

Shared storage management in the virtualization world

Liron Aravot, Senior Software Engineer
Red Hat

March 2016

- oVirt introduction
- oVirt basic storage concepts
- Shared storage management
- Q&A

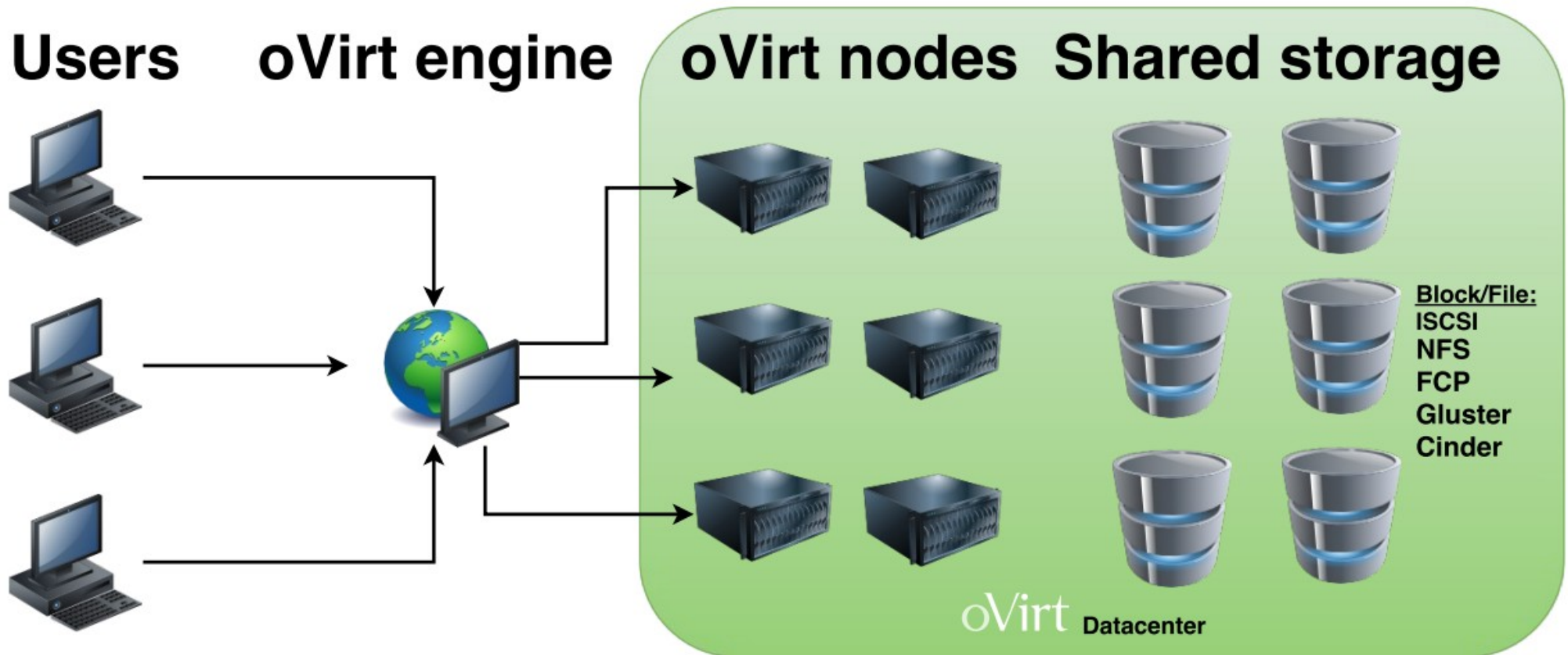
What is oVirt?



