

EMC² in the Real World

Bryan Province

Caterpillar Inc.

Building AD3321

600 W. Washington Street

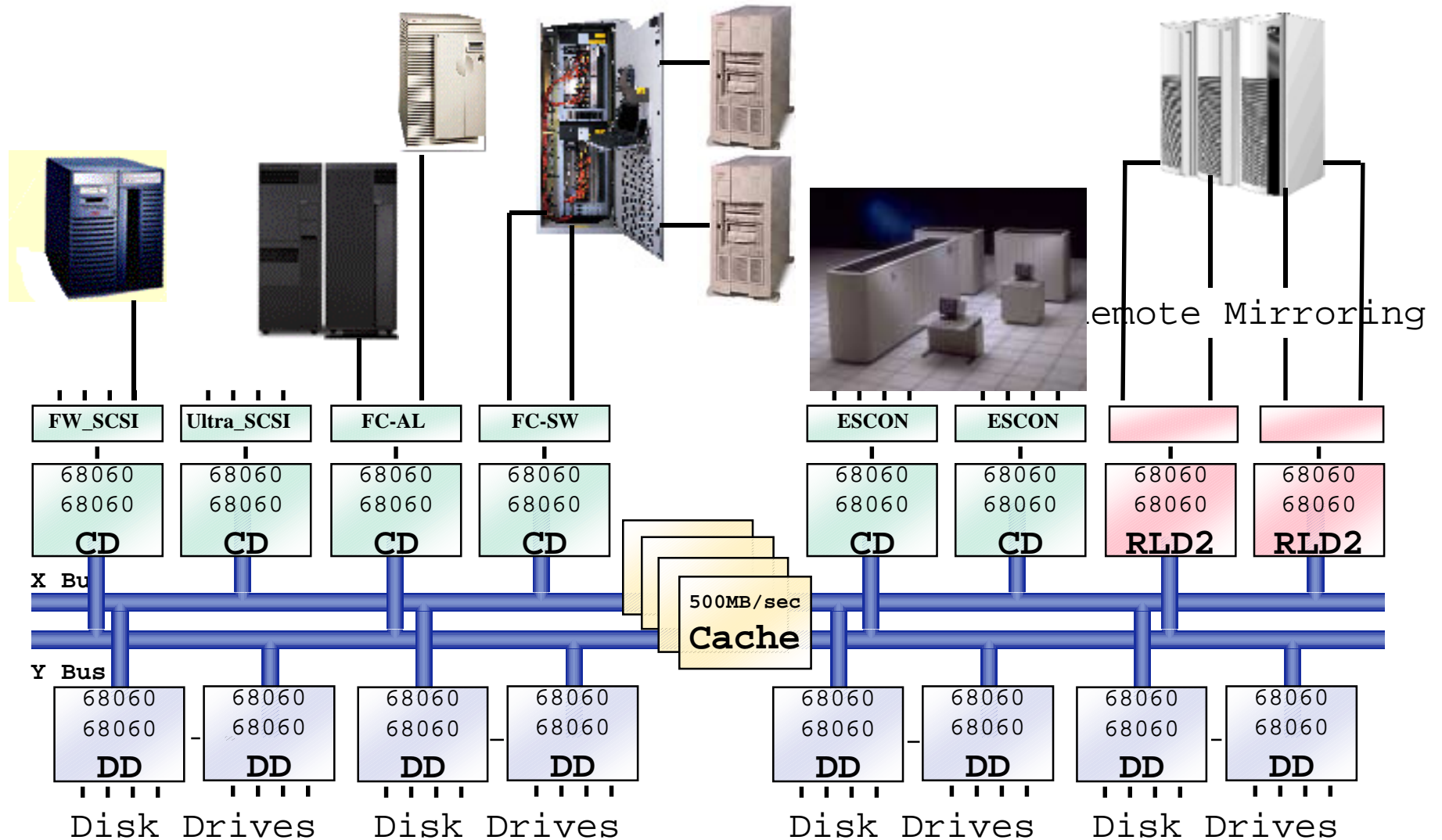
East Peoria, IL 61630-3321

Phone/Fax 309-675-2594

Province_Bryan_E@cat.com

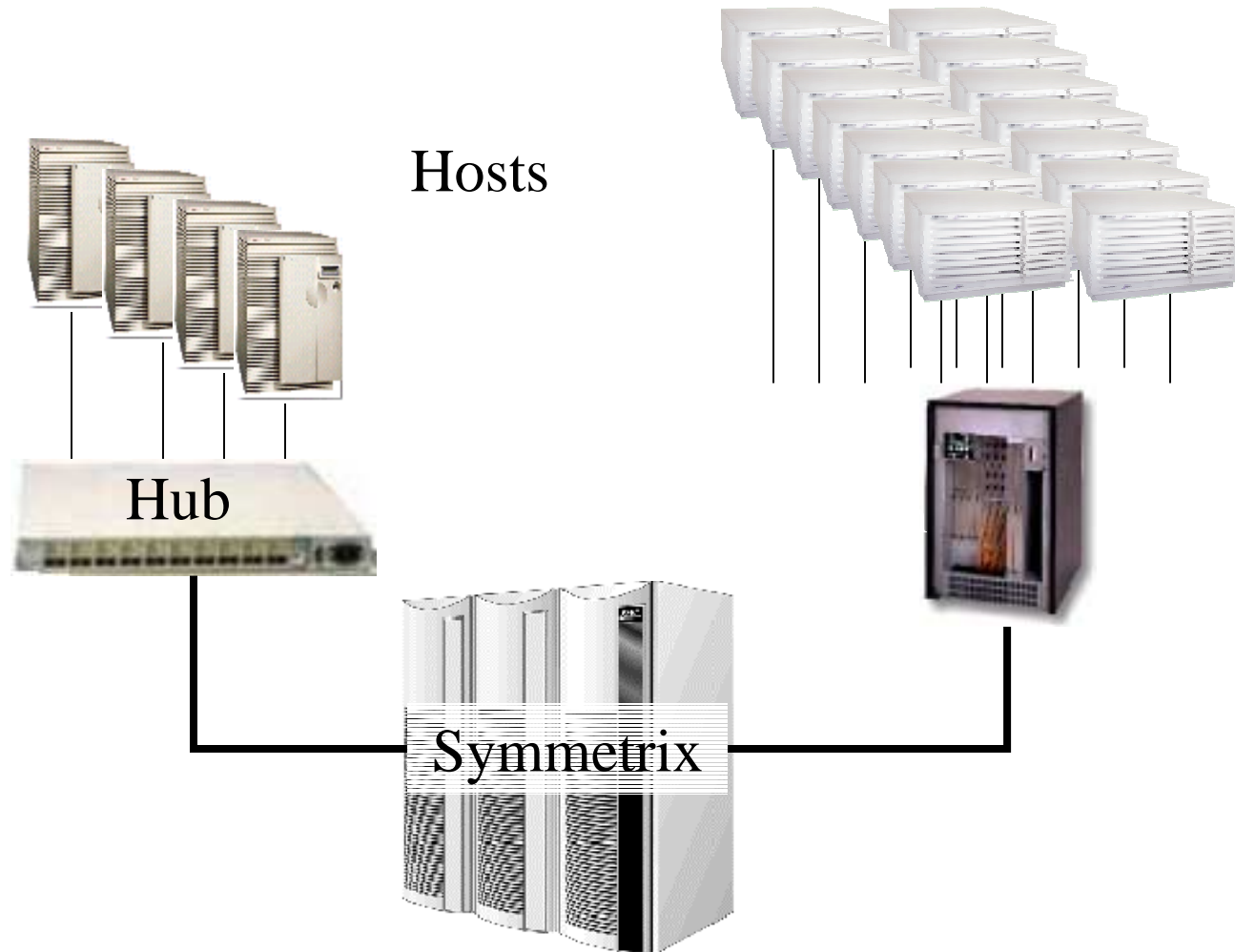
What is an EMC² Symmetrix?

