Tips & Tricks for Using LVM Effectively

Renay Gaye
Hewlett-Packard
renay.gaye@hp.com

HP WORLD 2003
Solutions and Technology Conference & Expo

Session Topics

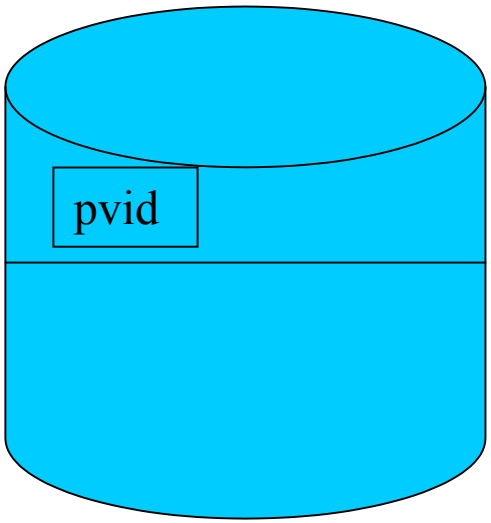
- LVM Concepts
- Multi-Pathing Solutions
- Moving Data in an LVM Environment
- Renaming LVM Objects
- Mirroring
- LVM Boot Disks
- Recovering Corrupted LVM Info
- LVM Performance Tips
- LVM in a MC/Service Guard Environment
- Disk Array Data Replication Issues

LVM Concepts

HP WORLD 2003
Solutions and Technology Conference & Expo

LVM Concepts-Physical Volume

- LVM Managed Disk
- Each PV is assigned a unique PVID

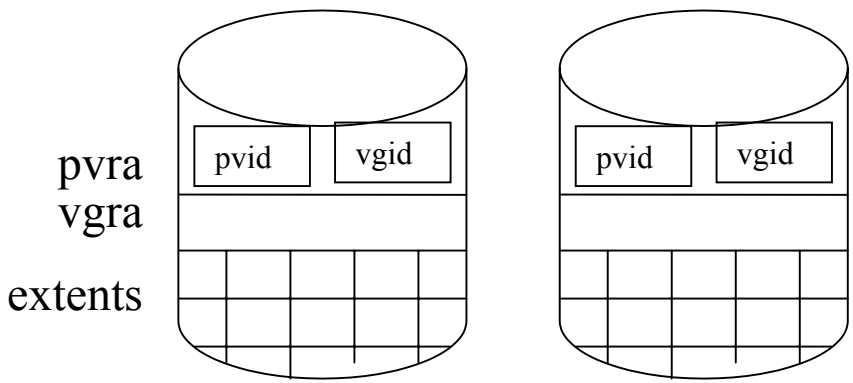


pvcreate /dev/rdsk/*cntndn*

pvcreate –f /dev/rdsk/*cntndn*

LVM Concepts-Volume Group

- One or more physical volumes
- Pool of physical extents

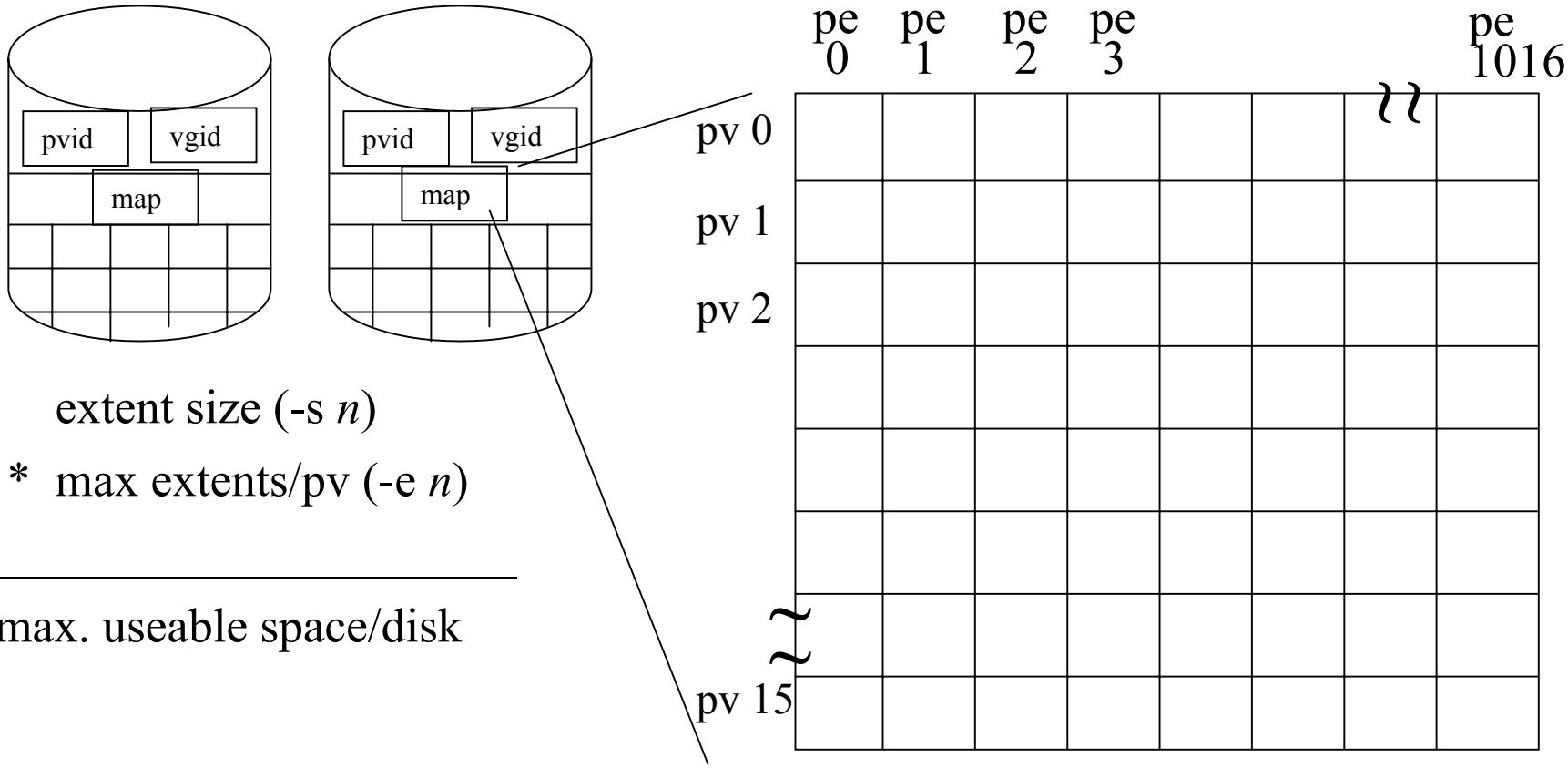


mkdir /dev/vg*name*

mknod /dev/vg*name*/group c 64
0x*nn*0000

vgcreate vg*name* /dev/dsk/c*n*t*n*d*n* ...

Volume Group Map



extent size (-s *n*)

* max extents/pv (-e *n*)

---

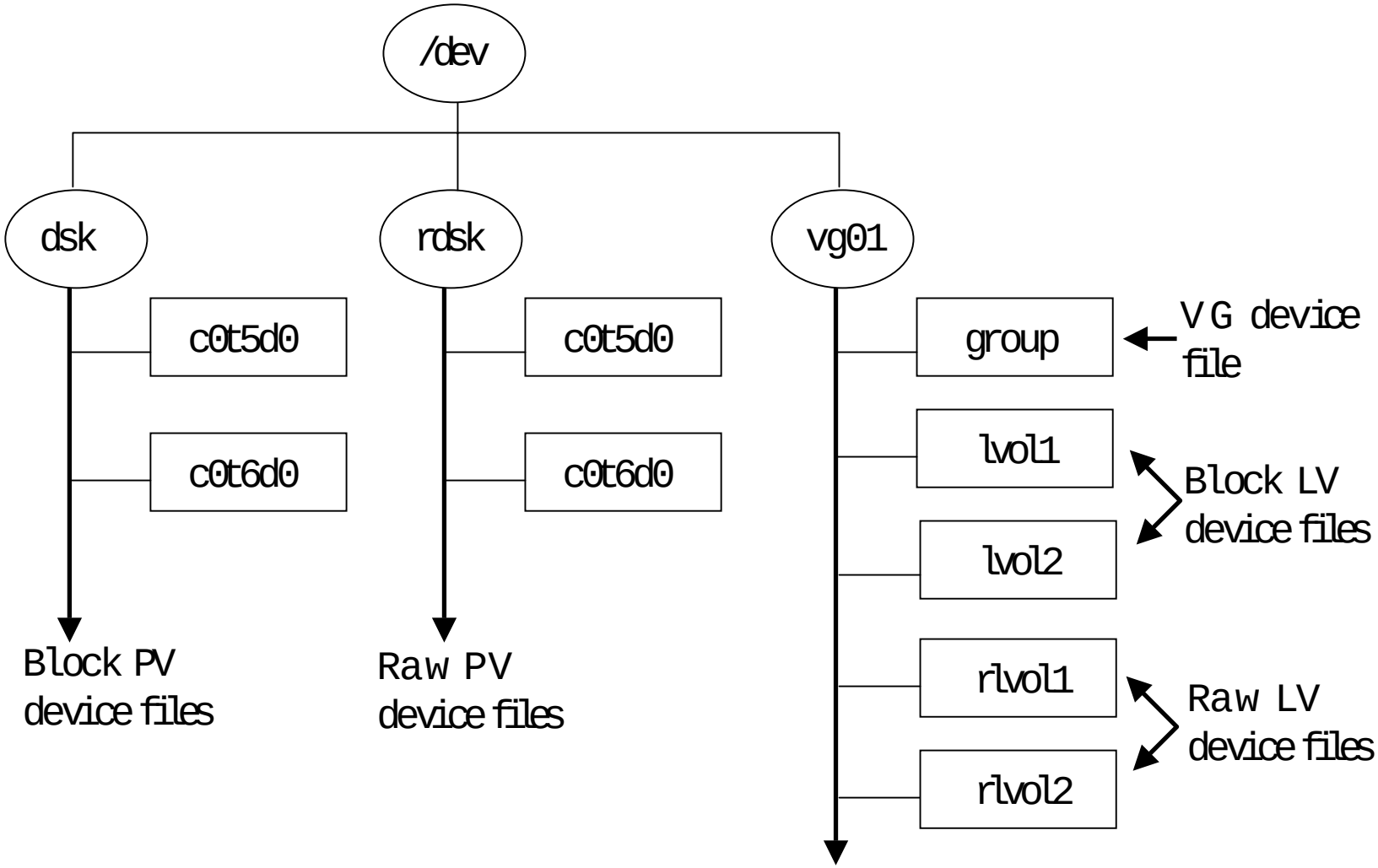
max. useable space/disk

Maximum Useable Space

Using default of 16 PVs per VG:

Extent Size	Max Extents	Max Disk Size
1MB	7676	8GB
2MB	15612	33GB
4MB	31484	132GB
8MB	63228	530GB

LVM Device Files



/etc/lvmtab

- Non-ascii file
- Records volume group/physical volume associations
- Used by many LVM commands for sanity checking

example:

/dev/vg00 (*vgid)*

/dev/dsk/c0t6d0 *(vgid,pvid)*

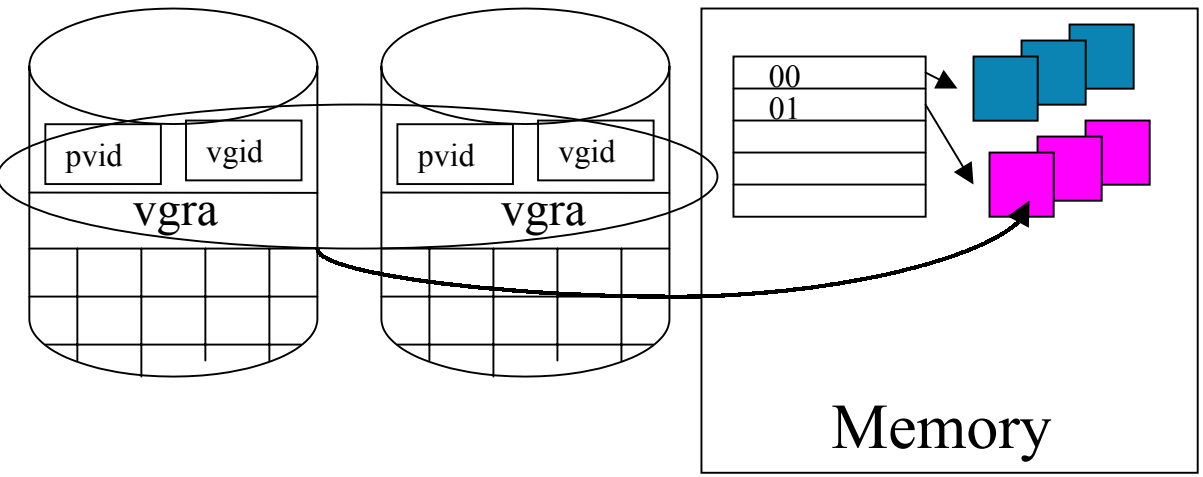
/dev/vg01 *(vgid)*

/dev/dsk/c1t2d1 *(vgid,pvid)*

/dev/dsk/c1t2d2 *(vgid,pvid)*

Volume Group Activation

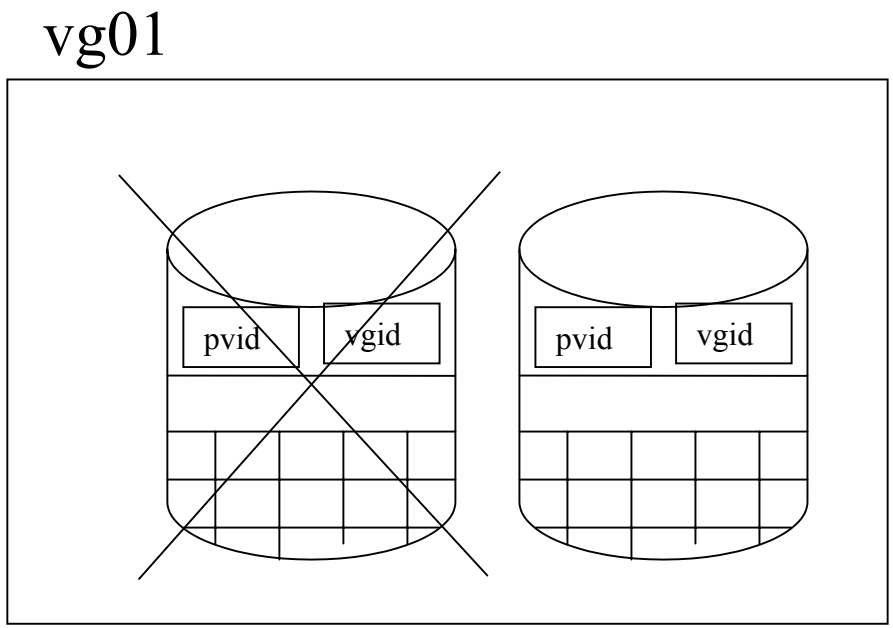
- Done automatically when vg is created
- automatically at boot (/etc/lvmrc)
- required in order to access any lvols



vgchange -a y vg*name*

vgchange -a r vg*name*

LVM Concepts-Quorum



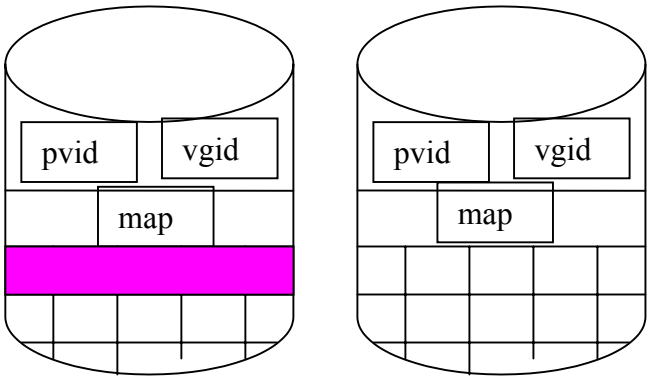
- More than 1/2 disks in a vg required to activate the volume group
- Can override:

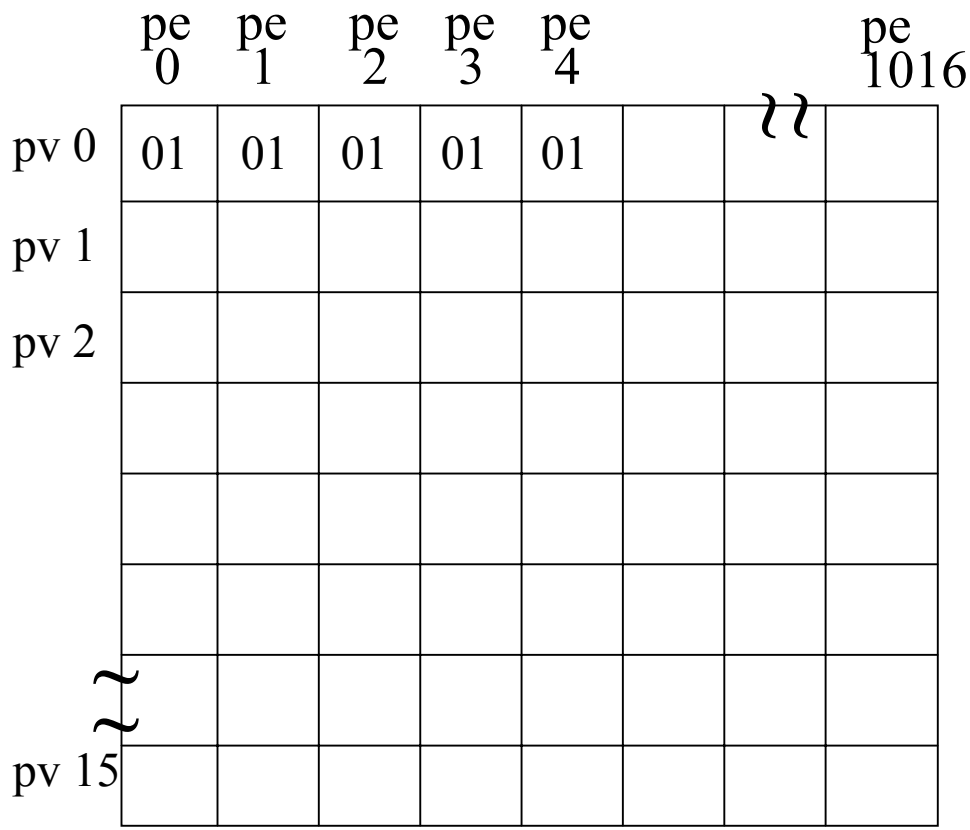
```
vgchange -a y -q n vg01
```

- Booting without quorum:

```
ISL> hpux -lq
```

LVM Concepts-Logical Volume





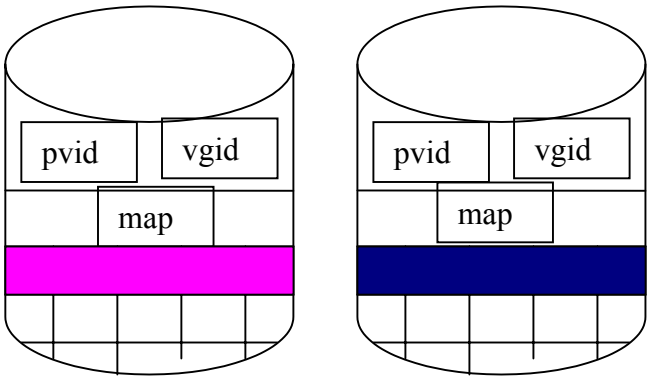
lvcreate -L *n* vg*name*

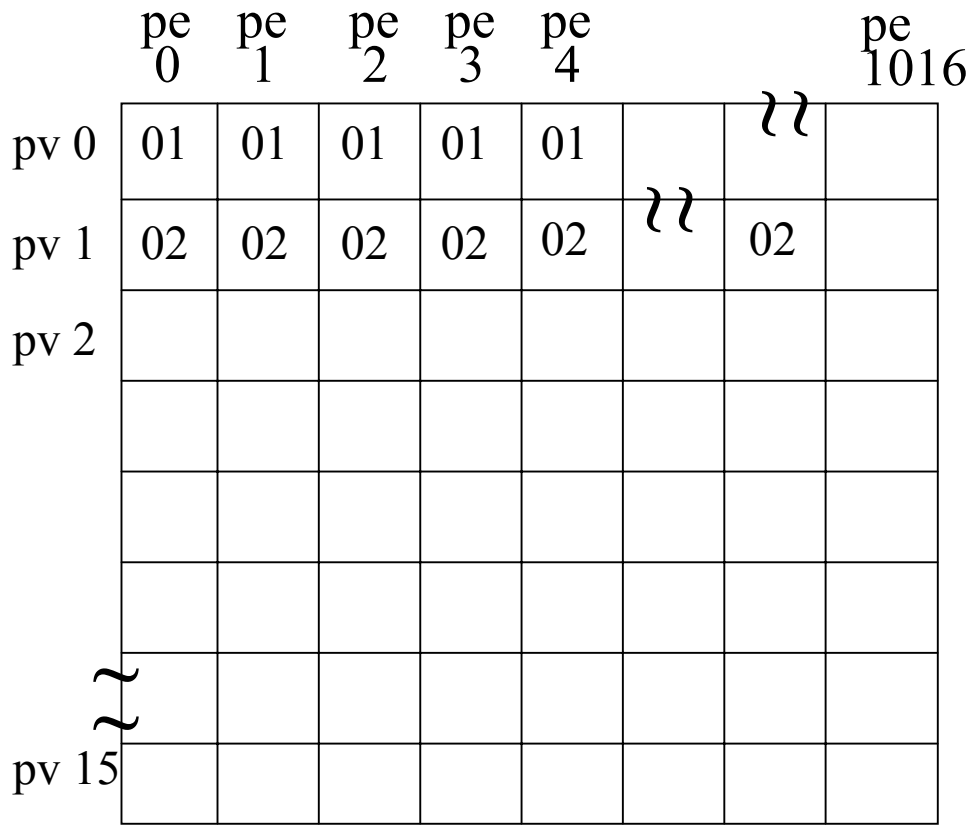
lvcreate -L *n* -C y vg*name*

lvcreate -l *n* vg*name*

lvcreate -L *n* -n *name* vg*name*

Placing logical volumes





lvcreate -n datalv vgdata

lvextend -L 500
/dev/vgdata/datalv

/dev/dsk/c4t2d0

Multi-Pathing Solutions

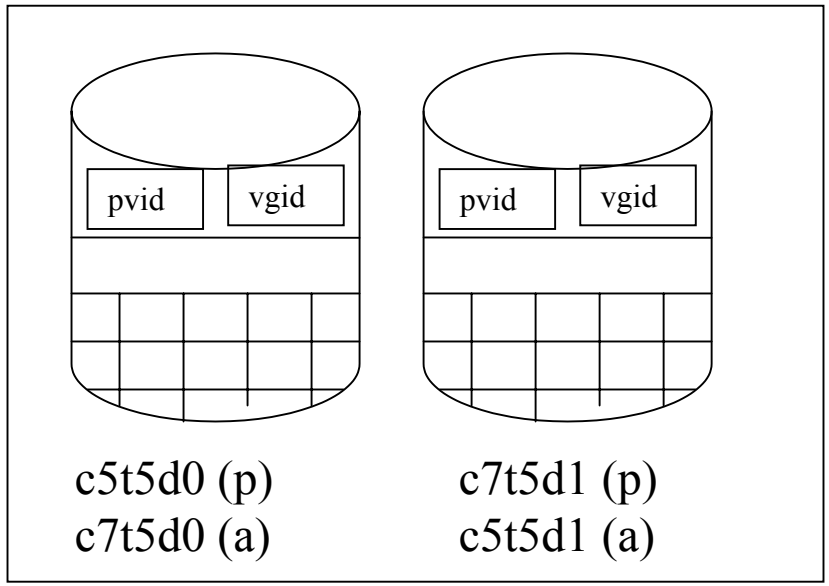
HP WORLD 2003
Solutions and Technology Conference & Expo

Multipathing Solutions with LVM

- LVM pvlinks
  - active/passive
  - no additional software required
- Auto path
  - active/active, load balancing
  - supported for XP and VA arrays
  - additional cost product
- Powerpath
  - active/active, load balancing
  - supported for EMC arrays
  - additional cost product

PV links

```
pvcreate /dev/rdsk/c5t5d0

pvcreate /dev/rdsk/c5t5d1

mkdir /dev/vg01

mknod /dev/vg01/group c 64
0x010000

vgcreate vg01
/dev/dsk/c5t5d0 /dev/dsk/c7t5d0
/dev/dsk/c7t5d1 /dev/dsk/c5t5d1
```



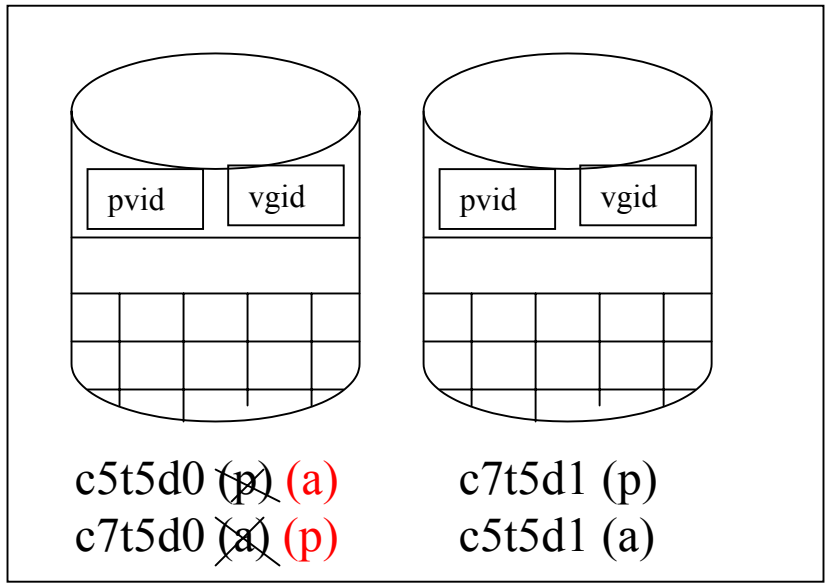
PV links-switching the order

- Temporary Change

```
pvchange -s /dev/dsk/c7t5d0
```

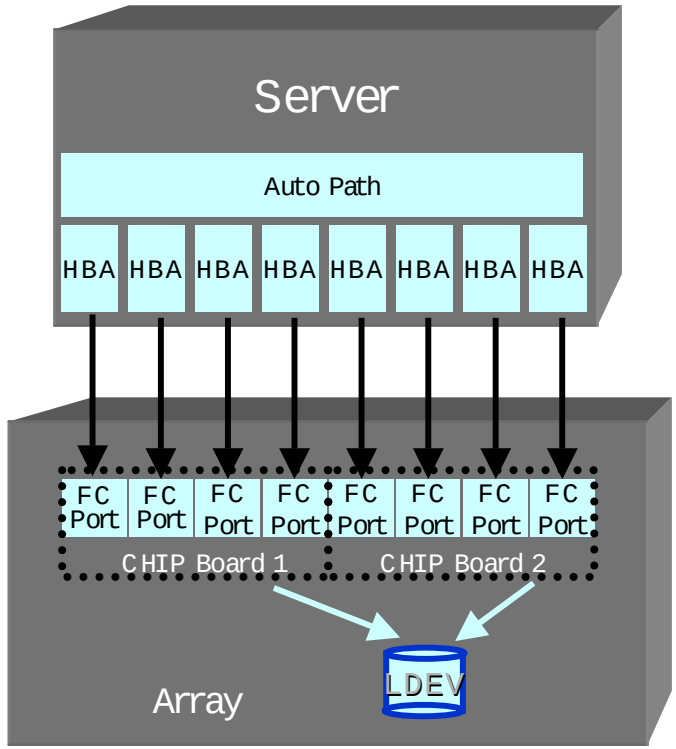
- Permanent Change

```
vgreduce vg01 /dev/dsk/c5t5d0
vgextend vg01 /dev/dsk/c5t5d0
```



