Introducing Linux to the Enterprise (Parts 1 and 2)

Joe Oak
Architect – State Auto Insurance
Columbus, Ohio

Joe.oak@stateauto.com
Introduction

Me –

Joe Oak – **Software Architect**

joe.oak@stateauto.com

Diver

Sheep & Lama Farmer
Introduction

- Overall, we’ll be talking about...
  - The Politics of introducing new technologies to a company.
  - Experiences while introducing Linux to the environment
  - Why Linux & Open Source
  - Virtualization
  - The zSeries Mainframe & z/VM
  - Building your first Linux system
  - Quickly deploying numerous Linux systems
  - Virtualization with Workstations and Servers
  - Experiences with introducing Open Source to your company
    - Also see: David Moskowitz – Managing Open Source
    - Donald Smith – Open Source presentations

- Section 1 – will cover mostly mainframe z/VM virtualization with Linux

- Section 2 – will explore virtualization on other platforms & Open Source experiences
Introduction

- Overview of Experiences
- Introduction to the z/Series (mainframe)
  - Overview & History
  - Z/VM the Operating System
    - Overview
    - Components
    - How it works to the scheme of things
  - Linux on the z/Series
    - Overview
    - Installations & Cloning
    - Tuning & Security
  - Summary of Part 1
- Virtualization with Servers & Workstations
  - Overview of virtualization on the non-mainframe platforms
  - Review a couple of popular products
  - Introducing Open Source products in your company
  - Summary of Part 2

About 10% of the Presentation

About 40% of the Presentation

About 50% of the Presentation
Introduction

What I hope you’ll get from this presentation...

- An updated understanding of z/VM & the zSeries complex
- General concepts of using Linux with the z/VM system
- Where virtualization was... and where it is today
- Thoughts on introducing new technologies to a corporation
- General experiences with Linux, z/VM & Open Source
I’ll be talking about several hardware and software products that are available as Open Source or are commercially produced.

Neither I nor my employer is endorsing or discounting the use of any company, product, or features presented.

Since there are so many products available, I cannot possibly mention everyone’s favorite. The presentation includes software I have encountered while working. Please do not take offense if I don’t mention one that you are fond of. Thanks!
Time Line – In One Year...

- CSS 2005
- Investigate – what’s Linux good for
- Play with Desktop Linux
- Mainframe –
  - Install z/VM
  - zLinux
  - Red Hat
  - SUSE 9 & 10 Enterprise
    - Tomcat 5.5 & Java 1.5
- Selling it... to Developers & Management
  - Help from our friends (Local User Groups & IBM)
- Our first app – small, stable
Time Line – In One Year...

- Let’s go Production
- Outages – Power & DASD
- Now everyone wants a Linux
- MySQL & Mono
- SourceSafe to... subVersion
- Groups discuss policies & procedures
- New Mainframe
- Networking fun
- SUSE 10
- WebSphere 6.0 scaling from Server to Mainframe
- Today – CSS 2006
A Brief History...

- Medium size insurance company
  - Two data centers
  - Numerous remote sites
  - A good number of customers & employees connecting in

- **Lots** of legacy code
  - COBOL
  - VB 5 & 6

- 225+ Intel servers
  - Windows 2003
  - A couple of Linux servers provided by vendor applications
A Brief History…

- A few thousand Intel Workstations
  - All within 1 month to 3 years old
  - Windows XP Professional
  - OS/2 Warp 4 – with lots of updates
    - Will be gone by Oct 1st

- VMWare – Intel
  - All running Windows 2003
  - Initial attempt at consolidating servers
  - You still pay for the software!
  - Not too good for Exchange/Mail or SQL
A Brief History…

- zSeries with MVS
  - COBOL
  - DB/2
  - CICS
  - Hardware
    - z/890 - 2086
    - 2 z/OS CP's
    - 1 IFL
    - 16 GIG memory
      - 13GB Production z/OS
      - 2GB z/VM Central
      - 1GB z/VM Expanded
A Brief History...

- At my company... growth every which-way!
  - Websphere Applications
    - 5.1 & 6.0 on Intel
    - Spanning two data centers
  - .Net Applications
    - Departmental and Corporate
  - Multi-tier architecture
    - Most application get their data/business function from the mainframe.
    - Some replicate from IBM DB/2™ (mainframe) to Microsoft SQL Server™
  - People are beginning to question software costs
  - Disaster Recovery & redundancy has high visibility
    - Setup for Testing & the Real thing takes time and infrastructure
  - Security has even higher visibility.
How Did We Start?

- **CSS 2005**
  - I received the .................................................................
    
    ... thanks Mike!

- At home, it was time to start looking into Linux
  - Downloaded a Desktop version and played
  - Started noting articles on Linux under z/VM
  - Help the Systems group build a VM System
  - Began to look into other Virtualization products

- As you proceed – remember: **Patience and perseverance have a magical effect before which difficulties disappear and obstacles vanish.**
  
  - John Quincy Adams
How Did We Start?

- Selling New Technology (Linux, Tools, etc.)

  - First ask yourself – Is there a place for it in your company?
    - Does anyone want it, other than to play with?
    - What is your motivation?
    - What will motivate others to help you?

  - Find the right people – developers, systems support, management
    - Start by talking one on one
    - **Develop relationships** – very hard to do with email or the phone
      - Have some Java with people in other departments 😊

  - Find a small *(very stable)* application for Linux *(Java is a very good candidate)*

  - Find in-roads for Open Source products – more in Part 2
Virtualization

What is it? What does it encompass?

