oVirt Architecture

Itamar Heim
Director, RHEV-M Engineering, Red Hat
oVirt Engine

Large scale, centralized management for server and desktop virtualization

Based on leading performance, scalability and security infrastructure technologies
Kenrel-based Virtual Machine (KVM)

- Included in Linux kernel since 2006
- Runs Linux, Windows and other operating system guests
- Advanced features
  - Live migration
  - Memory page sharing
  - Thin provisioning
  - PCI Pass-through
- KVM architecture provides high “feature-velocity” – leverages the power of Linux
Linux as a Hypervisor?

- What makes up a hypervisor?
  - Hardware management
  - Device drivers
  - I/O Stack
  - Resource Management
  - Scheduling
  - Access Control
  - Power Management
  - Memory Manager
  - Device Model (emulation)
  - Virtual Machine Monitor
Linux as a Hypervisor?

- What makes up a hypervisor?
  - Hardware management
  - Device drivers
  - I/O Stack
  - Resource Management
  - Scheduling
  - Access Control
  - Power Management
  - Memory Manager
  - Device Model (emulation)
  - Virtual Machine Monitor

Operating System Kernel
Linux as a Hypervisor?

How well does Linux perform as a hypervisor? Isn't Linux a general purpose operating system?

Linux is architected to scale from the smallest embedded systems through to the largest multi-socket servers

- From cell phones through to mainframes

KVM benefits from mature, time tested infrastructure

- Powerful, scalable memory manager
- Robust security infrastructure
- High performance network stack
- Versatile storage infrastructure – iSCSI, FC, NAS, multipath, etc
- Rich ecosystem of supported hardware systems
Linux as a Hypervisor?

How well does Linux perform as a hypervisor? Isn't Linux a general purpose operating system?

Over the last 4 years features have been added to Linux to provide a better infrastructure for a hypervisor

- Scheduler enhancements
  Improved scalability and reduced latency

- Enhancements to memory manager
  Advanced features such as memory page sharing and compression

- Improvements to Block I/O subsystem
  - Better performance, automated alignment, etc
SPECvirt_sc2010

Vendor neutral virtualization benchmarks
Comprised of application specific benchmarks running inside “tiles”

Each tile runs 6 virtual machines

- Application Server
- Database Server
- Mail Server
- Web Server
- Infrastructure Server
- Idle Server

Each VM runs a benchmark, eg SpecWeb, SPECjAppServer, SPECmail
and must meet specific
KVM leads the pack in 2, 4, 8 socket systems for SPECvirt
Including the largest benchmark results with over 400 Vms

Score : 7067 @ **432 VMs** (72 tiles)

Processor: Intel Xeon E7-4870 (80 cores, 8 chips, 10 cores/chip, 2 threads/core)
Memory: 2 TB (128 x 16 GB, Quad Rank x4 PC3-8500 CL7 ECC DDR3 1066MHz LP RDIMM)

http://www.spec.org/virt_sc2010/
Red Hat Enterprise Virtualization
Competitive Landscape

- InfoWorld “shootout” 2011
  - Independent analysis of leading virtualization platforms
  - After <18 months Red Hat has overtaken Citrix & Microsoft in performance and functionality


<table>
<thead>
<tr>
<th>Test Center Scorecard</th>
<th>InfoWorld</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Management</td>
</tr>
<tr>
<td></td>
<td>25%</td>
</tr>
<tr>
<td>Citrix XenServer 5.6.1</td>
<td>7</td>
</tr>
<tr>
<td>Microsoft Windows Server 2008 R2 Hyper-V</td>
<td>8</td>
</tr>
<tr>
<td>Red Hat Enterprise Virtualization for Servers 2.2</td>
<td>8</td>
</tr>
<tr>
<td>VMware vSphere 4.1</td>
<td>9</td>
</tr>
</tbody>
</table>
Security

oVirt inherits the security features of Linux

SELinux security policy infrastructure

  Provides protection and isolation for virtual machines and host

  Compromised virtual machine cannot access other VMs or host

sVirt Project

  Sub-project of NSA's SELinux community. Provides “hardened” hypervisors

  Multilevel security. Isolate guests

  Contain any hypervisor breaches
oVirt Node

- Standalone hypervisor
  - Small footprint < 100MB
  - Customized 'spin' of Fedora + KVM
  - 'Just enough' Fedora to run virtual machines
  - Runs on all RHEL hardware with Intel VT/AMD-V CPUs
  - Easy to install, configure and upgrade
  - PXE boot, USB boot, CD or Hard drive

oVirt Engine Architecture
oVirt Node vs. Full Host

**oVirt Node**
- Less than 100 MB
- Pre-configured, no Linux skills needed.

**Full Host**
- Flexible
- Add monitoring agents, scripts etc. Leverage existing Fedora infrastructure.
- Hybrid mode capable
Ovirt Engine