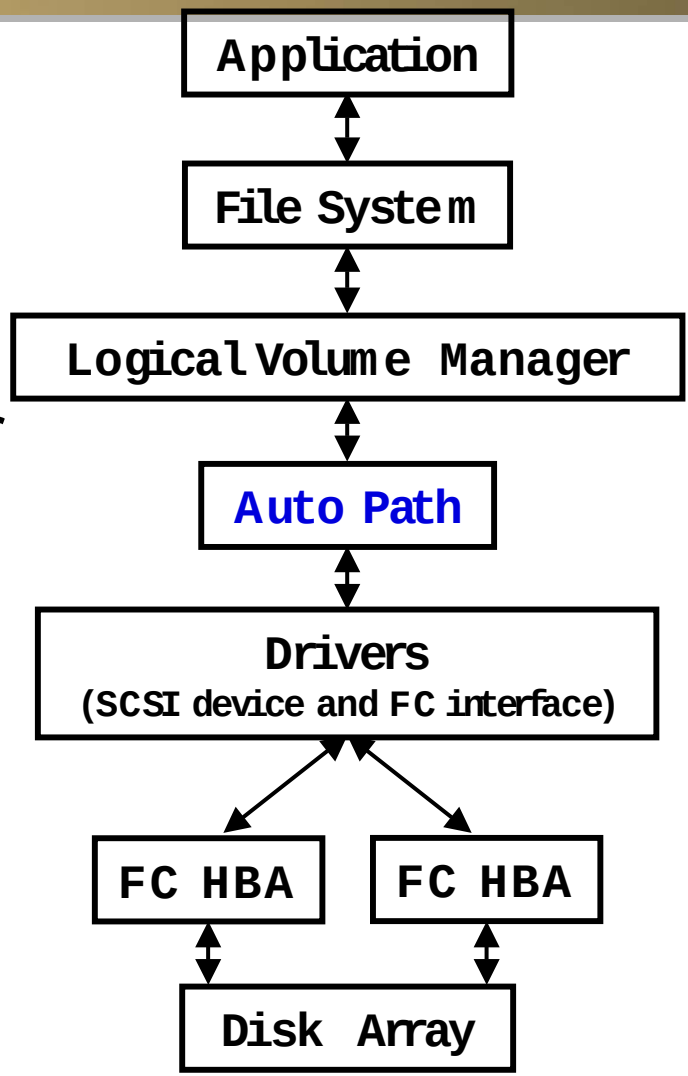
Auto Path-Dynamic load balancing

- Load balancing functionality
  - Supports up to 8 paths from a server to an end device
  - Provides dynamic load balancing across all paths to an end device
  - Choose from 4 load balancing policies, including “no load balancing”
  - Supports the XP and VA disk arrays
  - Load balancing supported in clustered environments



Auto Path driver

- A pseudo driver
- The Auto Path driver is layered between the LVM (Logical Volume Manager) and the SCSI device driver
- The driver provides the Command Line Interface



Moving Data

HP WORLD 2003
Solutions and Technology Conference & Expo

Moving Disks

Three Step Process

Remove definition of volume group

Move disk(s)

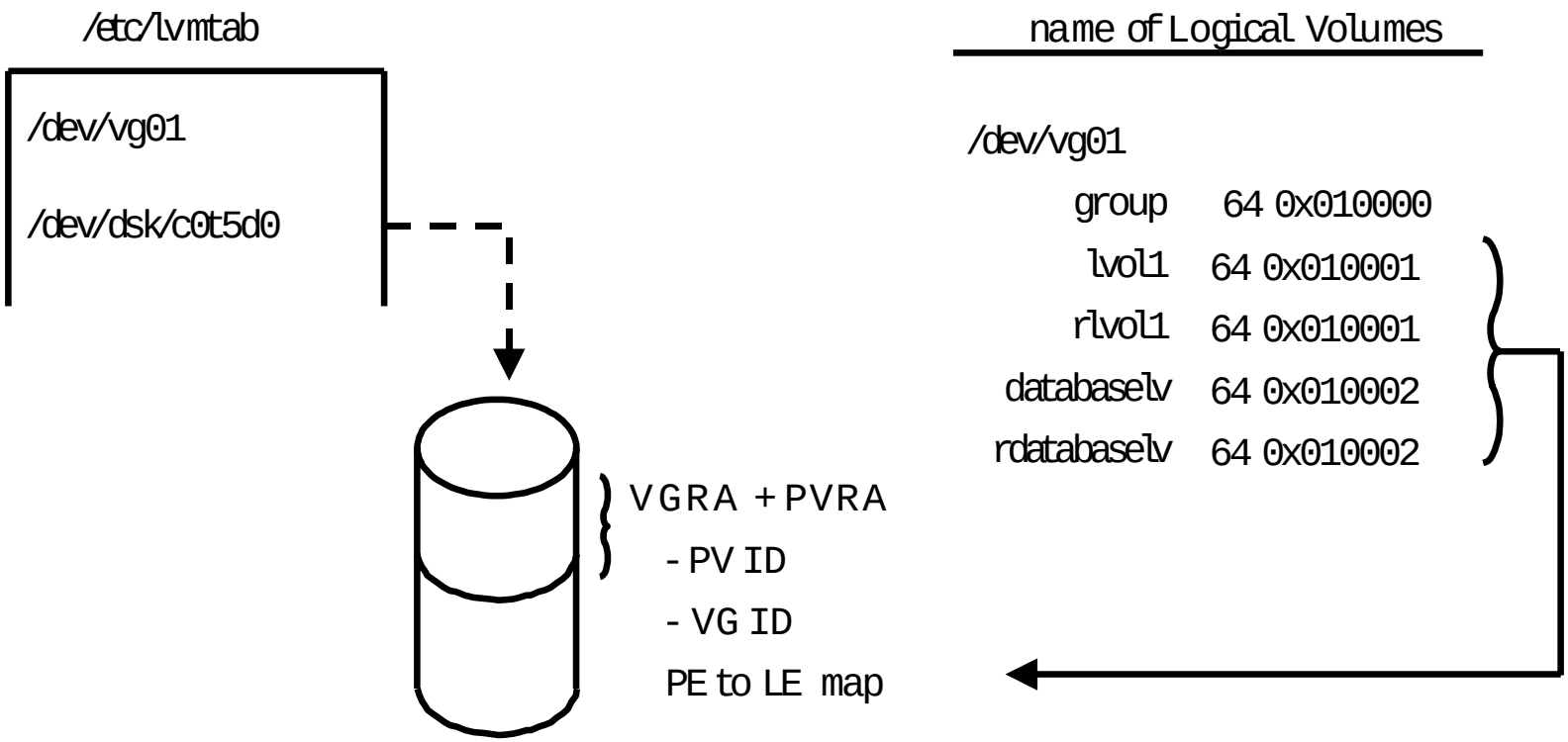
Add definition of volume group

Two commands

vgexport(1m)

vgimport(1m)

Volume Group Definition — Review



Exporting a Volume Group

Syntax:

`vgexport [-p][-v][-m` *`file`*`]VG`

- **`-p`** Preview actions only
- **`-v`** verbose
- **`-m`** used to specify a map file for logical volume names

- Removes volume group definition from the system completely by updating **`/etc/lvmtab`** and kernel memory.
- The volume group must first be deactivated with **`vgchange(1m)`.**

Example:

```
vgchange -a n /dev/vg01
vgexport  -v  -m  /etc/lvmconf/vg01.map  /dev/vg01
```

Importing a Volume Group

Syntax:

```
vgimport [-p][-v][-m file] VG PV [PV...]
```

`-p` Preview actions only

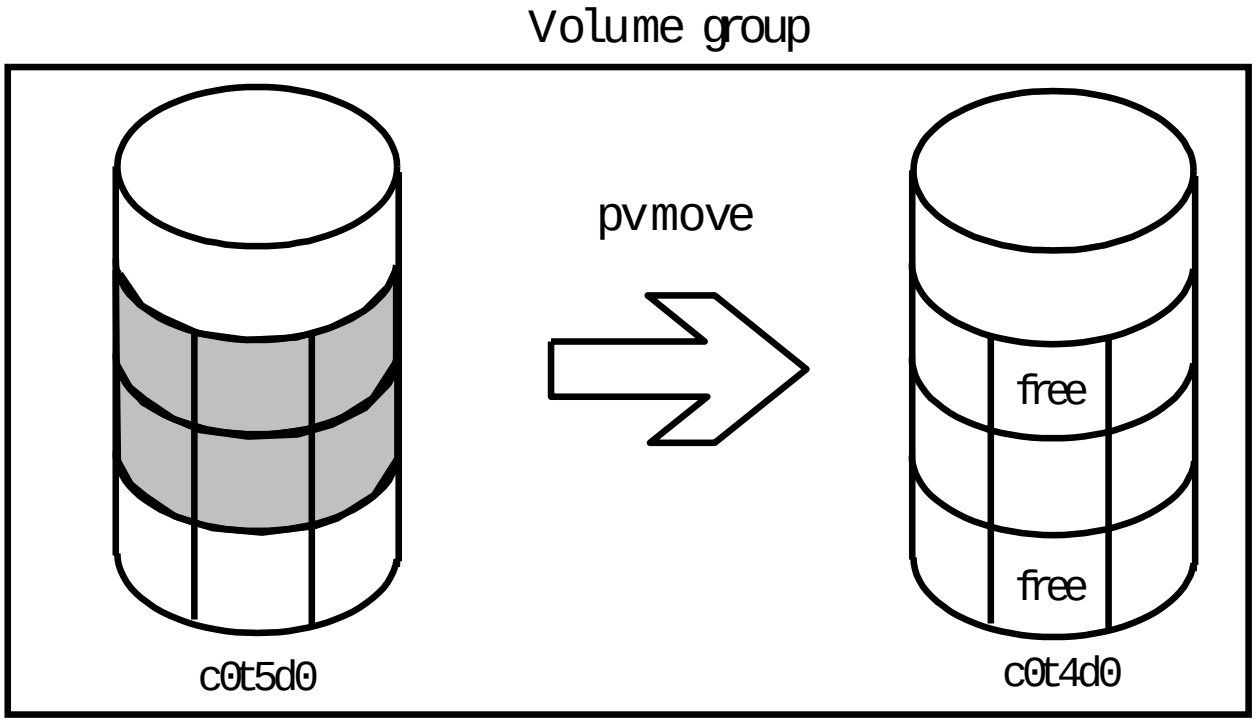
`-v` verbose

`-m` used to specify a map file for logical volume names

Example:

```
mkdir  /dev/vg01
mknod /dev/vg01/group c 64 0x010000
vgimport -v -m /etc/lvmconf/vg01.map /dev/vg01 /dev/dsk/c0t1d0
vgchange -a y /dev/vg01
vgcfgbackup vg01
```

Moving LVM Data



Syntax:

```
pvmove [-n lv] from_PV [to_PV]
```

Example:

```
pvmove -n /dev/vg01/lvol1 /dev/dsk/c0t5d0 /dev/dsk/c0t4d0
```

Renaming LVM Objects

HP WORLD 2003
Solutions and Technology Conference & Expo

Renaming Logical Volumes

- lvol names are not stored in the LVM maps or in /etc/lvmtab
- Simply rename device files and update /etc/fstab if necessary

Renaming Volume Groups

- Volume group name is kept in /etc/lvmtab and is used as the directory name to anchor the device files for the group
- Use vgexport and vgimport to rename volume group

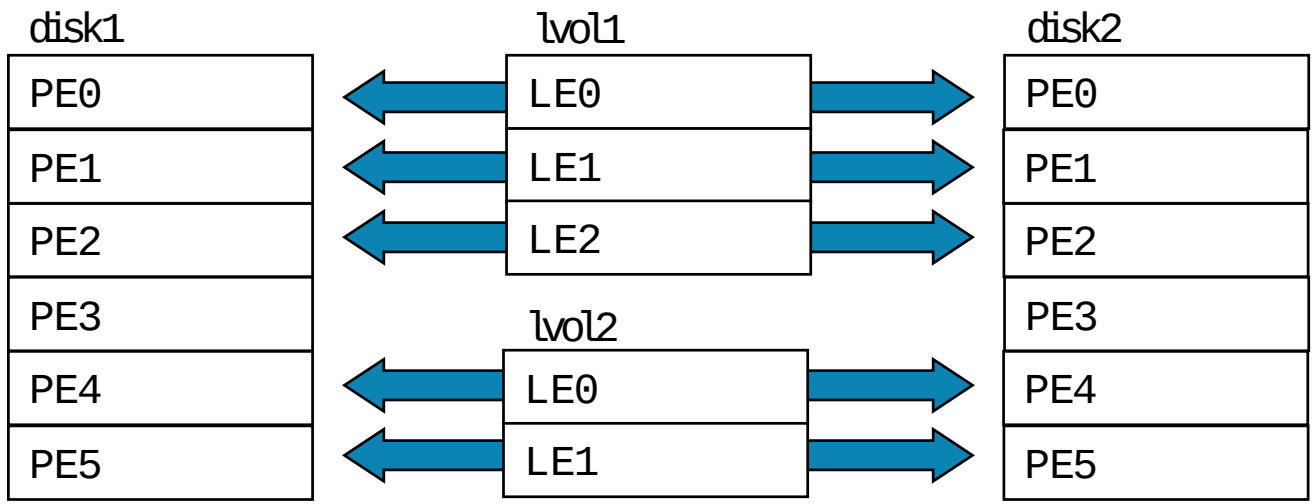
LVM Mirroring

HP WORLD 2003
Solutions and Technology Conference & Expo

Mirrored Volumes

In a mirrored logical volume ...

- Each logical volume consists of one or more Logical Extents (LEs).
- Each LE maps to two or three Physical Extents (PEs) on disk.
- The logical volume remains accessible even when one of the disks is unavailable.



