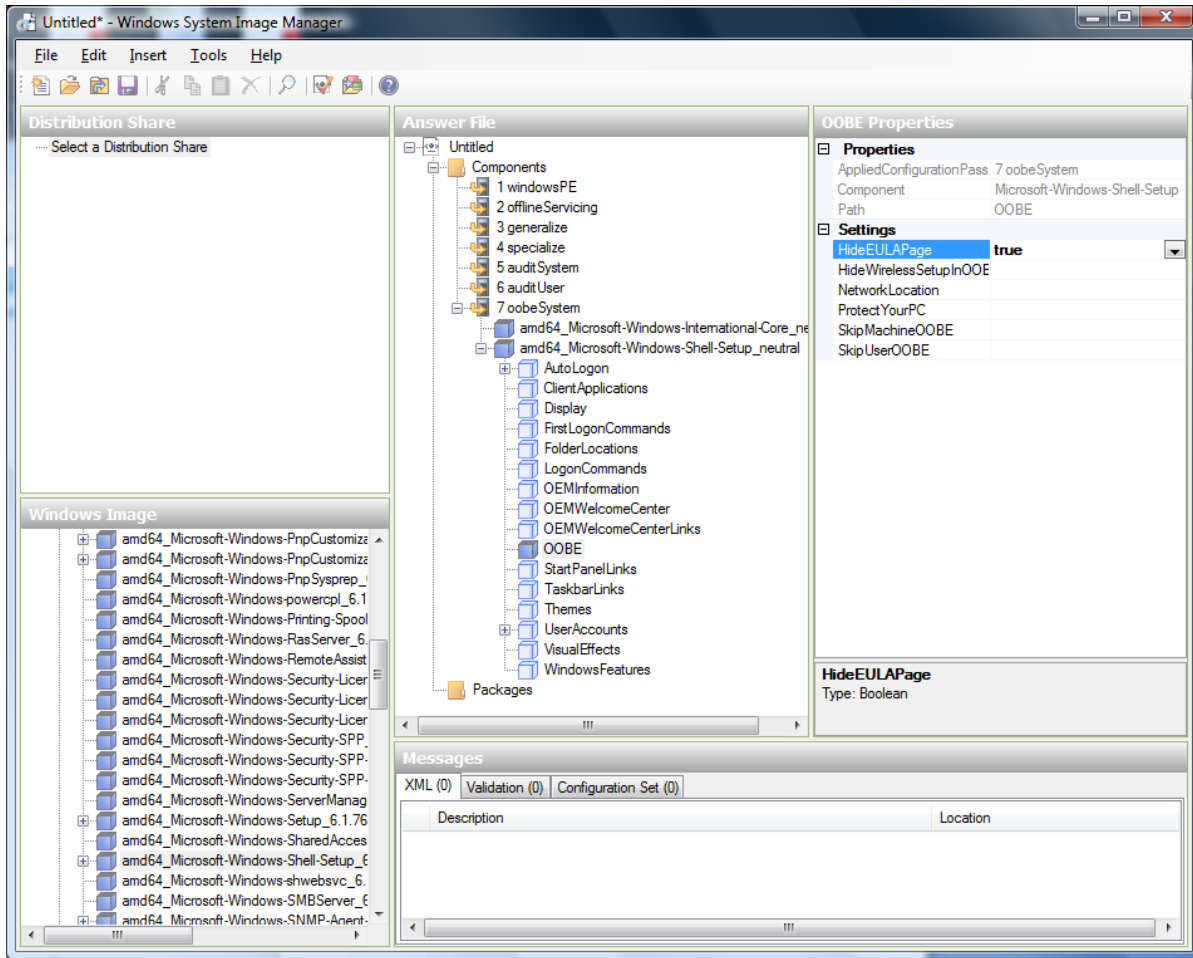
You may be prompted with a warning that the catalog file cannot be opened. Click Yes to create a new catalog file.

6. In the Answer File pane, right click to create a new answer file.
7. Generate the answer file from the Windows System Image Manager using the following steps.
 - a. The first page you need to automate is the Language and Country or Region Selection page. To automate this, expand Components in your Windows Image pane, right-click and add the Microsoft-Windows-International-Core setting to Pass 7 oobeSystem. In your Answer File pane, configure the InputLocale, SystemLocale, UILanguage, and UserLocale with the appropriate settings for your language and country or region. Should you have a question about any of these settings, you can right-click on the specific setting and select Help. This will open the appropriate CHM help file with more information, including examples on the setting you are attempting to configure.

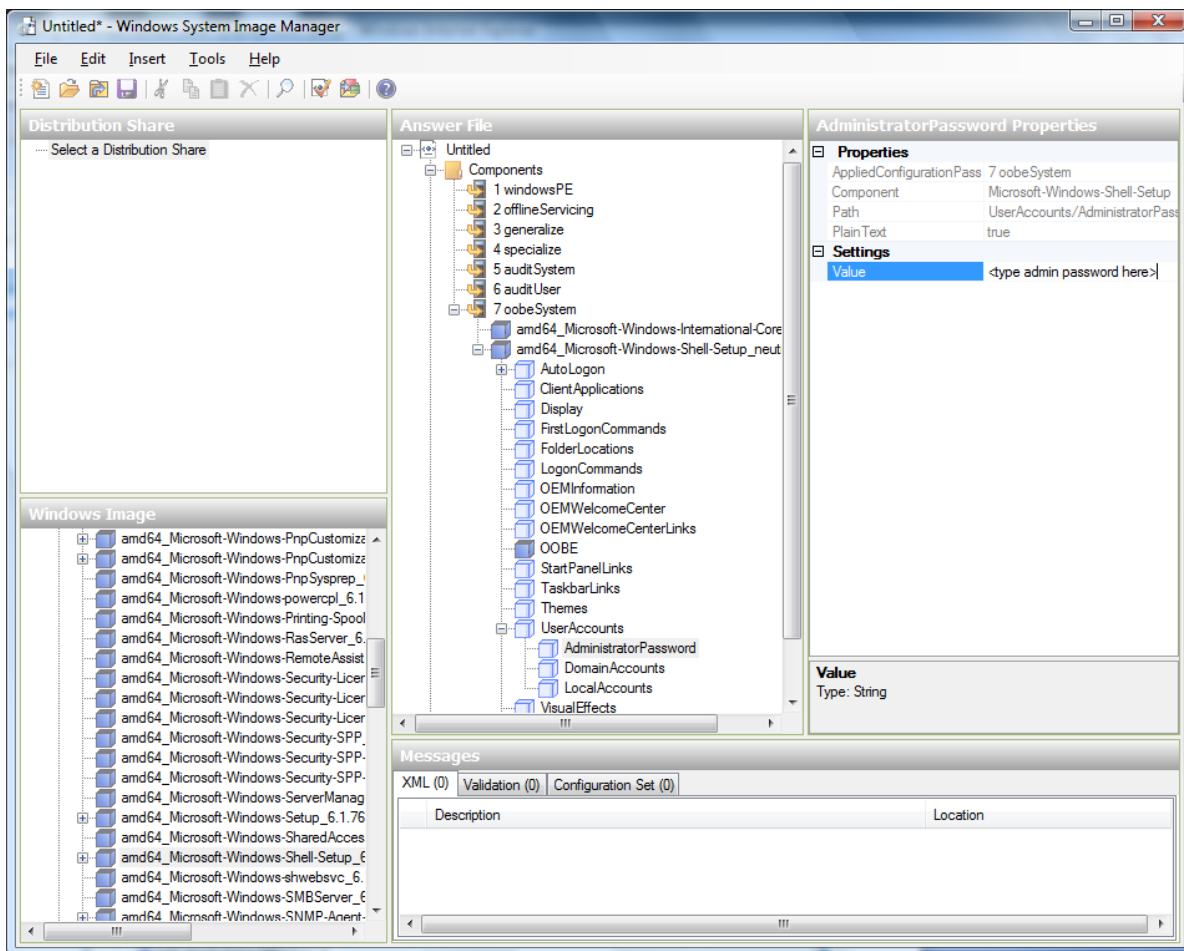
¹ The steps outlined here are derived from the excellent guide by Charity Shelbourne, originally published at the following URL.
<http://blogs.technet.com/askcore/archive/2008/10/31/automating-the-oobe-process-during-windows-server-2008-sysprep-mini-setup.aspx>



- b. You need to automate the Software License Terms Selection page, otherwise known as the End-User License Agreement (EULA). To do this, expand the Microsoft-Windows-Shell-Setup component. High-light the OobeSetting, and add the setting to the Pass 7 oobeSystem . Under Settings, select the drop down next to HideEULAPage and select true.