Ovirt Engine Architecture
Management Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Availability</td>
<td>Restart guest VMs from failed hosts automatically on other hosts</td>
</tr>
<tr>
<td>Live Migration</td>
<td>Move running VM between hosts with zero downtime</td>
</tr>
<tr>
<td>System Scheduler</td>
<td>Continuously load balance VMs based on resource usage/policies</td>
</tr>
<tr>
<td>Power Saver</td>
<td>Concentrate virtual machines on fewer servers during off-peak hours</td>
</tr>
<tr>
<td>Maintenance Manager</td>
<td>No downtime for virtual machines during planned maintenance windows.</td>
</tr>
<tr>
<td></td>
<td>Hypervisor patching</td>
</tr>
<tr>
<td>Image Management</td>
<td>Template based provisioning, thin provisioning and snapshots</td>
</tr>
<tr>
<td>Monitoring &amp; Reporting</td>
<td>For all objects in system – VM guests, hosts, networking, storage etc.</td>
</tr>
<tr>
<td>OVF Import/Export</td>
<td>Import and export VMs and templates using OVF files</td>
</tr>
<tr>
<td>V2V</td>
<td>Convert VMs from VMware and RHEL/Xen to RHEV</td>
</tr>
</tbody>
</table>
High Availability

- Build a highly available enterprise infrastructure
- Continually monitor host systems and virtual machines
- Automatically restart virtual machines in case of host failure
  - Restart virtual machine on another node in the cluster
- Use live migration to “fail-back” a VM to its original host when the server is restored

Edit Server Virtual Machine

<table>
<thead>
<tr>
<th>General</th>
<th>Highly Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Console</td>
<td></td>
</tr>
<tr>
<td>Host</td>
<td></td>
</tr>
<tr>
<td><strong>High Availability</strong></td>
<td></td>
</tr>
<tr>
<td>Priority for Run/Migrate Queue:</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>

Guests

Hypervisor

Servers

Hypervisor

Hypervisor

Storage
Live Migration

- Dynamically move virtual machines between hosts
  - No service interruption
  - Applications continue to run
- Migrate even I/O intensive workloads such as databases
  - Perform hardware maintenance without application downtime
- Dynamically balance workloads between host systems
System Scheduler

- Dynamically balance workloads in the data center.
- Automatically live migrate virtual machines based on resources.
- Define custom policies for distribution of virtual machines.

Maintain consistent resource usage across the enterprise data center.
Power Saver

Define policies to optimize workload on a fewer number of servers during “off-peak” hours
## Management Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Availability</td>
<td>Restart guest VMs from failed hosts automatically on other hosts</td>
</tr>
<tr>
<td>Live Migration</td>
<td>Move running VM between hosts with zero downtime</td>
</tr>
<tr>
<td>System Scheduler</td>
<td>Continuously load balance VMs based on resource usage/policies</td>
</tr>
<tr>
<td>Power Saver</td>
<td>Concentrate virtual machines on fewer servers during off-peak hours</td>
</tr>
<tr>
<td>Maintenance Manager</td>
<td>No downtime for virtual machines during planned maintenance windows.</td>
</tr>
<tr>
<td></td>
<td>Hypervisor patching</td>
</tr>
<tr>
<td>Image Management</td>
<td>Template based provisioning, thin provisioning and snapshots</td>
</tr>
<tr>
<td>Monitoring &amp; Reporting</td>
<td>For all objects in system – VM guests, hosts, networking, storage etc.</td>
</tr>
<tr>
<td>OVF Import/Export</td>
<td>Import and export VMs and templates using OVF files</td>
</tr>
<tr>
<td>V2V</td>
<td>Convert VMs from VMware and RHEL/Xen to RHEV</td>
</tr>
</tbody>
</table>
Virtual Desktop Infrastructure (VDI)