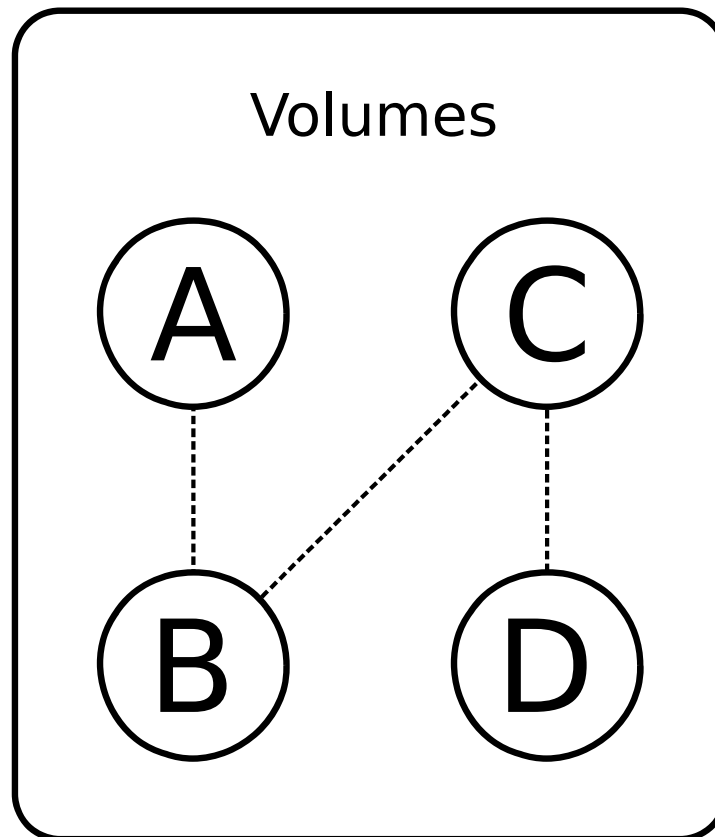
- Virtualization management application
- Manages nodes, storage and network resources
- Deploys, runs and monitors virtual machines

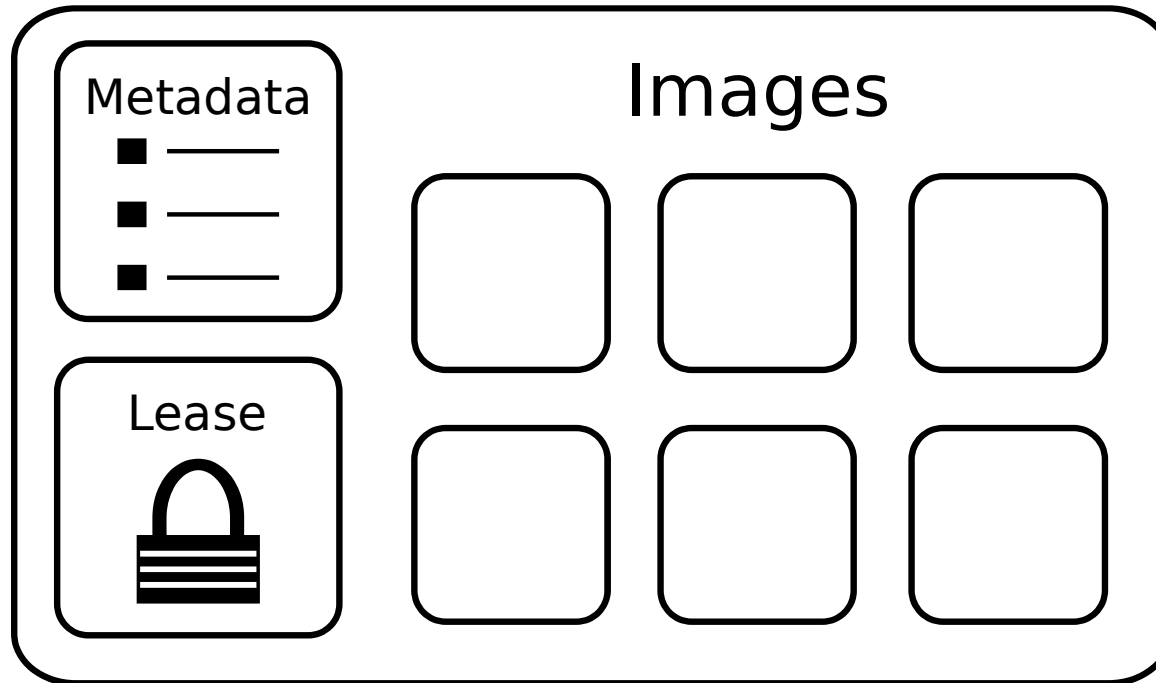
A screenshot of the oVirt Open Virtualization Manager web interface. The interface has a dark header bar with the oVirt logo and 'OPEN VIRTUALIZATION MANAGER' text. On the right of the header, there is a user dropdown menu showing 'admin@internal-authz', and links for 'Configure', 'Guide', 'About', and 'Feedback'. Below the header is a navigation bar with tabs for 'Data Centers', 'Clusters', 'Hosts', 'Networks', 'Storage', 'Disks', 'Virtual Machines', 'Pools', 'Templates', 'Volumes', and 'Users'. The 'Data Centers' tab is selected. On the left side, there is a sidebar with a tree view showing the system hierarchy: 'System' (expanded), 'Data Centers' (expanded), 'Default' (expanded), 'Storage' (expanded), 'Networks' (expanded), 'Templates' (expanded), 'Clusters' (expanded), 'Hosts' (expanded), 'External Providers' (expanded), 'asdasd', 'ovirt-image-repository', 'Errata', and 'Guest Information'. The main content area displays a table of Data Centers. The table has columns: 'Name', 'Comment', 'Storage Type', 'Status', 'Compatibility Version', and 'Description'. There are two rows: 'Default' (green triangle icon, 'Shared', 'Up', '4.0', 'The default Data Center') and 'tsfsfd' (red triangle icon, 'Shared', 'Non Responsive', '3.6'). At the bottom of the interface, there is a status bar with a 'Last Message' section showing a critical alert: 'Feb 29, 2016 6:52:19 PM Critical: Low disk space: sd777 domain has 2 GB of free space'. On the right of the status bar, there are links for 'Alerts (0)', 'Events', and 'Tasks (0)'.

