

KVM - Samepage merging and partition resizing

[http://www.thomas-krenn.com/de/wiki/KSM_\(Kernel_Samepage_Merging\)](http://www.thomas-krenn.com/de/wiki/KSM_(Kernel_Samepage_Merging))

/etc/init.d/ksm

```
#!/bin/sh#
# Author:          Marin Atanasov Nikolov <dnaeon@gmail.com>
#
# https://github.com/dnaeon/ksm-init.d-debian
#
### BEGIN INIT INFO
# Provides:        ksm
# Required-Start:
# Required-Stop:
# X-Start-Before:
# Default-Start:   2 3 4 5
# Default-Stop:    0 1 6
# Short-Description: Enable and disable KVM KSM
# Description:     Enables and disables the KVM Kernel Samepage Merging
#                  feature of the kernel
### END INIT INFO
set -e
. /lib/lsb/init-functions
do_start() {
    echo 1 > /sys/kernel/mm/ksm/run
    # default is 100
    echo 20 > /sys/kernel/mm/ksm/pages_to_scan
    # default is 20, fabian hat 250
    echo 100 > /sys/kernel/mm/ksm/sleep_millisecs
    #echo 20 > /sys/kernel/mm/ksm/sleep_millisecs
    log_success_msg "Enabling Kernel Samepage Merging"
}
do_stop() {
    echo 0 > /sys/kernel/mm/ksm/run
    log_success_msg "Disabling Kernel Samepage Merging"
}
do_status() {
    local ksm_status
    ksm_status=$( cat /sys/kernel/mm/ksm/run )
    if [ ${ksm_status} -eq 1 ]; then
        ksm_pages_shared=$(cat /sys/kernel/mm/ksm/pages_shared)
        log_success_msg "Kernel Samepage Merging is enabled, $ksm_pages_shared pages shared"
    else
        log_success_msg "Kernel Samepage Merging is disabled"
    fi
}
case "${1}" in
    start)
        do_start
        ;;
    reset|stop)
        do_stop
        ;;
    status)
        do_status
        ;;
    reload|restart|force-reload)
        do_stop
        do_start
        ;;
    *)
        log_success_msg "usage: ${0} {start|stop|status|reload|restart|force-reload|reset}" >&2
        ;;
esac
```

Resizing a partition within virtual disk

Short Story

```
# Shutdown VM
virsh shutdown sles11-resizetest
# Add 5GB to virtual disk
qemu-img resize disk0.raw +5G
# Change partition table of virtual disk. Partition 3 is / with ext3
fdisk disk0.raw (d,3,n,p,3,enter,enter,w)
# Create loop devices from partitions withing virtueal disk
kpartx -a -v disk0.raw
# Remove journal of root partition
tune2fs -O has_journal /dev/mapper/loop0p3
# Check consistency
e2fsck -f /dev/mapper/loop0p3
# Resize root partition to maximum
resize2fs /dev/mapper/loop0p3
# Check again
e2fsck -f /dev/mapper/loop0p3
# Activate journal
tune2fs -j /dev/mapper/loop0p3
# Remove loop devices
kpartx -d -v disk0.raw
# Start VM
virsh start sles11-resizetest
```

Long Story

```
kvm:/opt/kvm/images/sles11_resizetest # ll
total 11514072
-rw----- 1 root root 10737418240 Apr 11 10:27 disk0.raw

kvm:/opt/kvm/images/sles11_resizetest # fdisk -l disk0.raw
Disk disk0.raw: 10.7 GB, 10737418240 bytes
255 heads, 63 sectors/track, 1305 cylinders, total 20971520 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0000e90b
   Device Boot      Start         End      Blocks   Id  System
disk0.raw1   *          2048     1028095     513024   83   Linux
disk0.raw2             1028096     4177919     1574912   82   Linux swap / Solaris
disk0.raw3             4177920     20971519     8396800   83   Linux
qemu-img resize disk0.raw +5G
kvm:/opt/kvm/images/sles11_resizetest # ll
total 11514072
-rw----- 1 root root 16106127360 Apr 16 11:39 disk0.raw

kvm:/opt/kvm/images/sles11_resizetest # fdisk -l disk0.raw
Disk disk0.raw: 16.1 GB, 16106127360 bytes
255 heads, 63 sectors/track, 1958 cylinders, total 31457280 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0000e90b
   Device Boot      Start         End      Blocks   Id  System
disk0.raw1   *          2048     1028095     513024   83   Linux
disk0.raw2             1028096     4177919     1574912   82   Linux swap / Solaris
disk0.raw3             4177920     20971519     8396800   83   Linux
kvm:/opt/kvm/images/sles11_resizetest # fdisk disk0.raw
Command (m for help): p
Disk disk0.raw: 16.1 GB, 16106127360 bytes
255 heads, 63 sectors/track, 1958 cylinders, total 31457280 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
```

```

Disk identifier: 0x0000e90b
  Device Boot      Start          End      Blocks  Id System
disk0.raw1 *        2048          1028095    513024  83 Linux
disk0.raw2          1028096      4177919    1574912  82 Linux swap / Solaris
disk0.raw3          4177920      20971519   8396800  83 Linux

