USING RED HAT NETWORK SATELLITE TO DYNAMICALLY SCALE APPLICATIONS IN A PRIVATE CLOUD

ABSTRACT

Private cloud infrastructure has many clear advantages, not the least of which is the decoupling of an application from its physical resources. Once a private cloud is successfully implemented, one of the most significant advantages is the ability to dynamically scale a given application based on demand. By implementing Red Hat Network Satellite in a private cloud, organizations can create a dynamically scalable application environment.

INTRODUCTION

Cloud computing, using resources based outside the enterprise, offers a model of on-demand computing that can work equally well inside the corporate firewall and exclusively on internal resources. Users of Red Hat® technologies can build an internal or private cloud and configure Red Hat Network Satellite to seamlessly manage resources to meet demand changes—reacting to changes swiftly, in minutes. This private cloud can allow an organization to fully use its resources avoiding costly overbuilding of the hardware infrastructure.

For example, the Apache web server is the most popular web server on the Internet largely because of its high availability and scalability. Since April 1996, Apache has been the most popular HTTP server software in use. As of February 2010, Apache served over 54.46% of all websites and over 66% of the million busiest sites.[1]

Using Apache in a private cloud computing environment offers organizations the opportunity to increase the availability and scalability of Apache by dynamically adding and removing servers based upon demand. Using the features of Red Hat Network Satellite, system administrators can achieve this level of automation at a fraction of the cost of attempting this in a traditional computing environment.

NOTE: While Apache is used as the example application in this scenario, these same techniques could be applied to a wide variety of applications. Apache was chosen because of its wide deployment and familiarity.

To be able to automate application deployment in a private cloud computing environment, it is critical that an organization have the proper management tools—tools to automate as many tasks as possible, resulting in a more consistent and efficient environment. At its core, cloud computing is about being able to respond to changes in demand seamlessly. For example, if an additional instance of Apache is needed because of unexpected demand, the ideal system would provision it dynamically at that time rather than having a system administrator build it manually.

Instead of dynamically scaling its environment, most organizations simply over-subscribe their physical environments in terms of hardware and software in order to meet peak demands that may happen infrequently. While overbuilding an environment can address the problem, it can also be extremely costly to implement, manage, and maintain just to handle brief and infrequent peaks. Furthermore, no amount of hardware can meet fully unexpected demand. An example in the case of Apache might be a sudden breakthrough and massive media attention for a particular site causing too much demand on a physical or virtual server.

1 "February 2010 Web server Survey". Netcraft. February 22, 2010. http://news.netcraft.com/archives/2010/02/22/february_2010_web_server_survey.html.

Overbuilding creates excess capacity that will sit unused while waiting for bursts of activity. In addition, this excess capacity still has to be managed and be compliant with all internal and external standards, even if it is just sitting idle. In cloud computing, an application like Apache can be scaled up to meet peak demand and then those computing resources can be released for other applications to use. Organizations that build internal or private clouds can address unexpected demand with a much smaller hardware footprint and at much more manageable costs.

Consider an organization that derives the majority of its revenue from e-commerce sales. Their website needs maximum uptime and must process orders rapidly, regardless of the demands on the servers. If a consumer cannot complete a purchase in a timely manner, they'll look elsewhere, and the sale is lost. Rather than face the loss of revenue, many organizations invest in massive hardware server farms or pay for expensive external processing capabilities to ensure that peak demand can always be met. Cloud computing offers a better and more economical way to accomplish the same goal by allowing organizations to pay for the increased capacity on demand rather than overbuilding and owning significantly more capacity than they need for day-to-day operations.

The foundation of every cloud is a robust, scalable, and secure virtualization substrate. Decoupling the application from the physical hardware allows the organization to provision what is required to meet demand but does not necessarily give that organization the flexibility to dynamically reallocate services based upon the changes to that demand. In order to do that, you need a management tool that allows you to dynamically provision new instances of an application as needed.

Private, or internal cloud computing, offers a model of on-demand computing that can work inside the corporate firewall exclusively using internal resources. Red Hat offers a set of solutions to the problems of building out, managing, and scaling up multiple applications, such as Apache, on that private cloud. Red Hat provides the tools necessary to configure and automatically manage resources to meet demand changes in a matter of minutes. In addition, Red Hat's approach can allow an organization to fully use its resources without a costly overbuild of the hardware infrastructure.