- Large scale disk array
- I/O enhanced by large amounts of cache
- Optimized by caching algorithms
- Multiple host attachment via ESCON, FW SCSI, Ultra-SCSI, FC-AL (arbitrated loop), and FC-SW (switched fabric)



500MB/sec Cache Memory and Backplane
Dual Processor Channel and Disk Directors
Parallel, ESCON, FW SCSI , Ultra-SCSI,
FC-AL (arbitrated loop), and FC-SW (switched fabric) Host Attach

Storage Area Networks (SANS)



Storage Area Networks (SANS)

- Can connect 4 hosts to a hub (debatable)
- Can connect more hosts to a switch
- Older (Tachyon) HP FC cards do not support true switched fabric, only arbitrated loop
- There are fibre switches that emulate arbitrated loop as well as switched fabric

Switched Fabric with HP FC Cards

- Previous HP FC cards were based on the Tachyon chipset. Cannot do fabric logon.
- HP is now shipping a PCI version of the “Tach-Lite” chipset on N and L class machines. Supports full fabric connectivity.
- K-Class and D-Class HSC cards in May?
- Will also need HP-UX 11.00.

Switched Fabric with HP FC Cards

- Get TachLite drivers from www.software.hp.com
 - TachLite PCI cards (N & L Class) use A5158A
 - TachLite HSC cards for D-Class use A6684A
 - TachLite HSC cards for K-Class use A6685A
- Also need patch PHKL_21381 or latest equivalent.

Drivers for HP-UX 10.20

- J3630BA - FC Driver K class
- J1602AA_APZ - FC Driver D class
- Supports Arbitrated Loop only

CHECK YOUR FC CARDS!

- Look for the following label on the outside of your fibre channel cards

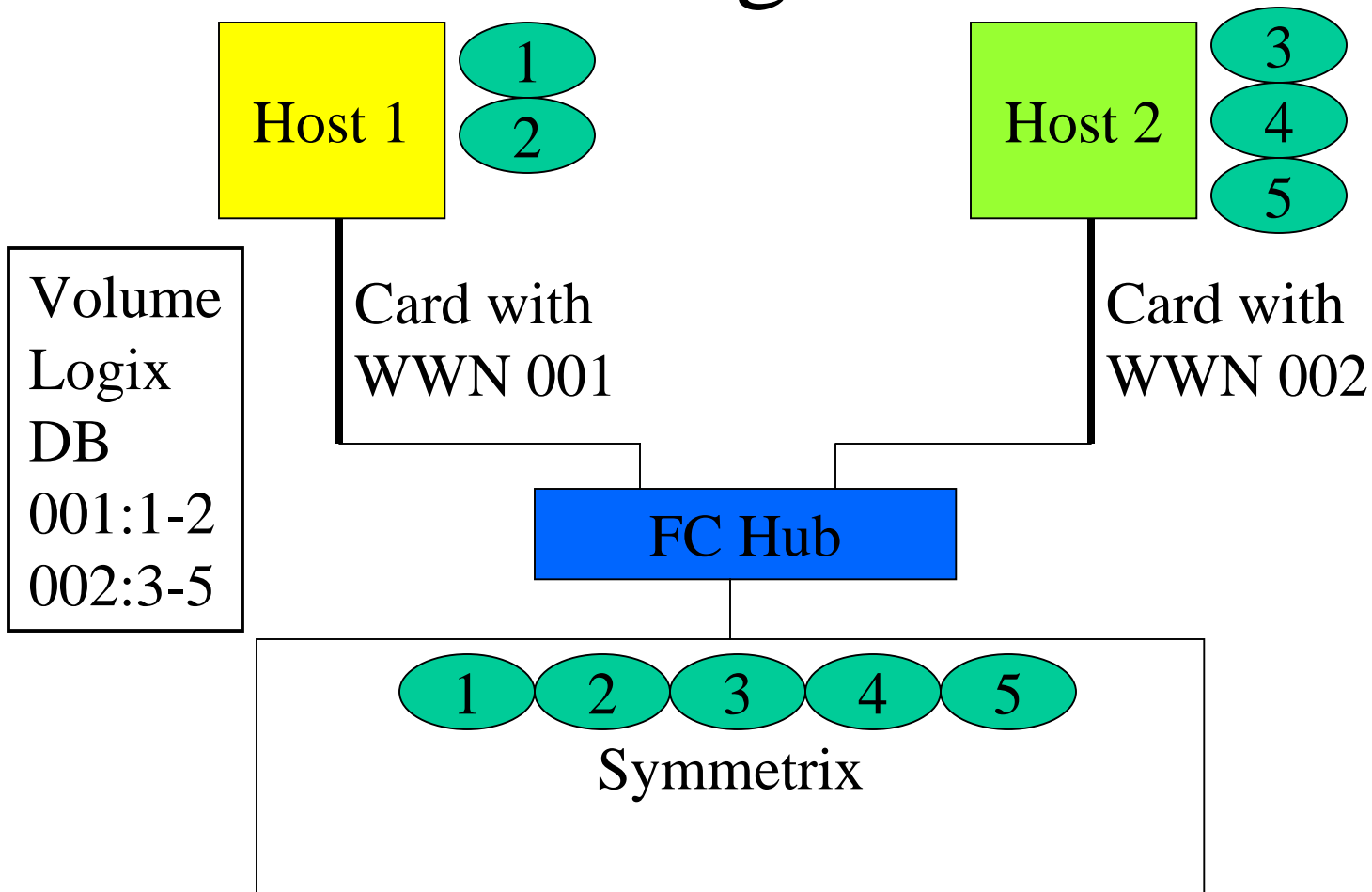


- Older style cards have a know problem of dropping connections
- You will see **POWERFAILED** messages from dmesg and in your syslog.log

Disk Security with Volume Logix

- Software that runs on the Symmetrix
- Provides disk masking based on a card's World Wide Name (WWN)
- Every card has a unique WWN
- Similar to an Ethernet MAC address
- Makes drives visible to only specific cards with the correct WWN

