Subject: vzdump snapshot problem and workaround for it
Posted by oleka on Fri, 06 Apr 2007 20:09:29 GMT
View Forum Message <> Reply to Message

Hi everyone,

Just would like to share with you my experience using vzdump with snapshot backups.

My setup:
OS: Fedora Core5
vzdump version: 0.4-1 on FC5, vzdump-0.4-1.noarch.rpm

First of all, I had to update lvm rpm package on FC5 up to lvm2-2.02.17-1.fc5 (yum -y update lvm2).

Doing backup with the following command:

vzdump --compress --dumpdir /backup/101 --snapshot 101

has given me this output:

```
starting backup for VPS 101 (/vz/private/101) Fri Apr  6 15:23:41 2007
creating lvm snapshot of /dev/mapper/VolGroup00-VZ ('/dev/VolGroup00/vzsnap')
  Rounding up size to full physical extent 512.00 MB
  Logical volume "vzsnap" created
mounting lvm snapshot
Use of uninitialized value in concatenation (.) or string at /usr/bin/vzdump line 463.
Use of uninitialized value in concatenation (.) or string at /usr/bin/vzdump line 465.
  Logical volume "vzsnap" successfully removed
creating backup for  VPS 101 failed (0.02 minutes): wrong lvm mount point '' at /usr/bin/vzdump
line 465.
```

Looking into vzdump perl scrip, I could not understand the logic those lines: 463 and 465 (variable $lvmpath cannot get any value by script design).

Commenting them out has fixed the problem. And output I've gotten after the "fix":

```
starting backup for VPS 101 (/vz/private/101) Fri Apr  6 15:56:47 2007
creating lvm snapshot of /dev/mapper/VolGroup00-VZ ('/dev/VolGroup00/vzsnap')
  Rounding up size to full physical extent 512.00 MB
  Logical volume "vzsnap" created
mounting lvm snapshot
Creating archive '/backup/101/vzdump-101.tgz' (/vz/private/101)
tar: ./var/run/dbus/system_bus_socket: socket ignored
tar: ./var/lib/mysql/mysql.sock: socket ignored
tar: ./dev/log: socket ignored
Total bytes written: 365987840 (350MiB, 7.5MiB/s)
  Logical volume "vzsnap" successfully removed
```

backup for VPS 101 finished successful (0.78 minutes)

Hope, this little trick will help.

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Wed, 11 Apr 2007 08:27:33 GMT
View Forum Message <> Reply to Message

Thanks for the post!

I am a little confused by the snapshot backup stuff. Is it creating an LVM snapshot that can later be mounted by a VPS, without first restoring? Or is the idea just to create a temporary snapshot to avoid downtime, then backup the snapshot?

Subject: Re: vzdump snapshot problem and workaround for it
Posted by oleka on Thu, 12 Apr 2007 02:29:56 GMT
View Forum Message <> Reply to Message

jarcher wrote on Wed, 11 April 2007 04:27I am a little confused by the snapshot backup stuff. Is it creating an LVM snapshot that can later be mounted by a VPS, without first restoring?
VPS has nothing to do with LVM snapshot. Snapshot is done on HN (hardware node). vzdump creates a tar/tgz file which later can be used to restore the VPS with an original VEID or a different VEID (aka VPS cloning).

Quote: Or is the idea just to create a temporary snapshot to avoid downtime, then backup the snapshot?

Snapshot is a part of LVM functionality and vzdump just uses it to do backup without VPS downtime. The script just does the procedure well explained in this article:
http://tldp.org/HOWTO/LVM-HOWTO/snapshots_backup.html

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Thu, 12 Apr 2007 05:26:12 GMT
View Forum Message <> Reply to Message

Okay, thanks for explaining. When I try this, it fails. Apparently there is some issue with volume group creation. I am running on Debian Etch 64 with the latest OpenVZ kernel (Apr 10).