- c. Make sure the license key is properly set. If you use MAK key, you can just enter the MAK key on the Windows 2008 R2 VM. You need not input the MAK into the Windows System Image Manager. If you use KMS host for activation you need not enter the Product Key. Details of Windows Volume Activation can be found here: <http://technet.microsoft.com/en-us/library/bb892849.aspx>
- d. You need to automate is the Change Administrator Password page. Expand the Microsoft-Windows-Shell-Setup component (if it is not still expanded), expand UserAccounts, right-click on AdministratorPassword, and add the setting to the Pass 7 oobeSystem configuration pass of your answer file. Under Settings, specify a password next to Value.



You may read the AIK documentation and set many more options that suit your deployment. The steps above are the minimum needed to make Windows unattended setup work.

8. Save the answer file as unattend.xml. You can ignore the warning messages that appear in the validation window.
9. Copy the unattend.xml file into the c:\windows\system32\sysprep directory of the Windows 2008 R2 Virtual Machine.
10. Once you place the unattend.xml file in c:\windows\system32\sysprep directory, you run the sysprep tool as follows:

```
cd c:\Windows\System32\sysprep
sysprep.exe /oobe /generalize /shutdown
```

The Windows 2008 R2 VM will automatically shut down after sysprep is complete

7.8.8.2 Sysprep for Windows Server 2003 R2

Earlier versions of Windows have a different sysprep tool. Follow these steps for Windows Server 2003 R2.

1. Extract the content of \support\tools\deploy.cab on the Windows installation CD into a directory called c:\sysprep on the Windows 2003 R2 VM.
2. Run c:\sysprep\setupmgr.exe to create the sysprep.inf file.
 - a. Select Create New to create a new Answer File.
 - b. Enter "Sysprep setup" for the Type of Setup.

- c. Select the appropriate OS version and edition.
- d. On the License Agreement screen, select “Yes fully automate the installation”.
- e. Provide your name and organization.
- f. Leave display settings at default.
- g. Set the appropriate time zone.
- h. Provide your product key.
- i. Select an appropriate license mode for your deployment.
- j. Select “Automatically generate computer name”.
- k. Type a default administrator password. If you enable the password reset feature, the users will not actually use this password. This password will be reset by the VMOps instance manager after the guest boots up.
- l. Leave Network Components at “Typical Settings”.
- m. Select the “WORKGROUP” option.
- n. Leave Telephony options at default.
- o. Select appropriate Regional Settings.
- p. Select appropriate language settings.
- q. Do not install printers.
- r. Do not specify “Run Once commands”.
- s. You need not specify an identification string.
- t. Save the Answer File as `c:\sysprep\sysprep.inf`.

3. Run

```
c:\sysprep\sysprep.exe -reseal -mini -activated
```

to sysprep the image. After this step the machine will automatically shut down.

7.8.8.3 Creating the Windows Template

Once your VM has shutdown, you now can create a template.

1. Click on My Instances and find your VM. Click on it.
2. Click on the volumes tab.
3. Click on the root disk.
4. In the Actions menu, choose Create Template.

7.8.9 Importing AMIs

The following procedures describe how to import an Amazon Machine Image (AMI) into the CloudStack platform when using the XenServer hypervisor.

Assume you have an AMI file and this file is called `CentOS_5.4_x64`. Assume further that you are working on a CentOS host. If the AMI is a Fedora image, you need to be working on a Fedora host initially.

Note: You need to have a XenServer host with a file-based storage repository (either a local ext3 SR or an NFS SR) to convert to a VHD once the image file has been customized on the Centos/Fedora host.

1. Set up loopback on image file:

```
# mkdir -p /mnt/loop/centos54
# mount -o loop CentOS_5.4_x64 /mnt/loop/centos54
```

2. Install the kernel-xen package into the image. This downloads the PV kernel and ramdisk to the image.

```
# yum -c /mnt/loop/centos54/etc/yum.conf --installroot=/mnt/loop/centos54/ -y
install kernel-xen
```

3. Create a grub entry in /boot/grub/grub.conf.

```
# mkdir -p /mnt/loop/centos54/boot/grub
# touch /mnt/loop/centos54/boot/grub/grub.conf
# echo "" > /mnt/loop/centos54/boot/grub/grub.conf
```

4. Determine the name of the PV kernel that has been installed into the image

```
# cd /mnt/loop/centos54
# ls lib/modules/
2.6.16.33-xenU 2.6.16-xenU 2.6.18-164.15.1.el5xen 2.6.18-164.6.1.el5.centos.plus
2.6.18-xenU-ec2-v1.0 2.6.21.7-2.fc8xen 2.6.31-302-ec2
# ls boot/initrd*
boot/initrd-2.6.18-164.6.1.el5.centos.plus.img boot/initrd-2.6.18-
164.15.1.el5xen.img
# ls boot/vmlinuz*
boot/vmlinuz-2.6.18-164.15.1.el5xen boot/vmlinuz-2.6.18-164.6.1.el5.centos.plus
boot/vmlinuz-2.6.18-xenU-ec2-v1.0 boot/vmlinuz-2.6.21-2952.fc8xen
```

Xen kernels/ramdisk always end with "xen". For the kernel version you choose, there has to be an entry for that version under lib/modules, there has to be an initrd and vmlinuz corresponding to that. Above, the only kernel that satisfies this condition is 2.6.18-164.15.1.el5xen

5. Based on your findings, create an entry in the grub.conf file. Below is an example entry.

```
default=0
timeout=5
hiddenmenu
title CentOS (2.6.18-164.15.1.el5xen)
    root (hd0,0)
    kernel /boot/vmlinuz-2.6.18-164.15.1.el5xen ro root=/dev/xvda
    initrd /boot/initrd-2.6.18-164.15.1.el5xen.img
```

6. Edit etc/fstab, changing "sda1" to "xvda" and changing "sdb" to "xvdb".

```
# cat etc/fstab
/dev/xvda / ext3 defaults 1 1
/dev/xvdb /mnt ext3 defaults 0 0
none /dev/pts devpts gid=5,mode=620 0 0
none /proc proc defaults 0 0
none /sys sysfs defaults 0 0
```

7. Enable login via the console. The default console device in a XenServer system is xvc0. Ensure that `etc/inittab` and `etc/securetty` have the following lines respectively:

```
# grep xvc0 etc/inittab
co:2345:respawn:/sbin/agetty xvc0 9600 vt100-nav
# grep xvc0 etc/securetty
xvc0
```

8. Ensure the ramdisk supports PV disk and PV network. Customize this for the kernel version you have determined above.

```
# chroot /mnt/loop/centos54
# cd /boot/
# mv initrd-2.6.18-164.15.1.el5xen.img initrd-2.6.18-164.15.1.el5xen.img.bak
# mkinitrd -f /boot/initrd-2.6.18-164.15.1.el5xen.img --with=xennet --
preload=xenblk --omit-scsi-modules 2.6.18-164.15.1.el5xen
```

9. Change the password.

```
# passwd
Changing password for user root.
New UNIX password:
Retype new UNIX password:
passwd: all authentication tokens updated successfully.
```

10. Exit out of chroot.

```
# exit
```

11. Check `etc/ssh/sshd_config` for lines allowing ssh login using a password.

```
# egrep "PermitRootLogin|PasswordAuthentication"
/mnt/loop/centos54/etc/ssh/sshd_config
PermitRootLogin yes
PasswordAuthentication yes
```

12. If you need the template to be enabled to reset passwords from the CloudStack UI or API, install the password change script into the image at this point. This can be found on <http://open.cloud.com>.

```
# wget -O - http://url/to/get_passwd_from_domr > /mnt/loop/centos54/etc/init.d/
# chroot /mnt/loop/centos54
# chmod a+x /etc/init.d/cloud-set-guest-password
# chkconfig --add cloud-set-guest-password
# chkconfig cloud-set-guest-password on
# exit
```

