Subject: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 01:31:20 GMT
View Forum Message <> Reply to Message

Howdy,

I've been tweaking out a cpanel template that we're using, and one of the most important things for cPanel is the ability to secure the /tmp and /var/tmp directories. Typicall running /scripts/securetmp on a server would do that. However, inside the VE there's no loopback support for mount -o loop so the /tmp directories cannot be secured properly.

Anyone have any ideas on how we can do this?

Thanks!

Subject: Re: /dev/loop support inside VEs?
Posted by kir on Tue, 07 Feb 2006 07:45:40 GMT
View Forum Message <> Reply to Message

Indeed there is no loop device support in VPS. This is because each loop device requires a dedicated kernel thread, and the number of loop devices is severily limited.

Still, there are several solutions for your problem exist. First, you can try to use 2.6.15 kernel which has advanced bind mount options, so you can try 'mount -obind,nosuid,noexec,nodev,rw /usr/tmp /tmp' (same for /var/tmp) and check if it works (I'm not sure).

The other solution is described in this SWsoft FAQ; it is for Virtuozzo/vzfs although, so you need to modify it accordingly (comment out mkvzfs, change vzfs to simfs etc.).

Subject: Re: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 13:57:17 GMT
View Forum Message <> Reply to Message

Hi Kir,

Could you please provide an example on what you mean to change? I changed vzfs to simfs and mkvzfs to mkfs on mine, but did not have much luck.

