RED HAT®
VIRTUALIZATION

redhat.

TECHNOLOGY DETAIL

BEST PRACTICES FOR RED HAT VIRTUALIZATION 4

Guidance, recommendations, and considerations for design and deployment

TABLE OF CONTENTS

facebook.com/redhatinc
@redhatnews
linkedin.com/company/red-hat

redhat.com

INTRODUCTION

This document provides guidance, recommendations, and considerations when designing or deploying Red Hat® Virtualization. This information comes from years of field-tested knowledge and comes from customers, developers, engineers, Red Hat consultants, Red Hat support, and Red Hat partners.

TARGET AUDIENCE

System engineers, system architects, system administrators, and anyone in the process of determining the layout of a virtualization environment will benefit from this document.

SCOPE

This document provides a baseline for administrators, engineers, and architects considering the design or deployment of a virtualization environment using Red Hat Virtualization. The information in this document will not align perfectly with all environments or use cases; however, with proper testing and validation, it will provide a solid base.

This document has two primary focal points:

- To provide guidance that leads to a proper design
- To provide recommendations that lead to a successful deployment

AVOIDING THE MOST COMMON MISTAKES

READ AND FOLLOW THE INSTALLATION AND ADMINISTRATION DOCUMENTATION

As straightforward as the deployment of Red Hat Virtualization is, it is still helpful to read the installation documentation for a solid understanding of the overall workflow. It is even more important to follow the installation recommendations to ensure a smooth installation. Even a brief review of the documentation will teach you several things about the product.

Note that all instructions in this document can also be found in the installation guide and the administration guide for Red Hat Virtualization. Deeper product knowledge is available in the technical reference for Red Hat Virtualization. Links to more online documentation can be found at the end of this document.

FULLY PREPARED ENVIRONMENT

After reviewing the installation documentation, be sure to have the environment completely ready for deployment. This may sound like an obvious recommendation, but many deployments are delayed because of items such as incorrect network configuration, waiting on server setup, or lack of storage provisioning. Additionally, other network-based services, like DNS, will play a major part. Commonly overlooked steps include:

- Not having fully qualified domain names (FQDN) for hosts and Red Hat Virtualization Manager, including full resolution in DNS (forward and reverse).
- Not having all of the underlying storage provisioned and configured properly for the storage domains.
- Not having the underlying network prepared correctly, including switch and virtual LAN (VLAN) configuration.
- Not having the right kind of hardware for the hypervisors (e.g., no remote power control).
- Not using the latest available versions of the software.
- Not having fully up-to-date hardware, firmware, BIOS, etc.
- Misconfigured remote management/out-of-band (OOB)/power management.
- Attempting to set up the environment in the wrong order (e.g., storage prior to networking).
- Not having proper file system layout (e.g., VM thinpool when using Red Hat Virtualization Hypervisor).
- Not having proper /var size requirements for hosted-engine deployment.
- Not having correct software channels subscribed or properly registered with Red Hat Content Delivery Network (CDN) or Red Hat Satellite.
- Not properly testing before going into production.
- Missing other configuration items that lengthen or frustrate the deployment process.

ENVIRONMENT

This is an often-overlooked area that bears emphasis. The environment in which Red Hat Virtualization is to be deployed needs to comparable to the workloads it will support. That is, if the workloads are mission-critical, then the overall environment should be mission-critical.

Things to consider in a mission-critical environment include:

- Redundant Ethernet switches for both data and storage.
- Redundant fibre channel (FC) switches.
- Redundant power to all components.
- Redundant power in the datacenter.
- Redundant network providers to the datacenter.
- Redundant network services (e.g., DNS, LDAP).

In addition, this means that Red Hat Virtualization itself needs to be accounted for. There are options in this document that provide for high availability of Red Hat Virtualization Manager, and if there is any notable service-level agreement (SLA) required for Red Hat Virtualization Manager, then one of those options (referred to as self-hosted engine, hosted engine, or HE for short) should be used.

Regardless of SLA requirements, Red Hat Virtualization needs to be tested in each environment. The product is rigorously tested by engineers, developers, quality assurance (QA), and many other individuals and teams at Red Hat, but that is no substitute for seeing how it will interact with each specific datacenter's configuration. Start with a subset of the expected production design and work out from there.

Finally, document absolutely everything, including:

- How are hosts deployed in your environment?
- How are VMs provisioned in your environment?
- How does application x get deployed in production?
- How does application y get updated in production?
- How does a developer get a new test environment for a week?
- How does that test environment get cleaned up?
- Where are IP addresses, DNS names, VLANs, and other network information documented?
- What are the security standards for Red Hat Enterprise Linux® VMs?
- What are the security standards for Red Hat Enterprise Linux virtualization hosts?

Essentially, any system administrator, engineer, architect, developer, or manager should be able to access the documentation (online or otherwise) and understand how to get things done. Especially when things go wrong.

RED HAT VIRTUALIZATION MANAGER

The purpose of this section is to provide guidance and consideration when deploying Red Hat Virtualization Manager.

RED HAT VIRTUALIZATION MANAGER

There are several options for how Red Hat Virtualization Manager, also referred to as the "engine," is deployed depending on technical requirements, business requirements, and available hardware.

Red Hat Virtualization Manager can be deployed in three primary ways:

- **Standard.** Red Hat Virtualization Manager is deployed on a Red Hat Enterprise Linux 7 host. That host can be a VM or a bare-metal host. This deployment method is perhaps the most simple conceptually, but it does not provide any means of high availability. It is fully supported for production environments, but compared to the self-hosted engine configuration, it may be more appropriate for lab environments, development environments, or other use cases that do not have the highest SLAs.

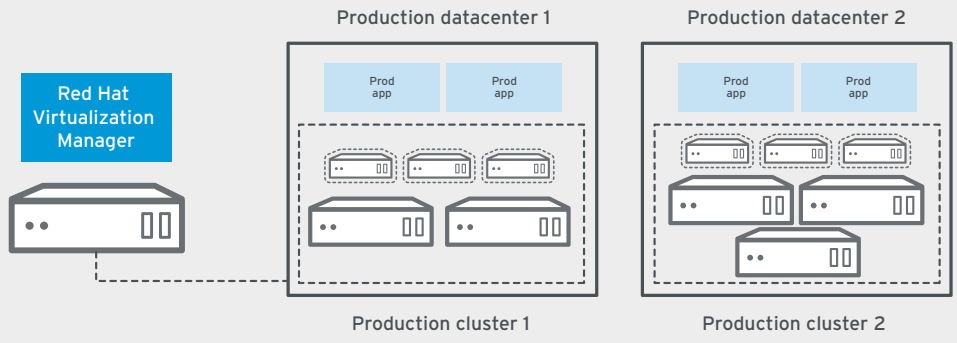
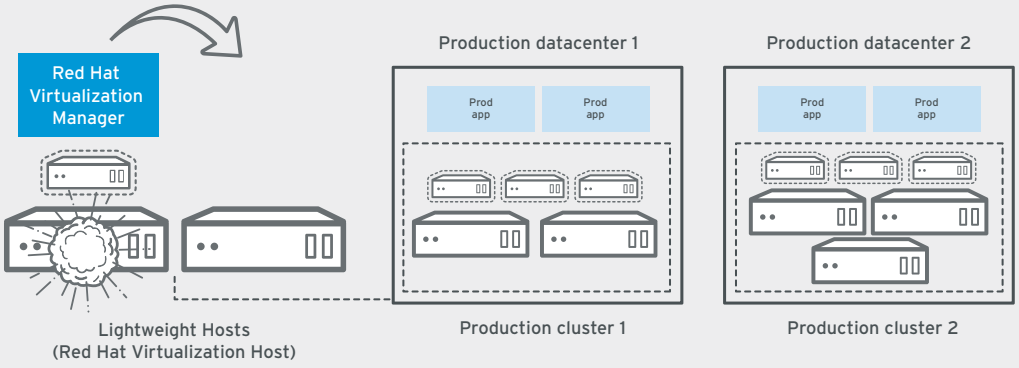