Extending and Reducing Mirrors

Create a new, mirrored logical volume:

```
# lvcreate -m 1 -L 16 -n myfs1 vg01
```

Mirror an existing logical volume:

```
# lvextend -m 1 /dev/vg01/myfs1
```

Mirror an existing logical volume to a specific disk:

```
# lvextend -m 1 /dev/vg01/myfs1 /dev/dsk/c0t3d0
```

Add a second mirror:

```
# lvextend -m 2 /dev/vg01/myfs1
```

Remove a logical volume's mirrors:

```
# lvreduce -m 0 /dev/vg01/myfs1 /dev/dsk/c0t3d0
```

Check a mirrored logical volume's status:

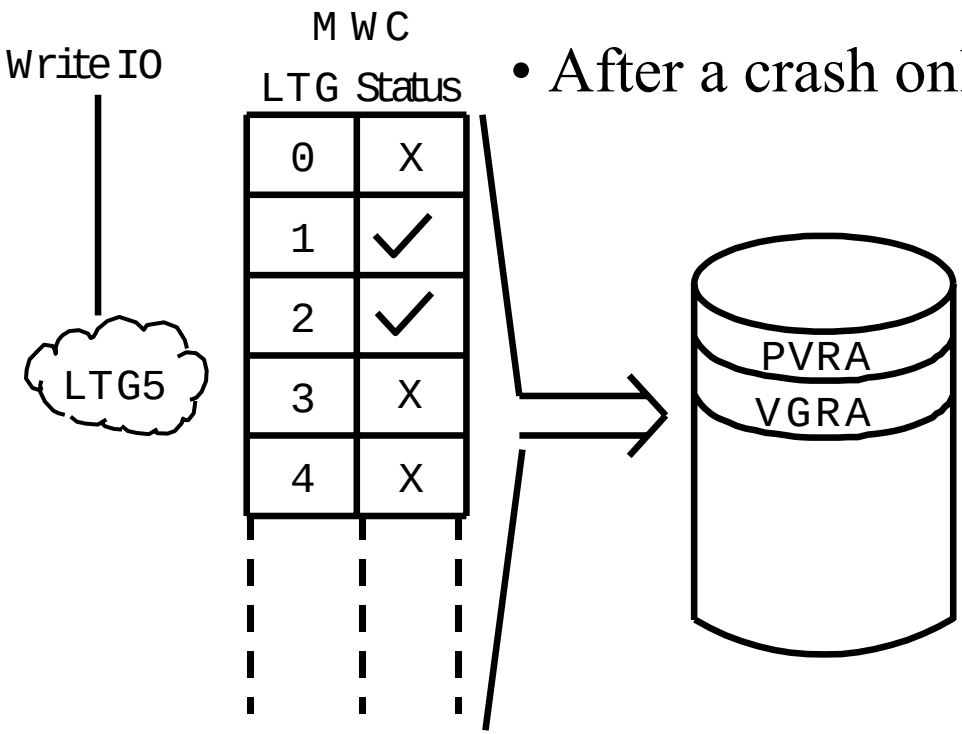
```
# lvdisplay -v /dev/vg01/myfs1
```

Mirrored I/O Scheduling

	Parallel (-d p)	Sequential (-d s)
Read	Access PV with lowest outstanding I/Os	Read in PV order
Write	Schedule writes simultaneously to all PVs	Schedule writes in PV order

MWC/MCR

- Writes are recorded in MWC in memory
- MCR record written to disk when a write is done to a logical track group not already recorded
- After a crash only "dirty" LTGs need be resynced



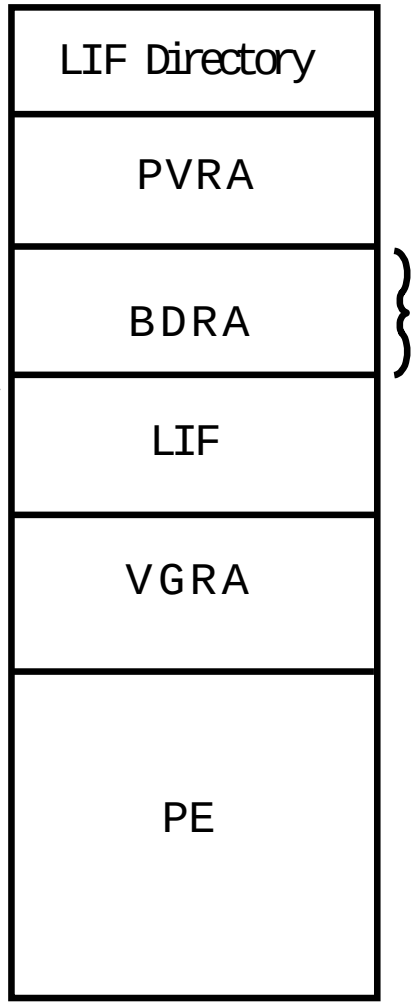
Mirror Consistency Recovery Options

		Advantage	Disadvantage
1	MWC/MCR **-M y**	Fast recovery on a system crash	Runtime overhead to write MCR record
2	NOMWC **-M n -c y**	No extra overhead at runtime	Slow recovery on a system crash
3	NONE **-M n -c n**	Application can do necessary recovery	No system managed recovery

LVM Boot Disks

HP WORLD 2003
Solutions and Technology Conference & Expo

LVM Boot volume structure



ISL,HPUX,AUTO,LABEL

- Hardware path/device file of disks in volume group
- boot, root, dump, swap lvols
- Maintained with the lvlnboot command

Mirroring the LVM Boot Disk

- Initialize disk for LVM. Leave room for BDRA/LIF

  `pvcreate -B /dev/rdsk/cntndn`

- Add disk to root VG

  `vgextend vg00 /dev/rdsk/cntndn`

- Install boot files in LIF area

  `mkboot /dev/rdsk/cntndn`

- Change the auto file on both the primary and alternate boot disk

  `mkboot -a "hpux -lq" /dev/rdsk/cntndn`

- Mirror each of the lvols in the root vg

  `lvextend -m 1 /dev/vg00/lvoln /dev/dsk/cntndn`

- Add the mirror disk definition to /stand/bootconf

Booting When BDRA is damaged

- Boot system into maintenance mode
  ISL> hpux -lm
- Activate vg00
  vgchange -a y vg00
- Use lvlnboot to examine/repair BDRA
  lvlnboot -v
  lvlnboot -b /dev/vg00/lvol1
  lvlnboot -r /dev/vg00/lvol3
  lvlnboot -s /dev/vg00/lvol2
  lvlnboot -d /dev/vg00/lvol2
- Reboot the system
  reboot

Moving the Boot Disk

- Problem: /etc/lvmtab contains the old device files for the boot disk
  - Solution: boot into maintenace mode, export and reimport volume group
- Problem: The BDRA and Label files contain the old device information
  - Solution: use lvlnboot to fix

Cookbook for moving root disk

Boot from new device. Reply Y to interact with IPL

ISL> hpux -lm

vgexport -v -m vg00.map vg00

mkdir /dev/vg00

mknod /dev/vg00/group c 64 0x000000

vgimport -v -m vg00.map vg00 /dev/dsk/*new_device_file*

vgchange -a y vg00

vgcfgbackup vg00

lvlnboot -R

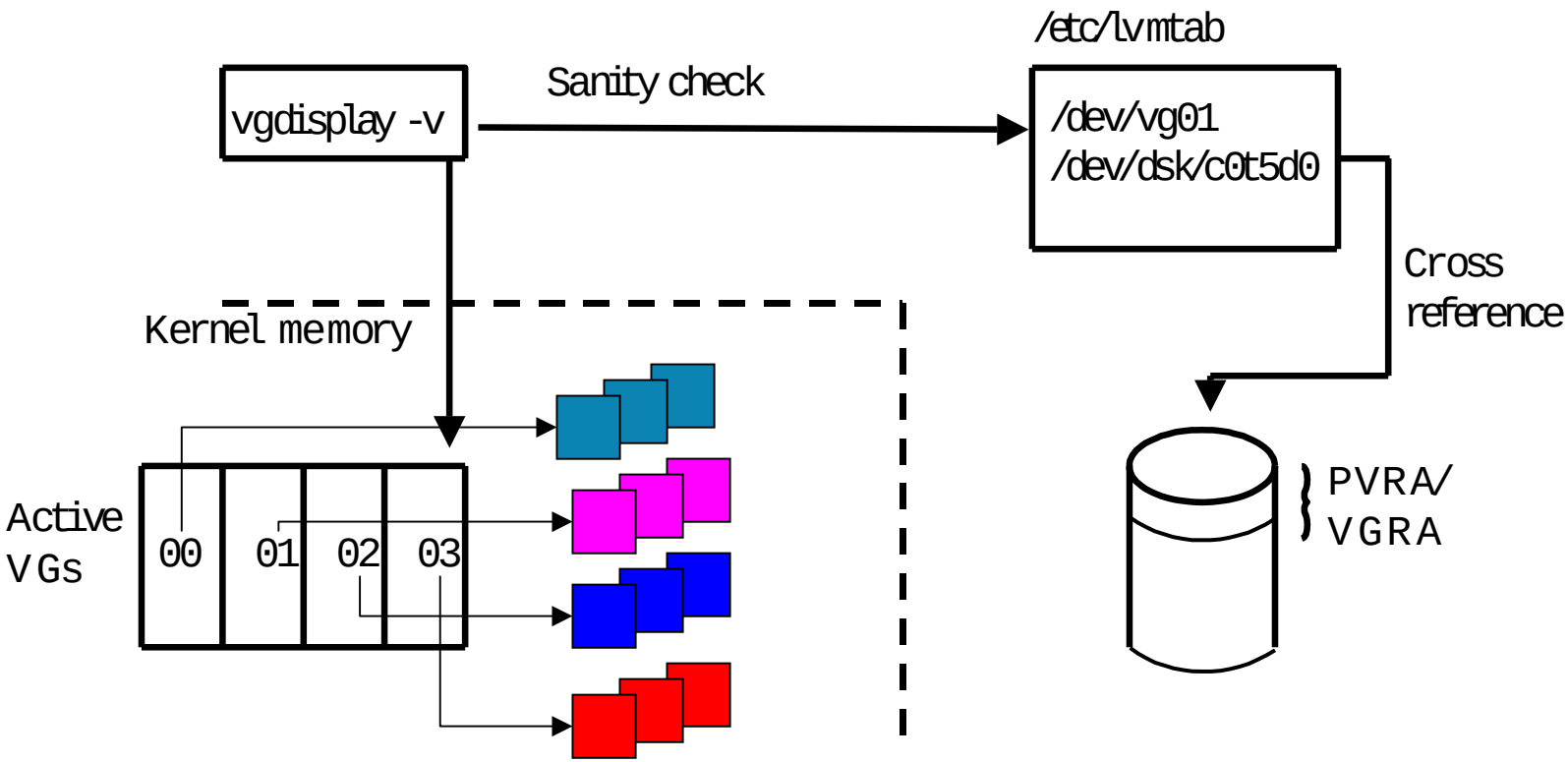
lvlnboot -v

Recovering Corrupt LVM Information

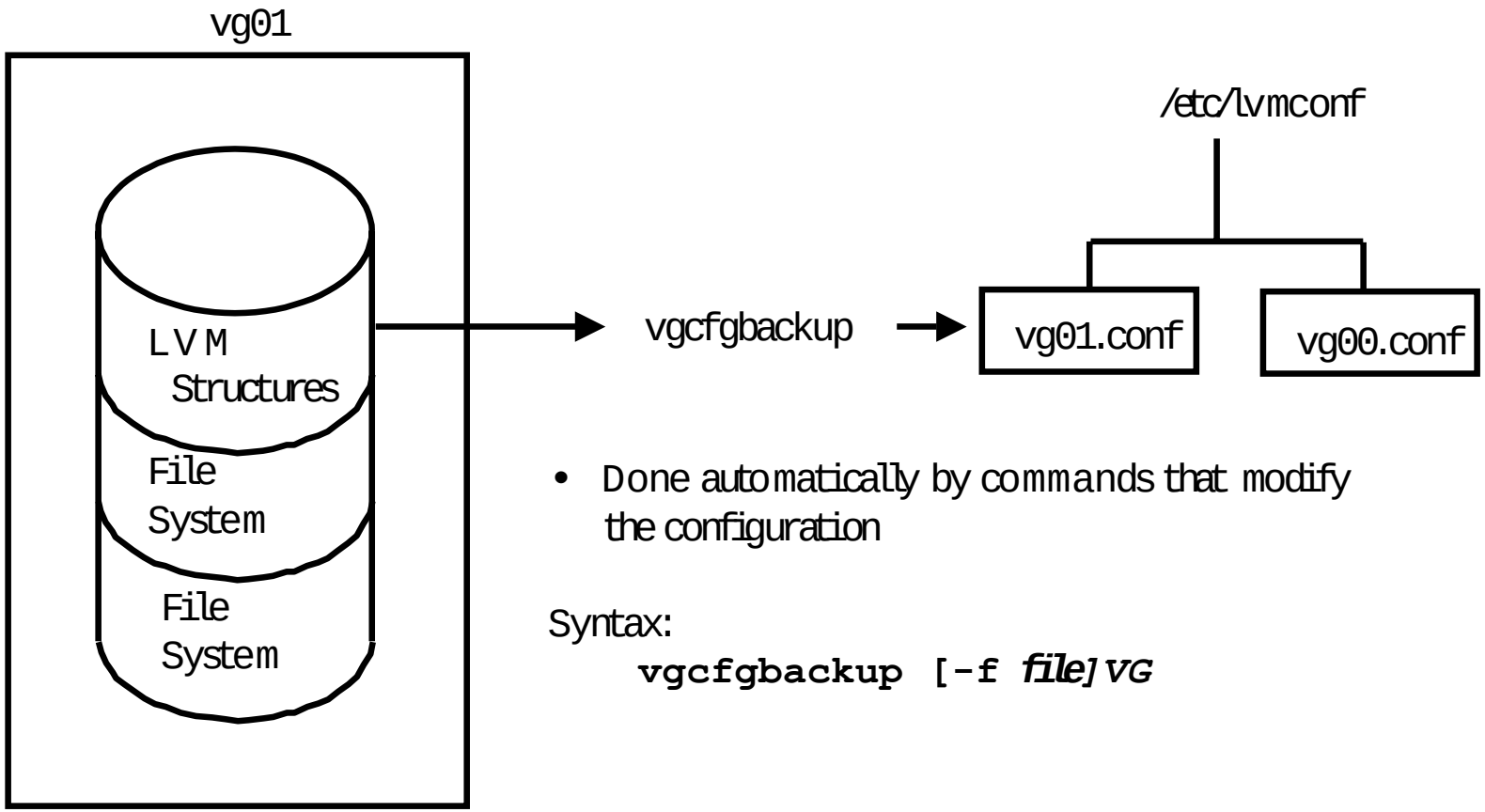
HP WORLD 2003
Solutions and Technology Conference & Expo

Valuable Data?