Centralized management, security and policy enforcement

Virtual desktops with user experience of a physical PC

- Multiple monitors
- HD quality video
- Bi-directional audio/video for VoIP or video-conferencing
- Smartcard support
- USB support

Industry leading density of virtual desktops/server
Red Hat Enterprise Virtualization
RHEV 3.0 Key Initiatives

- Move from proprietary to open technologies
- Remove dependency on Windows (But maintain interoperability with Windows)
- Deliver new features and releases in parallel
- Build Open Source community project around open virtualization
History

- Qumranet
  - KVM
  - SPICE
- SolidICE --> RHEV-M C# --> RHEV-M Java --> oVirt
- C# --> Java
  - using automatic conversion approach for core and UI[1]
- VDSM
- oVirt Node

Things have changed

Things have evolved

There are a lot of good ideas

There is a lot to refactor/change/do

This is where we are

Let's get to work...
### oVirt Engine Architecture

#### Admin Portal

Logged in user: vdcadmin | Sign out | Configure | About | Guide

**Vms:**

<table>
<thead>
<tr>
<th>Name</th>
<th>Cluster</th>
<th>Host</th>
<th>IP Address</th>
<th>Memory</th>
<th>CPU</th>
<th>Network</th>
<th>Display</th>
<th>Status</th>
<th>Uptime</th>
<th>Logged in User</th>
</tr>
</thead>
<tbody>
<tr>
<td>baz-xp-1</td>
<td>cluster-fc-23-N</td>
<td>nari12</td>
<td>0%</td>
<td>12%</td>
<td>0%</td>
<td>Spice</td>
<td>Up</td>
<td>2 h</td>
<td>vdcadmin</td>
<td></td>
</tr>
<tr>
<td>fc-p-bk-1</td>
<td>cluster-fc-23-N</td>
<td>white-vdsh</td>
<td>0%</td>
<td>6%</td>
<td>0%</td>
<td>Spice</td>
<td>Down</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iscsi-bk-1</td>
<td>cluster-iscsi-23-N</td>
<td></td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>Spice</td>
<td>Up</td>
<td>23 min</td>
<td>vdcadmin</td>
<td></td>
</tr>
<tr>
<td>jboss-srv-1</td>
<td>cluster-fc-23-N</td>
<td>white-vdsh</td>
<td>0%</td>
<td>3%</td>
<td>0%</td>
<td>Spice</td>
<td>Up</td>
<td>2 h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nfsvm</td>
<td>cluster-nfs-23-A</td>
<td></td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### General

- **Name:** baz-xp-1
- **Description:** xp-sg3-wu
- **Operating System:** Windows XP
- **Default Display Type:** Spice
- **Defined Memory:** 2048 MB
- **Physical Memory Guaranteed:** 2049 MB
- **Number of CPU Cores:** 1 (1 Cores, 1 Cores per Socket)
- **Number of Monitors:** 1
- **USB Policy:** Enabled
- **Resides on Storage Domain:** sd-fc-23-02

**Origin:** RHEV
**Run On:** Any Host in Cluster
**Custom Properties:** Not Configured
**Domain:** qa.lab.tlv.redhat.com
**Time Zone:** GMT Standard Time

Add Virtual Machine
Power User Portal - Resources
oVirt High Level Architecture

- Postgres
- AD
- IPA
- Shared Storage FC/iSCSI/NFS
- oVirt Engine
  - Java
  - REST
- Guest agent
- Linux VM
- Win VM
- libvirt
- VDSM
- Host | Node
- Local Storage
- Linux/Windows client
- Admin Portal
  - gwt
- SDK/CLI
  - python
- User Portal
  - gwt
- SPICE
Engine Core (Backend)

- VM & Template Life Cycle: create, schedule, snapshot
- Load Balancing
- HA
- Storage Configuration & Monitoring
- Network Configuration & Monitoring
- Host Register/Install
- Host Monitoring
- Host Maintenance
- Host Fencing
- Authentication, Authorization Audit
- Inventory
Authentication