Figure 1. Standard Red Hat Virtualization deployment

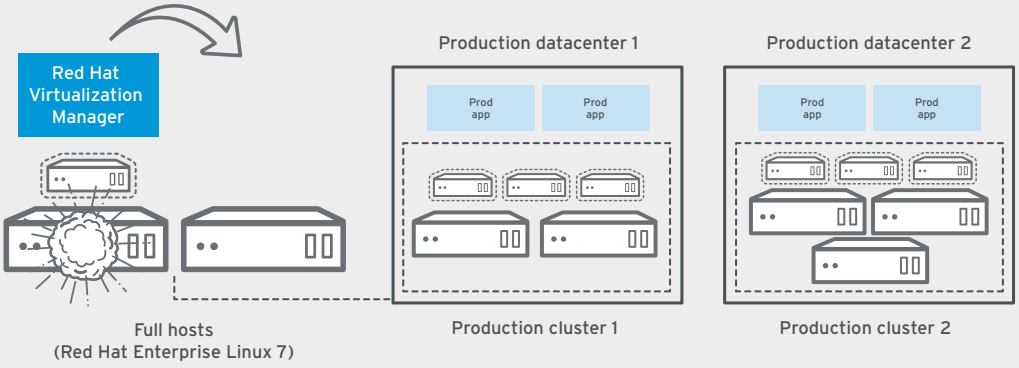
- **Self-Hosted engine with Red Hat Virtualization Host**. In this deployment configuration, Red Hat Virtualization Manager is deployed as a VM appliance on Red Hat Virtualization Host, a highly available (HA) cluster of lightweight hosts. The benefits are that Red Hat Virtualization Manager is provided with HA, and the overall hardware footprint is reduced. Additionally, the hosts are preconfigured and pretuned specifically for virtualization. Additional benefits of Red Hat Virtualization Host are explained later in the document.



Red Hat Virtualization Manager in self-hosted engine deployment
(provides high availability for Red Hat Virtualization Manager)

Figure 2. Red Hat Virtualization Hypervisor engine

- **Self-Hosted engine with Red Hat Enterprise Linux.** In this deployment configuration, Red Hat Virtualization Manager is deployed as a VM appliance on a highly available (HA) cluster of Red Hat Enterprise Linux 7 hosts. The benefits are that not only is Red Hat Virtualization Manager provided with HA, but the overall hardware footprint is reduced. Weighing Red Hat Enterprise Linux hosts against Red Hat Virtualization Host is explained later in this document.



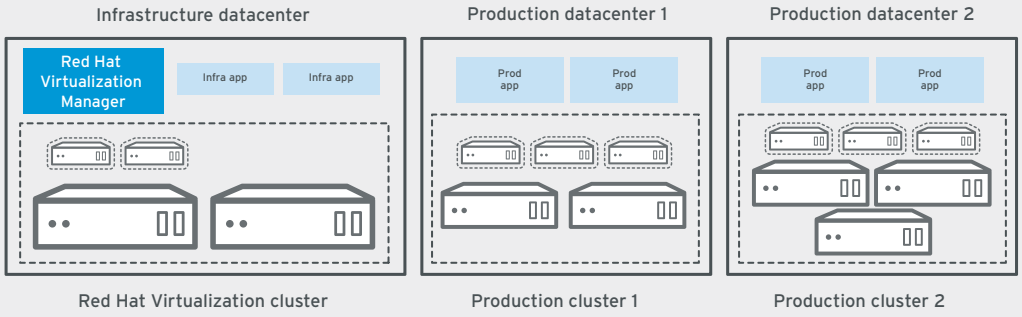
Red Hat Virtualization Manager in Linux-hosted engine deployment
(provides high availability for Red Hat Virtualization Manager)

Figure 3. Red Hat Enterprise Linux-hosted engine

Regardless of the deployment method chosen for Red Hat Virtualization Manager, a full backup should be done as soon as the Red Hat Virtualization Manager deployment is complete. That backup should be stored in a separate location from Red Hat Virtualization Manager. Regular backups of the engine, Red Hat Virtualization Manager, should be done thereafter.

Administrative or control plane datacenter

Regardless of the deployment choice for Red Hat Virtualization Manager, it is highly recommended to create a separate datacenter just for Red Hat Virtualization Manager and other infrastructure-level services. This separation will help facilitate backup schedules, performance, availability, and security.



Red Hat Virtualization Manager and other infrastructure services exist in a separate Red Hat Virtualization datacenter and cluster, segregated from other workloads.

Figure 4. Separate Red Hat Virtualization datacenter

Highest performance Red Hat Virtualization Manager

In most use cases, running Red Hat Virtualization Manager as an all-in-one (default) application is the preferred approach. However, in some scenarios, it may be desirable to have Red Hat Virtualization Manager run as fast as possible. For this, there is an option to break certain components out for performance reasons.

For example, it is possible to separate items such as the PostgreSQL database, the data warehouse, and the websocket proxy to other hosts. Note that because the hosted engine configuration uses an appliance version of Red Hat Virtualization Manager, the databases can only be made external post-installation; the websocket proxy will remain on the appliance. Also note that high-performance storage, such as solid-state drives (SSDs), should provide the performance required without separating the database. This is important to consider because having an external database host means having an additional host to provide high availability for as well as potential network latency.

Directory services

The use of directory services is highly recommended because it greatly enhances the capabilities of role-based access control (RBAC) within Red Hat Virtualization Manager. Additionally, user and group management is simplified, from both an operational and administrative standpoint. Operationally, having a single source of truth for users and groups is a best practice, regardless of the environment type or workload. Administratively, once the directory service is attached to Red Hat Virtualization, the process of adding users or groups from that service is a simple one.

The list of directory services that Red Hat Virtualization supports is extensive. And while one specific directory service is not recommended over another in this document, the most commonly used services encountered include Microsoft Active Directory, Red Hat Identity Management (IdM), and Red Hat Directory Server.