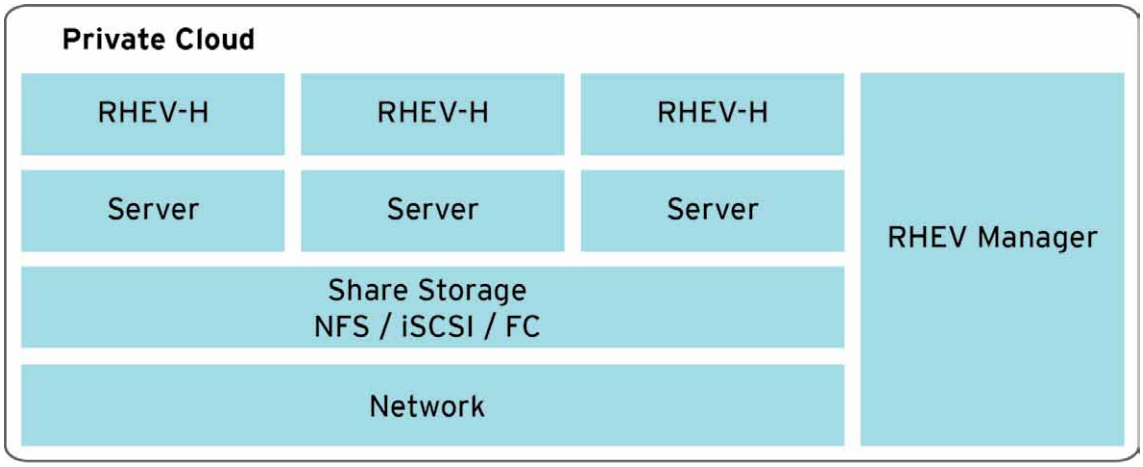
CONSTRUCTING THE PRIVATE CLOUD

The following example will describe a simple implementation. The example has been structured with a small number of servers for ease of explanation. The same principles can be applied to much larger scale implementations with hundreds of servers.

The first step is the implementation of Red Hat Enterprise Virtualization, which provides the virtualization substrate necessary for a private cloud.

This example system, as illustrated in Figure A, uses three Red Hat Enterprise Virtualization Hypervisors (RHEV-H) and a single Red Hat Enterprise Virtualization Manager (RHEV-M). NFS shared storage is configured to allow for live migration of virtual machines without interruption or downtime.