- Built-in user admin@internal
- AD, IPA integration
  - Kerberos authentication
  - LDAP - user info, group membership
  - Multiple domains, trusts, etc.
  - Cached for searches, not for login
- Next
  - Open LDAP (patch ready)
  - Internal users (picketlink?)
  - Linux users?
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Multi Level Admin

- Users
- Groups
- Roles
- Permissions

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SuperUser</td>
<td>Roles management administrator</td>
</tr>
<tr>
<td>RHEVMUser</td>
<td>RHEVM user</td>
</tr>
<tr>
<td>RHEVMPowerUser</td>
<td>RHEVM power user</td>
</tr>
<tr>
<td>ClusterAdmin</td>
<td>Cluster administrator</td>
</tr>
<tr>
<td>DataCenterAdmin</td>
<td>Data Center administrator</td>
</tr>
<tr>
<td>StorageAdmin</td>
<td>Storage administrator</td>
</tr>
<tr>
<td>HostAdmin</td>
<td>Host administrator</td>
</tr>
<tr>
<td>NetworkAdmin</td>
<td>Network administrator</td>
</tr>
<tr>
<td>VmAdmin</td>
<td>Vm administrator</td>
</tr>
<tr>
<td>VmPoolAdmin</td>
<td>Vm-Pool administrator</td>
</tr>
<tr>
<td>TemplateAdmin</td>
<td>Template administrator</td>
</tr>
<tr>
<td>TemplateUser</td>
<td>Template User</td>
</tr>
</tbody>
</table>
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Multi Level Admin

- Users
- Groups
- Roles
- Permissions
Database

- Moved from SQL Server to Postgres
- JDBC based
- Next
  - Hibernate
  - Scheme upgrade management
REST API

- New RESTful API for integration with oVirt Engine
  - REST interface exposed for all API functions
  - Developed in upstream RHEV-M API project (before oVirt)
RHEVM-API Upstream Project

Community project to deliver RESTful API for RHEV 2.2

https://fedorahosted.org/rhevm-api/

• Provides preview of 3.0 RESTful API
  • Draft implementation of new API
  • Runs on RHEL 2.2 wraps PowerShell
    • Allows early testing of API for customers and partners
    • 3.0 Implementation based on Java backend engine
  • Will be consolidated into oVirt
RHEV-M API Definition

This is an effort to define an official REST API for RHEV-M.

See here for the API reference guide.

Download a milestones release and follow this installation guide to get started.

There are three parts to the project:

1. The REST API definition - Java interfaces defining the API methods and base classes defining the API's object model. These are annotated with JAX-WS and JAXB annotations.
2. A PowerShell wrapper implementation - a servlet implementing the API by wrapping the RHEV-M 2.2 PowerShell API.
3. A mock implementation - a servlet implementing the API using mocked-up resources. This allows one to experiment with the API without needing a RHEV-M installation.

It is planned that a future version of RHEV-M will implement the API without wrapping the existing PowerShell API.

REST API Reference

- HTML
- Single Page HTML
- PDF

Python API

In addition to remote access protocols like the REST "API", we are developing a reference API library for coding access. These may use the REST protocol or not as needed. The goal of the API library is to present the RHEV-M objects in a form that will be most comfortable for the developer in the target language. As much as is reasonable communications and API internals should be hidden.

- Python API

Disclaimers

The API definition is in its early stages. It may yet change substantially before being becoming the the official API.

The PowerShell wrapper is for experimentation only and not a supported Red Hat solution.