- The various lvm data structures are crucial to the continued availability of logical volumes



Backing up LVM Structures

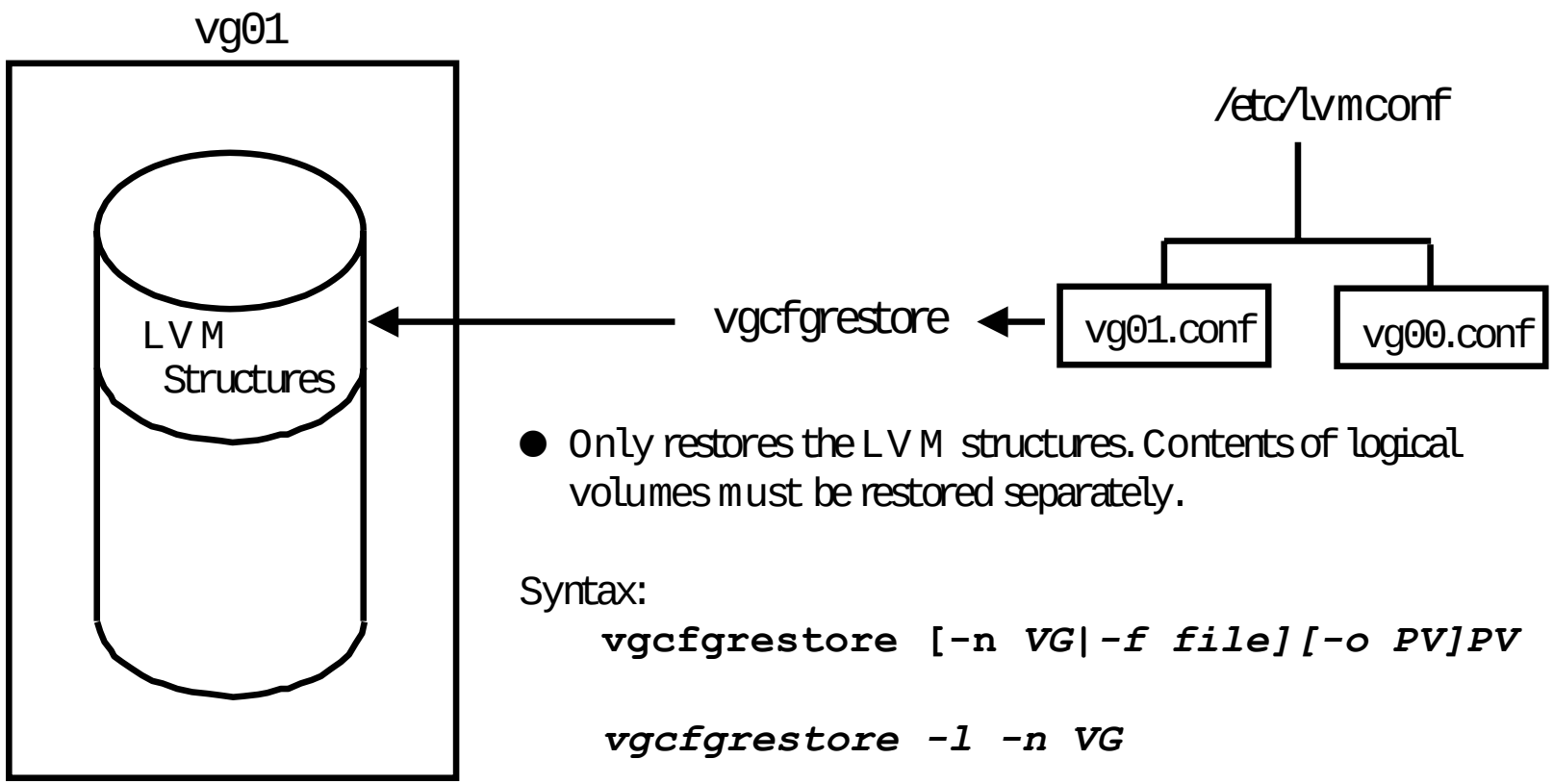


- Done automatically by commands that modify the configuration

Syntax:

`vgcfgbackup [-f` *`file`*`] VG`

Recovering LVM Structures

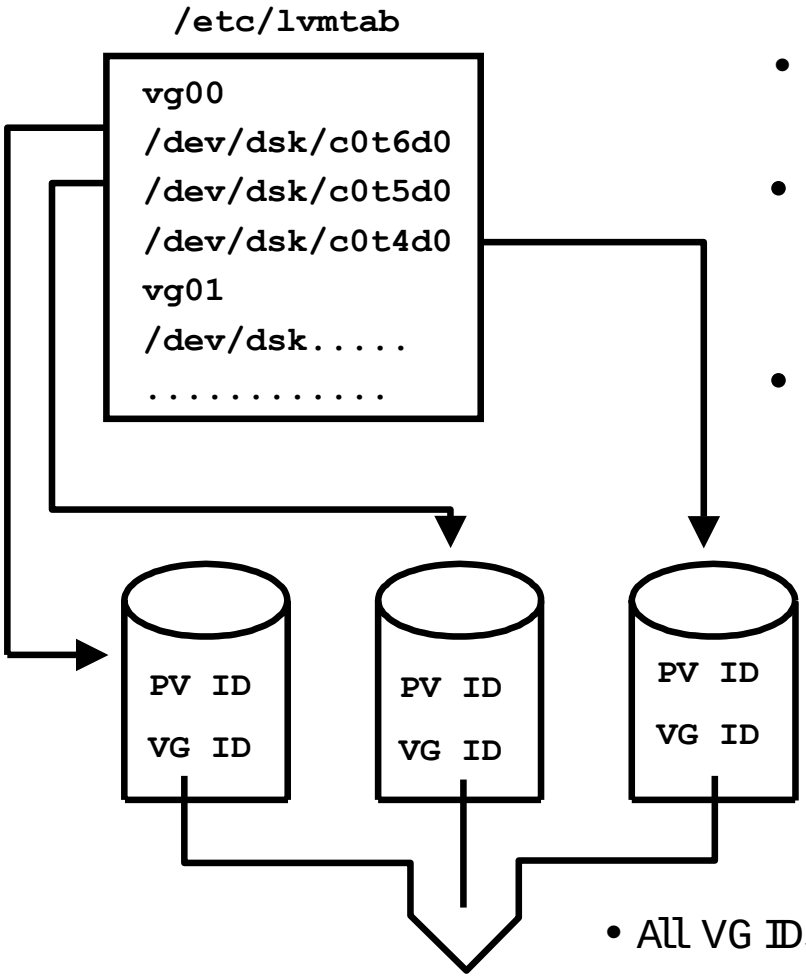


- Only restores the LVM structures. Contents of logical volumes must be restored separately.

Syntax:

```
vgcfgrestore [-n VG|-f file][-o PV]PV

vgcfgrestore -l -n VG
```

LVM Control File



- Not ascii data
  - can use strings (1) to read the ascii part
- Used primarily at boot up but also used to sanity check commands
- Can be rebuilt if lost

- All VG IDs the same - all disks in same vg

Recover or Repair
/etc/lvmtab

Syntax:

```
vgscan [-v][-p]
```

Semi-intelligent look at every disk to classify it

- imports those that are known to be on this system
- recommends the import of "new" disks
- rebuilds **/etc/lvmtab**

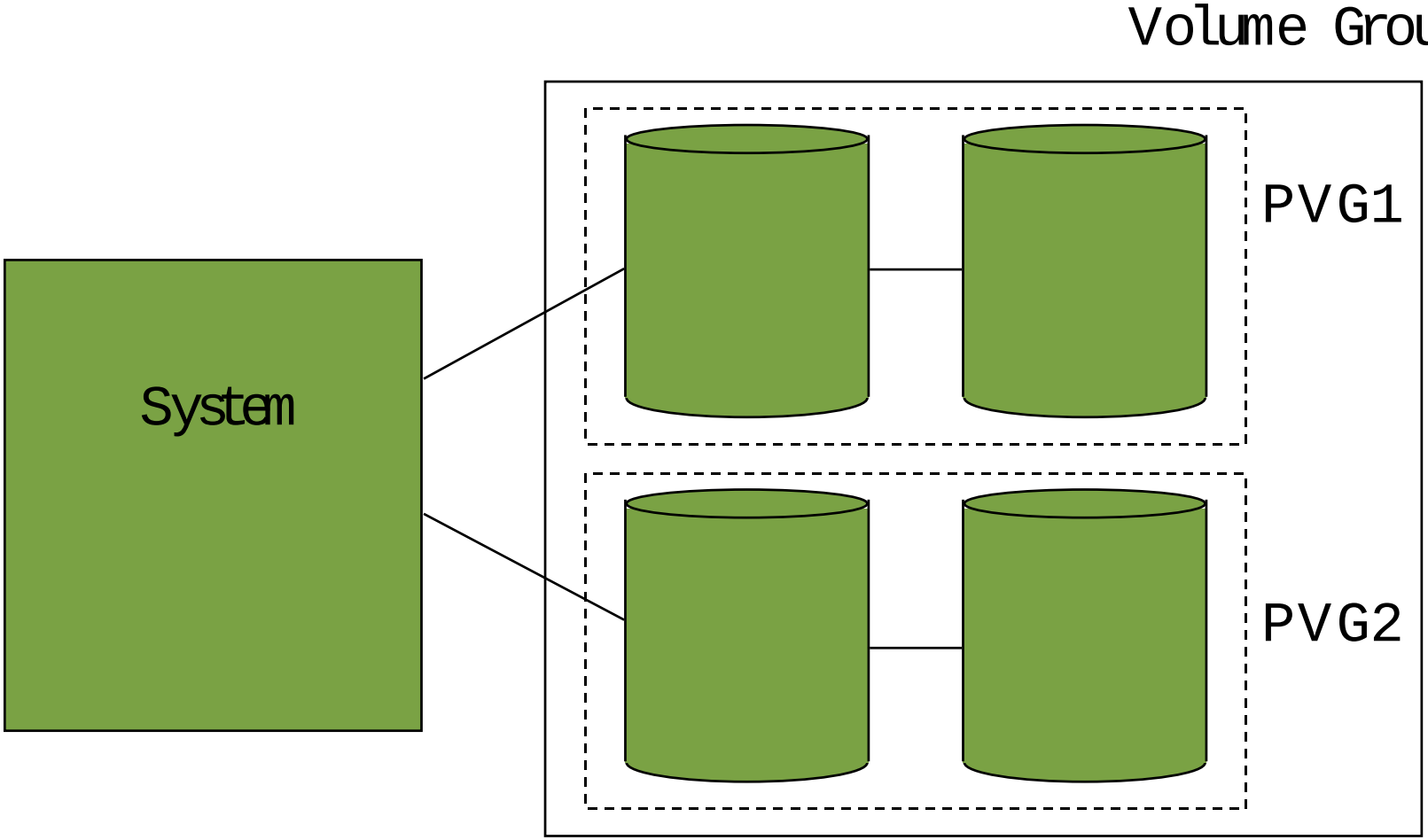
Ways to confuse vgscan

- redundant group files
- "old" LVM structures

LVM Performance Tips

HP WORLD 2003
Solutions and Technology Conference & Expo

Use PVGs to Offload Busy Controllers

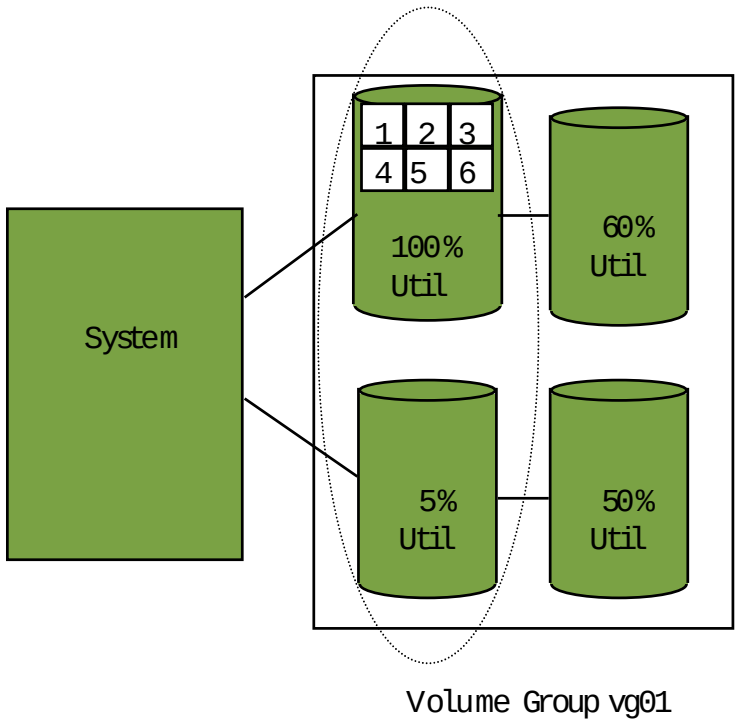


PVG Example

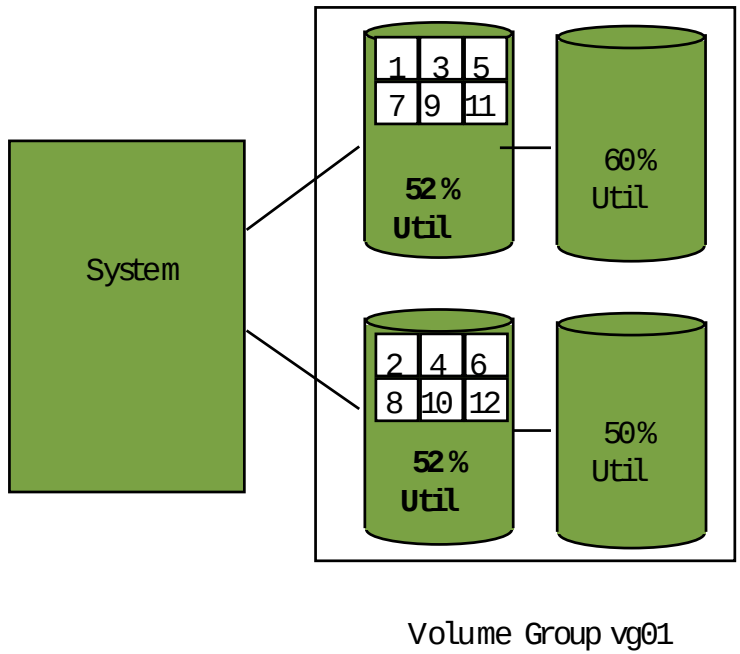
/etc/lvmpvg
VG /dev/vg01
PVG PVG1
/dev/dsk/c5t5d0
/dev/dsk/c5t5d1
PVG PVG2
/dev/dsk/c10t1d0
/dev/dsk/c10t1d1

```
# lvcreate -L 40 -m 1 -s g -n data1 vg01
# lvdisplay -v /dev/vg01/data1
   ....
--- Allocation                PVG-strict
```

