LVM & JFS Explained

Chris Wong
Cerius Technology Group
cwong@cerius.com
(425) 260-6501

InterWorks 2000 - Presentation #114

Should you be here?

- Stay if:
  - Mostly use SAM
  - New to HP-UX
  - New to LVM
  - Want LVM refresher
  - New to JFS or OnLine JFS
- Attend another session if:
  - Attended the HP LVM & Mirroring Class
  - Mostly use the command line and understand what the commands are doing
  - Attended this session at previous Interworks or HPWorld

Updates since printing

- **NEW SLIDE**
- **UPDATED**

Agenda

- Hardware
- Device Files
  - Major/Minor #s
- LVM
  - Physical Volume
  - Extents
  - Volume Group
  - Logical Volume
- Striped LV
- JFS
- Other LVM Tasks
- Root disk
- Mirroring
- Hot Spare
- OnLine JFS
- AutoRAID
- XP256
- Appendix A: ACLs
- Appendix B: Oracle
- Appendix C: MC/SG

Determine your system hardware

- What disks do you have and what kind are they?
- Hardware paths

```
#
# ioscan -fC disk
Class     I  H/W Path  Driver      S/W State H/W Type  Description
==========================================================================
disk      0  0.5.0     sdisk       CLAIMED   DEVICE    SEAGATE ST32550W
disk      1  0.8.0     sdisk       CLAIMED   DEVICE    SEAGATE ST34371W
disk      3  0.9.0     sdisk       CLAIMED   DEVICE    SEAGATE ST34371W
disk      4  0.10.0    sdisk       CLAIMED   DEVICE    SEAGATE ST32550W
disk      2  16/5.2.0  sdisk       CLAIMED   DEVICE    TOSHIBA CD-ROM XM-
#
```

Show device files associated with each device

```
#
# ioscan -fnC disk
Class     I  H/W Path  Driver      S/W State H/W Type  Description
===========================================================================
disk      0  0.5.0     sdisk       CLAIMED   DEVICE    SEAGATE ST32550W
                       /dev/dsk/c0t5d0   /dev/rdsk/c0t5d0
disk      1  0.8.0     sdisk       CLAIMED   DEVICE    SEAGATE ST34371W
                       /dev/dsk/c0t8d0   /dev/rdsk/c0t8d0
disk      3  0.9.0     sdisk       CLAIMED   DEVICE    SEAGATE ST34371W
                       /dev/dsk/c0t9d0   /dev/rdsk/c0t9d0
disk      4  0.10.0    sdisk       CLAIMED   DEVICE    SEAGATE ST32550W
                       /dev/dsk/c0t10d0   /dev/rdsk/c0t10d0
disk      2  16/5.2.0  sdisk       CLAIMED   DEVICE    TOSHIBA CD-ROM XM-
                       /dev/dsk/c1t2d0   /dev/rdsk/c1t2d0
#
```

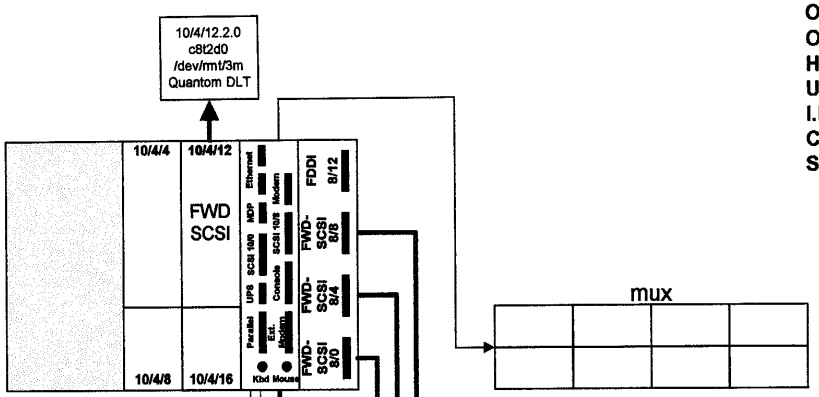
Information on physical disk

```
#
# diskinfo -b /dev/rdsk/c0t5d0
2082636
#
# diskinfo -v /dev/rdsk/c0t5d0
SCSI describe of /dev/rdsk/c0t5d0:
               vendor: SEAGATE
           product id: ST32550W
                 type: direct access
                 size: 2082636 Kbytes
    bytes per sector: 512
            rev level: HP09
     blocks per disk: 4165272
         ISO version: 0
        ECMA version: 0
        ANSI version: 2
     removable media: no
     response format: 2
    (Additional inquiry bytes: (32)31 (33)35 (
# )0 (40)0 (41)0 (42)0 (43)0 (44)0 (45)0 (46)
```

V I S I O

Host A

Model:	9000/K460
Memory:	1024MB
CPUs:	2 (180MHz)
OLTP Perf.:	?
OS Ver:	10.20
HW ID:	
User License:	Unlimited
I.P. Add:	
CPU Serial #:	
System Handle:	



mux

10/0.3.0 c2t3d0 A3352A 4.35GB vg01 lvol1:/opt	10/0.4.0 c2t4d0 A3352A 4.35GB vg01 lvol1:/opt		
10/0.5.0 c2t5d0 A3352A 4.35GB vg00 lvol1:/stand lvol2:swap lvol3:/ lvol4:/home lvol5:/tmp lvol6:/usr lvol7:/var lvol8:/var/mail lvol9:/var/spool lvol10:/var/tmp	10/0.6.0 c2t6d0 A3352A 4.35GB vg00 lvol1:/stand lvol2:swap lvol3:/ lvol4:/home lvol5:/tmp lvol6:/usr lvol7:/var lvol8:/var/mail lvol9:/var/spool lvol10:/var/tmp		
10/12/5.2.0 c4t2d0 cd-rom	10/12/5.0.0 c4t0d0 /dev/rmt/0hc DDS2		

High Availability Storage System			
A	8/8.1.0 c1t1d0 A5238A 9.1GB vg02 lvol1:/disk1	8/8.2.0 c1t2d0 A5238A 9.1GB vg02 lvol2:/disk2	8/8.3.0 c1t3d0 A5238A 9.1GB vg03 lvol1:/dtasave lvol2:/dtarea 280mb	8/8.4.0 c1t4d0 A5238A 9.1GB Unused 8680mb
B	8/4.1.0 c0t1d0 A5238A 9.1GB vg02 lvol3:/disk3	8/4.2.0 c0t2d0 A5238A 9.1GB vg02 lvol4:/disk4	8/4.3.0 c0t3d0 A5238A 9.1GB vg03 lvol3:/jutyme lvol4:/dscpol 2180mb	8/4.4.0 c0t4d0 A5238A 9.1GB Unused 8680mb

High Availability Storage System				
A	8/0.1.0 c9t1d0 A5238A 9.1GB vg02 lvol5:/disk5	8/0.2.0 c9t2d0 A5238A 9.1GB vg02 lvol6:/disk6	8/0.3.0 c9t3d0 A5238A 9.1GB vg03 lvol5:/dtasave lvol6:/dtarea 280mb	Open
B	10/8.1.0 c3t1d0 A5238A 9.1GB vg02 lvol7:/disk7	10/8.2.0 c3t2d0 A5238A 9.1GB vg02 lvol11:/disk11 lvol12:/disk12	10/8.3.0 c3t3d0 A5238A 9.1GB vg03 lvol7:/accept lvol8:/var/opt/uni 4156mb	Open

SCSI Priority

- 7 - Highest
- 6
- 5
- 4
- 3
- 2
- 1
- 0
- 15
- 14
- 13
- 12
- 11
- 10
- 9
- 8 - Lowest

Device Files

```
#
# ll /dev/rdsk
total 0
crw-r-----   1 bin        sys         188 0x00a000 Apr  9 14:00 c0t10d0
crw-r-----   1 root       sys         188 0x005000 Apr 20 20:02 c0t5d0
crw-r-----   1 root       sys         188 0x008000 Apr 20 16:06 c0t8d0
crw-r-----   1 bin        sys         188 0x009000 Apr 20 20:01 c0t9d0
crw-r-----   1 root       sys         188 0x012000 Jun  9  1996 c1t2d0
#
# ioscan -H 0.5.0
H/W Path  Class          Description
====================================
0.5.0           disk     SEAGATE ST32550W
#
# lssf /dev/rdsk/c0t5d0
sdisk card instance 0 SCSI target 5 SCSI LUN 0 section 0 at address 0.5.0 /dev
dsk/c0t5d0
#
#
```

Device file creation

- All hardware is probed as one of the many system initialization tasks during system boot
- Each auto-configurable device must be bound to a driver
- Device files are automatically created during the reboot process (10x+)
- Instance # are assigned in the order in which cards are bound to drivers

Device Files - DISKS

```
#
# ll /dev/dsk
total 0
brw-r-----   1 bin        sys          31 0x00a000 Apr  2 09:42 c0t10d0
brw-r-----   1 root       sys          31 0x005000 Jun 10  1996 c0t5d0
brw-r-----   1 root       sys          31 0x008000 Apr  9 17:30 c0t8d0
brw-r-----   1 bin        sys          31 0x009000 Apr  2 11:25 c0t9d0
brw-r-----   1 root       sys          31 0x012000 Jun  9  1996 c1t2d0
#
```

- Block (device file class) - b
- Raw or character (device file class) - c

```
#
# ll /dev/rdsk
total 0
crw-r-----   1 bin        sys         188 0x00a000 Apr  9 14:00 c0t10d0
crw-r-----   1 root       sys         188 0x005000 Apr 20 20:02 c0t5d0
crw-r-----   1 root       sys         188 0x008000 Apr 20 16:06 c0t8d0
crw-r-----   1 bin        sys         188 0x009000 Apr 20 20:01 c0t9d0
crw-r-----   1 root       sys         188 0x012000 Jun  9  1996 c1t2d0
#
```

Block vs. character devices

- BLOCK
  - Transfer data using system buffers
  - Want to treat the device as a file system
    - Tape drives
    - Disks
- CHARACTER
  - Transfer data one character at a time
  - Buffering controlled by applications
  - Raw
    - Printers
    - Terminals
    - Disks

He's quite a character, with a raw sense of humor.

Device files - DISKS

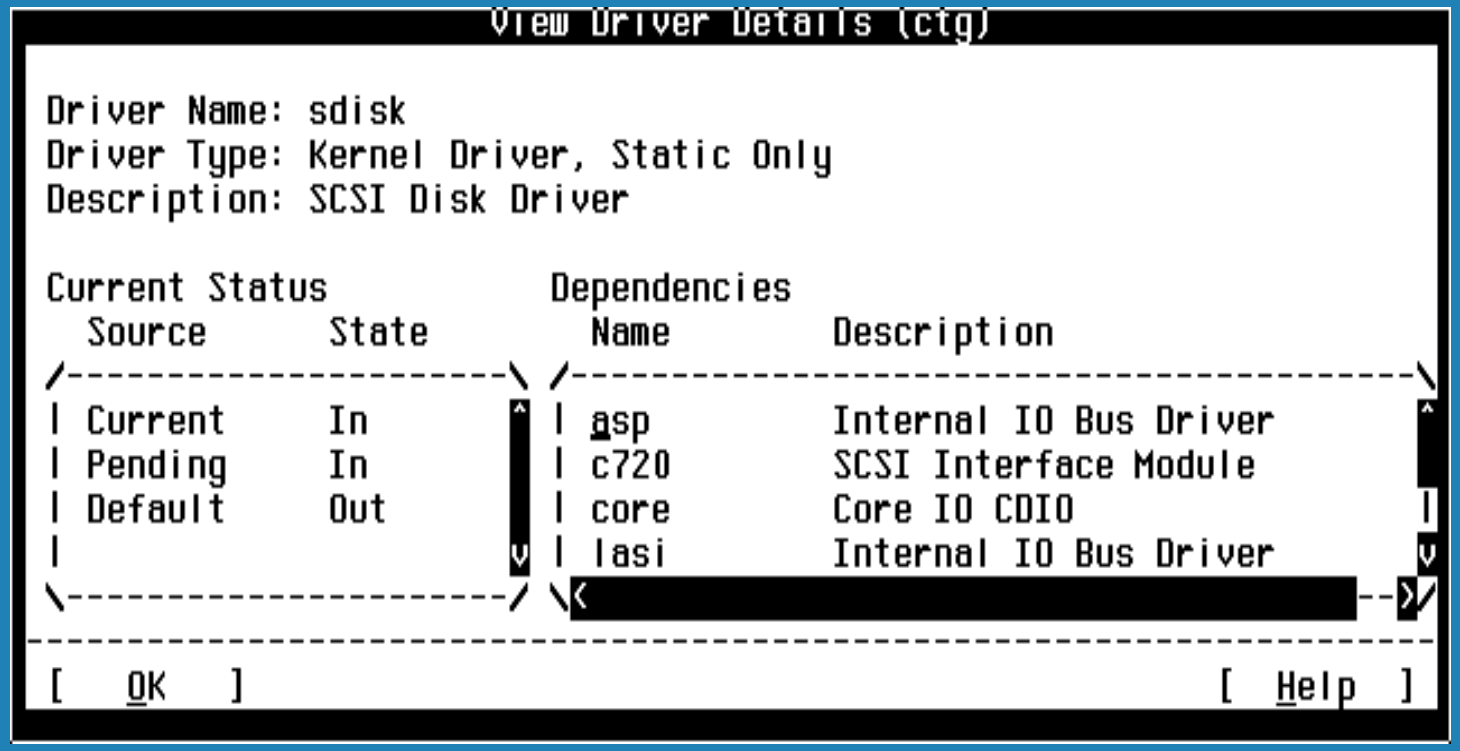
```
 1 root       sys          31 0x005000 Jun 10  1996 /dev/dsk/c0t5d0
 1 root       sys         188 0x005000 Apr 20 20:02 /dev/rdsk/c0t5d0
                           ^        ^
                           |        |
                         major    minor

Link count
Owner
Group
Major number
Minor number
Date & Time (modify)
```

Major numbers - device files

```
 1 root       sys          31 0x005000 Jun 10  1996 /dev/dsk/c0t5d0
 1 root       sys         188 0x005000 Apr 20 20:02 /dev/rdsk/c0t5d0
#
# lsdev -c 188
    Character     Block       Driver          Class
      188           31        sdisk           disk
#
```

- lsdev command lists the major device numbers
- Pointer to the kernel driver to use to communicate with the device
- Kernel driver controls I/O for a device

sdisk driver



Minor number - device files

- Physical location & optional characteristics
  - *Card Instance number*
  - *Target number*
  - Port number
  - HP-IB address
  - Device options
- 0xCCTPHD
- 0x005000

Hard link - 2 names for the same file "User Friendly"

```
#
# ll /dev/dsk
total 0
brw-r-----   1 bin        sys          31 0x005000 May 19 11:28 c0t5d0
brw-r-----   1 bin        sys          31 0x008000 May 19 11:28 c0t8d0
brw-r-----   2 bin        sys          31 0x012000 May 19 11:28 c1t2d0
#
# ln /dev/dsk/c1t2d0 /dev/dsk/cdrom
# ln /dev/rdsk/c1t2d0 /dev/rdsk/cdrom
# ll /dev/dsk
total 0
brw-r-----   1 bin        sys          31 0x005000 May 19 11:28 c0t5d0
brw-r-----   1 bin        sys          31 0x008000 May 19 11:28 c0t8d0
brw-r-----   3 bin        sys          31 0x012000 May 19 11:28 c1t2d0
brw-r-----   3 bin        sys          31 0x012000 May 19 11:28 cdrom
# mount /dev/dsk/cdrom /cdrom
#
```

- Great for cdrom

LVM History
Logical Volume Management

- Comes with HP-UX as of 9x on servers
- Comes with HP-UX as of 10x on workstations
- LVM originally designed by Open Software Foundation (OSF). Ported to HP-UX
- Available Veritas LVM
- LVM is a disk management subsystem

LVM objects

- PV (physical volume)
  - Individual Disk drive
- VG (volume group)
  - One or more PVs that creates one large logical disk
- LV (logical volume)
  - Section of VG used for file systems, swap, raw or dump

Physical Volumes in a VG

- One or more physical volumes (PV) create a volume group
- Entire disk must be used
- Minimum: 1 PV
- Maximum:255 (default is 16]

-p max_pv	Set the maximum number of physical volumes that the volume group is allowed to contain. The default value for max_pv is 16. The maximum number of physical volumes can be a value in the range 1 to 255.

Maximum Volume Groups

- Kernel parameter
- Default is 10

```
Parameter Name: maxvgs
Description:    Max Number of Volume Groups
```

```
The command used to create volume groups, /sbin/vgcreate, has failed.
The stderr output from the command is shown below.  The volume group
has not been created.


vgcreate: Cannot open the control file "/dev/vg10/group":
No such device

                              [[   OK   ]]
```

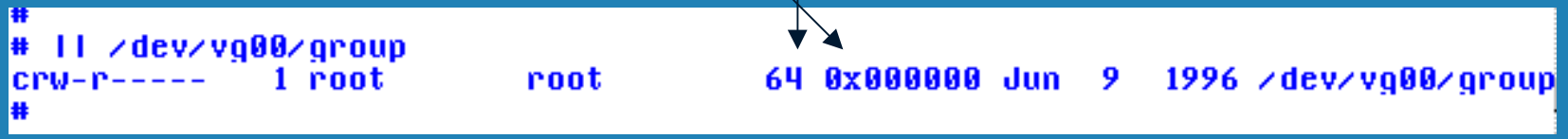
What makes a VG?

- Under the /dev directory start the structure for a VG
- Use vg00, vg01, etc..
- Or use names that make sense (oracle, home)
- MC/SG - Special Considerations (future slides)

```
#
# ll -d /dev/vg*
drwxr-xr-x   2 root       root          1024 Apr 20 20:01 /dev/vg00
#
```

Group file

- In the directory for the VG is a file (type character) called group
- Must be called group
- Minor number
- Major number must be 64



Major Number - VG group file Why 64?

- The major number must be 64 since this signifies the driver for LVM
- rdsk dsk

```
#
# ll /dev/*/group
crw-r-----   1 root       root        64 0x000000 Jun  9  1996 /dev/vg00/group
#
```

Minor numbers - VG group file

- Minor number must be unique
- 0xNN0000 (NN = unique for each VG)
- Hexadecimal
  - 0xNN0000
  - 0x0000000 - 0x090000 (0-9)
  - ox0a0000, 0x0b0000 (10,11)
  - 0xc80000 (200)

```
#
# typeset -i16 hex
# hex=11
# echo $hex
16#b
#
```

List used minor numbers

- See what #'s have been used
- ll /dev/*/group
- Good command to show all VGs
  - (Doesn't mean that VG still exists)

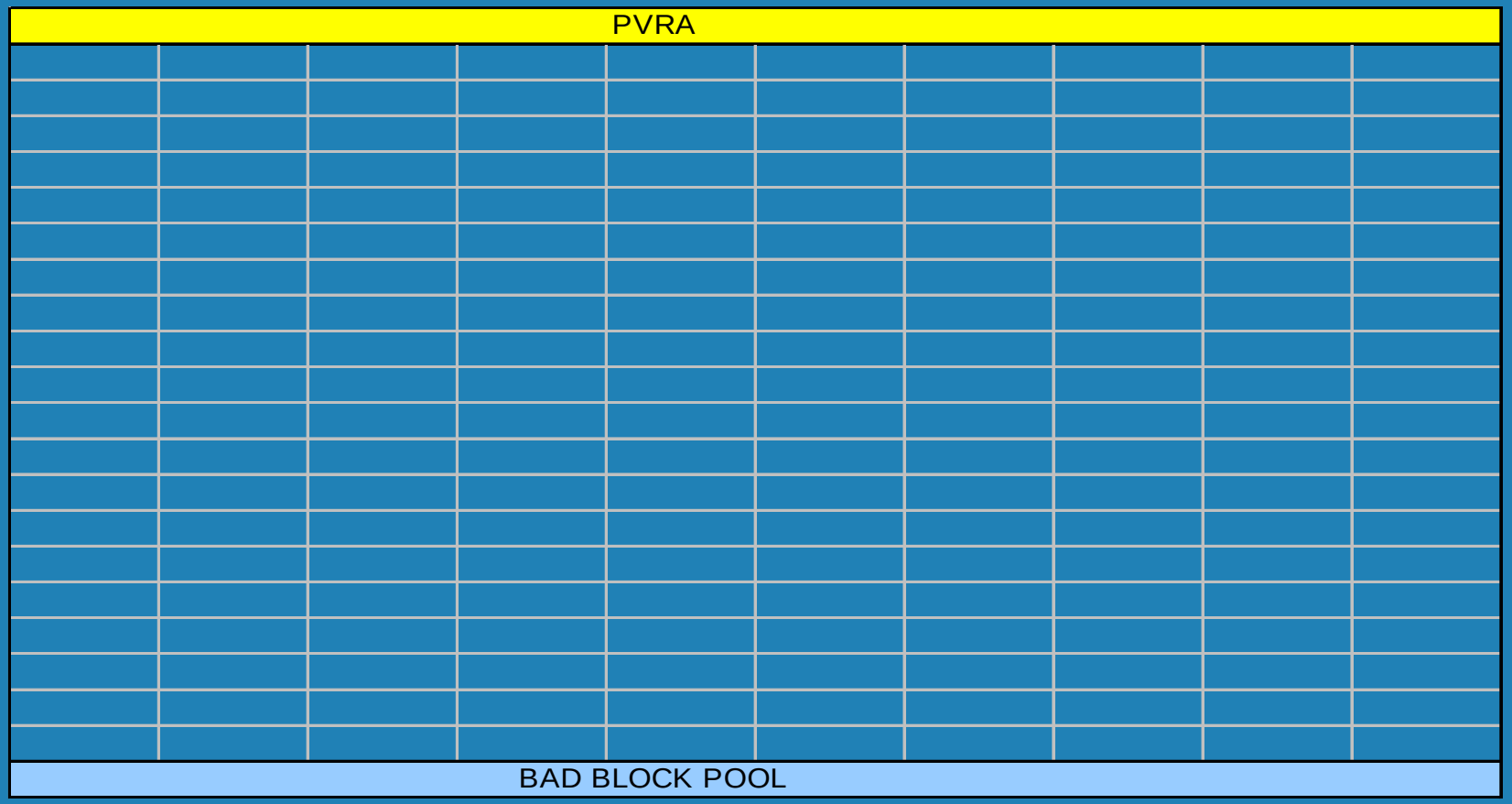
```
# ll /dev/*/group
crw-r-----   1 root       sys          64 0x000000 May 19 11:28 /dev/vg00/group
crw-r-----   1 root       sys          64 0x010000 May 28 10:37 /dev/vg01/group
```

Creating VG
Not bootable

- Disk can not be used in a volume group until it has been initialized with pvcreate
- Create PVRA (Physical Volume Reserve Area]

```
#
# pvcreate /dev/rdsk/c0t8d0
Physical volume "/dev/rdsk/c0t8d0" has been successfully created.
#
# pvcreate -f /dev/rdsk/c0t8d0
Physical volume "/dev/rdsk/c0t8d0" has been successfully created.
#
```

Disk layout after pvcreate



PVRA

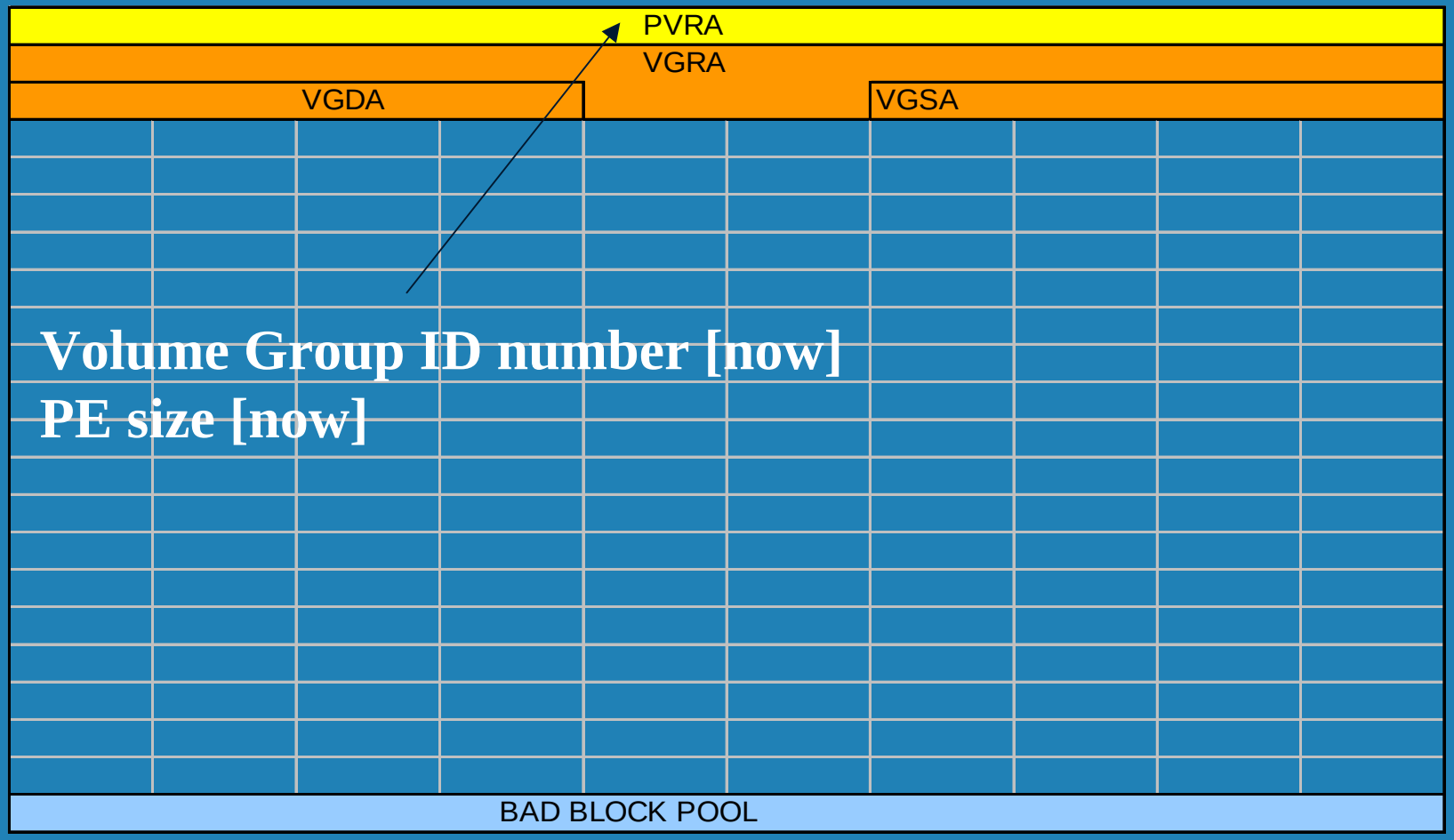
- Physical Volume ID number (CPU ID and time created)
- Volume Group ID number [later]
- PE size [later]
- PV size
- Bad block directory [map of good/bad]
- Pointers to start and size of other disk areas
- Now have PV