Get Involved

See our Roadmap and issues list

Done
RESTful Web Service

- Stands for Representational State Transfer
- Modeling entity actions around HTTP verbs
  - GET
  - PUT
  - POST
  - DELETE
- Still uses 'actions' for some state changes
- Self describes – entity navigation and actions
Welcome

```xml
<api>
  <link rel="capabilities" href="rhev-m-api/capabilities"/>
  <link rel="clusters" href="rhev-m-api/clusters"/>
  <link rel="clusters/search" href="rhev-m-api/clusters?search={query}"/>
  <link rel="datacenters" href="rhev-m-api/datacenters"/>
  <link rel="datacenters/search" href="rhev-m-api/datacenters?search={query}"/>
  <link rel="events" href="rhev-m-api/events"/>
  <link rel="events/search" href="rhev-m-api/events?search={query}"/>
  <link rel="hosts" href="rhev-m-api/hosts"/>
  <link rel="hosts/search" href="rhev-m-api/hosts?search={query}"/>
  <link rel="networks" href="rhev-m-api/networks"/>
  <link rel="roles" href="rhev-m-api/roles"/>
  <link rel="storagedomains" href="rhev-m-api/storagedomains"/>
  <link rel="storagedomains/search" href="rhev-m-api/storagedomains?search={query}"/>
  <link rel="tags" href="rhev-m-api/tags"/>
  <link rel="templates" href="rhev-m-api/templates"/>
  <link rel="templates/search" href="rhev-m-api/templates?search={query}"/>
  <link rel="users" href="rhev-m-api/users"/>
  <link rel="groups" href="rhev-m-api/groups"/>
  <link rel="domains" href="rhev-m-api/domains"/>
  <link rel="vm_pools" href="rhev-m-api/vm_pools"/>
  <link rel="vm_pools/search" href="rhev-m-api/vm_pools?search={query}"/>
  <link rel="vms" href="rhev-m-api/vms"/>
  <link rel="vms/search" href="rhev-m-api/vms?search={query}"/>
  <system_version revision="428" build="0" minor="6" major="4"/>
</api>

<summary>
  <vms>
    <total>22</total>
    <active>5</active>
  </vms>

  <hosts>
    <total>6</total>
    <active>5</active>
  </hosts>

  <users>
    <total>2</total>
  </users>
</summary>

ovirt Engine Architecture
<hosts>
  <host id="15896dce-edd0-415c-a524-c9b02f278895" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895">
    <name>nari1</name>
    <actions>
      <link rel="activate" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/activate"/>
      <link rel="fence" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/fence"/>
      <link rel="deactivate" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/deactivate"/>
      <link rel="approve" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/approve"/>
      <link rel="iscsitool" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/iscsitool"/>
      <link rel="iscsidisc" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/iscsidisc"/>
      <link rel="commitnetconfig" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/commitnetconfig"/>
    </actions>
    <link rel="storage" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/storage"/>
    <link rel="nics" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/nics"/>
    <link rel="tags" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/tags"/>
    <link rel="permissions" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/permissions"/>
    <link rel="statistics" href="/rhev/api/v1/hosts/15896dce-edd0-415c-a524-c9b02f278895/statistics"/>
    <address>nari1.englab.tvu.redhat.com</address>
    <status>UP</status>
    <cluster id="4a5ba0e-7c6d-4d75-9aba-d60f3aa188db" href="/rhev/api/v1/clusters/4a5ba0e-7c6d-4d75-9aba-d60f3aa188db">
      <port>54321</port>
      <storage_manager>false</storage_manager>
      <power_management>
        <enabled>false</enabled>
        <options/>
      </power_management>
      <ksm>
        <enabled>false</enabled>
        </ksm>
      <transparent_hugopages>
        <enabled>true</enabled>
        </transparent_hugopages>
      <iscsi>
        <initiator>iqn.1994-05.com.redhat:a7af41a35b0</initiator>
      </iscsi>
    </cluster>
  </host>
</hosts>
Host networks collection

This XML file does not appear to have any style information associated with it. The document tree is shown below.

```xml
<host_nics>
  <host_nic id="dbb39d06-3aeaf468c-83e6-88eae0a3f346" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/dbb39d06-3aeaf468c-83e6-88eae0a3f346"/>
    <name>eth0</name>
  </host_nic>
  <actions>
    <link rel="attach" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/dbb39d06-3aeaf468c-83e6-88eae0a3f346/attach"/>
    <link rel="detach" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/dbb39d06-3aeaf468c-83e6-88eae0a3f346/detach"/>
  </actions>
  <link rel="statistics" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/dbb39d06-3aeaf468c-83e6-88eae0a3f346/statistics"/>
  <host id="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/>
</host_nics>

<network>
  <name>rhevm</name>
</network>

<host_nic>
  <mac address="78:E7:D1:E4:8E:92"/>
  <ip netmask="255.255.252.0" address="10.35.16.151"/>
</host_nic>