Here is the output:

```
actual:/home/jim# vzdump --dumpdir /var/local/vps-bu --snapshot --compress 108
starting backup for VPS 108 (/mnt/compvps/vz/private/108) Thu Apr 12 01:23:17 2007
Use of uninitialized value in concatenation (.) or string at /usr/bin/vzdump line 455.
creating lvm snapshot of /dev/mapper/actual--vg001-vps--storage ('/dev//vzsnap')
Use of uninitialized value in concatenation (.) or string at /usr/bin/vzdump line 457.
Use of uninitialized value in concatenation (.) or string at /usr/bin/vzdump line 457.
  "/dev//": Invalid path for Logical Volume
  The origin name should include the volume group.
  lvcreate: Create a logical volume

lvcreate
        [-A|--autobackup {y|n}]
        [--addtag Tag]
        [--alloc AllocationPolicy]
        [-C|--contiguous {y|n}]
        [-d|--debug]
        [-h|-?|--help]
        [-i|--stripes Stripes [-I|--stripesize StripeSize]]
        {-l|--extents LogicalExtentsNumber |
         -L|--size LogicalVolumeSize[kKmMgGtT]}
        [-M|--persistent {y|n}] [--major major] [--minor minor]
        [-m|--mirrors Mirrors [--nosync] [--corelog]]
        [-n|--name LogicalVolumeName]
        [-p|--permission {r|rw}]
        [-r|--readahead ReadAheadSectors]
        [-R|--regionsize MirrorLogRegionSize]
        [-t|--test]
        [--type VolumeType]
        [-v|--verbose]
        [-Z|--zero {y|n}]
        [--version]
        VolumeGroupName [PhysicalVolumePath...]

lvcreate -s|--snapshot
        [-c|--chunksize]
        [-A|--autobackup {y|n}]
        [--addtag Tag]
        [--alloc AllocationPolicy]
        [-C|--contiguous {y|n}]
        [-d|--debug]
        [-h|-?|--help]
        [-i|--stripes Stripes [-I|--stripesize StripeSize]]
        {-l|--extents LogicalExtentsNumber |
         -L|--size LogicalVolumeSize[kKmMgGtT]}
        [-M|--persistent {y|n}] [--major major] [--minor minor]
        [-n|--name LogicalVolumeName]
        [-p|--permission {r|rw}]
        [-r|--readahead ReadAheadSectors]
```

[-t|--test]
[-v|--verbose]
[--version]
OriginalLogicalVolume[Path] [PhysicalVolumePath...]

umount: /vzsnap: not mounted
  Volume group "vzsnap" not found
creating backup for VPS 108 failed (0.00 minutes): command '/sbin/lvcreate --size 500m --snapshot --name vzsnap /dev//' failed with exit code 3 at /usr/bin/vzdump line 107.

actual:/home/jim#

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Thu, 12 Apr 2007 06:23:43 GMT
View Forum Message <> Reply to Message

I did some more checking, and have this to add:

I think I found the trouble, which is in the get_device function. Now, just as a warning, my PERL is really, really rusty. I did some many years ago and quickly decided I'll stick to C/C++. But I'll do my best to provide helpful information to the developer of vzdump.

My VPS private areas live on a LVM2 LV, which is ultimately on an iSCSI SAN.

I run this:

actual:/home/jim# vzdump --dumpdir /var/local/vps-bu --snapshot 108

As soon as get_device is entered, $dir is set to:

/mnt/compvps/vz/private/108

So running df against the private area for the VPS results in:

Filesystem   1K-blocks   Used Available Use% Mounted on
/dev/mapper/actual--vg001-vps--storage
                     100756920   6463816 89174948   7% /mnt/compvps

So of course $out becomes:

/dev/mapper/actual--vg001-vps--storage

Parsing $out with the specified regex results in:

res[0]: /dev/mapper/actual--vg001-vps--storage
res[5]:

At this point, I don't follow this:

($vg, $lv) = @{$devmapper->{$dev}} if defined $devmapper->{$dev};

I see $devmapper is defined as a global but I can't find $dev, so I think what is happening is that since $devmapper->$dev is undefined, $vg and $lv are never being set, hence the errors I see reported.

Also, I added a few print statements and I see that the get_device function is being called 3 times, at least on my system.

I hope this helps someone.

```
sub get_device {
    my $dir = shift;

    open (TMP, "df '$dir'|");
    <TMP>; #skip first line
    my $out = <TMP>;
    close (TMP);

    my @res = split (/\s+/, $out);

    my $dev = $res[0];
    my $mp = $res[5];
    my ($vg, $lv);

    ($vg, $lv) = @{$devmapper->{$dev}} if defined $devmapper->{$dev};

    return wantarray ? ($dev, $mp, $vg, $lv) : $dev;
}
```

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Thu, 12 Apr 2007 08:02:55 GMT
View Forum Message <> Reply to Message

After thinking about it more, I think one of the issues here is the way the output of df is being parsed. When I run df, the output contains a newline. The regex used to parse splits on spaces. Here is the output from df on my system:

actual:/var/local/vps-bu# df /mnt/compvps/vz/private/108/

Filesystem 1K-blocks Used Available Use% Mounted on
/dev/mapper/actual--vg001-vps--storage
100756920 6411820 89226944 7% /mnt/compvps

The output of lvscan may also be helpful:

actual:/var/local/vps-bu# lvscan
ACTIVE '/dev/actual-vg001/vps-storage' [97.62 GB] inherit
ACTIVE '/dev/VG-RAID/LV1' [144.35 GB] inherit

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Thu, 12 Apr 2007 08:48:58 GMT
View Forum Message <> Reply to Message

If you call df with the -P option it fixes the wraping problem:

actual:/# df -P /mnt/compvps/vz/private/108/
Filesystem 1024-blocks Used Available Capacity Mounted on
/dev/mapper/actual--vg001-vps--storage 100756920 6416292 89222472 7% /mnt/compvps
actual:/#

That fixes the parsing of result from df.

