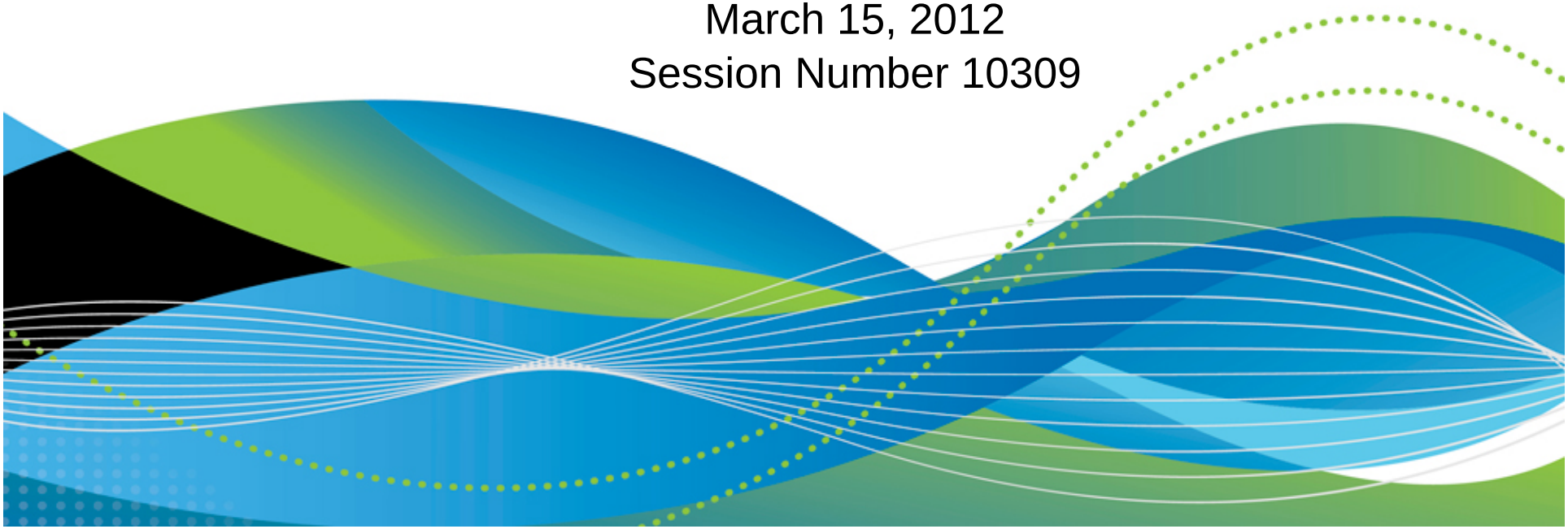
#SHAREorg

SHARE
Technology · Connections · Results

Replication Considerations for Linux on System z

Brad Hinson, Red Hat
Gail Riley, EMC

March 15, 2012
Session Number 10309

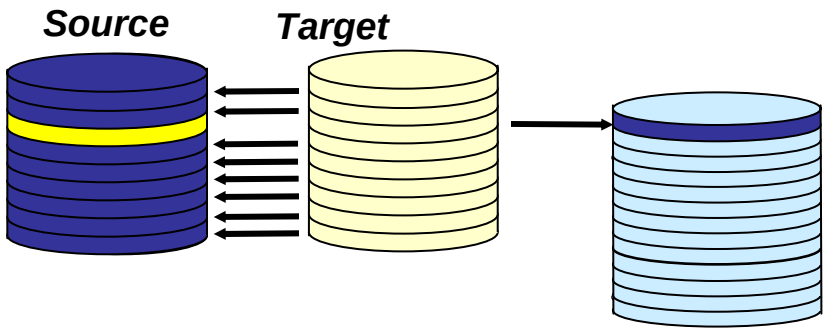


Objectives

After completing this session, you will be able to:

- Discuss the considerations when implementing replication
- Understand the Red Hat clone process
- Describe the tasks for accessing a Local and Remote replica in a Linux on System z environment