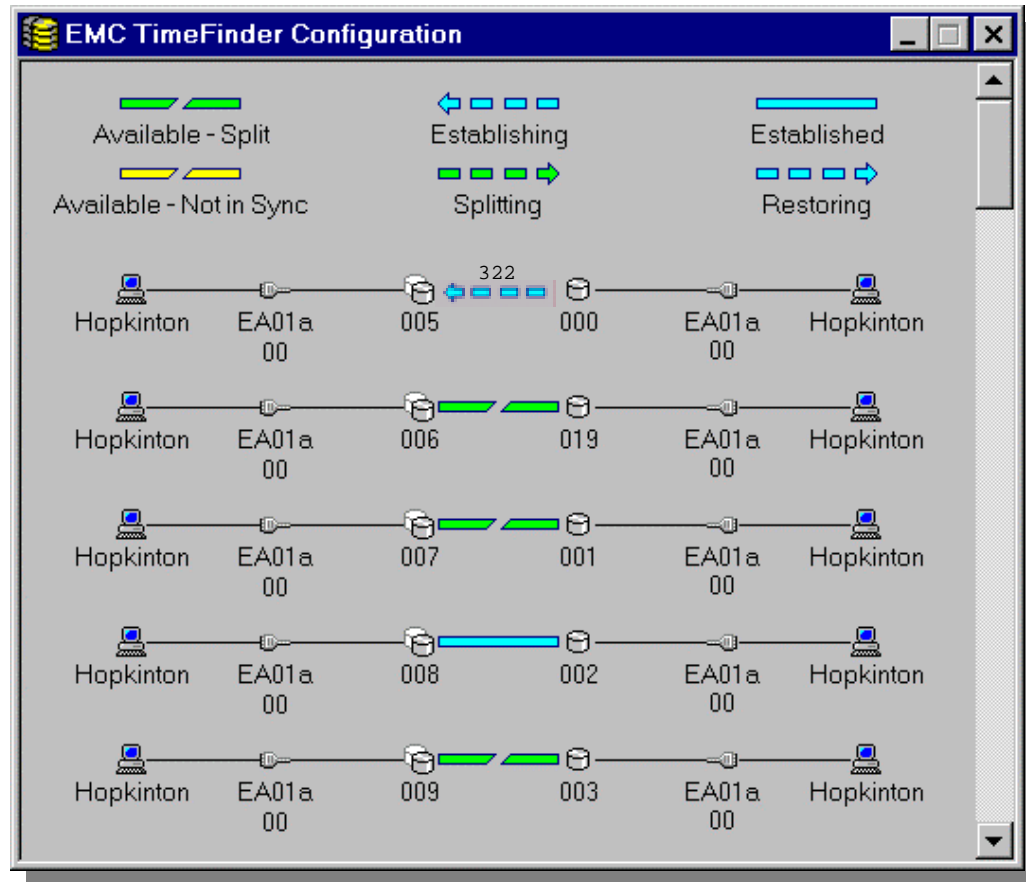
Disk Security with Volume Logix



Business Continuation Volumes (BCVs)

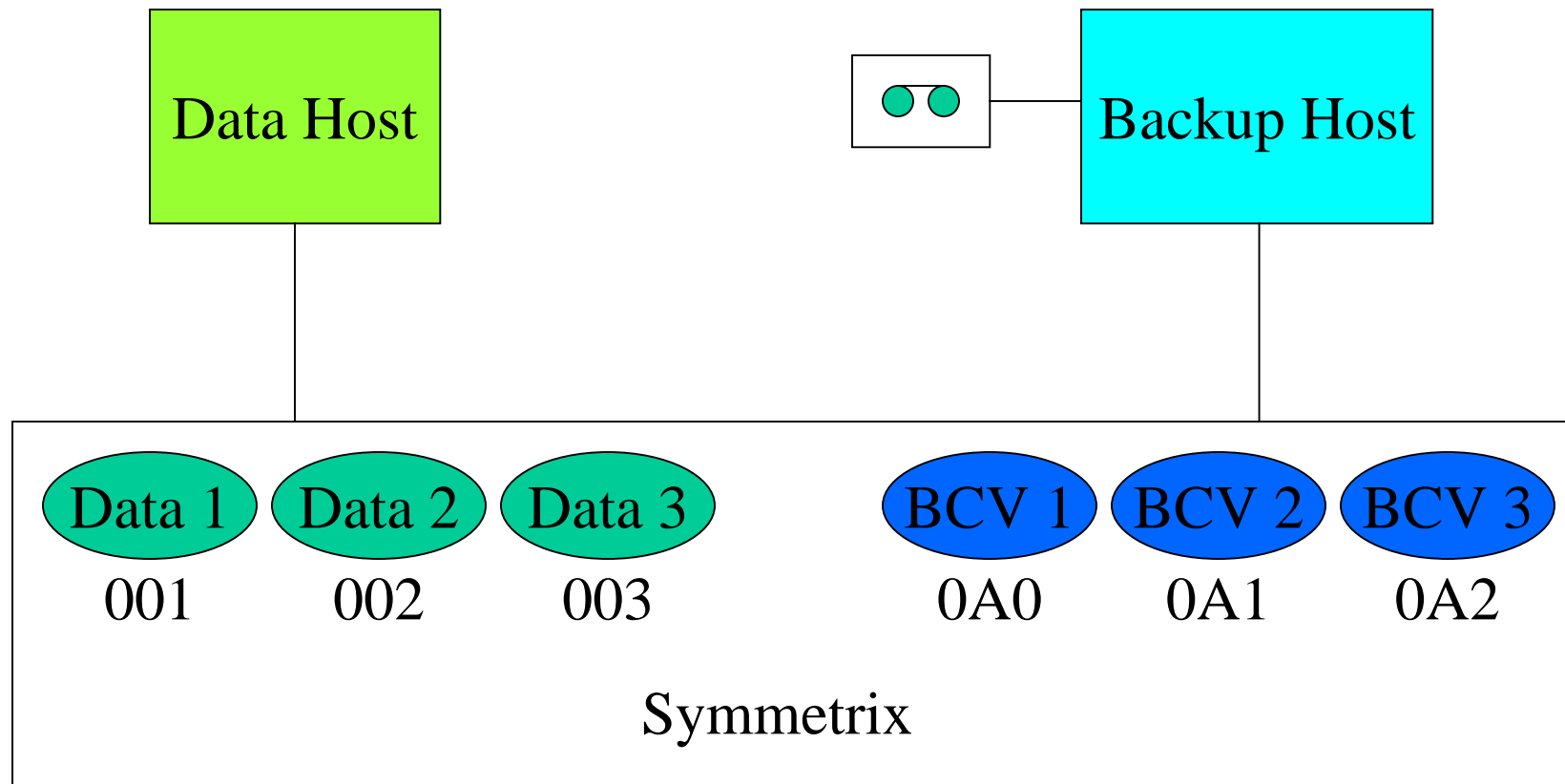
- Can use to create an additional mirrored copy of data
- Mirror can be established and split at will
- Mirrored data can be used for backups, testing, data moving, or other applications
- Can also restore data from BCVs to standard disks

TimeFinder Manager



- Monitor and manage TimeFinder operations:
 - Establish a BCV pair
 - Split a BCV pair
 - Re-establish a BCV pair incrementally
 - Restore from a BCV device
 - Incremental restore from a BCV device
 - Remote Management