But there is still an issue where it is misreading the name of the LV it needs to snapshot. df is not reporting the name of the LV correctly. Then again, I'm not sure if it is supposed to.

The VG name is: actual-vg001
The LV name is: vps-storage

So it seems df is adding an extra hyphen to any hyphen embedded in a name, and then delimiting the VG and LV with a hyphen, so VG-LG. I don't know where mapper comes from.

Subject: Re: vzdump snapshot problem and workaround for it
Posted by oleka on Thu, 12 Apr 2007 17:30:19 GMT
View Forum Message <> Reply to Message

Hi,

Looking at your original error output:
Quote:
umount: /vzsnap: not mounted
Volume group "vzsnap" not found
creating backup for VPS 108 failed (0.00 minutes): command '/sbin/lvcreate --size 500m
--snapshot --name vzsnap /dev//' failed with exit code 3 at /usr/bin/vzdump line 107

two thoughts:
- outdated perl - I run with v.5.8.8 (could explain /dev//)
- you do not have enough Free PEs in volume group. Run vgdisplay and check for "Free PE / Size". You should have at least 512Mb available. If you have less, you have to repartition the server. On my server I have:

```
# vgdisplay
  --- Volume group ---
  VG Name               VolGroup00
  System ID
  Format                lvm2
  Metadata Areas        1
  Metadata Sequence No  25
  VG Access             read/write
  VG Status             resizable
  MAX LV                0
  Cur LV                3
  Open LV               3
  Max PV                0
  Cur PV                1
  Act PV                1
  VG Size               465.62 GB
  PE Size               32.00 MB
  Total PE              14900
  Alloc PE / Size       13562 / 423.81 GB
  Free  PE / Size       1338 / 41.81 GB
```

And cannot tell about Debian specifics. I stick to FC so far

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Sat, 14 Apr 2007 20:14:45 GMT
View Forum Message <> Reply to Message

Thanks very much!

I'll set the /dev// issue aside for right now and write about the PE allocation in the volume group. As you thought, my VG has no more capacity:

```
prod03:~# vgdisplay
  --- Volume group ---
  VG Name               prod03-vg001
  System ID
  Format                lvm2
  Metadata Areas        2
  Metadata Sequence No  4
  VG Access             read/write
  VG Status             resizable
  MAX LV                0
  Cur LV                1
  Open LV               1
  Max PV                0
  Cur PV                2
  Act PV                2
  VG Size               195.24 GB
  PE Size               4.00 MB
  Total PE              49982
  Alloc PE / Size       49982 / 195.24 GB
  Free  PE / Size       0 / 0
  VG UUID               r9QqPU-rRQr-jIzr-U0Pa-bBle-2DYJ-fHcifB
```

If I understand correctly from investigating this, I can run:

pvchange –s 8m

And this will increase the size of each extent, adding capacity, but not actual storage space, to the VG. To get more storage space I would actually have to add an additional PV to the VG, correct?

There does not seem to be a command to increase the number of extents, just the size of each extent. Is that correct?

If I don't add another PV, will I be able to create another LV in the VG? If so, where would the data live?

Subject: Re: vzdump snapshot problem and workaround for it
Posted by oleka on Tue, 17 Apr 2007 15:52:11 GMT
View Forum Message <> Reply to Message

Hi there,

Quote:
If I understand correctly from investigating this, I can run:

pvchange –s 8m

And this will increase the size of each extent, adding capacity, but not actual storage space, to the VG. To get more storage space I would actually have to add an additional PV to the VG, correct?

There does not seem to be a command to increase the number of extents, just the size of each extent. Is that correct?

If I don't add another PV, will I be able to create another LV in the VG? If so, where would the data live?

Just have to mention to you that fiddling with logical volumes (e.g. changing partition size) can be very dangerous and you have to be absolutely sure what you're doing (there are tons of manuals/how-tos in the Inet). In your case I highly recommend to repartition your HN. More specific, create a separate partition for VZ VPSs and do not allocate all of your disk space (leave say several Gigs unoccupied). This will give you the needed Free PEs.

Subject: Re: vzdump snapshot problem and workaround for it
Posted by jarcher on Wed, 18 Apr 2007 06:37:28 GMT
View Forum Message <> Reply to Message

Thanks for the suggestion. Actually, my OpenVZ install lives on a partition but the VPSs all live on an iSCSI attached SAN. I am using OpenFiler, which itself is based upon LVM. So I have the extra little confusion that a LV on the SAN looks like a physical device on the initiator (client).

I have learned to expand my VGs by adding PVs to them and how to expand the file system while it is mounted. I also learned how to shrink a file system, which was big fun.

I have not tried it on my OpenVZ machine yet, as I am waiting for a second machine to arrive for redundancy. Once I have the ability to migrate my VPSs around I'll have more flexability.

Thanks for your help. I'll update the thread if I get a resolution.