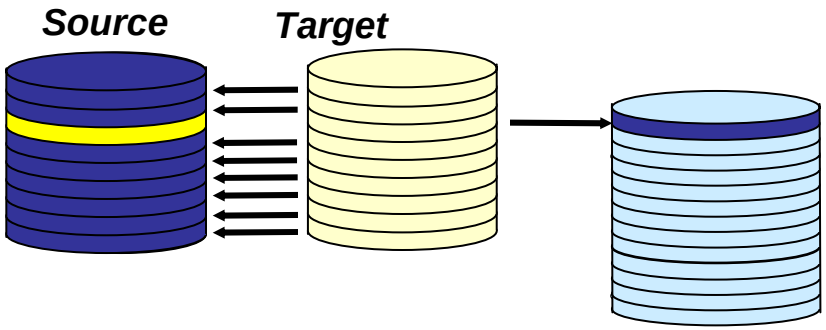
Local Replication



Forms of Local Replication

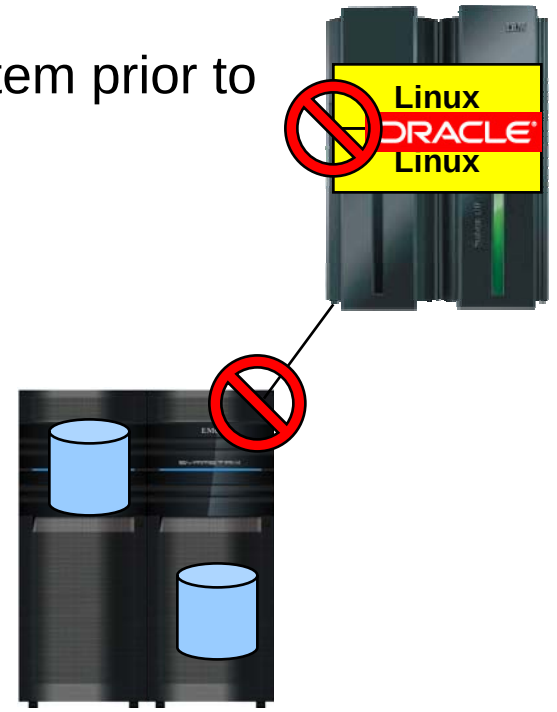
- Full Volume Copy - Clone
  - Data is copied from the Source Device to a Target Device of equal size and emulation

- Pointer Based Replication - Snap
  - The Target Device is a virtual device housing a collection of pointer between the Source and a reserve area for a point-in-time view



Creating a Local Consistent Copy

- Different options depending on application and host requirements
- Server
  - Pause I/O at the Server Level to provide a Consistent Point-in-Time Copy
- Application
  - Stop the application and unmount the file system prior to activate or split
  - Database hot backup mode
  - Database freeze/thaw
- Array
  - Holds IO at the array until replica completes
    - Symmetrix example: Enginuity Consistency Assist (ECA) holds IO at the Symmetrix until all Splits/Activate complete



