

SR-IOV PERFORMANCE ADVANTAGE: RED HAT ENTERPRISE LINUX 6

Red Hat Enterprise Linux® 6 improved virtualized database performance over 23 percent



OUR FINDINGS

Organizations virtualizing databases on their servers seek the best possible performance. Single Root I/O Virtualization (SR-IOV), utilized in Red Hat Enterprise Linux 6, is a pass-through technique that increases network storage performance by providing faster throughput and better latency than standard bridged connections, and allows multiple guests to use the same network attached storage. In Principled Technologies' tests in our labs, a server running a Red Hat Enterprise Linux 6 Kernel Virtual Machine (KVM) guest with SR-IOV increased database performance by 23.5 percent compared to a guest networked with a normal bridged connection.

OUR PROCESS

We tested the performance of PostgreSQL database instances in Red Hat Enterprise Linux 6 on a Dell™ PowerEdge™ R710 server. We connected the server to a Dell EqualLogic™ iSCSI storage area network (SAN) using an Intel® X520-SR2 10Gb Ethernet card. We used the open-source DVD Store Version 2 (DS2) benchmark to test the database performance of a Red Hat KVM guest using SR-IOV and bridged technologies. We used one-guest and four-guest configurations, running a single database instance in each guest. We tested a non-virtualized (bare-metal) configuration with both single and four database instances to establish a baseline.



PROJECT OVERVIEW

The purpose of this project is to show the performance gains of using SR-IOV in Red Hat Enterprise Linux 6 KVM guests rather than KVM guests using standard bridged network connections. We show the performance of the same server running bare metal (with no virtualization) as a baseline comparison.

SR-IOV provides better performance due to the way it communicates with the KVM guests and the network interface card (NIC). (To utilize SR-IOV with Red Hat Enterprise Linux 6 requires a network adapter that supports this feature, as well as compatible server hardware.) An SR-IOV enabled KVM guest communicates by sending data directly to the physical network adapter, which then relays it to the external iSCSI SAN. In addition, multiple guests (we used up to eight) can share the same network port with SR-IOV – a distinct improvement over standard PCI pass-through in which guests cannot share network ports, and only one guest can benefit from the increased communication performance. In a standard bridged network connection, the same data is sent to virtual NICs, which do additional processing before sending the data to a physical NIC and on to the iSCSI SAN. This processing takes additional system resources and hinders performance. The test results show this performance disparity.

Figure 1 shows DS2 results on Red Hat Enterprise Linux 6 running a single database instance for the following configurations: KVM bridged guest, KVM SR-IOV guest, and bare metal. In each case, we bound a single database instance to one socket on the computer using NUMA (non-uniform memory architecture) tools.

As Figure 1 shows, Red Hat KVM SR-IOV guest achieved 86,469 total orders per minute (OPM), which is a 23.5 percent performance increase over the Red Hat KVM bridged guest. The Red Hat KVM bridged guest achieved 69,984 total OPM.

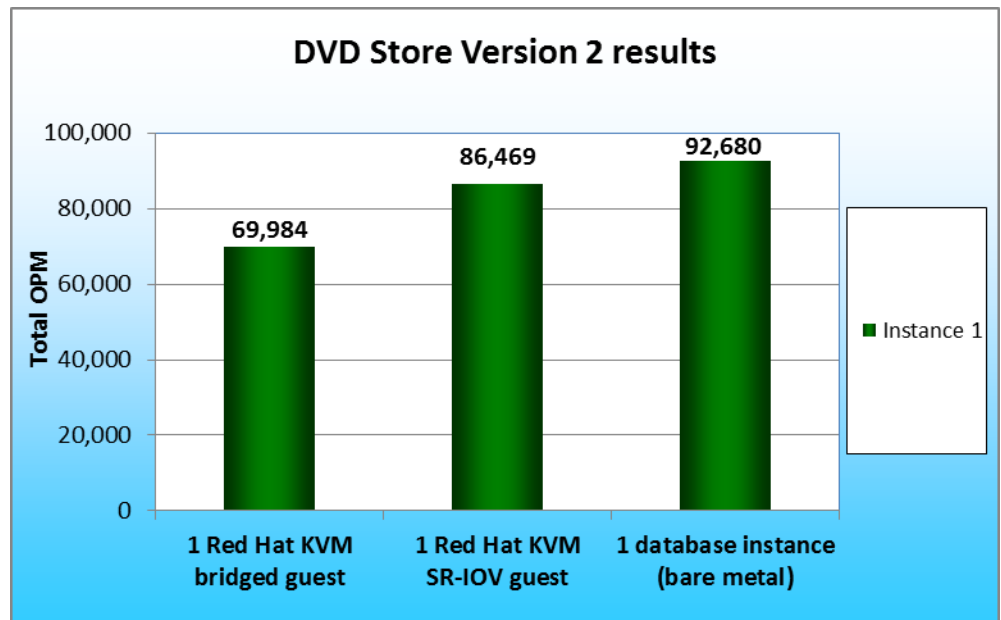


Figure 1: DVD Store Version 2 results for the server running Red Hat Enterprise Linux 6 with one database instance. Higher numbers are better.

While the Red Hat bare-metal database instance did provide higher total OPM, it is important to note the SR-IOV guest provided better performance than the standard bridged configuration and is closer to the bare-metal total OPM.

Figure 2 shows the total OPM results for the same configuration using four database instances. (The Red Hat KVM SR-IOV guest and Red Hat KVM bridged guest configurations ran four guests with a single database instance in each. The bare-metal configuration ran four database instances.) Note that we bound the instances to specific CPUs using NUMA tools for the bare-metal test, and NUMA plus libvirt for the two KVM guest configurations.