Creating VG - continued

- Create directory structure
- Create group file
- Create VG

```
#
# cd /dev
# mkdir vg01
# cd vg01
# mknod group c 64 0x010000
# vgcreate /dev/vg01 /dev/dsk/c0t8d0
Increased the number of physical extents per physical volume to 1023.
Volume group "/dev/vg01" has been successfully created.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
#
```

Disk layout after vgcreate



VGRA

- VGDA
  - Volume Group Description Area
    - # of LVs in this VG
    - How many PVs
    - PE to LE map
- VGSA
  - Volume Group Status Area
    - Status of each PV

Physical Extent - Basic unit of LVM

```
#
# vgdisplay /dev/vg01
--- Volume groups ---
VG Name                     /dev/vg01
VG Write Access             read/write
VG Status                   available
Max LV                      255
Cur LV                      0
Open LV                     0
Max PV                      16
Cur PV                      1
Act PV                      1
Max PE per PV               1023
VGDA                        2
PE Size (Mbytes)            4
Total PE                    1023
Alloc PE                    0
Free PE                     1023
Total PVG                   0
```

- 4 is the default. If want different, set during creation of VG
- Disk overhead increases with smaller extents

pvdisplay (can use after PV added to VG)

- Physical extent size 4MB x Total PE 1023 = 4092 MB

```
#
# pvdisplay /dev/dsk/c0t8d0
--- Physical volumes ---
PV Name                     /dev/dsk/c0t8d0
VG Name                     /dev/vg01
PV Status                   available
Allocatable                 yes
VGDA                        2
Cur LV                      0
PE Size (Mbytes)            4
Total PE                    1023
Free PE                     1023
Allocated PE                0
Stale PE                    0
IO Timeout                  default
```

```
pvdisplay -v /dev/dsk/c0t8d0


   --- Physical extents ---
   PE   Status   LV                    LE
   0000 free                          0000
   0001 free                          0000
   0002 free                          0000
   0003 free                          0000
   0004 free                          0000
   <removed>
   1019 free                          0000
   1020 free                          0000
   1021 free                          0000
   1022 free                          0000
   <end>
```

Physical Extents created during vgcreate

PVRA									
VGRA									
VGDA						VGSA			
0	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19
20	21	22	23	24	25	26	27	28	29
1010	1011	1012	1013	1014	1015	1016	1017	1018	1019
1020	1021	1022							
BAD BLOCK POOL									

Adding 9GB+ drive to existing VG

- MAXIMUM PE per PV (2000) X PE Size 4MB = 8GB
- Any disk added to this volume group can not exceed the maximum. Anything beyond 8GB will NOT be used.
- (If Max PE per PV is 1023 = 4 GB)

```
# vgextend /dev/vg00 /dev/dsk/c0t2d0
vgextend: Not enough physical extents per physical volume.
 Need: 2169, Have: 2000.
Volume group "/dev/vg00" has been successfully extended.
Volume Group configuration for /dev/vg00 has been saved in
f
# vgdisplay /dev/vg00 | grep Max
Max LV                      255
Max PV                      16
Max PE per PV               2000
```

Msg on HP-UX 11

```
# pvcreate /dev/rdsk/c0t2d0
Physical volume "/dev/rdsk/c0t2d0" has been successfully created.
# vgextend /dev/vg01 /dev/dsk/c0t2d0
Warning: Max_PE_per_PV for the volume group (1023) too small for this PV (2169).

         Using only 1023 PEs from this physical volume.
Volume group "/dev/vg01" has been successfully extended.
Warning: Logical Volume number "1073741827" found on physical volume not found i
n "/dev/vg01".
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.con
f
```

vgdisplay

- -v option will display information on each PV in the VG

```
vgdisplay -v /dev/vg01

   --- Physical volumes ---
   PV Name                     /dev/dsk/c0t8d0
   PV Status                   available
   Total PE                    1023
   Free PE                     1023


~
```

Information about all VGs

- What PV belongs to which VG
- Used at boot time by lvmrc
  - Copied into memory
- Includes verion # (9x can't be read by 10x)

```
# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c0t5d0
/dev/vg01
/dev/dsk/c0t8d0
/dev/dsk/c0t2d0
```

vgscan rebuilds /etc/lvmtab

```
#
# rm /etc/lvmtab
# vgscan -v
Creating "/etc/lvmtab".
vgscan: Couldn't access the list of physical volumes for volume group "/dev/vg00
vgscan: Couldn't access the list of physical volumes for volume group "/dev/vg01
vgscan: Couldn't access the list of physical volumes for volume group "/dev/vg02
Physical Volume "/dev/dsk/c1t2d0" contains no LVM information


/dev/vg00
/dev/dsk/c0t5d0
/dev/dsk/c0t11d0
```

- (Don't rm your /etc/lvmtab)

Creating Logical Volume

```
# lvcreate -L 500 vg01
Logical volume "/dev/vg01/lvol1" has been successfully created with
character device "/dev/vg01/rlvol1".
Logical volume "/dev/vg01/lvol1" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.con
f
```

0xVG00LV

```
# ll /dev/vg01
total 0
crw-rw-rw-   1 root       sys          64 0x010000 Apr 21 09:11 group
brw-r-----   1 root       sys          64 0x010001 Apr 22 14:26 lvol1
crw-r-----   1 root       sys          64 0x010001 Apr 22 14:26 rlvol1
# vgdisplay /dev/vg01
--- Volume groups ---
VG Name                     /dev/vg01
VG Write Access             read/write
VG Status                   available
Max LV                      255
Cur LV                      1
Open LV                     1
Max PV                      16
Cur PV                      1
Act PV                      1
Max PE per PV               1023
VGDA                        2
PE Size (Mbytes)            4
Total PE                    1023
Alloc PE                    125
Free PE                     898
Total PVG                   0
```

Config Files

- Volume Group configuration file is updated
- Backup configuration file created automatically
- lvchange, lvcreate, lvextend, lvlnboot, lvmerge, lvreduce, lvremove,
- lvrmboot, lvsplit, pvchange, pvmove, vgcreate, vgextend, vgreduce
- vgcfgbackup

```
#
# ll /etc/lvmconf
total 246
----------   1 root     sys            0 Jun 10  1996 lvm_lock
-rw-------   1 root     sys        34816 Apr 20 20:02 vg00.conf
-rw-------   1 root     sys        44032 Apr 20 20:01 vg00.conf.old
-rw-------   1 root     sys        23552 Apr 22 14:26 vg01.conf
-rw-------   1 root     sys        23552 Apr 21 09:12 vg01.conf.old
#
```

View configuration file

```
# vgcfgrestore -l -n vg00
Volume Group Configuration information in "/etc/lvmconf/vg00.conf"
VG Name /dev/vg00
 ---- Physical volumes : 1 ----
    /dev/rdsk/c0t5d0 (Bootable)
```

Extents - LE to PE

```
--- Physical extents ---
PE    Status    LV                          LE
0000  current   /dev/vg01/lvol1             0000
0001  current   /dev/vg01/lvol1             0001
0002  current   /dev/vg01/lvol1             0002
0003  current   /dev/vg01/lvol1             0003
0004  current   /dev/vg01/lvol1             0004
<removed>
0121  current   /dev/vg01/lvol1             0121
0122  current   /dev/vg01/lvol1             0122
0123  current   /dev/vg01/lvol1             0123
0124  current   /dev/vg01/lvol1             0124
0125  free                                  0000
<removed>
1021  free                                  0000
1022  free                                  0000
<end>
```

- LV always starts with LE 0
- By default, allocates PE to those that are free in the order you originally added disks to VG

Disk layout after lvcreate

of LVs
PE to LE Map

PVRA									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	26 - 26	27 - 27	28 - 28	29 - 29
< extents not displayed >				124 - 124					
1010 free	1011 free	1012 free	1013 free	1014 free	1015 free	1016 free	1017 free	1018 free	1019 free
1020 free	1021 free	1022 free							
BAD BLOCK POOL									

lvdisplay

```
#
# lvdisplay /dev/vg01/lvol1
--- Logical volumes ---
LV Name                     /dev/vg01/lvol1
VG Name                     /dev/vg01
LV Permission               read/write
LV Status                   available/syncd
Mirror copies               0
Consistency Recovery        MWC
Schedule                    parallel
LV Size (Mbytes)            500
Current LE                  125
Allocated PE                125
Stripes                     0
Stripe Size (Kbytes)        0
Bad block                   on
Allocation                  strict
```

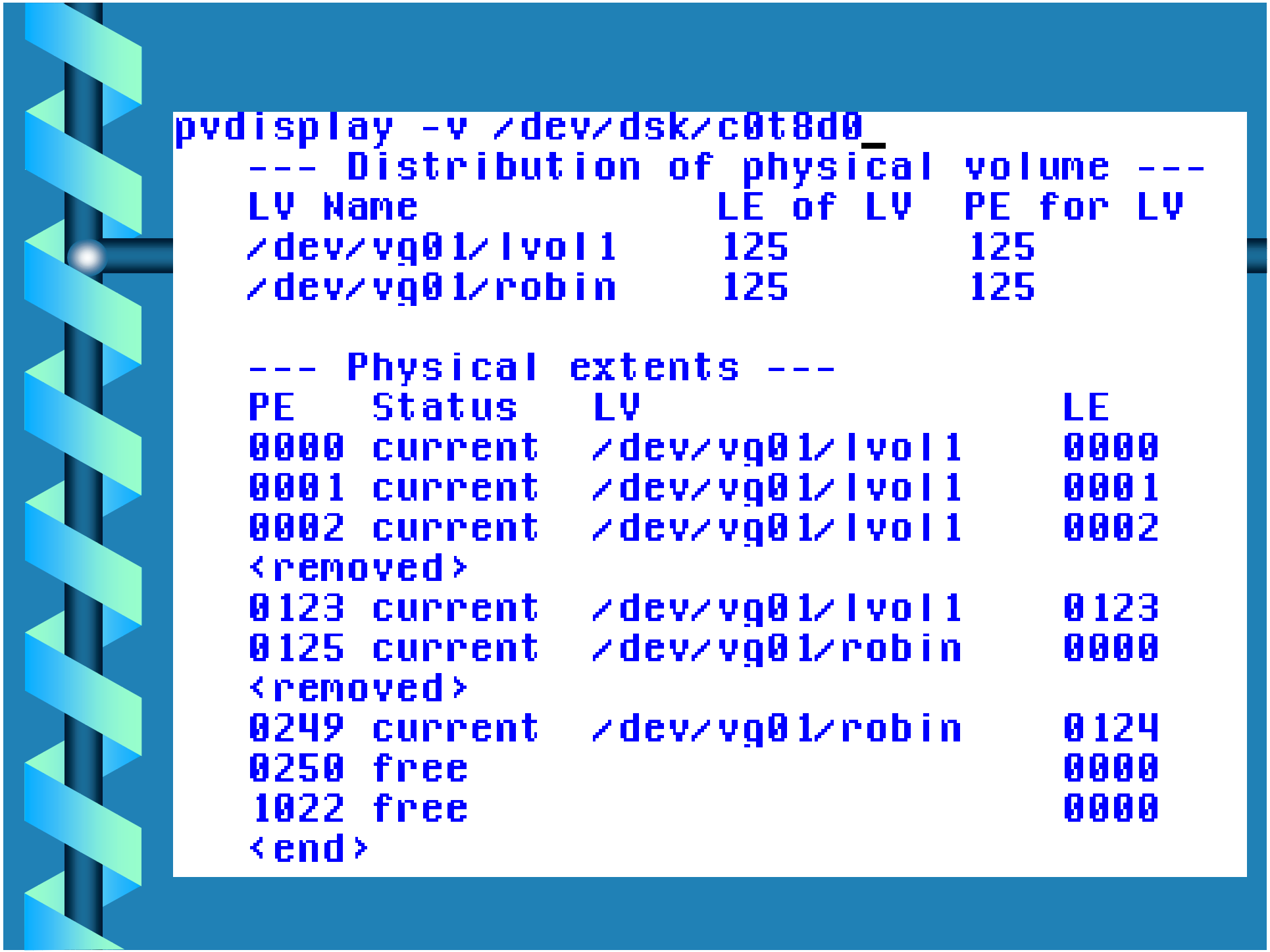
Create LV without default name

```
# lvcreate -L 500 -n robin /dev/vg01
Logical volume "/dev/vg01/robin" has been successfully created with
character device "/dev/vg01/rrobin".
Logical volume "/dev/vg01/robin" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# vgdisplay /dev/vg01
--- Volume groups ---
VG Name                     /dev/vg01
VG Write Access             read/write
VG Status                   available
Max LV                      255
Cur LV                      2
Open LV                     2
Max PV                      16
Cur PV                      1
Act PV                      1
Max PE per PV               1023
VGDA                        2
PE Size (Mbytes)            4
Total PE                    1023
Alloc PE                    250
Free PE                     773
# tal PVG                   0
```

-n to specify name

```
vgdisplay -v /dev/vg01
   --- Logical volumes ---

   LV Name                     /dev/vg01/lvol1
   LV Status                   available/syncd
   LV Size (Mbytes)            500
   Current LE                  125
   Allocated PE                125
   Used PV                     1

   LV Name                     /dev/vg01/robin
   LV Status                   available/syncd
   LV Size (Mbytes)            500
   Current LE                  125
   Allocated PE                125
   Used PV                     1


   --- Physical volumes ---
   PV Name                     /dev/dsk/c0t8d0
   PV Status                   available
   Total PE                    1023
```

```
#
# pvdisplay /dev/dsk/c0t8d0
--- Physical volumes ---
PV Name                     /dev/dsk/c0t8d0
VG Name                     /dev/vg01
PV Status                   available
Allocatable                 yes
VGDA                        2
Cur LV                      2
PE Size (Mbytes)            4
Total PE                    1023
Free PE                     773
Allocated PE                250
Stale PE                    0
IO Timeout                  default
```


pvdisplay -v /dev/dsk/c0t8d0
--- Distribution of physical volume ---
LV Name LE of LV PE for LV
/dev/vg01/lvol1 125 125
/dev/vg01/robin 125 125
--- Physical extents ---
PE Status LV LE
0000 current /dev/vg01/lvol1 0000
0001 current /dev/vg01/lvol1 0001
0002 current /dev/vg01/lvol1 0002
<removed>
0123 current /dev/vg01/lvol1 0123
0125 current /dev/vg01/robin 0000
<removed>
0249 current /dev/vg01/robin 0124
0250 free 0000
1022 free 0000
<end>

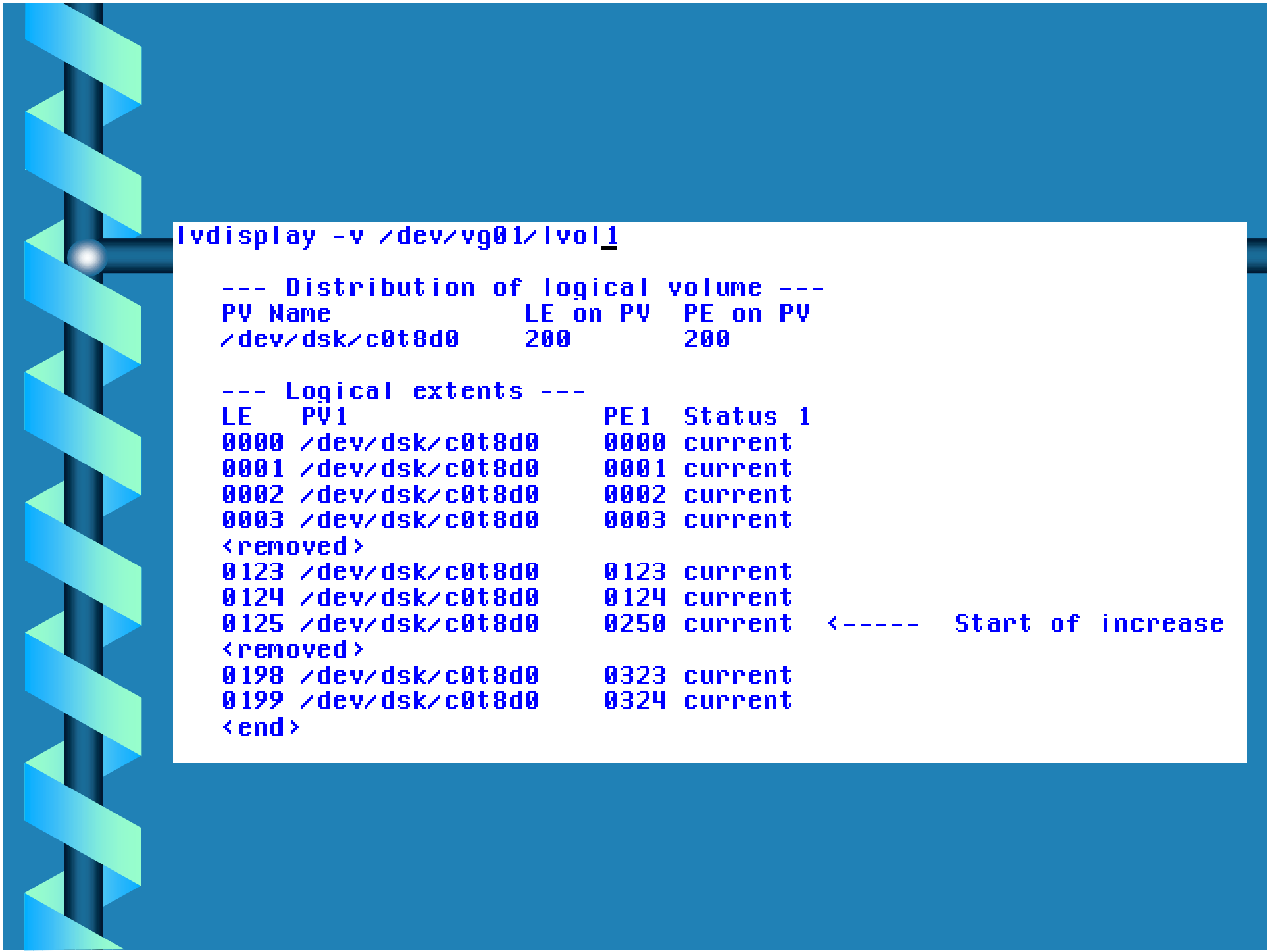
Disk layout after adding 2nd LV

PVRA									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	26 - 26	27 - 27	28 - 28	29 - 29
< extents not displayed >				124 - 124	125 - 0	126 - 1	127 - 2	128 - 3	129 - 4
130 - 5	131 - 6	132 - 7	133 - 8	134 - 9	< extents not displayed >				249 - 124
1010 free	1011 free	1012 free	1013 free	1014 free	1015 free	1016 free	1017 free	1018 free	1019 free
1020 free	1021 free	1022 free							
BAD BLOCK POOL									

Extend size of Logical Volume

```
#
# lvextend -L 800 /dev/vg01/lvol1
Logical volume "/dev/vg01/lvol1" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
#
#
```

New Size
-L (in MB)


lvdisplay -v /dev/vg01/lvol1
--- Distribution of logical volume ---
PV Name LE on PV PE on PV
/dev/dsk/c0t8d0 200 200
--- Logical extents ---
LE PV1 PE1 Status 1
0000 /dev/dsk/c0t8d0 0000 current
0001 /dev/dsk/c0t8d0 0001 current
0002 /dev/dsk/c0t8d0 0002 current
0003 /dev/dsk/c0t8d0 0003 current
<removed>
0123 /dev/dsk/c0t8d0 0123 current
0124 /dev/dsk/c0t8d0 0124 current
0125 /dev/dsk/c0t8d0 0250 current <----- Start of increase
<removed>
0198 /dev/dsk/c0t8d0 0323 current
0199 /dev/dsk/c0t8d0 0324 current
<end>

Disk layout after increasing first LV

PVRA									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	26 - 26	27 - 27	28 - 28	29 - 29
< extents not displayed >				124 - 124	125 - 0	126 - 1	127 - 2	128 - 3	129 - 4
130 - 5	131 - 6	132 - 7	133 - 8	134 - 9	< extents not displayed >				249 - 124
250 - 125	251 - 126	252 - 127	253 - 128	254 - 129	255 - 130	256 - 131	257 - 132	258 - 133	259 - 134
< extents not displayed >			323 - 198	324 - 199					
1010 free	1011 free	1012 free	1013 free	1014 free	1015 free	1016 free	1017 free	1018 free	1019 free
1020 free	1021 free	1022 free							
BAD BLOCK POOL									

End of Disk (PE)

Out of space in VG

```
#
# diskinfo /dev/rdsk/c0t8d0
SCSI describe of /dev/rdsk/c0t8d0:
               vendor: SEAGATE
           product id: ST34371W
                 type: direct access
                 size: 4194157 Kbytes
    bytes per sector: 512
# lvextend -L 3400 /dev/vg01/robin
lvextend: Not enough free physical extents available.
Logical volume "/dev/vg01/robin" could not be extended.
Failure possibly caused by strict allocation policy
#
```

Need 850 LE (3400 / 4)

VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	26 - 26	27 - 27	28 - 28	29 - 29
< extents not displayed >				124 - 124	125 - 0	126 - 1	127 - 2	128 - 3	129 - 4
130 - 5	131 - 6	132 - 7	133 - 8	134 - 9	< extents not displayed >				249 - 124
250 - 125	251 - 126	252 - 127	253 - 128	254 - 129	255 - 130	256 - 131	257 - 132	258 - 133	259 - 134
< extents not displayed >			323 - 198	324 - 199	325 - 125	326 - 126	327 - 127	328 - 128	329 - 129
330 - 130	331 - 131	332 - 132	333 - 133	334 - 134	335 - 135	336 - 136	< extents not displayed >		
1010 - 810	1011 - 811	1012 - 812	1013 - 813	1014 - 814	1015 - 815	1016 - 816	1017 - 817	1018 - 818	1019 - 819
1020 - 820	1021 - 821	1022 - 822							
BAD BLOCK POOL									

Extend the Volume Group

```
# ll /dev/dsk
total 0
brw-r-----   1 bin        sys          31 0x00a000 Apr  2 09:42 c0t10d0
brw-r-----   2 root       sys          31 0x005000 Jun 10  1996 c0t5d0
brw-r-----   1 root       sys          31 0x008000 Apr  9 17:30 c0t8d0
brw-r-----   1 bin        sys          31 0x009000 Apr  2 11:25 c0t9d0
brw-r-----   1 root       sys          31 0x012000 Jun  9  1996 c1t2d0
brw-r-----   2 root       sys          31 0x005000 Jun 10  1996 root
# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c0t5d0
/dev/vg01
/dev/dsk/c0t8d0
# vgextend /dev/vg01 /dev/dsk/c0t9d0
Volume group "/dev/vg01" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c0t5d0
/dev/vg01
/dev/dsk/c0t8d0
/dev/dsk/c0t9d0
```

```
# lvextend -L 3400 /dev/vg01/robin
Logical volume "/dev/vg01/robin" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
```

```
PV Name             LE on PV  PE on PV
/dev/dsk/c0t9d0     27        27
/dev/dsk/c0t8d0     823       823

--- Logical extents ---
LE   PV1                 PE1  Status 1
0000 /dev/dsk/c0t8d0     0125 current >------------ Original Creation
0001 /dev/dsk/c0t8d0     0126 current
0002 /dev/dsk/c0t8d0     0127 current
<removed>
0123 /dev/dsk/c0t8d0     0248 current
0124 /dev/dsk/c0t8d0     0249 current
0125 /dev/dsk/c0t8d0     0325 current >------------ Start of Extend
0126 /dev/dsk/c0t8d0     0326 current
<removed>
0821 /dev/dsk/c0t8d0     1021 current
0822 /dev/dsk/c0t8d0     1022 current
0823 /dev/dsk/c0t9d0     0000 current >------------ Continuation of
0824 /dev/dsk/c0t9d0     0001 current                extend on 2nd disk
<removed>
0848 /dev/dsk/c0t9d0     0025 current
0849 /dev/dsk/c0t9d0     0026 current
<end>
```

