Automating Oracle Database Resiliency on System Z Linux

Linux on System z Executive Advisory Customer Council
Friday March 15th, 2013
Michael Mac Isaac - mikemac at us.ibm.com

© 2012 IBM Corporation

Trademarks

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

IT Infrastructure Library is a registered trademark of the Central Computer and Telecommunications Agency which is now part of the Office of Government Commerce.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Windows Server and the Windows logo are trademarks of the Microsoft group of countries.

ITIL is a registered trademark, and a registered community trademark of the Office of Government Commerce, and is registered in the U.S. Patent and Trademark Office.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Java and all Java based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

* Other product and service names might be trademarks of IBM or other companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

© 2012 IBM Corporation

Abstract

This presentation describes how to automate z/VM and Linux® for Oracle® "Standalone" database and "Grid" (aka cluster). It will frame today's cloud terminology of *Infrastructure/Platform/Software as a Service* into concrete terms and steps that should be understandable to System z IT professionals. Look for the methodology described to be published in a forthcoming IBM Redbook.

© 2012 IBM Corporation

Agenda

- Introductions
- Cloud defined as I/P/SaaS
- Infrastructure as a Service
  - Configure z/VM
  - Define virtual machines
- Platform as a Service
  - Prepare to install Linux on the golden image
  - Install Linux on the golden image
  - Configure the 6.2 golden image
  - Cloning
- Software as a Service for Oracle Standalone
  - Configure a Linux system for the Oracle boot script
  - Clone a virtual server
  - Silently install Oracle database
- Software as a Service for Oracle Grid
- Miscellaneous

© 2012 IBM Corporation

Introductions

- Who am I?
  - Michael MacIsaac
- Who are you?
  - The audience

© 2012 IBM Corporation

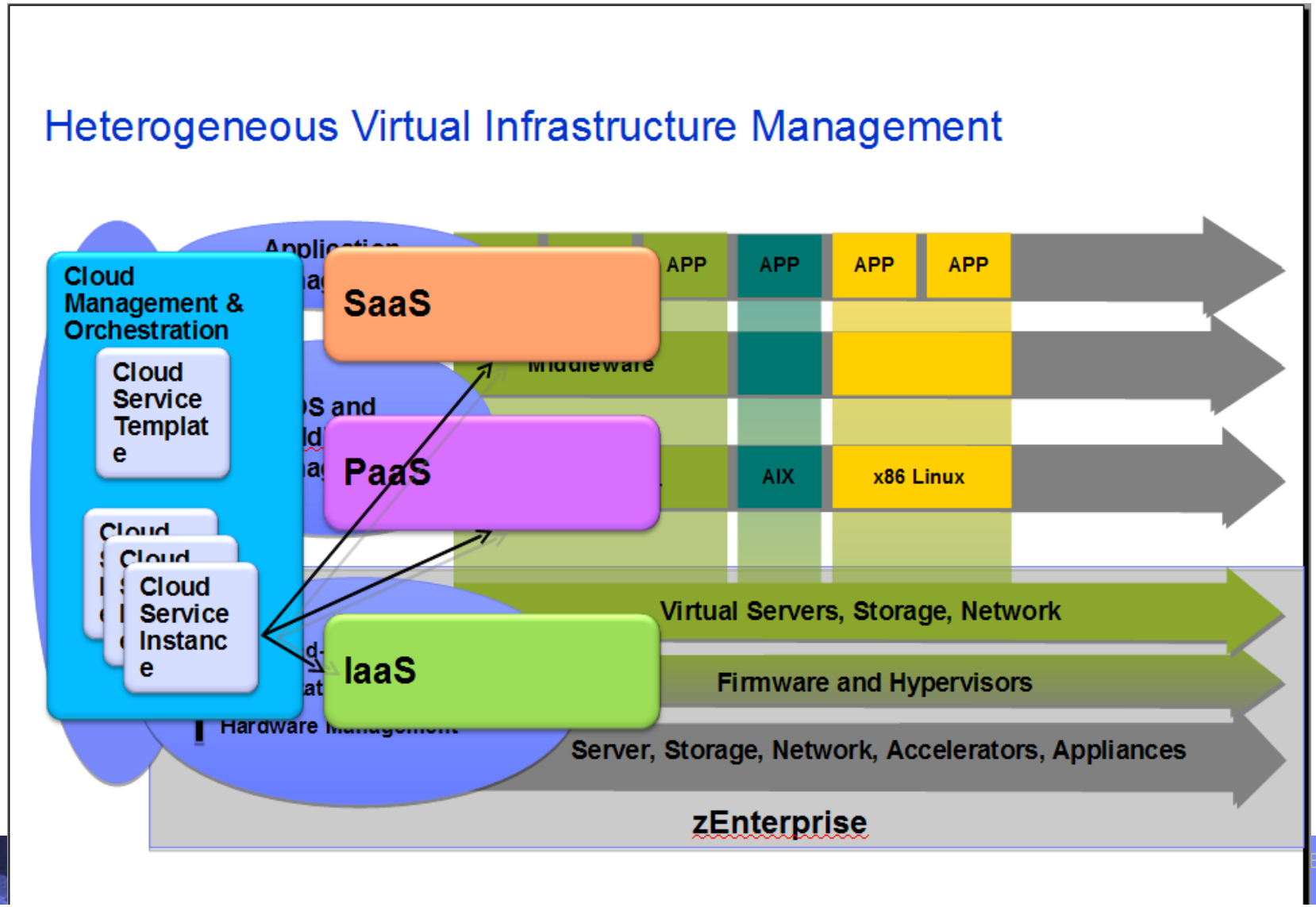
Agenda

- ~~Introductions~~
- **Cloud defined as I/P/SaaS**
- **Infrastructure as a Service**
  - **Configure z/VM**
  - **Define virtual machines**
- Platform as a Service
  - Prepare to install Linux on the golden image
  - Install Linux on the golden image
  - Configure the 6.2 golden image
  - Cloning
- Software as a Service for Oracle Standalone
  - Configure a Linux system for the Oracle boot script
  - Clone a virtual server
  - Silently install Oracle database
- Software as a Service for Oracle Grid
- Miscellaneous