13. Unmount and delete loopback mount.

```
# umount /mnt/loop/centos54
# losetup -d /dev/loop0
```

- Copy the image file to your XenServer host's file-based storage repository. In the example below, the XenServer is "xenhost". This XenServer has an NFS repository whose uuid is a9c5b8c8-536b-a193-a6dc-51af3e5ff799.

```
# scp CentOS_5.4_x64 xenhost:/var/run/sr-mount/a9c5b8c8-536b-a193-a6dc-51af3e5ff799/
```

- Log in to the XenServer and create a VDI the same size as the image.

```
[root@xenhost ~]# cd /var/run/sr-mount/a9c5b8c8-536b-a193-a6dc-51af3e5ff799
[root@xenhost a9c5b8c8-536b-a193-a6dc-51af3e5ff799]# ls -lh CentOS_5.4_x64
-rw-r--r-- 1 root root 10G Mar 16 16:49 CentOS_5.4_x64
[root@xenhost a9c5b8c8-536b-a193-a6dc-51af3e5ff799]# xe vdi-create virtual-size=10GiB sr-uuid=a9c5b8c8-536b-a193-a6dc-51af3e5ff799 type=user name-label="Centos 5.4 x86_64"
cad7317c-258b-4ef7-b207-cdf0283a7923
```

- Import the image file into the VDI. This may take 10–20 minutes.

```
[root@xenhost a9c5b8c8-536b-a193-a6dc-51af3e5ff799]# xe vdi-import filename=CentOS_5.4_x64 uuid=cad7317c-258b-4ef7-b207-cdf0283a7923
```

- Locate a the VHD file. This is the file with the VDI's UUID as its name. Compress it and upload it to your web server.

```
[root@xenhost a9c5b8c8-536b-a193-a6dc-51af3e5ff799]# bzip2 -c cad7317c-258b-4ef7-b207-cdf0283a7923.vhd > CentOS_5.4_x64.vhd.bz2
[root@xenhost a9c5b8c8-536b-a193-a6dc-51af3e5ff799]# scp CentOS_5.4_x64.vhd.bz2 webserver:/var/www/html/templates/
```

7.8.10 Creating an Ubuntu 10.04 LTS Template for XenServer

This section tells how to create an Ubuntu 10.04 LTS template so that you can create Ubuntu VM instances on the XenServer hypervisor.

- Create a running Ubuntu 10.04 VM by following these steps. (Copied from [Installing Ubuntu 10.04 LTS](http://community.citrix.com/display/xs/Installing+Ubuntu+Server+10.04+%2832bit+and+64bit%29+LTS) on the Citrix Developer Network. Check the website for the most up-to-date information: <http://community.citrix.com/display/xs/Installing+Ubuntu+Server+10.04+%2832bit+and+64bit%29+LTS>.)
 - Copy makeubuntu.sh script to the Pool Master (download the script from the Citrix Developer Network link above).
 - Execute makeubuntu.sh script to create Ubuntu Templates.
 - Create an Ubuntu VM with the new templates.
 - Perform install and reboot.
- Perform the following tests.
 - Make sure the VM is booted with one NIC (eth0).
 - Open the file /etc/network/interfaces and be sure that eth0 is set to use DHCP.
- Stop the Ubuntu VM.

In the next few steps, you will copy the virtual machine's virtual hard disk (VHD) to a web server.

4. From the XenServer command line, list the VMs with the following command. Note the UUID of your Ubuntu VM.

```
# xe vm-list
```

5. List the virtual block devices (VBDs) with the following command, passing in the VM UUID you discovered in the previous step. Note the VDI UUID for the VBD.

```
# xe vbd-list <your Ubuntu VM UUID>
```

6. Navigate to the mount point for primary storage to find the VHD file. The name of the VHD file is the VDI UUID you discovered in the previous step.
7. Copy the VHD file to a webserver.
8. In the CloudStack administrator UI, add a new template (see Creating Templates on page 36).
 - a. Select the URL of the VHD file on the web server as the location.
 - b. For the guest OS type, select Ubuntu if you are running XenServer 5.6 FP1 or greater (for earlier XenServer versions, select CentOS 5.4 x64).
9. Start a new VM from the template.
10. Make sure the VM was able to get an IP address. If not, follow these troubleshooting steps:
 - a. Start a Centos 5.3 x64 VM.
 - b. On the Centos VM, run this command to find the location of the DHCP client script.

```
# which dhclient
```

The location returned should be /sbin/modified-dhclient/dhclient.

- c. On the Ubuntu VM, create a new folder.

```
# mkdir /sbin/modified-dhclient
```

- d. Copy the dhclient script from the Centos VM to the Ubuntu VM at /sbin/modified-dhclient/dhclient.
- e. Add the new folder to the front of your VM's path.
- f. Log out of the VM and log in again.

7.8.11 Converting a Hyper-V VM to a Template

To convert a Hyper-V VM to a XenServer-compatible CloudStack template, you will need a standalone XenServer host with an attached NFS VHD SR. Use whatever XenServer version you are using with CloudStack, but use XenCenter 5.6 FP1 (it is backwards compatible to 5.6). Additionally, it may help to have an attached NFS ISO SR.

For Linux VMs, you may need to do some preparation in Hyper-V before trying to get the VM to work in XenServer. Clone the VM and work on the clone if you still want to use the VM in Hyper-V. Uninstall Hyper-V Integration Components and check for any references to device names in /etc/fstab:

1. From the linux_ic/drivers/dist directory, run make uninstall (where "linux_ic" is the path to the copied Hyper-V Integration Components files).
2. Restore the original initrd from backup in /boot/ (the backup is named *.backup0).
3. Remove the "hdX=noprobe" entries from /boot/grub/menu.lst.

4. Check `/etc/fstab` for any partitions mounted by device name. Change those entries (if any) to mount by LABEL or UUID (get that information with the "blkid" command).

The next step is make sure the VM is not running in Hyper-V, then get the VHD into XenServer. There are two options for doing this.

Option one:

1. Import the VHD using XenCenter. In XenCenter, go to Tools > Virtual Appliance Tools > Disk Image Import.
2. Choose the VHD, then click Next.
3. Name the VM, choose the NFS VHD SR under Storage, enable "Run Operating System Fixups" and choose the NFS ISO SR.
4. Click Next, then Finish. A VM should be created.

Option two:

1. Run XenConvert, under From choose VHD, under To choose XenServer. Click Next.
2. Choose the VHD, then click Next.
3. Input the XenServer host info, then click Next.
4. Name the VM, then click Next, then Convert. A VM should be created.

Once you have a VM created from the Hyper-V VHD, prepare it using the following steps.

1. Boot the VM, uninstall Hyper-V Integration Services, and reboot.
2. Install XenServer Tools, then reboot.
3. Prepare the VM as desired. For example, run sysprep on Windows VMs (see Running Sysprep for Windows Templates on page 38).

Either option above will create a VM in HVM mode. This is fine for Windows VMs, but Linux VMs may not perform optimally.

Converting a Linux VM to PV mode will require additional steps and will vary by distribution.

1. Shut down the VM and copy the VHD from the NFS storage to a web server; for example, mount the NFS share on the web server and copy it, or from the XenServer host use sftp or scp to upload it to the web server.
2. In CloudStack, create a new template using the following values:
 - **URL.** Give the URL for the VHD
 - **OS Type.** Use the appropriate OS.
 - **Hypervisor.** XenServer.
 - **Format.** VHD.

The template will be created and you can create instances from it.

7.8.12 Adding Password Management to Your Templates

The CloudStack platform provides an optional password reset feature that allows users to set a temporary admin or root password as well as reset the existing admin or root password from the CloudStack UI.

To enable the Reset Password feature, you will need to download an additional script to patch your template. When you later upload the template into the CloudStack platform, you can specify whether reset admin/root password feature should be enabled for this template.

The password management feature works always resets the account password on instance boot. The script does an HTTP call to the virtual router to retrieve the account password that should be set. As long as the virtual router is accessible the guest will have access to the account password that should be used. When the user requests a password reset the management server generates and sends a new password to the virtual router for the account. Thus an instance reboot is necessary to effect any password changes.

If the script is unable to contact the virtual router during instance boot it will not set the password but boot will continue normally.

7.8.12.1 Window OS Installation

Download the installer, CloudInstanceManager.msi, from <http://cloud.com/community/downloads> and run the installer in the newly created Windows VM.

7.8.12.2 Linux OS Installation

Use the following steps to begin the Linux OS installation.

