KVM Virtualization
Roadmap and Technology Update

Karen Noel
Senior Eng Manager
Red Hat, Inc.

Bhavna Sarathy
Senior Product Manager
Red Hat, Inc.

June 13, 2013
Why we believe KVM is the best virtualization platform

Performance
KVM holds the Top 6/11 virtual machine consolidation scores on SPECvirt (1)

Lower Cost
customers report up to 70% savings by using KVM (2)

Cross Platform
Support and certification for leading x86_64 operating systems including RHEL and Microsoft Windows (4)

Security
EAL4+ Certification (3) plus SE Linux enabling Mandatory Access Control between virtual machines

Cloud & Virtualization Management
Red Hat Open Stack for Cloud Virtualization and Red Hat Enterprise Virtualization for data-center Virtualization

(2) Source: Case study on Canary Islands Government migration from VMware to RHEV: http://www.redhat.com/resourcelibrary/case-studies/canary-islands-government-migrates-telecommunications-platform-from-vmware-to-red-hat
(3) Source: http://www.redhat.com/solutions/industry/government/certifications.html
(4) Source: http://www.redhat.com/resourcelibrary/articles/enterprise-linux-virtualization-support
KVM hypervisor in multiple Red Hat products

KVM is the foundation Virtualization technology in multiple Red Hat products
RHEV Hypervisor

- Prebuilt binary (ISO) with 300+ packages derived from RHEL
- Inherits performance, scalability, security and supportability of Red Hat Enterprise Linux
- Shares RHEL & KVM software and hardware ecosystem
KVM Architecture: Integrated Virtualization
KVM I/O Architecture

**Emulated Devices**
- Native drivers
- Compatibility over performance

**Virtio Devices**
- Paravirtualized
- Performance over compatibility

**Device Assignment**
- Native drivers
- Compatibility and Performance

---

**Components**
- Virtual Machine
- QEMU
- Tap
- Bridge
- Kernel
- Physical NIC
- vhost
Let's have a look at RHEL6
- **virtio-scsi**
  - Higher scaling
- **virtio-scsi**
  - & Passthru
- **Virtio-blk**
  - Data-Plane
- **Guest Agents**
  - RHEL & Win
- **Live Storage Migration**
- **Virtual PMU**
- **Live Snapshots Merge/Delete**
- **Para-virt End-of-int**
- **New CPU Models**
- **USB Re-direct**
- **USB 2.0 & Live Migration**
- **Virtual CPU In-Plug**
- **Largest x86 Guest (vCPU)**
- **Largest x86 Guest (vMemory)**
- **Hyper-V Relaxed Timing**
- **Windows 8 Windows 2012 Guests**
These were all introduced in RHEL 6.3 and RHEL 6.4 alone
SPECvirt_sc2010: RHEL 6 KVM Posts Industry Leading Results

RHEL (KVM) key enablers:
- SR-IOV
- Huge Pages
- NUMA
- Node Binding

http://www.spec.org/virt_sc2010/results/
Comparison based on best performing Red Hat and VMware solutions by CPU core count published at www.spec.org as of May 17, 2013. SPEC® and the benchmark name SPECvirt_sct® are registered trademarks of the Standard Performance Evaluation Corporation. For more information about SPECvirt_sc2010, see www.spec.org/virt_sc2010/.
Comparison based on best performing Red Hat and VMware solutions by cpu socket count published at www.spec.org as of May 30, 2013. SPEC® and the benchmark name SPECvirt_sc® are registered trademarks of the Standard Performance Evaluation Corporation. For more information about SPECvirt_sc2010, see www.spec.org/virt_sc2010/.
SPECvirt_sc2013 Tile

Single Tile Disk and Network I/O

Tile

Infrastructure Server Virtual Machine
- VM-to-VM File Server
- BeSim (DB simulator)
- Guest OS

