



Linux on VM

From Woe to Whoa!

Session 9213

Rick Barlow

Richard.Barlow@nationwide.com

August 24, 2005

SHARE 105



Overview and Disclaimer



Disclaimer:

The content of this presentation is for information only and is not intended to be an endorsement by Nationwide Insurance. Each site is responsible for their own use of the concepts and examples presented.

Overview:

With a few exceptions, this is an overview! Where possible there are technical details you may be able to use. As you frequently hear, when anyone asks for recommendations, "IT DEPENDS"! The information in this session is based on my experiences as a long-time VM-er adding virtual Linux. Interactive is good! Please ask questions. We'll all get the most out of this session that way.

Abstract:

You've probably heard that Linux is Linux - even on IBM eServers. That is true. However, there can be a lot of politics around choosing a distribution and methodology for installing and maintaining your virtual server farm. The major distributors provide working instructions for installing on z/VM. People who are familiar with UNIX and Linux on i386 may want to do it the same way they always have. Everyone wants to use the tool set that they understand best and with which they are most comfortable. Topics to be discussed include:

- What does it take to motivate server people to work with virtual servers?
- Setting up a server to provide the distribution media to your VM system
- Using kickstart to load a virtual server on VM
- Providing network access to your virtual server farm

Topics



- Our Environment
- Our Linux History
- What do we expect Linux to do for us
- Direction – What drives our project
- A learning process
- Conclusions

Our Environment

Environment



- Changing as we speak...
 - Before we got serious about Linux
 - § 3 z900 processors; mostly z/OS; models 104, 107, 1C8
The 104 has 24GB of storage and runs:
 - o 4 z/OS LPARs
 - o 1 Coupling Facility LPAR (ICF)
 - o 2 relatively small z/VM LPARs which use about 200 MIPs
 - » LPAR 1: 3 shared CP; 3 GB Central storage; 1 GB Expanded storage
 - » LPAR 2: 2 shared CP; 768 MB Central storage; 256 MB Expanded storage
 - » Mostly web services
 - » Some application development, support and cooperation with z/OS
 - » Business Recovery support

Environment



- For Linux pilot
 - There was capacity to create another small z/VM LPAR on the 104.
 - § 3 dedicated IFL engines
 - § 8 GB memory: 6 GB Central; 2 GB Expanded

- Moving forward – 2 new z990, each with:
 - § 3 IFL engines
 - § 40GB memory
 - § 4 LPARs (1 additional test LPAR on development box)

Our Linux Experience

Linux History



- 2000 – Marist Distribution (based on Red Hat)
 - First offering of install lab
- 2002 – SUSE 7
 - Basic demo of Apache and Samba
 - Little interest or direction
- 2004 – Red Hat
 - Intel and zSeries pilots planned
 - zSeries waned quickly
- 2005
 - New Emphasis on virtualization

What Do We Expect Linux to Do For Us?

Problems



- **Server Proliferation**
 - Space that previously was required to house a few mainframes is now mostly consumed by multitudes of all type of servers, network hardware, other support hardware
 - § Sun, HP, multiple brands of Intel
 - § Routers and switches
 - § SAN, NAS, data warehouse, etc

Problems



- Provisioning
 - Many requirements for stand-alone server
 - § Order and obtain hardware – several weeks
 - § Physical install
 - § Optional external disk subsystem configuration and connection
 - § Network configuration and connection
 - § OS load
 - § Middle-ware load
 - § Application load

Many hands and significant time; likely to be several weeks or more

Vision and Expectations



- **Physical space and environmental reduction**
 - One z990 IFL engine can support 10-30 (or more) virtual servers
 - A z990 can have up to 32 IFL engines so it could replace 300+ servers.
 - Significant savings in physical space, power, cooling
- **Reduce network complexity**
 - A small number of physical network connections (OSAs with VSWITCH) can support all of the virtual servers in contrast to every stand-alone server having 2 or more interfaces it must manage.
- **Quicker provisioning**
 - Setting up new server can be as fast as your disk copy tool
 - § Depends on software needed on server and amount of manual effort