Replication Options

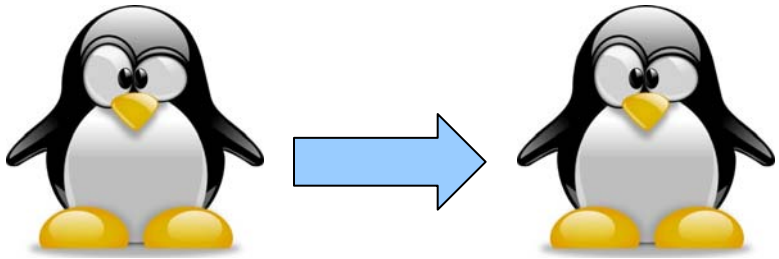
Linux Operating Systems utilities

- Red Hat clone rpm – local replication
- rsync for remote directory refresh

- Storage array supplied replication process for local and remote replication

- Create your own local replication process

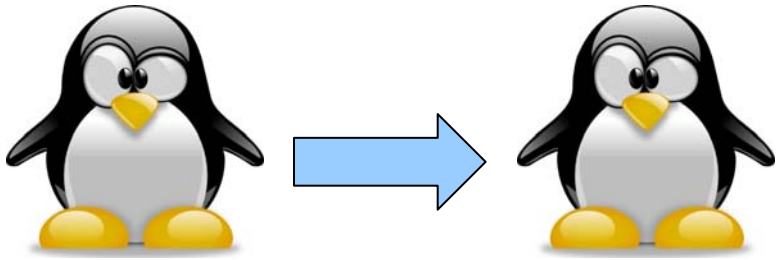
Red Hat Clone rpm



- Provided with RHEL Virtualization Cookbook
  - http://www.vm.ibm.com/devpages/mikemac/SG247932.tgz
  - http://people.redhat.com/bhinson/clone/ (latest copy)
- Requirements
  - Cloner guest, source guest (separate guests, cloner can't clone itself)
  - z/VM user definition for new/target clone must exist
  - Cloner must have privilege class B for FlashCopy and attach*
  - For “dd” options, cloner must LINK disks to copy
    - OPTION LNKNOPAS or
    - LINK password set to “ALL” for read & write
    - MDISK definitions for DASD, not DEDICATE
  - For LVM installs, cloner Volume Group name must be different from source

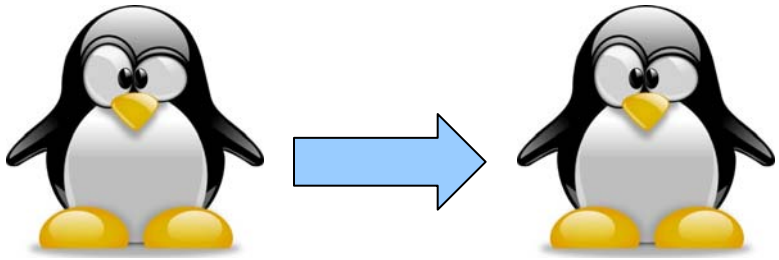
*attach is used for FCP port access

Red Hat Clone rpm



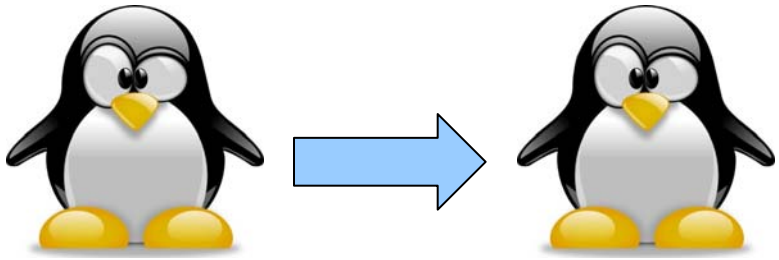
- Configuration file (/etc/sysconfig/clone)
  - AUTOLOG=
    - Boot guest automatically after cloning
  - CLONE_METHOD=
    - FlashCopy “auto” or Linux “dd”
  - CLONE_FCP=
    - symclone or Linux “dd”
- Clone configuration files (/etc/clone)
  - rhel.conf.sample: sample values. Copy to {target ID}.conf
  - Similar values can be copied to shared.conf

Red Hat Clone rpm



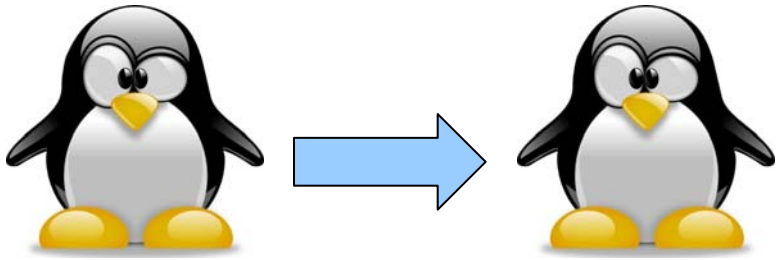
```
# rpm -ivh clone-1.0-12.s390x.rpm
Preparing...                ########################################### [100%]
   1:clone                  ########################################### [100%]

# cp /etc/clone/rhel.conf.sample /etc/clone/newguestID.conf
# vi /etc/clone/newguestID.conf

# clone -v masterguestID newguestID

This will copy disks from masterguestID to newguestID
Host name will be:  newguestID.s390.bos.redhat.com
IP address will be: 10.16.105.65
Do you want to continue? (y/n): y

[…]
Invoking Linux command: dasdfmt -p -b 4096 -y -F -f /dev/dasdd
cyl  3338 of  3338 |#################################################| 100%
Invoking Linux command: dd bs=4096 count=600840 if=/dev/dasdc of=/dev/dasdd
[…]
```