➤ Dozens of definitions – depending on use and product.

In computing, virtualization is the process of presenting a logical grouping or subset of computing resources so that they can be accessed in ways that give benefits over the original configuration. This new virtual view of the resources is not restricted by the implementation, geographic location, or the physical configuration of the underlying resources.

... Virtualization is a broad term that refers to the abstraction of resources in many different aspects of computing. As a perceived "hot term", virtualization has been claimed by IT marketers to refer to everything from virtual machines to systems management software, so as to become nearly meaningless.


Virtualization

- In general, what are some of the virtualization areas...

- We’ll be looking at some different possibilities for performing virtualization in the Mainframe and Server environments.

- Most have the same concepts...
  - Some have different approaches
  - Some, if not all have their own terminology

- Different approaches have differing results
Virtualization – Why?

- Workload consolidation
- Support for old systems (OS/2 – other legacy)
- Support migration by providing multiple versions of an OS or application
- Great developer play ground that is safe for the developer... and others!
- Develop / Debug operating systems
- Can simulate hardware
- Can “pre-package” system – new term: **appliances**
- Backup and DR (Disaster Recovery) can be easier
- Available for workstations though mainframes
- Why we’re looking at it....
  - All the above
  - Space, electric, cooling, postpone upgrades, ...
  - Saves LOTS of money
### Who’s Offering Virtualization?

<table>
<thead>
<tr>
<th>Company</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>VMWare</td>
<td><a href="http://www.vmware.com">http://www.vmware.com</a></td>
</tr>
</tbody>
</table>
| XEN             | [http://www.cl.cam.ac.uk/Research/SRG/netos/xen](http://www.cl.cam.ac.uk/Research/SRG/netos/xen)  
|                 | [http://www.xensource.com](http://www.xensource.com) |
| Microsoft       | [http://www.microsoft.com](http://www.microsoft.com) |
| Parallels       | [http://www.parallels.com](http://www.parallels.com) |
| QEMU            | [http://fabrice.bellard.free.fr/qemu/](http://fabrice.bellard.free.fr/qemu/) |
| Plex86          | [http://plex86.sourceforge.net](http://plex86.sourceforge.net) |
| WINE            | [http://www.winehq.org](http://www.winehq.org)  |
zSeries – A Very Brief History

- A very brief history of the virtualized mainframe

Virtualization is not new! *Control Program*
- CP/40, CP/67, VM/370, ... z/VM
  - Originally for OS Testing & multi-user systems
  - Spans 1960’s through today

360 to zSeries
- 360, 370, 308x Complex, S/390, zSeries
- IBM added VM/Assist & SIE (Start Interpretive Execution)
- For more history...
  - http://www-03.ibm.com/servers/eserver/zseries/timeline
zSeries – The Hardware

- Special hardware assists
  - IFL vs. CP
    - What is the IFL? *(Integrated Facility for Linux)*
    - IFL – additional processing for Linux – only runs Linux and z/VM
    - Doesn’t count as MIPS against your MVS systems
    - Can be used with Native Linux LPARs or Linux under z/VM (v4 & v5)
  - zAPP – *z9 Application Assist Processor*
    - JVM off loads Java execution to zAAP (MVS only ?? z/OS 1.6)
  - Others... SAP, zIIP, and ICF

- zSeries extremely stable
  - ... but, if VM goes down so do your guests (*underlying systems*)!
  - Incredible I/O subsystem
  - Entry level cost is around $100,000 (z9 BC)
zSeries – The Hardware

- Partition by Hardware, Logically, & Software
  - Hardware – items are allocate directly – not shared
  - Logically – The complex manages sharing of resources across partitions
  - Software – resources are managed by software like z/VM

- PR/SM – logical partitioning of the processor complex...
  - Processor Resource / System Manager
  - First level of mainframe virtualization
  - Isolates each defined LPAR
    - Some provisions for crossing boundaries –
      - CTCA’s, Communication Links (OSA), HiperSockets (Firmware)
    - Loosely, like VMWare ESX on servers
    - # of LPARs depends on the mainframe (e.g. zSeries 900 can have 15 LPARS)
zSeries – The Hardware

- LPARs can define:
  - CP’s (standard, IFL, zAAP, etc.)
  - Storage
  - Extended Storage
  - Channel Paths - ESCON and FICON
  - I/O Devices – networking, DASD, Tape, etc.

- Why would I use Linux in an LPAR rather than under z/VM?
- Can I use both?
- How does this impact my production MVS system?
z/VM Today

- Where is VM today?
  - It’s still here!
    - Can run very old legacy systems or run itself
    - Supports old hardware (real & virtually)

- Components
  - CP & CMS
  - Guests: MVS, DOS/VSE, VM (itself), Linux, R-Y-O
  - IPL – DASD, Tape, & even a Card Reader ..........
  - Simulates all type of equipment or can use Real devices
  - 32/64 bit
  - Dispatching and Scheduling
  - Real & Virtual Networking
  - 100’s of VM Guests
  - Incredible learning environment for systems programming
  - “Green Screens” – 3270’s – not fun with Linux!
z/VM Today

➢ The latest... Version 5.2
  • Enhancements for Linux
    ✓ New memory management for Linux (APAR VM63856)
  • Support for *Guests* access above 2Gb real storage
  • IPv6 HiperSocket support
  • With FICON Express4 – 4Gb/sec transfer rates
  • Easy upgrade from previous versions
  • The actual price of the VM operating system has dropped!
    ✓ Talk to your IBM Rep for exact details

➢ Bottom-line:

  IBM is putting a *lot* into VM to make it a premier Linux platform
z/VM Today & Yesterday

- VM/370 command cards from 1976 & 1979

The CP and CMS commands in these cards from 30 years ago...