Use Striping to Offload Busy Drives

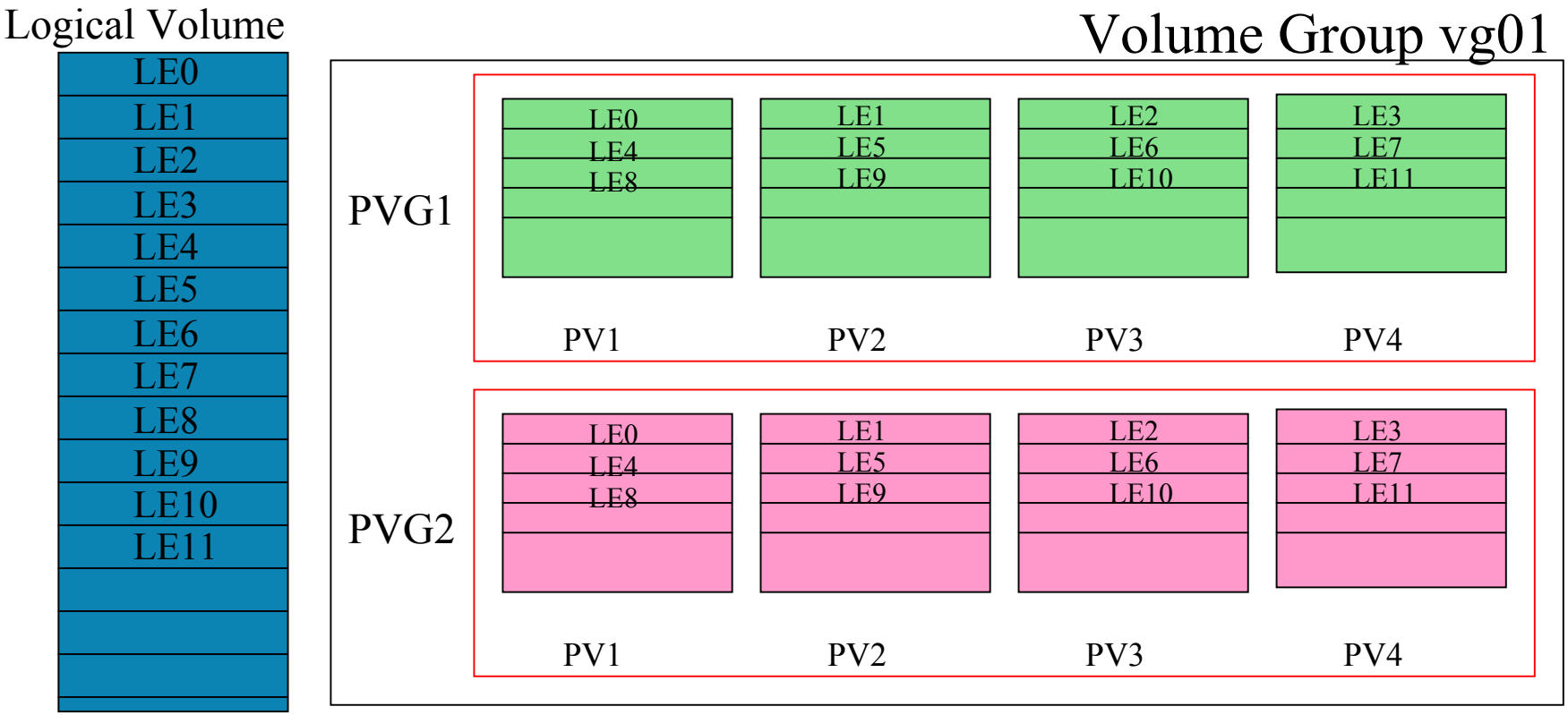


Without Striping



With Striping

Mirroring and Striping

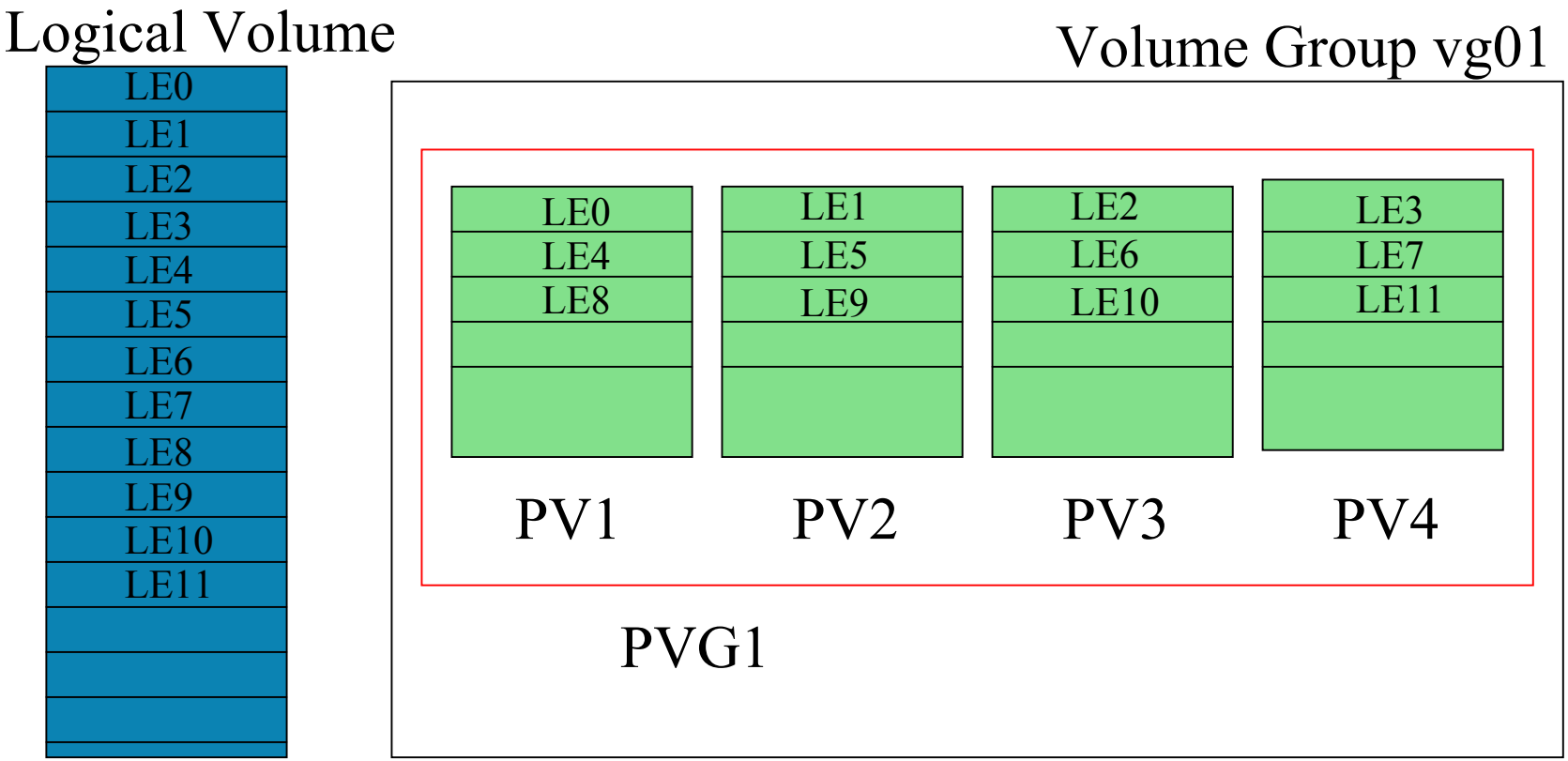


lvcreate -l 12 -m 1 -s g -D y /dev/vg01

Striping Example

```
# lvcreate -I 64 -i 4 -L 32 -n stripelv vg01
# lvdisplay -v /dev/vg01/stripelv
...
Schedule                 striped
LV Size (Mbytes)            32
Current LE                  8
Allocated PE                8
Stripes                     4
Stripe Size (Kbytes)        64
...
LE    PV1                PE1   Status 1
00000 /dev/dsk/c7t5d1    00003 current
00001 /dev/dsk/c5t5d0    00013 current
00002 /dev/dsk/c10t1d0   00013 current
00003 /dev/dsk/c9t1d1    00003 current
```

LVM Distributed Allocation



lvcreate -l 12 -s g -D y /dev/vg01

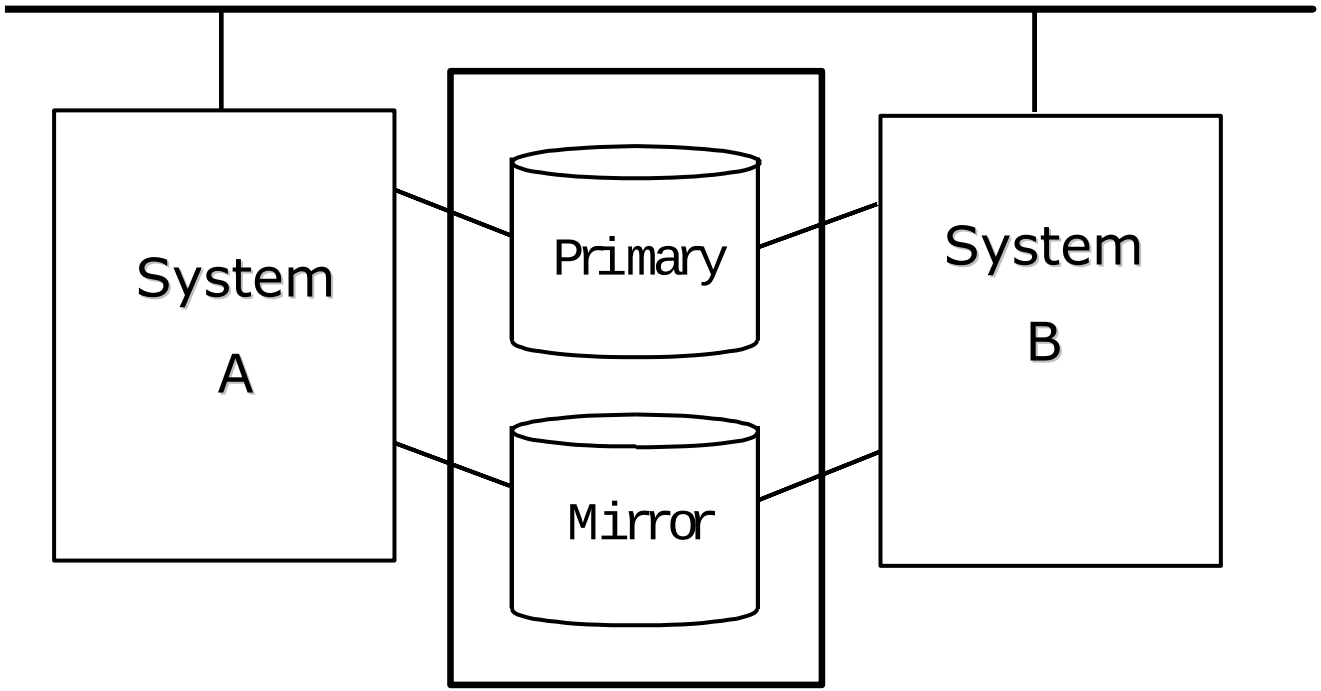
Distributed Allocation Example

```
/etc/lvmpvg.
# more /etc/lvmpvg
VG      /dev/vg01
PVG     PVG1
/dev/dsk/c5t5d0
/dev/dsk/c7t5d1
/dev/dsk/c10t1d0
/dev/dsk/c9t1d1
# lvcreate -l 12 -s g -D y -n distlv vg01
# lvdisplay -v /dev/vg01/distlv
...
Allocation                PVG-strict/distributed
...
 --- Logical extents ---
   LE    PV1                 PE1   Status 1
   00000 /dev/dsk/c5t5d0     00010 current
   00001 /dev/dsk/c7t5d1     00000 current
   00002 /dev/dsk/c10t1d0    00010 current
   00003 /dev/dsk/c9t1d1     00000 current
```

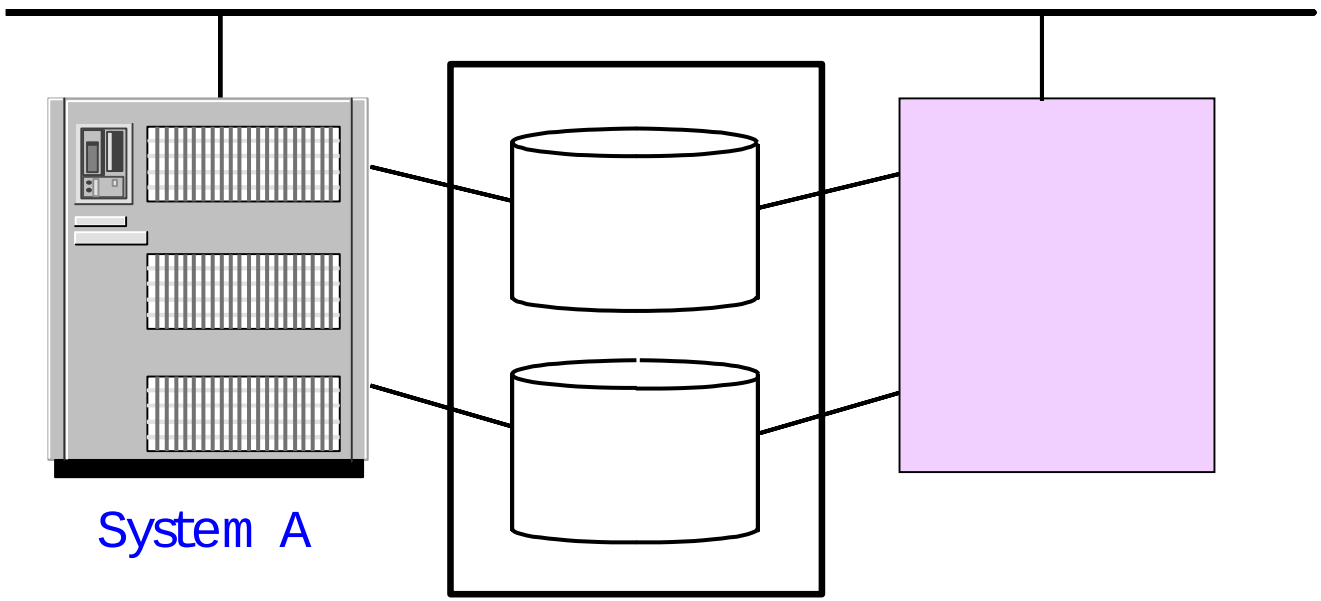
LVM in an MC/ServiceGuard Environment

HP WORLD 2003
Solutions and Technology Conference & Expo

MC/ServiceGuard Configuration



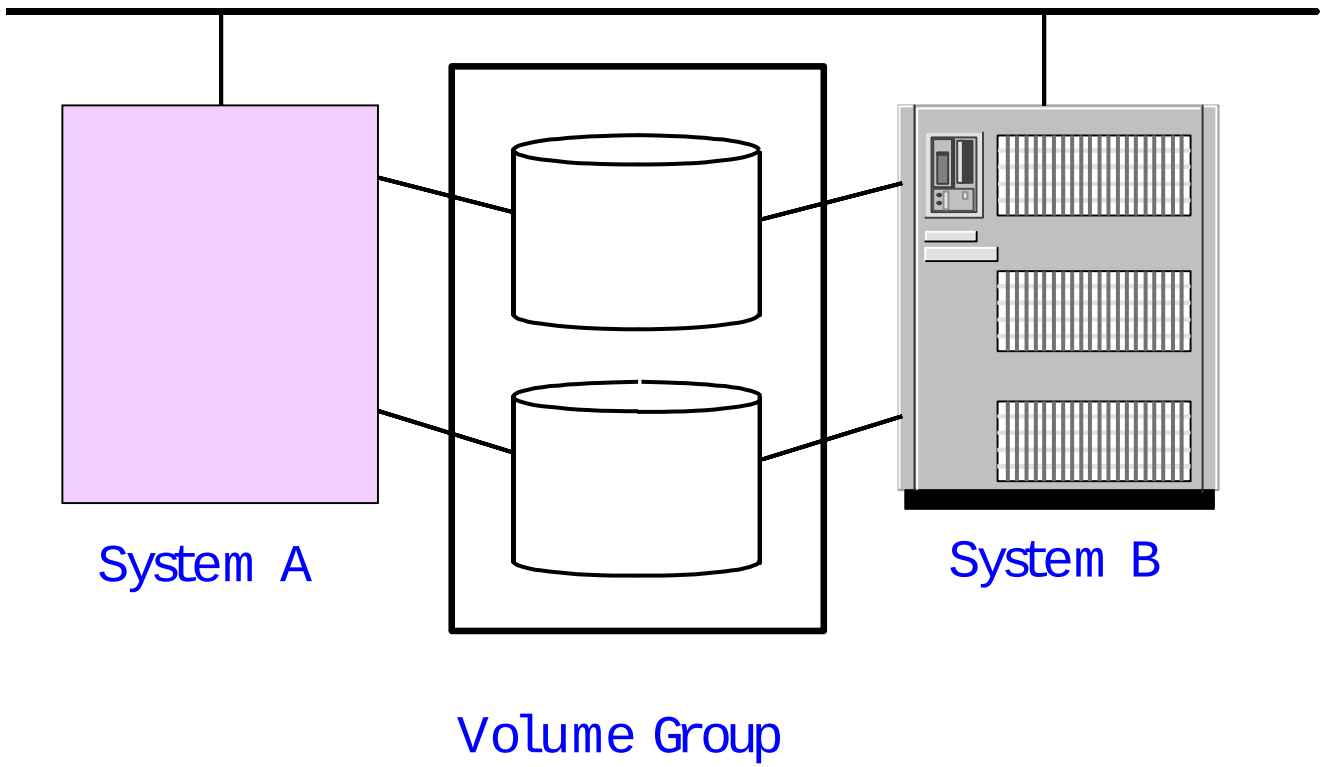
Step 1: Configure Volume Group on First System