Example of BCV backup



Example of BCV backup

- Setup BCV group file

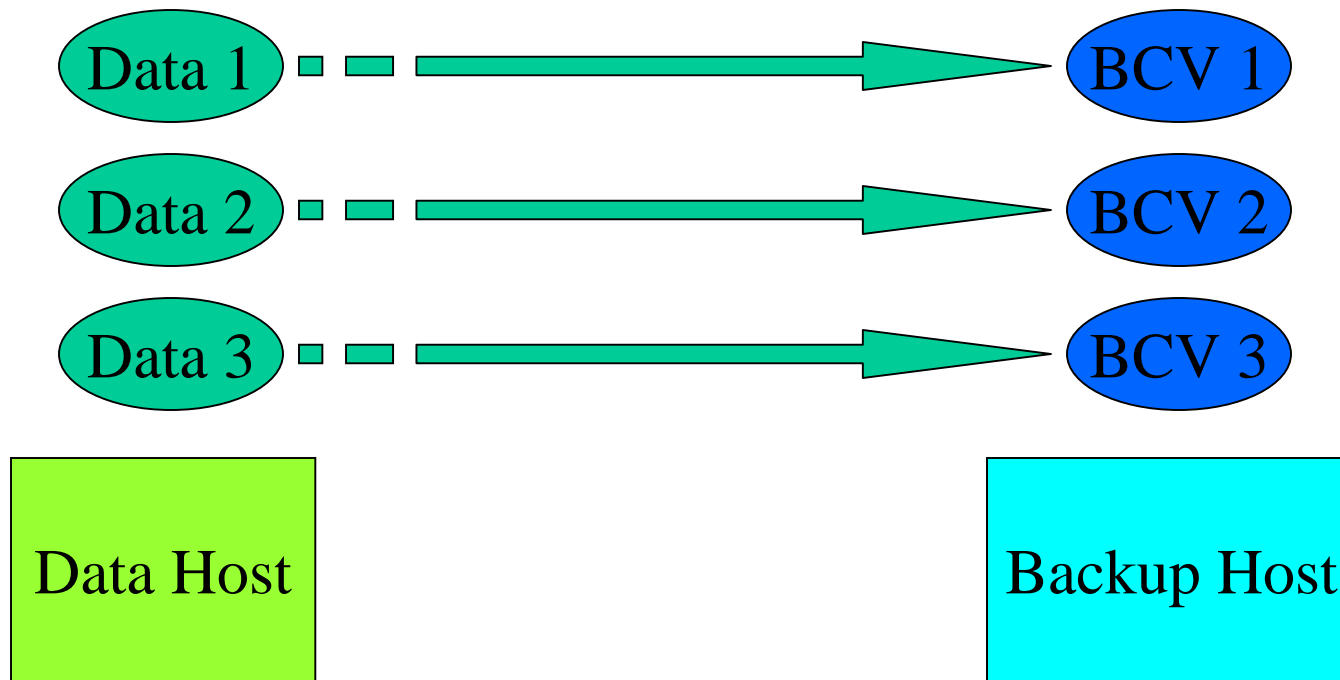
```
# SymmWin BCV File
# Reg Dev Num          BCV Dev Num
51001yyy 510a0yyy
51002yyy 510a1yyy
51003yyy 510a2yyy
```

Example of BCV backup

- Begin establish of BCVs to Data drives

```
echo "Establishing BCV pairs for ecq."  
bcv est -f backup.bcv -i  
if [[ $? -ne 0 ]]  
then  
    echo "NORMAL - Unable to do incremental establish."  
    echo " Trying full establish."  
    bcv est -f backup.bcv  
fi
```


Example of BCV backup



Example of BCV backup

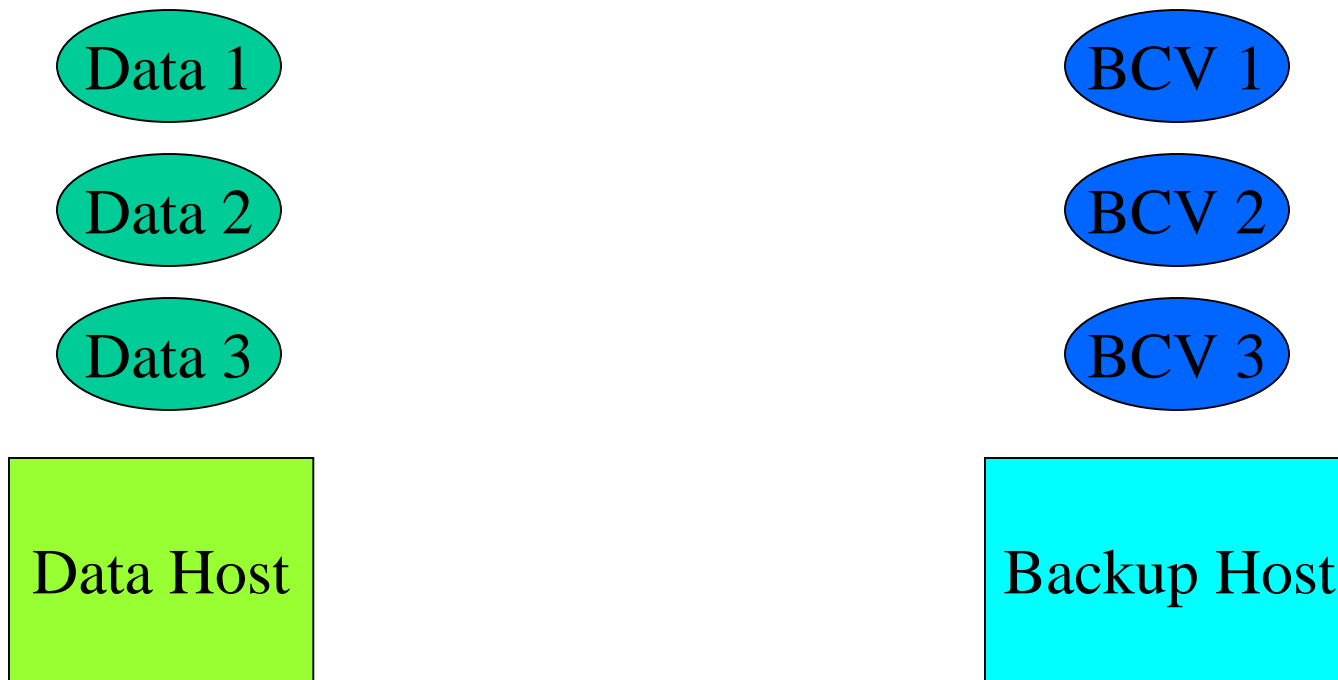
- Test progress of establish with bcv command

```
bcv verify -f backup.bcv
while [[ $? -ne 0 ]]
do
    sleep 360
    bcv verify -f backup.bcv
done
```

Example of BCV backup

- When establish complete:
 - shutdown application(s)
 - unmount data filesystems
 - run “vgchange -a n” to deactivate the volume group
 - split the BCVs
 - reactivate the volume group, mount the filesystems, and restart the application(s)

Example of BCV backup



Example of BCV backup

- On the backup host:
 - import the volume group
 - mount the filesystems
 - begin the backup
- Be sure to unmount the filesystems and deactivate the volume group on the backup host when the backup is complete.

Two Different BCV Commands

- Can control BCVs with “symmbcv” or “bcv”
- “bcv” is the older version but still available
- Need to take steps to make the “bcv” command work.

Using the “bcv” Command

- Create a link to the bcv executable

```
/usr/emc/ECC/bin/bcv -> /opt/emc/SYMCLI/4.0.2/bin/bcv
```

- Create the bcv_gatekeeper file

```
cat bcv_gatekeeper  
/dev/rdisk/c7t15d0
```