**WORK TODAY**

... and have been enhanced for newer systems & features.
z/VM – Virtualization

- Remember back to our earlier slide on what may get virtualized...

  - How does this apply to z/VM?
  - z/VM can virtualize all or nothing
  - It’s your choice
    - Of course that depends on what you’re trying to do!
**z/VM – Characteristics**

- **DASD**
  - Direct Access Storage Device (DAZ-dee)
  - Usually allocated in Block, Tracks, or Cylinders
  - Under VM – we have minidisks
    - Minidisk = Mdisk
    - Can be an entire volume – Dedicate statement
      - Dedicate 205 123 – Dedicate real DASD 205 and my virtual 123
    - Can be a piece of a volume – MDISK statement
      - MDISK 191 3390 2712 001 MYVOL1 MW READ WRITE MULTI
    - Can have mix and match several types
z/VM – Characteristics

- For your Linux DASD use 4096 blocksize

- General DASD Rules for VM & Linux
  - Use as many CHPIDs/CUs as possible
  - Spread the host adapters used across all host adapter bays
  - Try to make sure interconnected devices have the same transfer rates/speed.
  - Spread the used disks equally over all ranks and both clusters
  - If possible, avoid reusing the same resources (CHPID, host adapter, rank)
z/VM Characteristics

- Fancy Features
  - Communications
    - IUCV
    - Ethernet (via OSA)
    - HiperSockets (virtual and real)
      - Data never travels outside the mainframe
      - Very secure connection – doesn’t need extra encryption
      - Applications just see a LAN
      - Designed as a feature of the zSeries processors
      - Communicates across LPARS and Operating Systems
        - on the same mainframe
      - Basically, a memory to memory transfer of data
      - z/VM will simulate HiperSockets between guest systems
z/VM – Characteristics

- **Packaging**
  - Bare Bones VM
  - System Delivery Offering (SDO)
    - Preconfigured operating system and 1- several dozen products

- **Automating Tasks**
  - REXX
z/VM – Characteristics

- Sharing Data
  - Data that is defined on minidisks can be shared in a Read/Write mode.
  - The guest must know how to lock and preserve data integrity.
  - Minidisks can be a “full volume” or partial.

- My customer asked for 2Gb...
  - How do I handle that request?

<table>
<thead>
<tr>
<th>Disk Type</th>
<th>Cylinders</th>
<th>Track/Cylinder</th>
<th>Bytes/Track</th>
<th>Bytes/Cylinder</th>
<th>Bytes/Physical Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>3380-E</td>
<td>1113</td>
<td>15</td>
<td>47,476</td>
<td>712,140</td>
<td>1,260,487,800</td>
</tr>
<tr>
<td>3380-K</td>
<td>2226</td>
<td>15</td>
<td>47,476</td>
<td>712,140</td>
<td>1,890,731,700</td>
</tr>
<tr>
<td>3390-1</td>
<td>1113</td>
<td>15</td>
<td>56,664</td>
<td>849,960</td>
<td>946,005,480</td>
</tr>
<tr>
<td>3390-2</td>
<td>2226</td>
<td>15</td>
<td>56,664</td>
<td>849,960</td>
<td>1,892,010,960</td>
</tr>
<tr>
<td>3390-3</td>
<td>3339</td>
<td>15</td>
<td>56,664</td>
<td>849,960</td>
<td>2,838,016,440</td>
</tr>
<tr>
<td>3390-9</td>
<td>10017</td>
<td>15</td>
<td>56,664</td>
<td>849,960</td>
<td>8,514,049,320</td>
</tr>
<tr>
<td>3390-27</td>
<td>32760</td>
<td>15</td>
<td>56,664</td>
<td>849,960</td>
<td>27,844,689,600</td>
</tr>
</tbody>
</table>
z/VM – Characteristics

- Minidisk or Dedicated
  - The minidisk is the standard allocation of DASD under VM.
    - Provides virtualization of Real DASD to the end guest
    - Partitions a real DASD into tiny usable segments called mdisks
    - Major issue! – Be sure mdisks don’t overlap!
      - Overlap each other – you can be wiping another guest’s data away
      - Overlap Cylinder 0 – you can be wiping access to the entire volume
    - Use a utility like DIRMAP to review User Direct for overlaps
    - Minidisks do not span volumes- under Linux use LVM
**z/VM – Characteristics**

- **Minidisk Cache – a VM Performance feature**
  - MDC was created a while ago for *small CMS* guests, **not Linux**
  - **Don’t** use MDC for Linux Swap areas
  - Minidisk only – not dedicated volumes
  - Must be 3380, 3390, 9345, or FBA type DASD
  - No larger than a 3390-27 (32767 cylinders)
  - Certain other restrictions when using certain I/O
  - ✓ In general – Linux I/O is Okay
  - Caching is to Real and/or Expanded Storage
  - See tuning notes on disabling specific minidisk from the cache
z/VM – Characteristics

- Dedicated Volumes
  - Under VM, your guest is the only guest that has any access to the volume
  - Other LPARs can still access it...
  - Volumes are attached within the “User Direct” with Dedicate statements
  - Volumes can be dynamically attached with the CP Attach command
    - Requires CLASS B privilege level (Real Resource Operator)
  - For production systems, databases, etc., this is a good method.

