Running Genesys Inbound Voice Routing with VMware vMotion and HA on vSphere 4.1 and vSphere 5.0

February 2012

DEPLOYMENT AND TECHNICAL CONSIDERATIONS GUIDE
Table of Contents

Introduction ............................................................................................................................................. 3
Executive Summary ................................................................................................................................. 3
VMware and Genesys Overview ........................................................................................................... 4
  Genesys Suite Overview ...................................................................................................................... 4
  VMware Virtual Infrastructure ............................................................................................................. 5
Genesys Architecture and Deployment ................................................................................................ 6
Genesys Test Process and Results ......................................................................................................... 6
  vMotion .............................................................................................................................................. 7
  High Availability ................................................................................................................................. 14
  Hardware and Software Configuration ................................................................................................ 16
  Virtual Machine Configuration ............................................................................................................ 18
  Workload Used .................................................................................................................................. 18
Deployment Best Practices ..................................................................................................................... 19
Technical Support ................................................................................................................................ 21
Conclusions ........................................................................................................................................... 22
Resources ............................................................................................................................................. 23
Acknowledgements ............................................................................................................................... 23
Introduction

Virtualization is rapidly transforming the IT landscape and fundamentally changing the way that people compute. Today’s powerful x86 computer hardware was originally designed to run only a single OS. This approach introduced a tight coupling between Operating System (OS) and hardware. Virtualization breaks that bond between the OS and hardware by making it possible to run multiple OS and applications on the same computer at the same time, increasing the utilization and flexibility of hardware.

This document provides direction to those interested in running vMotion and High Availability (HA) using the Genesys Routing System on VMware® vSphere™ 4.1 and 5.0. It provides basic information on the architecture of Genesys Routing System, as well as the value of utilizing the VMware platform. Finally, the document outlines some best practices for utilizing the two product sets together in your datacenter.

Executive Summary

One of the largest challenges that customers have to contend with in any environment is how to deal with planned downtime. In the physical world servers always need to be powered down at some point, for various reasons: firmware upgrades, adding or removing hardware, or to upgrade to newer hardware. The same is true in the VMware virtual world.

VMware has pioneered a unique technology that minimizes planned downtime and allows workloads running on one physical ESX host server to be “hot” migrated (moved live) to another ESX host server, without any impact to end users. While this sounds like magic, VMware customers have been using vMotion technology successfully since it was first delivered in 2003.

VMware designed vMotion to handle three different use cases. It is the fundamental technology that enables each of these capabilities:

1. Minimizing planned downtime with operator managed (manual) vMotion.
3. Power savings with the automated power management subsystem VMware® Distributed Power Management (DPM).

In case of Server and operating system failures the best way to minimize downtime is by using VMware High Availability (HA).

1. Monitors virtual machines to detect operating system and hardware failures.
2. Restarts virtual machines on other physical servers in the resource pool without manual intervention when server failure is detected.
3. Protects applications from operating system failures by automatically restarting virtual machines when an operating system failure is detected.

You may want to know what impact VMware vMotion and HA will have on your applications when virtual machines are moved from one physical server to another or when there is server or operating system failure. The first question is around transactional integrity: will all your data be available after the migration completes? The second question is around understanding transaction rate and overall impact to SLAs: will your transactions per second and response times be negatively impacted by vMotion? This white paper provides data that will help answer these questions.
VMware and Genesys Overview

VMware (NYSE:VMW), the global leader in virtualization and cloud infrastructure, delivers customer-proven solutions that accelerate IT by reducing complexity and enabling more flexible, agile service delivery. VMware enables enterprises to adopt a cloud model that addresses their unique business challenges. VMware’s approach accelerates the transition to cloud computing while preserving existing investments and improving security and control. With more than 300,000 customers and 25,000 partners, VMware solutions help organizations of all sizes to lower costs, increase business agility and ensure freedom of choice.

Genesys Suite Overview

Genesys is the world’s leading provider of customer service and contact software — with more than 4,000 customers in 80 countries. With more than 20 years of customer service innovation and experience, Genesys is uniquely positioned to help companies bring their people, insights and customer channels together to effectively drive today’s customer conversation. Genesys software directs more than 100 million interactions every day, maximizing the value of customer engagement and differentiating the experience by driving personalization and multi-channel customer service – and extending customer service across the enterprise to optimize processes and the performance of customer-facing employees.