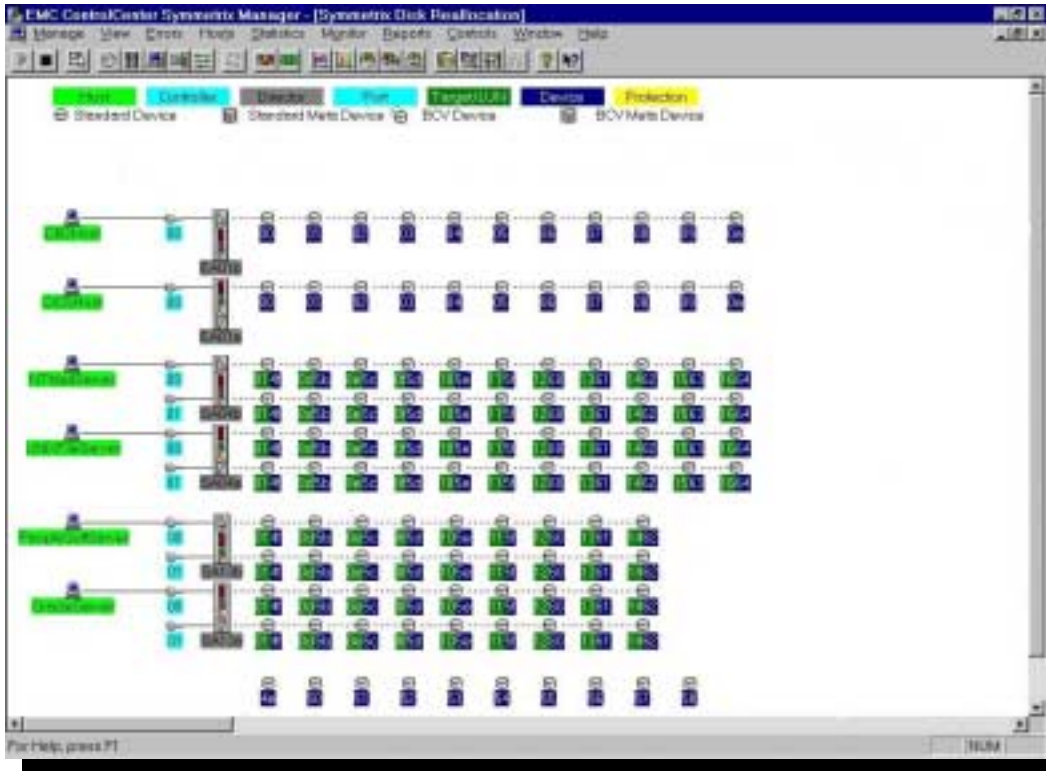
- Add the following lines to symmappsrc.sh

```
BCV_LOG=/usr/emc/ECC/bcv.log  
export BCV_LOG  
BCV_GK_FILE=/usr/emc/ECC/bcv_gatekeeper  
export BCV_GK_FILE  
BCV_ALL_GK=/usr/emc/ECC/bcv_gatekeeper  
export BCV_ALL_GK
```

Symmetrix Disk Reallocation (SDR)

Gives the ability to manage storage
within a Symmetrix

Symmetrix Disk Reallocation



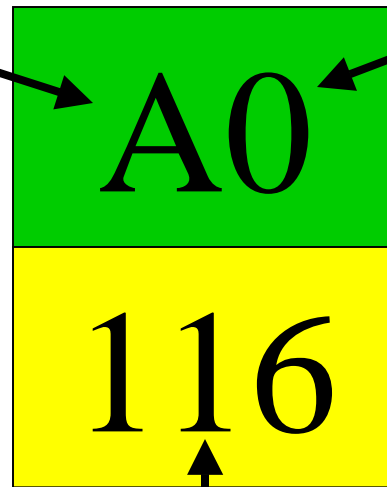
Assign or re-assign
Open System
Volumes
Drag and drop logical
volumes between:

- SCSI adapters
(workload
balancing)
- Host systems
(resource
allocation)

Disk view under Symmetrix Manager

Target (SCSI addr 10)

Logical Unit Number (zero)



Symmetrix Device Number

Precautions for Targets and LUNs for HP-UX

- LUNs can only be in the range of 0 to 7
- LUN chain for a particular SCSI address must start with 0. (Can't add a drive with target 1, LUN 1 without a target 1, LUN 0)
- Can have a total of 128 drives on one controller.

SOFTWARE VERSION WARNING!

- Do NOT use SDR software version less than 4.0.2.
- There is a known problem with SDR version 4.0.1 or less.
- When moving drives, ANY drive on ANY interface with target and LUN 00 could be removed from that interface.

Drive Appearance on HP-UX

- Make drives available to the interface with SDR and Volume Logix if necessary
- Run “ioscan” to discover the drives
- Run “insf -e” to setup the device files
- Confirm drives are available with “ioscan” and “inq”

Drive Appearance on HP-UX

- If using Volume Logix, first need to find the World Wide Name
- Can use either fcmsutil or vcmfind
- For non-fabric cards run “ioscan -fnC lan” and look for fcT1_cntl devices
- For fabric cards run “ioscan -fnC fc”

Drive Appearance on HP-UX (non-fabric)

```
ioscan -fnC lan
```

Class	I	H/W Path	Driver	S/W State	H/W Type	Description
lan	0	8/0.5	fcTl_cntl	CLAIMED	INTERFACE	HP Fibre Channel Mass Storage Cntl

```
/dev/fcms0
```

```
/opt/fcms/bin/fcmsutil /dev/fcms0
```

```
Local N_Port_ID is = 0x000001
N_Port Node World Wide Name = 0x10000060B03E42BB
N_Port Port World Wide Name = 0x10000060B03E42BB
Topology = IN_LOOP
Speed = 1062500000 (bps)
HPA of card = 0xFBF40000
EIM of card = 0xFFFA200D
Driver state = READY
Number of EDB's in use = 0
Number of OIB's in use = 0
Number of Active Outbound Exchanges = 1
Number of Active Login Sessions = 2
```

Drive Appearance on HP-UX (fabric)

```
ioscan -fnC fc
```

```
Class      I  H/W Path  Driver S/W State   H/W Type      Description
=====
fc          0  0/4/0/0  td   CLAIMED      INTERFACE      HP Tachyon TL/TS Fibre
Channel Mass Storage Adapter
                /dev/td0
/opt/fcms/bin/fcmsutil /dev/td0
```

```
Vendor ID is = 0x00103c
Device ID is = 0x001028
PCI Sub-system Vendor ID is = 0x00103c
PCI Sub-system ID is = 0x000006
Topology = PTTOPT_FABRIC
Local N_Port_id is = 0x282113
N_Port Node World Wide Name = 0x50060b0000077c09
N_Port Port World Wide Name = 0x50060b0000077c08
Driver state = ONLINE
Hardware Path is = 0/4/0/0
Number of Assisted IOs = 12572480
Number of Active Login Sessions = 2
```


Drive Appearance on HP-UX

- If already zoned through a switch, can use vcmfind to get World Wide Name for Volume Logix

```
/opt/emc/vcm/vcmfind
```

```
@(#) vcmfind , Copyright EMC Corp. 1998, 1999, and 2000 Revision: 2.2.1
```

```
ADAPTER          WWN                PATHNAME          FA
0-4-0-0 50060b0000077c08    /dev/rdisk/c4t0d0    3b
```

Drive Appearance on HP-UX

- The utility “inq” is your friend
- “inq” gives the best view of drives connected to a system whether EMC or otherwise

Drive Appearance on HP-UX

```
/opt/emc/SInquiry/V4.0.2/bin/inq -nodots
Inquiry utility, Version 7.17      (SIL Version 4.1.16)
Copyright (C) by EMC Corporation, all rights reserved.
For help type inq -h.
```

```
-----
DEVICE                :VEND      :PROD                :REV   :SER NUM      :CAP(kb)
-----
/dev/rdisk/c0t2d0     :HP        :DVD-ROM 305        :1.01  : 2000/09     :      0
/dev/rdisk/c1t6d0     :SEAGATE   :ST39204LC         :HP04  : 3BV0H35Z    :8891556
/dev/rdisk/c2t6d0     :SEAGATE   :ST39204LC         :HP04  : 3BV0LKH1    :8891556
/dev/rdisk/c4t0d0     :EMC       :SYMMETRIX         :5265  : 48000190    :   7680
/dev/rdisk/c4t0d1     :EMC       :SYMMETRIX         :5265  : 48001190    :88387200
/dev/rdisk/c4t0d2     :EMC       :SYMMETRIX         :5265  : 4800B190    :88387200
/dev/rdisk/c4t1d0     :EMC       :SYMMETRIX         :5265  : 48058190    :8838720
/dev/rdisk/c7t0d0     :EMC       :SYMMETRIX         :5265  : 99000000    :   7680
/dev/rdisk/c7t6d0     :EMC       :SYMMETRIX         :5265  : 99137000    :7619520
/dev/rdisk/c7t6d1     :EMC       :SYMMETRIX         :5265  : 99138000    :7619520
```

...

Drive Appearance on HP-UX

...