Web Server Virtual Machine
- Web Server
- Guest OS

Mail Server Virtual Machine
- IMAP Mail Server
- Guest OS

Application Server Virtual Machine
- J2EE Application Server
- Guest OS

Batch Server Virtual Machine
- Batch Server
- Guest OS

Database Server Virtual Machine
- Database Server
- Guest OS

Blue = disk I/O
Green = network I/O

Virtualization Layer and Hardware

System Under Test (SUT)
- Client Hardware

SPECweb2005 driver
SPECimap driver
SPECjAppServer2004 driver
SPEC CPU2006 driver

#redhat #rhsummit
Comparison based on best performing Red Hat and VMware solutions by cpu socket count published at www.spec.org as of May 22, 2013. SPEC® and the benchmark name SPECvirt_sc® are registered trademarks of the Standard Performance Evaluation Corporation. For more information about SPECvirt_sc2013, see www.spec.org/virt_sc2013.
Virtualization – Performance

Leadership Virtual Disk I/O performance (RHEL 6.4)

Single Virtual Machine - IOPS
Direct Random I/O at 4KB Block Size

Host Server = Intel E7-8870@2.4GHz, 40 Cores, 256GB

1,577,684 IOPS!
Virtualization – Performance

- Only virtualized TPC-C result (RHEL 6.4 with KVM)
  - Virtualized was approx. 88% of bare-metal
  - Great $/tpmC: $0.51
  - Feb 2013: 1,320,082 tpmC, 0.51 $/tpmC, RHEL6.4 w/KVM, DB2 ESE 9.7
  - Apr. 2012: 1,503,544 tpmC, 0.53 $/tpmC, RHEL 6.2, DB2 ESE 9.7

See the details for these results at: [http://www.tpc.org/tpcc/results/tpcc_last_ten_results.asp](http://www.tpc.org/tpcc/results/tpcc_last_ten_results.asp).
Results referenced are current as of May 30, 2013. To view all TPC results, visit [www.tpc.org](http://www.tpc.org)
Let's look at

**Scalability**

Once we accept our limits, we go beyond them.

Albert Einstein
Single Guest Scalability – Virtual CPUs

RHEL6.4 vCPU = 2.5 x * vSphere 5.1 vCPU limits

<table>
<thead>
<tr>
<th>RHEL Releases</th>
<th>Virtual CPUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHEL5.5</td>
<td>16</td>
</tr>
<tr>
<td>RHEL6.0-6.2</td>
<td>64</td>
</tr>
<tr>
<td>RHEL6.3-6.4</td>
<td>160</td>
</tr>
</tbody>
</table>
Single Guest Scalability – Virtual Memory

RHEL6.4 vMemory = 2 * vSphere 5.1 vMemory limits

The very important topic of Security
KVM Security and Isolation - sVirt

Attacks on guests are isolated from the host kernel, any associated storage, as well as other virtual machines.
RHEL 6.5: Cryptography

- Para-Virtual Random Number Generator (RNG)
  - Provide true randomness in the guest for cryptographic purposes
  - RHEL host or RHEV Hypervisor feeds entropy to the virtual machines
  - Helps alleviate entropy starvation in guests
## Certification status

http://www.redhat.com/solutions/industry/government/certifications.html

### Government standards

Red Hat is committed to making your certification and compliance resources below should help you comply with a variety of standards.

**On this page:**
- Certifications
- Requirements
- Standards
- Project FAQs

**Certifications**

- **Common Criteria**
  - Common Criteria is an internationally recognized certification framework for IT security products.

**FAQs**

- **Can I use a product if it's "in evaluation"?**
  - Under NSTISSP.11, government customers must have a USG-approved protection profile. Failing that, you can use the product, but you must ensure that the product is in evaluation.

- **We've been through the Common Criteria process. Are there any potential constraints to that?**
  - Depending on the evaluation, the results might sound. When we're in evaluation, we're confident that the product is a matter of time. If you have any trouble getting a product to work, please get in touch with your DAA.

