

OpenNebula - Feature #1796

Improve Ceph Integration

03/06/2013 03:47 PM - Bill Campbell

Status:	Closed	Start date:	03/06/2013
Priority:	High	Due date:	
Assignee:		% Done:	0%
Category:	Drivers - Storage	Estimated time:	0.00 hour
Target version:	Release 4.4	Pull request:	
Resolution:			

Description

I ran into an issue testing the Ceph drivers on a newer version of Libvirt (1.0.x in particular) and they are broken if CephX authentication is enabled (as noted in the documentation). However, it **should** work and function appropriately with CephX authentication, without requiring the ceph.conf and ceph.keyring present on each hypervisor, but we would need the following for each Ceph Datastore configured:

CEPHUUID= A generated UUID for a LibVirt secret (to hold the CephX authentication key in Libvirt on each hypervisor). This should be generated when creating the Ceph datastore in OpenNebula

CEPHUSER= The ceph client.user (this is 'admin' by default, but when configuring a cluster users can be added/removed, so ideally this should be defined)

CEPHKEY= A defined authentication Key for the user (extractable from the Ceph cluster)

CEPHMONS= The list of Monitors for the ceph cluster (either IP or hostnames)

CEPHPORT= The default port is '6789', this could be assumed, but it IS configurable in Ceph, so might need to define

What would need to happen from a deployment standpoint are 3 things:

- Generate a secret.xml file for defining the secret
- Check libvirt on the deployment hypervisor and verify the secret is defined. If not, define it.
- Inject the appropriate cluster information into the deployment.x file

The Secret XML would need to look like this (with the defined items replacing below)

```
<secret ephemeral='no' private='no'>
  <uuid>${CEPHUUID}</uuid>
  <usage type='ceph'>
    <name>anything here really</name>
  </usage>
</secret>
```

The Disk Section of the XML would need to look like this (with the defined items replacing the variables below):

```
<disk type='network' device='disk'>
  <source protocol='rbd' name='${SOURCE}'>
    <host name='${FIRST CEPHMON ENTRY}' port='${CEPHPORT}'/>
    <host name='${SECOND CEPHMON ENTRY}' port='${CEPHPORT}'/>
    <host name='${THIRD CEPHMON ENTRY}' port='${CEPHPORT}'/>
  </source>
  <auth username='${CEPHUSER}'>
    <secret type='ceph' uuid='${CEPHUUID}'>
  </target dev='${DEVICE}' bus='${BUS}'/>
```

These steps could potentially be done with a TM/DS MAD driver

To verify the secret exists, you could run something similar on the following on the hypervisor:

```
virsh secret-list | grep -c $CEPHUUID
```

To add the secret and inject the key, you must do the following on the hypervisor (should be done upon deployment of VM (possibly with 'clone' and 'ln' TM drivers). This should only need to be done once per Cluster/Datastore

```
virsh secret-define --file secret.xml
virsh secret-set-value --secret $CEPHUUID --base64 $CEPHKEY
```

After that it should just be a matter of injecting the appropriate information into the deployment.x file for the virtual machine.

This is NOT anything I expect to be worked on for 4.0, just want to supply some information to better integrate Ceph and future proof VM deployments with newer versions of Libvirt. Ubuntu 12.04 and CentOS 6x still use older versions of Libvirt that are not affected by the authentication changes, so for most deployments the existing implementation should work well.

Associated revisions

Revision 1a2b9238 - 11/15/2013 11:58 AM - Carlos Martín

Feature #1796: Add ceph host and secret to kvm deployment files

Revision c76a2cbb - 11/22/2013 10:53 PM - Ruben S. Montero

feature #1796: Add CEPH_USER to CEPH Datastore configuration attributes.

Revision 0203b7a3 - 11/25/2013 03:35 PM - Ruben S. Montero

feature #1796: Add CEPH_USER to CEPH Datastore configuration attributes.