oVirt architecture

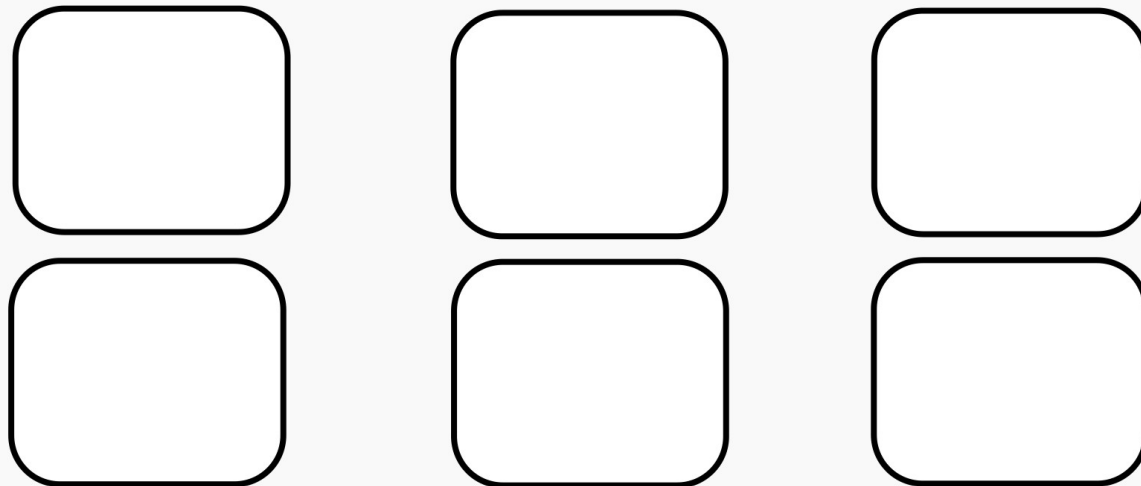


oVirt storage concepts





Storage Domains



oVirt Shared Storage Management

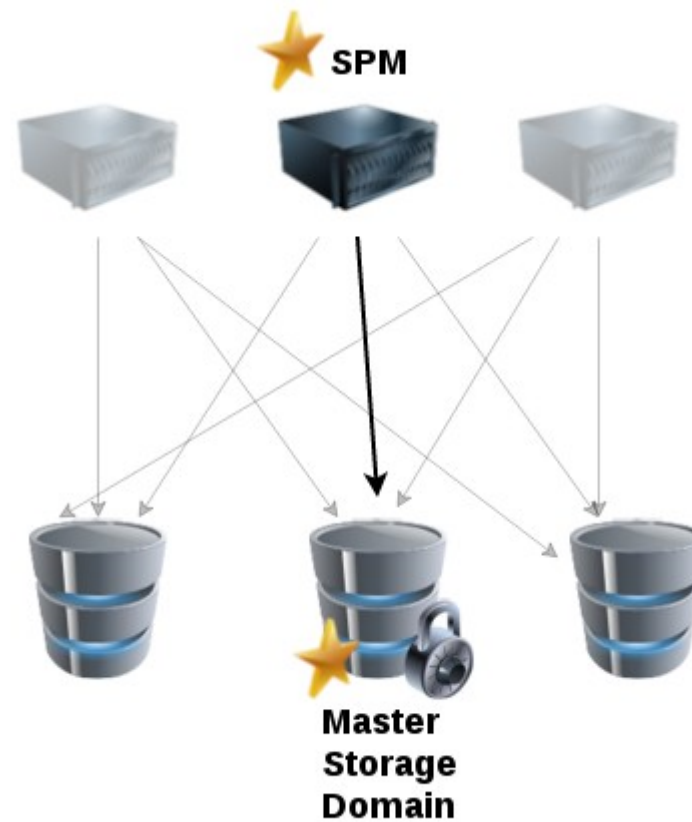
Why do we need management?