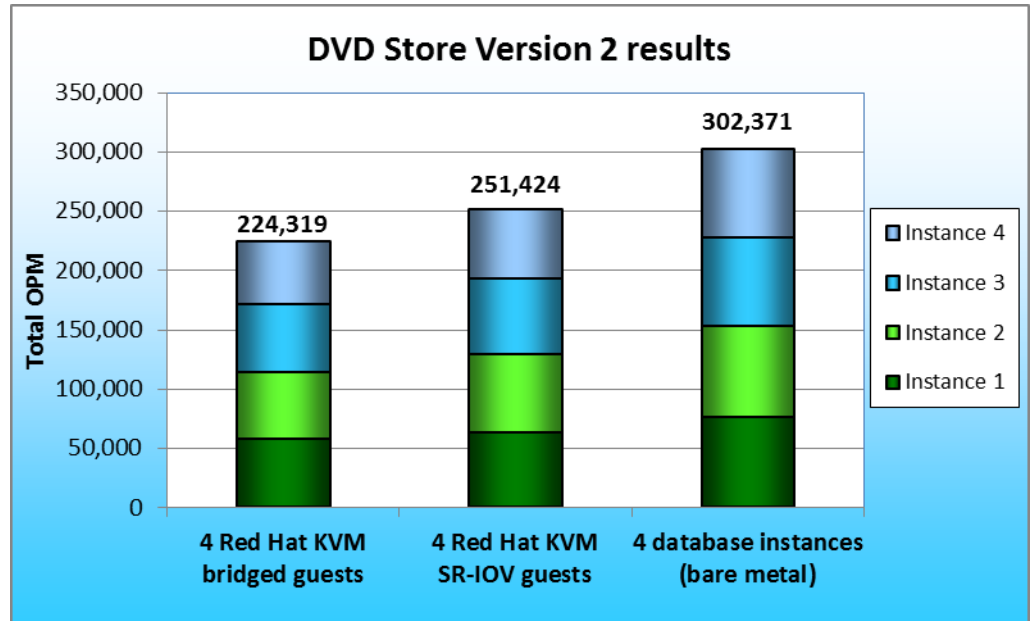


Figure 2: DVD Store Version 2 results for the server running Red Hat Enterprise Linux 6 with four database instances. Higher numbers are better.

As Figure 2 shows, the Red Hat KVM SR-IOV guests provided a 12.1 percent performance increase over Red Hat KVM bridged guests. The four Red Hat KVM SR-IOV guests configuration achieved 251,424 total OPM, while the four KVM bridged guests achieved 224,319 total OPM. Again, the bare-metal configuration provided higher OPM, but the Red Hat KVM SR-IOV guests gave better performance than the Red Hat KVM bridged guest configuration.

WHAT WE TESTED

We used DVD Store Version 2, an open-source application that models an online DVD store, to build the workload. DS2 has a back-end database component, a front-end Web application layer, and a driver layer that executes the workload. Because our goal was to test database server performance, we did not use the front-end Web client component of DS2, but ran the included driver program directly via the client machine's command-line interface. We used the default setup configuration for DS2, with exceptions we note in the DVD

Store setup section in the How We Tested Section of this report. We chose PostgreSQL version 8.4.4 running on Red Hat Enterprise Linux 6 as the back-end database.

DS2 reports results in orders per minute. One DS2 order consists of a customer login, a search for movies by title, actor, or category, and a purchase. The workload also performs other actions, such as adding new customers, to exercise a wide range of database functions.

We configured the Dell PowerEdge R710 server with a 4GB DS2 database. When running multiple instances of PostgreSQL, each instance contained a 4GB DS2 database. We used two clients against each instance of PostgreSQL.

Each client machine ran a single instance of the DS2 driver, with 20 threads, to simulate a heavily loaded environment. The client machines ran with no think time, processing requests as quickly as they were able. The workload ran for 15 minutes on each server, with the multiple instances on the Dell PowerEdge R710 running the workload simultaneously.

For more details about the DS2 tool, see <http://www.delltechcenter.com/page/DVD+Store>.

We tested DS2 performance on the server using Red Hat KVM virtualization. We first tested a single PostgreSQL instance using one Red Hat KVM bridged guest, one Red Hat KVM SR-IOV guest, and a single bare-metal database instance for comparison. In each case, we bound a single instance to one socket on the computer. We then tested four PostgreSQL instances in four Red Hat KVM bridged guests, four Red Hat KVM SR-IOV guests, and four bare-metal database instances.

WHAT WE FOUND

Figure 3 shows the DVD Store Version 2 OPM results, by client, for the Dell PowerEdge R710 running one Red Hat KVM bridged guest and one Red Hat KVM SR-IOV guest, and by instance for the Dell PowerEdge R710 running one instance in bare metal. Please note that we present the results in Figures 3 and 4 to show the performance benefits of using SR-IOV over standard bridged networking in guests, and that we present bare-metal results simply as a baseline for comparison.

	Red Hat KVM bridged guest	Red Hat KVM SR-IOV guest	Bare-metal instance
Client 1	34,894	43,205	46,227
Client 2	35,090	43,264	46,453
Total OPM	69,984	86,469	92,680

Figure 3: DVD Store Version 2 OPM results, by client, for the Dell PowerEdge R710 running one database instance. Higher numbers are better.

Figure 4 shows the DVD Store Version 2 OPM results, by client, for the Dell PowerEdge R710 running four Red Hat KVM bridged guests, four Red Hat KVM SR-IOV guests, and by instance for the Dell PowerEdge R710 running four instances in bare-metal mode.