Disk layout of 2 disks in VG

PVRA									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	26 - 26	27 - 27	28 - 28	29 - 29
< extents not displayed >				124 - 124	125 - 0	126 - 1	127 - 2	128 - 3	129 - 4
130 - 5	131 - 6	132 - 7	133 - 8	134 - 9	< extents not displayed >				249 - 124
250 - 125	251 - 126	252 - 127	253 - 128	254 - 129	255 - 130	256 - 131	257 - 132	258 - 133	259 - 134
< extents not displayed >			323 - 198	324 - 199	325 - 125	326 - 126	327 - 127	328 - 128	329 - 129
330 - 130	331 - 131	332 - 132	333 - 133	334 - 134	335 - 135	336 - 136	< extents not displayed >		
1010 - 810	1011 - 811	1012 - 812	1013 - 813	1014 - 814	1015 - 815	1016 - 816	1017 - 817	1018 - 818	1019 - 819
1020 - 820	1021 - 821	1022 - 822							
BAD BLOCK POOL									
PVRA									
VGRA									
VGDA						VGSA			
0 - 823	1 - 824	2 - 825	3 - 826	4 - 827	5 - 828	6 - 829	7 - 830	8 - 831	9 - 832
< extents not displayed >			23 - 846	24 - 847	25 - 848	26 - 849			
< extents not displayed >									
1020 - free	1021 -free	1023 - free							
BAD BLOCK POOL									

Remove every thing we've done!

```
# lvremove /dev/vg01/lvol1
The logical volume "/dev/vg01/lvol1" is not empty;
do you really want to delete the logical volume (y/n) : y
Logical volume "/dev/vg01/lvol1" has been successfully removed.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# lvremove /dev/vg01/robin
The logical volume "/dev/vg01/robin" is not empty;
do you really want to delete the logical volume (y/n) : y
Logical volume "/dev/vg01/robin" has been successfully removed.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# vgreduce /dev/vg01 /dev/dsk/c0t8d0
Volume group "/dev/vg01" has been successfully reduced.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# vgremove /dev/vg01
Volume group "/dev/vg01" has been successfully removed.
#
```

Create Striped Logical Volume

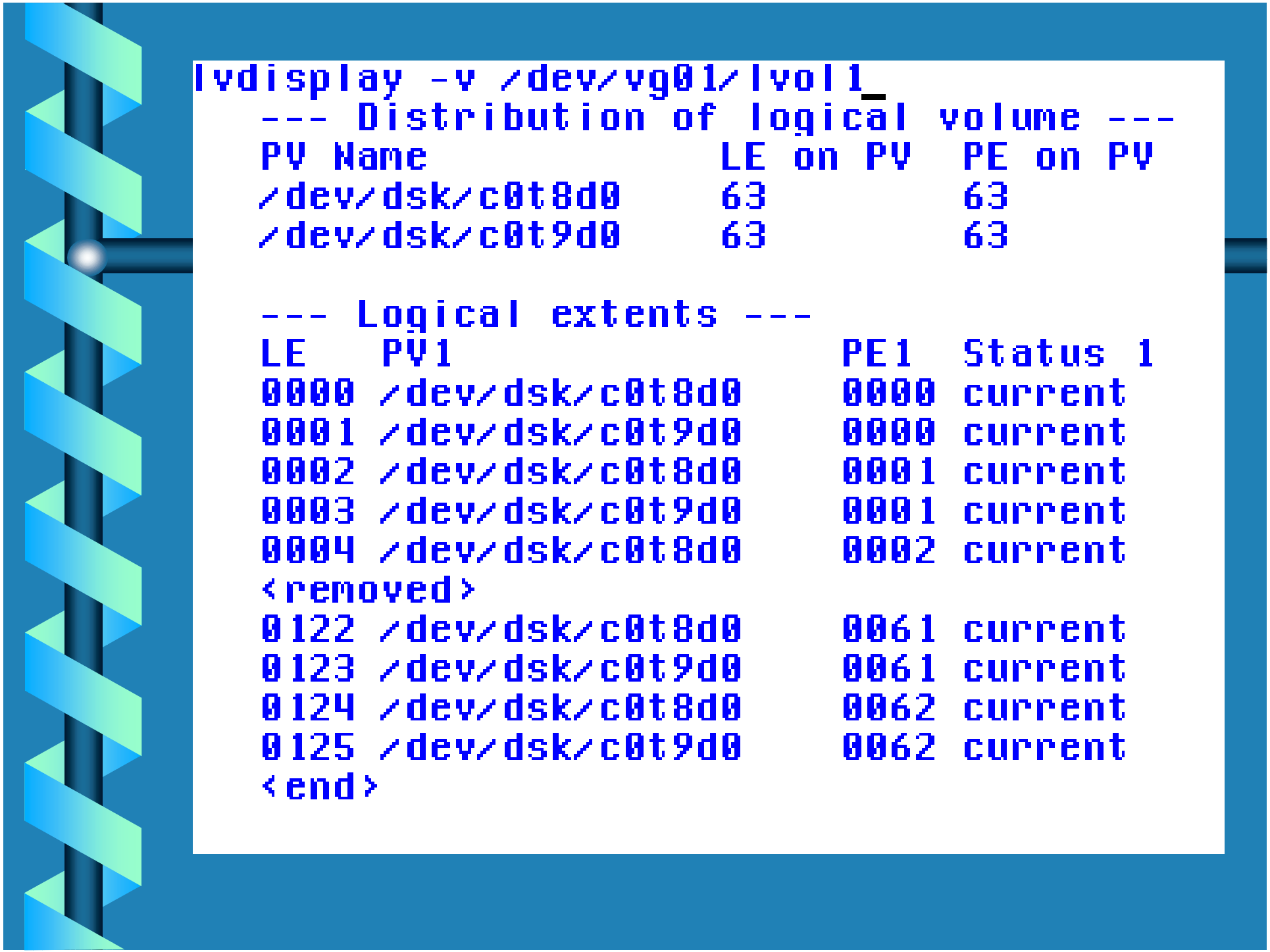
```
# vgcreate /dev/vg01 /dev/dsk/c0t8d0
Increased the number of physical extents per physical volume to 1023.
Volume group "/dev/vg01" has been successfully created.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# vgextend /dev/vg01 /dev/dsk/c0t9d0
Volume group "/dev/vg01" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# lvcreate -L 500 -i 2 /dev/vg01
Warning: rounding up logical volume size to extent boundary at size "504" MB f
 striping.
Logical volume "/dev/vg01/lvol1" has been successfully created with
character device "/dev/vg01/rlvol1".
Logical volume "/dev/vg01/lvol1" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
#
```

Default Block Size is 8Kb

```
#
# lvdisplay /dev/vg01/lvol1
--- Logical volumes ---
LV Name                     /dev/vg01/lvol1
VG Name                     /dev/vg01
LV Permission               read/write
LV Status                   available/syncd
Mirror copies               0
Consistency Recovery        MWC
Schedule                    striped
LV Size (Mbytes)            504
Current LE                  126
Allocated PE                126
Stripes                     2
Stripe Size (Kbytes)        8
Bad block                   on
Allocation                  strict
```

Striping for JFS File System

```
# lvcreate -L 100 -i 2 -I 64 vg01
Warning: rounding up logical volume size to extent boundary at size "104" MB for
 striping.
Logical volume "/dev/vg01/lvol3" has been successfully created with
character device "/dev/vg01/rlvol3".
Logical volume "/dev/vg01/lvol3" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.con
f
```


lvdisplay -v /dev/vg01/lvol1
--- Distribution of logical volume ---
PV Name LE on PV PE on PV
/dev/dsk/c0t8d0 63 63
/dev/dsk/c0t9d0 63 63
--- Logical extents ---
LE PV1 PE1 Status 1
0000 /dev/dsk/c0t8d0 0000 current
0001 /dev/dsk/c0t9d0 0000 current
0002 /dev/dsk/c0t8d0 0001 current
0003 /dev/dsk/c0t9d0 0001 current
0004 /dev/dsk/c0t8d0 0002 current
<removed>
0122 /dev/dsk/c0t8d0 0061 current
0123 /dev/dsk/c0t9d0 0061 current
0124 /dev/dsk/c0t8d0 0062 current
0125 /dev/dsk/c0t9d0 0062 current
<end>

Disk layout - striped LV Extents

PVRA (DISK 1)									
VGRA									
VGDA						VGSA			
0 - 0	1 - 2	3 - 4	3 - 6	4 - 8	5 - 10	6 - 12	7 - 14	8 - 16	9 - 18
< extents not displayed >									
60 - 120	61 - 122	62 - 124							
1020 - free	1021 -free	1023 - free							
BAD BLOCK POOL									
PVRA (DISK 2)									
VGRA									
VGDA						VGSA			
0 - 1	1 - 3	2 - 5	3 - 7	4 - 9	5 - 11	6 - 13	7 - 15	8 - 17	9 - 19
< extents not displayed >									
60 - 121	61 - 123	62 - 125							
1020 - free	1021 -free	1023 - free							
BAD BLOCK POOL									

Disk layout - striped LV

Stripes (blocks) 4MB / 8Kb = apx. 500

	Disk1		Disk2
LE0	B1,B3,B5,B7,B9,B11 B13,B15…….B499	LE1	B2,B4,B6,B8,B10,B12 B14,B17…….B500
LE2	B501,B503,B505, B507,B509…. B999	LE3	B502,B504,B506, B508,B510…. B1000
LE4	B1001,B1003,B1005, B1007…. B1499	LE5	B1002,B1004,B1006, B1008…. B1500

Disk Striping

- Performance
  - Read & write of large, sequentially accessed files
  - Best when use similar disks
  - Best when use more than one bus
- Reliability
  - Loss of one disk could possibly and most likely corrupt the entire volume group

Disk Striping Recommendations

- Don't stripe everything everywhere
- Keep striped VGs less than or equal to 4 physical disks
- HFS - stripe size = block size
- JFS - stripe size = average extent size
  - Source: HP-UX Tuning & Performance by Sauers & Weygant
  - ISBN: 0-13-102716-6

File Systems

- A newly created logical volume is considered raw
- You can use it as raw or place a file system on it
- Common file systems:
  - HFS, JFS

What is a File System?

- A file system structure contains:
  - A collection of files organized under a hierarchical or directory structure
- Auxiliary file systems are those that you can mount/umount (/opt, not /etc)
- Where you attach the file system to the HP-UX file system tree is called the mount point

JFS

- Journaling file system that uses an intent log
- Can not be used as root file system on 10.01 or 10.10
- /stand must be HFS on 10.20+
- Uses small blocks & extents rather than large blocks & fragments (HFS). Extent = adjacent disk blocks treated as a unit. Vary in size. No relation to LVM extent
- New technology added to JFS, not HFS
- Works with NFS
- Veritas (vxfs)
- ~300k additional memory
- Dynamically allocates inodes

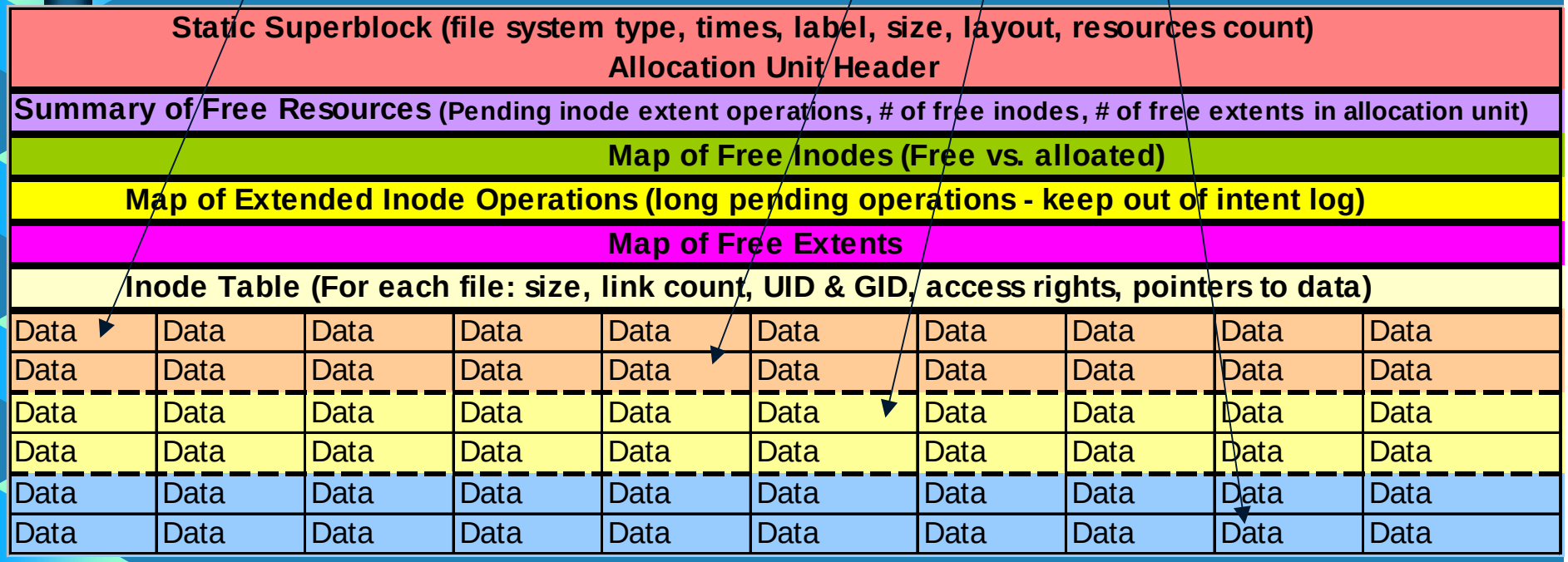
Parts of JFS File System

- Superblock - keeps track of the file system. Pointers to maps of free spaces. Multiple (static) copies of the superblock - since so important.
- Inodes
- Maps
- Directories
- Data Blocks
- Where these areas are depends on version of JFS (structural and unnamed filesets)

Default 1024 byte data blocks

Allocation Unit: Group of consecutive blocks

Extent: Contiguous area of data blocks
Example: 3



Header information now controlled in the structural fileset (JFS version 3 and 4)

Simplified view

Extents	Block # and count		100/5		File1		
	Block # and count		120/18		File2		
File1	File1	File1	File1	File1	Data	Data	Data	Data	Data
Data	Data	Data	Data	Data	Data	Data	Data	Data	Data
File2	File2	File2	File2	File2	File2	File2	File2	File2	File2
File2	File2	File2	File2	File2	File2	File2	File2	Data	Data

Extents	Block # and count		100/19		File1	Increase File1	
	Block # and count		120/18		File2		
File1	File1	File1	File1	File1	File1	File1	File1	File1	File1
File1	File1	File1	File1	File1	File1	File1	File1	File1	Data
File2	File2	File2	File2	File2	File2	File2	File2	File2	File2
File2	File2	File2	File2	File2	File2	File2	File2	Data	Data

Extents	Block # and count		100/20 + 140/19		File1	Increase File1	
	Block # and count		120/18		File2		
File1	File1	File1	File1	File1	File1	File1	File1	File1	File1
File1	File1	File1	File1	File1	File1	File1	File1	File1	File1
File2	File2		File2	File2	File2	File2	File2	File2
File2	File2		File2	File2	File2	File2	Data	Data
File1	File1		File1	File1	File1	File1	File1	File1
File1	File1		File1	File1	File1	File1	File1	Data

File 1:
2 extents

JFS much more sophisticated

Extents	Block # and count		100/5		File1	JFS: Extents
	Block # and count		120/18		File2	vary in size!
	Block # and count		160/59		File3	
File1	File1	File1	File1	File1	Data	Data	Data	Data	Data	Data
Data	Data	Data	Data	Data	Data	Data	Data	Data	Data	Data
File2	File2	File2	File2	File2	File2	File2	File2	File2	File2	File2
File2	File2	File2	File2	File2	File2	File2	File2	Data	Data	
File1	File1	File1	File1	File1	File1	File1	File1	File1	File1	File1
File1	File1	File1	File1	File1	File1	File1	File1	File1	Data	
File3	File3	File3	File3	File3	File3	File3	File3	File3	File3	File3
File3	File3	File3	File3	File3	File3	File3	File3	File3	File3	File3
File3	File3	File3	File3	File3	File3	File3	File3	File3	File3	File3
File3	File3	File3	File3	File3	File3	File3	File3	File3	File3	File3
File3	File3	File3	File3	File3	File3	File3	File3	File3	File3	File3
File3	File3	File3	File3	File3	File3	File1	File1	File1	Data	

```
super-block at 00000008.0000
magic a501fcf5  version 4
ctime 947969103 476150  (Sat Jan 15 13:45:03 2000 MDT)
log_version 9 logstart 0  logend 0
bsize  1024 size  204800 dsize  204800  ninode 0  nau 0
defiextsize 0  oilbsize 0  immedlen 96  ndaddr 10
aufirst 0  emap 0  imap 0  iextop 0  istart 0
bstart 0  femap 0  fimap 0  fiextop 0  fistart 0  fbstart 0
nindir 2048  aulen 32768  auimlen 0  auemlen 8
auilen 0  aupad 0  aublocks 32768  maxtier 15
inopb 4  inopau 0  ndiripau 0  iaddrlen 8   bshift 10
inoshift 2  bmask fffffc00  boffmask 3ff  checksum dd9169d6
free 192545  ifree 48
efree  5 2 0 3 2 3 2 0 1 1 1 1 0 1 1 1 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
flags 0 mod 0 clean 3c
time 948513586 0  (Fri Jan 21 20:59:46 2000 MDT)
oltext[0] 1285  oltext[1] 1286  oltsize 1
iauimlen 1  iausize 4  dinosize 256
  checksum2 b14
 checksum3 0
```

- Block size = 1024, Allocation unit size = 32MB
- Total amount free = 192545
- Free extents

Next Slide

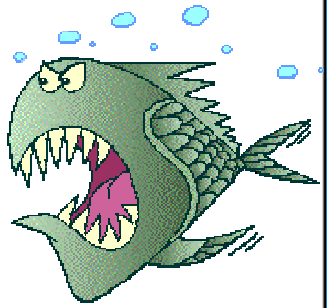
Object Location Table (Disk based data structures)

```
> 1285b; p oltext
OLT at 0x00000505.0000
OLT head entry:
        olt_magic 0xa504fcf5  olt_size 56  olt_totfree 872
        olt_time 947969103 476150  (Sat Jan 15 13:45:03 2000 MDT)
        olt_checksum 0x38883b5b
        olt_esize 1  olt_extents[1285 1286]
        olt_nsize 0  olt_next[0 0]
OLT fshead entry:
        olt_type 2  olt_size 16  olt_fsino[3 35]
OLT initial iext entry:
        olt_type 4  olt_size 16  olt_iext[1136 1144]
OLT cut entry:
        olt_type 3  olt_size 16  olt_cutino 6
OLT device entry:
        olt_type 5  olt_size 16  olt_devino[8 40]
OLT super-block entry:
        olt_type 6  olt_size 32  olt_sbino 33
        olt_logino[9 41]  olt_oltino[7 39]
OLT free entry:
        olt_type 1  olt_fsize 872
>
```

Intent log - key feature of JFS

```
# pwd
/jfs33
# mkdir zzz12
# uptime
  9:10pm  up 4 days, 15:10,  2 users,  load average: 0.21, 0.19, 0.23
# fsdb -F vxfs /dev/vg00/rjfs33 > myfile.log
fmtlog
q
```

```
00006180: id 156  func 1  ser 2  lser 3  len 292
Inode Modification  fset 999  ilist 0  dev/bno 0/1123  ino 12  osize 292
New Inode Contents:
type IFDIR mode 40700  nlink 2  uid 0  gid 3  size 96
atime 948514205 340001  (Fri Jan 21 21:10:05 2000 MDT)
mtime 948514205 340001  (Fri Jan 21 21:10:05 2000 MDT)
ctime 948514205 340001  (Fri Jan 21 21:10:05 2000 MDT)
aflags 0 orgtype 2 eopflags 0 eopdata 0
fixextsize/fsindex 0  rdev/reserve/dotdot/matchino 2
blocks 0  gen 0  version 0 1  iattrino 0

000062e0: id 156  func 2  ser 3  lser 3  len 55
directory fset 999  ilist 0  inode 2  bno 10  blen 1024  boff 204
previous  d_ino 11  d_reclen 840  d_namlen 10  d_hashnext 0000
added   d_ino 12  d_reclen 820  d_namlen 5  d_hashnext 0088
                z  z  z  1  2
```



Circular File

JFS

- JFS extents and LVM extents - NO RELATIONSHIP
- Learn more about JFS:
  - HP Inside HP-UX class (deep)
  - Veritas File System Administrator's Guide
    - docs.hp.com/hpux/os#papers
    - Very Good

JFS 10x vs. 11x - vxfs daemon

```
    root     65     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     66     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     67     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     68     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     69     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     70     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     71     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     72     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     73     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     74     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     75     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     76     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     77     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     78     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     79     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     80     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     81     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     82     0   0  Nov 22  ?          0:00 vx_inactive_thread
    root     83     0   0  Nov 22  ?          0:00 vx_inactive_thread
# uname -r
B.10.20
```

```
#
# ps -ef | grep vx
    root    28     0   0  Nov 15  ?          5:00 vxfsd
    root  9587  9572   1 16:16:12 pts/t0     0:00 grep vx
# uname -r
B.11.00
```

Creating a new file system

```
# ll /dev/vg01
total 0
crw-rw-rw-   1 root       sys          64 0x010000 Apr 21 09:11 group
brw-r-----   1 root       sys          64 0x010001 Apr 22 16:42 lvol1
brw-r-----   1 root       sys          64 0x010002 Apr 22 16:48 lvol2
crw-r-----   1 root       sys          64 0x010001 Apr 22 16:42 rlvol1
crw-r-----   1 root       sys          64 0x010002 Apr 22 16:48 rlvol2
# newfs /dev/vg01/rlvol1
    version 3 layout
    516096 sectors, 516096 blocks of size 1024, log size 1024 blocks
    unlimited inodes, 516096 data blocks, 514880 free data blocks
    16 allocation units of 32768 blocks, 32768 data blocks
    last allocation unit has 24576 data blocks
    first allocation unit starts at block 0
    overhead per allocation unit is 0 blocks
#
```

- newfs -F vxfs
- newfs -F hfs
- /etc/default/fs

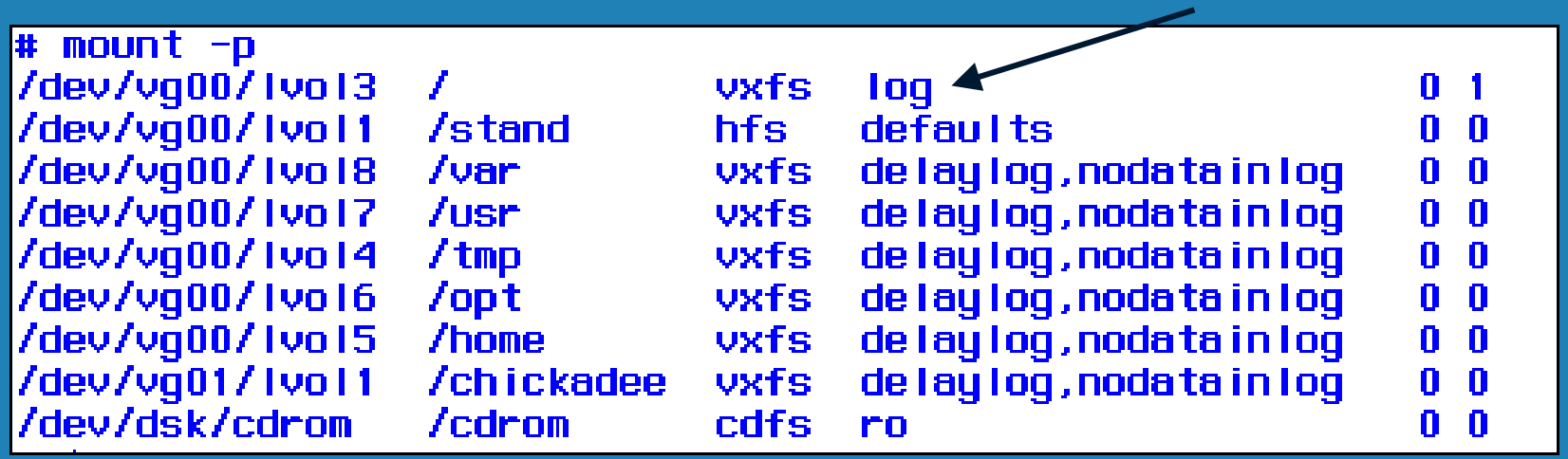
Mount the new file system

- Edit /etc/fstab to include new mount info

```
/dev/vg00/lvol3 / vxfs delaylog 0 1
/dev/vg00/lvol1 /stand hfs defaults 0 1
/dev/vg00/lvol4 /tmp vxfs delaylog 0 2
/dev/vg00/lvol5 /home vxfs delaylog 0 2
/dev/vg00/lvol6 /opt vxfs delaylog 0 2
/dev/vg00/lvol7 /usr vxfs delaylog 0 2
/dev/vg00/lvol8 /var vxfs delaylog 0 2
/dev/vg01/lvol1 /chickadee vxfs delaylog 0 3
#
```

mount -p

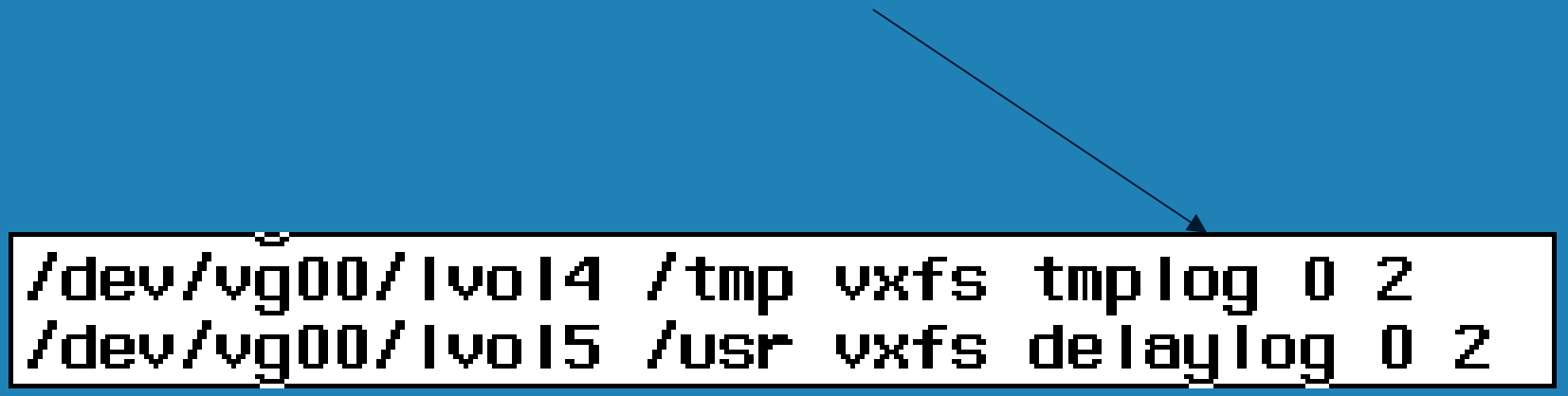
- Reads current mounts (/etc/mnttab)
- / is different since at boot does not read fstab, uses default of log
- -o remount will remount



```
# mount -p
/dev/vg00/lvol3  /           vxfs   log                      0 1
/dev/vg00/lvol1  /stand      hfs    defaults                 0 0
/dev/vg00/lvol8  /var        vxfs   delaylog,nodatainlog     0 0
/dev/vg00/lvol7  /usr        vxfs   delaylog,nodatainlog     0 0
/dev/vg00/lvol4  /tmp        vxfs   delaylog,nodatainlog     0 0
/dev/vg00/lvol6  /opt        vxfs   delaylog,nodatainlog     0 0
/dev/vg00/lvol5  /home       vxfs   delaylog,nodatainlog     0 0
/dev/vg01/lvol1  /chickadee  vxfs   delaylog,nodatainlog     0 0
/dev/dsk/cdrom   /cdrom      cdfs   ro                       0 0
```