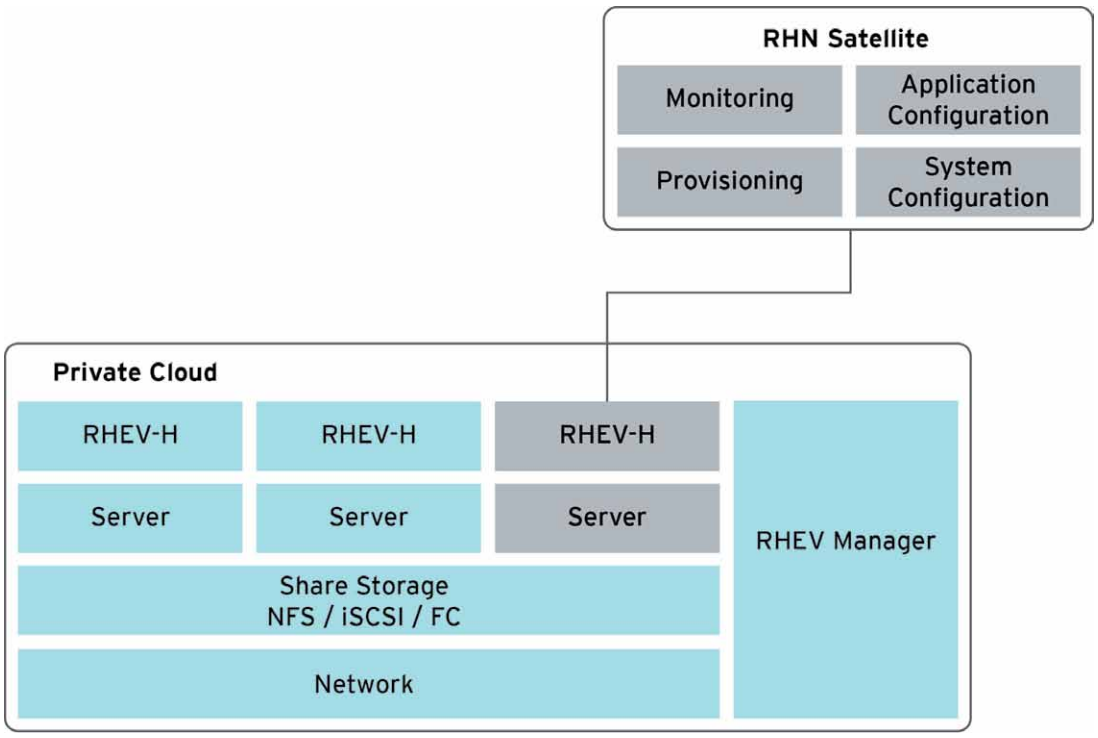
FIGURE A



The next step is to install an instance of Red Hat Network Satellite on one of the virtual machines. Red Hat Network Satellite will provide PXE/DHCP to allow for fast provisioning of guest operating systems without system administrator intervention in the Red Hat Enterprise Virtualization environment. In addition, Satellite will use its provisioning tools to create a profile and build the guest operating systems in an automated and repeatable manner.

Figure B illustrates the architecture after Red Hat Network Satellite is installed and configured.

FIGURE B



It is further recommended to make Red Hat Network Satellite highly available to reduce the time needed to recreate the environment by following the instructions in the High Availability Guide **(http://www.redhat.com/f/pdf/rhn/Satellite-HA.pdf)**.

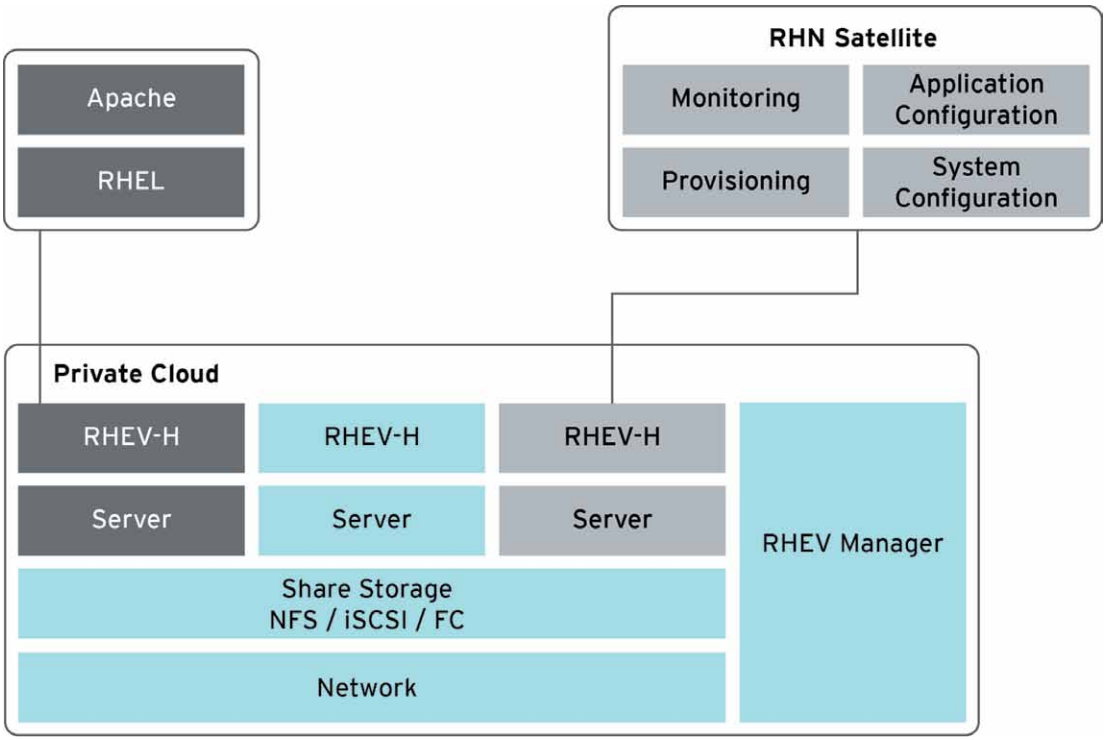
Creating an Apache web server as a virtual machine is the next step. Once again, Red Hat Network Satellite is used to provision this node.