	Red Hat KVM bridged guest	Red Hat KVM SR-IOV guest	Bare-metal instance
Client 1	28,651	31,434	38,245
Client 2	28,622	32,281	38,407
Client 3	28,201	32,936	37,811
Client 4	28,150	32,942	37,901
Client 5	28,939	31,409	37,542
Client 6	28,982	31,391	37,472
Client 7	26,392	29,544	37,492
Client 8	26,382	29,487	37,501
Total OPM	224,319	251,424	302,371

Figure 4: DVD Store Version 2 OPM results, by client, for the Dell PowerEdge R710 running four database instances. Higher numbers are better.

HOW WE TESTED

Setting up and configuring the storage

Configuring the internal storage

We configured two volumes on the internal storage. We configured the first volume as a RAID 1 set of two 73GB drives, and we installed the host operating system on this volume. We configured the second volume on the remaining 146GB drive and used that volume for scripts and utility files.

Configuring the Dell EqualLogic™ storage

Each of the two Dell EqualLogic arrays had 16 drives, for a total of 32 drives. We configured the six Dell EqualLogic arrays in regular RAID 10 mode (14 active drives and 2 spares each), for a total of 28 active drives.

We connected the Dell PowerEdge R710 server to the Dell EqualLogic storage using a Dell PowerConnect™ 6248 switch. We connected the 10Gb network card on our server to the 10GbE uplink module on the switch via a Cat6 Ethernet cable. We created two storage pools and then alternated the instances between the two pools.

We dedicated two LUNS to each instance: one 40GB LUN for data, and one 20GB LUN for logging.

Setting up the storage

1. If the Dell EqualLogic PS5000XV needs to be reset, connect through the serial port, and issue the following command upon login:

```
# reset
```

2. Using the command line wizard, proceed with the reset process, and provide a group IP address, subnet mask, and login information.
3. Perform the same procedure on the second Dell EqualLogic PS5000XV array, but provide the group IP address from Step 2 as the group to join when the wizard prompts you.
4. Using the Group Manager web application, set up a storage group with the following RAID policies:
 - PS5000XV-1: RAID-10
 - PS5000XV-2: RAID-10
 - PS5000XV-3: RAID-10
 - PS5000XV-4: RAID-10
 - PS5000XV-5: RAID-10
 - PS5000XV-6: RAID-10
5. Create a storage pool name of pool1 with the following members: PS5000XV-1, PS5000XV-2, and PS5000XV-3.
6. Create a storage pool name of pool2 with the following members: PS5000XV-4, PS5000XV-5, and PS5000XV-6.
7. Enable all network connections on each Dell EqualLogic PS5000XV using the following IP address scheme:
 - IP Address: 192.168.1.## (## being 11,13, and 15 for array 1, and 12, 14, 16 for array 2)
 - Subnet Mask: 255.255.255.0
8. Create four 40GB volumes with no snapshot reserve, alternating them between storage pools, beginning with pool1.
9. Create four 20GB volumes with no snapshot reserve, alternating them between storage pools, beginning with pool2.
10. Create an access control record for every volume without specifying any limitations.
11. Create a read-only SNMP community name to use for group monitoring.

Setting up the host server with Red Hat Enterprise Linux 6

We installed Red Hat Enterprise Linux 6 on the host server.

Installing Red Hat Enterprise Linux 6

1. Insert and boot from the Red Hat Enterprise Linux 6 install DVD.
2. Press Enter to install using graphical mode.
3. At the media test screen, select Skip.
4. At the Red Hat Enterprise Linux 6 title screen, click Next.
5. At the Choose a Language screen, select English, and click Next.
6. At the Keyboard Type screen, select U.S. English, and click Next.
7. At the Installation Devices screen, select Basic Storage Devices, and click Next.
8. Enter the hostname dbhost.ds2.lan, and click Next.
9. At the Time zone selection screen, select the appropriate time zone, and click Next.
10. Enter the root password in the Root Password and Confirm fields, and click Next.
11. At the type of installation screen, select Use All Space, check Review and modify partitioning layout, and click Next.

12. At the partitioning layout screen, select `lv_home`, and click Delete. At the Confirm Delete warning, click Delete.
13. Select `lv_root` and click Edit. In the Size field, enter the Max size displayed. Click OK.
14. Click Next. If a warning appears, click Write changes to disk.
15. At the Boot loader screen, click Next.
16. At the default installation screen, select Software Development Workstation, Customize now, and click Next.
17. Under the Base System software list, select iSCSI Storage Client.
18. Under the Databases software list, select PostgreSQL Database client and PostgreSQL Database server, and click Next.
19. Installation will begin automatically.
20. At the Congratulations screen, click Reboot.
21. After the system reboots and the Welcome screen appears, click Forward.
22. Agree to the EULA, and click Forward.
23. At the Set Up Software Updates screen, select No, I prefer to register at a later time, and click Forward.
24. If a pop-up screen appears asking if you are sure, click No thanks.
25. At the Finish Updates Setup, click Forward.
26. At the Create User Screen, enter username `testuser` and password `TestPass`, and click Forward.
27. If a warning appears, click Yes.
28. At the Date and Time screen, set the time and date, and click Forward.
29. At the Kdump screen, uncheck Enable Kdump, and click Finish.
30. If a warning appears, click Yes, and click OK.