Direction – What Drives Linux And Virtualization on zSeries

Direction



- Start with technicians then try to influence organization
 - Common with existing zSeries shops; especially those with z/VM ("skunk works")
 - § Build something and demonstrate function and don't bother to tell anyone what it is and where it is running.
 - Challenges
 - § Organization barriers
 - § Workload
 - § "Opinions"
- Start with CIO (upper management) and direct organization
 - More common as industry accepts zSeries virtualization solution
 - Driven by business need (e.g. space restraints, rapid growth, etc)
 - Typically causes more structured implementation

A Learning Process

Learning



- Mainframe methodology differs from non-mainframe methodology
 - Repeatability automated processes versus hands-on "hacker" install
 - § Typical mainframe person accustomed to well-documented, repeatable processes that permit automation for multiple installs.
 - § Preference for install once and copy rather than repeated installation.
 - § Expect software to be installed in one location and configuration in a common location.
 - Different philosophy for management:
 - § Privilege levels differ for installing OS, installing middle-ware, configuring
 - Different maintenance philosophy
 - § Expect to regularly upgrade software on all servers

Learning



- Documentation differs greatly

- Red Hat

- § Installation instructions begin at loading the RAM disk into memory. It appears to have been an afterthought.

- SUSE

- § Shows how to build the virtual server directory and copy the RAM disks to VM. It seems to understand the zSeries and z/VM environment.

- Both

- § Incomplete (inadequate) documentation of install parm information for all environments

- o Some not documented
 - o Little more than syntax
 - o Incomplete or no examples

Learning - Distributions



- Media
 - CDROM
 - § Can't load directly on z/VM
 - § Make distribution media available on existing Linux (or Unix) workstation
 - § Make RAM disk images available to FTP to VM for install
 - o Directly from CDROM
 - o Extract from mounted ISO file

Learning - Distributions



- Default package list

- Red Hat

- § Large list of packages in minimum load

- § Security template required omitting load, turning off or disabling many packages

- § Runtime compatibility for 31-bit not included in default 64-bit load

- SUSE

- § Smaller list of packages in minimum load – basic runnable Linux

- § Security template required only a handful of changes

Learning - Distributions



- Red Hat Enterprise Linux AS

- When installing Red Hat Enterprise Linux AS 3 64-bit, default RAM disk size was too small to build a complex DASD / LVM configuration.

Resolving this took several days and knowledgeable Linux experts to identify.

(Red Hat says this is fixed in AS 4)

- Working bugs out of kickstart was a time-consuming repetitive process.
 - § We started with a working kickstart script from Intel.
 - § It was difficult to identify packages that are not on the s390 and s390x CDROMs.
- Install (using kickstart) formatted DASD one-at-a-time (serially)
(This may be a restriction of the kickstart process.)

Learning - Distributions



- Red Hat

- **BOOT PARM example for RHEL AS 3 Update 3**

```
root=/dev/ram0 ro ip=off ramdisk_size=80000 DASD=0151-015F
HOSTNAME=linr2003.nwie.net MTU=1500 NETTYPE=eth
IPADDR=192.168.127.251
NETWORK=192.168.127.128 NETMASK=255.255.255.128
BROADCAST=192.168.127.255 GATEWAY=192.168.127.254
CHANDEV=qeth0,0x5708,0x5709,0x570A,0,0
SEARCHDNS=nwie.net:ent.nwie.net:cdc.ent.nwie.net:edc.ent.nwie.net
DNS=172.24.101.100:172.24.104.100:172.25.114.200
SWAP=/dev/dasdb1 INSTALL=default DTZ=EDT
QETHPARAM=add_parms,0x10,0x5708,0x570A,portname:Merlot
RUNKS=1 ks=http://192.168.122.226/kickstart/linr2003.ks
```