- Data protection
- Performance

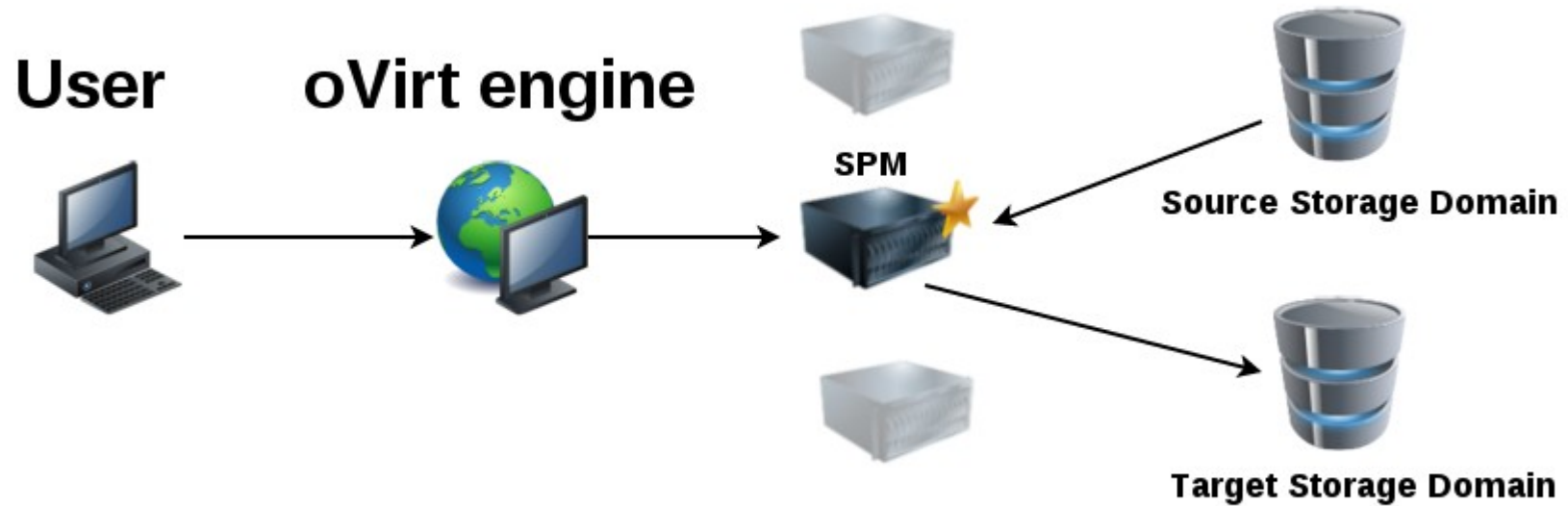


Storage Pool Manager

1 – Storage Pool Manager Architecture



Flow Example – Move Disk

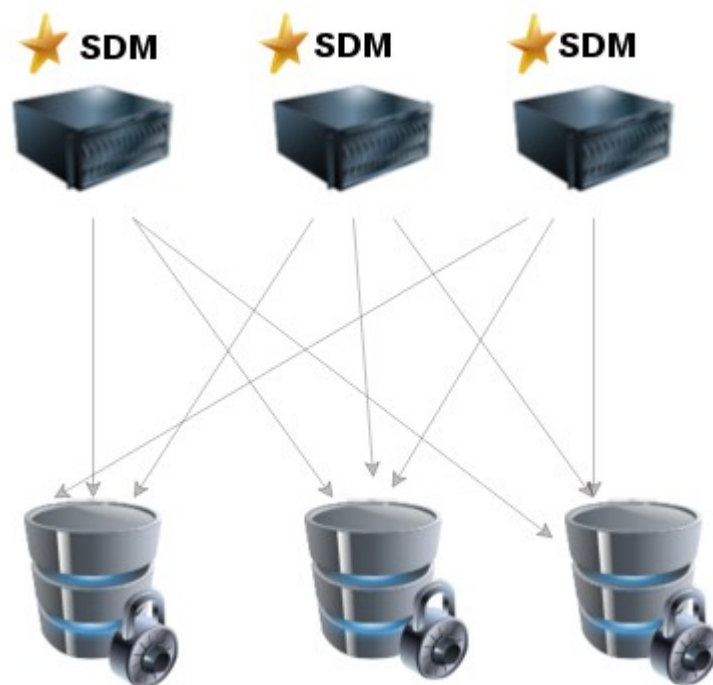


- SPM creates target image (all volumes)
- SPM copies data
- SPM deletes the source image

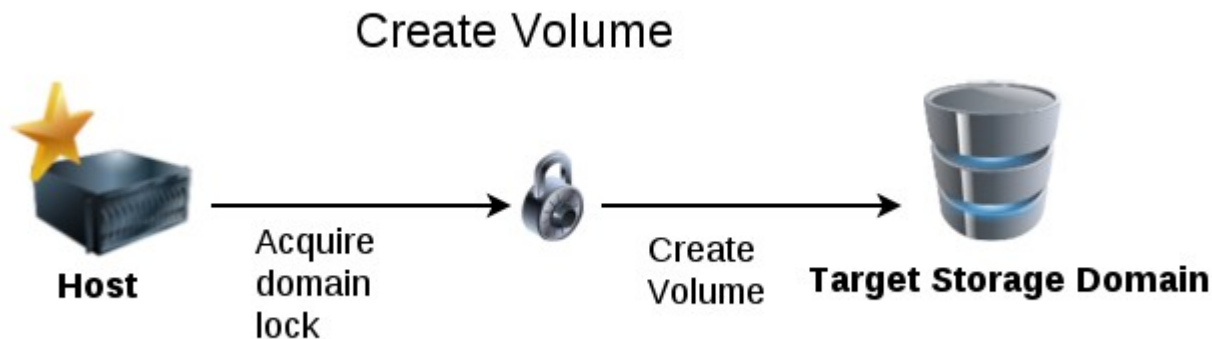
- Performs all shared storage operations
- Holds lock on the master storage domain
- SPOF
- Easy to manage