Setting up the network adapters

1. Log into the host, right-click the desktop, and choose Open Terminal.
2. Type the following command to edit the network configuration settings, where X is the relevant host NIC to modify:

```
# vi /etc/sysconfig/network-scripts/ifcfg-ethX
```

3. Modify the following lines to set the static IP address and netmask, where X is the remaining part of the relevant IP address:

```
BOOTPROTO=static
IPADDR=192.168.XXX.XXX
NETMASK=255.255.255.0
```

4. Save the file, and exit vi.
5. Type the following command to modify the remaining network settings:

```
# vi /etc/sysconfig/network
```

6. Modify the following lines to disable IPv6 and set the hostname, where “hostname” is the hostname of this server:

```
NETWORKING=yes
NETWORKING_IPV6=no
HOSTNAME=hostname.local.domain
```

7. Save the file, and exit vi.
8. Repeat steps 2 through 7 for the NIC dedicated to iSCSI traffic.
9. Reboot the server.

Configuring the iSCSI software initiator and multipathing in Red Hat Enterprise Linux 6

1. Log into the host, right-click the desktop, and choose Open Terminal.
2. Using a text editor, edit `/etc/iscsi/iscsid.conf`, and modify the following settings:

```
node.session.cmds_max = 1024
node.session.queue_depth = 128
```

3. Type the following sequence of commands to configure the iSCSI interface, configure iSCSI to start on boot, and to start it immediately:

```
# modprobe scsi_transport_iscsi
# modprobe bnx2i
# chkconfig iscsi on
# chkconfig iscsid on
# iscsiadm -m iface -o new -I iface1
# iscsiadm -m iface -o new -I iface2
# iscsiadm -m iface -o new -I iface3
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth4 -I
iface1
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth4 -I
iface2
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth4 -I
iface3
# iscsiadm -m discovery -t st -p 192.168.10.10:3260
# service iscsi start
```

Adding the storage volumes to Red Hat Enterprise Linux 6

1. Issue the `mkfs.ext4` command to make the file system:

```
# mkfs.ext4 -L log1 /dev/mapper/log1
# mkfs.ext4 -L log2 /dev/mapper/log2
# mkfs.ext4 -L log3 /dev/mapper/log3
# mkfs.ext4 -L log4 /dev/mapper/log4
# mkfs.ext4 -L db1 /dev/mapper/db1
# mkfs.ext4 -L db2 /dev/mapper/db2
# mkfs.ext4 -L db3 /dev/mapper/db3
```



```
# mkfs.ext4 -L db4 /dev/mapper/db4

# mkdir /vol/pglog1
# mkdir /vol/pglog2
# mkdir /vol/pglog3
# mkdir /vol/pglog4
# mkdir /vol/ds2data1
# mkdir /vol/ds2data2
# mkdir /vol/ds2data3
# mkdir /vol/ds2data4
```

2. Edit the fstab to mount the volumes on boot. Add a line for each volume, such as below.

```
/dev/mapper/log1 /vol/pglog1 ext4 _netdev,defaults 0 0
/dev/mapper/log2 /vol/pglog2 ext4 _netdev,defaults 0 0
/dev/mapper/log3 /vol/pglog3 ext4 _netdev,defaults 0 0
/dev/mapper/log4 /vol/pglog4 ext4 _netdev,defaults 0 0

/dev/mapper/db1 /vol/ds2data1 ext4 _netdev,defaults 0 0
/dev/mapper/db2 /vol/ds2data2 ext4 _netdev,defaults 0 0
/dev/mapper/db3 /vol/ds2data3 ext4 _netdev,defaults 0 0
/dev/mapper/db4 /vol/ds2data4 ext4 _netdev,defaults 0 0
```

3. Issue the following command to further configure the storage:

```
# mount -a
# tuned-adm profile enterprise-storage
```

Tuning the database

1. Edit the following lines of the postgresql.conf file:

```
shared_buffers = 4GB
effective_cache_size = 8GB
checkpoint_segments = 1024
checkpoint_timeout = 1h
```

Setting up and configuring the KVM guests

Creating one 40GB logical volume for each guest on the KVM host

1. Type the following command, replacing VolumeGroup and PhysicalVolumePath with their correct values, to create a logical volume named ds2_guest1:

```
# lvcreate -L 40G -n ds2_guest1 VolumeGroup PhysicalVolumePath
```

2. Repeat Step 1 for each guest.

Creating the first guest using the Virtual Machine Manager

1. Log onto the KVM host.
2. Go to Applications→System Tools→Virtual Machine Manager.
3. Click New.
4. Click Forward.
5. Choose a name for the guest. Click Forward.
6. Keep the default of Fully virtualized, choose the CPU architecture, and choose kvm as the hypervisor.
7. Assuming you have copied the Red Hat Enterprise Linux 6 media to the host machine, choose Local install media, Linux as the OS Type, and Red Hat Enterprise 6 as the OS Variant. Click Forward.
8. Click Browse to browse to the ISO Linux location.
9. Locate the ISO file, select it, and click Open. Click Forward.
10. Select managed or other storage, and click Browse. Click Browse local.
11. Locate the logical volume in /dev/mapper/ and click. Click Forward.
12. Set Max memory size and Startup memory size to 10240MB, and set the number of Virtual CPUs to 6. Click Forward.
13. Review the summary information, and click Finish.