Red Hat Clone rpm



- CLONE_FCP=dd
  - Read zFCP configuration on source system
  - Specify zFCP configuration of target system
    - /etc/clone/zfcp-{target}.conf
  - Attach source and target FCP port to cloner
  - Clone will bring both sets of LUNs online, use Linux “dd” to copy
- CLONE_FCP=symclone
  - Specify device group in configuration (SYMDG=)
  - Clone calls Symmetrix command-line utilities:
    - symclone {create, activate}
    - symclone {verify} gives updates until copy complete
    - symclone {terminate} to break connection

Red Hat Clone rpm



clone -v masterguestID newguestID
[…]

Calling symclone to copy FCP disks ...

Execute 'Create' operation for device group
'*clone-dg*' (y/[n]) ? **y**
[…]
Execute 'Activate' operation for device group
'*clone-dg*' (y/[n]) ? **y**
[…]
waiting for symclone to complete...
None of the devices in the group '*clone-dg*' are in 'Copied' state.
None of the devices in the group '*clone-dg*' are in 'Copied' state.
[…]
All devices in the group '*clone-dg*' are in 'Copied' state.

Execute 'Terminate' operation for device group
'*clone-dg*' (y/[n]) ? **y**

Clone Vendor Specific Prerequisites

- On z/VM or the Linux instance where the clone will be executed
  - IBM FLASH Copy (z/VM)
    - optional licensed feature on IBM array
    - authorized to execute the command
  - EMC Clone using Solutions Enabler is required (Linux)
    - Access to 1 gatekeeper
    - Device Group for command line operations
  - Native Linux commands (i.e. dd)
    - Target and source devices must be available to Linux
  - z/VM DDR
    - Target and source devices must be available to VM

CKD Local Replication Considerations

- Changes may be required for CKD local replication
- Create unique directory for local replication of Linux guest virtual machine
- Minidisks
  - ensure unique z/VM VOLSER
  - mdisk rdev devno – same as DEDICATEd devices
- DEDICATE/ATTACH
  - Change source (production) real device address to clone real device address
- Use virtual addresses to mask changes at the Linux layer