- Use LVM or ELVM to build your disk structure

- Other – Volatile
  - VDISK – good for paging (more later...)
  - TDISK – Good for temporary areas or paging
z/VM – Characteristics

- Wow – those volumes are small! What do I do for a 80Gb drive?
  - Use LVM and add volumes as required – three 3390-27’s give you ~90Gb

- FBA/ECKD volumes (DASD)
  - These tend to be small compared to what the server community deals with.
  - For Linux environments this causes more disk management

- Attach directly to SAN via FCP (Fibre Channel Protocol)
  - Allows for SCSI device access

- If you use a SAN –
  - If you need to perform a DR (Disaster Recovery), is the SAN and data available offsite?
  - Do you have the right Tools/Utilities to backup the SAN volumes?
  - VM can do it in some circumstances...
  - Is your SAN a single point of failure?
z/VM – Characteristics

➢ Watch out for other operating system utilities (MVS, z/OS)
  • For DR, it’s convenient to use Disk Technology that all your systems understand
    ✓ e.g. z/OS can take full pack dumps of Linux volumes
  • Some MVS utilities don’t play nicely with VM volumes
  • VM doesn’t use DSN’s and catalogs like MVS
  • Unless you know it works – take VM/Linux volume out of other OS scan and recovery packages

➢ If you have it in-house, use FICON rather than ESCON (mainframe only)
  • Increased number of concurrent connections
  • Increased distance
  • Increased link bandwidth
  • Increased channel device address support
  • Better yet – use FCP...
Using Barebones VM

- VM out of the box
  - TCP/IP
  - Directory maintenance
  - CMS
  - Other features
  - Optionally... DirMaint
  - Starting services with Autolog1
    - Wait for the "VM READ"
    - `Ac 191 a (noprof)` – otherwise it runs & logs off

- Some tuning is required for VM
  - How do you monitor what’s going on?
  - What commands do I need to know?
  - Watch what you adjust – not all zVM’s and Linux guests are the same.
  - As with any tuning – minor adjustments and review what happens.
    - More tuning later...
Devices – Real & Virtual

- The good news about virtual devices – they’re free and easy to define!

- What can you configure virtually?
  - HiperSockets – *Real & Virtual*
    - Real – LPAR transfer mechanism – only between LPARs on the same physical complex
    - Virtual – HiperSockets can be emulated Guest to Guest under VM
  - CTCA’s – Channel to Channel adapters – between VM Guests
  - IUCV – Inter-User Communications Vehicle – between VM Guests
  - Virtual Switches – emulate a physical network switch
    - Now included in z/VM *(as of R4.4)*
    - Don’t use customized Linux guest for routing
    - Play with MTU sizes (1492/8992) for your network’s best throughput
Devices – Real & Virtual

How?

- VM emulates dozens of pieces of mainframe hardware for you

- In addition, if you have a physical device you can share it or dedicate the device to a guest

- The User Directory can control devices
  - Real & Virtual devices are allocated when the Guest logs on to VM.
  - *See slide on Defining VM Guests*

- Many of the same devices can be dynamically attached with #CP commands.
  - Your guest must be aware of the potential devices
  - Your guest must know how to dynamically bring them on-line for Guest use.
  - Some device attachments required elevated VM Class privileges
  - You can also code the VM Diagnose instruction
Devices – Real & Virtual

A simplified view of combining the Real & Virtual networks

VM networks can be isolated from each other
Devices – Real & Virtual

Combining the Real & Virtual Network – a more complex view
z/VM – Starting Up...

- Start by installing your z/VM system
  - It’s not hard – but get some training!
  - Depending on the options purchased it can be as easy as an IPL
    - There will be some configuration though!
  - Your Linux & Z/VM systems support may or may not be the same person
- Once VM is up and running –
  - Some PTF’s & APARS
  - Directory (Guest) maintenance
  - Tuning
- You don’t have to know every VM command (CP/CMS) to use it effectively

So, we’ll assume from here that you have your VM system running...
z/VM Guests

- Guests are defined in the **User Direct** file.
  - Normally found on the MAINT guest ID – unless you have security packages

- **User Direct** is an editable text file
  - Use Xedit to manipulate

- Once edited use Diskmap to verify that DASD extents are not overlaid
  - Review the file using Xedit
    - `X user diskmap`
    - `Find overlap: /overlap`

- To put changes (dynamically) on-line: `DIRECTXA`
Defining a VM Guest

You can have included *profile* statements

```plaintext
PROFILE LINDFLT
  IPL CMS PARM AUTOCSR
  IUCV ANY PRIORITY MSGLIMIT 2000
  IUCV ALLOW
  MACH ESA
  SPOOL 00C 2540 READER *
  SPOOL 00D 2540 PUNCH A
  SPOOL 00E 1403 A
  CONSOLE 009 3215 T

NICDEF 300 TYPE QDIO LAN SYSTEM VSW1
  LINK MAINT 0190 0190 RR
  LINK MAINT 019D 019D RR
  LINK MAINT 019E 019E RR
```

Dedicated devices are “real” and can be used as they are in the *real* world – Linux in an LPAR.