/etc/sysconfig/vz-scripts/vps.mount

```
 #!/bin/bash
#
# This script is global and executed for every VPS at startup time.
# We are going to create and mount a temp area with nosuid, nodev and noexec,
# which will have vzquota configured and running.

# Current issues:
# 1) vzquota accepts only numeric and does it in a very weird way. Details below.
# 2) not clear how to handle on->off and off->on changes for tmp area--i.e. what to do with files
# under /tmp and /var/tmp.
# it's possible to move files back and forth on mount/umount stage--i.e.
#
# mv tmp temptmp
# mount tmparea
# tar -cf - -C temptmp . | tar xpf - -C tmp
#
# on mount and opposite action on umount but it may take considerable time - we have quotas already
# running, it's copying across mountpoits etc.
# 3) perhaps tmp should be added to /etc/fstab
# 4) completely unclear what to do with second-level quotas.

# script works with $VEID and $VE_CONFFILE vars which are passed as environment
# variables. All the rest can be defined
# a) in /etc/sysconfig/vz as a system-wide
# and b) in VE config file.

# tmp sizes/limits
VPSTMP_BLOCKS=$((150*1024))
VPSTMP_INODES=2000

# tmp 'path' - we might want have it outside
# of /vz
TMPPATH="/vz/private"
VPSTMP="$VEID-temparea"

# currently service VPS just doesn't work right
# with a dedicated nosuid / noexec TMP.

if [ $VEID -eq 1 ]; then
    exit 0
fi

# source configs.
if [ -f /etc/sysconfig/vz ]; then
    . /etc/sysconfig/vz
else
    exit 1
fi

if [ -f $VE_CONFFILE ]; then
    . $VE_CONFFILE
```

```
else
   exit 1
fi

# a special var from either global file or VPS config.
if [ -z "$VPS_TMP_AREA" ]; then
   # TMP area not configured in neither config.
   exit 0
fi

if [ "$VPS_TMP_AREA" != "yes" -a "$VPS_TMP_AREA" != "YES" ]; then
   # TMP area is disabled in either config
   exit 0
fi

# after sourcing configs we might have blocks/inodes in limit:barrier form
# and have to handle it. Perhaps we need to check that soft < hard here.

if [ "$VPSTMP_BLOCKS" = "${VPSTMP_BLOCKS/:/}" ]; then
   VPSTMP_BLOCKS_SOFT=$VPSTMP_BLOCKS
   VPSTMP_BLOCKS_HARD=$VPSTMP_BLOCKS
else
   VPSTMP_BLOCKS_SOFT=${VPSTMP_BLOCKS%%:*}
   VPSTMP_BLOCKS_HARD=${VPSTMP_BLOCKS##*:}
fi

if [ "${VPSTMP_INODES}" = "${VPSTMP_INODES/:/}" ]; then
   VPSTMP_INODES_SOFT=$VPSTMP_INODES
   VPSTMP_INODES_HARD=$VPSTMP_INODES
else
   VPSTMP_INODES_SOFT=${VPSTMP_INODES%%:*}
   VPSTMP_INODES_HARD=${VPSTMP_INODES##*:}
fi

# it seems that vzquota not only doesn't work with non-numeric but also silently
# removes non-numeric chars from supplied , without reporting errors.
# this indeed is very unfortunate since we have to use something like $00001
# instead of $VEID-tmparea for --otherwise there're some weird interaction
# between VPS and temparea quotas.

### WARNING!!!!!!#####
# VPS ID can not be more than 2^32-1, if you use "big" IDs for VPSs, you have to
# modify a var below to have VPSTMP_QUOTAID below the VPS ID "limit"
# (this limit also applies to quota IDs)

VPSTMP_QUOTAID=${VEID}1111

# other constants
# VZ_PRIVATE=/vz/private

# strip trailing slashes from TMPPATH
TMPPATH=${TMPPATH%%/?}

# extra sanity check
if [ "$TMPPATH/$VPSTMP" = "/" ]; then
   exit 1
fi

# if we don't have "vzfs filesystem" for the temp
# area, we have to create it, and init quota on it.
if [ ! -d "$TMPPATH/$VPSTMP" ]; then
   mkvzfs $TMPPATH/$VPSTMP
   RETVAL=$?
   if [ $RETVAL -ne 0 ]; then
      # some logging?
      exit $RETVAL
   fi
   vzquota init $VPSTMP_QUOTAID -p $TMPPATH/$VPSTMP \
      -c /var/vzquota/quota.$VPSTMP_QUOTAID \
      --block-softlimit $VPSTMP_BLOCKS_SOFT \
      --block-hardlimit $VPSTMP_BLOCKS_HARD \
      --block-exptime 0 \
      --inode-softlimit $VPSTMP_INODES_SOFT \
      --inode-hardlimit $VPSTMP_INODES_HARD \
      --inode-exptime 0
   RETVAL=$?
   if [ $RETVAL -ne 0 ]; then
      # some logging?
      exit $RETVAL
   fi
fi

# turning quota on.
vzquota on $VPSTMP_QUOTAID
RETVAL=$?
if [ $RETVAL -ne 0 ]; then
   # some logging
   exit $RETVAL
fi

# OK, assuming that everything is done. Now we need to mount tmp.
if [ ! -d "$TMPPATH/$VPSTMP" ]; then
   # something really is broken.
   exit 1
else
   mount -t vzfs \
        -o noatime,nosuid,noexec,nodev,rw,/vz/template:$TMPPATH/$VPSTMP \
        /vz/template:$TMPPATH/$VPSTMP $VE_ROOT/tmp
    RETVAL=$?
    if [ $RETVAL != 0 ]; then
        # some logging
        exit $RETVAL
    fi
    # we want tmp to have 1777 mode
    chmod 1777 $VE_ROOT/tmp
fi

# if we are here, everything is good so far
# we want to make /var/tmp to be symlink to /tmp.

if [ ! -L $VE_ROOT/var/tmp ]; then
    rm -rf $VE_ROOT/var/tmp
    ln -s /tmp $VE_ROOT/var/tmp
fi

exit 0
```

/etc/sysconfig/vz-scripts/vps.umount

```
 #!/bin/bash
#
# this script is global and executed for every VPS at stop time
# we're going to umount a temp area and stop vzquota for it.

# script works with $VEID and $VE_CONFFILE vars which are passed as environment
# variables. All the rest can be defined
# a) in /etc/sysconfig/vz as a system-wide
# and b) in VE config file.

TMPPATH="/vz/private"
VPSTMP="$VEID-temparea"

# currently service VPS just doesn't work right
# with a dedicated nosuid / noexec TMP.

if [ $VEID -eq 1 ]; then
    exit 0
fi

# source configs.
if [ -f /etc/sysconfig/vz ]; then
```

```
    . /etc/sysconfig/vz
else
    exit 1
fi

if [ -f $VE_CONFFILE ]; then
    . $VE_CONFFILE
else
    exit 1
fi

# script is really simple and most likely should be changed completely

VPSTMP_QUOTAID=${VEID}1111

if grep -q $VPSTMP /proc/mounts; then
    umount $VE_ROOT/tmp
    RETVAL=$?
    if [ $RETVAL -ne 0 ]; then
        # some logging?
        # do we need 'umount -f' here?
        exit $RETVAL
    fi
    vzquota off $VPSTMP_QUOTAID
    RETVAL=$?
    if [ $RETVAL -ne 0 ]; then
        # some logging?
        exit $RETVAL
    fi
fi

exit 0
```