© 2012 IBM Corporation

Tangent - Cloud defined as I/P/SaaS

- Software as a Service (SaaS) - the application(s)
- Platform as a Service (PaaS) - the (guest) operating system
- Infrastructure as a Service (IaaS) - the virtual machine or "container"



Infrastructure as a Service

- Configure z/VM
  - Define a virtual machine for a common 191 disk on all Linux virtual machines
    - `LNXMAINT 192`
  - Enable TCP/IP
  - Customize SYSTEM CONFIG
    - Define VSWITCHes (layer 2, w/ and w/o OSA), define MACPREFIX, allow VDISKs, etc.
- Define virtual machines (fairly well known tasks)
  - Sample User ID

```
USER LNXSA1 ORACLE 1G 6G G
 INCLUDE LNXDFLT
 MDISK 0100 3390 0001 10016 LX9A1A MR
 MDISK 0101 3390 0001 30050 LX6605 MR
 MDISK 0302 3390 0001 10016 LX9A1B MR
 DEDICATE 0400 B800
 DEDICATE 0500 B900
```

  - Sample PROFILE

```
PROFILE LNXDFLT
 COMMAND SET VSWITCH VSWITCH2 GRANT &USERID
 COMMAND DEFINE NIC 600 TYPE QDIO
 COMMAND COUPLE 600 TO SYSTEM VSWITCH2
 COMMAND SET VSWITCH VSWITCH3 GRANT &USERID
 COMMAND DEFINE NIC 700 TYPE QDIO
 COMMAND COUPLE 700 TO SYSTEM VSWITCH3
 CPU 00 BASE
 . . .
```

© 2012 IBM Corporation

Define virtual machines (cont'd)

- Sample User ID for Oracle cluster (1 of 2)

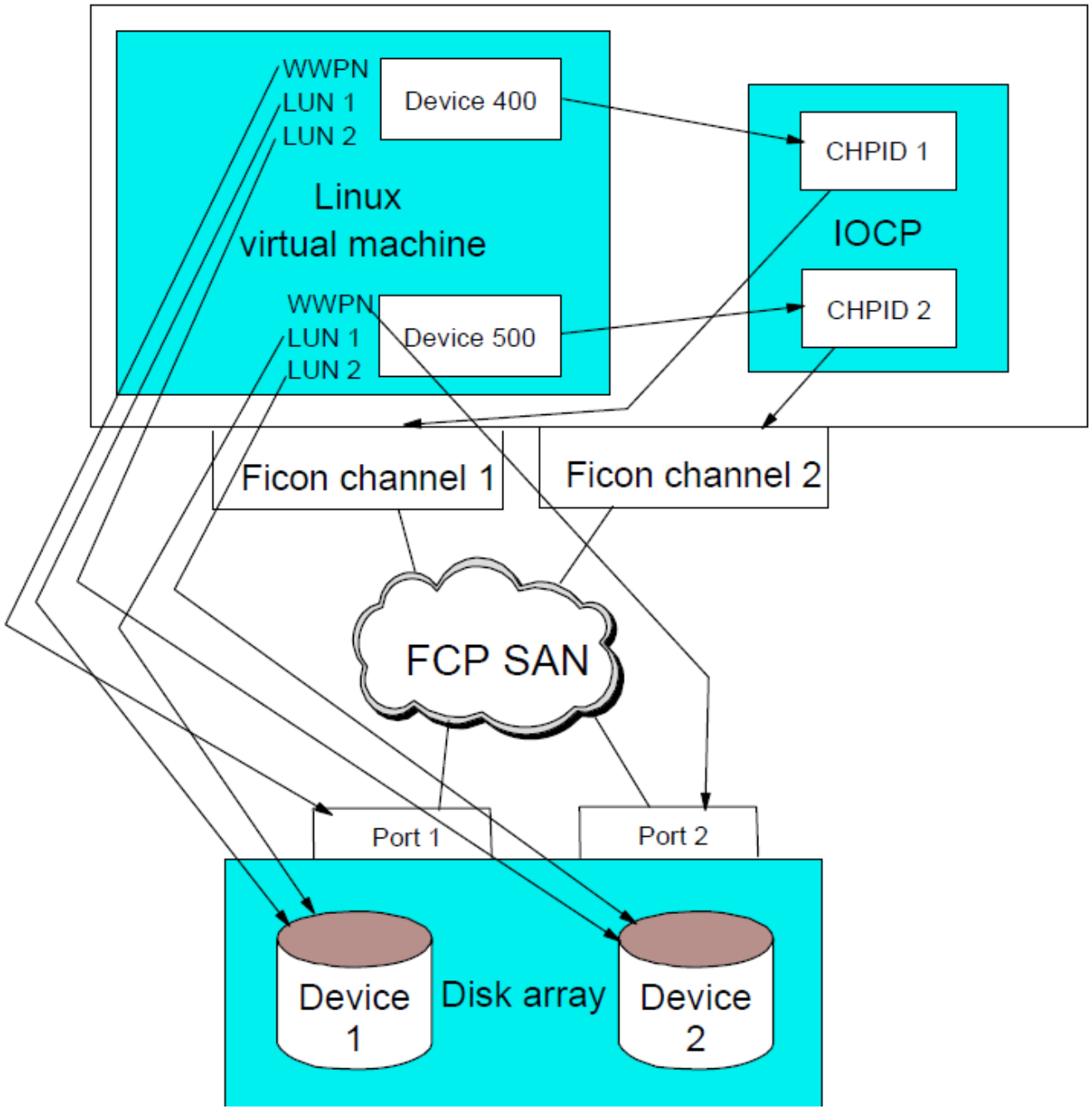
```
USER LNXC2N1 ORACLE 4G 6G G
 INCLUDE LNXDFLT
 MDISK 0100 3390 0001 10016 LX9A1D MR
 MDISK 0101 3390 0001 30050 LX6606 MR
 MDISK 0200 3390 1 1000 LX9A0E MR
 MINIOPT NOMDC
 MDISK 0201 3390 1001 1000 LX9A0E MR
 MINIOPT NOMDC
 MDISK 0202 3390 2001 1000 LX9A0E MR
 MINIOPT NOMDC
 MDISK 0302 3390 20033 10016 LX6705 MR
 MINIOPT NOMDC
 DEDICATE 0400 B803
 DEDICATE 0500 B903
```