Mount options

- Data integrity & performance
- Chapter 5 - Performance & Tuning - VxFS System Administrator's Guide



Converting to JFS

- See HPRC Document#:
  - UNX1030189
  - Converting root HFS to root JFS

New in JFS 3.3

- Access Control Lists
- Version 4 Disk Layout
- File System Shrink Enhancement
- vxtunefs command

Access Control List

- Additional access control mechanism
- Access permission at a finer level:
  - User
  - Group
  - Or combination of
- R,W,X with a particular User/Group combination
- (mouse.%,r-x) (% = any)
- See Appendix A for details

Tuning/Performance "With appropriate tuning, JFS outperforms HFS in all categories"

VxFs provides the following performance enhancements:

- extent based allocation
- enhanced mount options
- data synchronous I/O
- direct I/O and discovered direct I/O
- caching advisories
- enhanced directory features
- explicit file alignment, extent size, and preallocation controls
- tuneable I/O parameters
- tuneable indirect data extent size

From VxFS System Administrator's Guide

JFS Tunable Parameters

- None in HP-UX 10
- In HP-UX 11:
  - vx_ncsize
    - Normally don't need to change
  - vxfs_ra_per_disk
    - Normally don't need to change
  - vxfs_max_ra_kbytes
    - Increase if many large sequential I/Os from 1024 to 65536
- vxtunefs (chapter 5 Perf. & Tuning)

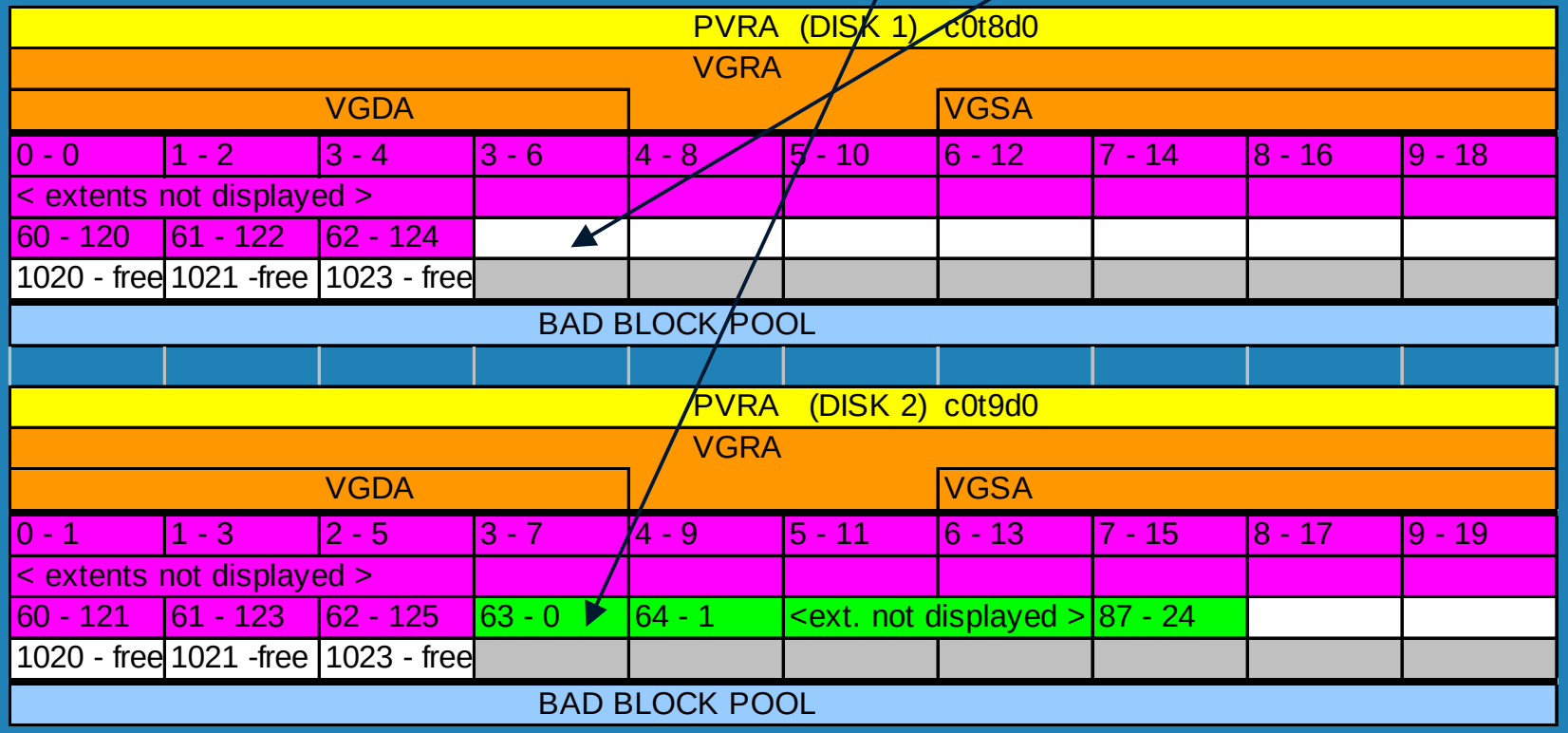
Management of LVM

- Move LVs around
- Replace PV
- Move PV
- Disaster Recovery

Create LV on specific disk

```
#
# lvcreate /dev/vg01
Logical volume "/dev/vg01/lvol2" has been successfully created with
character device "/dev/vg01/rlvol2".
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
# lvextend -L 100 /dev/vg01/lvol2 /dev/dsk/c0t9d0
Logical volume "/dev/vg01/lvol2" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
#
```

Instead of using next default available space, used next available on specific disk

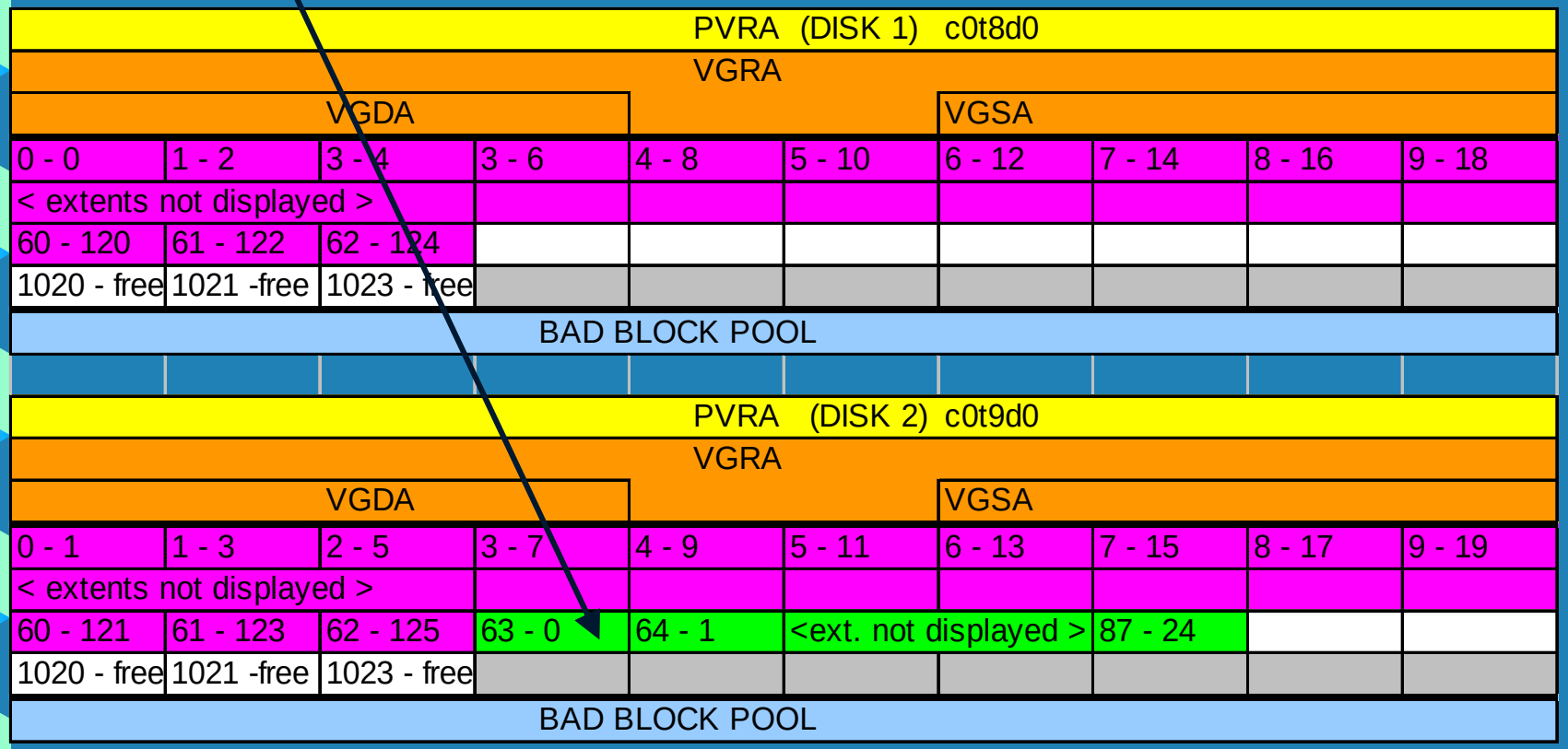


Controlling LV disk layout

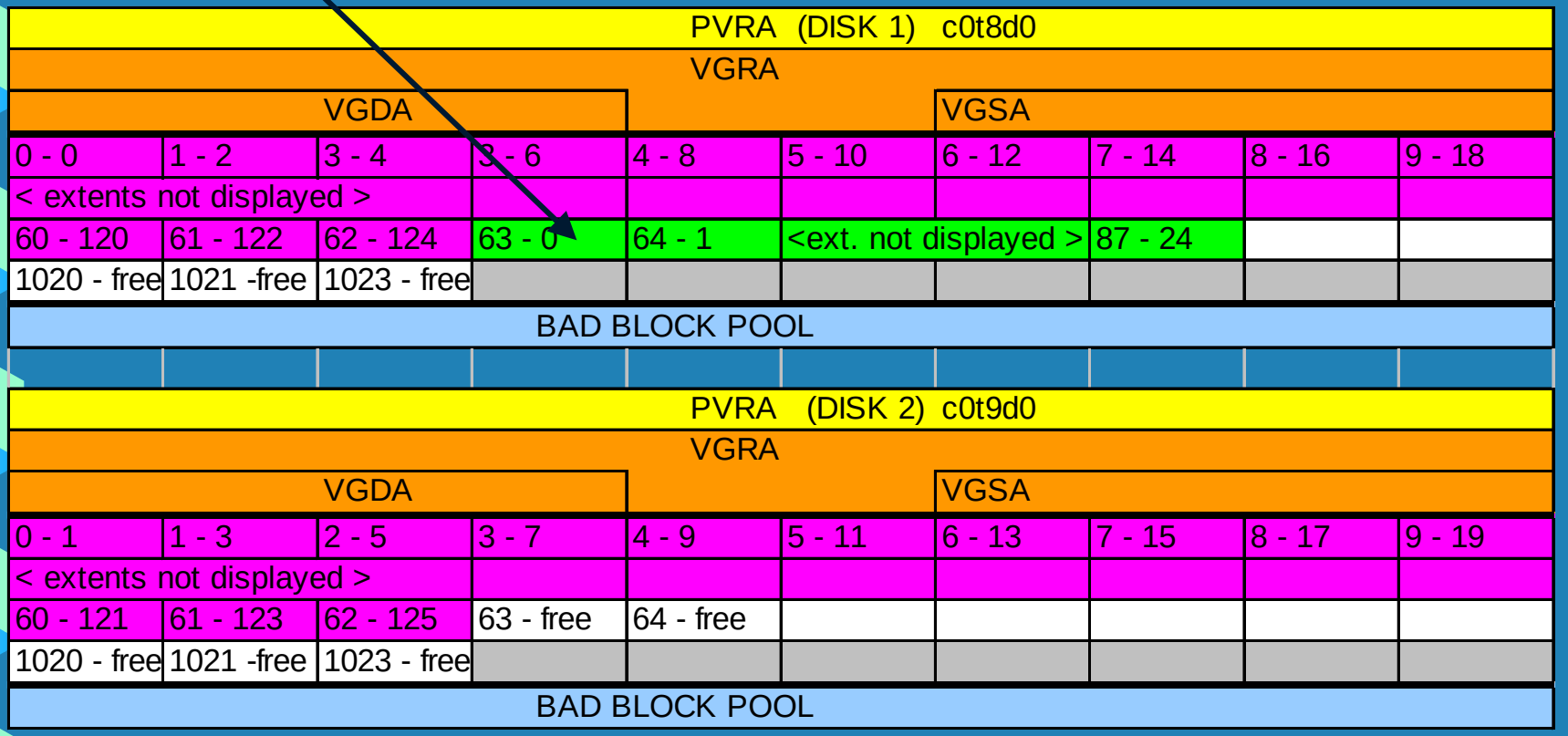
```
#
# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c0t5d0
/dev/dsk/c0t10d0
/dev/vg01
/dev/dsk/c0t8d0
/dev/dsk/c0t9d0
#                                   FROM            TO
# pvmove -n /dev/vg01/lvol2 /dev/dsk/c0t9d0 /dev/dsk/c0t8d0
Transferring logical extents of logical volume "/dev/vg01/lvol2"...
Physical volume "/dev/dsk/c0t9d0" has been successfully moved.
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
#
```

- Move lv to specific disk
- pvmove
- Allows you to move logical volume from one physical disk to another in the SAME volume group

Before PV move



After pvmove



Replace Failed Disk

- Put in new disk
- Run vgcfgrestore command
- Restore data from backup (if not mirroring)

```
#
# ll /etc/lvmconf
total 320
----------    1 root       sys              0 Jun 10  1996 lvm_lock
-rw-------    1 root       root         52224 Apr 23 16:01 vg00.conf
-rw-------    1 root       root         52224 Apr 23 16:01 vg00.conf.old
-rw-------    1 root       sys          29696 Apr 22 16:48 vg01.conf
-rw-------    1 root       sys          29696 Apr 22 16:48 vg01.conf.old
#
# vgcfgrestore -n /dev/vg01 /dev/rdsk/c0t9d0
Volume Group configuration has been restored to /dev/rdsk/c0t9d0
#
```

vgexport / vgimport

- Changing the hardware location of disk
- Add disks back after install
- Be sure to use mapfiles if non standard LV names

HPRC “must have” document

- Get HPRC document # UNX1001086
  - “Procedure for replacing an LVM disk in HP-UX 10.x and 11.x” 17 pages
- Replacing a non-boot disk without LVM-Mir
- Replacing a non-boot disk with LVM-Mir
- Replacing a boot disk without LVM-Mir
- Replacing a boot disk with LVM-Mir
- Preparing for the Recovery of LVM System
  - Includes script for LVM configuration recording

```
# vgchange -a n /dev/vg01
Volume group "/dev/vg01" has been successfully changed.
# vgexport /dev/vg01
# strings /etc/lvmtab
/dev/vg00
/dev/dsk/c0t5d0
/dev/dsk/c0t10d0
# mkdir /dev/vg01
# mknod /dev/vg01/group c 64 0x010000
# vgimport /dev/vg01 /dev/dsk/c1t8d0 /dev/dsk/c1t9d0  (new disk path)
# rning: A backup of this volume group may not exist on this machine.
Please remember to take a backup using the vgcfgbackup command after activatin
the volume group.
# vgchange -a y /dev/vg01
Activated volume group
Volume group "/dev/vg01" has been successfully changed.
# vgcfgbackup /dev/vg01
Volume Group configuration for /dev/vg01 has been saved in /etc/lvmconf/vg01.c
f
```

VG00

- VG00 aka boot disk aka root disk
- The following do not need to be on VG00:
  - /home
  - /opt
  - /var

Root disk - Separate LVs

- INSTEAD of
  - /var
- USE
  - /var/mail
  - /var/spool
  - /var/tmp

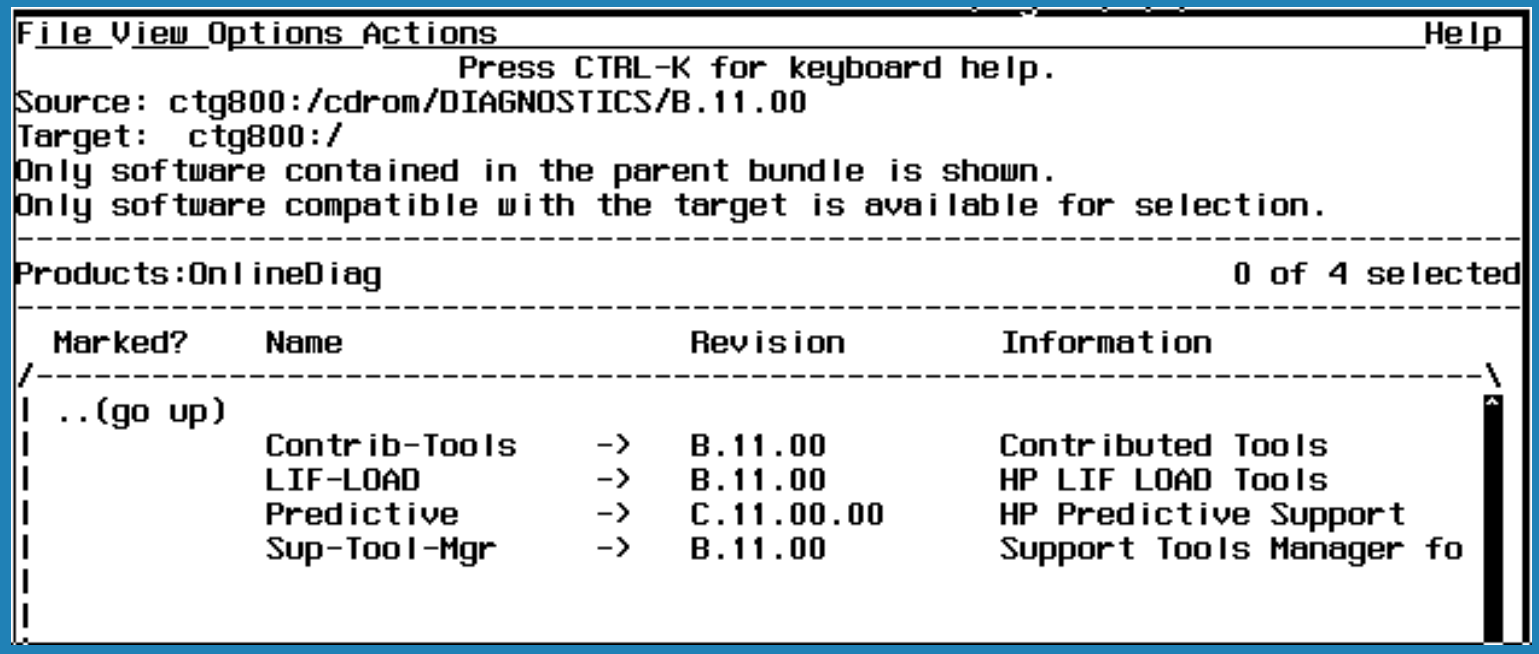
What happens when /var becomes full?

Increase system availability

Protect against mail bombs

Load LIF Tools from Diagnostic CD

- mount -r /dev/dsk/cdrom /cdrom



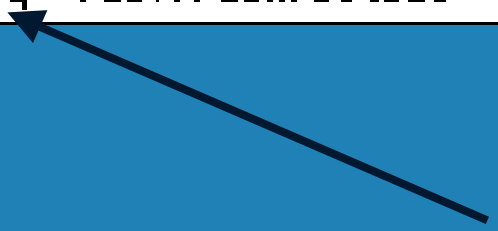
Mirroring

- $$
- Allows replication of disk resource
- System backup with very minimal downtime
- Does not protect against human error
- Mirror per LV not per disk
- 1-3% sw overhead
- Best read performance (up to 40%)
- Performance degradation for writes (down by 10%)
- LVM pseudo-driver

Mirror/UX

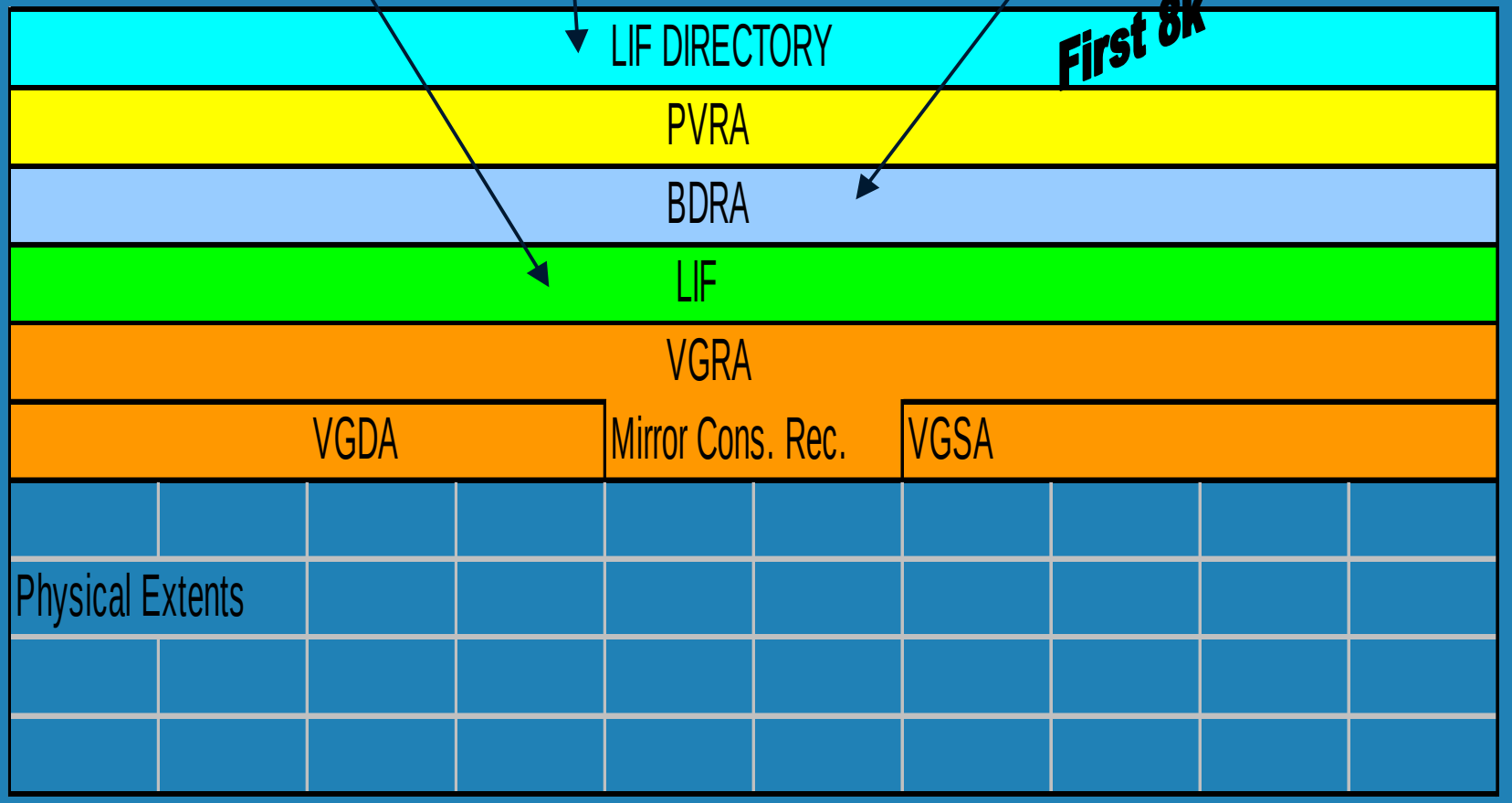
Make bootable PV

```
# pvcreate -B /dev/rdsk/c0t8d0
Physical volume "/dev/rdsk/c0t8d0" has been successfully created.
# vgextend /dev/vg00 /dev/dsk/c0t8d0
Volume group "/dev/vg00" has been successfully extended.
Volume Group configuration for /dev/vg00 has been saved in /etc/lvmco
f
# mkboot /dev/rdsk/c0t8d0
# IF HAVE DIAG:
mkboot -b /usr/sbin/diag/lif/updatediaglif -p ISL -p AUTO \
> -p HPUX -p LABEL /dev/rdsk/c0t8d0
# mkboot -a "hpux -lq" /dev/rdsk/c0t8d0
# mkboot -a "hpux -lq" /dev/rdsk/c0t5d0
```



Most important!