<host_nics>
  <host_nic id="0d98b08c-9b42-45a4-a226-b7dd3f0854cf" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/0d98b08c-9b42-45a4-a226-b7dd3f0854cf"/>
    <name>eth1</name>
  </host_nic>
  <actions>
    <link rel="attach" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/0d98b08c-9b42-45a4-a226-b7dd3f0854cf/attach"/>
    <link rel="detach" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/0d98b08c-9b42-45a4-a226-b7dd3f0854cf/detach"/>
  </actions>
  <link rel="statistics" href="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/nics/0d98b08c-9b42-45a4-a226-b7dd3f0854cf/statistics"/>
  <host id="/rhevm-api-hosts/15896dce-eddd-415c-a524-c9b02f278895/>
  <mac address="78:E7:D1:E4:8E:93"/>
  <ip netmask="" address=""/>
</host_nics>
```

Host networks collection

This XML file does not appear to have any style information associated with it. The document tree is shown below.
Create a Virtual Machine from a Template

POST http://10.35.1.1/rhevm-api/vms

<vm>
  <name>my_new_vm</name>
  <cluster id="99408929-82cf-4dc7-a532-9d998063fa95" />
  <template id="00000000-0000-0000-0000-000000000000" />
</vm>

curl -v -u "vdcadmin@qa.lab.tlv.redhat.com"
  -H "Content-type: application/xml"
  -d '<vm><name>my_new_vm</name><cluster id="99408929-82cf-4dc7-a532-9d998063fa95" /><template id="00000000-0000-0000-0000-000000000000" /></vm>'
'http://10.35.1.1/rhevm-api/vms'
Changing a property

PUT http://10.35.1.1/rhevm-api/vms/2496a177-e7c8-4f82-bf3d-2d0f73444990

<vm>
  <name>test_vm_new_name</name>
</vm>

echo "<vm><name>test_vm_new_name1</name></vm>" > /tmp/upload.xml
curl -v -u "vdcadmin@qa.lab.tlv.redhat.com"
  -H "Content-type: application/xml"
  -T /tmp/upload.xml
  'http://10.35.1.1/rhevm-api/vms/2496a177-e7c8-4f82-bf3d-2d0f73444990'
Adding a Virtual Disk

POST
http://10.35.1.1/rhevm-api/vms/2496a177-e7c8-4f82-bf3d-2d0f73444990/disks

<disk>
  <storage_domain id="3e1c96f0-8667-4a80-9689-af1337395dea" href="/rhevm-api/storagedomains/3e1c96f0-8667-4a80-9689-af1337395dea" />
  <size>1073741824</size>
  <type>system</type>
  <interface>virtio</interface>
  <format>raw</format>
  <sparse>true</sparse>
  <bootable>true</bootable>
  <wipe_after_delete>false</wipe_after_delete>
  <propagate_errors>false</propagate_errors>
</disk>

• curl -v -u "vdcadmin@qa.lab.tlv.redhat.com"
  -H "Content-type: application/xml"
  -d '<disk>...</disk>}' http://...
What Else?

- Data warehouse
- Reports (based on jasperforge.org)
- Tools
  - Notifications
  - Config
  - Iso uploader
  - Log collector
oVirt Data Warehouse

- ETL based on talendforge.org
- Periodic polling from operational DB
- Types of data
  - Config with version tracking
  - Statistics – aggregated hourly/daily
- API is view based
oVirt Reports

- Jasper allows to import/export reports definitions
- Rich reporting engine
  - Report scheduling
  - Filters
  - Export to various formats
  - Report creation studio
- Next
  - Integrated in web admin
# oVirt Reports

## Repository

<table>
<thead>
<tr>
<th>Folders</th>
<th>Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root</td>
<td></td>
</tr>
<tr>
<td>Organizations</td>
<td></td>
</tr>
<tr>
<td>oVirt Reports</td>
<td></td>
</tr>
<tr>
<td>Reports</td>
<td></td>
</tr>
<tr>
<td>Executive</td>
<td></td>
</tr>
<tr>
<td>Inventory</td>
<td></td>
</tr>
<tr>
<td>Service Level</td>
<td></td>
</tr>
<tr>
<td>Trend</td>
<td></td>
</tr>
<tr>
<td>Resources</td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td></td>
</tr>
<tr>
<td>Themes</td>
<td></td>
</tr>
<tr>
<td>default</td>
<td></td>
</tr>
<tr>
<td>ovirt-reports-theme</td>
<td></td>
</tr>
</tbody>
</table>

### Active Virtual Machines by OS (BR19)
`/organizations/ovirtreports/Reports/Executive/active_vms_by_os_br19`

The report contains comparative measurements number of running virtual machines and OS usage in for a selected cluster and a selected virtual machine's type within the requested period.

### Cluster Capacity Vs Usage (BR19)
`/organizations/ovirtreports/Reports/Executive/cluster_capacity_vs_usage_br19`

This report contains charts displaying host's resources usage measurements (CPU core; physical Memory) and charts displaying virtual machine's resources usage measurements (virtual machine's total vCPU, Virtual Memory size) for a selected cluster.

### Host OS Break Down (BR22)
`/organizations/ovirtreports/Reports/Executive/host_os_break_down_BR22`

This report contains a table and a chart displaying the number of hosts for each OS version for a selected cluster within a requested period.

### Summary of Host Usage Resources (BR17)
`/organizations/ovirtreports/Reports/Executive/summary_of_host_usage_resources_br17`