As Figure C illustrates, the user will select a new server from the Red Hat Enterprise Virtualization-Manager interface and boot the new server via PXE. Then using the services provided by Red Hat Network Satellite, a node will be created with the following qualities:

- The operating system will be automatically installed
- The node will be registered with the Satellite server
- All of the necessary Apache packages will be installed
- All necessary configuration files will be deployed

An added benefit of using Red Hat Network Satellite to provision the Apache web server is that it allows the flexibility to provision servers on both physical and virtual servers, independent of the hardware or virtualization platform (IBM, HP, Dell, VMware, Red Hat Enterprise Virtualization, etc). This provides the ability to extend the web tier beyond a single platform while still using a single provisioning method.

FIGURE C



At this stage, the administrator can grow the web services layer quickly and simply. As illustrated in Figure D, the administrator can use PXE booting to create as many virtual machines as needed, limited only by the underlying Red Hat Enterprise Virtualization infrastructure. This illustration shows the creation of a second Apache server but could just as easily show creating five or more.

As the virtual servers are booted, they are automatically provisioned with Apache and configured to serve content. Furthermore, Red Hat Network Satellite will be given the ability to automatically access the RHEV-M API to create virtual machines that will be PXE booted. This will be done via SSH public key authentication and the use of the RHEV-M Powershell API.

With this added functionality, it is possible for Red Hat Network Satellite to control the number of web servers that are provisioned based on metrics obtained through monitoring. This allows for dynamic scaling of services based upon demand.