Installing Red Hat Enterprise Linux 6.0 on the first guest

1. Double-click the new VM to connect to the console.
2. On the Hardware tab in Virtual Machine Manager, specify the ISO image on the host machine as the CD drive of the VM. Right-click the VM, and choose Run.
3. Power on the VM.
4. Press Enter to install using graphical mode.
5. At the media test screen, select Skip.
6. At the Red Hat Enterprise Linux 6 title screen, click Next.
7. At the Choose an Installation Language screen, select English, and click Next.
8. At the Keyboard Type screen, select U.S. English, and click Next.
9. At the Storage Devices screen, select Basic Storage Devices, and click Next.
10. If a warning for device initialization appears, select Re-initialize for every storage device needed the installation.
11. At the Name the Computer screen, type `##.domain.com`, where `##` is the host name, and click Next.
12. At the Time zone selection screen, select the appropriate time zone, and click Next.
13. Enter the root password in the Root Password and Confirm fields, and click Next.
14. At Partition selection screen, select Replace Existing Linux System(s), and click Next.
15. If a warning appears, click Write changes to disk.
16. At the default installation screen, click Next to begin the installation.
17. At the Congratulations screen, click Reboot.

Configuring the KVM host for guest networking

1. On the KVM host, create a network bridge on the same network as the DS2 clients.
2. Type the following command to edit the network bridge configuration:

```
# vi /etc/sysconfig/network-scripts/ifcfg-br0
```

3. Modify the file to set the static IP address and netmask, where XXX is the remaining portion of the IP address on the DS2 clients' network:

```
TYPE=Bridge
ONBOOT=yes
BOOTPROTO=static
IPADDR=192.168.137.XXX
NETMASK=255.255.255.0
DELAY=0
```

4. Save the file, and exit vi.
5. Type the following command to edit the configuration of the bridge's network adapter:

```
# vi /etc/sysconfig/network-scripts/ifcfg-eth0
```

6. Modify the following lines to add the adapter to the bridge:

```
BRIDGE=br0
ONBOOT=yes
```

7. Save the file, and exit vi.
8. Start this network by typing the following commands:

```
# ifup eth0
# ifup br0
```

9. For bridge/virtio networking from the guests to the database storage array, create a second network bridge by repeating steps 1 through 8 above.
10. For SR-IOV networking from the guests to the database storage array, enable the network adapter but do not set an IP address.
11. Type the following command to edit the configuration of the SR-IOV capable adaptor:

```
# vi /etc/sysconfig/network-scripts/ifcfg-eth6
```

12. Modify the following line to enable the adapter:

```
ONBOOT=yes
#IPADDR=192.168.XX.XXX
```

13. Save the file, and exit vi.
14. Start this network by typing the following command:

```
# ifup eth6
```

Configuring networking in the guest

1. Start virt-manager.
2. Right-click on the guest, and select Open.
3. On the Virtual Machine window, left-click View on the bar, and select Details.

4. Connect the guest to bridge br0 on the KVM host by clicking Add Hardware; then select Network for Hardware type. Click Forward.
5. Select br0 for Host Device, virtio for Device Model, click Forward, and then click Finish.
6. For bridge/virtio networking, connect the guest to bridge br1 on the KVM host by repeating steps 4 and 5.
7. For SR-IOV networking, we create a host device for the PCI device associated with the SR-IOV VF.
8. Determine the PCI address of the VF for this guest by executing the following command on the KVM host:

```
# lspci | grep -i ethernet
```

The PCI address we need is in the first field in the output of the command; for example, 06:10.0.

9. On the guest's Details window, click Add Hardware; then select Physical Host Device for Hardware type; and click Forward.
10. Select PCI Device for Device Type, and select the entry that contains the PCI address determined in Step 8; click Forward, and Finish.
11. Power on the VM, and open the console.
12. Log onto the VM, right-click the desktop, and choose Open Terminal.
13. Type the following command to edit the network configuration settings:

```
# vi /etc/sysconfig/network-scripts/ifcfg-eth0
```

14. Modify the following lines to set the static IP address and netmask, where XXX is the remaining portion of your IP Address:

```
BOOTPROTO=static
IPADDR=192.168.XXX.XXX
NETMASK=255.255.255.0
```

15. Save the file, and exit vi.
16. Repeat steps 13 and 14 for the second adapter, eth1.
17. Type the following command to modify the remaining network settings:

```
# vi /etc/sysconfig/network
```

18. Modify the following lines to disable IPv6 and set the hostname:

```
NETWORKING=yes
HOSTNAME=VM1.local.domain
```

19. Save the file, and exit vi.

Configuring additional tuning options in the VM

1. Log onto the VM using Virtual Machine Manager, or using an ssh client, such as Putty.
2. Disable SELinux as above for the KVM host.
3. Set the tuning profile to enterprise-storage:

```
# tuned-adm profile enterprise-storage
```

Creating additional guests on the host server

After installing and configuring the initial guest, we created the additional guest by using virt-clone on the Red Hat Enterprise Linux platform.

Cloning the guests on Red Hat Enterprise Linux

Run the following command to clone GUEST1 to the area reserved for GUEST2:

```
# virt-clone --original GUEST1 --name GUEST2 --file  
/var/lib/libvirt/images/MRG-GUEST2.img -nonsparse
```

Configuring the additional guests after cloning

Modify the IP addresses in each guest, Configuring networking in the guest.