Figure 1: Genesys Suite Overview
VMware Virtual Infrastructure

VMware’s leading virtualization solutions provide multiple benefits to IT administrators and users. VMware virtualization creates a layer of abstraction between the resources required by an application and operating system, and the underlying hardware that provides those resources. A summary of the value of this abstraction layer includes the following:

- **Consolidation**: VMware technology allows multiple application servers to be consolidated onto one physical server, with little or no decrease in overall performance.
- **Ease of Provisioning**: VMware virtualization encapsulates an application into an image that can be duplicated or moved, greatly reducing the cost of application provisioning and deployment.
- **Manageability**: Virtual machines may be moved from server to server with no downtime using VMware vMotion™, which simplifies common operations like hardware maintenance and reduces planned downtime.
- **Availability**: Unplanned downtime can be reduced and higher service levels can be provided to an application. VMware High Availability (HA) ensures that in the case of an unplanned hardware failure, any affected virtual machines are restarted on another host in a VMware cluster.

Figure 2: VMware vSphere Virtual Infrastructure
Genesys Architecture and Deployment

Genesys core components for Inbound Voice routing consist of Management Framework, T-Servers, Universal Routing Server (URS), and Stat Server. These components work together to form core architecture for routing inbound voice calls from the PSTN to the appropriate agent or other resource.

A brief description of Genesys core components:

- Management Framework configures all components, including agent and other place definitions.
- T-Server manages the switch interface to the PSTN, receives routing commands from URS, provides statistics to Stat Server, and registers agent logins.
- URS makes routing decisions and signals T-Server to make the appropriate connections according to the defined call flows.
- Stat Server collects real-time statistics and distributes them to other clients such as URS and Management Framework.

The above components can be deployed in the customer's network by hosted service providers and across single or multiple sites. Deployment using virtualization is possible for all components, and it can provide cost savings and management advantages.

Figure 3: Genesys Architecture

Genesys Test Process and Results

In this exercise, VMware worked with Genesys to test the behavior of Genesys Inbound Voice Routing with vMotion and High Availability. To make this testing relevant to real-world environments we used the Genesys Simulator which closely resembles industry standard benchmark. Given the mission-critical nature of the Routing System, they need to ensure that no calls were dropped during the VMware vMotion and whether all the services properly come up after VMware HA.

The primary objectives of testing was to check if vMotion and HA works resiliently on VMware virtual infrastructure. The joint tests were also designed to provide answers a common set of questions, and to...
derive and document a set of best practices that can be used by our customers. Specifically, key areas under investigation were:

- Impact on calls per second during VMware vMotion/HA operation.
- Test whether the Primary component properly switches as a backup and processes all the calls after VMware HA.

**vMotion**

VMware vMotion technology moves running virtual machines from one physical server to another with no impact on end users. VMware vMotion keeps your IT environment up and running, giving you unprecedented flexibility and availability to meet the increasing demands of your business and end use.

*Figure 4: VMware vMotion*

vMotion was tested with three different test scenarios as listed below with both the ESX hosts subjected to additional CPU load of 70% which was done using ‘mprime’ utility running on a 6 vcpu RHEL machine.

**25 Calls per second**

In this test the Simulator was configured to generate load of 25 calls per second. The vMotion feature was tested with TServer, StatServer and the URS.
Figure 5: VMware vMotion feature tests

Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion.
Observations and Highlights

1. Average vMotion time was approximately between 30 to 40 seconds on ESXi 4.1 and 15 to 20 seconds on ESXi 5.0.
2. No calls were dropped and there was no queuing of calls waiting to be processed. All requests were processed in a timely manner.
3. There was no impact on the call transaction due to the additional load on both the ESXi hosts.
4. There was no impact on agent login sessions.
5. There was no UDP traffic to the applications being tested.

50 Calls per second

In this test the Simulator was configured to generate load of 50 calls per second. The vMotion feature was tested with TServer, StatServer and the URS.
Figure 6: VMware vMotion feature tests

Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion.

TServer VMotion with 50 calls per Second

StatServer VMotion with 50 calls per Second

Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion.
Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion

Observations and Highlights

1. Average vMotion time was approximately between 30 to 45 seconds on ESXi 4.1 and 15 to 20 seconds on ESXi 5.0.
2. There were some calls in the “In Process” queue but eventually all the calls were processed without any problems.
3. No calls were dropped or abandoned.
4. There was no impact to agent login sessions.
5. There was no UDP traffic to the applications being tested.
75 Calls per second

In this test the Simulator was configured to generate load of 75 calls per second. The vMotion feature was tested with TServer, StatServer and the URS.

Figure 7: VMware vMotion Tests

Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion
Note: Graphs show CPU (%) utilization of primary and target hosts through vMotion

Observations and Highlights

1. Average vMotion time was approximately between 35 to 45 seconds on ESXi 4.1 and 18 to 23 seconds on ESXi 5.0.
2. There were some calls in the “In Process” queue but eventually all the calls were processed without any problems.
3. No calls were dropped or abandoned.
4. There was no impact on agent login sessions.
5. There was no UDP traffic to the applications being tested.
High Availability
vSphere High Availability (HA) delivers the availability needed by many applications running in virtual machines, independent of the operating system and application running in it. HA provides uniform, cost-effective failover protection against hardware and operating system failures within your virtualized IT environment.