### Red Hat Enterprise Linux

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Certification</th>
<th>Description</th>
<th>Evaluation Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>EAL4+ CAPP</td>
<td></td>
<td>Evaluated</td>
</tr>
<tr>
<td>IBM</td>
<td>EAL4+ CAPP/RBACPP/LSPP</td>
<td></td>
<td>Evaluated</td>
</tr>
<tr>
<td>IBM</td>
<td>EAL4+ with KVM virtualization</td>
<td></td>
<td>Evaluated</td>
</tr>
<tr>
<td>IBM</td>
<td>EAL4+ OSPP, including Labeled Security, Advanced Audit, Advanced Management, and Virtualization Extended Modules</td>
<td>Evaluated</td>
<td></td>
</tr>
<tr>
<td>IBM</td>
<td>EAL4+ OSPP, including Labeled Security, Advanced Audit, Advanced Management</td>
<td>Evaluated</td>
<td></td>
</tr>
<tr>
<td>Dell</td>
<td>EAL4+</td>
<td></td>
<td>Evaluated</td>
</tr>
<tr>
<td>HP</td>
<td>EAL4+</td>
<td></td>
<td>Evaluated</td>
</tr>
<tr>
<td>SGI</td>
<td>EAL4+</td>
<td></td>
<td>Evaluated</td>
</tr>
</tbody>
</table>

### Red Hat Enterprise Linux

**Certifications**

- **Common Criteria**
  - Common Criteria is an internationally recognized certification framework for IT security products.

**FAQs**

- **Can I use a product if it's "in evaluation"?**
  - Under NSTISSP.11, government customers must have a USG-approved protection profile. Failing that, you can use the product, but you must ensure that the product is in evaluation.

- **We've been through the Common Criteria process. Are there any potential constraints to that?**
  - Depending on the evaluation, the results might sound. When we're in evaluation, we're confident that the product is a matter of time. If you have any trouble getting a product to work, please get in touch with your DAA.
KVM

Mission Critical
Mission critical refers to any factor of a system (software, hardware) whose failure will result in the failure of business operations.
KVM is hardened to run mission critical workloads

**Integrated Virtualization**
KVM leverages RHEL kernel, trusted for mission critical workloads

**Hardware Abstraction**
KVM has the highest single guest scalability (vCPU = 160, vMemory = 2TB)

**Security**
EAL4+ Certification (1) plus SE Linux enabling Mandatory Access Control between virtual machines

**Server Consolidation**
KVM runs heavy IO workloads such as large databases and holds the Top 6/11 virtual machine consolidation scores on SPECvirt (2)

**Resource Management**
Cgroups helps manage resources in virtual environments too

---

What else?

RHEL 6.4 Features
RHEL6.4: virtio-scsi

New storage architecture for KVM!

- virtio-scsi device = SCSI host bus adapter (HBA)
- Allows arbitrary number of devices per guest
  - Virtual hard drives and CDs
  - Pass-through physical SCSI device
- Supports SCSI pass-through and SCSI reservations
- Rich features – depends on the target, not virtio-scsi
- Drop-in physical disk replacement, friendlier for P2V and V2V
- RHEL6.4 guests and Windows guests (excluding XP)
RHEL6.4: virtio-blk “data-plane” mode

Single guest 1,577,684 IOPS! (4kb random I/O)
RHEL6.4: Para-virt End-of-Interrupt (PV-EOI)

Improved performance!

- Optimization for interrupt-intensive workloads
  - Up to 10% less CPU usage in some scenarios
- Reduces the number of context switches between the VM and the hypervisor.

Works out-of-the-box with all I/O types

- Particularly useful for high incoming network traffic
- Guest OS = RHEL 6.4 (PV calls in guest kernel)
RHEL6.4: New Virtualized CPU models

• New virtualized CPU models – Intel and AMD
  • Intel 4th generation Intel Core processors (Haswell, Ivy Bridge)
  • AMD Opteron Series 6300 (Abu Dhabi, Seoul)
  • Performance!
    • Leverage new processor features
    • Use new instructions
  • New CPU model definitions in KVM
    • RHEL host support – new virt features
    • Virtualized guest benefits, too
RHEL 6.4: USB 2.0 Improvements

- USB 2.0 redirection – with Spice
- USB 2.0 Live Migration Support
  - System admins can relocate VDI desktops for efficient load balancing, transparent to the end user!
- Migration of VMs with USB devices attached - mostly useful for “migration to file” (save VM)
Reliability Availability Serviceability
Virtual Reliability, Availability, Serviceability (vRAS)

What is vRAS?