What disks are bootable, Where is the root filesystem
ISL, HPUX, AUTO, LABEL



BDRA

- Primary (sector #128) and secondary (sector #136) boot data record the points to:
  - root, swap and dump
- Primary and secondary PVOL list

Bad Block Relocation Pool on VG00

- Turned off for:
  - Logical Volume: /
  - Logical Volume: /stand
  - Primary swap
  - Any LV used for dump
- Turned on for:
  - Logical Volumes /var, /usr/, /tmp and /opt
- Ensures that LV is contiguous

Mirror each LV

```
# lvextend -m 1 /dev/vg00/lvol1 /dev/dsk/c0t8d0
The newly allocated mirrors are now being synchronized. This operation will
take some time. Please wait ....
Logical volume "/dev/vg00/lvol1" has been successfully extended.
Volume Group configuration for /dev/vg00 has been saved in /etc/lvmconf/vg00.con
f
#
# lvextend -m 1 /dev/vg00/lvol2 /dev/dsk/c0t8d0
The newly allocated mirrors are now being synchronized. This operation will
take some time. Please wait ....
```

Mirror/UX

lvdisplay -v of mirrored LV

```
--- Logical extents ---
LE    PV1                 PE1   Status 1 PV2                 PE2   Status 2
0000  /dev/dsk/c0t5d0     0000  current  /dev/dsk/c0t8d0     0000  current
0001  /dev/dsk/c0t5d0     0001  current  /dev/dsk/c0t8d0     0001  current
0002  /dev/dsk/c0t5d0     0002  current  /dev/dsk/c0t8d0     0002  current
0003  /dev/dsk/c0t5d0     0003  current  /dev/dsk/c0t8d0     0003  current
0004  /dev/dsk/c0t5d0     0004  current  /dev/dsk/c0t8d0     0004  current
0005  /dev/dsk/c0t5d0     0005  current  /dev/dsk/c0t8d0     0005  current
0006  /dev/dsk/c0t5d0     0006  current  /dev/dsk/c0t8d0     0006  current
0007  /dev/dsk/c0t5d0     0007  current  /dev/dsk/c0t8d0     0007  current
0008  /dev/dsk/c0t5d0     0008  current  /dev/dsk/c0t8d0     0008  current
0009  /dev/dsk/c0t5d0     0009  current  /dev/dsk/c0t8d0     0009  current
0010  /dev/dsk/c0t5d0     0010  current  /dev/dsk/c0t8d0     0010  current
0011  /dev/dsk/c0t5d0     0011  current  /dev/dsk/c0t8d0     0011  current
0012  /dev/dsk/c0t5d0     0012  current  /dev/dsk/c0t8d0     0012  current
0013  /dev/dsk/c0t5d0     0013  current  /dev/dsk/c0t8d0     0013  current
0014  /dev/dsk/c0t5d0     0014  current  /dev/dsk/c0t8d0     0014  current
0015  /dev/dsk/c0t5d0     0015  current  /dev/dsk/c0t8d0     0015  current
0016  /dev/dsk/c0t5d0     0016  current  /dev/dsk/c0t8d0     0016  current
```

Mirror/UX

Disk layout - mirroring

PVRA (DISK 1)									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	< extents not displayed >			
1020 - free	1021 -free	1023 - free							

PVRA (DISK 2)									
VGRA									
VGDA						VGSA			
0 - 0	1 -1	2 - 2	3 - 3	4 - 4	5 - 5	6 - 6	7 - 7	8 - 8	9 - 9
10 - 10	11 - 11	12 - 12	13 - 13	14 - 14	15 - 15	16 - 16	17 - 17	18 - 18	19 - 19
20 - 20	21 - 21	22 - 22	23 - 23	24 - 24	25 - 25	< extents not displayed >			
1020 - free	1021 -free	1023 - free							

Consistency Recover Policy

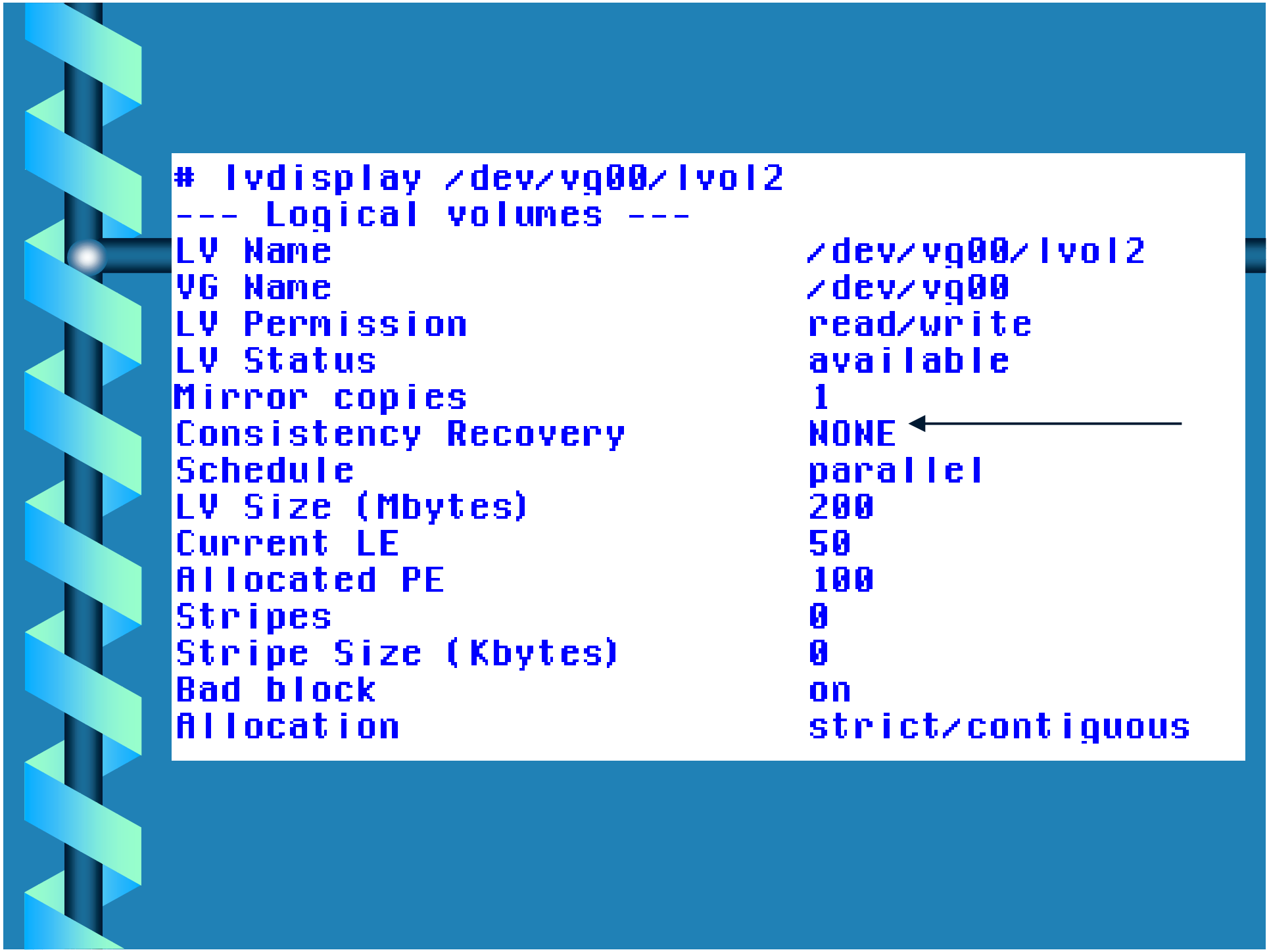
- MWC (Mirror Write Cache]
- No need to mirror swap data
- Turn off (on swap only) to increase performance
- Create more small VGs than fewer large VGs (MWC misses)
- Kept in LTG (Logical Track Group) in VGRA

```
#
# lvdisplay /dev/vg00/lvol2
--- Logical volumes ---
LV Name                     /dev/vg00/lvol2
VG Name                     /dev/vg00
LV Permission               read/write
LV Status                   available/syncd
Mirror copies               1
Consistency Recovery        MWC
Schedule                    parallel
LV Size (Mbytes)            200
Current LE                  50
Allocated PE                100
Stripes                     0
Stripe Size (Kbytes)        0
Bad block                   on
Allocation                  strict/contiguous
```

Mirror/UX

Must do in Maintenance Mode

- Shutdown -r
- Interupt boot process
- Enter Command> boot pri isl
- Interact with IPL?> Y
- ISL> hpux -lm
- vgchange -a y vg00
- lvchange -M n -c n /dev/vg00/lvol2
- lvlnboot -s /dev/vg00/lvol2
- lvlnboot -d /dev/vg00/lvol2
- reboot (DO NOT DO INIT)
- Boot from the alternate disk to test!


lvdisplay /dev/vg00/lvol2
--- Logical volumes ---
LV Name /dev/vg00/lvol2
VG Name /dev/vg00
LV Permission read/write
LV Status available
Mirror copies 1
Consistency Recovery NONE
Schedule parallel
LV Size (Mbytes) 200
Current LE 50
Allocated PE 100
Stripes 0
Stripe Size (Kbytes) 0
Bad block on
Allocation strict/contiguous

lvlnboot (root, boot, primary swap, dump)

```
# lvlnboot -v vg00
Boot Definitions for Volume Group /dev/vg00:
Physical Volumes belonging in Root Volume Group:
        /dev/dsk/c0t5d0 (8/4.5.0) -- Boot Disk
        /dev/dsk/c0t8d0 (8/4.8.0) -- Boot Disk
Boot:  lvol1      on:      /dev/dsk/c0t5d0
Root:  lvol3      on:      /dev/dsk/c0t5d0
Swap:  lvol2      on:      /dev/dsk/c0t5d0
Dump:  lvol2      on:      /dev/dsk/c0t5d0, 0
```

Quorum

- 51% + = quorum
- Only activate VG if quorum
- No Quorum
  - ISL> hpux -lq /stand/vmunix
- No Quorum & Single User Mode
  - ISL> hpux -is lq /stand/vmunix

Hot Spare for mirror (HP-UX 10.3+)

- Extra disk in the VG used as standby
- Every LV in the VG must be mirrored
- Every mirror must use strict allocation
- No hot spare for VG00 (won't move boot blocks)
  - Can set up, but won't use spare

```
# vgextend -z y vg01 /dev/dsk/c0t2d0
Volume group "vg01" has been successfully extended.
Volume Group configuration for /dev/vg01 has been saved
f
# vgdisplay /dev/vg01 | grep Spare
Total Spare PVs               1
Total Spare PVs in use        0
```

Mirror/UX

Hot Spare - When PV fails

```
PV Name                     /dev/dsk/c0t2d0
PV Status                   available/standby spare
Total PE                    2000
Free PE                     2000
```

- PF of > 5 minutes or I/O error encountered
- Status will change to available/active spare
- Status of failed PV: unavailable/data spared
- Check status using vgdisplay -v
- Takes more than 5 minutes

Mirror/UX

Replacing Failed Disk

- c0t2d0 (spare) c0t3d0 (failed)
- vgcfgrestore -n vg01 /dev/rdsk/c0t3d0
- vgchange -a y vg01
- pvchange -z y /dev/dsk/c0t3d0
- pvmove /dev/dsk/c0t2d0 /dev/dsk/c0t3d0

Mirror/UX

Other hot spare notes

- Can have multiple spares
- 1 spare disk replaces one failed PV
  - (9GB spare disk can't replace 2 failed 4 GB PVs)
- Spares are not included in "free"
- Spares can not be shared across VGs
- Spare should be as big as largest PV
- Non-clustered VGs

Mirror/UX

Mirroring & Striping

- Not supported on standard LVM
- Extent based striping is supported
  - aka pseudo-striped

```
# lvcreate
Usage: lvcreate
         [-A Autobackup]
         [-d Schedule]
         [-i Stripes -I StripeSize]
         {-l LogicalExtentsNumber |
          -L LogicalVolumeSize}
         [-m MirrorCopies]
         [-n LogicalVolumeName]
         [-p Permission]
         [-r Relocate]
         [-s Strict]
         [-C Contiguous]
         [-D Distributed]
         [-M MirrorWriteCache]
         [-c MirrorConsistency]
         VolumeGroupName
```

If don't have options, apply patches

Mirror/UX

/etc/lvmpvg

```
VG /dev/vg01
PVG PV0
/dev/dsk/c0t8d0
/dev/dsk/c0t9d0
PVG PV1
/dev/dsk/c1t2d0
/dev/dsk/c1t3d0
```

- Physical Volume Group
- PVG Strict
- # lvcreate -L 120 -D y -s g -m 1 vg01

Logical Volume Timeout

- I/O to a non-responsive disk will only be retried for the specific amount of time
- I/O "hung" state won't last longer than timeout

```
# lvchange -t 120 /dev/vg01/lvol5
Logical volume "/dev/vg01/lvol5" has been successfully changed.
Volume Group configuration for /dev/vg01 has been saved in /etc/
f
# lvdisplay /dev/vg01/lvol5
--- Logical volumes ---
LV Name                     /dev/vg01/lvol5
VG Name                     /dev/vg01
# _Permission               read/write
LV Status                   available/syncd
Mirror copies               0
Consistency Recovery        MWC
Schedule                    parallel
LV Size (Mbytes)            20
Current LE                  5
Allocated PE                5
Stripes                     0
Stripe Size (Kbytes)        0
Bad block                   on
Allocation                  strict
IO Timeout (Seconds)        120
```

Length of time I/O will be retried (in seconds)

OnLine JFS

- $$
- Online backup [snapshot]
- Online fs increase and reduce
- Defragmentation [fsadm]
- Patches

OnLine JFS

Table 1. VxFS File System Feature Comparison

Feature	JFS 3.3	OnLineJFS 3.3
extent-based allocation	*	*
extent attributes	*	*
fast file system recovery	*	*
access control list (ACL) support	*	*
enhanced application interface	*	*
enhanced mount options	*	*
improved synchronous write performance	*	*
support for large files (up to one terabyte)	*	*
support for large file systems (up to one terabyte)	*	*
enhanced I/O performance	*	*
support for BSD-style quotas	*	*
unlimited number of inodes	*	*
file system tuning [`vxtunefs(1M)`]	*	*
online administration		*
ability to reserve space for a file and set fixed extent sizes and allocation flags		*
online snapshot file system for backup		*
data synchronous I/O		*
DMAPI		*

Snapshot backup (Online JFS)

```
/dev/vg00/lvol4       20480    1119   18155    6% /home
```

```
# pwd
/home/chris
# ll
total 18
-rw-r--r--    1 chris      users            814 Jan  4 14:13 .cshrc
-rw-r--r--    1 chris      users            347 Jan  4 14:13 .exrc
-rw-r--r--    1 chris      users            341 Jan  4 14:13 .login
-rw-r--r--    1 chris      users            446 Jan  4 14:13 .profile
-rw-r--r--    1 chris      users            512 Jan  4 14:12 apple
-rw-r--r--    1 chris      users            659 Jan  4 14:14 banana
-rw-r--r--    1 chris      users            512 Jan  4 14:12 hosts
-rw-r--r--    1 chris      users            702 Jan  4 14:14 orange
-rw-r--r--    1 chris      users            805 Jan  4 14:14 pear
#
```

OnLine JFS

Create separate LV
Mount using snapof option

```
#
# lvcreate -L 20 -n homesnap /dev/vg00
Logical volume "/dev/vg00/homesnap" has been successfully created with
character device "/dev/vg00/rhomesnap".
Logical volume "/dev/vg00/homesnap" has been successfully extended.
Volume Group configuration for /dev/vg00 has been saved in /etc/lvmconf/vg00.con
f
```

```
# mkdir /homesnap
# mount -F vxfs -o snapof=/dev/vg00/lvol4 /dev/vg00/homesnap /homesnap
#
# bdf
Filesystem          kbytes    used   avail %used Mounted on
/dev/vg00/lvol3      86016   16076   65528   20% /
/dev/vg00/lvol1      47829   16696   26350   39% /stand
/dev/vg00/lvol7     163840    9044  145007    6% /var
/dev/vg00/lvol9      20480    1388   17953    7% /var/spool
/dev/vg00/lvol8      20480    1111   18165    6% /var/mail
/dev/vg00/lvol6     491520  227225  247738   48% /usr
/dev/vg00/lvol5      90112    1140   83404    1% /tmp
/dev/vg01/lvol1    2093056  130627 1839830    7% /opt
/dev/vg00/lvol4      20480    1119   18155    6% /home
/dev/vg00/lvol10     81920    1139   75733    1% /var/tmp
/dev/vg00/homesnap   20480    1119   18150    6% /homesnap
#
```

OnLine JFS

Snapshot file system will not change (read only FS)

```
# diff /homesnap/chris /home/chris
# rm /home/chris/pear
# diff /homesnap/chris /home/chris
Only in /homesnap/chris: pear
# cp /homesnap/chris/pear /home/chris/pear
#
```

```
# cp /stand/vmunix /home/chris/kern
#
# bdf | grep home
/dev/vg00/lvol4       20480     9449    10339    48% /home
/dev/vg00/homesnap    20480     1119    18150     6% /homesnap
#
```

OnLine JFS

Snapshot File System

- Must have OnLine JFS
- FS must be JFS
- fbackup will not work (does not support read only FS). Most other backup utilites/software will work
- vxdump (replaces dump - HFS only)
- Only previous version of block is copied to snapshot when block is modified. LV only requires enough for changed blocks. (10-20%)

Defragmentation - Directories

```
#
# fsadm -F vxfs -D -d /usr

  Directory Fragmentation Report
             Dirs        Total      Immed    Immeds   Dirs to   Blocks to
             Searched    Blocks     Dirs     to Add   Reduce    Reduce
  total           806       546      471          1        5          90


  Directory Fragmentation Report
             Dirs        Total      Immed    Immeds   Dirs to   Blocks to
             Searched    Blocks     Dirs     to Add   Reduce    Reduce
  total           806       545      472          0        4          89

#
```

OnLine JFS

Defragmentation - Extents

```
# fsadm -F vxfs -e -E /opt
  Extent Fragmentation Report
        Total    Average      Average     Total
        Files    File Blks    # Extents   Free Blks
         1230         104            1     1962429
    blocks used for indirects: 0
    % Free blocks in extents smaller than 64 blks: 1.53
    % Free blocks in extents smaller than  8 blks: 0.10
    % blks allocated to extents 64 blks or larger: 92.39
    Free Extents By Size
        1:          3     2:          5     4:          4     8:          6
       16:         11    32:          6    64:          4   128:          2
      256:          0   512:          1  1024:          1  2048:          1
     4096:          0  8192:          1 16384:          1 32768:          1
```

OnLine JFS

Defragmentation

- Directory
  - Reordered to place subdirectories 1st, then all other entries in decreasing order by time of last access. Compacted to remove free space
- Extents
  - Aged files are moved to the end of the allocation units. Files are reorganized to have the least amount of possible extents.
- fsadm -F vxfs -d -D -e -E /home
- cron

OnLine JFS

Increase Logical Volume & FS

```
# lvextend -L 44 /dev/vg00/lvol8
Logical volume "/dev/vg00/lvol8" has been successfully extended.
Volume Group configuration for /dev/vg00 has been saved in /etc/lvmconf/vg00.con
f
# extendfs /dev/vg00/rlvol8
vxfs extendfs: /dev/vg00/rlvol8 is mounted, cannot extend.
#
# fsadm -F vxfs -b 45056 /var/mail
fsadm: /dev/vg00/rlvol8 is currently 20480 sectors - size will be increased
```

- # of MB x 1024 = blocks (44 x 1024 = 45056)
- Without OnLine JFS, must umount the file system before using extendfs
- Can not resize LV if using snapshot

OnLine JFS

Reduce Logical Volume

- Defrag before reducing
- JFS 3.3+, attempts to move extents off the area you want to shrink

```
# fsadm -F vxfs -d -e /usr
# fsadm -F vxfs -b 512000 /usr
fsadm: /dev/vg00/rlvol6 is currently 614400 sectors - size will be reduced
# lvreduce -L 500 /dev/vg00/lvol6
When a logical volume is reduced useful data might get lost;
do you really want the command to proceed (y/n) : y
Logical volume "/dev/vg00/lvol6" has been successfully reduced.
Volume Group configuration for /dev/vg00 has been saved in /etc/lvmconf/vg00.con
f
```

OnLine JFS

AutoRAID

- Auto ”magically” switches between RAID levels 0/1 (mirroring & striping) and RAID level 5 (data protection by parity)
- Most actively written data is stored in RAID 0/1. (Never less than 10%)
- No control of data placement (APT)

AutoRAID with 2 20GB PVs

Disk Devices 0 of 30 select

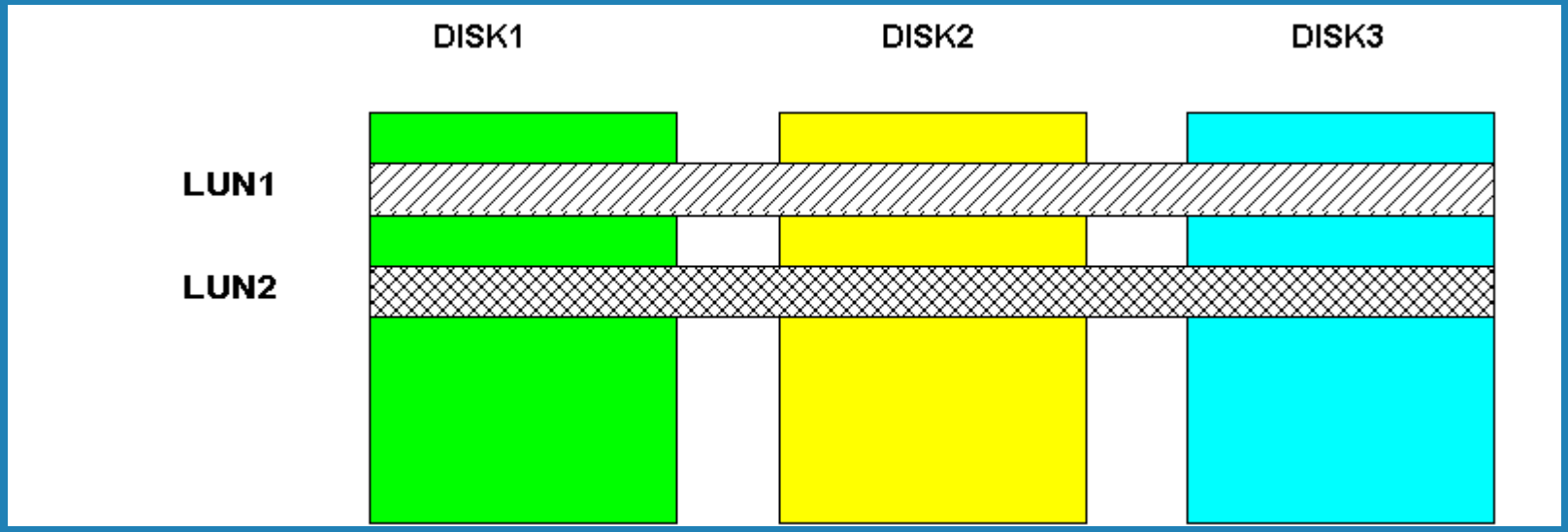
Hardware Path	Number of Paths	Use	Volume Group	Total Mbytes	Description
16.0	2	--	--	--	HP AutoRAID Disk Array
16.0.0	2	LVM	vg08	20000	HP AutoRAID LUN
16.0.1	2	LVM	vg09	20000	HP AutoRAID LUN