(cherry picked from commit c76a2cbb8e9e8dd1d26ee9fa909de51d8cc270d7)

History

#1 - 03/06/2013 07:06 PM - Bill Campbell

Not sure how I messed that up, but the correct disk entry is below:

```
<disk type='network' device='disk'>
  <source protocol='rbd' name='$SOURCE'>
    <host name='$FIRST CEPHMON ENTRY' port='$CEPHPORT'/>
    <host name='$SECOND CEPHMON ENTRY' port='$CEPHPORT'/>
    <host name='$THIRD CEPHMON ENTRY' port='$CEPHPORT'/>
  </source>
  <auth username='$CEPHUSER'>
    <secret type='ceph' uuid='$CEPHUUID'/>
  </auth>
  <target dev='$DEVICE' bus='$BUS'/>
```

#2 - 03/15/2013 10:36 PM - Ruben S. Montero

- *Tracker changed from Request to Feature*
- *Assignee set to Jaime Melis*
- *Target version set to Release 4.0*

#3 - 03/16/2013 04:39 PM - Bill Campbell

- *File opennebula4.0-rbd-format2-transfer-datastore-drivers.tar.gz added*

I added these to the original request which is closed, but these are updated drivers that support format 2 images, which takes advantage of the copy-on-write clones for non-persistent images. Minimizes copy operations and allows for rapid deployment of non-persistent instances.

#4 - 03/19/2013 09:20 PM - Ruben S. Montero

- *Target version changed from Release 4.0 to Release 4.2*

OK thanks Bill for the update. We agree that for 4.0 the current configuration should be enough (i.e. not using cephx or the keyring must be placed in the expected path).

What we need to decide is where to place the configuration options, e.g. where to place the secret file... Moving this to 4.2

Thanks again for your contributions!!!

#5 - 05/09/2013 02:44 PM - Ruben S. Montero

- *Category changed from Drivers - Auth to Drivers - Storage*

#6 - 05/13/2013 09:45 PM - Ruben S. Montero

- *Tracker changed from Feature to Backlog*
- *Status changed from New to Pending*
- *Assignee deleted (Jaime Melis)*
- *Priority changed from Normal to High*
- *Target version deleted (Release 4.2)*

#7 - 07/24/2013 10:29 AM - Matthew Richardson

Hi - just wondering what the status of this report is?

Currently it appears impossible to use ceph with libvirt > 1.0 and opennebula, unless cephx is disabled, which is not recommended by the ceph project.

Is this still on track for inclusion in 4.2 (I notice the target has been deleted).

Is there anything that I can do to help progress this issue?

Thanks!

#8 - 07/24/2013 05:59 PM - Jaime Melis

This feature has been postponed to the next release OpenNebula 4.4. The rationale behind this decision is as follows:

- OpenNebula 4.2 is mostly about services, elasticity and vmware drivers, and although improving the Ceph integration would have been ideal, other things were prioritised instead.
 - Disabling cephx, even though not recommended by default by the Ceph team, is a very sensible thing to do in the context of an OpenNebula cluster. OpenNebula assumes it's running in a safe network and with exclusive access to the nodes, thus no authentication is required for Ceph, NFS or other storage solutions.
 - Libvirt 1.x is only available in 2 of the OpenNebula supported platforms: Ubuntu 13.04 and OpenSUSE 12.3 (in Ubuntu 12.04, CentOS and Debian it does work with cephx).
 - Ceph continues to work perfectly without cephx, and there are other mechanisms to provide additional security, such as iptables
- Of course this doesn't mean that this issue won't be resolved, and that we don't consider this to be a top priority issue. OpenNebula 4.4, which will be here before long (~ 3 or 4 months) will feature this and many other improvements.
- OpenNebula 4.4 will be here very shortly and it will include this and many other improvements

#9 - 11/13/2013 12:06 PM - Ruben S. Montero

- *Target version set to Release 4.4*

#10 - 11/28/2013 04:10 PM - Jaime Melis

- *Status changed from Pending to Closed*

#11 - 10/02/2014 11:49 AM - Ruben S. Montero

- *Tracker changed from Backlog to Feature*

Files

opennebula4.0-rbd-format2-transfer-datastore-drivers.tar.gz	3.86 KB	03/16/2013	Bill Campbell
---	---------	------------	---------------