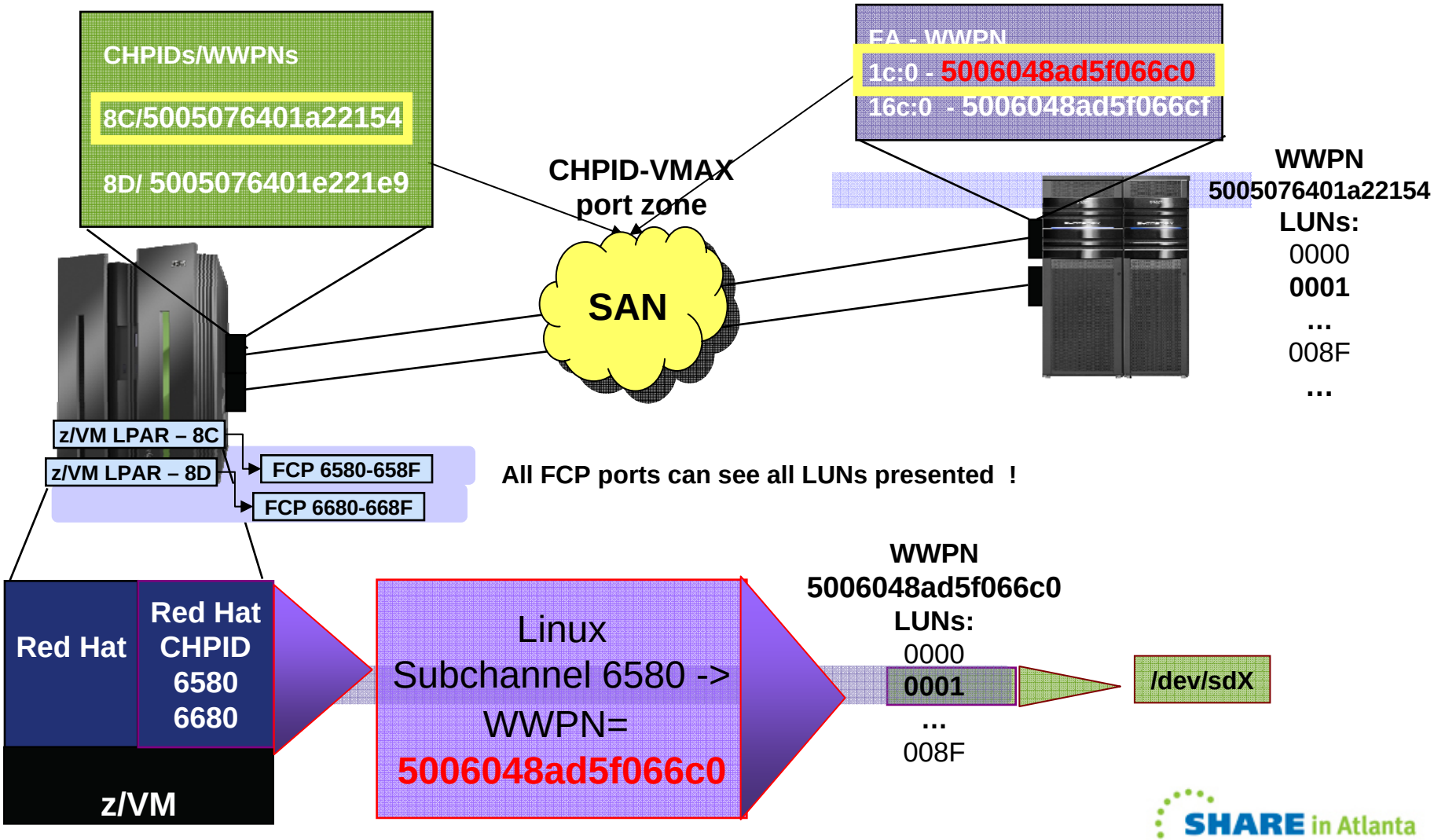
SCSI Replication Considerations

- Storage Area Network (SAN)
  - Zoning – Provides connectivity between end points
    - CHPID to Array Front-end port
  - Mapping – Array presentation of the SCSI device for host access
  - Masking – Provides access to host for specific devices on the channel
- Node-Port ID Virtualization (NPIV)
  - Managing the changing or consistency with port World Wide Names (WWN)