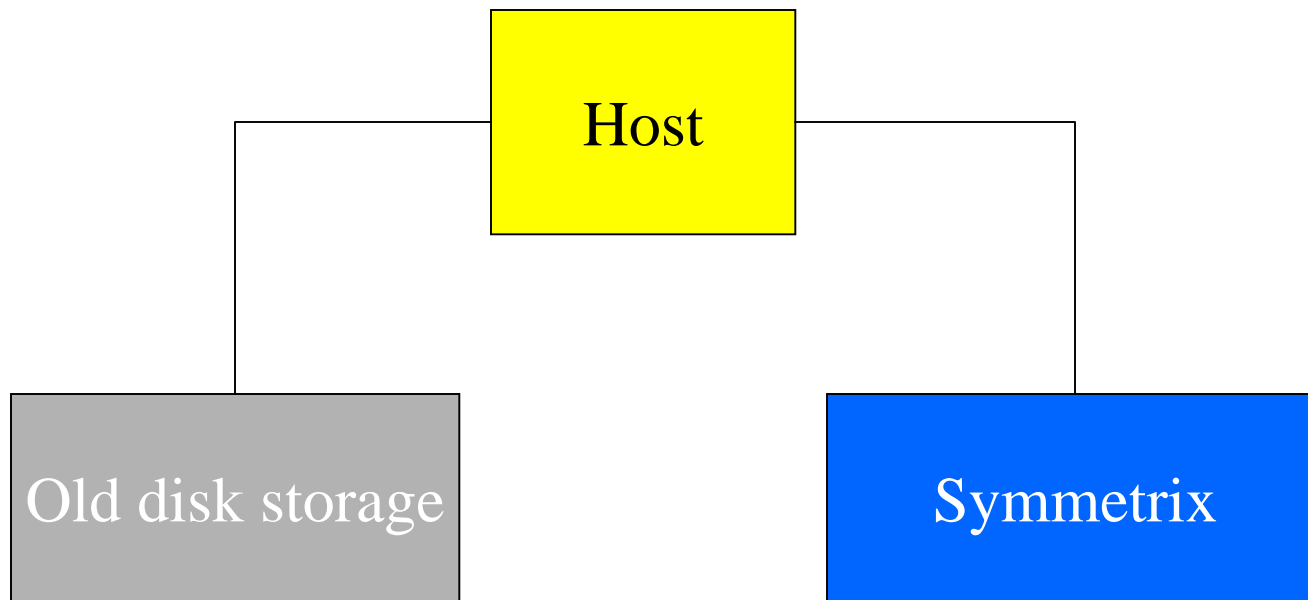
/dev/rdisk/c7t6d2	:EMC	:SYMMETRIX	:5265	:99139000	:7619520
/dev/rdisk/c7t6d3	:EMC	:SYMMETRIX	:5265	:9913A000	:7619520
/dev/rdisk/c7t6d4	:EMC	:SYMMETRIX	:5265	:9917C000	:7619520
/dev/rdisk/c7t6d5	:EMC	:SYMMETRIX	:5265	:99079000	:76195200
/dev/rdisk/c7t6d6	:EMC	:SYMMETRIX	:5265	:9908D000	:76195200
/dev/rdisk/c7t6d7	:EMC	:SYMMETRIX	:5265	:990BF000	:76195200
/dev/rdisk/c7t7d0	:EMC	:SYMMETRIX	:5265	:9913D000	:76195200
/dev/rdisk/c7t7d1	:EMC	:SYMMETRIX	:5265	:99147000	:76195200
/dev/rdisk/c7t15d0	:EMC	:SYMMETRIX	:5265	:99191000	: 2880

How to Move your Data

Use LVM mirroring to seamlessly move data from one storage device to another.

First, establish a connection to both devices.

How to Move your Data



How to Move your Data

For example, you have an existing volume group on drive c1t5d0 and your new Symmetrix disk is at c8t1d0.

Add the new disk to the volume group.

```
vgextend vg01 /dev/dsk/c8t1d0
```

Mirror all of the logical volumes to the new disk.

```
vgdisplay -v vg01 | grep "LV Name" | \  
awk '{ print "lvextend -m 1 " $3 " /dev/dsk/c8t1d0" }' \  
| sh
```

How to Move your Data

After the logical volumes are all synced, remove the first mirrored drive.

```
vgdisplay -v vg01 | grep "LV Name" | \  
awk '{ print "lvreduce -m 0 " $3 " /dev/dsk/c1t5d0" }' \  
| sh
```

Then remove the first drive from the volume group.

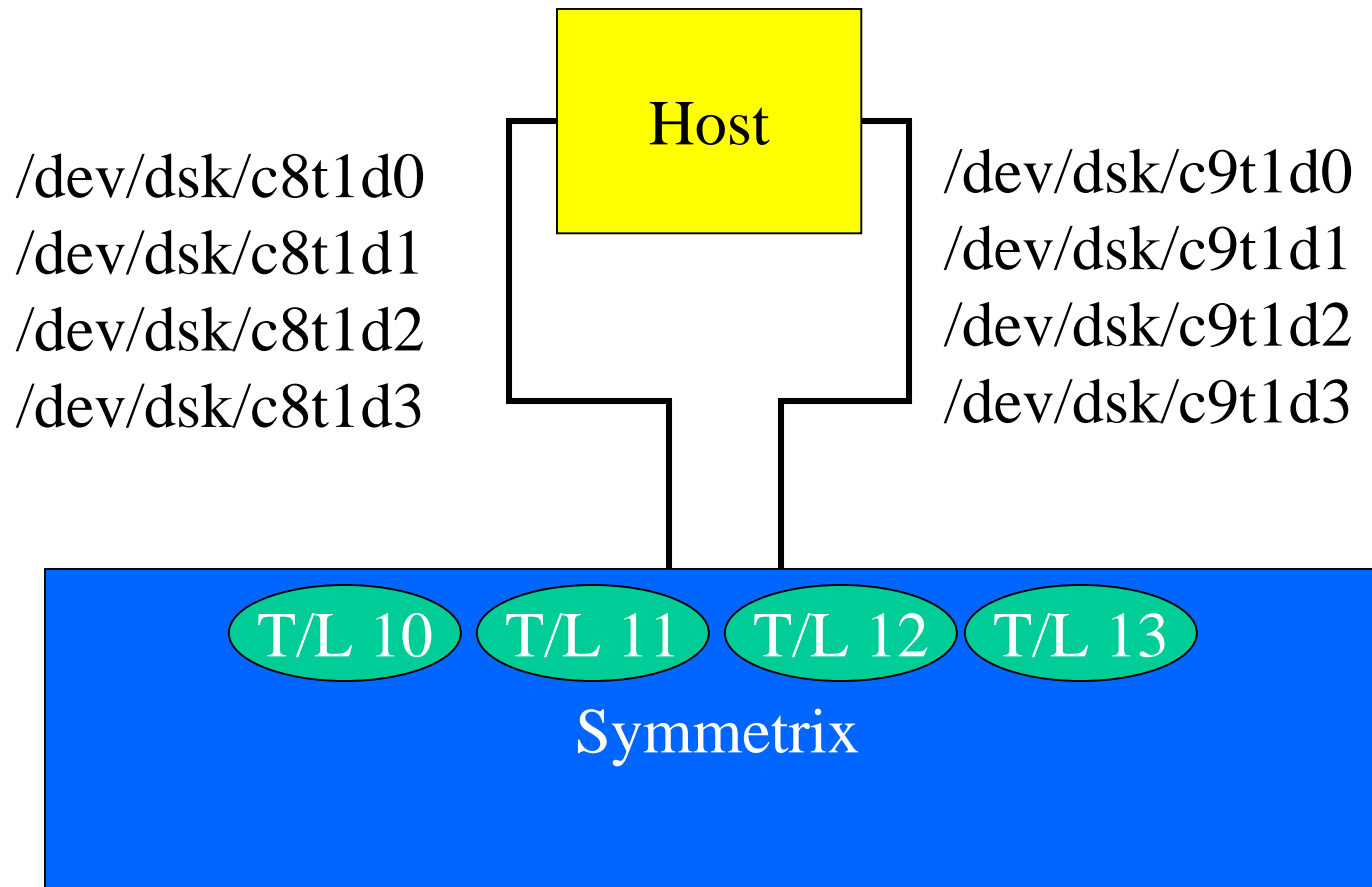
```
vgreduce vg01 /dev/dsk/c1t5d0
```

Now you are ready to disconnect from the old storage.