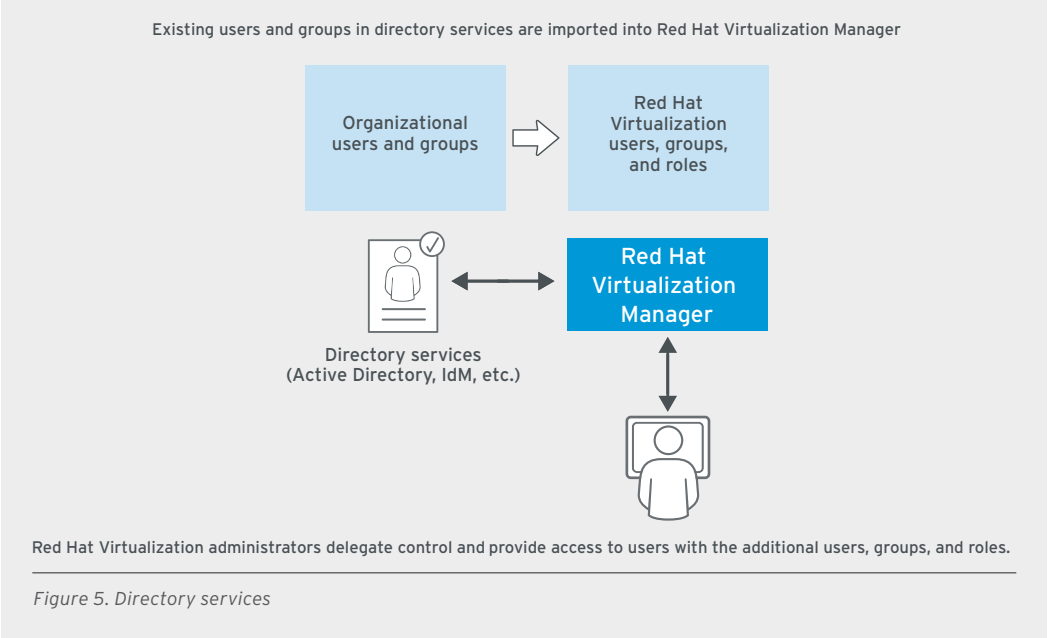


Figure 5. Directory services

Recommended users and roles

In addition to the administrator user (super user) for Red Hat Virtualization, these additional users and roles are recommended for various environments:

- **Power user.** A power user can connect to, manage, and monitor virtual resources. This role can be assigned to one or more users to offload certain administrative tasks.
- **Datacenter admin.** A datacenter admin can perform all duties except storage-related changes. By easily changing the default setting, datacenter admins can make storage-related changes. This role can be assigned to one or more users to create "tenant administrator" type delegations for assigned datacenters.
- **Cluster admin**. A cluster admin can only manage resources in assigned clusters. This role is effective for operator-level users that need to perform duties such as taking snapshots and editing, adding, or removing virtual resources such as disks.

To ensure proper tracking of activities for security and compliance, it is paramount that the individual administrator accounts be created as opposed to many people using the default admin account.

Other required network services

- **Network time protocol (NTP).** NTP should be used on all hosts and virtual machines in the environment to synchronize time. Authentication and certificates are particularly sensitive to time skew.
- **Domain name service (DNS).** DNS is required for Red Hat Virtualization Manager. Previous versions of Red Hat Virtualization allowed for simple name services by way of the /etc/hosts file, but full backward and forward resolution is *required* for Red Hat Virtualization 4. Additionally, hosts used in the hosted engine will also require fully resolvable DNS (forward and reverse).

SECURITY RECOMMENDATIONS IN *RED HAT VIRTUALIZATION*

There are security considerations in three different areas of the Red Hat Virtualization environment:

- Red Hat Virtualization Manager
- Red Hat Virtualization hosts (full and lightweight)
- Red Hat Virtualization guests

It is important to not disable the Red Hat Virtualization security features that are built into these respective areas. For example, in Red Hat Virtualization, the default connection method for the portal is via secure HTTP/HTTPS. Do not attempt to circumvent this. Use an external directory service to delegate authority to discrete admin accounts to avoid people sharing a single account. The underlying host for Red Hat Virtualization Manager should also be properly secured with IPtables and SELinux.

Regardless of whether full or lightweight hosts are used, the firewall and SELinux will be configured automatically. Do not disable either service. Add to the existing rules or controls if additional security is needed.

Red Hat Virtualization guests should also maintain some form of host-level firewall and mandatory access control, such as IPtables and SELinux in the case of Red Hat Enterprise Linux guests. In addition to a host-level firewall, Windows guests should also have an up-to-date antivirus program installed.

All hosts and Red Hat Enterprise Linux guests should be registered to either the Red Hat CDN or Red Hat Satellite in order to receive the latest security updates and errata.

VIRTUALIZATION HOSTS

This section will focus on recommendations and considerations for virtualization hosts. Red Hat Virtualization 4 supports the Red Hat Virtualization Hypervisor and Red Hat Enterprise Linux 7 as lightweight and full hosts, respectively.

STANDARDIZING ON HARDWARE

It is highly recommended to standardize on the same make and model server within the same Red Hat Virtualization cluster. This is for consistent performance results and ease of troubleshooting.

The importance of having consistent performance results in the same cluster of hosts cannot be overstated, especially in production but also in development and staging. A virtualized workload migrated from one host to another is expected to perform equally well and for that to happen in any service-level agreement, the underlying hardware needs to be the same. In most cases, the make and model of the servers within a cluster of hosts should be the same.

As much as possible, in addition to the same make and model of the server, the CPU families should be the same, the network interfaces should be the same, the host bus interfaces (HBA) should be identical, including firmware, RAID cards, and any other hardware that is common among the hosts. For production workloads, it is not recommended to mix different makes of servers within the same cluster because it does not provide consistent performance from host to host.

With that said, there are valid use cases in which some hosts should be more or less powerful than others. Virtual machines within the cluster are then configured to have affinity or non-affinity to certain hosts. However, because those use cases typically involve ensuring that a tiered application is grouped in the same cluster but still properly resourced, the consistent results argument is still valid.

Because budgets are a major concern and consideration in most environments, it is extremely important to prioritize certain components and features. For hosts, the weighting is typically as follows, starting with the most important:

- **Memory.** In a virtualization environment, the amount of memory (RAM) is arguably the most important piece because it will limit the number of VMs that can be supported. Also, hosts with not enough RAM will force memory swapping, which will significantly degrade VM performance.
- **CPU count.** The number of CPUs (i.e., sockets, cores, threads, etc.) is more important than the actual chipset, but not as important as the amount of RAM. In other words, if having to balance budget, go for more RAM and select a slightly older CPU.
- **Storage interfaces.** Storage interfaces should be used in pairs in order to enable multipathing and remove/reduce single points of failure. This includes both host bus adaptors (HBA) for fibre channel SAN and hardware-based iSCSI initiators for iSCSI SAN. Additionally, HBAs within the same storage cluster should be of the same make, model, and firmware level.
- **Network interface cards.** Network interface cards (NIC) should not be overlooked. If your workload is important, so is your NIC. See the "Network Considerations" section.
- **Management interface/remote access/out of band (OOB).** A management interface to remotely power cycle the host is critical for virtualized workloads that depend on high availability. It eases troubleshooting, but more importantly, it is a requirement for fencing operations as part of the high availability feature set.

- **CPU chipset.** Unless there is a specific instruction set or feature in a new or newer CPU or chipset, a slightly older CPU can be chosen. This can save budget for more important requirements. However, while different CPU families of the same CPU make can be mixed in the same cluster, only the capabilities of the oldest chipset will be used.
- **Local disk.** The local disk requirements are minimal; they only need the space required by the host image and Red Hat Virtualization Manager appliance if deploying the hosted-engine configuration. Once the host is running, only log files and configuration changes are being written, which does not require a fast local disk. Even a RAID configuration is optional as network installs are fast, and VMs are easily configured to restart on other hosts. For applications that have an SLA with a higher level uptime, a RAID configuration on the host boot disks make sense in order to reduce the chance of the VMs going down. However, not all applications require that SLA level, so not all hosts require the RAID.