Storage Domain Management

2 – Storage Domain Manager Architecture



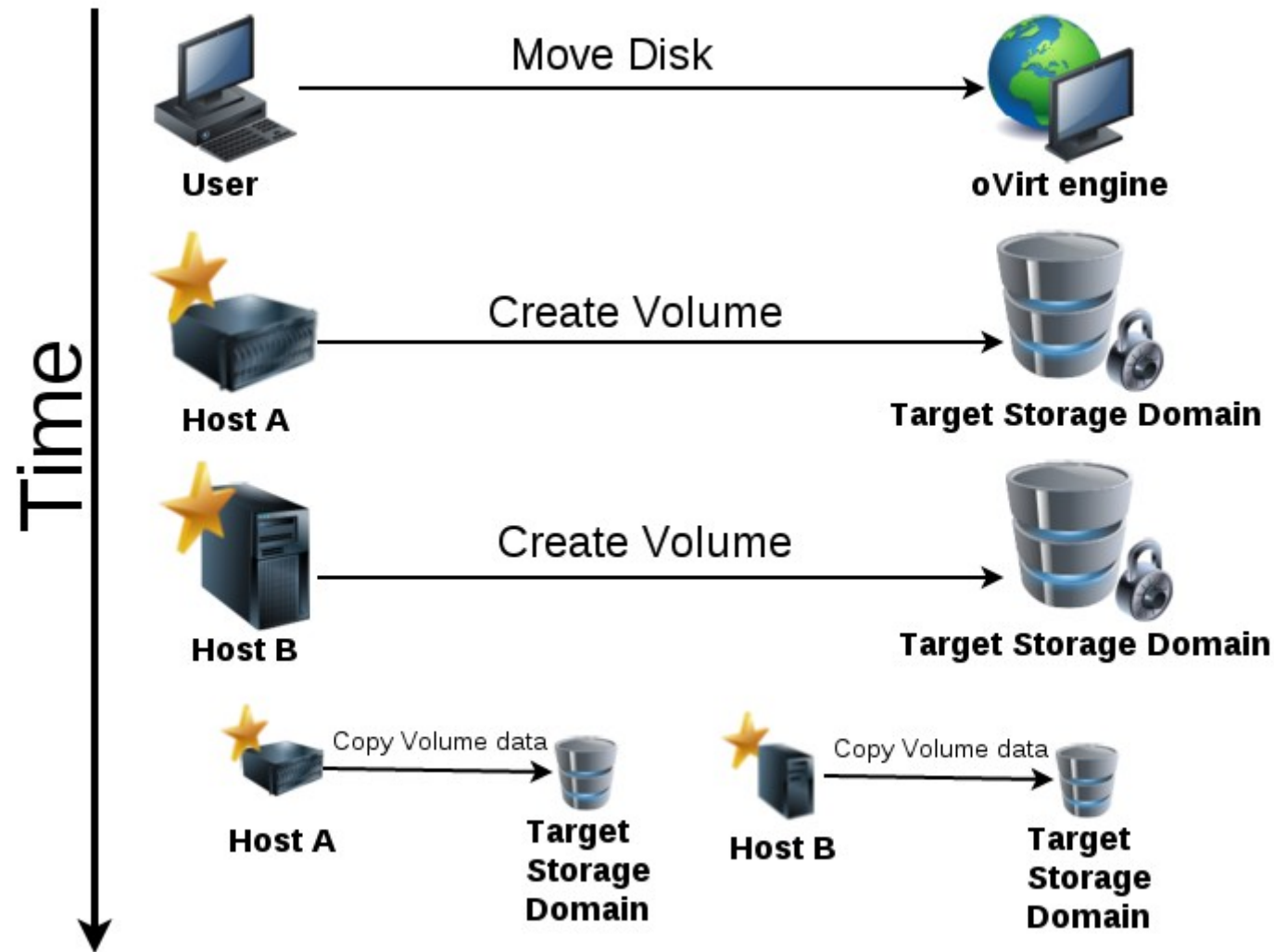
- For consistency and preventing corruption
- Engine level locking
- Storage domain metadata lock (SANLock)
- Volume level lock (SANLock)
- Example: volume creation



- Long operations
- Involve lots of I/O