1. Download the script file `cloud-set-guest-password` from <http://cloud.com/community/downloads>
2. Copy this file to `/etc/init.d`.
3. On some Linux distributions, you will need to copy the file to `/etc/rc.d/init.d`.
4. Run the following command to make the script executable.

```
chmod +x /etc/init.d/cloud-set-guest-password
```

5. Depending on the Linux distribution, continue with the appropriate step.
 - a. **Fedora, CentOS/RHEL, and Debian.** Run `chkconfig --add cloud-set-guest-password`.
 - b. **Ubuntu.** Run `sudo update-rc.d cloud-set-guest-password defaults 98`. Then run `mkpasswd` and check that it is generating a new password. If the `mkpasswd` command does not exist, run `sudo apt-get install whois` and repeat.

7.9 Working with Snapshots

The CloudStack platform supports snapshots of disk volumes. Snapshots are a point-in-time capture of virtual machine disks. Memory and CPU states are not captured.

Users can create snapshots manually, or by setting up automatic recurring snapshot policies. Users can also create disk volumes from snapshots, which may be attached to a VM as any other disk volume. Snapshots of both root disks and data disks is supported. However, the software does not currently support booting of a VM from a recovered root disk. A disk recovered from snapshot of a root disk is treated as a regular data disk; the data on recovered disk can be accessed by attaching the disk to a VM.

7.9.1 Automatic Snapshot Creation and Retention

Users can set up a recurring snapshot policy to automatically create multiple snapshots of a disk at regular intervals. Snapshots can be created on an hourly, daily, weekly, or monthly intervals. Multiple policies can be set up per disk volume as long as the policies

are not of the same internal type. For example, a user can set up hourly snapshots to be taken every fifteenth minute of the hour, and a daily snapshots at every 02:30 hours of the day. A user cannot set up hourly snapshots at both fifteenth and thirtieth minute of the hour.

With each snapshot schedule, users can also specify the number of snapshots to be retained. Older snapshots that exceed the retention limit are automatically deleted.

7.9.2 Incremental Snapshots and Backup

Snapshots are created on primary storage where a disk resides. After a snapshot is created, it is immediately backed up to secondary storage and removed from primary storage for optimal utilization of space on primary storage.

CloudStack platform does incremental backups for some hypervisors. When incremental backups are supported, every N backup is a full backup.

Hypervisor	VMware vSphere	Citrix XenServer	KVM
Support incremental backup	N	Y	N

7.9.3 Volume Status

When a snapshot operation is triggered by means of a recurring snapshot policy, a snapshot is skipped if a volume has remained inactive since its last snapshot was taken. A volume is considered to be inactive if it is either detached or attached to a VM that is not running. The CloudStack platform ensures that at least one snapshot is taken since the volume last became inactive.

When a snapshot is taken manually, a snapshot is always created regardless of whether a volume has been active or not.

7.9.4 Snapshot Restore

There are two paths to restoring snapshots. Users can create volumes from the snapshot. The volume can then be mounted to a VM and files recovered as needed. A template may be created from the snapshot of a root disk. The user can then boot a VM from this template to effect recovery of the enter root disk.

7.9.5 Performance Considerations

Snapshots not only consume space in secondary storage, but can take up significant CPU cycles and network bandwidth as the snapshots are moved between primary and secondary storage. This is something to be factored in for capacity planning and end-user pricing of snapshot operations.

8 Working with System Virtual Machines

The CloudStack platform uses several types of system virtual machines to perform tasks in the cloud. In general the CloudStack platform manages these system VMs and creates, starts, and stops them as needed based on scale and immediate needs. However, the administrator should be aware of them and their roles to assist in debugging issues.

The System VMs come from a single template. The System VM has the following characteristics:

- Debian 6.0 ("Squeeze"), 2.6.32 kernel with the latest security patches from the Debian security APT repository
- Has a minimal set of packages installed thereby reducing the attack surface
- 32-bit for enhanced performance on Xen/VMWare
- pvops kernel with Xen PV drivers, KVM virtio drivers, and VMware tools for optimum performance on all hypervisors
- Xen tools inclusion allows performance monitoring
- Latest versions of haproxy, iptables, ipsec, apache from debian repository ensures improved security and speed
- Latest version of JRE from Sun/Oracle ensures improved security and speed

8.1 Console Proxy

The Console Proxy has a role in presenting a console view via the web UI. It connects the user's browser to the VNC port made available via the hypervisor for the console of the guest. Both the administrator and end user web UIs offer a console connection.

Clicking on a console icon brings up a new window. The AJAX code downloaded into that window refers to the public IP address of a console proxy VM. There is exactly one public IP address allocated per console proxy VM. The AJAX application connects to this IP. The console proxy then proxies the connection to the VNC port for the requested VM on the Host hosting the guest.

Note: The hypervisors will have many ports assigned to VNC usage so that multiple VNC sessions can occur simultaneously.

There is never any traffic to the guest virtual IP, and there is no need to enable VNC within the guest.

The console proxy VM (available in Premium Edition only) will periodically report its active session count to the Management Server. The default reporting interval is five seconds. This can be changed through standard Management Server configuration with the parameter `consoleproxy.loadscan.interval`. Assignment of guest VM to console proxy is determined by first determining if the guest VM has a previous session associated with a console proxy. If it does, the Management Server will assign the guest VM to the target Console Proxy VM regardless of the load on the proxy VM. Failing that, the first available running Console Proxy VM that has the capacity to handle new sessions is used.

Console proxies can be restarted by administrators but this will interrupt existing console sessions for users.

The console viewing functionality uses a dynamic DNS service under the domain name `realhostip.com` to assist in providing SSL security to console sessions. The console proxy is assigned a public IP address. In order to avoid browser warnings for mismatched SSL certificates, the URL for the new console window is set to the form of `https://a-b-c-d.realhostip.com`. Customers will see this URL during console session creation. The CloudStack includes the `realhostip.com` SSL certificate in the console proxy VM. Of course the CloudStack cannot know about DNS A records for our customers' public IPs prior to shipping the software. As a result Cloud.com runs a dynamic DNS server that is authoritative for the `realhostip.com` domain. It maps the `a-b-c-d` part of the DNS name to the IP

address of a.b.c.d on lookups. This allows the browser to correctly connect to the console proxy's public IP, where it then expects and receives a SSL certificate for realhostip.com, and SSL is set up without browser warnings.

8.1.1 Changing the Console Proxy SSL Certificate and Domain

The use of realhostip.com is configurable. The administrator may select a different domain and upload a new SSL certificate that will be used. The other domain must run a DNS service that is capable of resolving queries for addresses of the form a-b-c-d.some.other.domain as above.

The domain is controlled by a global configuration parameter:

consoleproxy.url.domain - this parameter may be changed, for example to company.com. The Management Server will generate URLs of the form "a-b-c-d.company.com" after this change is made and the Management Server is restarted.

The certificate may be uploaded. Go to Resources -> Physical Resources and select "Upload SSL Certificate".

8.2 Virtual Router

The function of the virtual router was explained in section 1.4. The end user has no direct access to the virtual router. They can ping it and take actions that impact it (E.g. setting up port forwarding) but they do not have SSH access into the virtual router.

There is no mechanism for the administrator to log in to the virtual router. Virtual routers can be restarted by administrators, but this will interrupt public network access and other services for end users.

A basic test in debugging networking issues is to attempt to ping the virtual router from a guest VM.

8.3 Secondary Storage VM

The secondary storage VM provides a background task that copies templates from one Availability Zone to another.

The administrator can log in to the secondary storage VM if needed. The procedure for this is documented in the Troubleshooting section of the Installation Guide.

9 System Reliability and HA

9.1 Management Server

The CloudStack Management Server should be deployed on a server farm such that it is not susceptible to individual server failures. The Management Server itself (as distinct from the MySQL database) is stateless and may be placed behind a load balancer.

Normal operation of Hosts is not impacted by an outage of all Management Servers. All guest VMs will continue to work.

When the Management Server is down, no new VMs can be created, and the end user and admin UI, API, dynamic load distribution, and HA will cease to work.

9.2 Host

When Hosts are down, the CloudStack platform will restart impacted HA-enabled VMs automatically, assuming that other Hosts have sufficient resources available. When the Host comes back online it will be marked as available and newly started VMs may be allocated to it. VMs previously migrated from it will not be migrated back. VMs that were running on it but did not have HA enabled will not be started automatically.