AutoRAID

LUN - Logical Drive

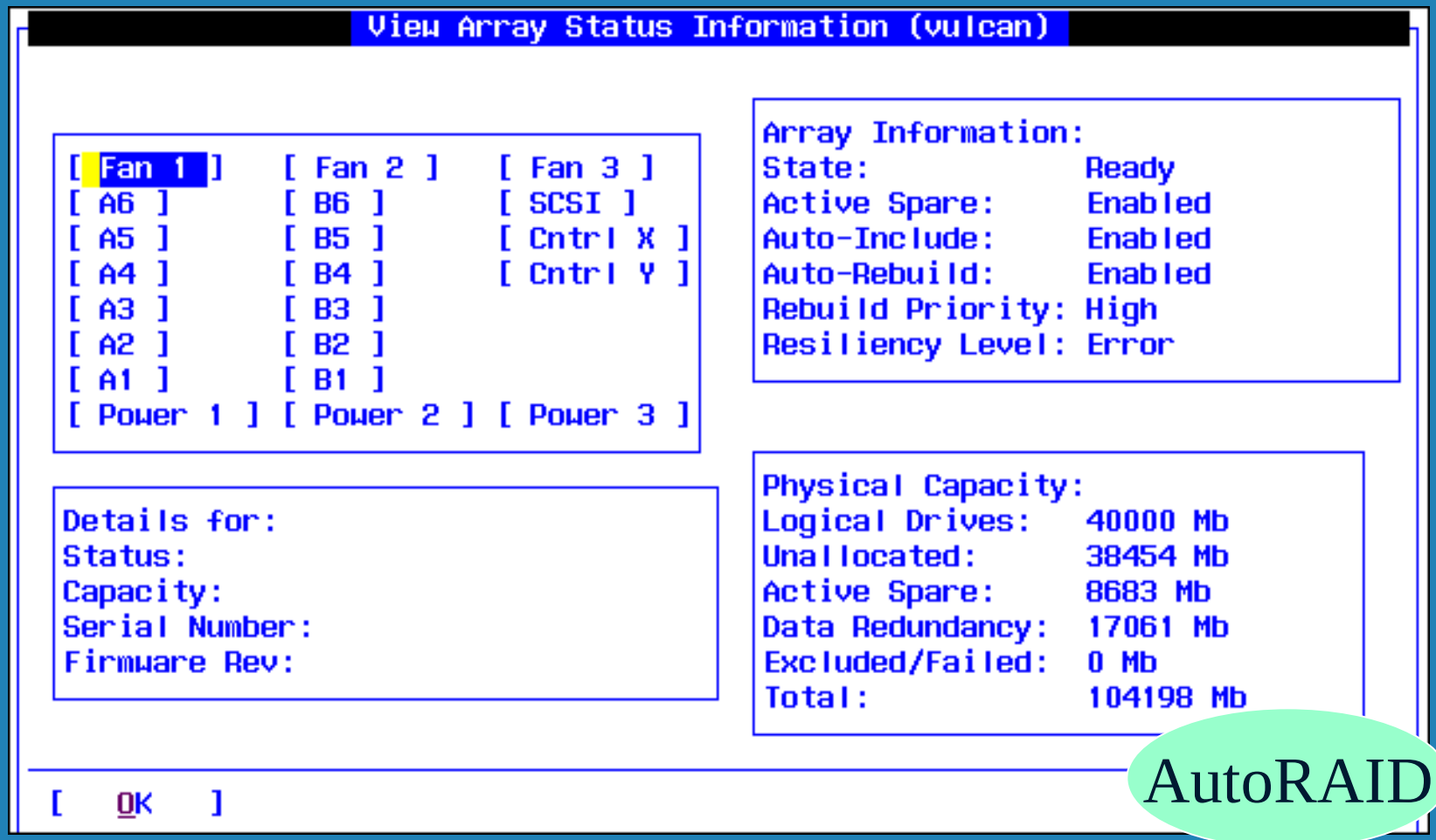
- Up to 8 LUNs
- HP-UX looks at a LUN as a physical individual disk
- LUNs do NOT correspond to individual disks in the AutoRAID
- LUN = Physical Volume in LVM
- Physical Volume(s) = Volume Group

LUN is created across all disks



AutoRAID

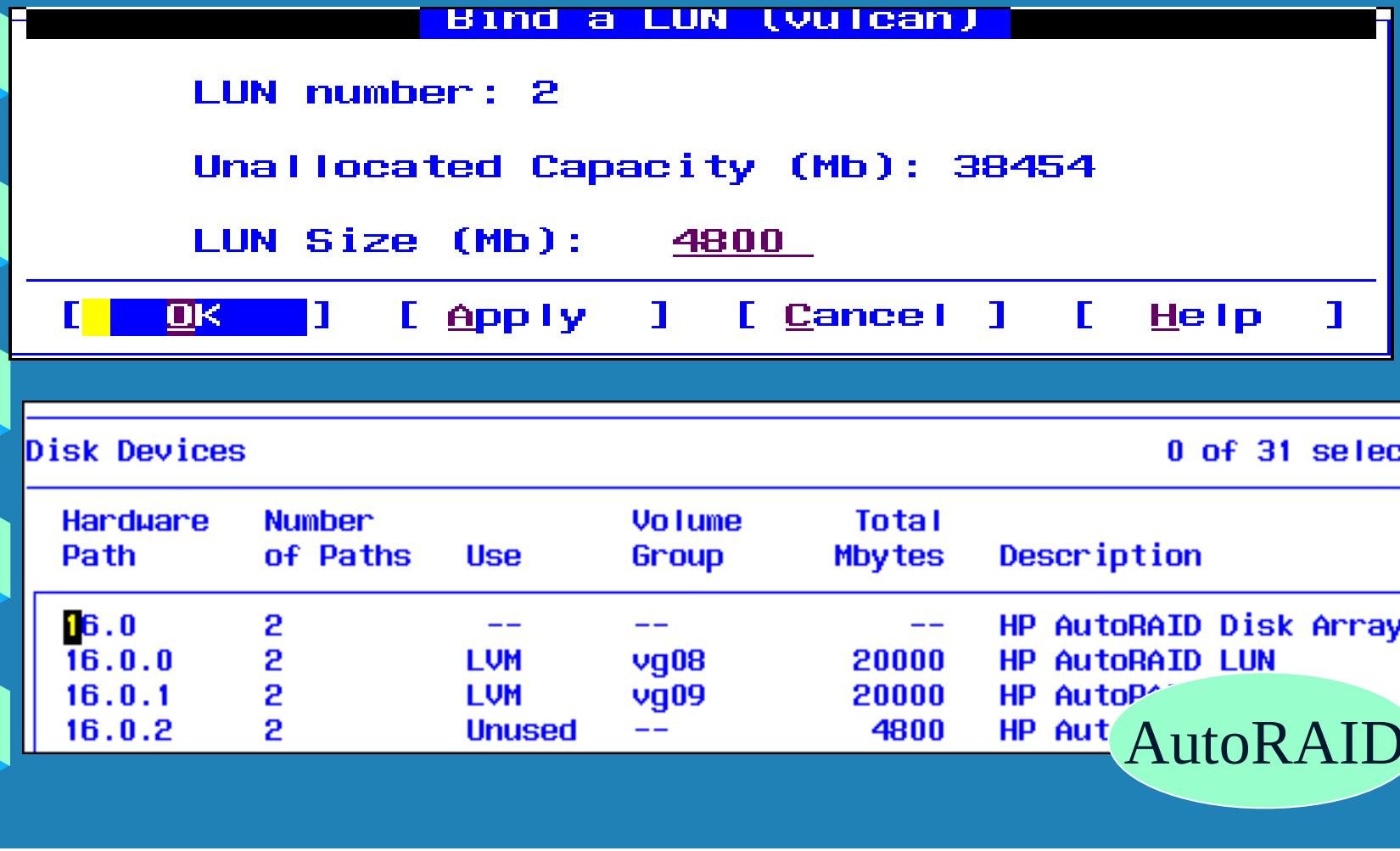
Disk space allocations



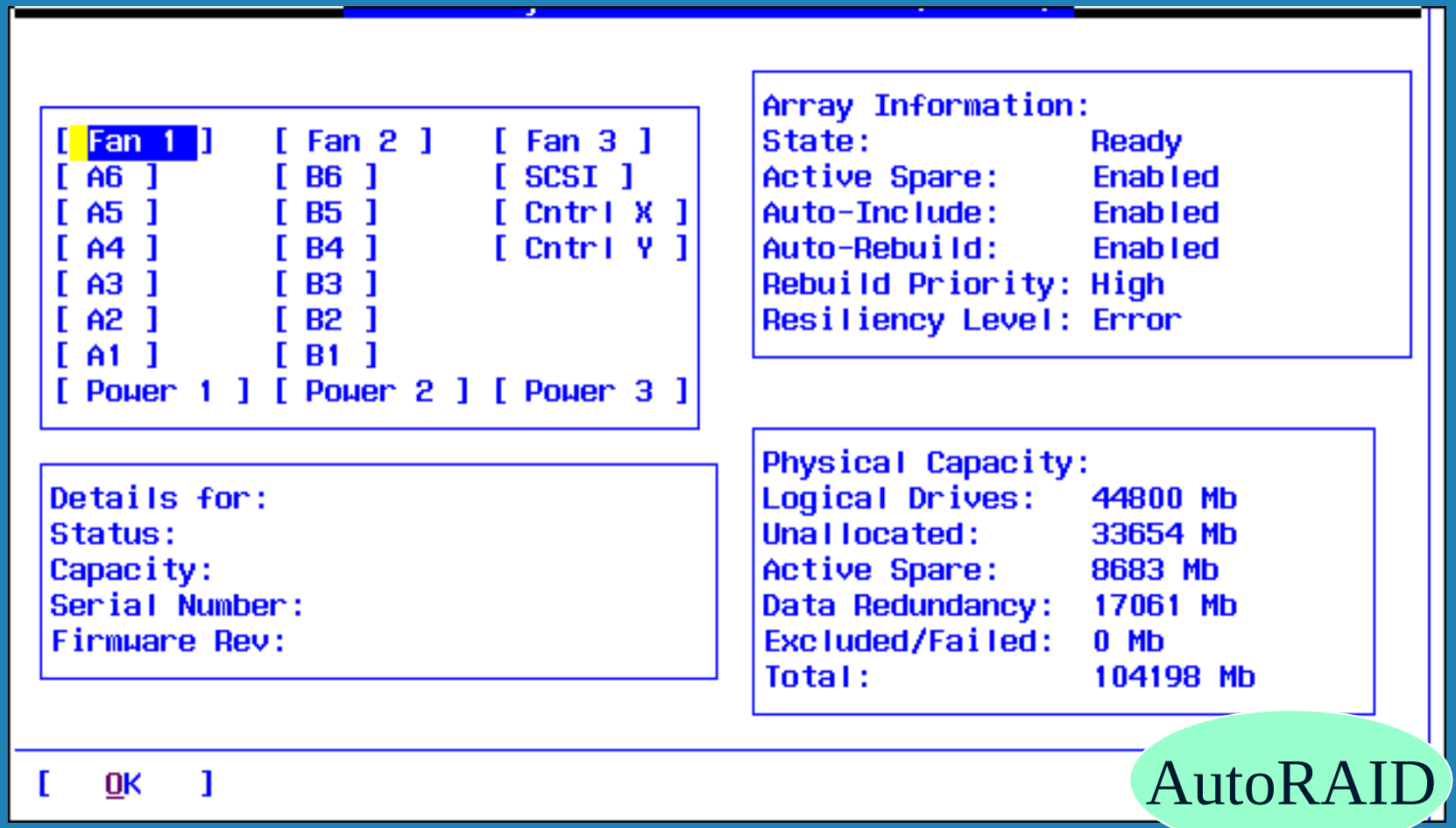
Add VG to AutoRAID
Need disk (PV) to create VG

AutoRAID

Created 4.8 GB LUN

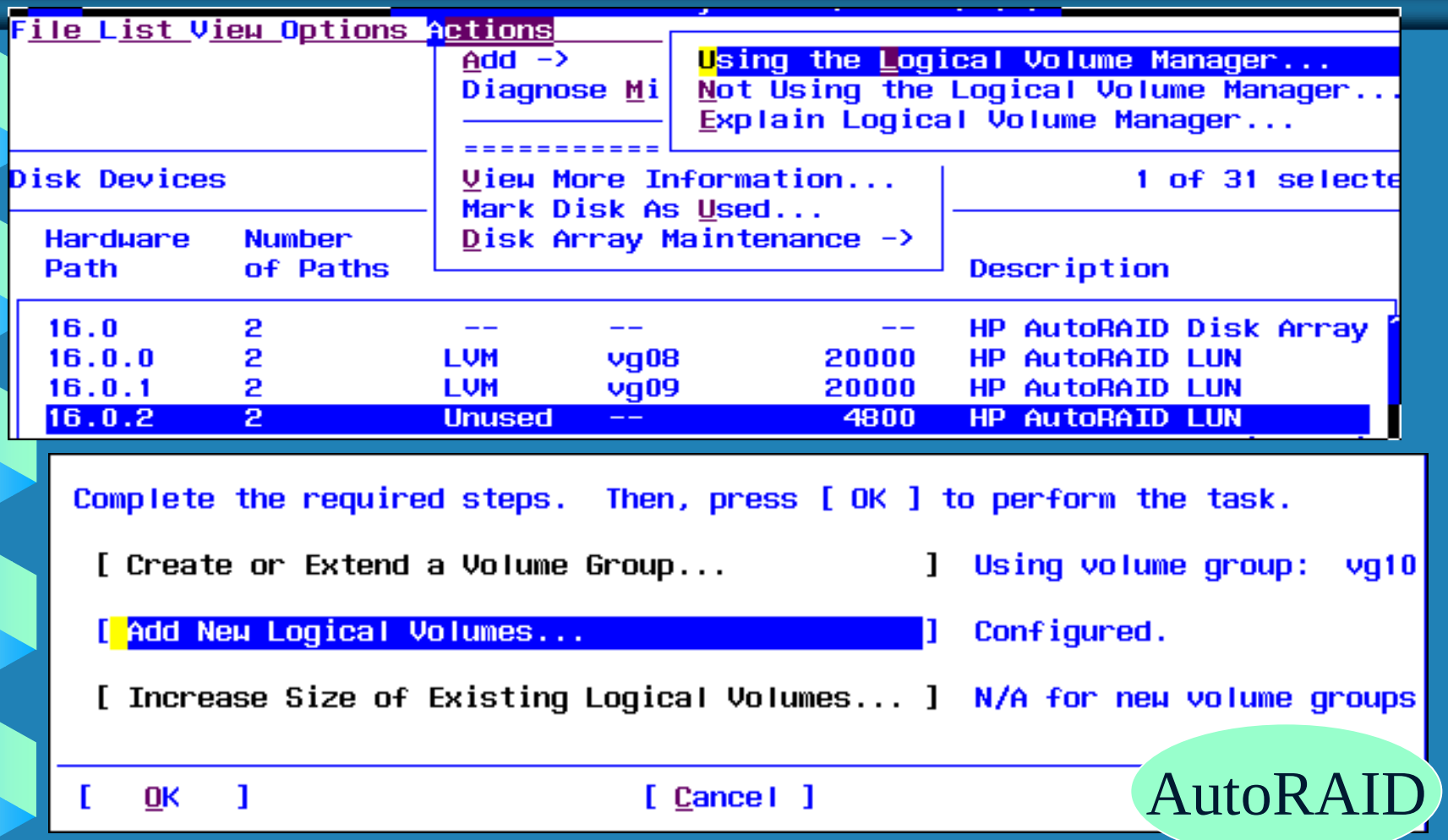


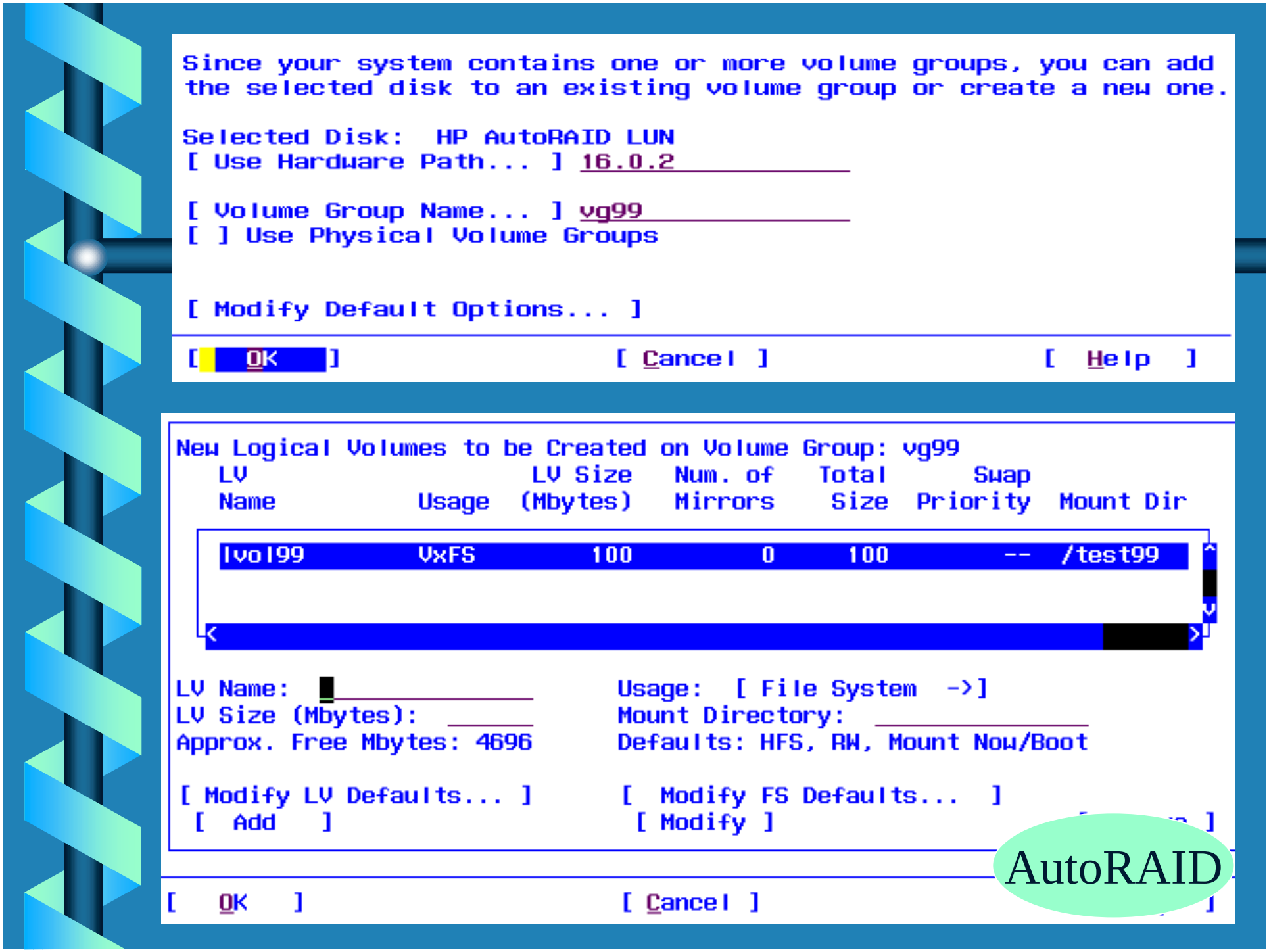
Logical drives increased by 4.8GB, Unallocated decreased

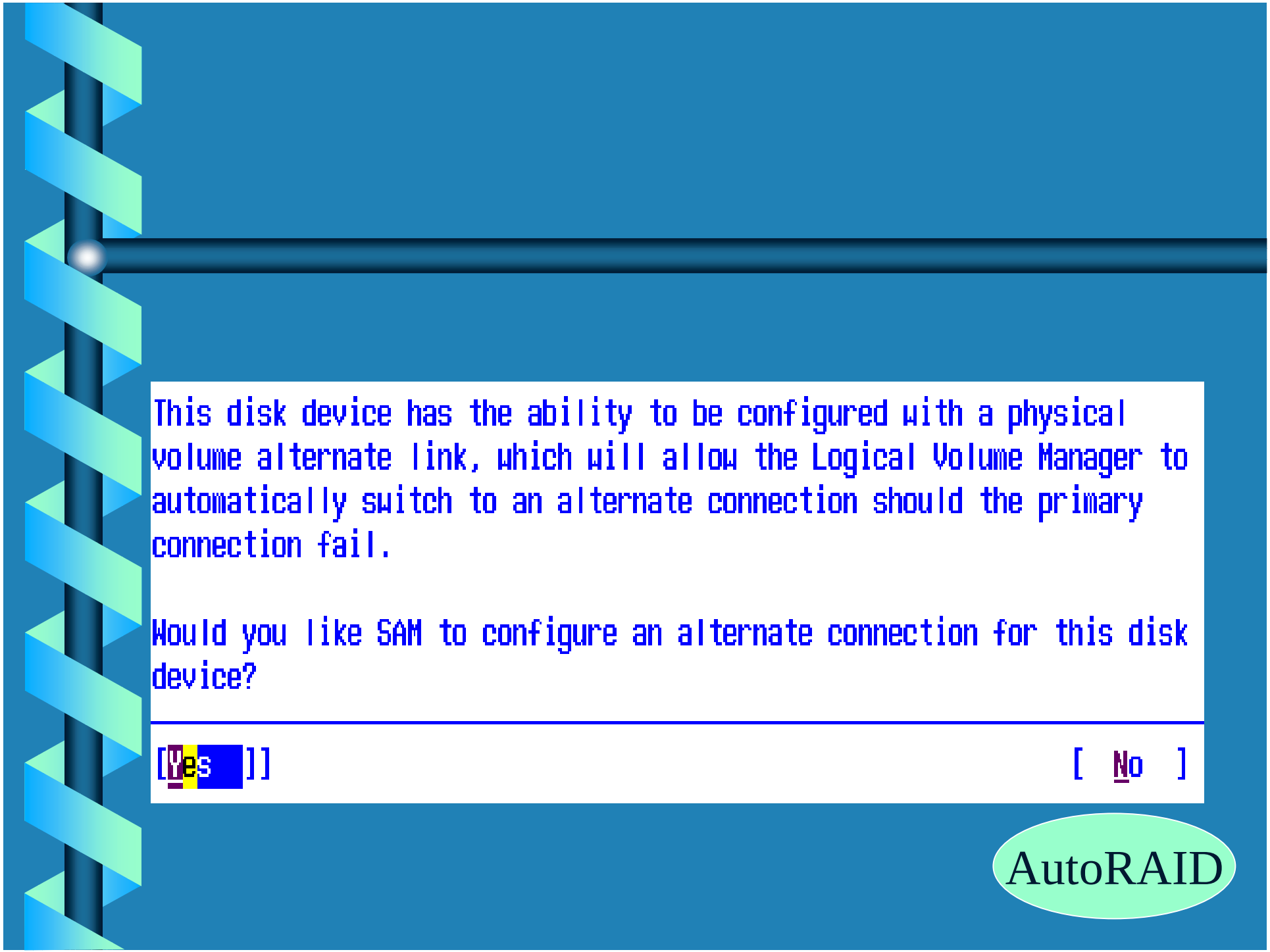


AutoRAID

Create VG and LVOL




Since your system contains one or more volume groups, you can add
the selected disk to an existing volume group or create a new one.
Selected Disk: HP AutoRAID LUN
[Use Hardware Path...] 16.0.2
[Volume Group Name...] vg99
[ ] Use Physical Volume Groups
[Modify Default Options...]
[OK]
[Cancel]
[Help]
New Logical Volumes to be Created on Volume Group: vg99
LV Name
Usage
LV Size (Mbytes)
Num. of Mirrors
Total Size
Swap Priority
Mount Dir
lvol99 VxFS 100 0 100 -- /test99
LV Name:
LV Size (Mbytes):
Approx. Free Mbytes: 4696
Usage: [File System ->]
Mount Directory:
Defaults: HFS, RW, Mount Now/Boot
[Modify LV Defaults...]
[Add]
[Modify FS Defaults...]
[Modify]
[OK]
[Cancel]
AutoRAID


This disk device has the ability to be configured with a physical volume alternate link, which will allow the Logical Volume Manager to automatically switch to an alternate connection should the primary connection fail.
Would you like SAM to configure an alternate connection for this disk device?
[Yes]]
[No]
AutoRAID

```
--- Logical volumes ---
LV Name                     /dev/vg99/lvol99
LV Status                   available/syncd
LV Size (Mbytes)            100
Current LE                  25
Allocated PE                25
Used PV                     1


--- Physical volumes ---
PV Name                     /dev/dsk/c5t0d2
PV Name                     /dev/dsk/c5t1d2  Alternate Link
PV Status                   available
Total PE                    1199
Free PE                     1174
```

AutoRAID

Everything is the same - device file just points to the LUN

```
   --- Distribution of logical volume ---
   PV Name                  LE on PV  PE on PV
   /dev/dsk/c5t0d2          25        25

   --- Logical extents ---
   LE    PV1                        PE1   Status 1
   0000  /dev/dsk/c5t0d2            0000  current
   0001  /dev/dsk/c5t0d2            0001  current
   0002  /dev/dsk/c5t0d2            0002  current
   0003  /dev/dsk/c5t0d2            0003  current
   0004  /dev/dsk/c5t0d2            0004  current
<removed>
   0021  /dev/dsk/c5t0d2            0021  current
   0022  /dev/dsk/c5t0d2            0022  current
   0023  /dev/dsk/c5t0d2            0023  current
   0024  /dev/dsk/c5t0d2            0024  current
<end>
```

AutoRAID

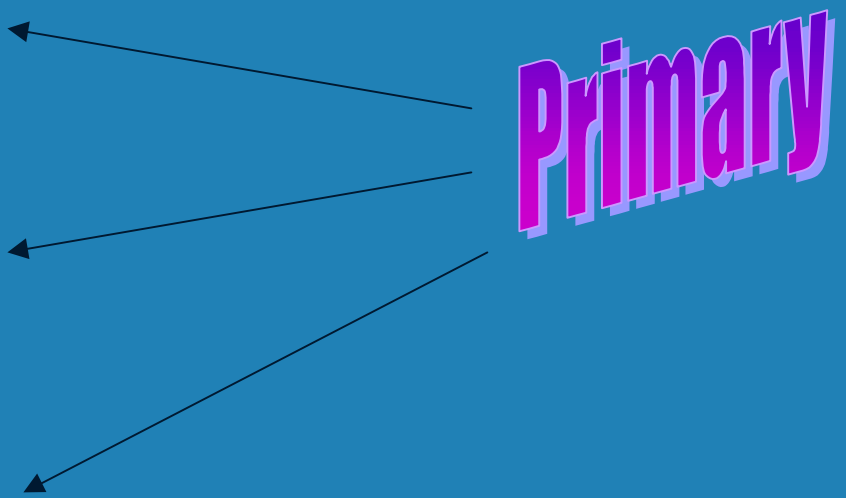
Primary / Alternate

```
strings /etc/lvmtab
/dev/vg01
/dev/dsk/c3t1d0
/dev/dsk/c2t1d0
/dev/dsk/c2t5d0
/dev/vg00
/dev/dsk/c4t6d0
/dev/vg08               (AutoRAID)
/dev/dsk/c5t0d0
/dev/dsk/c5t1d0
/dev/dsk/c5t0d3
/dev/dsk/c5t1d3
/dev/vg09               (AutoRAID)
/dev/dsk/c5t0d1
/dev/dsk/c5t1d1
/dev/vg99               (AutoRAID)
/dev/dsk/c5t0d2
/dev/dsk/c5t1d2
~
```