- Sample User ID for Oracle cluster (2 of 2)

```
USER LNXC2N2 ORACLE 4G 6G G
INCLUDE LNXDFLT
MDISK 0100 3390 0001 10016 LX9A0A MR
MDISK 0101 3390 0001 30050 LX6702 MR
LINK LNXC1N1 0200 0200 MW
LINK LNXC1N1 0201 0201 MW
LINK LNXC1N1 0202 0202 MW
MDISK 0302 3390 0001 10016 LX9A0B MR
DEDICATE 0400 B804
DEDICATE 0500 B904
```

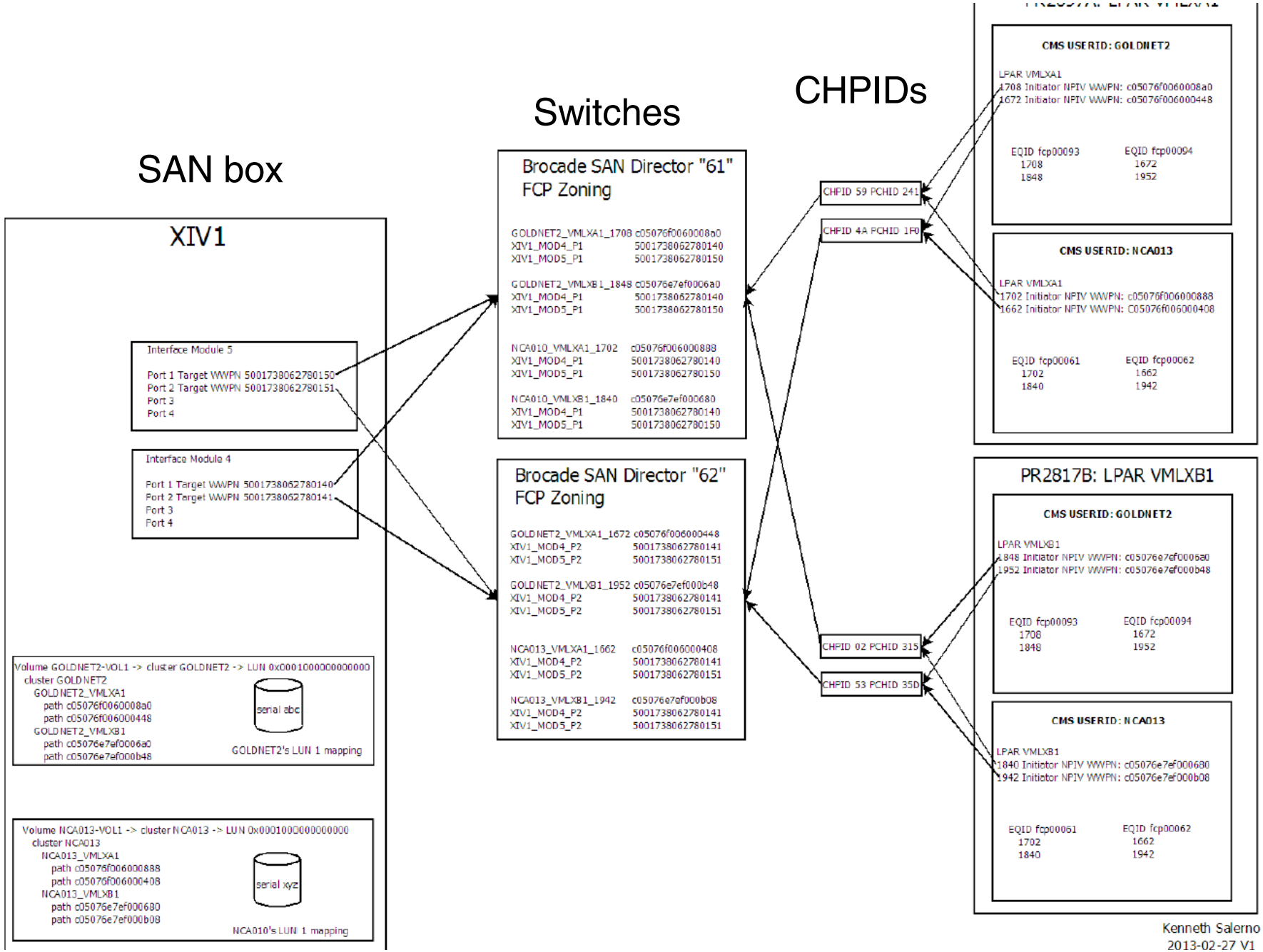
© 2012 IBM Corporation

FCP/SCSI diagrams



© 2012 IBM Corporation

FCP/SCSI diagrams (cont'd)