Configuring the iSCSI software initiator and multipathing in Red Hat Enterprise Linux 6 guests

1. Log into the first guest, right-click the desktop, and choose Open Terminal.
2. Using a text editor, edit `/etc/iscsi/iscsid.conf`, and modify the following settings:

```
node.session.cmds_max = 1024  
node.session.queue_depth = 128
```

3. Type the following sequence of commands to configure the iSCSI interface, configure iSCSI to start on boot, and to start it immediately:

```
# modprobe scsi_transport_iscsi  
# chkconfig iscsi on  
# chkconfig iscsid on  
# iscsiadm -m iface -o new -I iface1  
# iscsiadm -m iface -o new -I iface2  
# iscsiadm -m iface -o new -I iface3  
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth1 -I  
iface1  
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth1 -I  
iface2  
# iscsiadm -m iface -o update -n iface.net_ifacename -v eth1 -I  
iface3  
# iscsiadm -m discovery -t st -p 192.168.10.10:3260  
# service iscsi start
```

4. Repeat Steps 1 through 3 for each guest.

Adding the storage volumes to the Red Hat Enterprise Linux 6 guests

1. Issue the `mkfs.ext4` command to make the file system on the four guests:

```
Guest 1  
# mkfs.ext4 -L log1 /dev/mapper/log1  
# mkfs.ext4 -L log2 /dev/mapper/log2
```

```
# mkdir /vol/pglog1 /vol/ds2data1

Guest 2
# mkfs.ext4 -L log3 /dev/mapper/log3
# mkfs.ext4 -L log4 /dev/mapper/log4
# mkdir /vol/pglog2 /vol/ds2data2

Guest 3
# mkfs.ext4 -L db1 /dev/mapper/db1
# mkfs.ext4 -L db2 /dev/mapper/db2
# mkdir /vol/pglog3 /vol/ds2data3

Guest 4
# mkfs.ext4 -L db3 /dev/mapper/db3
# mkfs.ext4 -L db4 /dev/mapper/db4
# mkdir /vol/pglog4 /vol/ds2data4
```

2. Edit the fstab to mount the volumes on boot. Add a line for each volume, such as below.

```
## Guest 1
/dev/mapper/log1 /vol/pglog1 ext4 _netdev,defaults 0 0
/dev/mapper/db1 /vol/ds2data1 ext4 _netdev,defaults 0 0

## Guest 2
/dev/mapper/log2 /vol/pglog2 ext4 _netdev,defaults 0 0
/dev/mapper/db2 /vol/ds2data2 ext4 _netdev,defaults 0 0

## Guest 3
/dev/mapper/log3 /vol/pglog3 ext4 _netdev,defaults 0 0
/dev/mapper/db3 /vol/ds2data3 ext4 _netdev,defaults 0 0

## Guest 4
/dev/mapper/log4 /vol/pglog4 ext4 _netdev,defaults 0 0
/dev/mapper/db4 /vol/ds2data4 ext4 _netdev,defaults 0 0
```

3. Issue the following commands on each guest to further configure the storage:

```
# mount -a
# tuned-adm profile enterprise-storage
```

4. Issue the following command on the host server to further configure the storage:

```
# tuned-adm profile default
```

Installing and configuring PostgreSQL on each guest

1. Install the PostgreSQL server from the Red Hat Enterprise Linux 6 media.
2. Edit the following lines of the postgresql.conf file:

```
shared_buffers = 4GB
effective_cache_size = 8GB
checkpoint_segments = 1024
checkpoint_timeout = 1h
```

Setting up DVD Store

Data generation overview

We built the database schema using the scripts in the DS2 distribution package, although we modified several as we note below. The DS2 stress tool provides options to generate 10MB, 1GB, or 100GB datasets, but not 4 GB of user data that we wished to use in the test. Accordingly, we modified the data-generation source code and the DVD Store application's scripts to generate our user data. Note: We created our test data on a Linux system to take advantage of the larger RAND MAX.

Editing the ds2_create_orders.c module

The module ds2_create_orders.c defines constants that specify the maximum values for the customer ID and the product ID. Because the constants for the 4GB database size did not exist, we added the constants for this size.

The available command-line options for specifying the size were S (small), M (medium), and L (large). We added the option R to create the 4GB database. In the switch statement that sets the values for the variables max_cust_id and max_prod_id, we added cases that assigned them the proper values for the 4GB database size.

We recompiled the ds2_create_orders.c module on Linux, following the instructions in the header comments. We used the following command line:

```
gcc -o ds2_create_orders ds2_create_orders.c -lm
```

Editing the ds2_create_cust.c module

We modified the ds2_create_cust.c module to permit 4GB sizes, following the changes we made to the ds2_create_orders.c module. We added the command-line option R, for creating 4 GB databases, to the existing options, S (small), M (medium), and L (large). In the switch statement that sets the values for the

variables `max_cust_id` and `max_prod_id`, we added cases that assigned them the proper values for the 4GB database size.

We recompiled the `ds2_create_cust.c` module on Linux, following the instructions in the header comments. We used the following command line:

```
gcc -o ds2_create_cust ds2_create_cust.c -lm
```

Generating the data for the 4GB database

We used shell scripts to run the four executables that generate user data. Because the distribution did not include shell scripts for the 4GB size, we created shell scripts based on the `ds2_create_cust_large.sh` and `ds2_create_orders_large.sh` scripts. The `ds2_create_prod` and `ds2_create_inv` executables did not ship with associated shell scripts, so we created shell scripts using the instructions in the readme files. We ran the shell scripts in the following order to generate the data for the 4GB database:

1. `ds2_create_orders_4gb.sh`
2. `ds2_create_inv_4gb.sh`
3. `ds2_create_prod_4gb.sh`
4. `ds2_create_cust_4gb.sh`

We waited until the processes finished before we moved onto the next step.

Creating the database

After creating the PostgreSQL user as we note above, we placed the generated CSV data into the appropriate folders in the DVD Store distribution and ran the `pgsqlds2_create_all.sh` shell script to create the database.

After configuring and creating the PostgreSQL DS2 database, we performed a cold backup. We stopped the PostgreSQL service, copied all relevant files to our utility partition, and restarted the PostgreSQL service.