- What you’re getting here is a mix of “real” & “virtual”
- We’ve found a starting memory size of 385Mb to be reasonable for small applications & developers
- Don’t define more CPUs than the physical system has
Defining a VM Guest

- Profiles Exec
  - Can be unique to each Guest
  - Runs under CMS only
  - Simple CP or CMS commands
  - Complex Rexx script
  - Edit with Xedit
  - To bypass:  *Access 191 A (noprof)* at VM Read prompt
  - For Linux
    - Set the CP environment
    - Attach Devices
    - Couple to other Guests
    - IPL the Linux system

- For this you’d want the User Direct guest entry to *IPL CMS* rather than 200
  *(previous example)*

/
*/ A simple Linux - Profile exec */
'VMFCLEAR'
'CP SET MSG ON'
'CP SET EMSG ON'
'SET IMCP ON'
'SET RUN ON'
'CP TERM CHARDEL OFF'
'CP SET PF10 RETR'
'CP SET RETR MAX'
'CP TERM HOLD OFF'
'CP TERM MORE 5 3'
'CP IPL 200 CLEAR'
So Why VM?

So why Linux under VM?

- Just like interactive CMS of the past, you can deploy 100’s of interactive Linux guests today.
- VM’s scheduler and Dispatcher have been tweaked for the new Linux workloads
- Like other hypervisors, you can prioritize your guests and distribute your resources
- Can the same be accomplished with big VMWare Intel Server – yes... to a point
  - If the I/O subsystem is going to be exploited, you might consider VM
  - If you’re accessing a quantity of mainframe data (VSAM, DB/2, ...) VM might be better
So Why VM?

- If you already have a mainframe – look into it
  - Ask IBM about trials and pricing

- Is it for everyone?
  - No, but for many it is probably worth looking into.

- How do I compare MIPs and Ghz?
  - Using examples on the Internet or from a vendor – you really cannot.
  - Too many varying areas to compare... like the old Apples :: Oranges
  - Much better to test an application
    - See if IBM will test your application for you
    - See if VM is available for a multi-month trial (Linux for s390 is available for trial)
    - If it needs to be a top performer – run Linux in an LPAR by itself

- Don’t just measure processor speed – other factors need to be reviewed
Linux – OS’s for the zSeries

- Open Source or Vendor Purchase
  - Still Free
    - Debian
    - CentOS
    - Slack/390
    - Compile your own
  - Commercial
    - Novell/SUSE Enterprise 9 & 10
    - Red Hat Enterprise
    - Others?
Software – Questions

- What is your company’s attitude?
  - Is Linux and/or Open Source safe to use ??
  - Do they just “feel” safe paying for a product?
  - Do you have the staff and experience?
  - How fast do you need this running?
  - You’ll eventually need standards and policies
Linux – In the Enterprise

What we installed...

- SUSE Enterprise 9.1 & 10
  - Getting NFS directories set up was difficult
    - SLES 10 corrects those problems
    - Standard Ethernet connection
  - Samba
    - MD /mnt/images
    - Smbmount //netname/resource /mnt/images –o username=xxxxx workgroup=domain
- IBM Java 1.5.0 X SDK (64-bit)
- Tomcat 5.5.16 & Axis 1.3
- A little minor Linux tuning...
  - On Demand Timer patch
    - sysctl -w kernel.hz_timer=0
  - ZIPl & fsTab
  - LVM & DASD
- Kerberos Authentication with our Windows Active Directory™
Linux – What’s It Good for?

- Running as a guest under z/VM –
  - You can have dozens or even hundreds of Linux guests
    - The total concurrently running guests does depend on real available resources
  - These guests can be developers or production servers
    - With z/VM Scheduling you can adjust performance

- Some services to consider...

  - Web Site/Apps ➔ Apache/Tomcat
  - E-Mail ➔ Scalix/Cyrus/Postfix
  - Firewall ➔ Firestarter/IPCop
  - File & Print ➔ Samba/CUPS
  - Database ➔ mySQL/PostgreSQL
  - Directory Svc ➔ LDAP
First Linux Setup – z/VM

Initially

- You need to setup a NFS on a workstation or server that has a CDROM
  - SMB and other connections are also available

- Extract CD images to NFS according to documentation
  - SUSE 9.1 wasn’t too good in this area. Documentation was a bit poor for the NFS layout
  - SuSE 10 is great!

- Once your z/VM is set up, you’ll use it to:
  - Punch to a virtual card punch
  - IPL from a virtual reader
  - Build your initial system
  - Back it up!! *(DDR or Flashcopy type tool)*
  - Once built, there are not too many times you’ll need 3270 access again
Dealing with the VM Console

VM Status

CP READ – Basically the Guest is in a hold stating running.

Running – The Guest is running.

MORE... – The screen buffer needs to be cleared.

VM Read – a CMS Read prompt.
What’s My Guest Look Like?

- Poking through the Virtual Machine
  - `#CP Q V ALL`
  - First column is the Unit Type
  - Second – Address of the unit
  - Followed by additional unit features
  - Dozens of CP Query/Set commands
Building the First Linux Guest

- Your first guest will be built from scratch – it takes more than 20 minutes!
- You’ll need to use a few VM commands, but nothing too in depth
- Create a SuSE Parm file and LinLoad Exec – Use Xedit to create both
  - SuSE Parm – contains Network & Linux media location information
    ```
    ramdisk_size=65536 root=/dev/ram1 ro init=/linuxrc term=dumb
    BROADCAST=??..??..?? GATEWAY=??..??..?? HOSTIP=??..??..??
    hostname=tlnx10.corp.stateauto.com
    INSTALL=NFS://??..??..??/NFS/SLES10DVD
    NAMESERVER=??..??..?? NETMASK=255.255.0.0 PASSWORD=MYPASS
    USERNAME=ANONYMOUS USEVNC=1 VNCPASSWORD=VNCPASS
    layer2=0
    portname=eth0
    readchannel=0.0.0300 writechannel=0.0.0301 datachannel=0.0.0302
    ```
  - Use Xedit to create it
    - X linload exec
      ```
      /* Linload exec*/
      'close rdr'
      'purge rdr all'
      'spool punch * rdr'
      'PUNCH SUSE IMAGE A (NOH)
      'PUNCH SUSE PARM A (NOH)
      'PUNCH SUSE INITRD A (NOH)
      'change rdr all keep nohold'
      'ipl 00c clear'
      ```
Run the Linload exec – just type **LINLOAD** and press the **enter** key