The priorities listed above are not absolute and may differ from site to site, customer to customer, and even application to application.

FULL OR LIGHTWEIGHT HOST

Whether to deploy virtualized workloads on the lightweight Red Hat Virtualization Hypervisor or the full Red Hat Enterprise Linux host has several considerations. Both full and lightweight hypervisor hosts are built on Red Hat Enterprise Linux 7 with KVM and have IPtables and SELinux preconfigured to protect the host and the VMs.

The lightweight host has these advantages over the full host:

- **Cost.** The lightweight host is included in the Red Hat Virtualization subscription, regardless of the level of support and regardless of how guest subscriptions are purchased. In contrast, using Red Hat Enterprise Linux hosts may incur or use additional subscriptions.
- **Ease of overall management.** The lightweight host is deployed as a single image, which streamlines the update process. The entire image is updated as a whole instead of packages being updated individually.
- **Prescribed package and service selection.** The lightweight host has been designed specifically to support and protect VMs. Only the packages and services needed to service and protect the VMs, or manage the hypervisor itself, are included. This streamlines operations and reduces the overall attack vector since unnecessary packages and services are not deployed and therefore cannot be exploited.
- **Preconfigured.** Because the lightweight host has been designed as a single image, with packages and services predefined and pre-tuned, it is already at a "recommended practice" state.
- **Tools.** The Cockpit web administration tool is pre-installed and it in turn includes extensions specific to Red Hat Virtualization, including a GUI installer for the hosted engine installer as well as VM monitoring tools.

The full host has these advantages over the lightweight host:

- **Highest level of customization.** If the host requires specific file system layouts, the full host may be a better fit.
- **Better for higher rate of updates.** If additional packages are required and frequent updates to the host are expected, the full host may be a better solution. While the lightweight host can accept additional packages outside of the base image, those packages must be re-installed when the image is updated. For once-a-quarter updates, it may not be an operational or administrative issue, especially if configuration management is used. However, for multiple-times-per-month updates, the full host may hold an operational advantage.

Mixing full and lightweight hosts

An environment can have a mix of full and lightweight hosts. Both are supported within the same environment and even within the same cluster. In a mixed-design environment, there should be a documented business or technical use case for it.

For example, a valid mixed-design use case is having the web front end of an application virtualized on lightweight hosts for quick deployment while having the virtualized database back end on full hosts with a heavily customized file system layout.

DEPLOYING HOSTS

This section highlights recommendations and considerations for designing environments and deploying hosts.

Lightweight hosts:

- All lightweight hosts supporting hosted-engine configurations must have fully qualified domain names along with full DNS resolution (forwards and backwards).
- Be sure to enable networking and configure at least one interface to greatly reduce further configuration post-deployment via the Cockpit administrative console.
- Automatic partitioning (if doing a manual installation) is highly recommended. However, for custom layouts keep in mind that Red Hat Virtualization Hypervisor requires "LVM Thin Provisioning" with "LVM Thin pools" and a separate /var partition of at least 5GB. The deployment will fail if you ignore these requirements.
- The use of "NetworkManager" in Red Hat Virtualization Hypervisor is not currently supported; use the static ifcfg files to manually configure additional interfaces.
- IPtables and SELinux are preconfigured with Red Hat Virtualization Host, the lightweight host. They must remain enabled.
- Automated deployment via PXE, Kickstart, Red Hat Satellite, Red Hat CloudForms, or Ansible by Red Hat is highly recommended.
- Configuring fencing devices at deployment time is a requirement for proper functionality of the VM high-availability feature.

Full hosts:

- All full hosts supporting hosted-engine configurations must have fully qualified domain names along with full DNS resolution (forwards and backwards).
- Be sure to install, start, and enable the Cockpit administrative console; it is not installed by default on full hosts.
- Be sure to enable networking and configure at least one interface to reduce further configuration post-deployment via the Cockpit administrative console.
- Automated deployment via PXE, Kickstart, Red Hat Satellite, Red Hat CloudForms, or Ansible by Red Hat is highly recommended.
- Configuring fencing devices at deployment time is a requirement for proper functionality of the VM high-availability feature.

DEPLOYMENT RECOMMENDATIONS AND CONSIDERATIONS FOR HOSTS

When deploying a host, there is an option to burn an ISO image to a DVD or USB key, and then to boot the server to the media. However, there is a more automated, streamlined, and repeatable approach. Although there are numerous options, we recommend three—or combination of two or more—of these:

- Red Hat Satellite
- PXE server
- SAN boot/cloning

This recommendation is based on many customer successes with it.

Red Hat Satellite. Red Hat Satellite allows administrators to deploy, configure, and maintain physical and virtual systems. It synchronizes the content from the Red Hat Customer Portal, provides life-cycle management, and many other features. The focus in this section is on the ability to use the following methods to deploy bare-metal hosts for Red Hat Virtualization:

- **Unattended provisioning.** Identify a host using a MAC address and the Satellite server provisions it using a PXE boot process. The MAC address of the bare-metal server is known and shared with the Satellite server.
- **Unattended provisioning with discovery.** New hosts use PXE boot to load the Satellite discovery service. This service identifies hardware information about the host and lists it as an available host to provision. The MAC address of the bare-metal server is not known, and therefore a special discovery phase is used to grab the MAC address.
- **PXE-less provisioning.** The ability to provision new hosts using a boot disk or PXE-less discovery image that the Satellite server generates.
- **PXE-less provisioning with discovery.** New hosts use an ISO boot disk that loads the Satellite discovery service. This service identifies hardware information about the host and lists it as an available host to provision.

The benefits of this option are that not only will the bare-metal host be deployed in an automatic and streamlined fashion, but the hosts will be managed post-deployment as well. No separate process is necessary to bring the host(s) into a managed process for software life-cycle management.

PXE server. Identify a host using a MAC address and the PXE server provisions it using a PXE boot process. This is separate from the above mentioned PXE configuration related to the Red Hat Satellite. It's a more traditional PXE configuration that allows for multiple choices to select from in regards to installing an operating system over a network.

SAN boot/cloning. If a SAN is already used in the datacenter, it can be used to centralize the boot disks of the hosts. The potential downside is that each host requires an HBA (FC HBA or iSCSI initiator NIC), but the benefits include no local disks, faster recovery times on hosts as compared to hosts on local disks, and the ability to potentially use the SAN to clone the boot disks as a means of streamlining operations. Many enterprise storage vendors provide cloning technologies that can be used to quickly copy files (boot disks in this case) that can be quickly zoned to new hosts so hosts can be created on demand. Additionally, many enterprise storage vendors offer storage-side snapshots that allow administrators to easily revert to an earlier state of a host.

As suggested, the these three options are not mutually exclusive. It's possible to have a design that includes hypervisors that are deployed by Satellite, over PXE, but boot from SAN.