- Can be executed by any host in the cluster
- Doesn't the Storage Domain metadata lock
- Example: Data copy between volumes

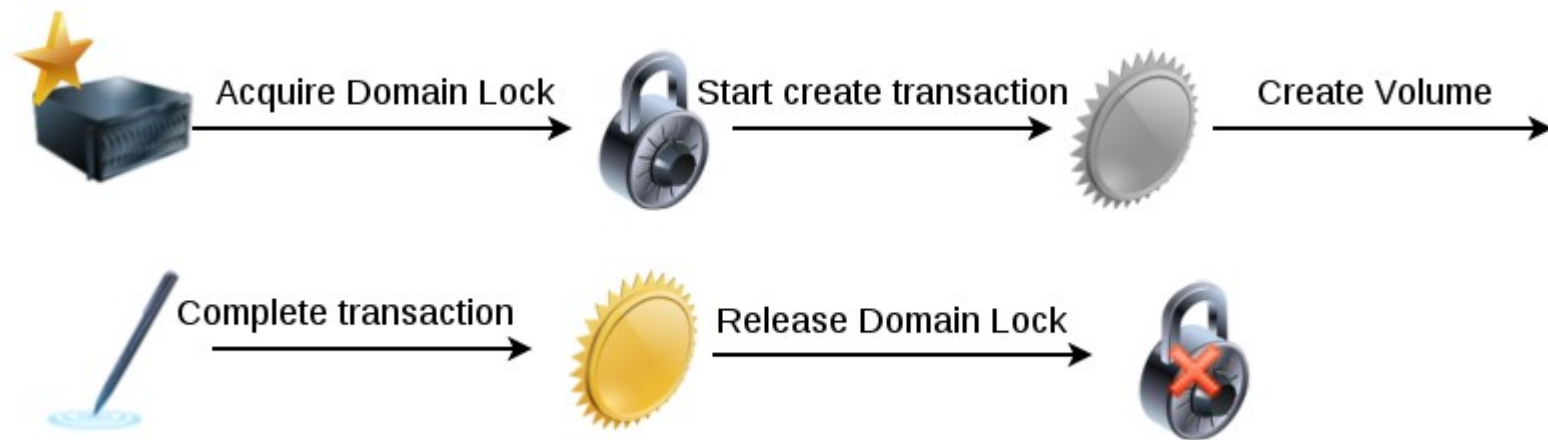
- Short operations
- Creation/Update/Deletion of oVirt storage objects
- Can be executed by any host in the cluster
- Requires the Storage domain metadata lock
- Example: Volume creation