Step 1 Details

```
pvcreate /dev/rdsk/c1t8d0
pvcreate /dev/dsk/c2t8d0
mkdir /dev/vg01
mknod /dev/vg01/group c 64 0x010000
vgcreate vg01 /dev/dsk/c1t8d0 /dev/dsk/c2t8d0
lvcreate -m 1 -n mydata -L 300 vg01
```

Step 2: Import Volume Group to Second System



Step 2 Details

On System A:
vgexport -p -m /tmp/vg01.map vg01 # create a map file
rcp /tmp/vg01.map systemb:/tmp/vg01.map # copy to other system

On System B:
mkdir /dev/vg01
mknod /dev/vg01/group c 64 0x010000
vgimport vg01 /dev/dsk/c1t8d0 /dev/dsk/c2t8d0
vgchange -a y vg01
vgcfgbackup vg01
vgchange -a n vg01

Caution: The device files may not be the same on both systems!

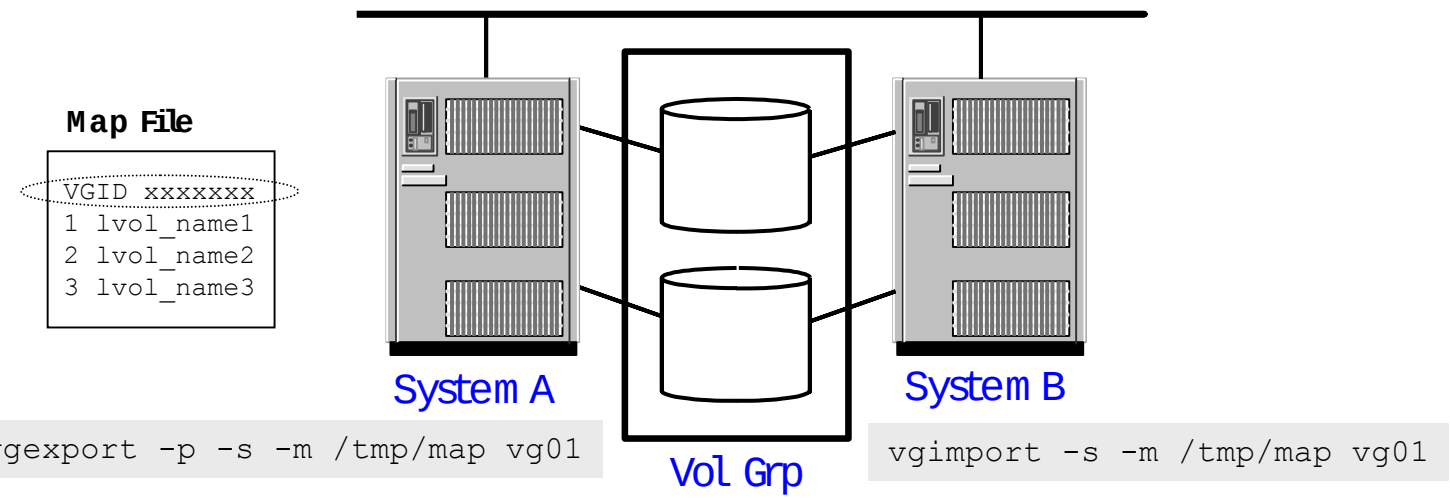
VG Export and Import -s Option

vgexport -p -s -m <*mapfilename*> <*vg_name*>

create *mapfile* without removing the *VG*,
and save the VGID for the vgimport command

vgimport -s -m <*mapfilename*> <*vg_name*>

scan for disks that have the same *VG ID* as in the *mapfile*



LVM Definitions-Both Nodes

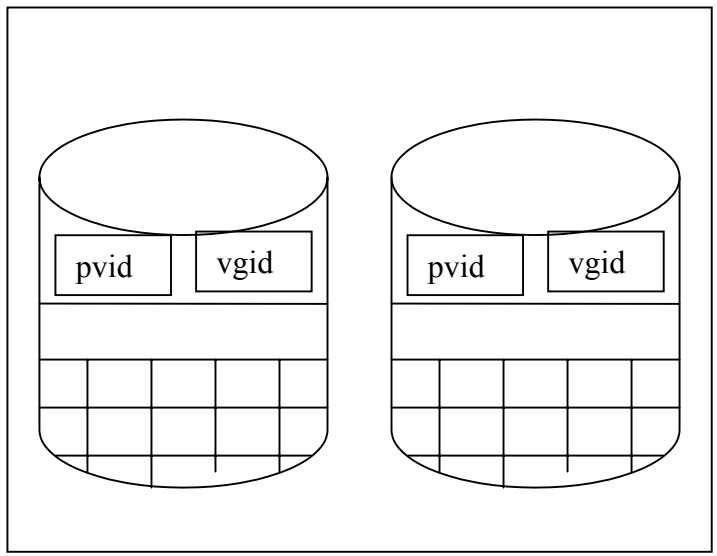
Node 1

/dev/vg01

```
group        64 0x010000
lvol1        64 0x010001
rlvol1       64 0x010001
databaselv   64 0x010002
rdatabaselv  64 0x010002
```

/etc/lvmtab

```
/dev/vg01
/dev/dsk/c5t5d0
/dev/dsk/c7t5d0
/dev/dsk/c7t5d1
/dev/dsk/c5t5d1
```



Node 2

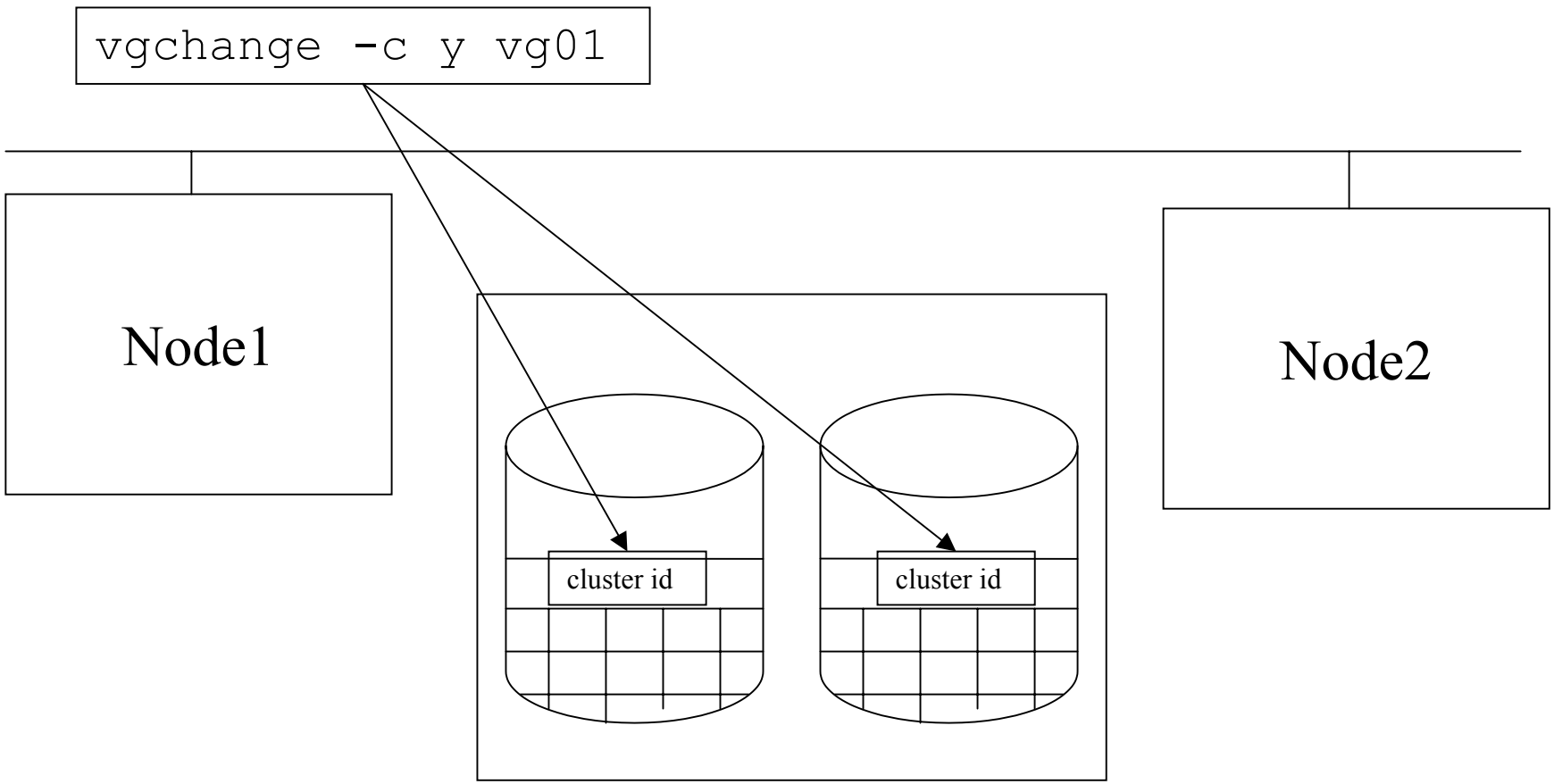
/dev/vg01

```
group        64 0x010000
lvol1        64 0x010001
rlvol1       64 0x010001
databaselv   64 0x010002
rdatabaselv  64 0x010002
```

/etc/lvmtab

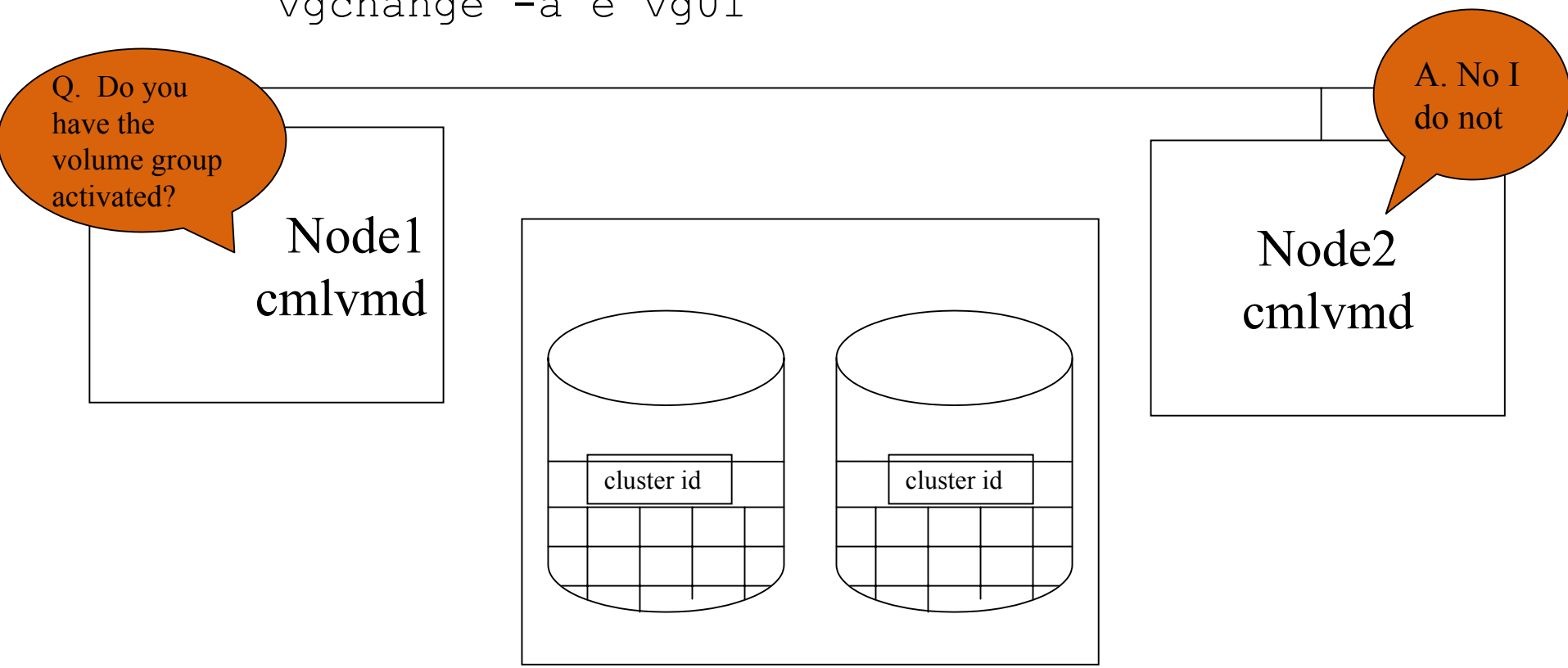
```
/dev/vg01
/dev/dsk/c5t5d0
/dev/dsk/c7t5d0
/dev/dsk/c7t5d1
/dev/dsk/c5t5d1
```