SECURITY CONSIDERATIONS IN HOSTS

Entire volumes have been written on security. However, here are some Red Hat-specific considerations:

- **Red Hat Virtualization host.** Use the lightweight Red Hat Virtualization Host whenever possible. Compared to the full Red Hat Enterprise Linux host, the lightweight host has fewer packages installed and fewer services running, which reduces the overall security footprint. Additionally, because Red Hat Virtualization Host is deployed as an appliance, it is also maintained and updated as an appliance. A simple update command will result in the latest available Red Hat Virtualization Host image being deployed quickly, instead of individual packages being updated.
- **Red Hat Enterprise Linux host.** When deploying the Red Hat Enterprise Linux hosts, only install packages and services required to satisfy virtualization, performance, security, and monitoring requirements. Production hosts should not have additional packages such as analyzers, compilers, or other pieces that will unnecessarily add security risks.
- **Considerations for both hosts.** Both Red Hat Virtualization Hypervisor and Red Hat Enterprise Linux use SELinux and IPtables as a Red Hat best practice. Do not disable either. If a service needs to get through, learn the procedure or contact Red Hat support for assistance. Do not disable the security features.
- **Errata updating.** Subscribe all hosts to Red Hat CDN or Red Hat Satellite and keep them updated with the latest security patches. Monitor all errata for relevant updates.
- **Access control.** Do not share access. Limit access to the hosts and create separate logins. Do not create a single root login for everyone to use. Do not create untrusted users for Red Hat Virtualization hosts. Only Red Hat Virtualization administrators or operators need access to a Red Hat Virtualization host. Most administrative activities occur via Red Hat Virtualization Manager, so there is no need for application owners or developers to access the hosts directly.

NETWORK CONSIDERATIONS IN RED HAT VIRTUALIZATION

Arguably, storage and networking are the most important factors in virtualization when looking at overall performance, followed quickly by host memory. These items should be considered requirements for production workloads and highly recommended for most other environments:

- **Bonding.** Network interfaces on production hosts should be bonded, preferably using LACP. This will contribute to the overall availability of service as well as network bandwidth. If high availability is a concern or requirement, then bonding must be a configuration requirement. Note that while Red Hat Enterprise Linux and Red Hat Virtualization support all bonding modes (0-6), logical networks that support VM traffic require modes 1, 2, 3, or 4 only.
- **Ethernet.** 1GbE should only be used for management traffic. VMs and Ethernet-based storage should use 10GbE or even 40GbE if it is available. For any reasonable performance by applications in the VMs, 10GbE (at least) will be required for the VM traffic and any Ethernet storage traffic.
- **VLANs.** Virtual LANs should be used extensively to segregate different traffic types and make the best use of the 10GbE (or 40GbE) networks. This lets quality of service (QoS) be used to limit how the network resources are used. VLANs are required to properly manage and segregate traffic on 10GbE and 40GbE networks.

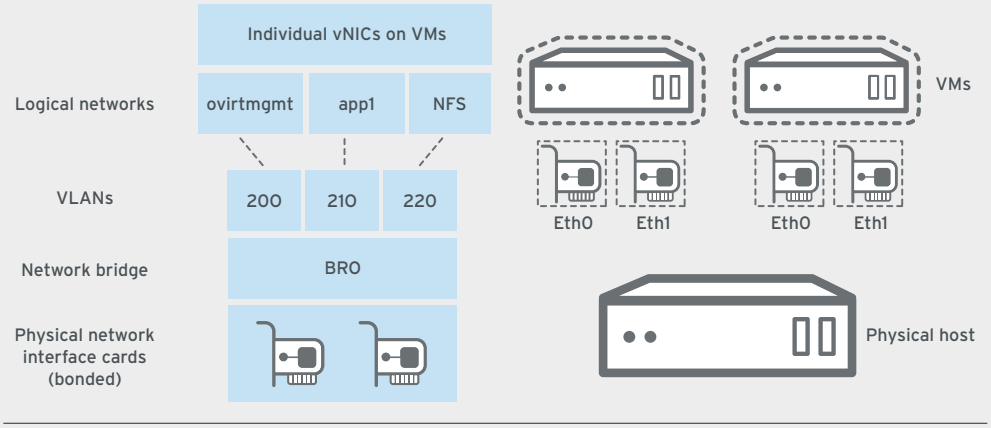


Figure 6. Logical networks

Bandwidth and VLAN considerations

With VLANs, it is necessary to determine priority and relative size of bandwidth requirements for production hosts carrying tier-one workloads. It is suggested that hosts in those environments have at least these four types of VLANs:

- Low-bandwidth monitoring network for communication with engine (Red Hat Virtualization Manager) and SPICE/VNC (VM consoles)
- High-bandwidth network for VM live migration
- Medium/high-bandwidth for user-to-VM communication (can be more depending on app requirements)
- High-bandwidth for Ethernet storage, if used

For the highest-availability SLA, fencing traffic should use completely separate hardware switches. Monitoring and fencing should never use the same switch, or that switch becomes a single point of failure for all high-availability capabilities.

Summary of hardware considerations for SR-IOV

When networking with PCI Express (PCIe) devices with single root I/O virtualization (SR-IOV), consider:

- The host firmware (BIOS or UEFI) must also support SR-IOV, and the function may need to be enabled in the BIOS or UEFI.
- Root ports, or ports immediately upstream of the PCIe device (such as a PCIe switch), must support alternative routing ID interpretation (ARI).

STORAGE CONSIDERATIONS IN RED HAT VIRTUALIZATION

Storage will have a huge impact on the overall performance of the virtual environment, so consider these storage-related items:

- **Bonding (iSCSI).** If using iSCSI or NFS, then multiple NICs should be used and bonded.
- **Host bus adapters.** Using a SAN, multiple HBAs (FC) or initiators (iSCSI) should be used to provide multiple paths to the SAN. When using HBAs (FC) or initiators (iSCSI), the same make, model, firmware version, and driver versions should be used in the same systems and clusters. This provides consistent and predictable performance, aids in troubleshooting, and allows usable baseline testing. Different makes and models can be used throughout the datacenter, just not in the same host or cluster.
- **SAN boot.** If there is already a SAN available to store VMs, then consider using SAN boot. This can save resources (no local disks) and the hypervisor images are easily cloned on SAN, speeding up deployment times.

CONSIDERING STORAGE PROTOCOLS

Selecting the right storage platform and protocol is a critical design stage decision. If the underlying storage is not right, no amount of memory or CPU at the host or VM level will make up for it. Here are some high-level questions to consider about storage:

- *Will the storage need to be backed up or replicated?* If the data is worth keeping, it is worth backing up. Consider how fast it needs to be available again. Enterprise storage can typically be backed up and replicated off-site easily. If there is already a disaster recovery site, network requirements are likely already in place, too. Then, determine the level of replication and how often to make backups between sites.
- *Will the storage be used to replicate data, VMs, datastores?* Storing data is the simplest scenario. However, if offloading activities such as VM cloning or datastore cloning is a priority, then application programming interfaces (APIs), software development kits (SDKs), and integration are Red Hat Virtualization Manager considerations as well.
- *What types of workloads will the virtualization platform be supporting?* If the applications do not depend on one type of storage over another, then this is less of a consideration. However, if there is a directive to use a particular protocol or format, Red Hat Virtualization supports all storage protocols that are POSIX-compliant and also support direct I/O.

Here are some additional considerations from a financial perspective:

- *Fibre channel*. If there is already a large investment in something like FC, then it is financially unwise to suggest a design around iSCSI, especially if there is already a level of comfort with FC switches and HBAs.
- *NFS storage*. Likewise, if NFS is the storage of choice, it would be better to just show how to ensure the security and speed of NFS over 10GbE.