```

Command (m for help): d

Partition number (1-4): 3

Command (m for help): n

Command action
 e extended

p primary partition (1-4)

P

Partition number (1-4, default 3): 3

First sector (4177920-31457279, default 4177920):

Using default value 4177920

Last sector, +sectors or +size{K,M,G} (4177920-31457279, default 31457279):

Using default value 31457279

Command (m for help): p

Disk disk0.raw: 16.1 GB, 16106127360 bytes

255 heads, 63 sectors/track, 1958 cylinders, total 31457280 sectors

Units = sectors of 1 * 512 = 512 bytes

Sector size (logical/physical): 512 bytes / 512 bytes

I/O size (minimum/optimal): 512 bytes / 512 bytes

Disk identifier: 0x0000e90b

```

  Device Boot      Start          End      Blocks  Id System
disk0.raw1 *        2048          1028095    513024  83 Linux
disk0.raw2          1028096      4177919    1574912  82 Linux swap / Solaris
disk0.raw3          4177920      31457279   13639680  83 Linux

```

Command (m for help): w

The partition table has been altered!

Syncing disks.

```
kvm:/opt/kvm/images/sles11_resizetest # kpartx -a -v disk0.raw
```

```
add map loop0p1 (253:0): 0 1026048 linear /dev/loop0 2048
```

```
add map loop0p2 (253:1): 0 3149824 linear /dev/loop0 1028096
```

```
add map loop0p3 (253:2): 0 27279360 linear /dev/loop0 4177920
```

```
kvm:/opt/kvm/images/sles11_resizetest # tune2fs -O ^has_journal /dev/mapper/loop0p3
```

```
tune2fs 1.41.9 (22-Aug-2009)
```

```
kvm:/opt/kvm/images/sles11_resizetest # e2fsck -f /dev/mapper/loop0p3
```

```
e2fsck 1.41.9 (22-Aug-2009)
```

```
Pass 1: Checking inodes, blocks, and sizes
```

```
Pass 2: Checking directory structure
```

```
Pass 3: Checking directory connectivity
```

```
Pass 4: Checking reference counts
```

```
Pass 5: Checking group summary information
```

```
/dev/mapper/loop0p3: 105712/525200 files (0.3% non-contiguous), 604662/2099200 blocks
```

```
kvm:/opt/kvm/images/sles11_resizetest # resize2fs /dev/mapper/loop0p3
```

```
resize2fs 1.41.9 (22-Aug-2009)
```

```
Resizing the filesystem on /dev/mapper/loop0p3 to 3409920 (4k) blocks.
```

```
The filesystem on /dev/mapper/loop0p3 is now 3409920 blocks long.
```

```
kvm:/opt/kvm/images/sles11_resizetest # e2fsck -f /dev/mapper/loop0p3
```

```
e2fsck 1.41.9 (22-Aug-2009)
```

```
Pass 1: Checking inodes, blocks, and sizes
```

```
Pass 2: Checking directory structure
```

```
Pass 3: Checking directory connectivity
```

```
Pass 4: Checking reference counts
```

```
Pass 5: Checking group summary information
```

```
/dev/mapper/loop0p3: 105712/848400 files (0.3% non-contiguous), 625456/3409920 blocks
```

```
kvm:/opt/kvm/images/sles11_resizetest # tune2fs -j /dev/mapper/loop0p3
```

```
tune2fs 1.41.9 (22-Aug-2009)
```

```
Creating journal inode: done
```

```
This filesystem will be automatically checked every 28 mounts or 180 days, whichever comes first. Use tune2fs -c or -i to override.
```

```
kvm:/opt/kvm/images/sles11_resizetest # kpartx -d -v disk0.raw
```

```
del devmap : loop0p3
```

```
del devmap : loop0p2
```

```
del devmap : loop0p1
```

```
loop deleted : /dev/loop0
```

Monitoring

Munin

`/usr/lib/munin/plugins/ksm`

```
#!/bin/sh
# https://www.kernel.org/doc/Documentation/vm/ksm.txt
# http://munin-monitoring.org/wiki/HowToWritePlugins
case $1 in
  config)
    cat <<'EOM'
graph_title KSM pages shared
graph_vlabel Number of memory pages shared
pages_shared.label pages shared
pages_shared.info how many shared pages are being used
pages_shareing.label pages shareing
pages_shareing.info how many more sites are sharing them i.e. how much saved
pages_unshared.label pages unshared
pages_unshared.info how many pages unique but repeatedly checked for merging
pages_volatile.label pages volatile
pages_volatile.info how many pages changing too fast to be placed in a tree
graph_args --base 1000 -l 0
graph_scale no
graph_category system
EOM
    exit 0;;
esac

printf "pages_shared.value "
cat /sys/kernel/mm/ksm/pages_shared
printf "pages_shareing.value "
cat /sys/kernel/mm/ksm/pages_sharing
printf "pages_unshared.value "
cat /sys/kernel/mm/ksm/pages_unshared
printf "pages_volatile.value "
cat /sys/kernel/mm/ksm/pages_volatile
```

