

OpenNebula - Bug #5032

Datastores with TARGET = SELF (e.g. Ceph) needs to properly account image DS usage (as done in disk-resize operations for persistent images)

02/18/2017 03:18 AM - Vy Nguyen Tan

Status:	Closed	Start date:	02/18/2017
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	Drivers - Storage	Estimated time:	0.00 hour
Target version:	Release 5.4	Pull request:	
Resolution:	fixed		
Affected Versions:	OpenNebula 5.2		

Description

Hello,

I'm using OpenNebula 5.2.1, CentOS 7.3, Ceph datastore. I found a bug that user can create a VM with overdisk quota (please check video).

Associated revisions

Revision d90cd64c - 03/30/2017 04:58 PM - Ruben S. Montero

B #5032: Add datastore capacity usage in quota calculations for storage drivers that clone to SELF (e.g. Ceph)

Revision d1ad6a0c - 03/31/2017 06:09 PM - Ruben S. Montero

B #5032: Further fixes for SELF DS (e.g. Ceph) for disks with resizes and snapshots. Also updates delete-recreate quota computation

History

#1 - 03/22/2017 10:47 AM - Arnaud Abélard

Same problem here.

Although my users have a quota affected to the ceph system datastore, they can create system disk larger than their quota.

The problem doesn't occur with volatile disks which are properly accounted in the user's quota.

When I create a 2TB system disk from a ceph based image, no system disk is actually affected to my user:

```
~# onevm show 409
VIRTUAL MACHINE 409 INFORMATION
ID           : 409
NAME         : testaa3
USER         : abelard-a
...
VM DISKS
ID DATASTORE TARGET IMAGE          SIZE  TYPE SAVE
0 ceph-image vda  debian8.4-univnantes-v7  /2T  rbd  NO
```

~# oneuser show abelard-a

NUMBER OF VMS	MEMORY	CPU	SYSTEM_DISK_SIZE
1 / 5	1024M / 9.8G	1.00 / 8.00	0M / 1.9T

DATASTORE ID	IMAGES	SIZE
100	1 / -	2.2G / -

I'm puzzled by the fact the datastore 100 is the ceph image pool and there's no mention of the datastore 101 which is the ceph system datastore:

~# onedatastore show 100

DATASTORE 100 INFORMATION

ID : 100
NAME : ceph-image
USER : oneadmin
GROUP : oneadmin
CLUSTERS : 0
TYPE : IMAGE
DS_MAD : ceph
TM_MAD : ceph
BASE PATH : /var/lib/one//datastores/100
DISK_TYPE : RBD
STATE : READY

DATASTORE CAPACITY

TOTAL: : 57.5T
FREE: : 56.8T
USED: : 761.3G
LIMIT: :-

PERMISSIONS

OWNER : um-
GROUP : u--
OTHER : u--

DATASTORE TEMPLATE

BRIDGE_LIST=""
CEPH_HOST=""
CEPH_SECRET=""
CEPH_USER="opennebula"
CLONE_TARGET="SELF"
DATASTORE_CAPACITY_CHECK="YES"
DISK_TYPE="RBD"
DS_MAD="ceph"
LN_TARGET="NONE"
POOL_NAME="opennebula"
TM_MAD="ceph"
TYPE="IMAGE_DS"

IMAGES

17
18

19
42
46
49
58
59
62
63
65
67
69
71

~# onedastore show 101

DATASTORE 101 INFORMATION

ID : 101
NAME : ceph-system
USER : oneadmin
GROUP : oneadmin
CLUSTERS : 0
TYPE : SYSTEM
DS_MAD : -
TM_MAD : ceph
BASE PATH : /var/lib/one//datastores/101
DISK_TYPE : RBD
STATE : READY

DATASTORE CAPACITY

TOTAL: : 57.5T
FREE: : 56.8T
USED: : 760.5G
LIMIT: :-

PERMISSIONS

OWNER : uma
GROUP : u--
OTHER : ---

DATASTORE TEMPLATE

BRIDGE_LIST=""
CEPH_HOST=""
CEPH_SECRET=""
CEPH_USER="opennebula"
DATASTORE_CAPACITY_CHECK="YES"
DISK_TYPE="RBD"
DS_MIGRATE="NO"
POOL_NAME="opennebula"
RESTRICTED_DIRS=""
SAFE_DIRS="/var/tmp"
SHARED="YES"
TM_MAD="ceph"
TYPE="SYSTEM_DS"

IMAGES

#2 - 03/22/2017 05:59 PM - Ruben S. Montero

- Category set to Drivers - Storage
- Target version set to Release 5.4

#3 - 03/23/2017 10:55 AM - Ruben S. Montero

- Status changed from Pending to Closed
- Resolution set to worksforme

OK, I've the opportunity to look at the video. Note that System Disks is for the disks created in the system datastore, in ceph disks are created using the same ceph pool as the images, so only volatile disks use storage from the system datastore. To limit the size, you need to add a quota to the Datastore.