Soon you’ll get this prompt..............
Building the First Linux Guest

You’ll need to respond to a few prompts

- Press Enter to take the value you had coded in the SuSE Parm file, or enter a new value

Please choose the SCSI bar interface
1) CD16
2) LSC

Please choose the physical medium
1) Ethernet
2) Token Ring

Device address for modem channel (8 0 0010): 0 0 0800
Device address for video channel (0 1 0001): 0 7 0101
Device address for disk channel (9 0 0001) 0 0 1302
Device to use:

Enable OSI Layer 2 support?
1) Yes
2) No

Please enter:
- Link type: eth1
- 0.0.0.0 (default)
- 
Expected first login: root

Enter the IP address of your host server. If you do not use a host server, then use the following IP address:
- 192.168.0.10

Press Enter to take the value you had coded in the SuSE Parm file, or enter a new value
Building the First Linux Guest

Problems may occur...

- If your Virtual Switch is not set up correctly, or
- Your Linux Guest has not been granted access to the VSwitch...

```
geth: received an IDI TERMINATE with cause code 0xf5
geth: sense data available on channel 0.0.0.0
geth: set 0x0

dest 0x6
geth: arb 00 c2 61 17 16 53 e0 38 be 00 10 00 00 00 00
geth: arb 01 c2 00 00 00 00 00 00 00 00 00 00 00 00
geth: sense data 02 00 00 00 00 00 00 00 00 00 00 00 00 00
geth: sense data 02 00 00 00 00 00 00 00 00 00 00 00 00 00
geth: sense data 02 00 00 00 00 00 00 00 00 00 00 00 00 00
Geth: Initialization in handset failed! rc=-5
sit0: Disabled Privacy Extensions
```

*Set VSwitch VSW1 Grant Linux1*

- Rather than running Linload again, you will be presented with a menu:
  - Use the menu to start the installation again (#4)
  - Worst case –
    - #CP IPL CMS
    - Enter: LINLOAD [enter]
    - Reply to the prompts again!
Building the First Linux Guest

If everything is working correctly...

- You’ll see the following display and be directed to the correct Guest.

  ![Linux boot process]

- Use VNC Viewer on port 5901  
  aaa.bbb.ccc.ddd:5901

- Use a browser on port 5801  
  http://aaa.bbb.ccc.ddd:5801

- Use the password specified in the SuSE Parm file for VNC connection

![VNC Viewer screenshot]
Building the First Linux Guest

- From the VNC Viewer/Browser you will be prompted to continue the installation.

- Use the DASD Disk Management dialog to assign your Linux areas.
- **ALL** of your VM DASD are displayed
  - Only use the ones that were allocated to the Linux system

**Finally...**
- After your system is configured, add it to the Autolog1 guest’s Profile Exec.
- This allows it to startup automatically with the z/VM system.
Building the First Linux Guest

- Add a Shell script to modify:
  
  - Makes it easier when you initialize a cloned guest
  
  - Files that need to be modified would contain:
    - Network Names
    - Network IP Addresses
    - Server specific settings
  
  - Some of the files to modify might/will include:
    - etc/HOSTNAME
    - etc/hosts
    - etc/install.inf
    - etc/hosts.YaST2save
    - etc/sysconfig/network/ifcfg-qeth-bus-ccw-0.0.xxxx
    - Scan for others...
  
  - If you’re attached to an Active Directory remember to join the domain
    - Net ads join
Making a Backup

DDR – DASD Dump & Restore

- Comes with all versions of VM
  - Can be used as command or IPL'ed
  - Dump/Restore – Tape & DASD
  - Partial volumes or whole volumes
  - Minidisks or Real DASD

- Command Line example..........................

- If you’re lucky... you might have an instantaneous DASD copy feature...
  - Flashcopy™ – a feature of IBM TotalStorage Enterprise Storage server
  - Snapshot™ – feature of Storage Tek’s Shared Virtual Array™ systems
  - TimeFinder™ – feature of EMC
  - ShadowImage™ – feature of Hitachi
    - These can provide a near instantaneous copy of full volumes.
    - The IBM Flashcopy is a CP command within VM, but you need the hardware!
      - FLASHCOPY 101 0 END to 201 0 END

- Others:
  - FDR Upstream
  - Tivioli
  - CA

```
1/VM DASD DUMP/RESTORE PROGRAM
ENTER: in 200 3336
ENTER: out 400 3336
ENTER: copy all
HCP00? valid read is 0x0200
DO YOU WISH TO CONTINUE? RESPOND YES, NO OR REREAD: yes
HCP000? valid read is 0x0200
DO YOU WISH TO CONTINUE? RESPOND YES, NO OR REREAD: yes
COPYING 0x0200
```
Create a Second, Third, ... Guest