AutoRAID

Primary is always the first listed in /etc/lvmtab
Alternate is ONLY used for failover

- /dev/vg02
- /dev/dsk/c6t0d0
- /dev/dsk/c7t1d0
- /dev/vg03
- /dev/dsk/c6t0d1
- /dev/dsk/c7t1d1
- /dev/vg04
- /dev/dsk/c6t0d2
- /dev/dsk/c7t1d2
- vgdisplay -v
- --- Physical volumes ---
- PV Name /dev/dsk/c6t0d0
- PV Name /dev/dsk/c7t1d0 Alternate Link



AutoRAID

Use both controllers - Temporary

- pvchange -s
  - Temporary
- /dev/vg02
- /dev/dsk/c6t0d0
- /dev/dsk/c7t1d0
  - pvchange -s /dev/dsk/c7t1d0
- Primary for vg02 is now c7
- Alternate is now c6

AutoRAID

Use both controllers - Permanent

- vgreduce
  - Permanent
- /dev/vg02
- /dev/dsk/c6t0d0
- /dev/dsk/c7t1d0
  - vgreduce /dev/vg02 /dev/dsk/c6t0d0
- c7 becomes Primary (and only)
  - vgextend /dev/vg02 /dev/dsk/c6t0d0
- c6 becomes Alternate
- Changes /etc/lvmtab

AutoRAID

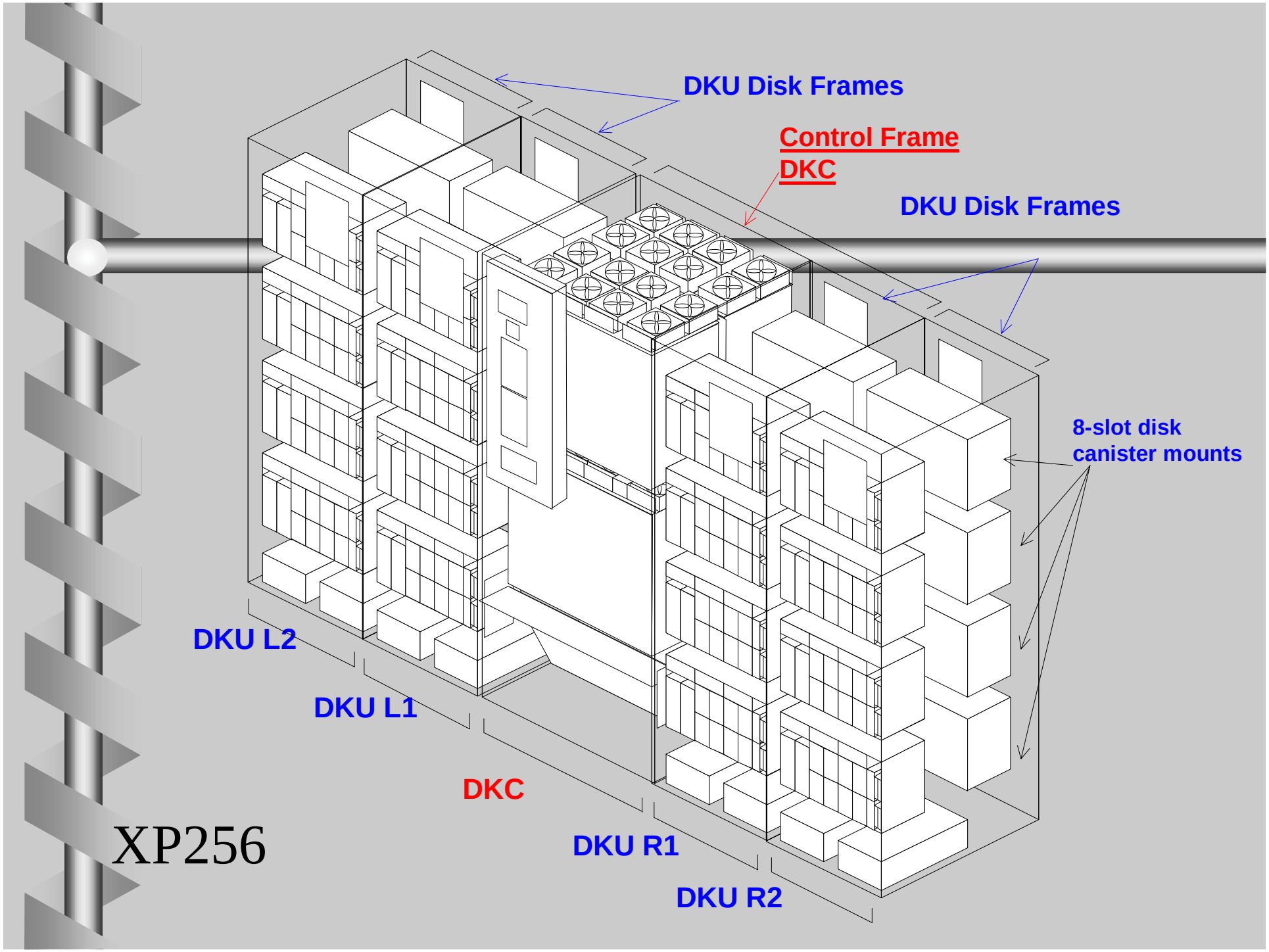
AutoRAID and performance

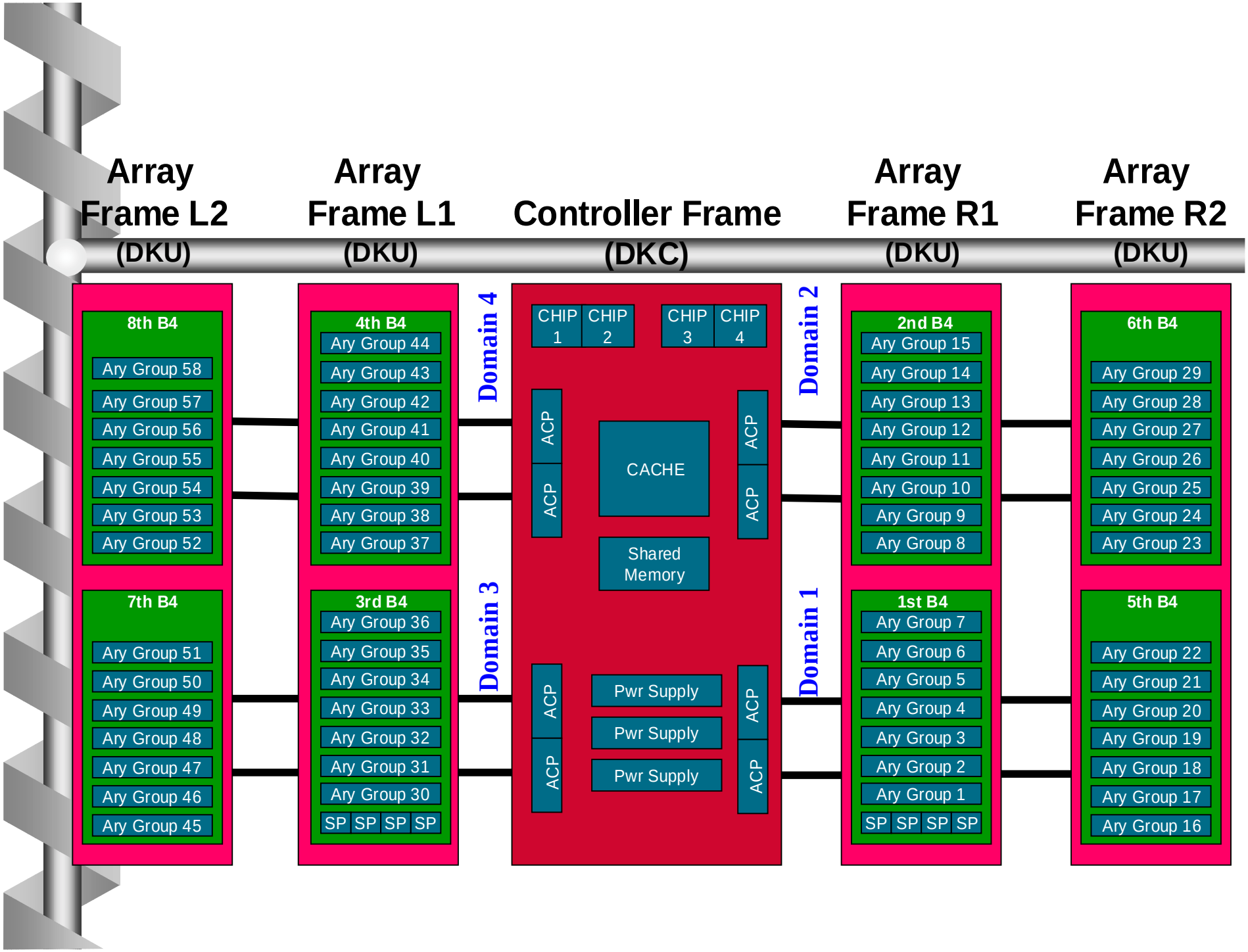
- Use both controllers
- Create 4-6 LUNs
  - More LUNs increases the size of the I/O command queue
- Allow more RAID 0/1

AutoRAID

LVM - AutoRAID and XP256

- LVM commands are all the same
- HP-UX sees the LUN or LDEV as a disk
- Doesn't care or know that it may be only a part of a disk or a portion of many disks
- Limitation of 120 devices per controller
- Using pvlinks? Really only 60 per (60 primary + 60 alternate)


DKU Disk Frames
Control Frame
DKC
DKU Disk Frames
8-slot disk
canister mounts
DKU L2
DKU L1
DKC
DKU R1
DKU R2
XP256


Array Frame L2 (DKU)
Array Frame L1 (DKU)
Controller Frame (DKC)
Array Frame R1 (DKU)
Array Frame R2 (DKU)
8th B4
Ary Group 58
Ary Group 57
Ary Group 56
Ary Group 55
Ary Group 54
Ary Group 53
Ary Group 52
7th B4
Ary Group 51
Ary Group 50
Ary Group 49
Ary Group 48
Ary Group 47
Ary Group 46
Ary Group 45
4th B4
Ary Group 44
Ary Group 43
Ary Group 42
Ary Group 41
Ary Group 40
Ary Group 39
Ary Group 38
Ary Group 37
3rd B4
Ary Group 36
Ary Group 35
Ary Group 34
Ary Group 33
Ary Group 32
Ary Group 31
Ary Group 30
SP SP SP SP
Domain 4
Domain 3
CHIP 1
CHIP 2
CHIP 3
CHIP 4
ACP
ACP
ACP
ACP
CACHE
Shared Memory
ACP
ACP
ACP
ACP
Pwr Supply
Pwr Supply
Pwr Supply
Domain 2
Domain 1
2nd B4
Ary Group 15
Ary Group 14
Ary Group 13
Ary Group 12
Ary Group 11
Ary Group 10
Ary Group 9
Ary Group 8
1st B4
Ary Group 7
Ary Group 6
Ary Group 5
Ary Group 4
Ary Group 3
Ary Group 2
Ary Group 1
SP SP SP SP
6th B4
Ary Group 29
Ary Group 28
Ary Group 27
Ary Group 26
Ary Group 25
Ary Group 24
Ary Group 23
5th B4
Ary Group 22
Ary Group 21
Ary Group 20
Ary Group 19
Ary Group 18
Ary Group 17
Ary Group 16

LDEVs and XP256

- Open level is important for 3 reasons:
  - Limited number of LDEVs on one controller (120)
    - Example: Open-3 uses 12 LDEVs while for virtually the same space Open-8 uses 4
    - 2.4 GB vs. 7.3 GB
  - XP256 only support 1024 LDEVS (1/00)

	RAID1		RAID5		
OPEN-3	OPEN-8	OPEN-9	OPEN-3	OPEN-8	OPEN-9
28/12	27/4	28/4	41/18	41/6	41/6
Usable Space in GB / LDEVS					

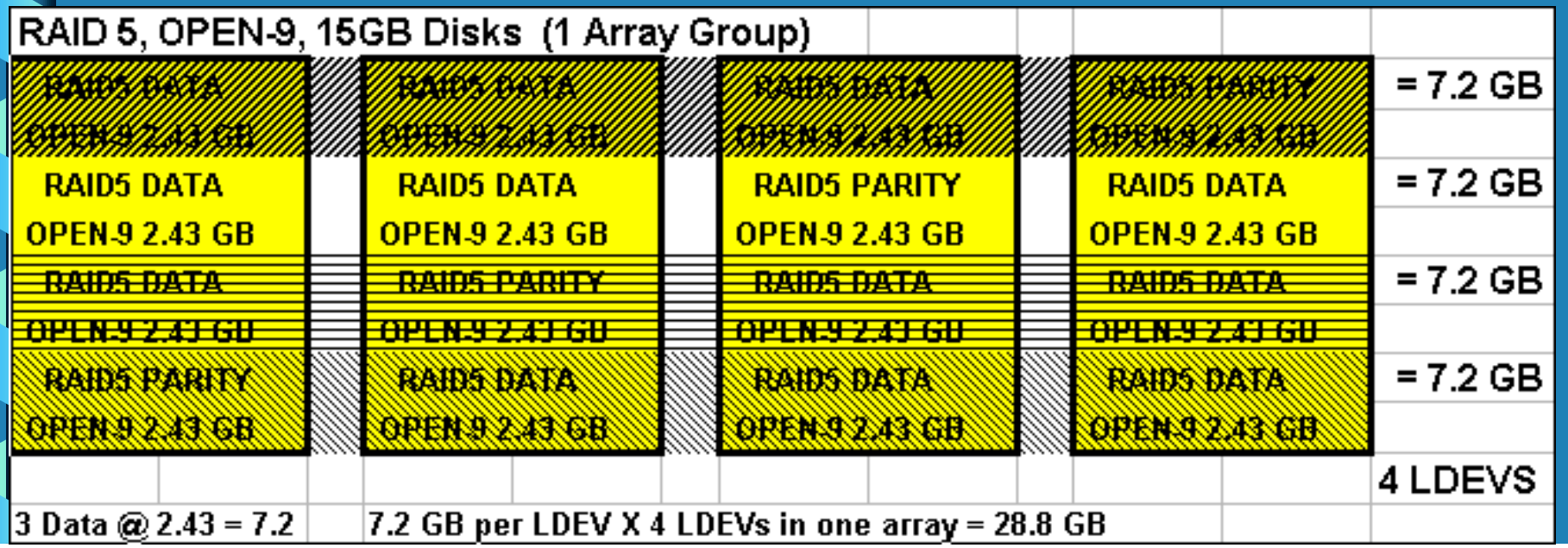
LDEVs and XP256 Write queue depth

Decrease/Increase Based on # of device files

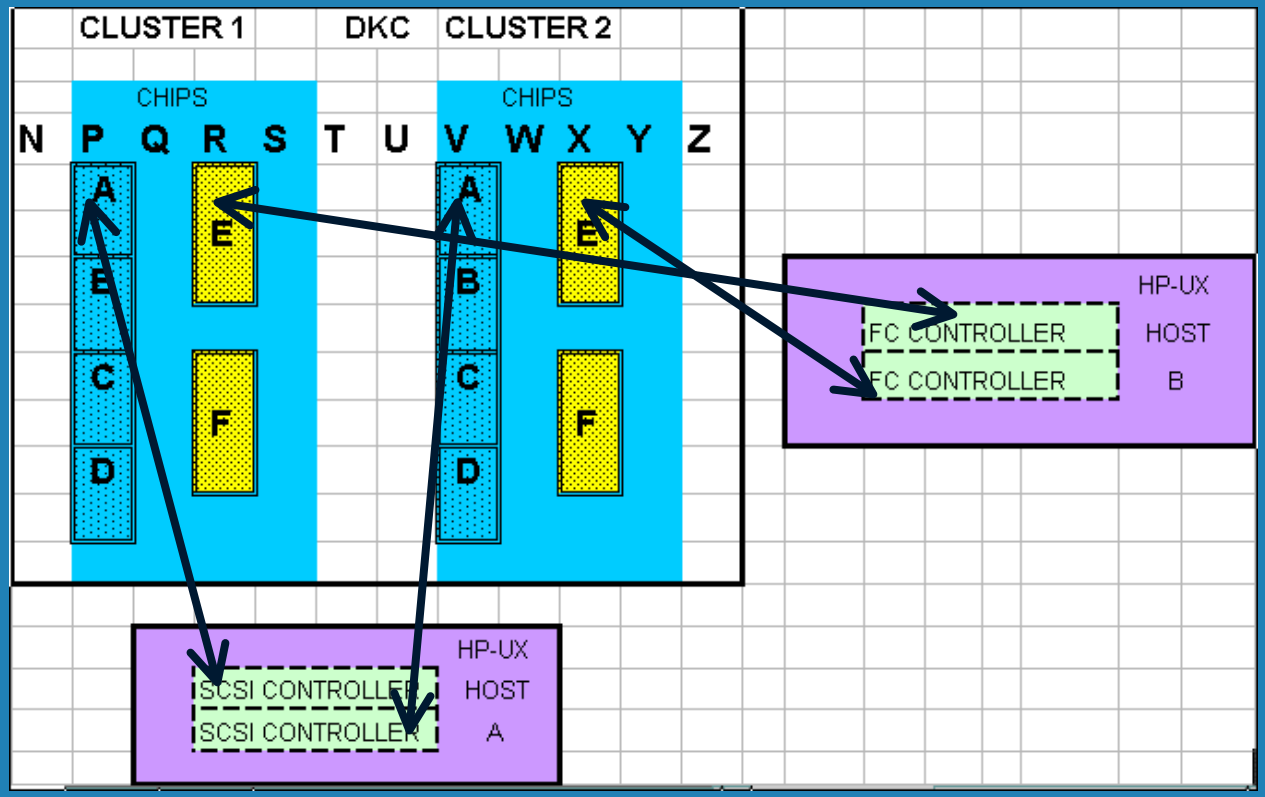
- The 3rd reason the OPEN level is important is the write queue depth
- Max number of I/Os on one XP256 CHIP port
  - 256 per SCSI 1024 per FC
- HP-UX supports 8 per device file (LDEV)
- 8 array groups
  - OPEN-3: 96 LDEVs * 8 = 768 (overflow for SCSI)
  - OPEN-8/9: 32 LDEVs * 8 = 256
- Set up on HP-UX by editing queue start up configuration
- /usr/sbin/scsictl -m queue_depth=# /dev/rdsk/c6t0d0

Disks and their relationship to HP-UX

- HASS, internal, etc.. - 1 device file per disk
- AutoRAID - 1 device file per LUN - LUN is across ALL disks
- XP256 - 1 device file per LDEV - LDEV is across disks in array group



CHIP
Client Host Interface Processor



CL1A to c3 on HostA

CL2V to c5 on HostA

CL1E to c7 on HostB

CL2E to c9 on HostB

Host A: 2 paths to XP256 - c3 to Cluster1, slot P, port A and c5 to Cluster2, slot V, port A

Assign LDEV to CHIP scsi

Hi! I'm LDEV 0:01

RAID5 DATA OPEN-9 2.43 GB	RAID5 DATA OPEN-9 2.43 GB	RAID5 PARITY OPEN-9 2.43 GB	RAID5 DATA OPEN-9 2.43 GB	= 7.2 GB LDEV
CONFIGURE HOST ACCESS TO THE LDEV:		Pick CHIP that you want to access the LDEV from.		

Access LDEV #0:01

From CL1-A

Using SCSI ID 1

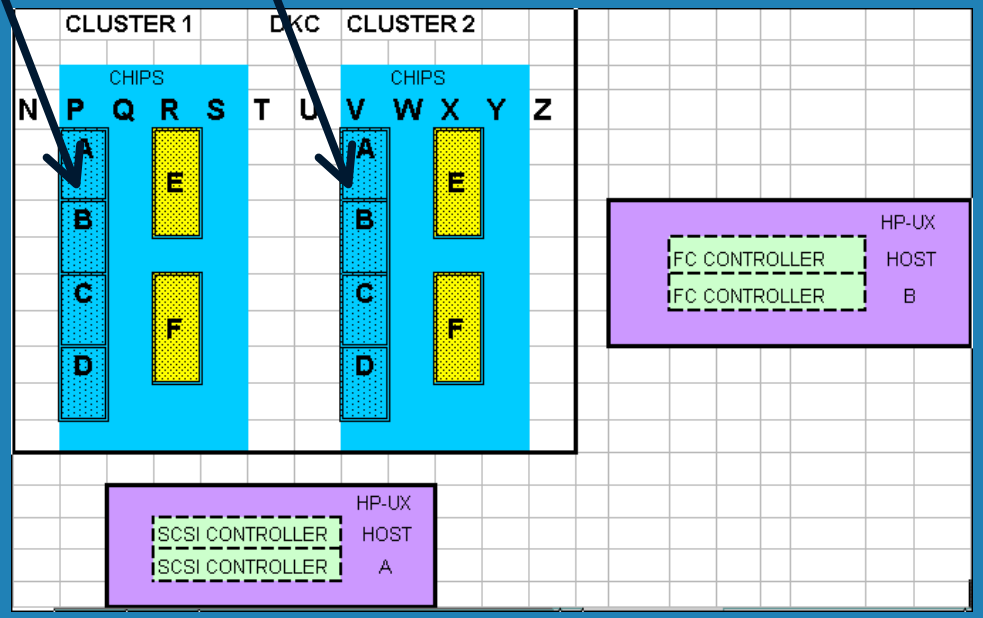
and LUN 00

And access this same LDEV

From CL2-A

Using SCSI ID 9

and LUN 00



ioscan OPEN-#

```
disk     109   0/2/0/0.1.0              sdisk     CLAIMED      DEVICE        HP
OPEN-9
                                       /dev/dsk/c3t1d0    /dev/rdsk/c3t1d0
disk     110   1/8/0/0.9.0              sdisk     CLAIMED      DEVICE        HP
OPEN-9
                                       /dev/dsk/c5t9d0    /dev/rdsk/c5t9d0
```

- On HP-UX Host A, 2 device files pointing to the same LDEV on the XP256.
- T1 and T9 was set by the SCSI ID# we gave the LDEV, d0 was set by LUN#

```
--- Physical volumes ---
PV Name                     /dev/dsk/c3t1d0
PV Name                     /dev/dsk/c5t9d0   Alternate Link
PV Status                   available
Total PE                    24
Free PE                     24
```

Assign LDEV to CHIP Fibre Channel

Hi! I'm LDEV 0:02

RAID5 DATA OPEN-9 2.43 GB	RAID5 DATA OPEN-9 2.43 GB	RAID5 PARITY OPEN-9 2.43 GB	RAID5 DATA OPEN-9 2.43 GB	= 7.2 GB LDEV
CONFIGURE HOST ACCESS TO THE LDEV:		Pick CHIP that you want to access the LDEV from.		