The user will not lose the virtual machine disk image during a Host outage. However, the guest OS may perceive its disk image as corrupt (and needing fsck or equivalent) on restart.

9.3 Primary Storage Outage and Data Loss

When a primary storage outage occurs the hypervisor immediately stops all VMs stored on that storage device. Guests that are marked for HA will be restarted as soon as practical when the primary storage comes back on line. With NFS, the hypervisor may allow the virtual machines to continue running depending on the nature of the issue. For example, an NFS hang will cause the guest VMs to be suspended until storage connectivity is restored.

Primary storage is not designed to be backed up. Individual volumes in primary storage can be backed up using snapshots.

9.4 Secondary Storage Outage and Data Loss

A secondary storage outage will have feature level impact to the system but will not impact running guest VMs. It may become impossible to create a VM with the selected template for a user. A user may also not be able to save snapshots or examine/restore saved snapshots. These features will automatically be available when the secondary storage comes back online.

Secondary storage data loss will impact recently added user data including templates, snapshots, and ISO images.

Secondary storage should be backed up periodically.

9.5 HA-Enabled VM (Service Provider and Enterprise Edition)

The user can specify a virtual machine as HA-enabled.

The system detects HA-enabled VM crashes and restarts the VM automatically within the same Availability Zone. HA is never performed across different Availability Zones.

The CloudStack platform has a conservative policy towards restarting VMs and ensures that there will never be two instances of the same VM running at the same time. The Management Server attempts to start the VM on another Host in the same Cluster.



All virtual router VMs and system VMs are automatically configured as HA-enabled.

HA features work with iSCSI or NFS primary storage. HA with local storage is not supported.

10 Management Features

10.1 Users, Accounts, Administrators, and Domains

There are several types of entities in the system: users, accounts, administrators, and domains.

An account is the unit of isolation in the CloudStack when virtual networking is in place. Typically an account is the entity that is provisioned by the administrator. With the Usage Server, usage records are emitted on a per-account basis.

Multiple users can exist within an account. Conceptually users are like aliases on the account. Users can have different login names, but they have access to the same resources as other users in the same account. Many customers do not expose the concept of users to their customers; they just expose the notion of an account with one login.

An account exists in a domain. Multiple accounts can be in a domain. A domain can contain other child domains. Arbitrary depth is allowed for domain nesting. In the future the CloudStack will provide additional administrative privileges around the concept of a domain.

Administrators are accounts with special privileges in the system. There may be multiple administrators in the system. Administrators can create or delete other administrators.

10.1.1 Root Administrators

Root administrators have complete access to the system, including managing templates, service offerings, customer care administrators, and domains.

10.1.2 Domain Administrators

Domain administrators can perform administrative operations for users who belong to that domain. Domain administrators do not have visibility into physical servers or other domains.

10.2 Provisioning

10.2.1 Register

Users and Accounts must be provisioned and modified through the provisioning API. The following user attributes are specified through the provisioning API:

- Email address and user name (may be the same)
- First and last names
- A callback authentication function or encrypted password
- Reseller ID

10.3 Changing User and Administrator Passwords

The CloudStack includes a "Test Provisioning Tool" in the Administrator UI. This tool includes a change password function that will allow the administrator to change the password for any user in the system.

10.4 VM Lifecycle Management

The CloudStack platform provides administrators with complete control over the lifecycle of all guest VMs executing in the cloud.

10.4.1 VM Creation

Virtual machines are usually created from a template. They may be created from a blank VM booted off an ISO.

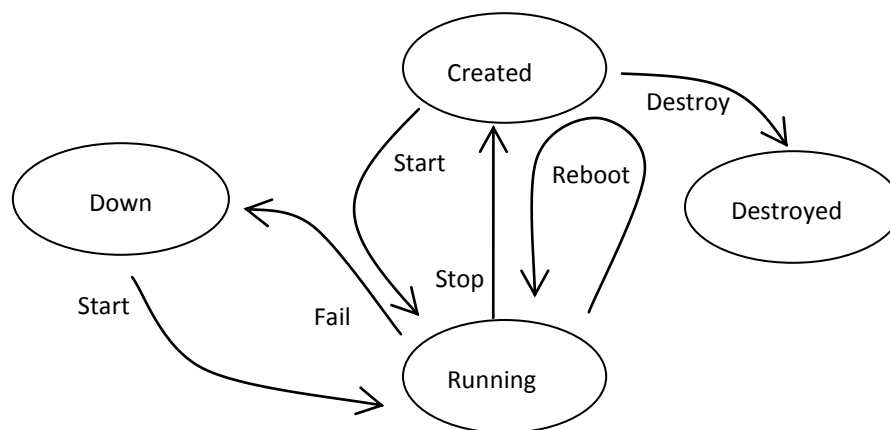
10.4.2 VM Deletion

Users can delete their own virtual machines. A running virtual machine will be abruptly stopped before it is deleted.

Administrators can delete any virtual machines.

10.4.3 VM Lifecycle

Virtual machines can be in the following states:



Once a virtual machine is destroyed, it cannot be recovered. All the resources used by the virtual machine will be reclaimed by the system. This includes the virtual machine's IP address.

A stop will attempt to gracefully shut down the operating system, which typically involves terminating all the running applications. If the operation system cannot be stopped, it will be forcefully terminated. This has the same effect as pulling the power cord to a physical machine.

A reboot is a stop followed by a start.

Unlike Amazon EC2, the system preserves the state of the virtual machine hard disk until the machine is destroyed.

A running virtual machine may fail because of hardware or network issues. A failed virtual machine is in the down state.

The system places the virtual machine into the down state if it does not receive the heartbeat from the hypervisor for three minutes.

The hard disk image is preserved when a virtual machine enters the down state.

The user can manually restart the virtual machine from the down state.

The system will start the virtual machine from the down state automatically if the virtual machine is marked as HA-enabled.

10.4.4 Remote Access

The user is able to access virtual machine console through the web management UI.

Administrators can access the virtual machine consoles that belong to any users for the purpose of support and troubleshooting. The administrators will be subject to providing the root (or other) password for the guest.

10.5 Changing the Database Configuration

The CloudStack Management Server stores database configuration information (e.g., hostname, port, credentials) in the file `/etc/cloud/management/db.properties`. To effect a change, edit this file on each Management Server, then restart the Management Server.

10.6 PV Drivers

For XenServer, Windows VMs require PV drivers to be added in either the template or after install for the CloudStack platform management functions to work properly. The PV drivers allow functions such as mounting additional volumes and ISO images, live migration, and graceful shutdown.

10.7 Administrator Alerts

Emails will be sent to administrators under the following circumstances:

1. The Management Server cluster runs low on CPU, memory, or storage resources
2. The Management Server loses heartbeat from a Host for more than 3 minutes
3. The Host cluster runs low on CPU, memory, or storage resources

10.8 Limits

The CloudStack platform provides several administrator control points for capping resource usage by users. Some of these limits are global configuration parameters. Others are applied at the ROOT domain and may be overridden on a per-account basis.

Aggregate limits may be set on a per-domain basis. For example, you may limit a domain and all subdomains to the creation of 100 VMs.

10.8.1 Configuration Limits

On a Zone the guest virtual network has a 24 bit CIDR by default. This limits the guest virtual network to 254 running instances. It can be adjusted as needed, but this must be done before any instances are created in the Zone. For example, 10.1.1.0/22 would provide for ~1000 addresses.

The following table lists limits set on in the Global Configuration.

Parameter Name	Definition
max.template.iso.size	Maximum size for a downloaded template or ISO in GB
max.volume.size.gb	Maximum size for a volume in GB
network.throttling.rate	Default data transfer rate in megabits per second allowed per user
snapshot.max.hourly	Maximum hourly snapshots for a volume
snapshot.max.daily	Maximum daily snapshots for a volume
snapshot.max.weekly	Maximum weekly snapshots for a volume
snapshot.max.monthly	Maximum monthly snapshots for a volume

10.8.2 ROOT Domain Limits


The following limits are set at the ROOT domain, inherited by all other domains, and may be overridden on a per-account basis.

- Number of instances a user can create.
- Number of public IP addresses a user can own.
- Number of disk volumes a user can create.
- Number of snapshots a user can create.
- Number of templates that a user can register/create.