Here are some benefits of the major storage protocols that Red Hat Virtualization supports:

NFS

NFS is easy to manage, easy to administer, and easily works with Red Hat Virtualization. Production workloads require an enterprise-grade NFS server, unless NFS is only being used to store ISO images. NFS can be used by data, ISO, and export domain types. NFS versions 3 and 4 are supported by Red Hat Virtualization 4.

For security, configure NFS so that individual services use specific ports; firewalls such as IPtables can then monitor and allow only those services. Additionally, Ethernet storage traffic should be segregated by VLANs. When enterprise NFS is deployed over 10GbE, segregated with VLANs, and individual services are configured to use specific ports, it is both fast and secure.

As NFS exports grow to accommodate more storage needs, Red Hat Virtualization recognizes the larger datastore immediately. No additional configuration is necessary on the hosts or from within Red Hat Virtualization. This gives NFS a slight edge over SAN from a scale and operational perspective.

NFS must be used for production ISO datastores. For lab and sandbox environments, an internal (to Red Hat Virtualization Manager) NFS ISO domain can be used. Do not use the internal NFS for production.

SAN (FC/FCoE)

FC-based SAN is both fast and secure. If fibre channel-based SAN is already in use in the target datacenter, then it should be taken advantage of for Red Hat Virtualization data domains. However, because FC SAN requires specialized network gear and training, it is a large capital expense to make if it is not already in place. Additionally, fiber channel switches are not known to be easily automated.

FC/FCoE SAN is required when using direct LUNS. It also has the advantage of low CPU overhead as compared to iSCSI and NFS.

FC SAN is only usable for data domains; NFS is still needed for the ISO domains. However, if FC SAN is available, it should be seriously considered to SAN boot the hosts as well. Enterprise FC SANs frequently have cloning features that make deploying new hosts based on host templates very fast.

SAN (iSCSI)

iSCSI-based SAN is easy to manage, easy to administer, and easily works with Red Hat Virtualization. Note that production workloads require an enterprise-grade iSCSI server.

When enterprise iSCSI SAN is deployed over 10GbE, segregated with VLANs, and uses CHAP authentication, it is both fast and secure.

iSCSI SAN is only usable for data domains; NFS is still needed for the ISO domains. However, if iSCSI SAN is used, it should be seriously considered to SAN boot the hosts as well. Enterprise iSCSI SANs frequently have cloning features that make deploying new hosts based on host templates very fast. This requires a hardware-based iSCSI initiator for each host.

Red Hat Gluster Storage

Red Hat Gluster Storage is a POSIX-compatible and open source file system that supports storing VM images, snapshots, and templates. Three or more servers are configured as a Red Hat Gluster Storage server cluster instead of network-attached storage (NAS) appliances or a SAN array.

Red Hat Storage should be used over 10GbE and segregated with VLANs.

Red Hat Gluster Storage is only supported for data domains; NFS is still required for the ISO domains.

Important: Do not assume that a particular version of Red Hat Virtualization works with a particular version of Gluster. Always refer to the Red Hat Gluster Storage Version Compatibility and Support document to be sure. For example, Red Hat Gluster Storage (as of mid-October, 2016) does not include the version of VDSM supported by Red Hat Virtualization 4. Red Hat Virtualization 4 can still be used; however, it requires that datacenters using Gluster use a compatibility version of 3.6.

Local storage

Local storage should only be considered for small lab environments. Do not use local storage for production workloads. It will preclude the use of live migration, snapshots, and the flexibility afforded by virtualization.

VIRTUAL RESOURCE CONSIDERATIONS

The purpose of this section is to provide guidance on the compute, network, and storage resources managed by Red Hat Virtualization Manager. The focus is on the virtual aspects of the environment.

Hosts, clusters, and datacenters

It is important to understand the terminology and the implications of how the objects in a Red Hat Virtualization environment fit together. This has a large impact on the overall design of the environment, the policies used to schedule resources, how storage is attached and defined, and how VMs are deployed.

While it is easy to deploy Red Hat Virtualization Manager and a few hosts, this deeper knowledge assists in making the best use of available resources.

Host

A host is a physical server running either Red Hat Enterprise Linux 7 (full) or Red Hat Virtualization Host 7 (lightweight) and is registered to Red Hat Virtualization Manager with the sole purpose of supporting VMs. Both full and lightweight hosts use full hardware virtualization by way of Kernel-based Virtual Machine (KVM).

Hosts are grouped into clusters where VMs can be dynamically allocated. Hosts must be part of a cluster, and a cluster part of a datacenter before VMs can be deployed.

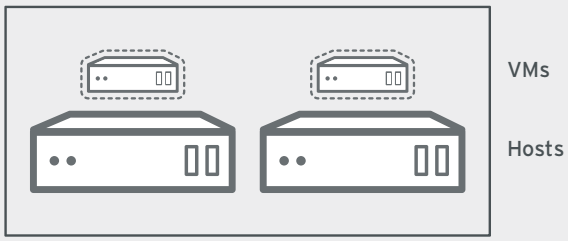


Figure 7. VMs on a host

Cluster

A cluster is a group of hosts that share the same make of CPU. If they do not share the same chipset, they will be stepped down to the lowest common set of features. A cluster belongs to a datacenter, and there may be multiple clusters in a datacenter. VMs can only live migrate between hosts of the same cluster. Because clusters in the same datacenter share the same underlying storage, it may be a design decision to build an application at the cluster level.

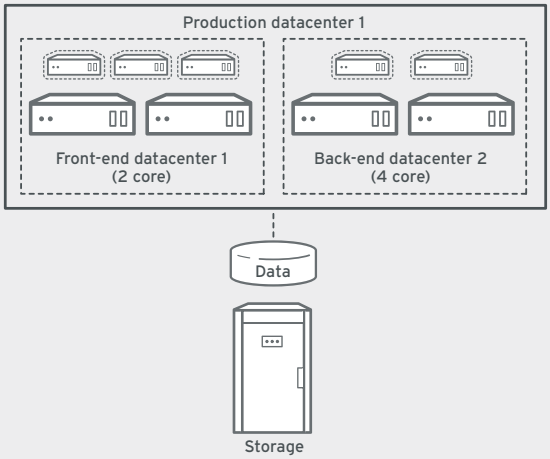


Figure 8. Hosts in a cluster in a datacenter

Datacenter

A datacenter is a logical container for clusters, hosts, storage, and networks. A Red Hat Virtualization datacenter can contain one cluster or many clusters. Red Hat Virtualization can also support many datacenters that contain many clusters.

One of the defining features of a Red Hat Virtualization datacenter is that all clusters within the same datacenter share the same underlying storage. However, different Red Hat Virtualization datacenters maintain separate storage, even within the same Red Hat Virtualization environment. This has a big impact on design decisions.

DESIGN CONSIDERATIONS FOR DATACENTERS AND CLUSTERS

Here are some considerations for designing or deploying datacenters and clusters in Red Hat Virtualization:

- Red Hat Virtualization datacenter layer
  - Use Red Hat Virtualization datacenters to segregate customers that are not allowed to share data or resources.
  - Use Red Hat Virtualization datacenters to separate applications that should not be on the same back-up schedule.
- Red Hat Virtualization cluster layer
  - Use Red Hat Virtualization clusters to segregate application components (two-core CPU from four-core CPU, or front-end from back-end, etc.).
  - Use Red Hat Virtualization clusters to create natural live migration barriers or to use affinity/anti-affinity.
  - Use Red Hat Virtualization clusters to group similar, but not identical, hardware together for investment protection.