Learning - Distributions



- Red Hat

- Kickstart example for RHEL AS 3 Update 3 64-bit (base stuff)

```
#Install Red Hat Linux instead of upgrade
install
url --url http://172.25.103.235/linux/frame/as3_u4_64
text
lang en_US.UTF-8
langsupport --default en_US.UTF-8 en_US.UTF-8
keyboard us
mouse none
network --bootproto=static --device=hsi0 --hostname=nndcas821
firewall -disabled
authconfig --enablesshadow --enablemd5 -enablecache
timezone America/New_York
bootloader --location=mbr
zerombr yes
reboot
cmdline
```

Learning - Distributions



- Red Hat

- Kickstart example for RHEL AS 3 Update 3 64-bit (partitioning stuff)

```
clearpart --initlabel
clearpart --all
part /boot --size=1 --grow --ondisk=dasda
part swap --size=1 --grow --ondisk=dasdb
part pv.01 --size=1 --grow --ondisk=dasdc
part pv.02 --size=1 --grow --ondisk=dasdd
part pv.03 --size=1 --grow --ondisk=dasde
part pv.04 --size=1 --grow --ondisk=dasdf
part pv.05 --size=1 --grow --ondisk=dasdg
part pv.06 --size=1 --grow --ondisk=dasdh
#
volgroup rootvg pv.01 pv.02 pv.03 pv.04
volgroup datavg pv.05 pv.06
logvol / --vgname=rootvg --size=256 --name=rootlv
logvol /usr --vgname=rootvg --size=1024 --name=usrlv
logvol /opt --vgname=rootvg --size=256 --name=optlv
logvol /tmp --vgname=rootvg --size=512 --name=tmplv
logvol /var --vgname=rootvg --size=1024 --name=varlv
logvol /home --vgname=rootvg --size=512 --name=homelv
logvol /u01 --vgname=rootvg --size=2048 --name=u01lv
logvol /webdata --vgname=datavg --size=4096 --name=wdlv
```

Learning - Distributions



- Red Hat
 - Kickstart example for RHEL AS 3 Update 3 64-bit (package stuff)

```
%packages
@ compat-arch-development
-compat-gcc
-compat-gcc-c++
-compat-libstdc++-devel
... packages to exclude ...
kernel
net-snmp
... packages to add ...
nss_ldap
```


Learning - Distributions



- Red Hat

- Kickstart example for RHEL AS 3 Update 3 64-bit (post-install script)

```
%post
```

```
/sbin/chkconfig gpm off
/sbin/chkconfig cups off/sbin/chkconfig auth off
rpm -i http://172.25.103.235/linux/frame/as3_u4_64/RedHat/RPMS/libstdc++-
.2.3-47.s390.rpm
cd /etc/rc.d/rc3.d
/usr/bin/wget http://172.25.103.235/linux/S98postinstall4
chmod 755 /etc/rc.d/rc3.d/S98postinstall4
echo "# Do not remove the following line, or various programs" > /etc/hosts
echo "# that require network functionality will fail." >> /etc/hosts
echo "127.0.0.1          localhost.localdomain localhost" >>
/etc/hosts
echo "192.168.127.162      nndcas821.nwie.net nndcas821" >> /etc/hosts
echo "alias hsi0 geth" >> /etc/modules.conf
rpm -e fontconfig-devel freetype-devel gdb glibc-devel glibc-headers glibc-
kernheaders libcap-devel libstdc++-devel ncurses-devel pam-devel XFree86-
devel zlib-devel

echo "echo \"0\" > /proc/sys/kernel/hz_timer" >> /etc/rc.d/rc.local
rpm -i http://192.168.122.226/TSM/TIVsm-API.s390.rpm
rpm -i http://192.168.122.226/TSM/TIVsm-BA.s390.rpm
```

Learning - Distributions



- SUSE Linux Enterprise Server
 - Install processes formatted multiple DASD in parallel
 - Never completely got AutoYaST to work
 - § Realized that cloning / copying servers make this less important