Figure 8: vSphere High Availability (HA)

In this scenario, TServer, StatServer and URS were enabled for VMware HA. After a host failure the backup nodes on the other host in the cluster automatically become primary nodes and re-establish connection with different components in the System. The scenario was tested as follows:

1. HA enabled for the TServer, StatServer and URS and its backup components.
2. vSphere host powered off.
3. TServer, StatServer and URS failover and the back node starts acting as a primary node.
4. All the components are back in Sync.
Figure 9: vSphere High Availability (HA) Tests
Observation and Highlights

1. When the HA was triggered, the primary VMs were terminated and TServer did not process any calls during ADDP timeout because it was in the backup mode. This resulted in some lost calls. To minimize the number of lost calls, the ADDP timeout was reduced to 4 seconds.
2. The above diagram depicts the approximate time until the Backup server becomes active on the other host.
3. Three iterations of HA tests were triggered to check the consistency of the results.
4. In the three iterations we observed average of 60 lost calls which according to Genesys is the expected Genesys HA behavior.

Note: “Lost” calls refer to transient incoming calls only. Genesys HA preserves stable calls already connected to agents or resources.

Hardware and Software Configuration

The following section provides details on the hardware and software used in the testing.

Figure 10: Genesys Test Configuration

2 ESX Hosts configured with ESX 4.1 (X5560 @2.00 Ghz and 40 GB RAM)
The following table describes the configuration of ESX host servers and storage in the Genesys test configurations.

Table 1: ESX Host Hardware

<table>
<thead>
<tr>
<th>HARDWARE</th>
<th>CONFIGURATION</th>
</tr>
</thead>
</table>
| HP Proliant BL460c G6 (Nehalem Boxes) | Two HP Proliant BL 460c G6 servers. Each server is equipped with:  
- Intel(R) Xeon(R) Nehalem CPU X5560 2.80 GHz  
- Sockets: 2  
- Number of Cores per Socket: 4  
- Hyperthreading: Enabled  
- 48 GB RAM  
- 4 X 1 GB NICS per ESX host |
| Storage | NetApp FAS3020:  
- 28 disks, 10K RPM  
- RAID DP  
- 4 Gbps Fibre Channel |

Installed Software

Table 2 lists the software used for the Genesys solution running on VMware virtual infrastructure.

Table 2: Software Installed for Genesys Testing on VMware Virtual Infrastructure

<table>
<thead>
<tr>
<th>INSTALLED SOFTWARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMware</td>
</tr>
</tbody>
</table>
| VMware vSphere ESX 4.1(vSphere) Build 260247  
VMware vSphere ESX 5.0(vSphere) Build 469512 |
| Microsoft           |
| Windows 2003 Enterprise Edition 32 bit  
URS  
TServer  
Stat Server  
Simulator |
Virtual Machine Configuration

The following table describes the configuration of virtual machines running on ESX host servers in the Genesys test configurations.

Table 3: Virtual Machine Configuration

<table>
<thead>
<tr>
<th>VIRTUAL MACHINE</th>
<th>HARDWARE CONFIGURATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 virtual machine instances</td>
<td>Each configured with:-</td>
</tr>
<tr>
<td>TServer</td>
<td>• 1 vCPUs</td>
</tr>
<tr>
<td>TServer_bk</td>
<td>• 1 GB memory</td>
</tr>
<tr>
<td>StatServer</td>
<td>• Ethernet card (1 Gbps Network)</td>
</tr>
<tr>
<td>StatServer_bk</td>
<td>• Disk storage 1 x12 GB</td>
</tr>
<tr>
<td>URS</td>
<td>• vmdk only</td>
</tr>
<tr>
<td>URS_bk</td>
<td></td>
</tr>
</tbody>
</table>

One Simulator virtual machine

- 2 vCPUs
- 3 GB memory
- Ethernet card (1 Gbps Network)
- Disk storage 1 X 20 GB vmdk only

2 RHEL x64 VM (This VM was used to drive additional load on both the ESX hosts)

- 6 vCPUs
- 20 GB memory
- Ethernet card (1 Gbps Network)
- Disk storage 1 X 20 GB
- vmdk only

Workload Used

To accomplish the testing, Genesys tools were used to simulate the loads:

- Avaya G3 Switch simulator: Simulates CTI link messaging and Inbound callers
- CCAS/Autoserv: Simulates Agent activity
- Smart Proxy: Simulates network delays
- Wireshark: Network traffic monitoring

These tools were able to generate call loads to simulate inbound voice routing scenarios in excess of 75 calls per second.
Deployment Best Practices

Here are some of the best practices derived from configuration and testing of the Genesys application in a VMware virtual infrastructure environment:

1. When using Intel-based systems, ensure that the BIOS settings enable VT and EPT options on all ESX hosts. Hyperthreading should be turned on.