The report contains a scattered chart of CPU and memory usage date within a requested period and for a selected cluster.
Notification Service

- oVirt allows registration to certain audit events
- The notification service sends emails per audit message to relevant users
- Also monitors engine itself
Configuration tool

• The configuration utility allows changing oVirt advanced configuration options

• Sample commands
  • engine-config --list
  • engine-config --get <key_name>
  • engine-config -all
  • engine-config --set <key_name>=<value>

• Special config for authentication domains: manage-domains
ISO Uploader

• Iso uploader is a utility to upload iso files to the iso domain, to allow bootstrapping guests from them
• Admin can just copy the files to the iso domain
• Supports both scp and nfs based copies
• Integrates with the REST API to allow using storage domain name instead of specific NFS path
Log Collector

- The log collector utility helps collecting logs and configuration data for troubleshooting
- Written as a linux script launching sos plugins
- Collects the data from engine and nodes
oVirt Guest Agent

- The guest agent provides additional information to oVirt Engine, such as guest memory usage, guest ip address, installed applications and sso.
- Python code, available for both linux and windows guests
- Communication is done over virtio-serial
- SSO for windows is based on a gina module for XP and a credential provider for windows 7
- SSO for RHEL 6 is based on a PAM module with support for both KDE and Gnome
Guest

- balloon
- Virtio-net
- Virtio-block
- USB
- Spice driver
- guest Agent

SSO

RHEL / RHEV-H
RHEV-M Guest Agent - SSO for RHEL

Diagram:
- virtio-serial
  - RHEV-agent
    - D-Bus
      - Plug-in
        - Greeter's API
          - Greeter
            - PAM's API
              - PAM
    - Abstract socket
      - PAM Module
        - PAM Stack
oVirt Host Agent - VDSM

- Covers all functionality required by oVirt Engine
- Configures host, networking and shared storage
- Uses libvirt for VM life cycle operations
oVirt Host Agent - VDSM
oVirt Storage

- VDSM manages a Storage Pool, comprised of Storage Domains
- **Storage Pool** - a VM repository that contains meta data about storage domains, storage tasks, VMs, locks, etc.
- **Storage Domain** - a disk image repository
- **Disk Image** - a collection of volumes (chain of snapshots)
- **Volume** - stored as files in NFS, and as Logical Volumes for FC/iScsi
- Thin provisioning for SAN supported (storage mailbox based)
Storage Pool Manager

The SPM runs on an arbitrary host (chosen by oVirt Engine)

oVirt Engine requires SPM to be running in order to add storage

If SPM host dies/disappears, RHEV-M causes SPM to start on a different host
oVirt Storage “Clustering”

- A Storage Pool is implemented as a managed cluster
- Manager is oVirt Engine, running on a node external to the hosts using the storage pool
- Heartbeats and fencing are used in case of node failures
- Storage based leased locks used as another layer of protection
- Clustering wise - VMs are mostly single reader/writer - locks mostly needed to handle failures
- Can easily create a cluster of >100 nodes
Hooks

• “Hook” mechanism for customization
  • Allows administrator to define scripts to modify VM operation
    • eg. Add extra options such as CPU pinning, watchdog device, direct LUN access, etc
  • Allows oVirt to be extended for new KVM features before full integration is done
  • An easy way to test a new kvm/libvirt/linux feature
Hooks
Hooks
Hooks

- Hook scripts are called at specific VM lifecycle events
  - VDSM (management agent) Start
  - Before VM start
  - After VM start
  - Before VM migration in/out
  - After VM migration in/out
  - Before and After VM Pause
  - Before and After VM Continue
  - Before and After VM Hibernate
  - Before and After VM resume from hibernate
  - On VM stop
  - On VDSM Stop
- Hooks can modify a virtual machines XML definition before VM start
- Hooks can run system commands – eg. Apply firewall rule to VM
# Hooks

**Edit Server Virtual Machine**

<table>
<thead>
<tr>
<th>General</th>
<th>Custom Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Console</td>
<td></td>
</tr>
<tr>
<td>Host</td>
<td></td>
</tr>
<tr>
<td>High Availability</td>
<td></td>
</tr>
<tr>
<td>Resource Allocation</td>
<td></td>
</tr>
<tr>
<td>Boot Options</td>
<td></td>
</tr>
</tbody>
</table>

**Custom Properties**

![Image of Edit Server Virtual Machine dialog box](image_url)
Hooks

Hooks installed in /usr/libexec/vdsm/hooks