Example multitier application on Red Hat Virtualization

The front-end and middleware VMs of an application can exist in a cluster of two-core hosts, while the back-end database of that same application's VMs may exist in a cluster of four-core hosts. This provides a natural means of segregating the workloads; however, the underlying storage is still shared allowing for back-up schedules to still be properly coordinated.

Example multitenant in Red Hat Virtualization

Multiple lines of business that cannot share data may exist alongside each other. Therefore, they cannot be on the same storage or networks. Placing them on separate datacenters with separate VLANs will achieve the same results as physically separating them because datacenters do not share storage or logical networks.

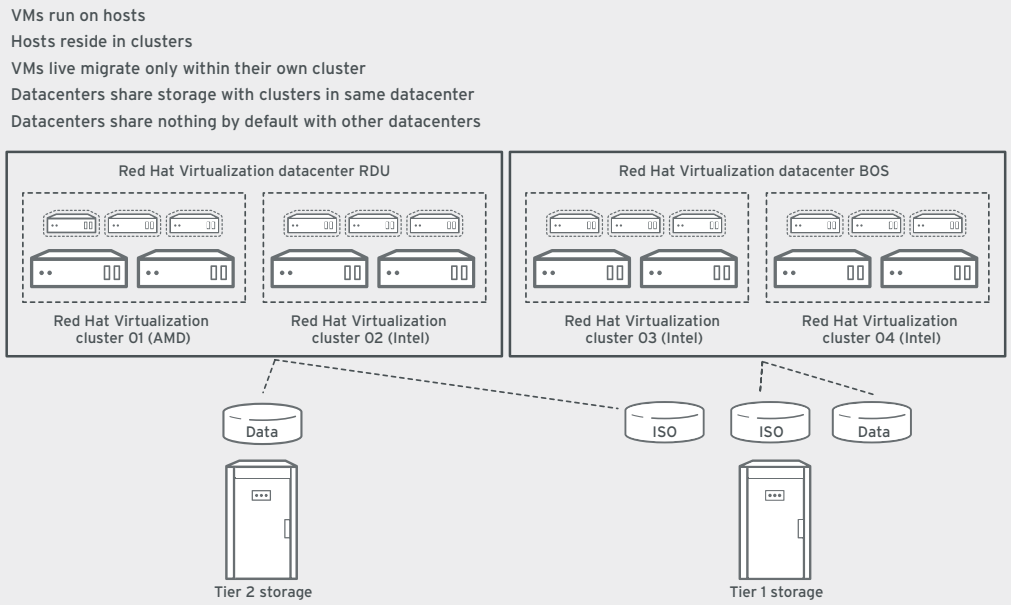


Figure 9. Two datacenters

COMPUTE (VMs)

Building a VM should not be just a copy or abstraction of a physical server, because that negates many of the benefits associated with virtualization such as faster deployment and provisioning, faster backups, faster recovery, reduced overall storage, and ease of management.

VM templates

The design of the environment must take into consideration how the VMs and the applications will be deployed. It is highly recommended that a combination of tools be used to automate and orchestrate the deployment and configuration of both VMs and the their applications.

The workflow for deploying an application must be completely understood and documented. Whether that application is backed by a single VM or multiple VMs, it should be automated. Once the VM is deployed with all of the patches and configurations that are safe to be cloned, any configurations that are not safe to be cloned (e.g., SSH host keys, MAC addresses) must be removed. A template should then created from this "golden image."

Configuration tools (e.g., Ansible, Puppet) can be used to automate the deployment of the application. And as recommend earlier, Red Hat Satellite should be employed to manage the software life cycle and security patching. However, a self-service portal (e.g., Red Hat CloudForms, ServiceNow) should also be employed to orchestrate all of the moving parts—the Red Hat Virtualization template, host configuration, application configuration, and automated patch management.

To reap all of the benefits of virtualization, VMs design should include thinking through automation, repeatability, streamlined operations, recovery, and retirement. It is highly recommended that VM design be done with templates, configuration management, software life-cycle management, and self-service.

VM disks

A VM should have its own local boot disk. The VM operating system and application binaries will be on a virtual disk, managed by Red Hat Virtualization and presented to the VM, which is perceived as a local disk drive to the VM. It is highly recommended that any additional disks required for application data be managed externally on either an NFS share or an iSCSI disk and accessed directly by the VM.

This method provides multiple benefits:

- The VM can be created as a template.
- The application can be backed up separately from the central storage (e.g., SAN, NAS).
- Performance is significantly better for the application than having the application disk also managed by Red Hat Virtualization.
- The VM can be rebuilt quickly from template or snapshot, then the application data remounted from the SAN or NAS.
- Overall operations are faster because new VMs can be created on demand from a template rather than individual VMs per specific workload or application.

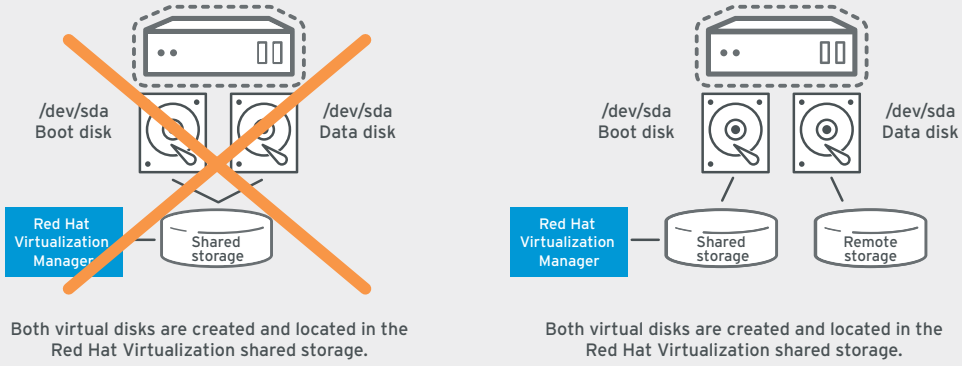


Figure 10. Local VM disk and remote app disk

Provisioning thick or thin virtual disks

Whether to provision thick or thin virtual disks depends on the application use case. If the purpose of the VM is to be a web front end to other servers, it most likely will not write much data other than log files, making it a prime candidate for thin provisioning even if it is a production server. Even a tier-one database server could potentially get away with being thinly provisioned if all of its tablespaces, database log files, and other important data were remotely accessed over NFS or iSCSI as opposed to being local to the VM

However, there is a performance consideration for thin provisioning. As the underlying storage is allocated, a wait time is involved. If the VM is only writing to log files and the occasional configuration file, then the performance is negligible.

Thin provisioning is faster from a deployment, backup, replication, and recovery standpoint. Thick provisioning is faster from a performance standpoint, but takes more space. If the underlying storage is able to deduplicate data in the VMs, then the recommendation is to thin provision the underlying storage, enable deduplication, and thick provision the VMs.

The best way to know for sure is to test in your environment, on your servers, with your applications.

Provisioning direct LUNs

Using Direct LUNs is supported by both Red Hat Enterprise Linux and Windows VMs and requires use of VIRTIO drivers. Red Hat Enterprise Linux 5 and above include the required drivers automatically. Windows VMs require a separate installation from either the tools ISO or virtual floppy disk (VFD).