2. Configure the VMs and the VMkernel on a separate NIC on a separate vSwitch.

3. Use VMware paravirtualized vmxnet3 adapter for better network throughput.

4. Install VMware tools on the virtual machines. The VMware Tools package provides support required for shared folders and for drag and drop operations. Other tools in the package support synchronization of time in the guest operating system and the host, automatic grabbing and releasing of the mouse cursor, copying and pasting between guest and host, and improved mouse performance in some guest operating systems.

5. Consider using server-class network interface cards (NICs) for the best performance and configure paravirtualized vmxnet3 adapters for better network throughput.

6. Virtual machines dependent on each other should be configured on same ESX host and same vSwitch. This will reduce traffic flow through the wire and avoid unnecessary CPU and network overhead. An example would be having SQLSERVER and one of the APPS on the same ESX host.

7. Disconnect or disable unused or unnecessary physical hardware devices, such as:
- COM ports
- LPT ports
- USB controllers
- Floppy drives
- Optical drives (that is, CD or DVD drives)

Disconnecting or disabling devices will help free up interrupt resources. For example, traditionally, some devices, such as USB controllers, operate on a polling scheme that consumes extra CPU resources. Some PCI devices reserve blocks of memory, making that memory unavailable to ESXi. HA Failover Architecture.

8. To reduce the number of the abandoned/unprocessed calls during Genesys HA recovery, ADDP parameters should be set for 4 seconds on the host object in CME, as shown below:

Figure 12: ADDP Parameters
9. Genesys HA and VMware HA are complementary, and may be used together, provided that anti-affinity rules are followed (see #10 below).
   - Genesys HA should always be used to provide application level protection, and ensure that calls and contacts in progress are not interrupted. (DRS anti-affinity required.)
   - VMware HA adds an additional layer of hardware failure protection. If there is a physical server failure, VMware HA will restart any failed "Genesys resiliency HA VM's impacted by the server failure" on different servers automatically. Without VMware HA, the "Genesys resiliency HA VM's" would remain on the failed server, powered down, disabling Genesys HA's functionality.
   - The combination of VMware HA, anti-infinity rules and Genesys HA allow Genesys HA to be brought back to full protection within a few minutes of a failed physical server that is hosting Genesys HA components, while at the same time facilitating the repair / replacement of the failed server.

10. Anti-Affinity Rules

In implementing VMware HA in conjunction with Genesys active/standby HA, VM-VM anti-affinity rules must be properly defined to ensure that the active and standby copies do not end up on the same hardware.
   - Since DRS controls initial placement of VMs when they are powered on, DRS groups must be set up in order to define anti-affinity rules, even if DRS is not used.
   - If automatic DRS is not being used, DRS should be set to semi-automatic mode -- this will allow it to place the VMs as per the rules on power-on, but not vMotion them later.

There is an issue where HA will not respect the anti-affinity rules when re-powering-on the VMs after a failure. VMware expects to fix that in the future. In the meantime, the following workarounds may be used:
   - Use VM to host group rules (VM-a on hosts 1,3,5,7 and VM-b on hosts 2,4,6,8 etc) -- these are respected by HA
   - Use a "failover host" so all VMs are re-powered on the same new host, therefore the logical VM grouping won't change
   - Let the VMs be placed improperly on HW failure (should be a rare occurrence anyway) and then if that happens, power down the standby VM, and re-power-on. The re-power-on will cause DRS to place it properly.

Technical Support

For technical support on virtualization issues, existing Genesys customers should use the following contacts:

North America support@genesyslab.com
Europe, Middle East, and Africa support@genesyslab.co.uk
Asia Pacific, Malaysia, India support@genesyslab.com.au
Japan support@genesyslab.co.jp

For general questions on Genesys virtualization support, contact virtualizationsupport@genesyslab.com
Conclusions

Overall, testing results show that, running VMware vMotion and HA with Genesys Inbound Voice Routing performs well. Furthermore, it has potential to reduce cost, increase service levels and simplify the manageability of applications. It is critical that new virtualization server capabilities are used, since this will significantly increase performance of Genesys. It is equally important that you follow hardware recommendations and best practices as mentioned in this guide.
Resources

Customers can find more information about VMware and Genesys products using the links listed below.

**VMware Resources**

- VMware official website: http://www.vmware.com/
- VMware download Web site: https://www.vmware.com/download/
- VMware support Web site: http://www.vmware.com/vmtn/

**Genesys Resources**

- Genesys Web site: www.Genesyslab.com

Acknowledgements

The following individuals contributed to the creation of this deployment guide:

- Balayya Kamanboina, ISV Validation Engineer, VMware
- Leonid Mitnik, Senior Software Engineer, Quality Assurance, Genesys
- Leonid Ryklansky, Senior Manager, Quality Assurance, Genesys