10.8.3 Per-Domain Limits

The CloudStack allows the configuration of limits on a domain basis. With a domain limit in place all users still have their account limits. They are additionally limited, as a group, to not exceed the resource limits set on their domain. Domain limits aggregate the usage of all accounts in the domain as well as all accounts in all subdomains of that domain.

To set a domain limit, go to Domains in the admin UI, then find the target domain in the tree of domains. Once there select the Resource Limits tab. Choose "Edit Resource Limits" in the Actions menu. The displayed values will become editable. A value of -1 shows that there is no limit in place.

 Domains

Details Admin Accounts **Resource Limits**

Instance Limits:	<input type="text" value="-1"/>
Public IP Limits:	<input type="text" value="-1"/>
Volume Limits:	<input type="text" value="-1"/>
Snapshot Limits:	<input type="text" value="-1"/>
Template Limits:	<input type="text" value="-1"/>

The limits available are the same as described for the ROOT domain previously.

11 Working with Hosts

11.1 Adding Hosts to a Cluster

Additional Hosts may be added at any time up to the limit of nodes in a cluster for the selected Hypervisor type.

The administrator may use the Add Host function in the System -> Physical Resources -> Zone section in the admin UI to add a new host.

11.1.1 vSphere Host Addition

For vSphere, vSphere clusters may be added using this function as well.

vCenter may also be used to add individual hosts to an existing Cluster.

11.1.2 XenServer Host Addition

If network bonding is in use, the administrator must cable the Host identically to other Hosts in the Cluster.

For all additional hosts to be added to the Cluster execute this step. This will cause the host to join the master in a XenServer pool.

```
# xe pool-join master-address=[master IP] master-username=root master-  
password=[your password]
```

With all hosts added to the XenServer pool, run the cloud-setup-bond script. This script will complete the configuration and set up of the bonds on the new Hosts in the Cluster.

1. Copy the script from the Management Server in `/usr/lib64/cloud/agent/scripts/vm/hypervisor/xenserver/cloud-setup-bonding.sh` to the master Host and ensure it is executable.
2. Run the script:

```
# ./cloud-setup-bonding.sh
```

11.1.3 KVM Host Addition

If shared mountpoint storage is in use, the administrator should ensure that the new Host has all the same mountpoints (with storage mounted) as the other Hosts in the Cluster.

11.2 Scheduled Maintenance and Maintenance Mode

To perform maintenance on a host, both vCenter and CloudStack must be used in concert. CloudStack and vCenter have separate maintenance modes that work closely together.

1. Place the host into CloudStack's "scheduled maintenance" mode. This does not invoke the vCenter maintenance mode, but only causes VMs to be migrated off the host.

When the CloudStack maintenance mode is requested, the host first moves into the Prepare for Maintenance state. In this state it cannot be the target of new guest VM starts. Then all VMs will be migrated off the server. Live migration will be used to move VMs off the host. This allows the guests to be migrated to other hosts with no disruption to the guests. After this migration is completed, the host will enter the Ready for Maintenance mode.

2. Wait for the "Ready for Maintenance" indicator to appear in the UI.
3. Now use vCenter to perform whatever actions are necessary to maintain the host. During this time, the host cannot be the target of new VM allocations.
4. When the maintenance tasks are complete, take the host out of maintenance mode as follows:
 - a. First use vCenter to exit the vCenter maintenance mode.
This makes the host ready for CloudStack to reactivate it.
 - b. Then use CloudStack's administrator UI to cancel the CloudStack maintenance mode.

When the host comes back online, the VMs that were migrated off of it are migrated back to it and new VMs can be added.

Some host in the cluster must be up in order for a host to be put into maintenance mode. With XenServer, maintenance mode ejects the host from the XenServer pool. If no hosts are up then the command to eject the host cannot be processed by any XenServer, and maintenance mode will fail.

With XenServer the administrator should confirm that the host has been ejected from the pool after maintenance mode has been entered. To check this, on the removed host:

```
# xe host-list
```

If this returns more than one host then you should manually eject this host:

```
# xe pool-eject uuid={uuid of this host}
```

11.3 Removing Hosts

Hosts can be removed from the cloud as needed. The procedure to remove a Host varies depending on hypervisor type.

11.3.1 XenServer and KVM Hosts

A node cannot be removed from a Cluster until it has been placed in maintenance mode. This will ensure that all of the VMs on it have been migrated to other Hosts. To remove a Host from the cloud:

1. Place the node in maintenance mode (see "Scheduled Maintenance and Maintenance Mode" on page 60)
2. For KVM, stop the cloud-agent service
3. Use the UI option to remove the node.

Then you may power down the Host, re-use its IP address, re-install it, etc.

11.3.2 vSphere Hosts

To remove this type of host, first place it in maintenance mode, as described in Scheduled Maintenance and Maintenance Mode on page 60. Then use CloudStack to remove the host. CloudStack will not direct commands to a host that has been removed using CloudStack. However, the host may still exist in the vCenter cluster.

11.4 Re-installing Hosts

From time to time it may be necessary to re-install a Host. If a node must be re-installed it should first be placed in maintenance mode and then removed. If a node is down and cannot be placed in maintenance mode it should be removed before the re-install. See "Scheduled Maintenance and Maintenance Mode" on page 60.

For XenServer, a node that has been re-installed will have a different UUID than it had previously. Presenting the CloudStack and other Hosts with a Host with the same IP address but changed UUID can cause unpredictable results. This should be avoided by removing XenServer hosts before re-installing them.

11.5 Changing Host IP Address

A Host's IP address can be changed. It should be placed into maintenance mode and removed from the CloudStack. Then it can have its IP address changed. Then it may be added back into the CloudStack. See "Scheduled Maintenance and Maintenance Mode" on page 60.

11.6 Changing Host Password

The password for a XenServer Node, KVM Node, or vSphere Node may be changed in the database. Note that all Nodes in a Cluster must have the same password. To change a Node's password:

1. Identify all Nodes in the Cluster
2. Change the password on all Nodes in the Cluster. At this time the password for the Node and the password known to the CloudStack will not match. Operations on the Cluster will fail until the two passwords match.
3. Get the list of host IDs for the host in the cluster where you are changing the password. You will need to access the database to determine these host IDs. For each hostname "h" (or vSphere cluster) that you are changing the password for, execute;

```
mysql> select id from cloud.host where name like '%h%';
```

This should return a single ID. Record the set of such IDs for these hosts.

4. Update the passwords for the host in the database. In this example we change the passwords for Nodes with IDs 5, 10, and 12 to "password".

```
mysql> update cloud.host set password='password' where id=5 or id=10 or id=12;
```

11.7 Host Allocation

The system automatically picks the most appropriate Host to host virtual machines. End users may specify the Availability Zone in which the virtual machine will be created. End users do not have control over which Host will host the virtual machine instance.

11.7.1 OS Preferences

The CloudStack platform allows administrators to specify that certain Hosts should have a preference for particular types of guest instances. For example, an administrator could state that a Host should have a preference to run Windows guests. If this is set the default Host allocator will attempt to place guests of that OS type on such nodes first. If no such node is available the allocator will place the node wherever there is sufficient physical capacity.

11.7.2 Over-Provisioning and Service Offering Limits

The CloudStack platform does not perform memory over provisioning.

The CloudStack platform performs CPU over-provisioning based on an over-provisioning ratio configured by the administrator. This is defined by the `cpu.overprovisioning.factor` global configuration variable.

CPU over-provisioning allows the sum total of the gigahertz of CPU speed allocated to guests to exceed the physically available gigahertz. For example, if a Host had 2 cores at 2 GHz each, it would have 4 GHz total. With a CPU over provisioning factor of 1.5, the CloudStack would allocate VMs up to 6 GHz total on the Host.

Service offerings limits (e.g. 1 GHz, 1 core) are strictly enforced for core count. For example, a guest with a service offering of one core will have only one core available to it regardless of other activity on the Host.

Service offering limits for gigahertz are enforced only in the presence of contention for CPU resources. For example, suppose that a guest was created with a service offering of 1 GHz on a Host that has 2 GHz cores, and that guest is the only guest running on the Host. The guest will have the full 2 GHz available to it. When multiple guests are attempting to use the CPU a weighting factor is used to schedule CPU resources. The weight is based on the clock speed in the service offering. Guests receive a CPU allocation that is proportionate to the GHz in the service offering. For example, a guest created from a 2 GHz service offering will receive twice the CPU allocation as a guest created from a 1 GHz service offering.