Learning - Distributions



- SUSE Linux Enterprise Server

- **BOOT PARM example for SLES 9 64-bit**

```
ramdisk_size=65536 root=/dev/ram1 ro init=/linuxrc TERM=dumb
IP_INTERFACE=hsi READ_DEVNO=5708 WRITE_DEVNO=5709 DATA_DEVNO=570a
AUTOYAST=nfs://172.25.103.235/export/autoinstall/nndcas809.xml
IP_HOST=nndcas809.nwie.net IP_ADDR=192.168.127.147 INST_INFO=nfs
IP_NETMASK=255.255.255.128 IP_BROADCAST=192.168.127.255
IP_GATEWAY=192.168.127.254 IP_DNS=172.24.101.100 IP_SEARCH=nwie.net
IP_MTU=1500 INST_IP_ADDR=172.25.122.90 INST_PASSWORD=suse4u
INST_SCREEN=VNC DASD_Parameter=dasd=0151-0158
INSTALL=nfs://172.25.122.90/export/ INST_IP_DIR=/export
```

Linux Basics



- Even z/VM Sysprogs need to understand what Linux is up to.
 - What we know about running VM applications has an impact on how Linux is built on VM and knowing Linux to some degree helps get the points across to the Linux admins.
 - Apply some mainframe disciplines / history / concepts to virtual Linux
- Learn how new devices are added, defined and identified in Linux for the distribution you are using.
- Learn LVM
 - Differs depending on distribution (kernel level and/or LVM level)
 - Using CKD DASD for virtual servers is likely to cause you to use it.

Linux Basics



- Take careful notes about what you learn so you can use them later
 - Pick up on Linux tips/tricks to make your life easier (refer to note 1).
- § Something as simple as adding a "&" to the end of a dasdfmt command lets you run things in the background and not have to wait for them to complete to do something else.
 - § RPM - learn how to search for packages, display information on them and how to install/uninstall them.
 - § Learn the file attributes and what they mean, along with the decimal representations of them. For chmod commands, you need to understand them.
 - § know how to use tar and gzip
 - § Keep cheat sheets on VI (VIM), ED, and one on Linux System Admin (www.cactus.org/~dak/sysadmin.html)
 - § Symbolic Links are useful to know
 - § killproc is your friend. Know how and when to use it.
 - § Another tip - you can create environment variables for common locations you "cd" to often. For instance, if you always cd to "/opt/var/html/web/tsm/www/logs/public" you can create an environment variable called \$TLOG and then cd \$TLOG to get there. Using the "tab" trick is okay, but it still takes longer than typing "cd \$xxx..."

Linux Basics



- Cloning:
 - There are numerous ways to clone Linux images. PICK ONE and stick to it.
 - If you have hardware disk duplication available (IBM Flashcopy, STK Snapshot), it can be a huge benefit to cloning. If not, DDR can be used (slow).
 - You may want to create "standby" Linux images for quick deployment.

Linux workloads



- Linux on zSeries virtual servers *may* be able to run with small memory (storage) sizes.
 - "It depends" on what will run there.
 - A basic Linux virtual server can easily run on 64MB of memory.
 - An IBM HTTP Server can probably run in 256-512MB depending on the number of static pages versus CGIs, etc
 - An IBM WAS Server probably needs 512MB-1GB
 - An IBM WAS Portal server probably needs 2G or more.
 - There are a lot of simpler application options
 - § Firewall
 - § DNS
 - § Web server
 - § File and Print serving

Learning



- zSeries Hardware
 - CPU
 - § Effectual speed much higher than raw speed
 - Storage (aka memory)
 - § Maximum on a 32-engine z990 is 256 Gigabytes
 - § Fixed allocation to each z/VM LPAR
 - § z/VM dynamic allocates to virtual servers
 - DASD (aka disk)
 - § Count Key Data (CKD) – traditional zSeries
 - § SAN – used on many stand-alone servers

Learning



- zSeries Hardware
 - Open System Adapter (OSA) Express 2 Gigabit Ethernet
 - § Gigabit adapter with a smart network controller
 - § zSeries LPAR microcode allows:
 - o Sharing of the same OSA across LPARs
 - o Multiple Read/Write/Data groups to be attached to virtual server or defined as a VSWITCH