Flow Example – Move Disk



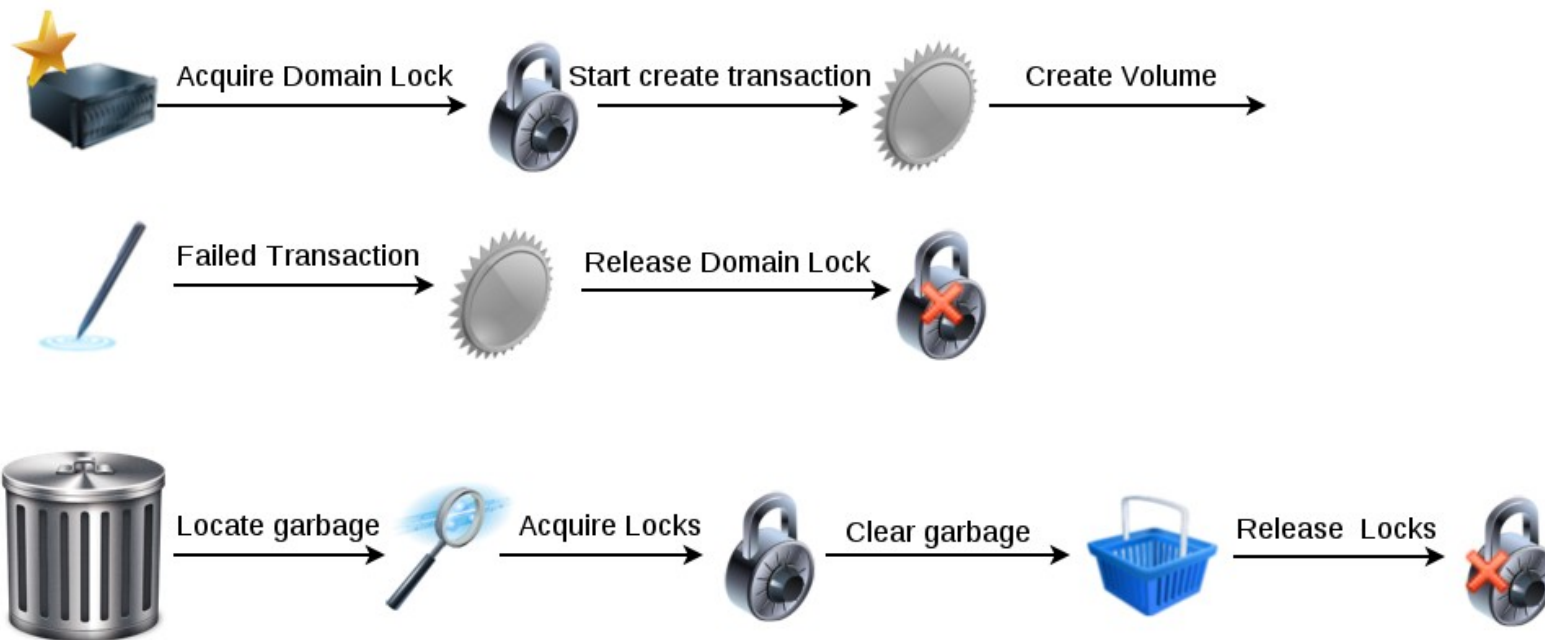
Storage transactions

- Used for metadata operations
- A transaction is opened with a marker operation
- A transaction is committed by converting the marker
- The domain lock is acquired during the transaction



Garbage Collection

- Runs periodically on an arbitrary host
- Identifies candidates by finding markers
- Acquires necessary locks for the candidate
- Cleans garbage associated with the marker



Garbage Collector

- Operations can be performed by any host
- Separation of metadata and data operations
- Load balancing between the hosts
- No SPOF

- What is oVirt?
- oVirt storage concepts
- SPM management
- SDM management

Questions?

THANK YOU !

<http://www.ovirt.org>

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

#ovirt irc.oftc.net

laravot@redhat.com