Agenda

- ~~Introductions~~
- ~~Cloud defined as I/P/SaaS~~
- ~~Infrastructure as a Service~~
  - ~~Configure z/VM~~
  - ~~Define virtual machines~~
- **Platform as a Service**
  - **Prepare to install Linux on the golden image**
  - **Install Linux on the golden image**
  - **Configure the 6.2 golden image**
  - **Cloning**
- Software as a Service for Oracle Standalone
  - Configure a Linux system for the Oracle boot script
  - Clone a virtual server
  - Silently install Oracle database
- Software as a Service for Oracle Grid
- Miscellaneous

© 2012 IBM Corporation

Platform as a Service

- Prepare to install Linux on the golden image
  - Set up CONF and PARM files

```
DASD=100-101,300-302
HOSTNAME=rh62gold.itso.ibm.com
NETTYPE=qeth
IPADDR=9.12.7.2
SUBCHANNELS=0.0.0600,0.0.0601,0.0.0602
NETMASK=255.255.240.0
SEARCHDNS=itso.ibm.com
GATEWAY=9.12.4.1
DNS=9.12.6.7
MTU=1500
PORTNAME=DONTCARE
LAYER2=1
IPADDR2=10.1.1.2
```

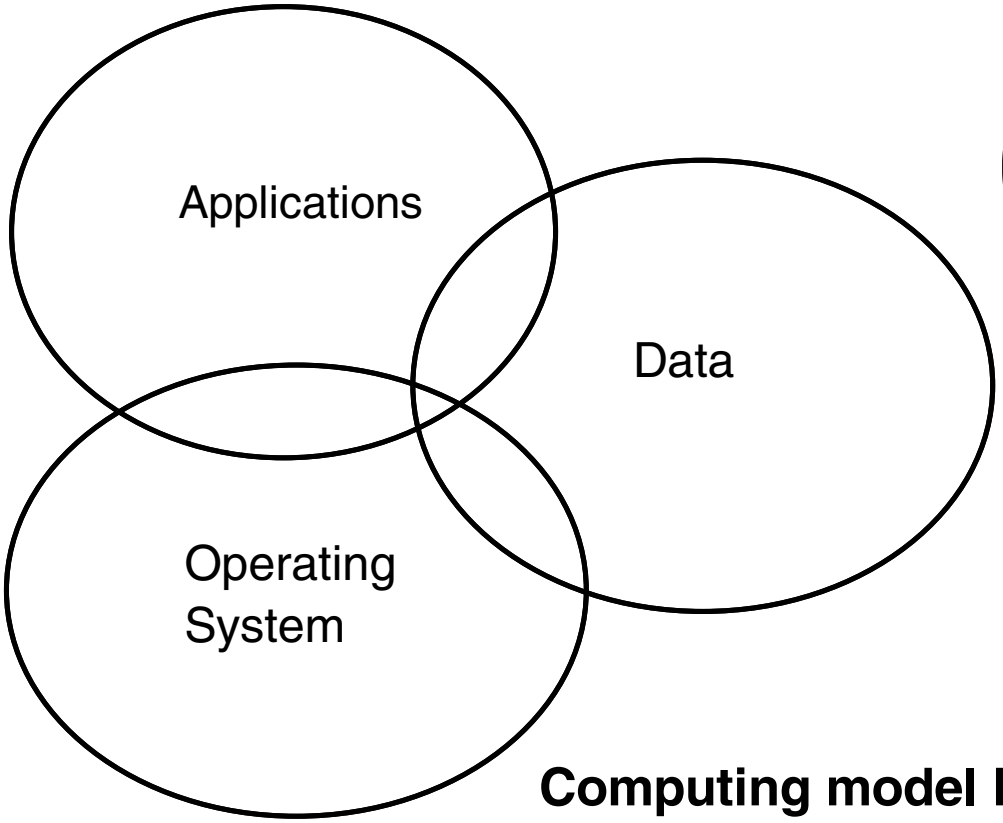
  - Decide on a file system layout:

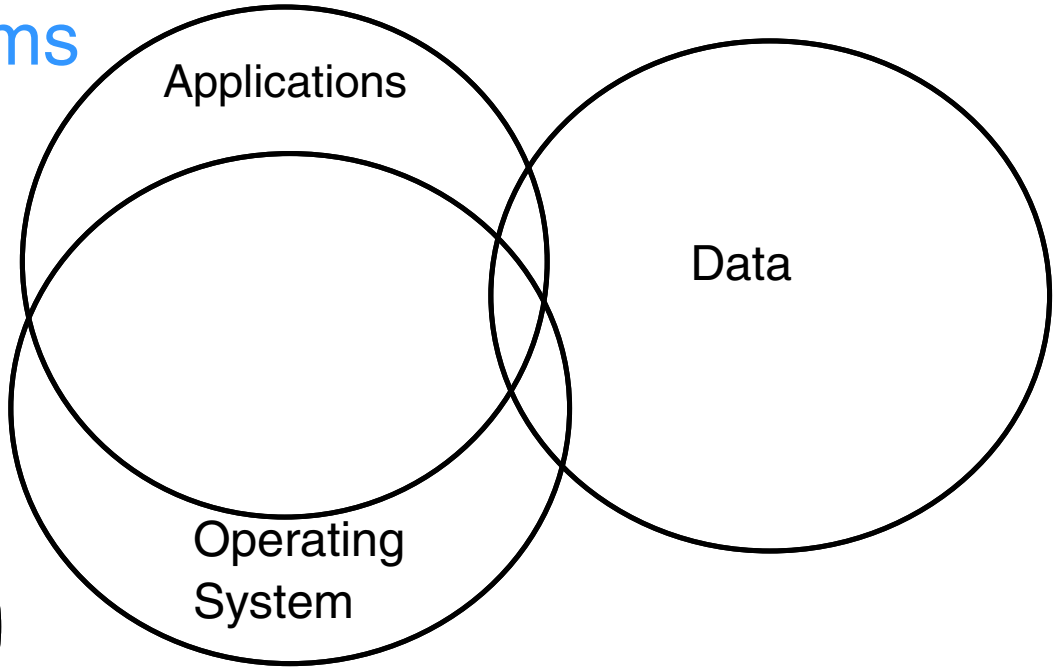
Mount point	Size	Volume group name	Logical volume name	Minidisk
/	512 MB	None	None	100
/tmp/	1 GB	system_vg	tmp_lv	100
/usr/	3 GB	system_vg	usr_lv	100
/var/	512 MB	system_vg	var_lv	100
/opt/	20 GB	opt_vg	opt_lv	101
swap	7 GB	None	None	302

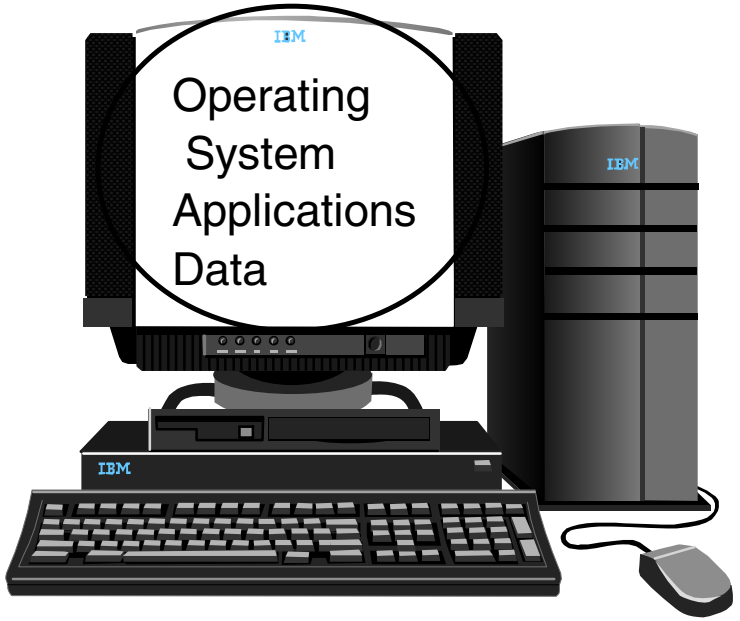
2 IBM Corporation

Tangent - thinking about systems

- Computing model today
  - How do your systems look?







Computing model hierarchy:

Hardware
- Operating System
  - Applications
    - Data

© 2012 IBM Corporation

FHS summary

Directory	Description
/	Root file system (must be able to boot/repair)
/bin	Essential commands (static)
/boot	Static files of the boot loader (static)
/dev	Device files (static, maintained by OS)
/etc	Host-specific system configuration (static)
/etc/opt	Add-on application configuration (site specific)
/home	User's home directories (site-specific, optional)
/lib	Essential shared libraries and kernel modules
/media	Mount point for removeable media (N/A on System z)
/mnt	Temporary mount point (usually empty)
/opt	Add-on application software packages (site-specific)
/root	Root user home directory (static, optional)
/sbin	Essential system binaries (static)
/srv	Data for services provided by this system (site specific)
/tmp	Temporary files (delete when system is booted?)
/usr	Sharable read-only data
/usr/bin	Most user commands
/usr/include	Header files included by C programs
/usr/lib	Libraries
/usr/local	Local hierarchy (empty after main installation)
/usr/sbin	Non-vital system binaries
/usr/share	Architecture-independent data
/var	Variable data
/var/cache	Application cache data
/var/lib	Variable state information
/var/local	Variable data for /usr/local
/var/lock	Lock files
/var/log	Log files and directories
/var/opt	Variable data for /opt
/var/run	Data relevant to running processes
/var/spool	Application spool data
/var/tmp	Temporary files preserved between system reboots