Setting Up Alternate Links

- Alternate links, or PV links, identify dual paths to the same disk
- Provides failover and can provide some load balancing

Setting Up Alternate Links



Setting Up Alternate Links

- For example, you want 1 logical volume that exists across the 4 drives
- Add the drives to the volume group such that the drives are staggered across the two paths
- For the best performance, stripe the logical volume across the 4 drives and both interfaces

Setting Up Alternate Links

```
pvcreate /dev/dsk/c8t1d0  
pvcreate /dev/dsk/c8t1d1  
pvcreate /dev/dsk/c8t1d2  
pvcreate /dev/dsk/c8t1d3
```

```
vgcreate -p 90 -e 5000 vg01 /dev/dsk/c8t1d0  
vgextend vg01 /dev/dsk/c9t1d1  
vgextend vg01 /dev/dsk/c8t1d2  
vgextend vg01 /dev/dsk/c9t1d3
```

```
vgextend vg01 /dev/dsk/c9t1d0 /dev/dsk/c8t1d1 \  
/dev/dsk/c9t1d2 /dev/dsk/c8t1d3
```

```
lvcreate -r N -L 1024 -i 4 -I 64 -n lvoll1 vg01
```

Setting Up Alternate Links

```
vgdisplay -v vg01
```

```
...
```

```
PV Name          /dev/dsk/c8t1d0
PV Name          /dev/dsk/c9t1d0 Alternate Link
PV Status        available
Total PE         1859
Free PE          1795

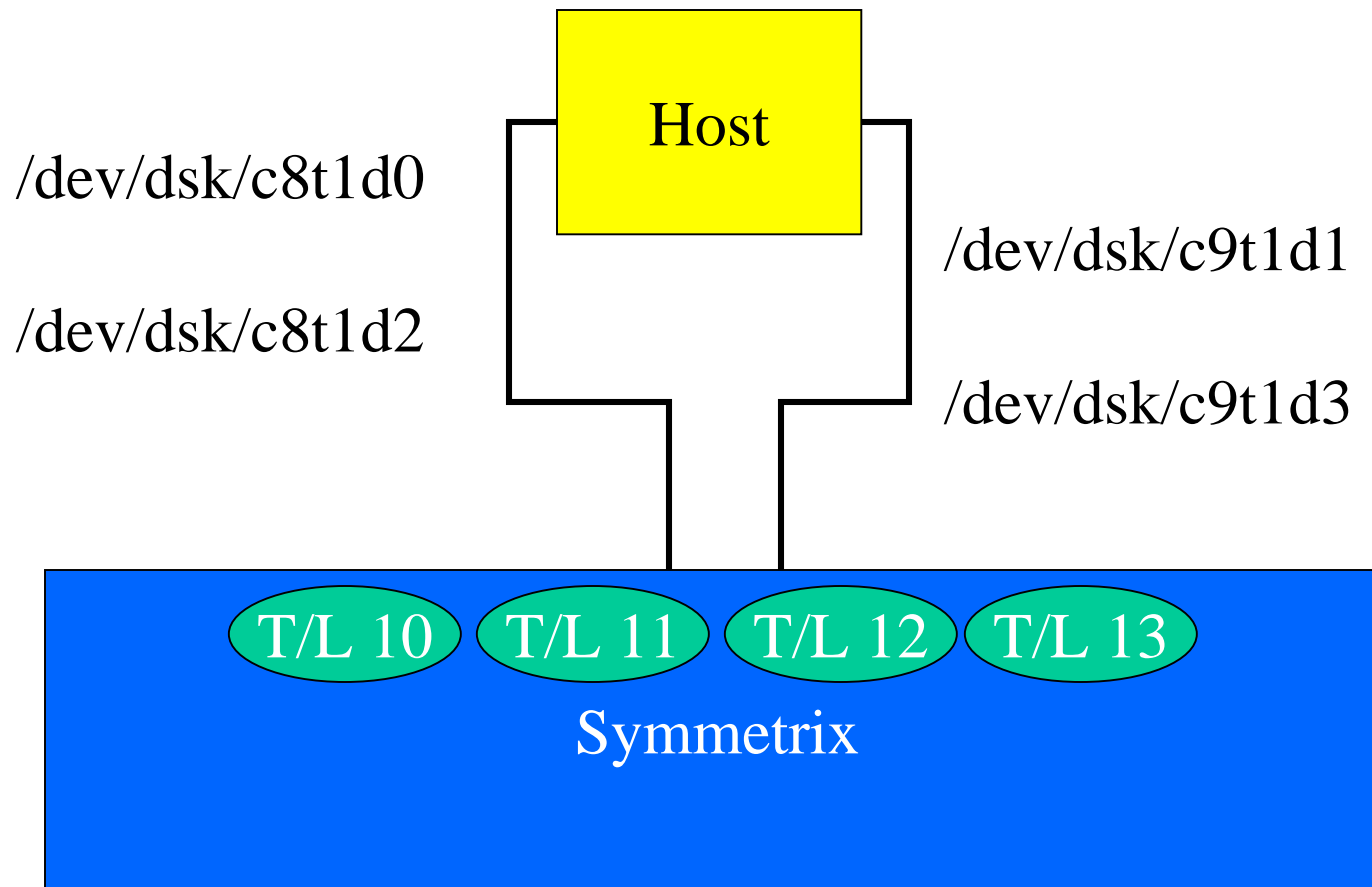
PV Name          /dev/dsk/c9t1d1
PV Name          /dev/dsk/c8t1d1 Alternate Link
PV Status        available
Total PE         1859
Free PE          1795
```