- What’s on your “Gold” image?
  - You might have one – you might have a dozen
  - Determine what your run-time images need & what’s in common.
  - Don’t make too many – you’ll have to maintain them!
  - Some ideas – Developer images, Production Web, Database, Firewall, ...
  - Attempt to keep the disk layout similar – *Standards!*
  - LVM is great for VM (standard DASD) environments
  - Save your gold images with non-production IP addresses
    - Networks don’t like duplicate names & addresses
    - Possibly use DHCP for your image
Create a Second, Third, … Guest

- If you have bare-bones VM use DDR
  - If you have Flashcopy like tool – use it (saves 15-20 minutes)

- Utilities
  - Have a utility on your image to allow easy installation of the new IP & server name
  - For your first IPL of the clone –
  - Don’t IPL on the production network – avoid duplicate names
  - Possibly detach your NIC under VM (#cp det xxx), then IPL
  - Run Shell Script to alter Names & Addresses

- If possible... a quick demo!
Accessing Your Linux Guest

A few things we’ve looked at...

- **PuTTY** ([http://www.putty.nl](http://www.putty.nl)) Telnet/SSH Client for Windows
  - Very fast – if you’re good at command line entry

- **VNC** – use your favorite, we’ve used the VNC Client (server not required)
  - Connect to Port 5901
  - If connection fails, enable *Version 3.3 protocol only* in options
  - Reasonable response time

- **VNC Browser**
  - Connect to Port 5801
  - You’ll need to original VNC password from installation
  - Really really slow!

- **Xming + Plink** = A nice GUI desktop interface to Linux
  - Fastest GUI we’ve encountered to date
  - Xming - [http://sourceforge.net/projects/xming](http://sourceforge.net/projects/xming)
  - Plink.exe - [http://www.putty.nl](http://www.putty.nl)
Accessing Your Linux Guest

- **Xming setup**
  - Download and install Xming & Plink
  - Select your Display settings
  - Session Type – *Start a program*

- Start Program –
  - Select using Plink.exe
  - Enter your GUI of choice – `/usr/bin/X11/kde`
  - IP or DNS of Linux Guest
  - Linux User ID

**CPU’s (IFL’s) & Memory…**
- Tuning the OS is important – how’s your code?
  - Don’t get in the habit of feeding good resources to a badly designed/coded application.
- Can be shared or dedicated if you have more than one
- Dynamic
- How to Share the processor(s)
  - Just let VM do it all... Probably not the best choice if you have production Linux guests
  - You can pre-allocate an absolute / relative amount of processing power to individual guests
  - You can fix the processing amount or let it grow... to a pre-determined maximum
  - The excess (if any) is available for everyone else

**Memory – Central & XStore**
- Rule of thumb – XStore no more than 25% of your overall memory.
A Bit of Tuning

- If this is a Critical Server consider running Linux in a private LPAR
  - z/VM can extract 1-12% of overhead

The Share Command

- Query Share userid
- Default for a VM Guest is Relative 100 nolimit
- Not available to all guests... must be class A

<table>
<thead>
<tr>
<th>Clas</th>
<th>Type Guest &amp; Command Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>System Operator</td>
</tr>
<tr>
<td>B</td>
<td>Real Resource Operator – can attach/detach real devices</td>
</tr>
<tr>
<td>C</td>
<td>System Settings – Can set global z/VM settings</td>
</tr>
<tr>
<td>D</td>
<td>Spooling Operator</td>
</tr>
<tr>
<td>E</td>
<td>System Storage Analysis</td>
</tr>
<tr>
<td>F</td>
<td>IBM Field Rep. – examines devices</td>
</tr>
<tr>
<td>G</td>
<td>General User – (typical guest)</td>
</tr>
</tbody>
</table>

Privilege class can be set in combinations AB, BCG, etc.
A Bit of Tuning

- VM has multiple Run & Dispatch Queues to handle interactive and long running (batch-like) guests
- You are allowed to over commit the resources, but do it carefully!
- To perform tuning – with or without a package – you’ll need some privileges...
  - A – G
- Set QuickDSP looks good, but:
  - QuickDSP can tie up your entire z/VM system – understand it before you use it!
  - Usually used for time critical guests like TCP/IP
- Determine what resources go with categories of Linux systems
  - Database, Web, File Share, Developer :: DASD, Memory, Paging, CPU cycles
- Linux may appear hung if a sever memory shortage occurs. Be sure you’re not too over committed on memory.
A Bit of Tuning

- VM Paging
  - Do you have Expanded Storage allocated? – You might want to...
    - VM’s paging is setup in a hierarchy – Expanded Storage is at the top.
    - Make DASD Page Area contiguous
    - Linux & other Guests OS’s can use XStore directly

- For Linux Swap Area – consider using VDISK
  - Use for critical applications
  - It maps a virtual mdisk to memory, not DASD
  - Read VDISK Performance tips first!  
  