- Maintain the Service Application Level Agreement (SLA)
- Meet workload demands, spikes at peak hours
- Maintain running virtual services and applications
- Perform live operations on running VMs and storage with no downtime
Virtual RAS Features

- Virtual CPU Hot-Plug
- Live Snapshots
- Live Snapshots Merge and Delete
- Live Storage Migration
RHEL 6.5: Virtual CPU Hot-Plug

- Flexibility - add additional vCPUs to running VMs
- No application downtime to adjust VM's compute capacity
RHEL 6.5: Virtual CPU Hot In-Plug

VM1

8 virtual CPUs

VM2

8 virtual CPUs
RHEL 6.5: Virtual CPU Enable/Disable

VM1

8 virtual CPUs
- 8 enabled

VM2

8 virtual CPUs
- 4 enabled
RHEL 6.5: Virtual CPU Enable/Disable

VM1

- 8 virtual CPUs
- 6 enabled

VM2

- 8 virtual CPUs
- 6 enabled
KVM-enabled features in RHEV 3.1
RHEV 3.2 Storage data-center features
Live Snapshots, Merge & Delete

- Snapshot a VM while the guest is running, preserving state and data of the VM at a given point in time

- Sample use cases:
  - Data-center admin saves snapshot of a running VM prior to disruptive upgrades to system
  - Backup scenario with periodic incremental snapshot/backup

Simple image chain:

- Merge toward root:
  - Root Base
  - Snap-1
  - Snap-2
  - Snap-3
  - Snap-4

- Merge toward active layer:
  - Root Base
  - Snap-2
  - Snap-3
  - Snap-4

Sample use cases:
- Data-center admin saves snapshot of a running VM prior to disruptive upgrades to system
- Backup scenario with periodic incremental snapshot/backup
RHEV 3.1 Storage data-center features
Live Storage Migration

Migrate a VM's storage files across storage arrays – no application downtime!
Let's have a look at Red Hat KVM Futures
KVM Futures – 2H 2013

- Device Mapper For QCOW2 Files
- Native Gluster FS Support
- virtio-trace
- I/O Throttling
- Serial Device Hot-plug
- Flow Control Improvements
KVM Futures – 2014

- Automatic NUMA Balancing
- Live Migration Thread
- Live Migration Optimizations
- USB 3.0
- QEMU Sandboxing
- PCI Bridging
- PCI-express Bus in guest
- Para-virt Page Faults
- QCOW2 Improvements
- VFIO – Device Assignment
- Assigned VGA Devices
- Para-virt Time vsyscall
- Data-Plane
- Multi-queue virtio
- Memory Hot-Plug
- Virtual PMU
KVM Community and Ecosystem
KVM: Strong hypervisor foundation for RHEV, Open Stack, IBM Cloud and many others.
Open Virtualization Alliance

An alliance that includes leading...
• Virtualization
• Datacenter and
• Cloud Solution Providers

244 Members and counting!

Governing Members

- Increase overall awareness of KVM
- Drive the adoption of KVM based solutions
- Foster an ecosystem of third-party solutions around KVM
- Promote Best Practices and highlight Customer Successes

to help...
Bringing the Community, Vendors and Users Together
What makes KVM the best Virtualization platform?

- Performance
- Features
- Open Source
- OpenStack
- RHEV
- Partners
- Security
Questions?
Related Summit Sessions

- Hypervisor Technology Comparison & Migration
  - Fri 9:45, Room 313
- Partner Pavilion
  - Demo of the latest KVM features
  - In the RHEV booth
Thank you