Setting Up Alternate Links

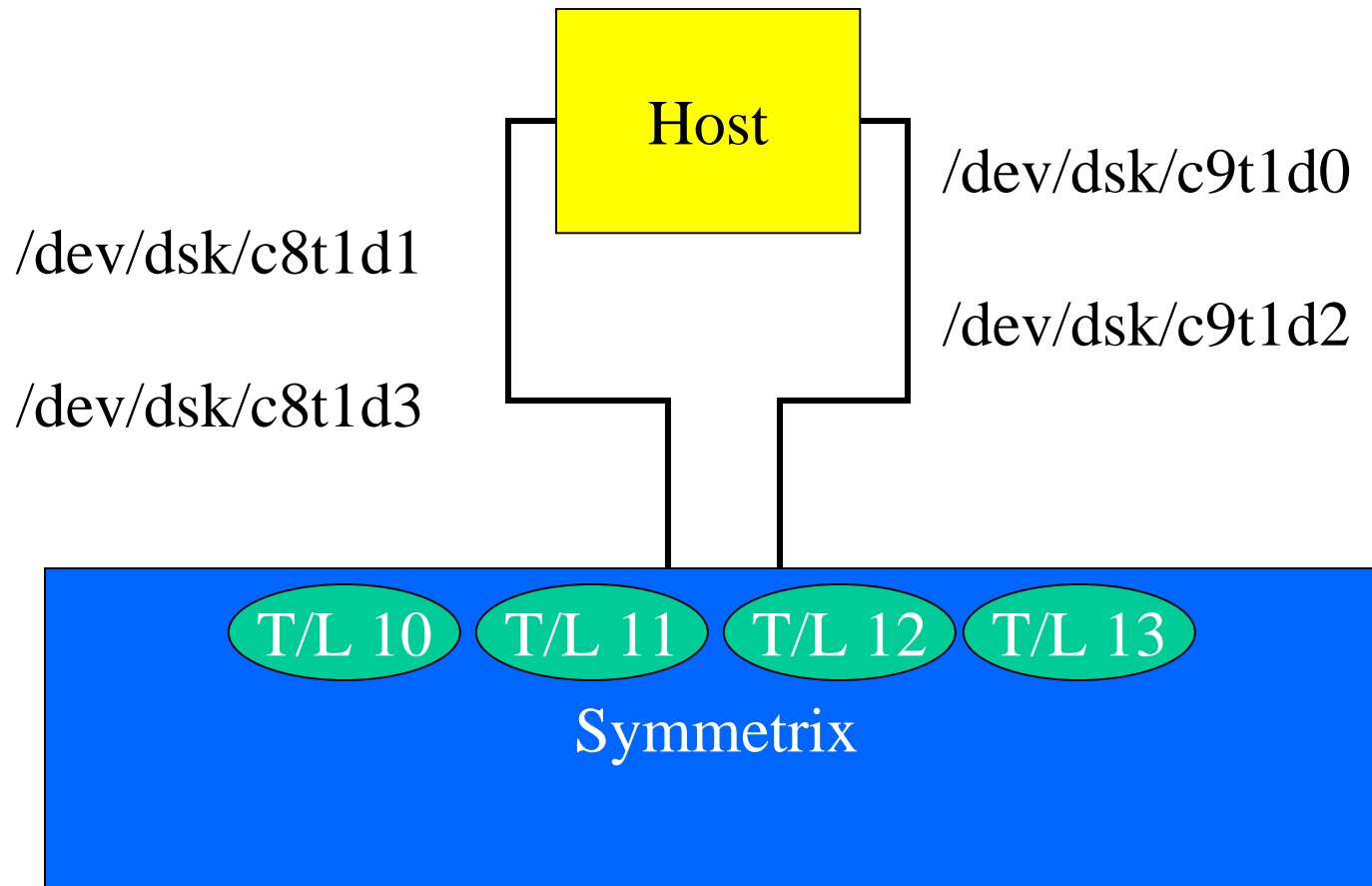
...

PV Name	/dev/dsk/c8t1d2
PV Name	/dev/dsk/c9t1d2 Alternate Link
PV Status	available
Total PE	1859
Free PE	1795
PV Name	/dev/dsk/c9t1d3
PV Name	/dev/dsk/c8t1d3 Alternate Link
PV Status	available
Total PE	1859
Free PE	1795

View of Primary Links

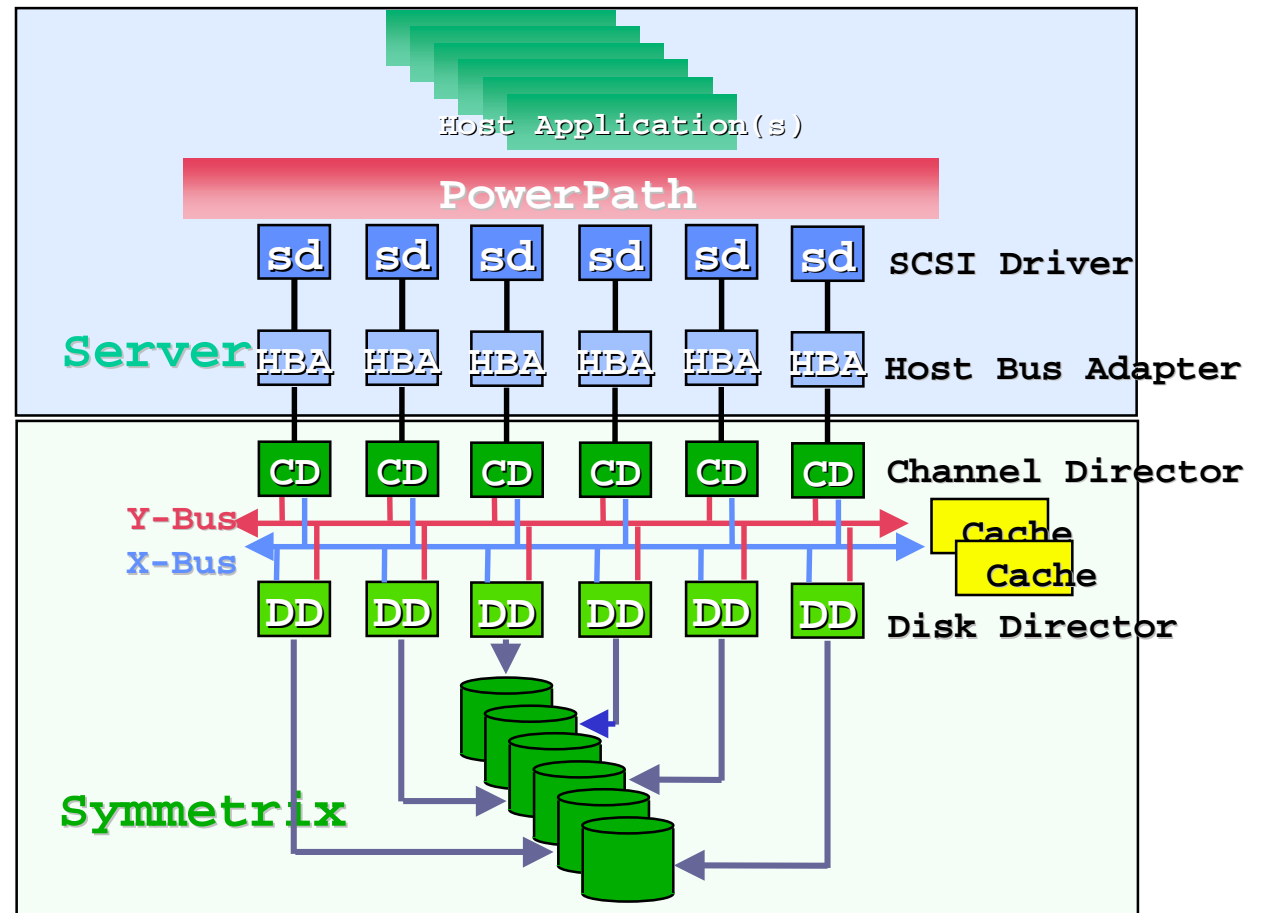


View of Alternate Links



EMC² PowerPath

- Multiple paths to a logical device
- Dynamic load management across paths
- On-line configuration and management
- Automatic detection and management of failed paths
- On-line path recovery after repair



It “Never” Goes Down

But what to do when it does

- The Symmetrix needs to go through a scheduled shutdown for several reasons
 - Upgrading a family of microcode
 - Certain configuration changes
 - Fragmented cache
 - Physical move
- You can use LVM mirroring techniques to keep your system online

It “Never” Goes Down

But what to do when it does

- If using alternate links, connect one of them to another disk array. (Ok, not everyone can afford two Symmetrixes)
- Configure an equal amount of storage on the other box
- Run ioscan to see the new drives
- Use LVM mirroring to create a copy of your data on the new box

Questions?