Access LDEV #0:02

From CL1-E

Using SCSI ID F

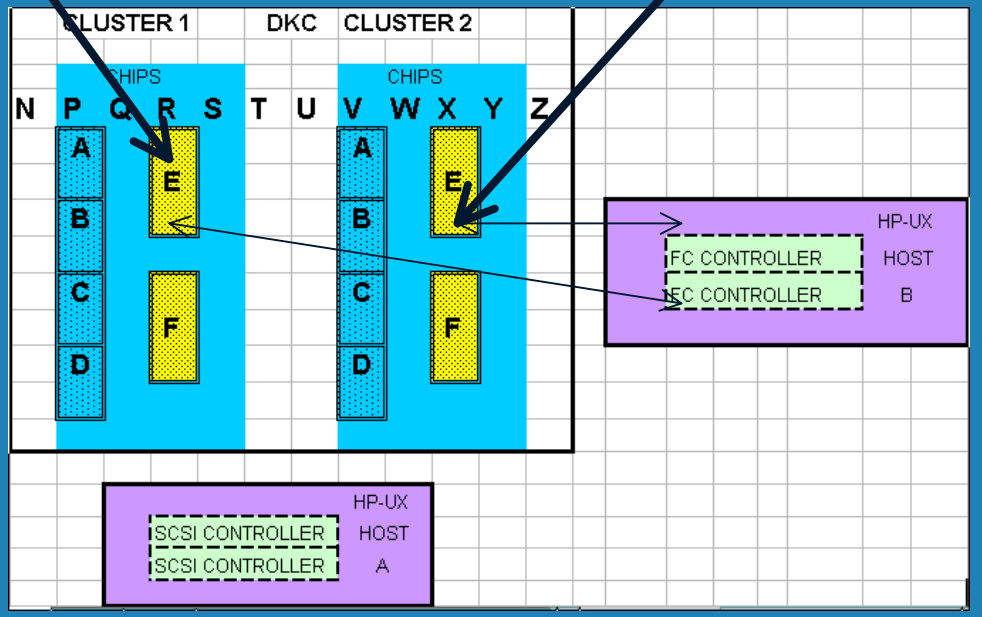
and LUN 60

And access this same LDEV

From CL2-E

Using SCSI ID F

and LUN 20



XP256	HP-UX	
LUN	Target	LUN
0-7	0	0-7
8-F	1	0-7
10-17	2	0-7
18-1F	3	0-7

LUN 60 on XP256 will be: SCSI 12, LUN0

LUN 61 on XP256 will be: SCSI 12, LUN1

LUN 20 on XP256 will be: SCSI 4, LUN0

LUN 7 on XP256 will be: SCSI 0, LUN 7

XP256
ioscan - Fibre Channel

```
0/10/0/0                                fc                          HP Fibre Channel
Mass Storage Adapter
0/10/0/0.5                              lan                         HP Fibre Channel
Mass Storage Cntl
0/10/0/0.8                              fcp                         FCP Protocol
Adapter
0/10/0/0.8.0.0.0                                ext_bus             FCP Array
Interface
0/10/0/0.8.0.0.0.0                                target
0/10/0/0.8.0.0.0.0.0                                disk    HP      OPEN-8
0/10/0/0.8.0.0.0.1                                target
0/10/0/0.8.0.0.0.1.0                                disk    HP      OPEN-9-CVS
0/10/0/0.8.0.0.0.2                                target
0/10/0/0.8.0.0.0.2.0                                disk    HP      DISK-SUBSYSTEM
0/10/0/0.8.0.0.0.3                                target
0/10/0/0.8.0.0.0.3.0                                disk    HP      DISK-SUBSYSTEM
0/10/0/0.8.0.0.0.4                                target
0/10/0/0.8.0.0.0.4.0                                disk    HP      DISK-SUBSYSTEM
```

FC Hardware Path

- 0/10/0/0.8.0.0.0.12.0
- 0/10/0/0 FC Adapter (will be c7)
- 8=Protocol Adapter (8=emulation)
- 0=Private loop
- 0=FC Port
- 0=Bus #
- 12=Target (SCSI ID# assigned on XP256)
- 0=LUN# assigned on XP256

```
disk      18  0/10/0/0.8.0.0.0.12.0  sdisk    CLAIMED     DEVICE       HP
DISK-SUBSYSTEM
                                     /dev/dsk/c7t12d0   /dev/rdsk/c7t12d0
```

Set IDs in a pattern

My target is always high for the primary

```
--- Physical volumes ---
PV Name                     /dev/dsk/c7t12d0
PV Name                     /dev/dsk/c9t4d0   Alternate Link
PV Status                   available
Total PE                    1751
Free PE                     1751

PV Name                     /dev/dsk/c9t12d1
PV Name                     /dev/dsk/c7t4d1   Alternate Link
PV Status                   available
Total PE                    1751
Free PE                     1751
```

Use target ID to know which SHOULD be the primary. Alternate should always be +or- 8

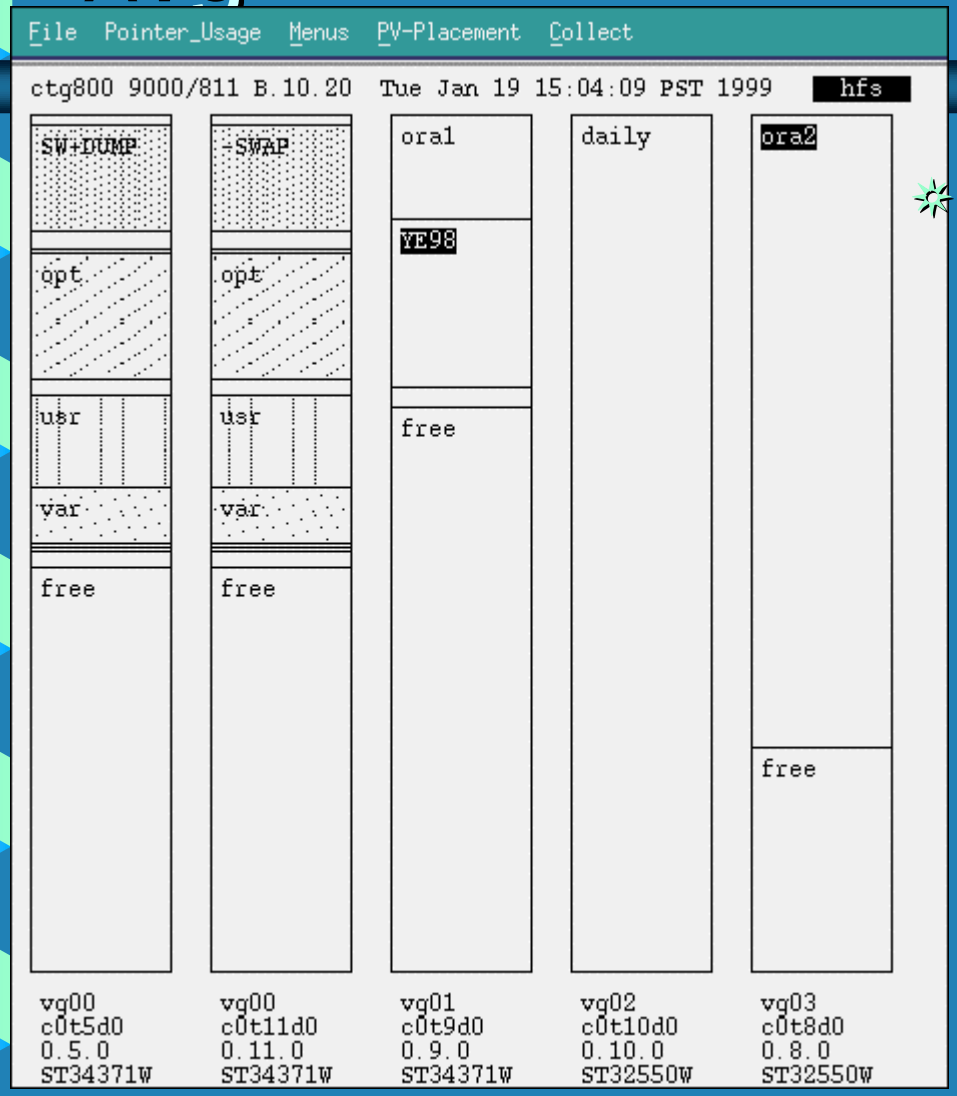
Get inquiry256.ksh from the HPRC

- Tool that maps between a single device file and the associated port and ldev on the XP256

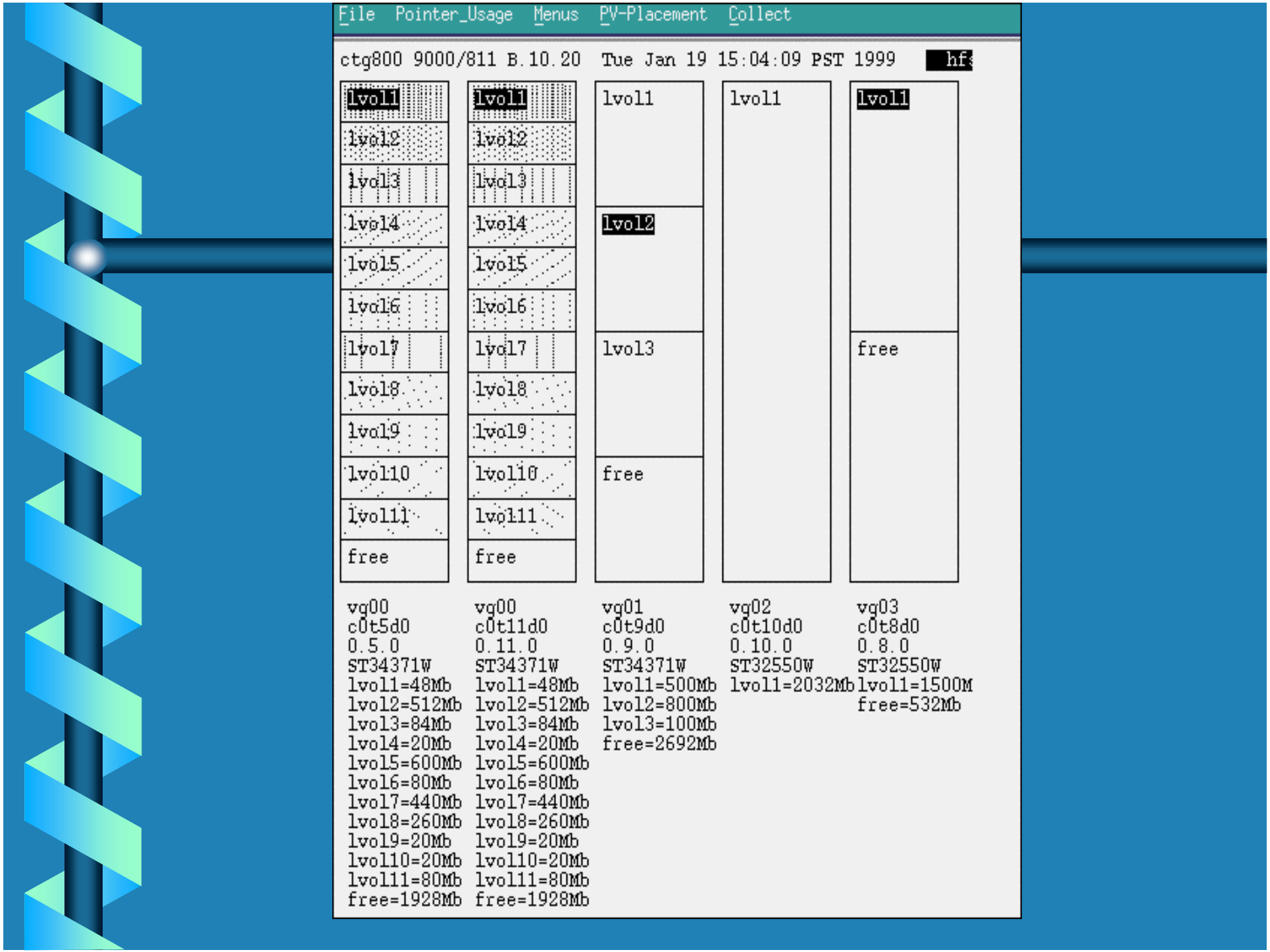
```
 Device File    ---> Port Target Lun CU:Ldev Serial#  Type      Size

 /dev/rdsk/c9t0d0 : CL1E    0    0   0:00  00035013 OPEN-3*2  4806720
Kbytes
 /dev/rdsk/c9t0d1 : CL1E    0    1   0:02  00035013 OPEN-3     2403360
Kbytes
 /dev/rdsk/c9t0d2 : CL1E    0    2   0:03  00035013 OPEN-3     2403360
Kbytes
/dev/rdsk/c18t0d0 : CL2E    0    0   0:00  00035013 OPEN-3*2  4806720
Kbytes
```

ftp://contrib:9unsupp8@hprc.external.hp.com/sysadmin/xvg/

Xvg



- Display:
  - Graphical
    - LV size
    - PV size


File Pointer_Usage Menus PV-Placement Collect
ctg800 9000/811 B.10.20 Tue Jan 19 15:04:09 PST 1999 hf
lvol1
lvol2
lvol3
lvol4
lvol5
lvol6
lvol7
lvol8
lvol9
lvol10
lvol11
free
lvol1
lvol2
lvol3
lvol4
lvol5
lvol6
lvol7
lvol8
lvol9
lvol10
lvol11
free
lvol1
lvol2
lvol3
free
lvol1
lvol1
free
vg00
c0t5d0
0.5.0
ST34371W
lvol1=48Mb
lvol2=512Mb
lvol3=84Mb
lvol4=20Mb
lvol5=600Mb
lvol6=80Mb
lvol7=440Mb
lvol8=260Mb
lvol9=20Mb
lvol10=20Mb
lvol11=80Mb
free=1928Mb
vg00
c0t11d0
0.11.0
ST34371W
lvol1=48Mb
lvol2=512Mb
lvol3=84Mb
lvol4=20Mb
lvol5=600Mb
lvol6=80Mb
lvol7=440Mb
lvol8=260Mb
lvol9=20Mb
lvol10=20Mb
lvol11=80Mb
free=1928Mb
vg01
c0t9d0
0.9.0
ST34371W
lvol1=500Mb
lvol2=800Mb
lvol3=100Mb
free=2692Mb
vg02
c0t10d0
0.10.0
ST32550W
lvol1=2032Mb
vg03
c0t8d0
0.8.0
ST32550W
lvol1=1500M
free=532Mb

More information

- Disk & File Management Tasks on HP-UX by Tom Madell (ISBN 0-13-518861-X)
- HP Education:
  - Hands on with LVM & MirrorDisk/UX
  - HP-UX Troubleshooting
  - Inside HP-UX

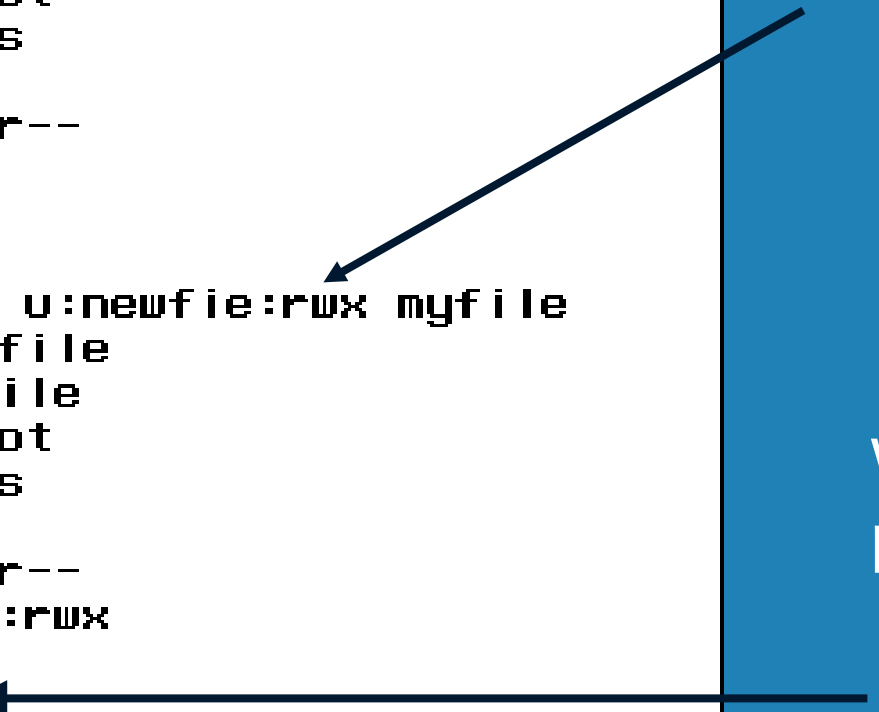
Appendix A: ACLs
JFS 3.3/ HP-UX 11+

```
# chmod 750 myfile
# ll myfile
-rwxr-x---   1 root       sys             24 Jan 15 14:11 myfile
# getacl myfile
# file: myfile
# owner: root
# group: sys
user::rwx
group::r-x
class:r-x
other:---
```

- Group and class entry are the same if no ACL has been set

setacl

```
# setacl -m u:sassy:r-- myfile
# getacl myfile
# file: myfile
# owner: root
# group: sys
user::rwx
user:sassy:r--
group::r-x
class:r-x
other:---
# setacl -m u:newfie:rwx myfile
# getacl myfile
# file: myfile
# owner: root
# group: sys
user::rwx
user:sassy:r--
user:newfie:rwx
group::r-x
class:rwx
other:---
```



W is now part of class

```
$ cd /jfs33
su: /jfs33: Permission denied.
$ more /jfs33/myfile
/jfs33/myfile: Permission denied
$
$ exit
logout
# ll -d /jfs33
drwxr-x---   3 root       root           96 Jan 15 14:11 /jfs33
# ll /jfs33/myfile
-rwxrwx---+  1 root       sys            24 Jan 15 14:11 /jfs33/myfile
# setacl -m u:newfie:rwx /jfs33
# ll -d /jfs33
drwxrwx---+  3 root       root           96 Jan 15 14:11 /jfs33
# su - newfie
```

```
$ whoami
newfie
$ ll /jfs33
total 2
drwxr-xr-x   2 root       root           96 Jan 15 14:05 lost+found
-rwxrwx---+  1 root       sys            24 Jan 15 14:11 myfile
$ more /jfs33/myfile
Hello, my name is Chris
$ rm /jfs33/myfile
$ ll /jfs33
total 0
drwxr-xr-x   2 root       root           96 Jan 15 14:05 lost+found
$
```

setacl

```
$ whoami
sassy
$ ll -d /jfs33
drwxrwx---+  3 root       root            96 Jan 15 14:23 /jfs33
$ ll /jfs33
/jfs33 unreadable
total 0
$ exit
logout
# setacl -m u:sassy:r-x /jfs33
```

```
$ whoami
sassy
$
$ ll /jfs33
total 2
drwxr-xr-x   2 root       root            96 Jan 15 14:05 lost+found
-rw-r-x---+  1 root       sys             16 Jan 15 14:25 myfile
$
$ more /jfs33/myfile
This is my file
$ rm /jfs33/myfile
/jfs33/myfile: 650+ mode ? (y/n) y
rm: /jfs33/myfile not removed.  Permission denied
```

Default ACLs

```
# setacl -m default:u:nking:rwx /jfs33
# getacl /jfs33
# file: /jfs33
# owner: root
# group: root
user::rwx
user:sassy:r-x
user:newfie:rwx
group::r-x
class:rwx
other:---
default:user:nking:rwx
#
# touch /jfs33/file1
# getacl /jfs33/file1
# file: /jfs33/file1
# owner: root
# group: sys
user::rw-
user:nking:rwx  #effective:---
group::---
class:---
```

Correct combo for ACLs on JFS, Trusted

- JFS 3.3 installed
- HP-UX 11+
- File system - version 4
- NOT /, /usr, /var, or /opt

```
# /usr/lbin/getprdef -r
NO, 0, 8, 0, 0, -1, 0, YES, YES, NO, NO, NO, YES, 3, 10, 2, 0
# grep nking /etc/passwd
nking:*:202:20:,,,:/home/nking:/opt/perf/bin/glance
# ll /tcb/files/auth/n/nking
-rw-rw-r--   1 root       root           143 Jan 15 13:45 /tcb/files/auth/n/nkin
g
# swlist -l fileset | grep "JFS 3.3 base"
# JFS                                    3.3            JFS 3.3 base filesystem
# uname -a
HP-UX ctg800 B.11.00 A 9000/803 2000767436 two-user license
# vxupgrade /jfs33
/jfs33: vxfs file system version 4 layout
```

```
# vxupgrade -n 4 /jfs33
# vxupgrade /jfs33
/jfs33: vxfs file system version 4 layout
```

Appendix B: Oracle
Oracle on raw

- Used when want the maximum performance
  - Bypasses traditional disk writing. 20-30% +
- Administrative headache
- Less popular than in past
- Very intensive I/O application
- Use LVM
- Add entries in fstab as comments
- Do not put archive logs on raw

Oracle

Oracle with Asynchronous I/O

- Oracle does parallel writes without using multiple DBWRs
- Supported on raw devices only
- async driver must be in kernel
- /dev/async owned by oracle

Oracle

Oracle on HFS

- Tuning
  - Block size
  - Fragment size
  - minfree
- mkfs -m
  - Use this command to display command that was used to create the file system

Oracle

Recommended Mount Options with JFS

- Mount -F vxfs -o delaylog,mincache=direct,
- convosync=direct,nodatainlog
- Uses the “direct” feature
- SLOW if use other methods for accessing Oracle data other than Oracle itself. (Example: tar).
- Direct Writes
  - Simulates raw, in that writes bypass the file system buffer cache

Requires OnLine JFS

Oracle

Oracle and OFA and disk layout

- OFA (Oracle Flexible Architecture)
  - Archive log files
  - Rollback segment data files, export files
  - Executables, copy of the control file, redo logs, SYSTEM data files
  - Data files, temp user data files, copy of control file
  - Index data files, copy of control file

Oracle

Oracle
Table Striping

- Splitting a tablespace into many physical files on more than one disk
- Reduce I/O bottlenecks
  - CREATE TABLESPACE PROD
    - DATAFILE "/disk1/file1.dbf" SIZE 50M
    - DATAFILE "/disk2/file1.dbf" SIZE 50M
  - CREATE TABLE MYTABLE
    - STORAGE (INITIAL 45M NEXT 45M)

Oracle

Oracle and LVM Striping

- Recommended over DATAFILE striping
- 50-500% increase over non striped tables
- Random access (OLTP)
  - Smaller stripe size (1MB)
- Sequentially accessed data
  - Larger stripe size

Oracle

Oracle and Database Writers

- If not using asynchronous I/O
- One DBWR per disk with DB files
- 0-15% +

Oracle

Oracle and AutoRAID/XP256

- OK for AutoRAID
  - Datafiles
  - Archive Logs
- Best not on AutoRAID
  - Redo logs
  - Should still be mirrored
- XP256
  - RAID1 vs. RAID5
  - Use Target ID (SCSI) for specific usage:
    - 5 (Database)
    - 4 - Lun 0+1 (Redo)
    - 4 - Lun 2 (Archive logs)
    - 3- Lun 0 (Programs)
    - 3 -Lun 1 (Workspace)

Oracle

Appendix C: MC/ServiceGuard

- CLVM (Cluster LVM)
  - Superset of LVM
  - Exclusive VG activation
    - vgchange -a e (SG daemons must be running)
  - Cluster ID
    - vgchange -c y/n
  - SG or LockManager enables functionality

MC/SG

Reserve SCSI IDs

- Host A - 7 <----------> Host B - 6
- What if add 2 more hosts?
- Reserve 5 and 4
- Start disks at 3
- Want the highest priorities on the controller cards
- Good to do even if not currently using MC/SG

MC/SG

Minor # (0xVG00LV), VG name and MC/SG

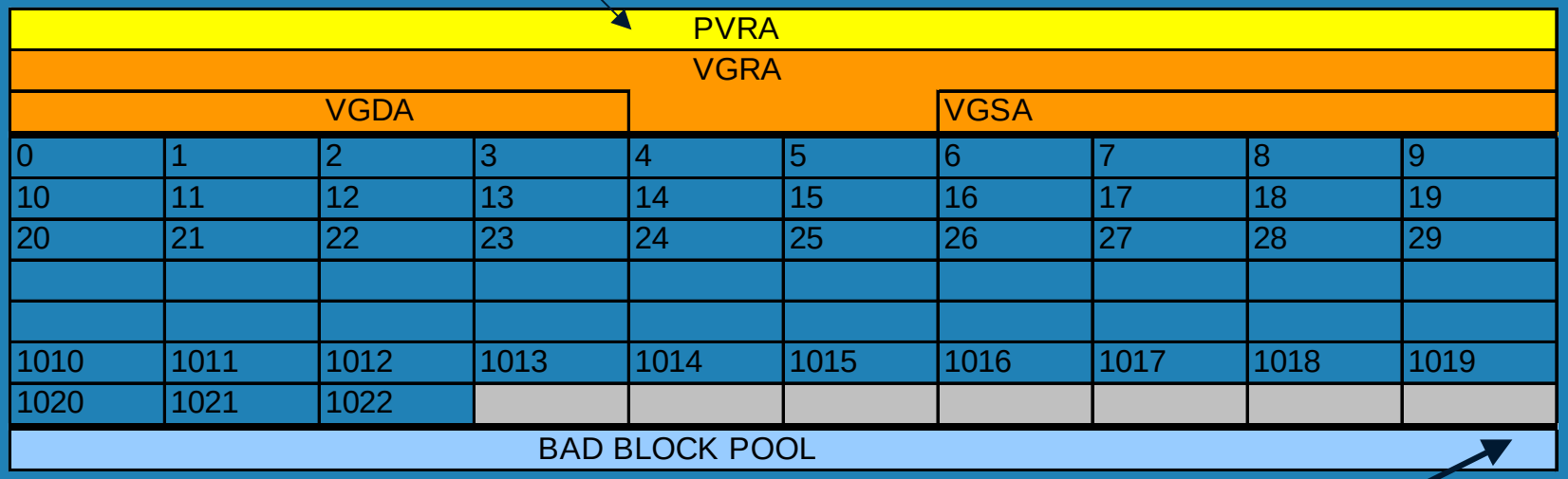
- Use naming convention to be able to tell what VG is primary for a node
- VGs that begin with 0 are local to the node
  - 0x000000, 0x010000, 0x020000, vg00, vg01
- VGs that begin with 1 belong to node A
  - 0x100000, 0x110000, 0x120000, vg10,vg11,vg12
- VGs that being with 2 belong to node B
  - 0x200000, 0x210000, 0x220000, vg20,vg21
- Increase KP: maxvgs
- Warning: VOLGRP structure is large

MC/SG

Where is the lock disk?

Exclusive

Cluster ID



LOCK DISK

MC/SG

Lock disk and sw mirroring

- Only a problem if needs to use lock disk
- Disk A / Disk B
- Disk A crashes
- Replace Disk A with new disk
- vgcfgrestore
  - Does not restore BBRT
- vgchange - a y (vgsync)
- Lock disk not initialized
- Halt the cluster
- Bring up cluster
- Reinitializes lock disk

MC/SG

MC/Service Guard

- /etc/lmvrc: AUTO_VG_ACTIVATE=0
- VG names & minor numbers need to be the same
- Don't need to halt if adding or modifying LVs, must halt cluster if adding VG
- See Appendix "D" for information on LVM tasks on MC/SG

MC/SG

Need to access data with application down

- Backups and other maintenance
  - Halt cluster, remove VG from cluster, activate & mount
  - Halt package, activate VG, mount
  - “While true do” loop
    - Test for application maintenance switch
    - Stop the app the normal way (not with SG)
    - Run maintenance
    - Restart app
    - Remove maintenance switch

MC/SG