11.8 VLAN Provisioning

The CloudStack automatically creates and destroys interfaces bridged to VLANs on the Hosts. In general the administrator does not need to manage this process. Zone VLANs are allocated sequentially (1,2,3,4, ...) by the CloudStack to accounts as demand warrants.

The CloudStack manages VLANs differently based on hypervisor type. For XenServer or KVM, the VLANs are created on only the hosts where they will be used and then they are destroyed when all guests that require them have been terminated or moved to another host.

For vSphere the VLANs are provisioned on all hosts in the Cluster even if there is no guest running on a particular Host that requires the VLAN. This allows the administrator to perform live migration and other functions in vCenter without having to create the VLAN on the destination Host. Additionally, the VLANs are not removed from the Hosts when they are no longer needed.



12 Working with Usage

The Usage Server provides aggregated usage records. It provides billing integration for the CloudStack platform. The Usage Server works by taking data from the events log, and then creating summary usage records for access via the listUsageRecords API call.

The Usage Server runs at least once per day. It can be configured to run multiple times per day. Its behavior is controlled by the following Configuration table settings.

Parameter Name	Parameter Definition
usage.stats.job.exec.time	<p>This is the time that the Usage Server processing will start. It is specified in 24-hour format in the time zone of the server, which should be GMT. For example, to start the Usage job at 10:30 GMT, enter "10:30".</p> <p>Note: The Usage server processing may run multiple times per day depending on the value of usage.stats.job.aggregation.range.</p>
usage.execution.timezone	<p>Defines that time zone that the usage execution will start in. That is, the time zone of usage.stats.job.exec.time. Valid values for the time zone are specified in Appendix A—Time Zones.</p>
usage.stats.job.aggregation.range	<p>This is the time period in minutes between Usage server processing jobs. For example, if you set it to 1440 the Usage server will run once per day. If you set it to 600 it will run every ten hours. In general, when a Usage server job runs it processes all events generated since usage was last run.</p> <p>There is special handling for the case of 1440 (once per day). In this case the Usage Server does not necessarily process all records since Usage was last run. The CloudStack platform assumes that you require processing once per day for the previous, complete day's records. For example, if the current day is October 7 then you would like to process records for October 6, from midnight to midnight. The CloudStack platform assumes this "midnight to midnight" is relative to the usage.execution.timezone.</p>

For example, suppose that your server is in GMT, your user population is predominantly in the East Coast of the United States, and you would like to process usage records every night at 2 AM local (EST) time. Choose these settings:

- **usage.stats.job.exec.time** = 07:00. This will run the Usage job at 2:00 AM EST. Note that this will shift by an hour as the East Coast of the U.S. enters and exits Daylight Savings Time.
- **usage.execution.timezone** = America/New_York
- **usage.stats.job.aggregation.range** = 1440

With this configuration, the Usage job will run every night at 2 AM EST and will process records for the previous day's midnight-midnight as defined by the EST (America/New_York) time zone.

Note: Because the special value 1440 has been used for usage.stats.job.aggregation.range the Usage Server will ignore the data between midnight and 2 AM.

13 User Interface and API

13.1 User Interface

The system supports both an administrator interface and an end user interface.

13.1.1 Admin User Interface

The Admin UI supports the following functionalities:

- Service offering management
- User management
- Template management
- Virtual machine management
- Server management
- Storage management
- Network management
- Events
- Initial set up
- Dashboard

13.1.2 End User Interface

The end user UI is an AJAX-based UI available in popular browsers including IE7, IE8, IE9, Firefox 3.5+, Firefox 4, Safari 4, and Safari 5. It offers a number of features for end users:

- Service Offering Description
- Template Management
- Virtual Machine Management
- Network Management
- Event Logs
- Snapshot Management
- Dashboard

13.2 API

The CloudStack API is a low level API that has been used to implement the web UI's. It is also a good basis for implementing other popular API's such as EC2/S3 and emerging DMTF standards.

The 2.0 API has had structural changes to make many of the calls that were previously synchronous are now asynchronous. These calls will return a Job ID immediately when called. This Job ID can be used to query the status of the job later. Also, status calls on impacted resources will provide some indication of their state.

The API has a REST-like query basis and returns results in XML or JSON.

The complete API is available at <http://open.cloud.com>.

13.2.1 Provisioning and Authentication API

The CloudStack platform expects that a customer will have their own user provisioning infrastructure. It provides APIs to integrate with these existing systems where the systems call out to the CloudStack platform to add/remove users.

The CloudStack platform support pluggable authenticators. By default the CloudStack platform assumes it is provisioned with the user's password, and as a result authentication is done locally. However, external authentication (E.g. via LDAP) is possible as well.

13.2.2 Allocators

The CloudStack platform enables administrators to write custom allocators that will choose the Host to place a new guest and the storage host from which to allocate guest virtual disk images.

13.2.3 User Data and Meta Data

The CloudStack platform provides API access to attach user data to a deployed VM. Deployed VMs also have access to instance metadata via the virtual router.

User data can be accessed once the IP address of the virtual router is known. Once the IP address is known, use the following steps to access the user data:

1. Run the following command to find the virtual router.

```
# cat /var/lib/dhclient/dhclient-eth0.leases | grep dhcp-server-identifier | tail -1
```

2. Access user data by running the following command using the result of the above command.

```
# curl http://10.1.1.1/latest/user-data
```

Meta Data can be accessed similarly, using a URL of the form <http://10.1.1.1/latest/{metadata type}>. The following are the possible metadata types.

- **service-offering.** A description of the VMs service offering.
- **availability-zone.** The Zone name.
- **local-ipv4.** The guest IP of the VM.

- **local-hostname.** The hostname of the VM.
- **public-ipv4.** The first public IP for the router. (E.g. the first IP of eth2)
- **public-hostname.** This is the same as public-ipv4.
- **instance-id.** The instance name of the VM.

14 Tuning

This section provides tips on how to improve the performance of your cloud.

14.1 Increase Management Server Maximum Memory

If the Management Server is subject to high demand, the default maximum JVM memory allocation can be insufficient. To increase the memory:

1. Edit the Tomcat configuration file `/etc/cloud/management/tomcat6.conf`.
2. Change the command-line parameter `-XmxNNNm` to a higher value of *N*. For example, if the current value is `-Xmx128m`, change it to `-Xmx1024m` or higher.
3. To put the new setting into effect, restart the Management Server.

```
# service cloud-management restart
```

For more information about memory issues, see "FAQ: Memory" in the Tomcat Wiki at <http://wiki.apache.org/tomcat/FAQ/Memory>.

14.2 Set Database Buffer Pool Size

It is important to provide enough memory space for the MySQL database to cache data and indexes.

1. Edit the MySQL configuration file `/etc/my.cnf`.
2. Insert the following line in the `[mysqld]` section, below the `datadir` line. Use a value that is appropriate for your situation. We recommend setting the buffer pool at 40% of RAM if MySQL is on the same server as the management server or 70% of RAM if MySQL has a dedicated server. The following example assumes a dedicated server with 1024M of RAM.

```
innodb_buffer_pool_size=700M
```

3. Restart the MySQL service:

```
# service mysqld restart
```

For more information about the buffer pool, see "The InnoDB Buffer Pool" in the MySQL Reference Manual at <http://dev.mysql.com/doc/refman/5.5/en/innodb-buffer-pool.html>.

15 Troubleshooting

This section describes how to diagnose and remedy runtime issues.

15.1 Event Logs

There are two types of events logged in the Cloud.com CloudStack Event Log. Standard events log the success or failure of an event and can be used to identify jobs or processes that have failed. There are also long running job events. Events for asynchronous jobs log when a job is scheduled, when it starts, and when it completes. Other long running synchronous jobs log when a job starts, and when it completes. Long running synchronous and asynchronous event logs can be used to gain more information on the status of a pending job or can be used to identify a job that is hanging or has not started. The following sections provide more information on these events.

15.1.1 Standard Events

The events log records three types of standard events.

- **INFO.** This event is generated when an operation has been successfully performed.
- **WARN.** This event is generated in the following circumstances.
 - When a network is disconnected while monitoring a template download.
 - When a template download is abandoned.
 - When an issue on the storage server causes the volumes to fail over to the mirror storage server.
- **ERROR.** This event is generated when an operation has not been successfully performed.