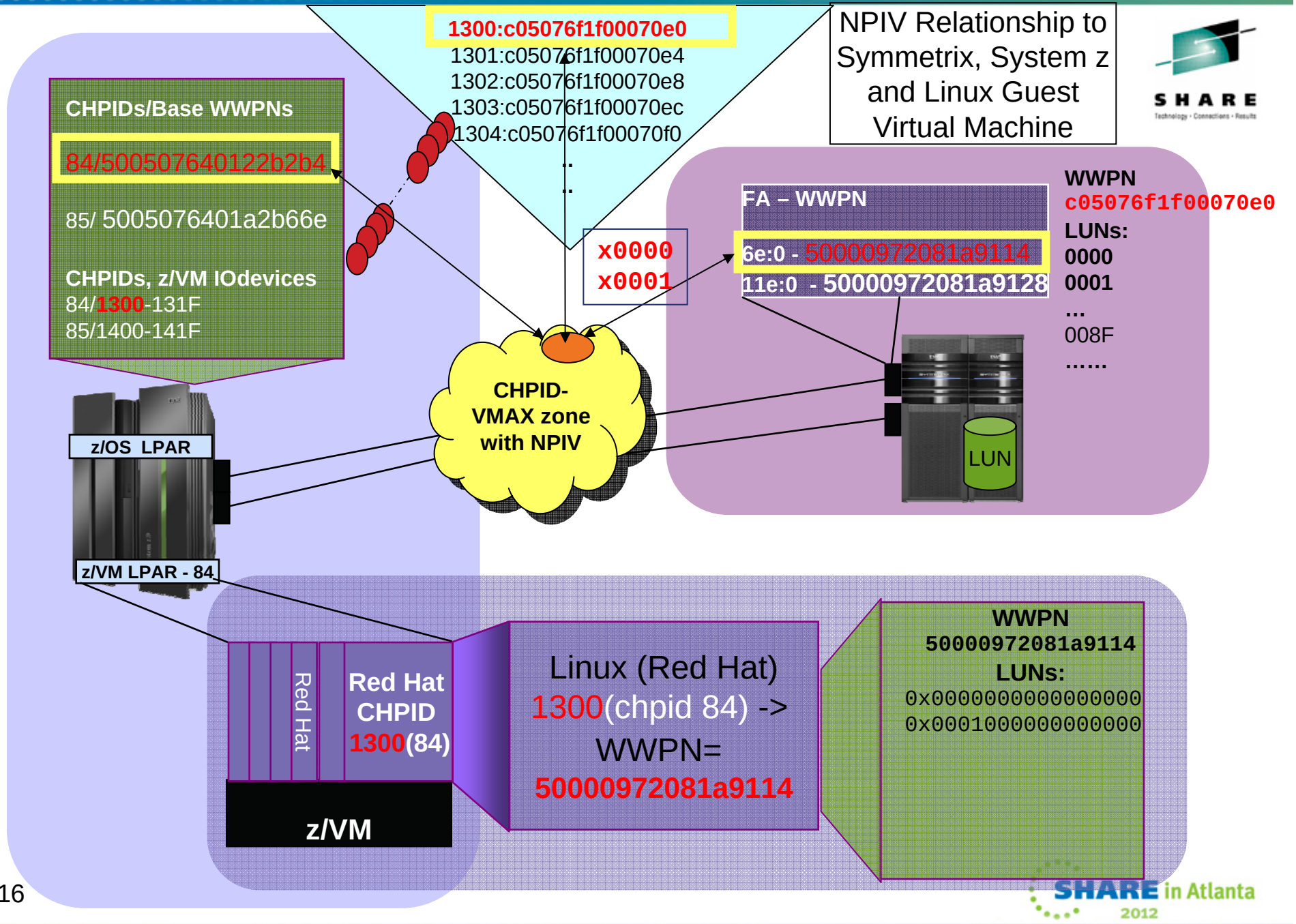
FCP Path Relationship without NPIV

(z/VM Channel/subchannel device) + (Symmetrix port WWPN + LUN (Symmetrix Logical Volume))
(6580) + (5006048ad5f066c0 + 0001) = /dev/sdX

6581-658F



NPIV Relationship to Symmetrix, System z and Linux Guest Virtual Machine



SCSI Local Replication Considerations

- Use a different, unique WWxN for your locally replicated SCSI devices
- WWxN will change
  - With NPIV
    - a different FCP port on the same CHPID
    - the same FCP port on a different LPAR
    - any FCP port on another CHPID
  - Without NPIV - different CHPID
- WWxN will not change, with no NPIV, and any port on same CHPID
  - This means *ALL* LUNs mapped and masked to a CHPID (WWxN) may be seen through all FCP ports/subchannels on the CHPID (thus all attached virtual machines)

FCP/SCSI - VM USER Directory

- Create a separate USER directory for the production and clone instances
- Make sure the WWxN is unique between production and clone instances
- When using NPIV allocate different FCP port(s) for unique WWxN
- IF not using NPIV, use a different CHPID

- Production Site 1 and/or 2

```
USER PR192166
........
* FCP for R1 site - Prod
dedicate 1330 1330
dedicate 1430 1430
* FCP for R2 site
dedicate 1010 1010
dedicate 1011 1011
........
```

