

oVirt 3.5 Storage Features Overview

Allon Mureinik
amureini@redhat.com
Red Hat

Federico Simoncelli
fsimonce@redhat.com
Red Hat

April 2014

- Import Existing Data Domain
 - And the required Store OVF on Any Domain
- SANLock Fencing
- Live Merge (delete snapshot)
 - And the required VM Async Tasks
- Snapshot Overview in the Webadmin
- Q&A

Live Merge – The Problem



- Large chains of snapshots may degrade a VM's performance
 - Not to mention the disk space they waste
- Often, taking a live snapshot is unintended – it's just a side effect of using another feature
 - E.g., Live Storage Migration, using Backup API
- Shutting down a VM (for a possibly prolonged time) just to merge these snapshots is **unacceptable**.

Live Merge - Solution



- Simple – merge a snapshot without bringing the VM down
- This has been supported in QEMU for a while now...
- All we need to do is to consume it...

- VM Tasks
 - We need a new type of operation - a long running task that's handled by HSM
 - Somewhat parallel effort to future plans for redesigning the tasks framework
 - Leveraging the current work on Command Coordination
- Several VDSM calls that need to be coordinated by a single Engine flow
 - Another good candidate for SEAT
- Was blocked till we had an implementation of single disk snapshots (oVirt 3.4)
 - Critical in order to handle error flows

Snapshot Overview – The Problem



- When you create a disk, you get to choose which Storage Domain you'll use
 - And we even report the free space on each one
- Eventually, you'd want to manage this space effectively
 - And free up some space in your critical domains
- There's no reasonable way of doing this

Snapshot Overview – The Solution



... Storage ...

Name	...	Total Space	Free Space
Data1	...	50GB	30GB
...			

... Snapshots ...

Remove

	Size	Creation Date	Disk Alias	Snapshot Description	VM Name	...
	10GB					
	5GB					
	2GB					

Snapshot Overview – The Solution



... Storage ...

Name	...	Total Space	Free Space
Data1	...	50GB	30GB
...			

... Snapshots ...

Remove

	Size	Creation Date	Disk Alias	Snapshot Description	VM Name	...
	10GB					
	5GB					
	2GB					

Snapshot Overview – Implementation Details

- Enabling multiple selection on snapshots
 - A mix of MultipleActionRunner and SEAT
- Doesn't quite fit our REST approach
 - Missing a disk snapshot business entity
 - And there isn't a decent way of implementing multiple selection
- Since we're managing snapshots in a storage context, and not a VM context, this feature was also blocked by single disk snapshots (oVirt 3.4)

- The problem – operations not supported today:
 - Detach a Storage Domain from a Data Center (if it contains VMs/Templates/Disks)
 - Move a Storage Domain from a Data Center to another
 - Import a Storage Domain from a known storage
- These operations have few implications on existing VMs
 - How to handle VMs with disks on the Storage Domain we are detaching
 - How to select the target Clusters for the VMs we are importing
 - How to handle the import of VMs with additional disks on an another Storage Domain (not present in the same Data Center)

- Allow to detach a Storage Domain from a Data Center
 - The VM definitions will be kept in the database (but not visible in their tabs) and on a special volume on the Storage Domain
 - The Storage Domain will be still visible in the Storage tab
 - The VMs can be inspected in the Storage subtabs
- Allow to attach a Storage Domain to a Data Center
 - The VM will be available to be imported in the new Clusters
- Import a Storage Domain (Disaster Recovery)
 - The VM definitions will be imported from the special volume located on the storage domain

Import Existing Data Domains

... Storage ...

New Domain ...

Name	...
Data1	
Data2	
Data3	

... VM Import Template Import ...

Import

Name	Memory	CPUs	Architecture	...	Cluster
VM1	1GB	2	x86_64		
VM2	2GB	1	PPC		
VM3					

Import Existing Data Domains

Import Virtual Machine(s) [-] [] [X]

Default Cluster: ▼

	Name	Memory	CPUs	Architecture	...	Cluster
	VM1	1GB	2	x86_64		<input type="text" value="Cluster1"/> ▼
	VM2	2GB	1	PPC		<input type="text" value="Cluster2"/> ▼
	VM3					<input type="text" value="Cluster1"/> ▼

Name: VM1
Description:
OS: F20

“Fencing is the process of isolating a node of a computer cluster or protecting shared resources when that node appears to be malfunctioning”

- oVirt Engine fences an host to:
 - Ensure that the host released its resources (VMs)
 - Try to make the host responsive again
- oVirt supports Soft Fencing (ssh) and Hard Fencing (hardware)
 - Soft Fencing depends on the host cooperation and network availability
 - Hard Fencing requires dedicated hardware

- Sanlock Fencing will provide an additional fencing mechanism communicating through the storage
 - Engine will send the fencing request picking a proxy host
 - The request will be forwarded to the target host through the storage
 - Target host will provide an ACK and will stop to *pet the watchdog*

“A watchdog timer (WDT) is an electronic timer that is used to detect and recover from computer malfunctions.

If software hangs or is lost, a WDT resets the system.”

- After the watchdog timeout expired we know that the host has been rebooted

- Sanlock Fencing will leverage the existing fencing API
- Sanlock Fencing requirements:
 - Both proxy and target hosts should run Sanlock with the fencing support (version provided in oVirt 3.5)
 - Target host must reach the storage and be connected to the Storage Pool

THANK YOU !

<http://www.ovirt.org>

<http://lists.ovirt.org/mailman/listinfo>

devel@ovirt.org

users@ovirt.org