#4 - 03/23/2017 11:45 AM - Arnaud Abélard

actually I do have a quota on all my datastores:

```
~# oneuser show abelard-a
```

RESOURCE USAGE & QUOTAS

NUMBER OF VMS	MEMORY	CPU	SYSTEM_DISK_SIZE
2 / 5	2G / 9.8G	2.00 / 8.00	0M / 1.9T

DATASTORE ID	IMAGES	SIZE
0	0 / 0	0M / 1.9T
1	0 / -	0M / 1.9T
2	0 / -	0M / 1.9T
101	0 / -	0M / 1.9T
100	0 / -	0M / 1.9T

But as you can notice, none of my VM actual consume anything

even though I have a 2TB VM running:

```
~# onevm show 409
```

VIRTUAL MACHINE 409 INFORMATION

```
ID          : 409
NAME        : testaa3
USER        : abelard-a
```

...

VM DISKS

ID	DATASTORE	TARGET IMAGE	SIZE	TYPE	SAVE
0	ceph-image	vda	debian8.4-univnantes-v7	-/2T	rbd NO
1	-	hda	CONTEXT	-/-	-

Notice that the used size is not available...

Anyway, I do use the 2TB so the volume should take some place in my quota,

```
root@testaa3:~# df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/vda1       2,0T  2,0T   0 100% /
```

ceph-image being the datastore 100:

```
~# onedatastore show 100
DATASTORE 100 INFORMATION
ID        : 100
NAME      : ceph-image
USER      : oneadmin
GROUP     : oneadmin
CLUSTERS  : 0
TYPE      : IMAGE
DS_MAD    : ceph
TM_MAD    : ceph
BASE PATH : /var/lib/one//datastores/100
DISK_TYPE : RBD
STATE     : READY
...
IMAGES
17
18
19
42
46
49
59
62
63
65
67
69
71
79
83
```

my VM is using the image with id 42 on that datastore:

```
root@one-ctrl-1:~# oneimage show 42
IMAGE 42 INFORMATION
ID        : 42
NAME      : debian8.4-univnantes-v7
USER      : oneadmin
GROUP     : oneadmin
DATASTORE : ceph-image
TYPE      : OS
REGISTER TIME : 11/10 16:28:31
```

PERSISTENT : No
SOURCE : opennebula/one-42
PATH : /var/tmp/debian8-5.0.1-univnantes-v7.qcow2
FSTYPE : raw
SIZE : 2G
STATE : used
RUNNING_VMS : 22

PERMISSIONS
OWNER : um-
GROUP : u--
OTHER : u--

IMAGE TEMPLATE
DEV_PREFIX="vd"
DRIVER="raw"

VIRTUAL MACHINES

ID	USER	GROUP	NAME	STAT	UCPU	UMEM	HOST	TIME
...								
409	abelard-	DSI-IRTS	testaa3	runn	0.0	1.1G	iaas-vm-3.	1d 01h34
...								

As you can see, I have 2TB of data on my ceph-image datastore that aren't being accounted at all. I might have missed something, but I really wonder why since volatile disk are properly accounted.

Thanks

#5 - 03/23/2017 03:26 PM - Ruben S. Montero

- Status changed from Closed to Pending

Let me double check this, I remember we fix something related with quotas for TM drivers using SELF (like ceph). So maybe you are being hit by that issue.

#6 - 03/27/2017 04:33 PM - Ruben S. Montero

- Subject changed from *Can't limit system disk with Ceph datastore* to *Datastores with TARGET = SELF (e.g. Ceph) needs to properly account image DS usage (as done in disk-resize operations for persistent images)*

- Status changed from Pending to New

- Resolution deleted (worksforme)

TL;DR There is a bug that only affects VM creation with a resize. Ceph (an other clone to SELF DS) control DS usage through the VM (max. number of virtual machines). As the scheduler is already using the right check I'm reopening the issue to properly account this datastores.

Sorry, my bad. I was confused about the bug I mentioned in my previous comment.

The idea is the following:

- The space used by VMs in the image datastore (as in the case of Ceph) is controlled by the VMs quota number. So you can "indirectly" limit the space consumed by the users in the image datastore (by creating virtual machines) with "NUMBER OF VMS".

- The space used by Images is controlled by the SIZE quota of the DATASTORE.

However I found a bug, in the resize operation. In allocation for Ceph (any target SELF datastore, for the matter of fact) and non-persistent images:

- You can create VMs and request any size for the disk (with non-persistent images), effectively bypassing the VM quota. (Note that persistent images cannot be resized)

Note that this only occurs when allocating a VM, the resize operation is properly controlled.

As a side note, the usage of system_ds is aggregated in SYSTEM_DS_USAGE, that's the reason you do not see it under the Datastore list (which is intended for images).

#7 - 03/31/2017 06:30 PM - Ruben S. Montero

This is now fixed in master, a migrator for next release is still needed to recompute the quotas for TARGET=SELF Datastores

#8 - 07/05/2017 09:57 AM - Ruben S. Montero

- *Status changed from New to Closed*

- *Resolution set to fixed*

Files

opennebula-can-limit-disk-quota-with-ceph.mov	6.74 MB	02/18/2017	Vy Nguyen Tan
---	---------	------------	---------------