- Clone Site 1 and/or 2

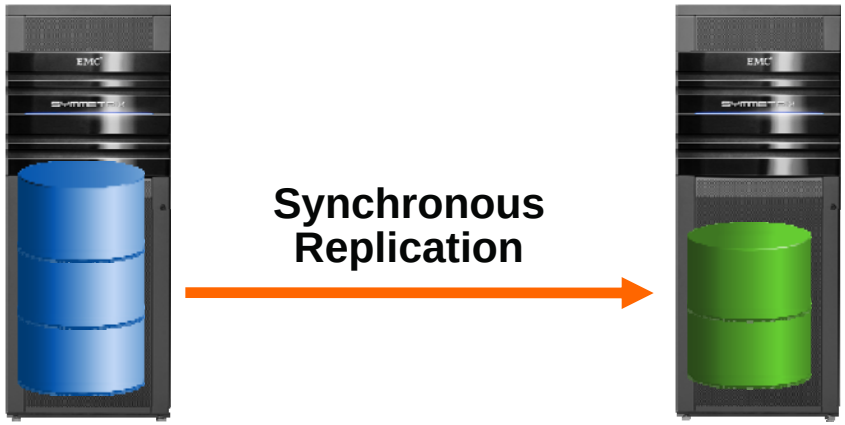
```
USER CL192166
........
* FCP for R1 site - R1 CLONE
  dedicate 1331 1331
  dedicate 1431 1431
* FCP for Site 2 - R2 Clone
 dedicate 101a 101a
 dedicate 111a 111A
........
```

SCSI Local Replication Considerations

- Goal is to make replica unique from production instance
- Update specific Linux files
- Use scripting at the Linux level
- Use Logical Volume Manager (LVM)

- Same considerations for remote replication

Remote Replication

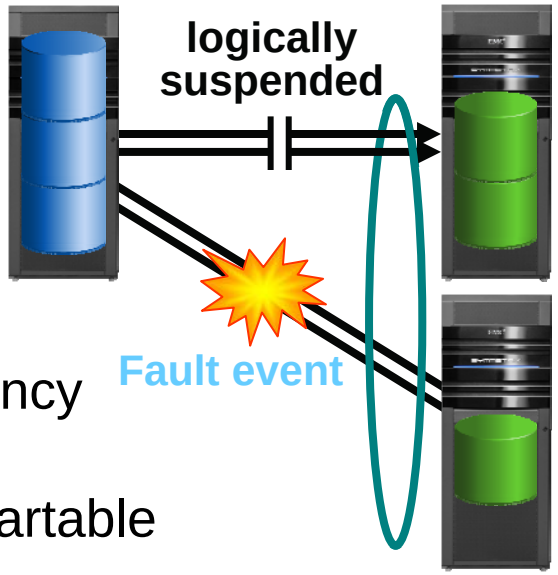


Forms of Remote Replication

- ***Synchronous Replication***
  - Identical copies of data across storage systems where writes are committed across to remote systems/sites first which increases execution time
    - ***Source = Target***
- ***Asynchronous Replication***
  - Data is a point-in-time consistent copy but writes happen locally and are sent across to remote systems/sites at a periodic interval
    - ***Source ≅ Target***
- Data Distribution – Data Migration ONLY!
  - Data is copied from one storage system to another without maintaining a consistent recoverable copy
    - ***Source ≠ Target***

Remote Replication Consistency

- Preserves dependent-write consistency of devices
  - Ensures application dependent write consistency of the application data remotely mirrored operations in the event of a disaster
- Host
  - I/O held at host
- Array
  - I/O held at Array
    - Symmetrix example: SRDF/CG (Consistency Groups) use host software to suspend replication guaranteeing a consistent restartable image at remote site



Ensures dependent-write consistency of the data remotely mirrored

CKD Remote Replication Considerations

- Changes may be required for CKD remote replication, but it depends…..
- Minidisks
  - Full or partial – if replicating all z/VM volumes also, no directory changes needed at remote site
  - VOLSER will be the same at local and remote site
  - mdisk rdev – same as DEDICATE
- DEDICATE/ATTACH
  - No change if real device address is the same at the primary and backup site
- Use virtual addresses to mask changes at the Linux layer

SCSI Remote Replication Considerations

- WWxN will change for remotely replicated SCSI devices because there is a physically different FCP port and CHPID on a different CEC
  - This applies to NPIV and non NPIV environments
- Linux to recognize the new WWxN and find its data
  - Same issue as local replica access
  - Update specific Linux files
  - Use scripting at the Linux level
  - Use Logical Volume Manager (LVM)