Learning



- zSeries Hardware

- Virtual Switch (VSWITCH)

- § Combination of zSeries microcode and z/VM CP code to create an extension of a network switch

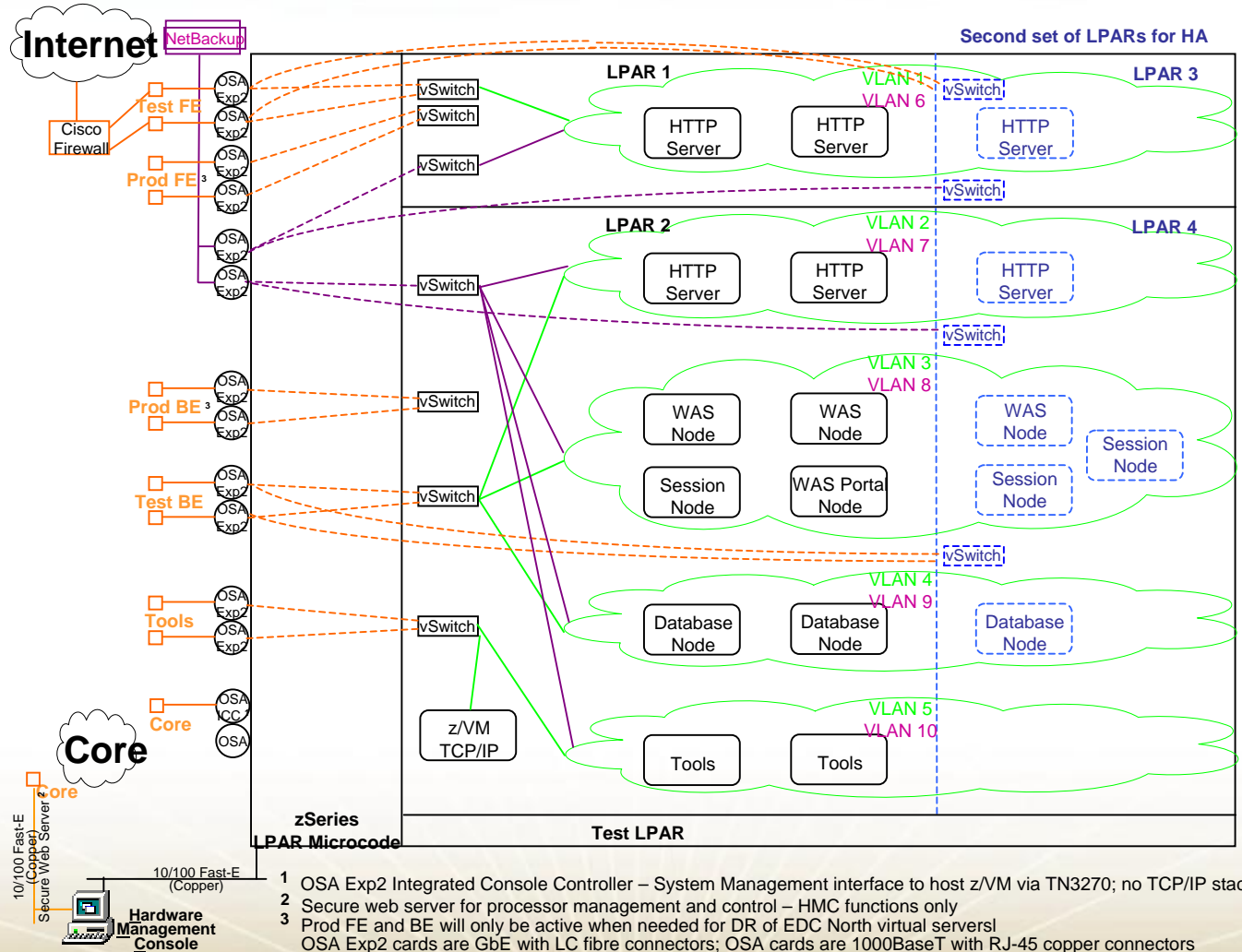
- § Layer 3

- o Defined as "IP"
 - o Sufficient for most implementations
 - o Common MAC included for all guests