Data → /home

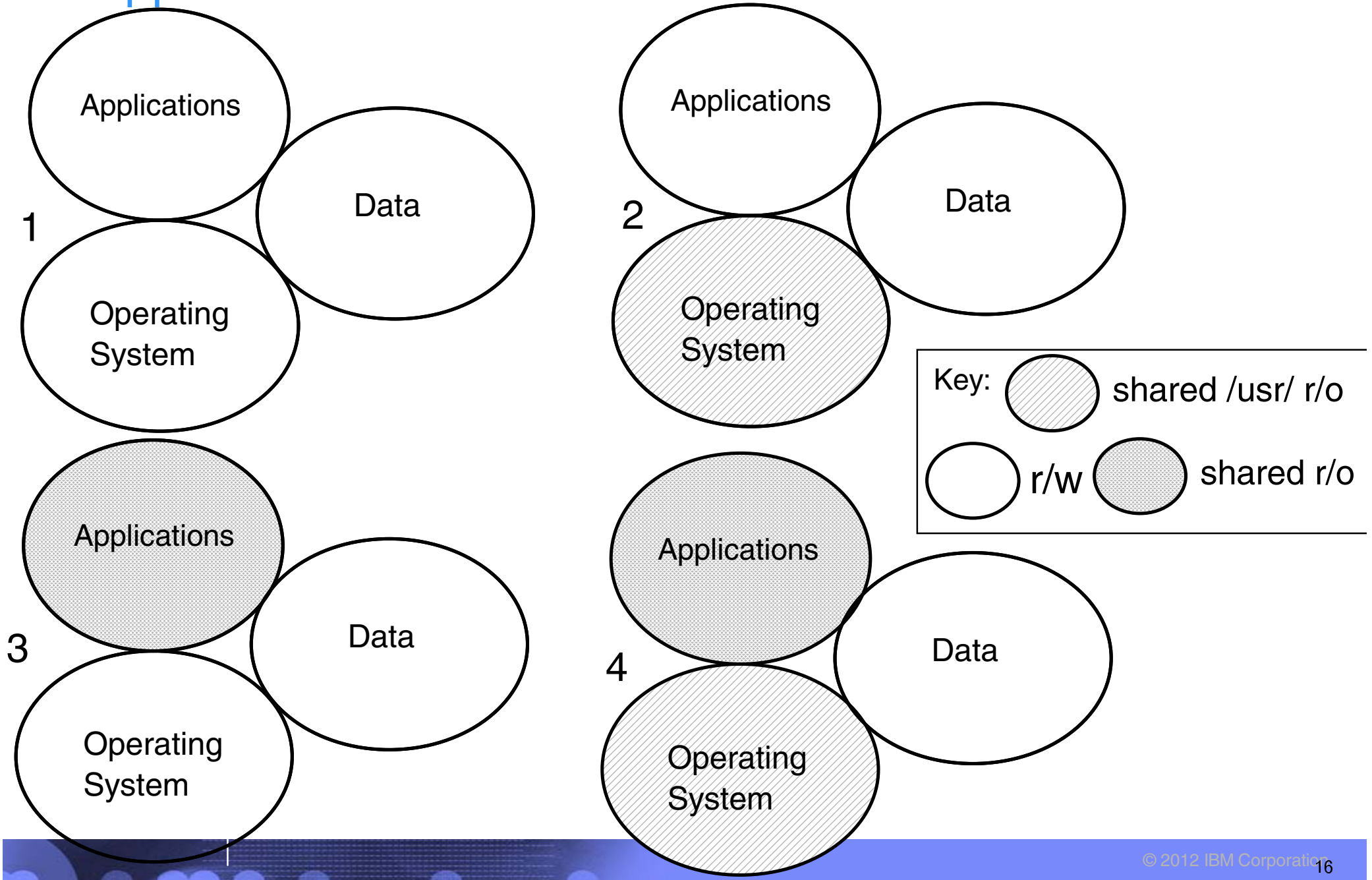
Apps → /opt

Data → /srv

Data or OS? → /var

© 2012 IBM Corporation

OS/Apps/Data view and virtualization - some models



© 2012 IBM Corporation

Platform as a Service (cont'd)

- Install Linux
- Configure Linux
  - Add a network interface to the private interconnect
  - Configure yum on the RHEL 6.2 golden image
  - Prepare for multipathing
  - Install and configure the VNC server
  - Copy files to Linux
  - Customize for Velocity software
  - Turn SE Linux off (??)
  - Customize rc.local
  - Shut down the RHEL 6.2 golden image
  - Add a network interface to the private interconnect

© 2012 IBM Corporation

Platform as a Service

- Cloning: Linux 'boot.firstone' service script used to set IP address and hostname
- REXX EXEC on z/VM for cloning

```
Parse Arg sourceID targetID .
If sourceID = '' | sourceID = '?' | targetID = '' Then Do
  say 'Syntax is:'
  say 'CLONE sourceID targetID'
  exit 1
End
/* verify that the source ID is logged off */
'CP QUERY' sourceID
If rc <> 45 Then Do
  Say sourceID 'does not exist or is not logged off?'
  exit 2
End
Say 'Are you sure you want to overwrite disks on' targetID '(y/n)?'
Parse upper pull answer .
If answer <> 'Y' then
  exit 3
/* FLASHCOY the 100, 101 and 302 disks from sourceID to targetID */
call copyDisk sourceID '100 1100' targetID '100 2100'
call copyDisk sourceID '101 1101' targetID '101 2101'
call copyDisk sourceID '302 1302' targetID '302 2302'
/* start the target virtual machine */
say "Starting new clone" targetID
'CP XAUTOLOG' targetID
exit
```