Using LUNs directly as the backing for the VM disk images removes a layer of abstraction between the VM and the data underneath. These considerations must be made when using a direct LUN:

- Live storage migration of direct LUN hard disk images is not supported.
- Direct LUN disks are not included in VM exports.
- Direct LUN disks are not included in VM snapshots.

NETWORK (LOGICAL NETWORKS)

The network recommendations for VMs are an extension of those given for the physical hosts. Essentially, segregate all of the traffic possible. This allows for better security and easier troubleshooting. A misbehaving piece on one network (VLAN) will only affect neighbors on that particular network (VLAN).

Logical networks are created at the datacenter layer, but ultimately managed at the cluster layer. This provides a sound level of control and access for VMs and applications and a natural barrier when datacenters are used as tenants. Logical networks provide flexibility in the configuration of the physical devices.

Considerations when designing or deploying logical networks in Red Hat Virtualization include:

- Red Hat Virtualization datacenter layer
  - If additional physical interfaces are added to a host for storage use, uncheck VM traffic so that the VLAN will be assigned directly to the physical interface. No logical interface will be used and only storage traffic will be allowed on that VLAN.
  - Create a VLAN just for management traffic and assign it to the interface with the smallest bandwidth (e.g., onboard 1GbE). There is no need to waste a VLAN on a 10GbE or 40GbE for management traffic.
  - Create a VLAN for each Ethernet storage connection. If possible, have a single VLAN for all NFS traffic and another one just for iSCSI traffic. Create individual NFS VLANs and individual iSCSI VLANs. The exception here would be for the NFS ISO domain.
- Red Hat Virtualization cluster layer
  - Create VLANs for each application.
  - Create a VLAN for external traffic.
  - If Red Hat OpenStack® Platform is deployed, use its Neutron integration to have SDN capabilities (Open vSwitch). This allows for ease of administration because all SDNs (Red Hat OpenStack Platform and Red Hat Virtualization) can be managed from Red Hat OpenStack Platform (web or API).

STORAGE (STORAGE DOMAINS)

Storage domain recommendations include:

- **Data domains.** Data domains are the most important datastore type in Red Hat Virtualization because they store the VM as well as the VM snapshots. Data domains may be backed by SAN (iSCSI or FC/FCoE), NFS, Red Hat Gluster Storage, or POSIX-compliant file system that supports direct I/O.

  Performance in virtualization starts with storage, so matching the storage tier to the application is very important. It is also common to use multiple storage tiers within a Red Hat Virtualization environment and to attach them to the appropriate Red Hat Virtualization datacenters.

- **ISO domains.** ISO domains are used to store ISO images and VFD. ISO images are used to deploy operating systems and VFDs are used to deploy device drivers. Even if VM deployments will be done over the network, an ISO domain is highly recommended as both a backup deployment

method and a means to slipstream device drivers when converting VM images from other virtualization platforms to Red Hat Virtualization. ISO domains currently only support NFS; however, a lower-tier NFS server will suffice since performance is not a factor.

- **Export domains.** Export domains are used as transient/temporary storage as data is moved between Red Hat Virtualization datacenters or environments. Export domains currently only support NFS; however, a lower-tier NFS server will likely suffice since performance is not typically a factor for export domains.

Storage design or deployment recommendations with Red Hat Virtualization include:

- Match the storage tier to the application SLA. Place your tier 1, virtualized relational database on EMC, NetApp, Hitachi, or similar. An ISO domain that gets used twice a month, for example, can safelybe on a third-tier NFS.
- If all VM provisioning will be done from PXE, or some other external source, you may not need an ISO domain.
- If you expect to import VMs from VMware, including Windows guests, the drivers are on virtual floppy disks (VFD) images, and should be kept in the ISO domain.
- If using Red Hat OpenStack Platform, use Glance integration to share cloud OS images. They come up very fast on Red Hat Virtualization and can be critical to reacting on demand.

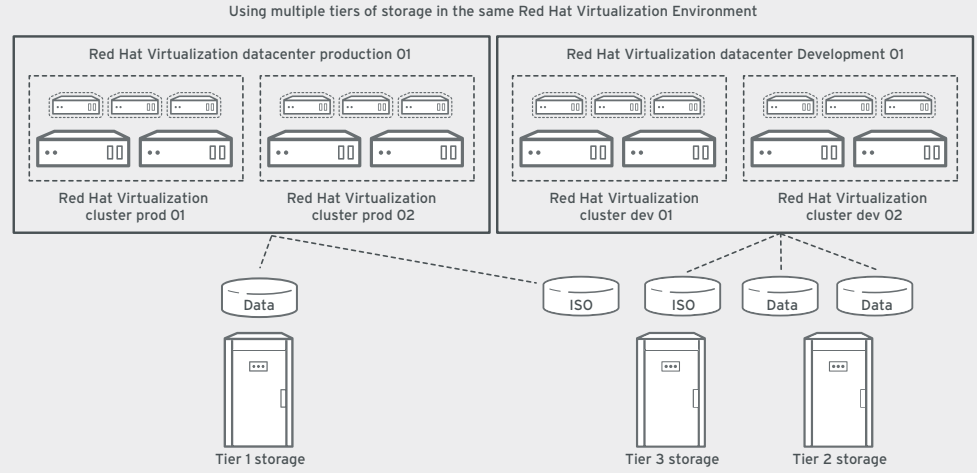


Figure 11. Storage tiers in Red Hat Virtualization

ADDITIONAL REFERENCES

All Red Hat Virtualization 4 Documentation
https://access.redhat.com/documentation/en/red-hat-virtualization/

Red Hat Virtualization 4 Installation Documentation
https://access.redhat.com/documentation/en/red-hat-virtualization/4.0/paged/installation-guide/

Red Hat Virtualization 4 Self-Hosted Engine Documentation
https://access.redhat.com/documentation/en/red-hat-virtualization/4.0/paged/self-hosted-engine-guide/

Red Hat Virtualization Technical Reference
https://access.redhat.com/documentation/en/red-hat-virtualization/4.0/paged/technical-reference/

OTHER RED HAT STORAGE RESOURCES

It is highly recommended to be familiar with the Red Hat Storage Administration Guide:
https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/7/html/Storage_Administration_Guide/

The official guide for configuring storage in Red Hat Virtualization can be found at:
https://access.redhat.com/documentation/en/red-hat-virtualization/4.0/single/administration-guide/#chap-Storage

RHGS Compatibility Matrix:
https://access.redhat.com/articles/2356261

ABOUT RED HAT

Red Hat is the world's leading provider of open source software solutions, using a community-powered approach to provide reliable and high-performing cloud, Linux, middleware, storage, and virtualization technologies. Red Hat also offers award-winning support, training, and consulting services. As a connective hub in a global network of enterprises, partners, and open source communities, Red Hat helps create relevant, innovative technologies that liberate resources for growth and prepare customers for the future of IT.

NORTH AMERICA
1 888 REDHAT1

EUROPE, MIDDLE EAST, AND AFRICA
00800 7334 2835
europe@redhat.com

ASIA PACIFIC
+65 6490 4200
apac@redhat.com

LATIN AMERICA
+54 11 4329 7300
info-latam@redhat.com

facebook.com/redhatinc
@redhatnews
linkedin.com/company/red-hat

redhat.com
#f7659_0617