FCP/SCSI - VM USER Directory

- if replicating all z/VM volumes add Site 1 and Site 2 FCP ports to User directory
  - FCP port device addresses should be unique for each site

- Production Site 1 and/or 2

```
USER PR192166
........
* FCP for R1 site
dedicate 1330 1330
dedicate 1430 1430
* FCP for R2 site
dedicate 1010 1010
dedicate 1011 1011
........
```

Minimize changes to Linux for failover

- Use Linux facilities already in place when using NPIV
  - /etc/zfcp.conf - List second site (DR) entries also along with Site 1
  - Correct paths will be found at each site
  - Pro: Updates are made in one location
  - Con: Linux will still search for FCP port of other site at startup

```
# site 1 R1 path
0.0.1330 0x50000972081a9114 0x0000000000000000
0.0.1330 0x50000972081a9114 0x0001000000000000
.....................
#
# site 1 R1 path
0.0.1430 0x50000972081a9128 0x0000000000000000
0.0.1430 0x50000972081a9128 0x0001000000000000
.......
#
# site 2 R2 path
0.0.1010 0x50000972081acd59 0x0000000000000000
0.0.1010 0x50000972081acd59 0x0001000000000000
.....
# site 2 R2 path
0.0.1110 0x50000972081acd65 0x0000000000000000
0.0.1110 0x50000972081acd65 0x0001000000000000
```

Red Hat Multipathing

- /etc/multipath.conf – basic configuration file
  - Created and maintained by the multipath program
  - /etc/multipath/bindings
  - /etc/multipath/wwids
- Both files contain wwid for each device with different entries for Site 1 and Site 2 ➔ different physical device
  - Site1

    360000970000192601700533030383737
  - Site2

    360000970000192601715533030333032

Use LVM with Replicated Copies

- LVM masks the changing SCSI multipath information
- Volume groups (VG) are made up of LVM physical volumes (PVs)
- LVM physical volumes are identified by PV UUID, not multipath device UUID/WWID
- Logical volumes(LVs) are associated to LVM volume groups
- Filesystems are associated to logical volumes in /etc/fstab
- All LVM entities are found, brought online and the filesystem mounted at Site 2, no different than Site 1

Test the Replication Environment

- Clones/Snaps can be used at the Primary or DR site
  - Ensure consistency across all devices at time of clone creation
- System Considerations - Make sure you have a unique environment for your clone
  - Create a separate VM directory entry for clone use
  - CKD minidisks
    - make sure the VOLSER is unique if using minidisks
  - DEDICATE/ATTACH
    - make sure the same *virtual* address is used
  - Change the network – IP address, DNS as appropriate
  - Use different NPIV/WWxn ports than the production environment

Application Considerations when Cloning

- Does it start up automatically?
- Does it connect to another application, IP address?
- Does it use a NFS mounted filesystem?
- Does it export information when it starts?
- Does it download or upload information when it starts or sometime during its instantiation?
- Does the application rely on a specific
  - Hostname
  - IP address
  - raw device
- Identify any application interdependencies

Linux Replication Considerations

- Both Local and Remote Replication have device considerations
  - CKD and/or FBA devices are supported
  - Use device-by-path, not device-id for device setup
  - Use the same virtual address for replicated devices
  - Make sure SCSI LUN mapping is the same at both sites
  - Let LVM assist you in reducing changes for replicated copies
- Other considerations
  - Automate the process wherever possible
  - Standardize wherever possible, i.e., addressing scheme for operating system, application, other devices
  - Shared R/O Linux kernel –
    - May create unintended interdependencies between (application) environments
    - One environment can force another to upgrade
  - Are there cron jobs you need to disable on the clone
  - Don’t forget about backups at the DR site

Discussion Topic Recap

- Replication methods
  - Home grown (i.e. DDR) vs. Native Red Hat Linux clone
  - Synchronous vs asynchronous
- Local vs. Remote replication considerations
- CKD and SCSI replication considerations
- Use of LVM to handle replication failover
- Linux considerations