15.1.2 Long Running Job Events

In addition to the three standard event types, the events log also records the three following events for long running jobs.

- **SCHEDULED.** (Asynchronous jobs only) This event is generated when an asynchronous job is submitted.
- **STARTED.** This event is generated when a job begins execution.
- **COMPLETED.** This event is generated when a job is completed.

Both the Started and Completed events are logged for all long running job types. The Scheduled event is only logged for asynchronous events. When an action is initiated synchronously or as part of another asynchronous job, the Scheduled event won't be logged.

15.1.3 Event Log Queries

Database logs can be queried from the user interface. The list of events captured by the system includes:

- Virtual machine creation, deletion, and on-going management operations
- Virtual router creation, deletion, and on-going management operations

- Template creation and deletion
- Network/load balancer rules creation and deletion
- Storage volume creation and deletion
- User login and logout

The following is the full list of event types.

VM.CREATE	USER.DISABLE	DISK.OFFERING.CREATE
VM.DESTROY	TEMPLATE.CREATE	DISK.OFFERING.EDIT
VM.START	TEMPLATE.DELETE	DISK.OFFERING.DELETE
VM.STOP	TEMPLATE.UPDATE	NETWORK.OFFERING.CREATE
VM.REBOOT	TEMPLATE.COPY	NETWORK.OFFERING.EDIT
VM.UPGRADE	TEMPLATE.DOWNLOAD.START	NETWORK.OFFERING.DELETE
VM.RESETPASSWORD	TEMPLATE.DOWNLOAD.SUCCESS	POD.CREATE
ROUTER.CREATE	TEMPLATE.DOWNLOAD.FAILED	POD.EDIT
ROUTER.DESTROY	TEMPLATE.EXTRACT	POD.DELETE
ROUTER.START	TEMPLATE.UPLOAD	ZONE.CREATE
ROUTER.STOP	TEMPLATE.CLEANUP	ZONE.EDIT
ROUTER.REBOOT	VOLUME.CREATE	ZONE.DELETE
ROUTER.HA	VOLUME.DELETE	VLAN.IP.RANGE.CREATE
PROXY.CREATE	VOLUME.ATTACH	VLAN.IP.RANGE.DELETE
PROXY.DESTROY	VOLUME.DETACH	CONFIGURATION.VALUE.EDIT
PROXY.START	VOLUME.EXTRACT	SG.AUTH.INGRESS
PROXY.STOP	VOLUME.UPLOAD	SG.REVOKE.INGRESS
PROXY.REBOOT	SERVICEOFFERING.CREATE	HOST.RECONNECT
PROXY.HA	SERVICEOFFERING.UPDATE	MAINT.CANCEL
VNC.CONNECT	SERVICEOFFERING.DELETE	MAINT.CANCEL.PS
VNC.DISCONNECT	DOMAIN.CREATE	MAINT.PREPARE
NET.IPASSIGN	DOMAIN.DELETE	MAINT.PREPARE.PS
NET.IPRELEASE	DOMAIN.UPDATE	VPN.REMOTE.ACCESS.CREATE
NET.RULEADD	SNAPSHOT.CREATE	VPN.REMOTE.ACCESS.DESTROY
NET.RULEDELETE	SNAPSHOT.DELETE	VPN.USER.ADD
NET.RULEMODIFY	SNAPSHOTPOLICY.CREATE	VPN.USER.REMOVE
NETWORK.CREATE	SNAPSHOTPOLICY.UPDATE	NETWORK.RESTART
NETWORK.DELETE	SNAPSHOTPOLICY.DELETE	UPLOAD.CUSTOM.CERTIFICATE
LB.ASSIGN.TO.RULE	ISO.CREATE	STATICNAT.ENABLE
LB.REMOVE.FROM.RULE	ISO.DELETE	STATICNAT.DISABLE
LB.CREATE	ISO.COPY	SSVM.CREATE
LB.DELETE	ISO.ATTACH	SSVM.DESTROY
LB.UPDATE	ISO.DETACH	SSVM.START
USER.LOGIN	ISO.EXTRACT	SSVM.STOP
USER.LOGOUT	ISO.UPLOAD	SSVM.REBOOT
USER.CREATE	SERVICE.OFFERING.CREATE	SSVM.HA
USER.DELETE	SERVICE.OFFERING.EDIT	
USER.UPDATE	SERVICE.OFFERING.DELETE	

15.2 Working with Server Logs

The CloudStack Management Server logs all web site, middle tier, and database activities for diagnostics purposes in `/var/log/cloud/management/`. The CloudStack logs a variety of error messages. We recommend this command to find the problematic output in the Management Server log:

```
grep -i -E 'exc|unable|fail|invalid|leak|invalid|warn|error'
/var/log/cloud/management/management-server.log
```

The CloudStack processes requests with a Job ID. If you find an error in the logs and you are interested in debugging the issue you can `grep` for this job ID in the management server log. For example, suppose that you find the following ERROR message:

```
2010-10-04 13:49:32,595 ERROR [cloud.vm.UserVmManagerImpl] (Job-Executor-11:job-
1076) Unable to find any host for [User|i-8-42-VM-untagged]
```

Note that the job ID is 1076. You can track back the events relating to job 1076 with the following `grep`:

```
grep "job-1076)" management-server.log
```

The CloudStack Agent Server (present in the Community Edition) logs its activities in `/var/log/cloud/agent/`.

15.3 Data Loss on Exported Primary Storage

Symptom

Loss of existing data on primary storage which has been exposed as a Linux NFS server export on an iSCSI volume.

Cause

It is possible that a client from outside the intended pool has mounted the storage. When this occurs, the LVM is wiped and all data in the volume is lost.

Solution

When setting up LUN exports, restrict the range of IP addresses that are allowed access by specifying a subnet mask. For example:

```
echo "/export 192.168.1.0/24(rw,async,no_root_squash)" > /etc/exports
```

Adjust the above command to suit your deployment needs.

More Information

See the export procedure in the "Secondary Storage" section of the CloudStack Installation Guide.

15.4 Maintenance mode not working on vCenter

Symptom

Host was placed in maintenance mode, but still appears live in vCenter.

Cause

The CloudStack administrator UI was used to place the host in scheduled maintenance mode. This mode is separate from vCenter's maintenance mode.

Solution

Use vCenter to place the host in maintenance mode.

More Information

See "Scheduled Maintenance and Maintenance Mode" on page 60.

15.5 Unable to deploy VMs from uploaded vSphere template

Symptom

When attempting to create a VM, the VM will not deploy.

Cause

If the template was created by uploading an OVA file that was created using vSphere Client, it is possible the OVA contained an ISO image. If it does, the deployment of VMs from the template will fail.

Solution

Remove the ISO and re-upload the template.

16 Appendix A—Time Zones

The following time zone identifiers are accepted by the CloudStack API. There are several places that have a time zone as a required or optional parameter. These include scheduling recurring snapshots, creating a user, and specifying the usage time zone in the Configuration table.

Etc/GMT+12	America/La_Paz	Asia/Jerusalem
Etc/GMT+11	America/Santiago	Europe/Minsk
Pacific/Samoa	America/St_Johns	Europe/Moscow
Pacific/Honolulu	America/Araguaina	Africa/Nairobi
US/Alaska	America/Argentina/Buenos_Aires	Asia/Karachi
America/Los_Angeles	America/Cayenne	Asia/Kolkata
Mexico/BajaNorte	America/Godthab	Asia/Bangkok
US/Arizona	America/Montevideo	Asia/Shanghai
US/Mountain	Etc/GMT+2	Asia/Kuala_Lumpur
America/Chihuahua	Atlantic/Azores	Australia/Perth
America/Chicago	Atlantic/Cape_Verde	Asia/Taipei
America/Costa_Rica	Africa/Casablanca	Asia/Tokyo
America/Mexico_City	Etc/UTC	Asia/Seoul
Canada/Saskatchewan	Atlantic/Reykjavik	Australia/Adelaide
America/Bogota	Europe/London	Australia/Darwin
America/New_York	CET	Australia/Brisbane
America/Caracas	Europe/Bucharest	Australia/Canberra
America/Asuncion	Africa/Johannesburg	Pacific/Guam
America/Cuiaba	Asia/Beirut	Pacific/Auckland
America/Halifax	Africa/Cairo	