Editing the workload script - `ds2xdriver.cs` module

To use the 4GB database we created earlier, we changed the following constants in the DVD Store client application:

- In the routine `Controller()`, we changed the string sizes. We added the `R` option for the 4GB database size. DS2 uses the sizes string to interpret the `db_size_str` option.
- In the class `Controller`, we changed the arrays `MAX_CUSTOMER` and `MAX_PRODUCT`. To each, we added values specifying the bounds for the customer and product IDs. The `Controller()` routine uses these arrays.

Recompiling the ds2pgsqldriver.exe executable

We recompiled the ds2xdriver.cs and ds2pgsqlfns.cs module on the Windows server by following the instructions in the header comments. Because the DS2 instructions were for compiling from the command line, we performed the following steps:

1. Install the Npgsql connector on the development workstation.
2. Copy the Npgsql.dll and Mono.Security.dll files to the same directory as the DVD Store source files.
3. Open a command prompt.
4. Use the cd command to change directories to that containing our sources.
5. Execute the following command:

```
C:\Windows\Microsoft.NET\Framework\v2.0.50727\csc.exe
/out:ds2pgsqlserverdriver.exe ds2xdriver.cs ds2pgsqlserverfns.cs
/d:USE_WIN32_TIMER /d:GEN_PERF_CTRS /r:Npgsql.dll
/r:Mono.Security.dll /debug
```

Creating the DS2 client machines

We used eight Dell PowerEdge M600 Blade Systems for the DS2 clients. Each blade server had two Intel® Xeon® Processor E5345s 2.33GHz, 4 GB of PC2-5300 memory, two 73GB 10K SAS drives in RAID 1 configuration, and Microsoft® Windows Server® 2008 R2 Enterprise Edition.

For the DVD Store scripts, we used a number of client machines to simulate a number of users putting a load on the server. For our client machines, we installed Windows Server 2008 R2 Enterprise Edition, and copied the DVD Store files to the appropriate folders. We created a folder we called C:\ClientShare to store workload scripts, and shared this folder for access from our controller machine. We followed this process for each installation.

Installing Windows Server 2008 R2 Enterprise Edition

1. Boot the server, and insert the Windows Server 2008 R2 installation DVD in the DVD-ROM drive.
2. At the Language Selection Screen, click Next.
3. Click Install Now.
4. Select Windows Server 2008 R2 Enterprise (Full Installation), and click Next.
5. Click the I accept the license terms check box, and click Next.
6. Click Custom.
7. Click Drive options (advanced).
8. Ensure you select the proper drive, and click New.
9. Click Apply.
10. Click Next.
11. At the User's password must be changed before logging on warning screen, click OK.
12. Type TestPass as the new password in both fields, and click the arrow to continue.
13. At the Your password has been changed screen, click OK.

Configuring the IP Address

1. Click Start→Control Panel→Network Connection→Local Area Connection.
2. Double-click Internet Protocol (TCP/IP).
3. Select Use the following IP address.
4. Type 192 . 168 . 137 . X for the IP address, X being the appropriate IP address.
5. Type 255 . 255 . 255 . 0 for the Subnet mask, and click OK.
6. Click OK to close the Local Area Connection Properties.

Testing procedure

To perform the test, we used a series of batch files and shell scripts, which we executed from a controller machine. We stored batch files respective to each system on that system; and we used the PsExec and Plink utilities to coordinate the test procedures amongst the client machine, server machine, and controller. We used simple file sharing and WinSCP to copy files from machine to machine.

The testing procedure consisted of the following steps:

1. Execute batch files and shell scripts to clean up prior outputs on client machines, the server physical machine, and the controller.
2. Stop the PostgreSQL service.
3. Delete all prior database files.
4. Copy all original database files from the backup utility partition.
5. Reboot the client machines.
6. Reboot the server under test.
7. Wait for a ping response from all client machines.
8. Wait for a ping response from the physical server machine.
9. Start the KVM guests (one to four, depending on the test) with the appropriate NUMA node bindings.
10. Mount all necessary partitions.
11. Start the PostgreSQL services (either one or four instances, depending on the test) on the KVM guests.
12. Start the workload connections.
13. Start the workload ramp up period.
14. Start the workload.
15. Stop the workload.
16. Copy all output files to the controller.

APPENDIX A – SERVER CONFIGURATION INFORMATION

Figure 5 provides detailed configuration information for the test server.

System	Dell PowerEdge R710
Power supplies	
Total number	2
Vendor and model number	Dell A570P-00
Wattage of each (W)	570
Cooling fans	
Total number	5
Dimensions (h x w) of each	2-1/2" x 2-1/2"
Volts	12
Amps	1.6
General	
Number of processor packages	2
Number of cores per processor	6
Number of hardware threads per core	2
CPU	
Vendor	Intel
Name	Xeon
Model number	X5670
Stepping	01
Socket type	LGA1366
Core frequency (GHz)	2.93
Bus frequency (GT/s)	6.40
L1 cache (KB)	32 + 32 (per core)
L2 cache (KB)	256 (per core)
L3 cache (MB)	12
Platform	
Vendor and model number	Dell PowerEdge R710
Motherboard model number	PWB9YY69
Motherboard chipset	Intel 5520
BIOS name and version	Dell Incorporated 2.2.2 (9/21/2010) Revision 2.2
BIOS settings	Default
Memory module(s)	
Total RAM in system (GB)	24
Vendor and model number	Crucial CT51272BB1339.36SFD1
Type	PC3-10600R
Speed (MHz)	1,333
Speed running in the system (MHz)	1,333
Timing/Latency (tCL-tRCD-tRP-tRASmin)	9-9-9-24
Size (GB)	4