Also, I am assuming we can put this inside the /etc/sysconfig/vz-scripts/VEID.conf:

```
VPS_TMP_AREA="yes"
```

Subject: Re: /dev/loop support inside VEs?
Posted by kir on Tue, 07 Feb 2006 14:01:29 GMT

View Forum Message <> Reply to Message

(1) mkvzfs - just comment it out, no need at all.
(2) change 'mount -t vzfs' to 'mount -t simfs'

What exact problem do you have? Is script gets executed by vzctl mount or vzctl start? If yes - trace it (set -x) and find out then it does something wrong

---

Subject: Re: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 14:08:33 GMT

View Forum Message <> Reply to Message

Hi Kir,

I have done as you instructed:

```
[root@spv505 vz-scripts]# vzctl start 1050
Starting VPS ...
Error executing mount script /etc/sysconfig/vz-scripts/vps.mount
```

Here's the relevant info from the strace:

```
[root@spv505 vz-scripts]# strace -o /root/strace.log vzctl start 1050
+ strace -o /root/strace.log vzctl start 1050
Starting VPS ...
+ VPSTMP_BLOCKS=153600
+ VPSTMP_INODES=2000
+ TMPPATH=/vz/private
+ VPSTMP=1050-temparea
+ '[' 1050 -eq 1 ']'
+ '[' -f /etc/sysconfig/vz ']'
+ . /etc/sysconfig/vz
++ VIRTUOZZO=yes
++ LOCKDIR=/vz/lock
++ VE0CPUUNITS=1000
++ LOGGING=yes
++ LOGFILE=/var/log/vzctl.log
++ LOG_LEVEL=0
++ DISK_QUOTA=yes
++ VZFASTBOOT=no
++ TEMPLATE=/vz/template
++ VE_ROOT=/vz/root/1050
++ VE_PRIVATE=/vz/private/1050
++ CONFIGFILE=vps.basic
++ DEF_OSTEMPLATE=centos-4-i386-minimal
++ VZWDOG=no
++ IPTABLES='ipt_REJECT ipt_tos ipt_limit ipt_multiport iptable_filter iptable_mangle
ipt_TCPMSS ipt_tcpmss ipt_ttl ipt_length'
+ '[' -f /usr/lib/vzctl/scripts/1050.conf ']'
+ exit 1
Error executing mount script /etc/sysconfig/vz-scripts/vps.mount
```

Subject: Re: /dev/loop support inside VEs?
Posted by kir on Tue, 07 Feb 2006 14:21:01 GMT

View Forum Message <> Reply to Message

Ughm, now I see the problem is in incorrect VE_CONFFILE set in vzctl.

Which distro are you on? Which vzctl version?

Subject: Re: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 14:23:56 GMT

View Forum Message <> Reply to Message

Hi,

Its CentOS 4.2 and the vzctl version is: vzctl-2.7.0-25

I probably should upgrade eh?

Subject: Re: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 14:26:33 GMT

View Forum Message <> Reply to Message

I bumped up to the vzctl-2.7.0-26 version and still no luck.

Subject: Re: /dev/loop support inside VEs?
Posted by kir on Tue, 07 Feb 2006 14:49:41 GMT

View Forum Message <> Reply to Message

This is definitely a bug in vzctl; thanks for catching this, and sorry for spotting this!

See bug #100; you can add yourself to Cc: for this bug to track its progress -- it should be fixed in the next vzctl release.

For now, you should change
if [-f $VE_CONFFILE]; then
to
VE_CONFFILE=/etc/sysconfig/vz-scripts/${VEID}.conf
if [-f $VE_CONFFILE]; then
in both vps.mount and vps.umount scripts. I.e. you should add a single line (VE_CONFFILE=...) to the both scripts.

Subject: Re: /dev/loop support inside VEs?
Posted by phpfreak on Tue, 07 Feb 2006 15:29:42 GMT

View Forum Message <> Reply to Message

Would it be easier to just symlink:

/etc/sysconfig/vz-scripts to /usr/lib/vzctl/scripts for now? This way when the patched version of vzctl comes out, we won't have to modify our vps.mount scripts again.

Just a thought.

Edit:
Ahh, nevermind. /usr/lib/vzctl/scripts already exists. I'll modify the mount scripts.