- § Layer 2

- o New on z990 with OSA Express 2
 - o Defined as "Ethernet" – allows non-IP protocols like NETBIOS or IPX
 - o Recommended by IBM
 - o Unique MAC for each virtual server
 - » Local MAC addressing must be administered
 - o z/VM TCPIP cannot connect to a Layer 2 VSWITCH

Network



VSWITCH Detail



- Our OSA / VSWITCH configuration
 - 6 OSA Express 2 Gigabit Ethernet cards (12 Gb ports)
 - 1 OSA Express 1000BaseT (2 ports: 1 ICC)
 - 6 different network zones; 12 VSWITCHes defined
 - § 2 VSWITCHes on each pair of OSA ports for redundancy and load distribution
 - o Paired OSA ports are on separate cards for redundancy
 - § Each pair of ports is in a specific network zone
 - o Each OSA port in a pair is connected to a different physical switch

VSWITCH Detail

- Defining VSWITCH

- In SYSTEM CONFIG or via CP command by authorized user (same syntax in both places)

§ Example of a pair of VSWITCHes:

```
CP DEFINE VSWITCH NWZONE1A RDEV C100 C204 CONTROLLER * IP VLAN 4094
CP DEFINE VSWITCH NWZONE1B RDEV C200 C104 CONTROLLER * IP VLAN 4094
```

VLAN on the VSWITCH is the default VLAN used by the hardware switches.

- Authorizing virtual servers to use VSWITCH

- SYSTEM CONFIG format

§ Example of 2 virtual servers in same zone on opposite VSWITCHes

```
MODIFY VSWITCH NWZONE1A GRANT LINSERV1 VLAN 1001
MODIFY VSWITCH NWZONE2A GRANT LINSERV2 VLAN 1001
```

- CP command format

§ Example of 1 virtual server on 2 VSWITCHes in different zones

```
CP SET VSWITCH NWZONE1A GRANT LINSERV1 VLAN 1001
CP SET VSWITCH NWZONE2B GRANT LINSERV1 VLAN 2001
```

"Experts"?



- "Experts": Do they really exist?
 - There are many people with varying levels of experience in specific areas.
 - There are few (if any) who know enough about everything.
 - Make friends with people with expertise in:
 - § Mainframe disciplines
 - § Linux
 - § Network
 - Learn as much as possible about all of these

Conclusions

Conclusions



- Linux definitely is Linux – the same on all platforms
 - Code written for Linux on any platform can usually be used on any other platform that supports Linux with only a recompile.
- Linux virtualization on zSeries can:
 - Reduce cost
 - Reduce complexity
 - Accelerate provisioning
 - Reduce human error of manual installation and configuration

Conclusions



- Not every workload is suited to Linux on zSeries
- Not all software is ready for Linux on zSeries
- Things are changing rapidly

- Be careful what you ask for because you may get it!

References



- Other sessions this week:

- Mon 11:00 9125 Virtual Networking with z/VM Guest LANs and the z/VM Virtual Switch
- Mon 04:30 9112 z/VM TCP/IP Stack Configuration
- Tue 09:30 9131 TCP/IP Routing
- Thu 01:30 9132 MPRoute Configuration for z/VM
- Thu 01:30 9216 From LPAR to Virtual Servers in Two Days: Day 1 - z/VM
- Thu 03:00 9217 From LPAR to Virtual Servers in Two Days: Day 2 - z/VM

- Documentation

- REDP3719 **Linux on IBM eServer zSeries and S/390:
VSWITCH and VLAN Features of z/VM 4.4**

Contact Information



Rick Barlow
Systems Engineering Consultant

Phone: (614) 249-5213

Internet: Richard.Barlow@nationwide.com