System	Dell PowerEdge R710
Number of RAM module(s)	6
Chip organization	Double-sided
Rank	Dual
Hard disk	
Vendor and model number	Seagate ST9146802SS
Number of disks in system	2
Size (GB)	146
Buffer size (MB)	16
RPM	10,000
Type	SAS
Disk controller	
Vendor and model	LSI Logic / Symbios Logic SAS1068E
Controller driver (module)	2.6.32-71.el6.x86_64
Controller driver version	0B4D557979D0BC8F39D9984
Controller firmware	0.25.47.00-IR
RAID configuration	RAID 1
Operating system	
Name	Red Hat Enterprise Linux 6
File system	ext4
Kernel	2.6.32-71.el6.x86_64
Language	English
Graphics	
Vendor and model number	Matrox® G200eW
Ethernet	
Vendor and model number	Broadcom® NetXtreme® II BCM57710
Type	PCI Express
Driver (module)	2.6.32-71.el6.x86_64
Driver version	1.52.1-7
Ethernet	
Vendor and model number	Intel® Ethernet Server Adapter X520-SR2
Type	PCI Express
Driver (Module)	2.6.32-71.el6.x86_64
Driver Version	2.0.62-k2
Optical drive(s)	
Vendor and model number	TEAC DV-28S
Type	DVD ROM
USB ports	
Number	4
Type	2.0

Figure 5: Detailed configuration information for the test server.

APPENDIX B – TEST STORAGE INFORMATION

Figure 6 presents detailed information for the test storage we used in this report.

Enclosure	Dell EqualLogic PS5000XV
General dimension information	
Height (inches)	5.25
Width (inches)	19.0
Depth (inches)	21.5
U size in server rack	3
Power supplies	
Total number	2
Wattage of each	450
Cooling fans	
Total number	4
Dimensions (H x W) of each	4" x 4"
Voltage (V)	12.0
Amps (A)	1.8
Hard disk	
Vendor and model number	Seagate ST3146855SS Maxtor 8k147S0087511
Number of disks in system	16
Size (GB)	146
Buffer size (MB)	16
RPM	15,000
Type	SAS
Storage connectivity (Fibre Channel, SAS, SATA, iSCSI, SCSI)	iSCSI
Firmware revision	V5.0.2
Storage cache (GB)	1
Number of storage controllers	2 (1 active)
Number of storage ports	3 per controller
Maximum network bandwidth of storage connectivity to host	1 Gb x 3
Switch type/model/firmware revision	Dell PowerConnect 6248 V3.2.0.7

Figure 6: Detailed configuration information for the test storage.

ABOUT PRINCIPLED TECHNOLOGIES



Principled Technologies, Inc.
1007 Slater Road, Suite 300
Durham, NC, 27703
www.principledtechnologies.com

We provide industry-leading technology assessment and fact-based marketing services. We bring to every assignment extensive experience with and expertise in all aspects of technology testing and analysis, from researching new technologies, to developing new methodologies, to testing with existing and new tools.

When the assessment is complete, we know how to present the results to a broad range of target audiences. We provide our clients with the materials they need, from market-focused data to use in their own collateral to custom sales aids, such as test reports, performance assessments, and white papers. Every document reflects the results of our trusted independent analysis.

We provide customized services that focus on our clients' individual requirements. Whether the technology involves hardware, software, Web sites, or services, we offer the experience, expertise, and tools to help our clients assess how it will fare against its competition, its performance, its market readiness, and its quality and reliability.

Our founders, Mark L. Van Name and Bill Catchings, have worked together in technology assessment for over 20 years. As journalists, they published over a thousand articles on a wide array of technology subjects. They created and led the Ziff-Davis Benchmark Operation, which developed such industry-standard benchmarks as Ziff Davis Media's Winstone and WebBench. They founded and led eTesting Labs, and after the acquisition of that company by Lionbridge Technologies were the head and CTO of VeriTest.

Principled Technologies is a registered trademark of Principled Technologies, Inc.
All other product names are the trademarks of their respective owners.

Disclaimer of Warranties; Limitation of Liability:

PRINCIPLED TECHNOLOGIES, INC. HAS MADE REASONABLE EFFORTS TO ENSURE THE ACCURACY AND VALIDITY OF ITS TESTING, HOWEVER, PRINCIPLED TECHNOLOGIES, INC. SPECIFICALLY DISCLAIMS ANY WARRANTY, EXPRESSED OR IMPLIED, RELATING TO THE TEST RESULTS AND ANALYSIS, THEIR ACCURACY, COMPLETENESS OR QUALITY, INCLUDING ANY IMPLIED WARRANTY OF FITNESS FOR ANY PARTICULAR PURPOSE. ALL PERSONS OR ENTITIES RELYING ON THE RESULTS OF ANY TESTING DO SO AT THEIR OWN RISK, AND AGREE THAT PRINCIPLED TECHNOLOGIES, INC., ITS EMPLOYEES AND ITS SUBCONTRACTORS SHALL HAVE NO LIABILITY WHATSOEVER FROM ANY CLAIM OF LOSS OR DAMAGE ON ACCOUNT OF ANY ALLEGED ERROR OR DEFECT IN ANY TESTING PROCEDURE OR RESULT.

IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC. BE LIABLE FOR INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES IN CONNECTION WITH ITS TESTING, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT SHALL PRINCIPLED TECHNOLOGIES, INC.'S LIABILITY, INCLUDING FOR DIRECT DAMAGES, EXCEED THE AMOUNTS PAID IN CONNECTION WITH PRINCIPLED TECHNOLOGIES, INC.'S TESTING. CUSTOMER'S SOLE AND EXCLUSIVE REMEDIES ARE AS SET FORTH HEREIN.