© 2012 IBM Corporation

Platform as a Service

- REXX EXEC on z/VM for cloning (cont'd)

```
copyDisk:
Arg sourceID vdev1 vdev2 targetID vdev3 vdev4 .
/* Link source disk read-only then target disk read-write */
'CP LINK' sourceID vdev1 vdev2 'RR'
If rc <> 0 Then Do
  say 'CP LINK' sourceID vdev1 vdev2 'RR failed with' rc
  exit 4
End
'CP LINK' targetID vdev3 vdev4 'MR'
If rc <> 0 Then Do
  say 'CP LINK' targetID vdev3 vdev4 'MR failed with' rc
  exit 5
End
Say 'Trying FLASHCOPY of' vdev2 'to' vdev4 '...'
'CP FLASHCOPY' vdev2 '0 END' vdev4 '0 END'
If (rc <> 0) Then Do /* Fallback to DDR */
  Say 'FLASHCOPY failed, falling back to DDR ...'
  Queue 'SYSPRINT CONS' /* Don't print to file */
  Queue 'PROMPTS OFF' /* Don't ask 'Are you sure?' */
  Queue 'IN' vdev2 '3390' /* Input minidisk */
  Queue 'OUT' vdev4 '3390' /* Output minidisk */
  Queue 'COPY ALL' /* Copy all contents */
  Queue ' ' /* Empty record ends DDR */
  'DDR'
  retVal = rc
End
```

© 2012 IBM Corporation

Platform as a Service

- REXX EXEC on z/VM for cloning (cont'd)

```
Else retVal = rc
/* Detach the source and target disks */
'CP DETACH' vdev2
'CP DETACH' vdev4
If retVal <> 0 Then
  Say 'Return value from COPYDISK' source target '=' retVal
```

© 2012 IBM Corporation

Agenda

- ~~Introductions~~
- ~~Cloud defined as I/P/SaaS~~
- ~~Infrastructure as a Service~~
  - ~~Configure z/VM~~
  - ~~Define virtual machines~~
- ~~Platform as a Service~~
  - ~~Prepare to install Linux on the golden image~~
  - ~~Install Linux on the golden image~~
  - ~~Configure the 6.2 golden image~~
  - ~~Cloning~~
- **Software as a Service for Oracle Standalone**
  - **Configure a Linux system for the Oracle boot script**
  - **Clone a virtual server**
  - **Silently install Oracle database**
- Software as a Service for Oracle Grid
- Miscellaneous

© 2012 IBM Corporation

Software as a Service for Oracle Standalone

- Configure a Linux system for the Oracle boot script
  - Copy **boot.oracle** to golden image **/etc/init.d/** directory.
  - Set the script to be executable with the **chmod +x** command
  - Set the script to start at boot time with the **chkconfig** command
  - Shut down the golden image
- Clone a virtual server

```
==> clone rh62gold lnxsa2
HCPCQU045E RH62GOLD not logged on
Are you sure you want to overwrite disks on lnxsa2 (y/n)?
y
Trying FLASHCOPY of 1100 to 2100 ...
Command complete: FLASHCOPY 1100 0 10015 TO 2100 0 10015
DASD 1100 DETACHED
DASD 2100 DETACHED
Trying FLASHCOPY of 1101 to 2101 ...
Command complete: FLASHCOPY 1101 0 30049 TO 2101 0 30049
DASD 1101 DETACHED
DASD 2101 DETACHED
Trying FLASHCOPY of 1302 to 2302 ...
Command complete: FLASHCOPY 1302 0 10015 TO 2302 0 10015
DASD 1302 DETACHED
DASD 2302 DETACHED
Starting new clone LNXSA2
```

© 2012 IBM Corporation

Software as a Service for Oracle Standalone

- Log on to the new Linux and watch boot

```
...
S01boot.onetime: this userID = LNXSA2
...
```

- Later, you should see:

```
...
S98boot.oracle: Searching for SOFTWARE variable in parameter file LNXSA2.CONF-RH6
S98boot.oracle: SOFTWARE variable = OracleStandalone
S98boot.oracle: Preparing for Oracle standalone installation
S98boot.oracle: Creating Oracle groups and users
S98boot.oracle: cmd: groupadd -g 198 oinstall
S98boot.oracle: cmd: groupadd -g 199 asmadmin
S98boot.oracle: cmd: groupadd -g 201 dba
S98boot.oracle: cmd: groupadd -g 205 asmdba
S98boot.oracle: cmd: groupadd -g 207 asmoper
...
S98boot.oracle: cmd: chown -R oracle.oinstall /opt/oracle
S98boot.oracle: *
S98boot.oracle: ************************************************
S98boot.oracle: Sucessfully completed!
S98boot.oracle: ************************************************
```

© 2012 IBM Corporation

Software as a Service

- One more file system for data: /oradata
  - Snippet of code from **boot.oracle** script:

```
dataName="oradata"                          # Oracle data mount point, vg name,
...
# for Oracle standalone, make a logical volume of the LUNs then mount it
  if [ "$type" = "ora" ]; then # make LV and mount it
    mkLogicalVolume /dev/mapper/mpatha /dev/mapper/mpathb
    mountLogicalVolume /dev/${dataName}_vg/${dataName}_lv /$dataName
  else # voting disks and data FCP LUNs will be controlled by ASM
    setDiskOwnership
  fi
```

© 2012 IBM Corporation

Software as a Service for Oracle Standalone

- At end of run level, **boot.oracle** should run:
  - Defines users and groups for Oracle
  - Installs co-requisite RPMs
  - Configures the Network Time Protocol (NTP)
  - Sets limits for the system, then the oracle and grid users
  - Sets kernel parameters
  - Configures FCP disks
  - Creates a logical volume from the two FCP disks
  - Makes a directory is made for Oracle data
  - Mounts the logical volume is mounted over the new directory
- Required variables:

```
===> x lnxsa2 conf-rh6 d
...
FCP400WWPN=0x500507630500c74c
FCP500WWPN=0x500507630508c74c
FCPLUN1=0x4010401200000000
FCPLUN2=0x4011401200000000
SOFTWARE=OracleStandalone
```