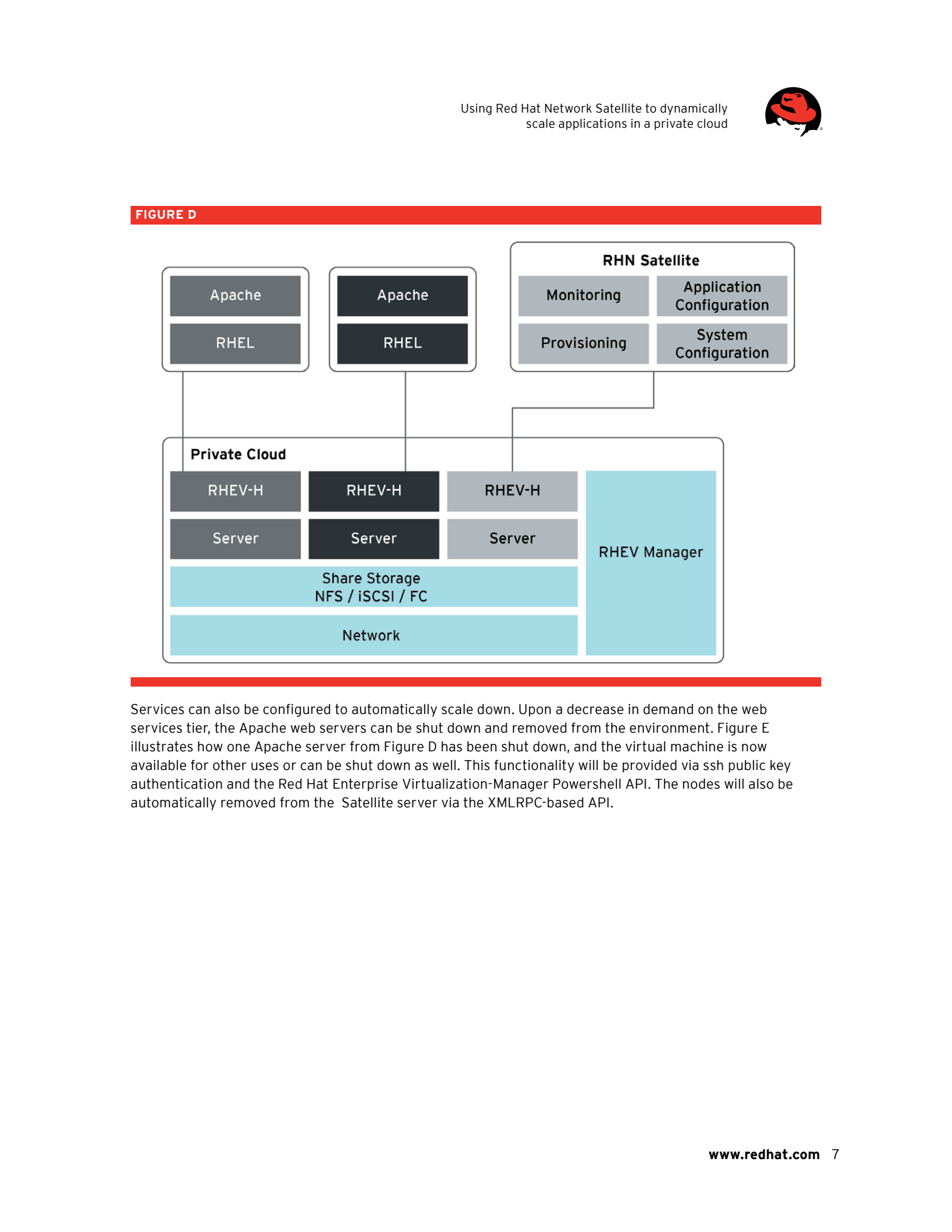
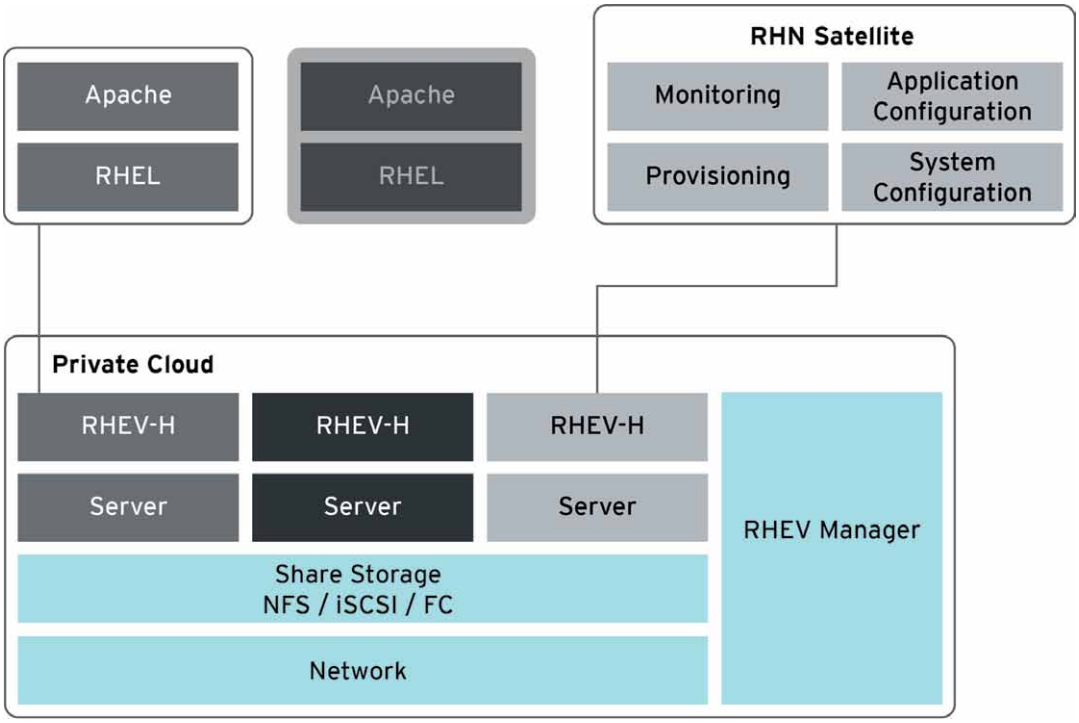


FIGURE D

Services can also be configured to automatically scale down. Upon a decrease in demand on the web services tier, the Apache web servers can be shut down and removed from the environment. Figure E illustrates how one Apache server from Figure D has been shut down, and the virtual machine is now available for other uses or can be shut down as well. This functionality will be provided via ssh public key authentication and the Red Hat Enterprise Virtualization-Manager Powershell API. The nodes will also be automatically removed from the Satellite server via the XMLRPC-based API.