```
[root@host1 ~]# cd /usr/libexec/vdsm/hooks/
[root@host1 hooks]# ls -l
```
```
total 68
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vdsm_stop
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_cont
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_dehbernate
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_destroy
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_hibernate
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_migrate_destination
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_migrate_source
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_pause
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 after_vm_start
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vdsm_stop
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_cont
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_dehbernate
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_destroy
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_hibernate
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_migrate_destination
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_migrate_source
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_pause
drwxr-xr-x. 2 root root 4096 Apr 12 03:55 before_vm_start
```

```
[root@host1 hooks]#
```

<table>
<thead>
<tr>
<th>Event Name</th>
<th>Script Name</th>
<th>Property Name</th>
<th>Property Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>before_vm_start</td>
<td>10_faqemu</td>
<td>md5</td>
<td>2c352c04ecf994</td>
</tr>
</tbody>
</table>
Hooks

[Diagram showing the relationship between VDSM, libvirt, hook script, and script]
Hooks

```python
#!/usr/bin/python
import os
import sys
import hooking
import traceback

pinvcpu = "0" (use the first cpu)
pinvcpu = "1-4" (use cpus 1-4)
pinvcpu = "^3" (don't use cpu 3)
pinvcpu = "1-4,^3,6" (or all together)

if os.environ.has_key('pinvcpu'):
    try:
        domxml = hooking.read_domxml()

        vcpu = domxml.getElementsByTagName('vcpu')[0]

        if not vcpu.getAttribute('pinvcpu'):
            sys.stderr.write('pinvcpu: pinning cpu to: %s
                           %s.environ['pinvcpu'])
            vcpu.setAttribute('cpuset', os.environ['pinvcpu'])
            hooking.write_domxml(domxml)
        else:
            sys.stderr.write('pinvcpu: cpuset attribute is present in vcpu, doing nothing\n')
    except:
        sys.stderr.write('pinvcpu: [unexpected error]: %s\n                            traceback.format_exc()')
        sys.exit(2)
```
Sample Hooks

- CPU pinning
- SR/IOV
- Smart card
- Direct LUN
- Hugepages
- Promiscuous mode network interface
- Cisco VN-Link
- Fileinject
- Floppy
- Hostusb
- Isolatedprivatevlan
- Numa
- Qos
- Scratchpad
- smbios
On the Horizon - Infra

- Engine – JBoss AS 7, modular lighter engine
- Engine – custom hooks
- Engine – vdsm communication protocol and transport
- API – non admin api
- Reports – integrated in web admin
- Code cleanups, refactoring, unitests, etc
On the Horizon - Features

- Live snapshots
- Live storage migration
- Quotas
- Hot plug
- Multiple storage domains
- Shared disks
- iScsi disk
- Shared file system support
- Storage array integration
- Qbg/Qbh
- virt-resize, pv-resize
- Progress bars
- Stable pci addresses
- Network types
- Backup API
- SLA
- SDM
- Many many more...
THANK YOU!

http://www.ovirt.org