© 2012 IBM Corporation

Software as a service for Oracle Standalone

- Silently install Oracle database
  - Prepare the response file ("xxxx" values replaced by variables)

```
ORACLE_HOSTNAME=xxxx
UNIX_GROUP_NAME=oinstall
INVENTORY_LOCATION=/opt/oraInventory
SELECTED_LANGUAGES=en
ORACLE_HOME=/opt/oracle/11.2
ORACLE_BASE=/opt/oracle
...
oracle.all.db.DBA_GROUP=dba
oracle.all.db.OPER_GROUP=dba
oracle.all.db.isRACOneall=false
oracle.all.db.config.starterdb.type=GENERAL_PURPOSE
...
oracle.all.db.config.starterdb.password.ALL=xxxx
oracle.all.db.config.starterdb.control=DB_CONTROL
oracle.all.db.config.starterdb.automatedBackup.enable=false
oracle.all.db.config.starterdb.storageType=FILE_SYSTEM_STORAGE
oracle.all.db.config.starterdb.fileSystemStorage.dataLocation=/oradata
oracle.all.db.config.asm.ASMSNMPPassword=xxxx
SECURITY_UPDATES_VIA_MYORACLESUPPORT=false
DECLINE_SECURITY_UPDATES=true
oracle.installer.autoupdates.option=SKIP_UPDATES
```

© 2012 IBM Corporation

Software as a service for Oracle Standalone

- Silently install Oracle database
  - Run the silent installer:

```
# mount 9.12.5.131:/zCode /mnt -o vers=4
# su - oracle
$ cd /mnt/database
$ ./runInstaller -silent -force -ignorePrereq -responseFile ~/database.rsp
Starting Oracle Universal installer...
Checking Temp space: must be greater than 80 MB. Actual 923 MB Passed
Checking swap space: must be greater than 150 MB. Actual 7803 MB Passed
Preparing to launch Oracle Universal installer from
/tmp/Oraall2012-11-10_06-07-16AM. Please wait ...$ You can find the log of this
all session at:
/opt/oraInventory/logs/installactions2012-11-10_06-07-16AM.log
```

© 2012 IBM Corporation

Agenda

- ~~Introductions~~
- ~~Cloud defined as I/P/SaaS~~
- ~~Infrastructure as a Service~~
  - ~~Configure z/VM~~
  - ~~Define virtual machines~~
- ~~Platform as a Service~~
  - ~~Prepare to install Linux on the golden image~~
  - ~~Install Linux on the golden image~~
  - ~~Configure the 6.2 golden image~~
  - ~~Cloning~~
- ~~Software as a Service for Oracle Standalone~~
  - ~~Configure a Linux system for the Oracle boot script~~
  - ~~Clone a virtual server~~
  - ~~Silently install Oracle database~~
- **Software as a Service for Oracle Grid**
- Miscellaneous

© 2012 IBM Corporation

Software as a service for Oracle Grid

- Did not make it into Redbook :(
- Extra steps (if it had)
  - Decide on architecture
  - Clone multiple nodes
  - Set up key-based authentication between all nodes
    - **grid** and **oracle** users must be able SSH without passwords.
    - A helper script named **setsshkeys** will be available
  - Verify nodes are prepared
  - Silently install Oracle grid on one system
  - Silently install Oracle database on all systems
  - Test the cluster
    - HA
    - DR
    - Document the failover/failback steps

© 2012 IBM Corporation

Miscellaneous

- The Redbook
  - Title: *Experiences with Oracle 11gR2 on Linux for System z*
  - Order number: **SG24-8104**
  - From project leader:

  "ITSO legal is working on getting 'yes' from Oracle, Novell and RedHat, I think that we wil have it in 2 weeks. Before that I'm not allowed to publish the draft even internally."
- Additional material - one tar file:

```
# tar xzvf SG248104.tgz
oracleRedbook-SG248104/
oracleRedbook-SG248104/linux/
oracleRedbook-SG248104/linux/boot.oracle
oracleRedbook-SG248104/linux/boot.onetime
oracleRedbook-SG248104/vm/
oracleRedbook-SG248104/vm/CLONE.EXEC
oracleRedbook-SG248104/README.txt
```

© 2012 IBM Corporation

Resources

- All *Virtualization Cookbooks* and other papers:
  - http://www.vm.ibm.com/devpages/mikemac/
- *The* Linux for zSeries and S/390 portal
  - http://linuxvm.org/
- The linux-390 list server
  - http://www2.marist.edu/htbin/wlvindex?linux-390
- The IBMVM list server
  - http://www.lsoft.com/scripts/wl.exe?SL1=IBMVM&H=LISTSERV.UARK.EDU
- Linux for zSeries and S/390 developerWorks®
  - http://awlinux1.alphaworks.ibm.com/developerworks/linux390/index.shtml
- Red Hat Enterprise Linux evaluation
  - http://www.redhat.com/rhel/server/mainframe/
- SUSE LINUX Enterprise Server evaluation
  - http://www.novell.com/products/linuxenterpriseserver/eval.html
- z/VM publications
  - http://www.vm.ibm.com/pubs/
- z/VM performance tips
  - http://www.vm.ibm.com/perf/tips/

© 2012 IBM Corporation

Questions

- Are there any questions?

© 2012 IBM Corporation