FIGURE E



Using Red Hat Enterprise Virtualization and Red Hat Network Satellite, it is clear that any organization can build a dynamic and cost-effective virtual environment that can be structured to provide services when they are needed—even allowing for seasonal peaks—without significantly overbuilding the physical environment.

For example, an organization with strong e-commerce sales can predict when the bulk of its daily traffic occurs and use Red Hat Network Satellite to automatically provision additional Apache web servers at a particular time of day or a day of the week, based upon usage demand. In the case of unexpected demand, Red Hat Network Satellite can be configured to provision additional Apache servers based upon CPU utilization or in this case, specified TCP metrics. When the servers are no longer needed, they can be shut down and the resources allocated to other services.

To extend this example, consider this same organization that knows that once every year, for a period of three weeks in the beginning of December, its web traffic spikes to ten times the average daily traffic. In a physical environment, the organization would be required to purchase and install enough servers to meet this peak traffic. The remaining forty-nine weeks of the year, the servers would be massively overbuilt—a costly and inefficient approach. By using Red Hat Network Satellite on a private cloud, the organization would make a significantly smaller hardware investment and software subscription and structure their physical environment to provide enough hardware to support its typical need and allow the dynamic provisioning to temporarily shift resources to meet short-term demand.

Any metric monitored by Red Hat Network Satellite can be used to determine when to scale a service up or down. Within minutes, the enterprise environment can automatically scale itself to meet current demand. In a physical environment, this could take hours or possibly days. Even a private cloud environment without the use of Satellite would require manual provisioning, decreasing staff productivity and increasing the risk of human error.

OTHER ALTERNATIVES

While this level of dynamic provisioning is based on Red Hat technologies, other approaches can be used to build dynamically allocated virtual environments. Both proprietary commercial packages and open source solutions are available; however, neither provides the simple and comprehensive solution described in this brief using Red Hat technologies.

For example, open source solutions require numerous tools to accomplish similar benefits, as well as considerable staff time to "home brew" a solution. For example, a system administrator would need to use Puppet, Cobbler, Kickstart, Nagios, Eucalyptus, and a significant amount of custom scripting to structure a basic private cloud environment. Using this homegrown approach also means relying on community support rather than professional and certified support and training. In addition, Red Hat Satellite's graphical user interface (GUI) and ease of use makes installation and operation fast and inexpensive when compared to a do-it-yourself solution.

Commercial, proprietary solutions are frequently incomplete and typically lock an organization into a single platform. For example, one of the leading commercial options lacks integrated DNS/DHCP services. The Red Hat Network Satellite server provides DNS/DHCP, enabling organizations to manage the lifecycle of a server instance from provisioning to retirement automatically. In addition, Red Hat Network Satellite can provision across most virtualization substrates, including VMware and Red Hat Enterprise Virtualization—even bare metal.

Red Hat Network Satellite in a private cloud environment provides a clear and cost effective way to address the changing requirements for enterprise services.

SUMMARY

Red Hat Network Satellite provides significant benefits in the private cloud, including:

- Reduction of capital expenditures through the use of a smaller hardware footprint needed to meet occasional peak demand. Physical infrastructure and software subscription budgets can be significantly reduced.
- Limiting the need for system administrator time required to provision and de-provision services, reducing overtime requirements and allowing staff to focus on core projects rather than continual system tuning.
- Automation and dynamic scaling can help to eliminate human errors, making systems more reliable and available.
- Automatic response to changing demands, within minutes, dynamically changes the available resources to a service.
- Fully supported, standardized approach backed by Red Hat services, support and training.

We invite you to see how Red Hat Network Satellite Server can help you implement a dynamically allocated virtual environment using this approach.

- For more information on Red Hat Network Satellite, please go to http://www.redhat.com/red_hat_network/
- Or contact your local Red Hat Sales Representation or reseller.

RED HAT SALES AND INQUIRIES

NORTH AMERICA
1-888-REDHAT1
www.redhat.com

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

ASIA PACIFIC
+65 6490 4200
www.apac.redhat.com
apac@redhat.com

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

www.redhat.com
#3488087_0710