Cluster Volume Group



Volume Group Exclusive Activation

```
vgchange -a e vg01
```



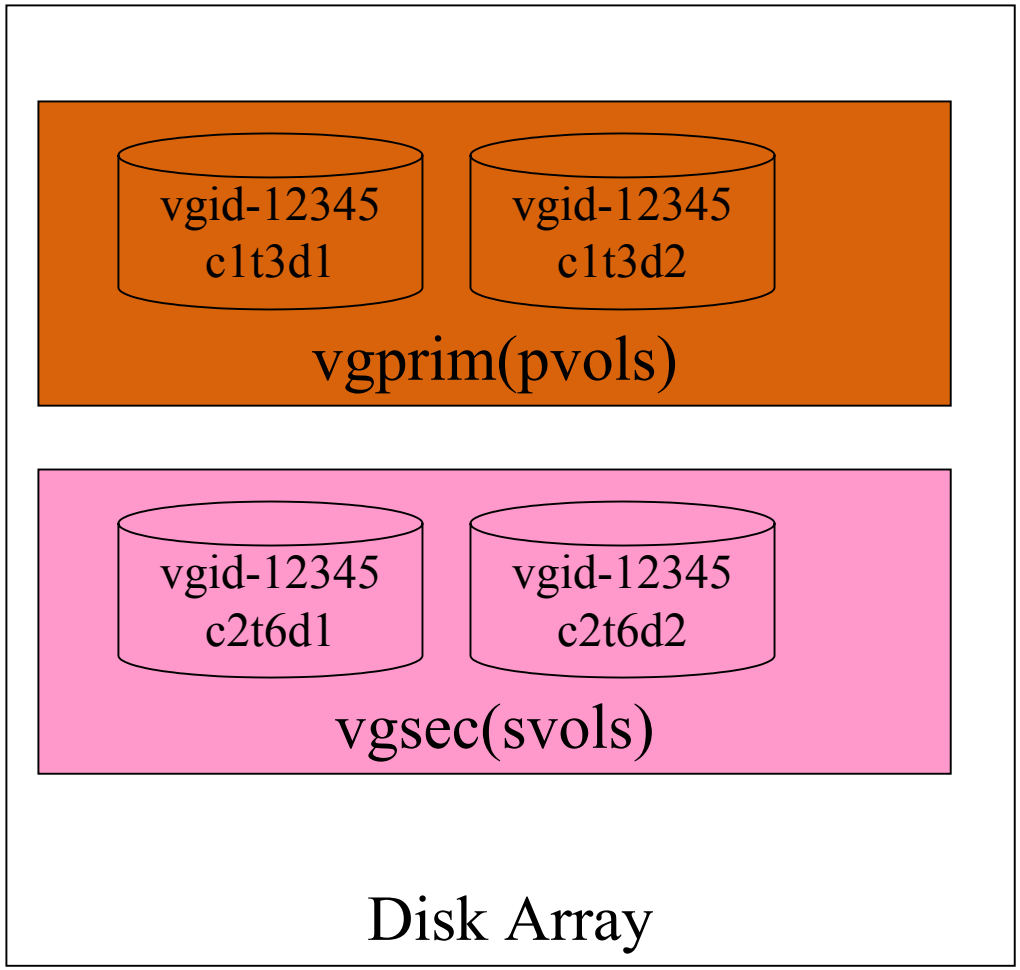
Marking Volume Groups as MC/ServiceGuard Volume Groups

Marking Volume Group for MC/ServiceGuard	**vgchange -c y VGName**
Marking Volume Group as non-MC/ServiceGuard	**vgchange -c n VGName**
Standard Volume Group Activation	**vgchange -a y VGName**
Exclusive Volume Group Activation	**vgchange -a e VGName**

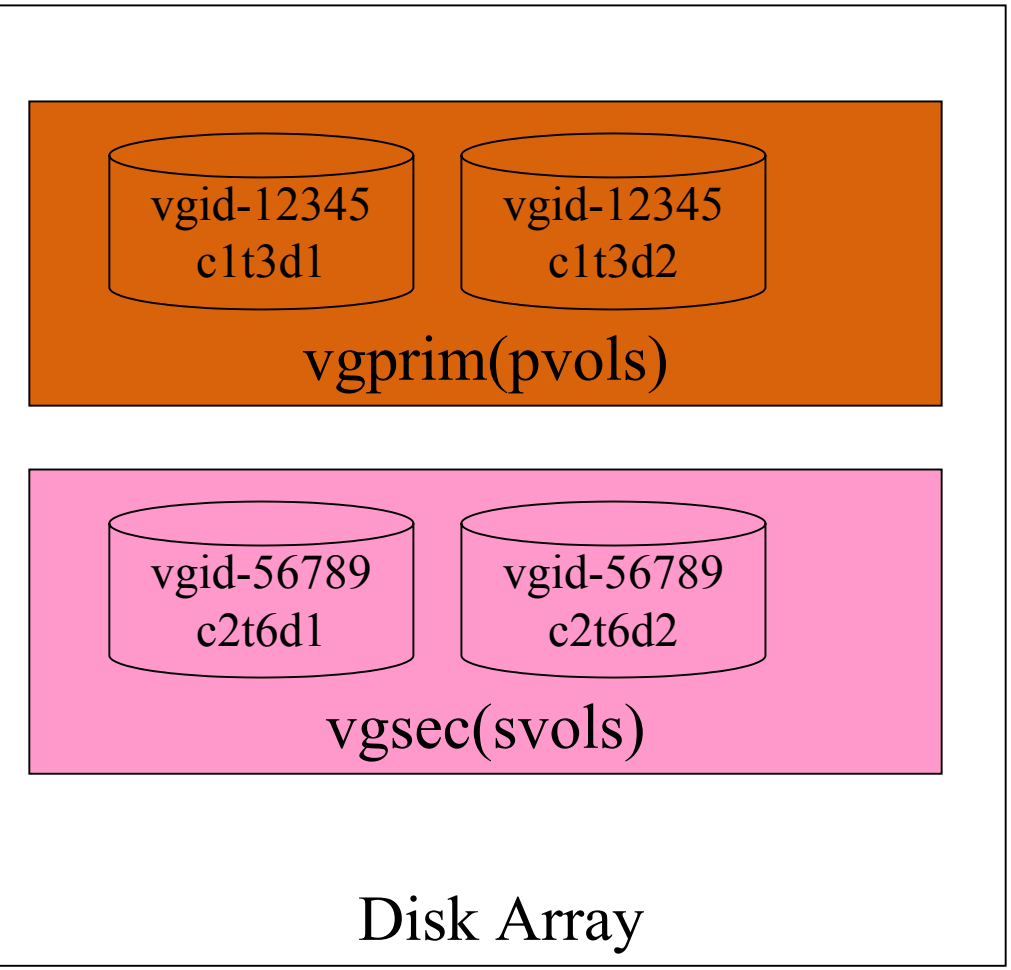
Disk Array Data Replication Issues

HP WORLD 2003
Solutions and Technology Conference & Expo

Problem-Duplicate vgid

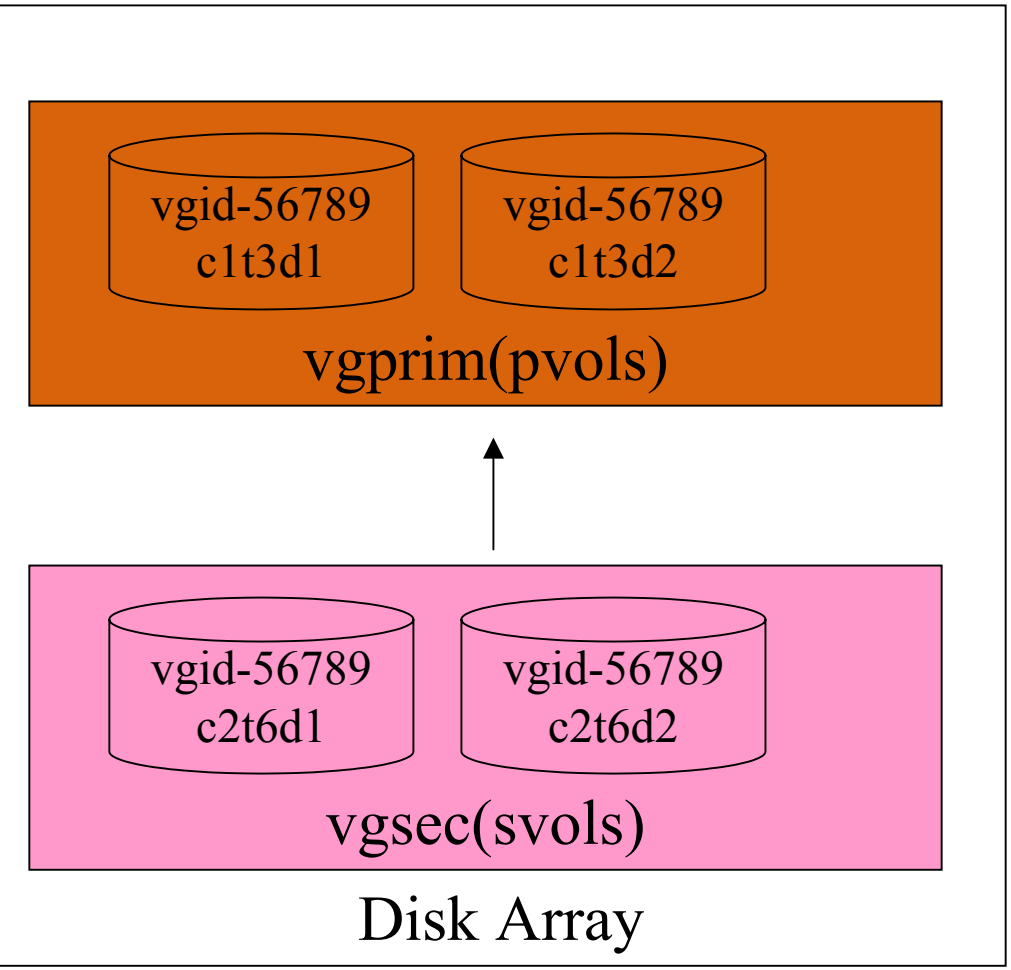


Solution-vgchgid




```
vgchgid /dev/rdsk/c2t6d1
/dev/rdsk/c2t6d2

mkdir /dev/vgsec

mknod /dev/vgsec/group c 64
0x020000

vgexport -pm mapfile vgprim

vgimport -m mapfile vgsec
/dev/dsk/c2t6d1
/dev/dsk/c2t6d2

vgchange -a y vgsec

vgcfgbackup vgsec
```

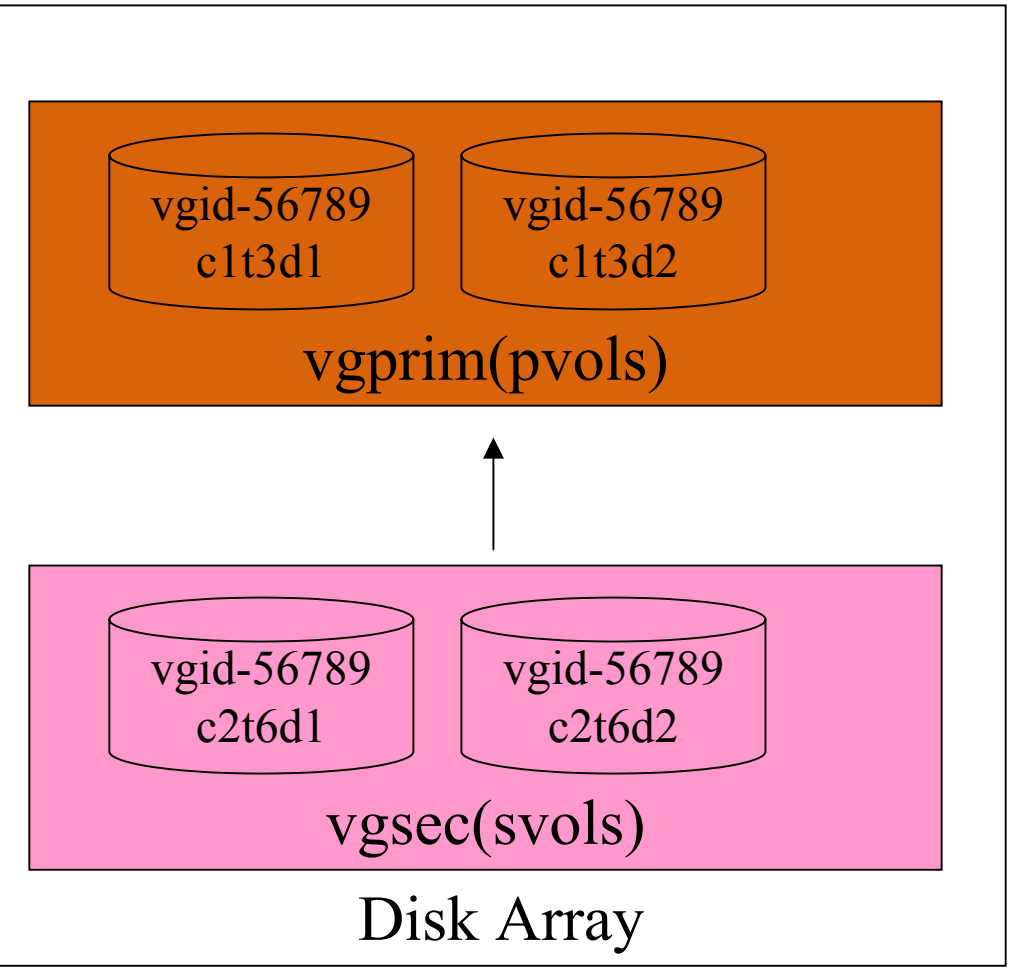
Problem after a reverse sync



/etc/lvmtab

/dev/vgpriv (*vgid:12345)*
/dev/dsk/c1t3d1
/dev/dsk/c1t3d2

Solution



- export volume group
- reimport volume group

/etc/lvmtab

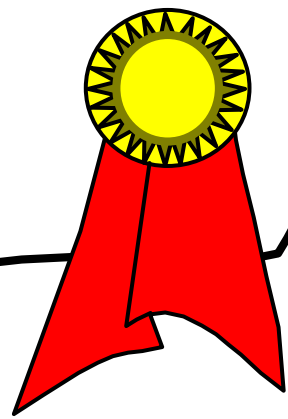
/dev/vgpriv (*vgid:56789)*

/dev/dsk/c1t3d1

/dev/dsk/c1t3d2

MVM Degree (Masters of Volume Managemet)

Please complete the Session Survey



Interex, Encompass and HP bring you a powerful new HP World.