  
  Goes away when you logoff or issue a Detach command

```bash
define vfb=512 as 600 blk 100
DASD 0600 DEFINED
Ready; T=0.01/0.01 21:41:34
q.v da
DASD 0190 3390 VV34E0 R/O  107 CYL ON DASD 34E0 SUBCHANNEL = 0009
DASD 0191 3390 PLN01 R/W  50 CYL ON DASD 36F0 SUBCHANNEL = 0000
DASD 019D 3390 VV34E3 R/O  146 CYL ON DASD 34E3 SUBCHANNEL = 0000
DASD 019E 3390 VV34E3 R/O  250 CYL ON DASD 34E3 SUBCHANNEL = 0000
DASD 0200 3390 LDAK01 R/W  10016 CYL ON DASD 37FF SUBCHANNEL = 0001
DASD 0592 3390 VV34E3 R/O  67 CYL ON DASD 34E3 SUBCHANNEL = 0000
DASD 0600 3336 (VDSK) R/W  104 BLK ON DASD VDSK SUBCHANNEL = 0000
Ready; T=0.01/0.01 21:41:40
```
A Bit of Tuning

- Dedicated Volume – use if you have large swap area
  
- TDisk – a temporary MDISK can be used...
  - It requires setup each time (under CMS) and is on real DASD

- MDISK – A standard Minidisk is good
  - Requires setup only once or when your swap size changes
  - Be sure to disable MDISK cache on the swap minidisk `MINIOPT NOMDC`
    - This immediately follows the minidisk you want it to affect, in the USER DIRECT file.

- Linux Disk Format
  - For 3390 like devices the current recommendation is EXT3
    - Same features as EXT2 but, with Journaling
    - Based on performance, recovery, and better use of DASD space
    - Keep an eye on Reiser4 – see:  `http://www.namesys.com`
A Bit of Tuning

- **DCSS – Discontiguous Shared Segment**
  - Linux uses the XIP (Execute-In-Place) file system driver (*like* ext2)
  - Improves performance by sharing “same Read-Only” data with multiple systems
    - Files in the DCSS must be EXACTLY the same, down to the last bit!
    - Cannot share any file that might be updated, created, or deleted
      - Not even timestamps can change!
      - Avoid certain “specific” system files in /etc
  - The DCSS is a z/VM feature that resides in a shared & specific range of memory
  - Any change to a DCSS requires that all Guests using it need to change.
  - Refer to IBM document: SC33-8287
    - “How to Use Execute-in-Place Technology with Linux on z/VM” Dec. 14, 2005
    - The 2004 version is for Linux Kernel 2.4
A Bit of Tuning

- How do you know how much a mainframe can hold?
  - Remember – not every developer needs their own private Linux guest
    - Just because it’s under VM, doesn’t mean it doesn’t support multiple users

- IBM will help you
  - Just ask your local System z Solutions Technical Specialist

- IBM Processor Capacity Reference for IBM System z9 and eServer zSeries (zPCR)
  - Provides capacity planning for z/OS, z/VM, & Linux on zSeries
  - Newly available to the general public (Aug. 2006) – Free!
  - http://www-03.ibm.com/support/techdocs/atsmastr.nsf/WebIndex/PRS1381
  - If bad link, Google: “Getting Started with zPCR”

- Several monitors and analysis packages are available from commercial vendors
  - Capacity Planning
  - Performance & Trending
  - Real-time Performance
  - IBM Tivoli OMEGAMON for VM
  - Velocity Software
  - IBM Performance Toolkit for z/VM
A Bit of Tuning

- Don’t forget built-in commands – not as pretty, but useful!

- CP Indicate
  - Numerous parameters
  - Help Indicate
  - Different output depending on privilege class (B,C, E or G)
  - #CP Ind Load

- Monitor command – (Classes A & E) collects events for later analysis
  - Requires external application to analyze and report the data

```plaintext
CP IND LOAD
AVGPROC-054% 02
XSTORE-000000/SEC MIGRATE-0000/SEC
MDC READS-000005/SEC WRITES-000001/SEC HIT RATIO-049%
PAGING-0/SEC STEAL-000%
Q0-00001(00000) DORMANT-00006
Q1-00000(00000) E1-00000(00000)
Q2-00002(00000) EXPAN-002 E2-00000(00000)
Q3-00010(00000) EXPAN-002 E3-00000(00000)
PROC 0000-056% PROC 0001-053%
LIMITED-00000
```
A Bit of Tuning

- CP Ind User ???? Expanded
  - Provide more details on a specific user

```bash
ind user xprtp1 exp

Userid=XPRTP1 Mach=ESA V=V Attached xstore=NONE
Iplsys=DEV 0200 Devnum=16
Spool: Reads=0 Writes=0
Owned spaces: Number=1 Owned size=384M
Pages: LockedReal=13 LockedLogical=2
  Primary space: ID=XPRTP1:BASE PRIVATE
    Defined size=384M Address limit=384M
  Private spaces: Number=1 Owned size=384M
    Pages: Main=97581 Xstore=0 Dasd=0
      WS=97568 Reserved=0
        ResidentLogical=54 LockedLogical=2
    Shared spaces: Number=0 Owned size=0
      Pages: Main=0 Xstore=0 Dasd=0
        ResidentLogical=0 LockedLogical=0
  Private paging:
    Xstore: Reads=0 Writes=0 Migrates=0
    Dasd: Reads=1 Writes=1
  Shared paging:
    Xstore: Reads=0 Writes=0 Migrates=0
    Dasd: Reads=0 Writes=0

CPU 00: Ctime=18 18:55:17 Vtime=0 03:52:07 Ttime=0 03:55:36
  Rdr=0 Prt=0 Pch=0 IO=1095029
```
A Bit of Tuning

- CP Ind Q
  - Show all users and what Queue they are currently in
    - Running, Page wait, I/O Wait, Idle, Ready, Enabled Wait, etc.
    - Number of pages in real storage
    - Estimate of guests Working Set Size
    - Priority in the Eligible or Dispatch list

<table>
<thead>
<tr>
<th>ind q</th>
<th>Q1</th>
<th>R00</th>
<th>00000696/00000676</th>
<th>CNXP1</th>
<th>Q3</th>
<th>PS</th>
<th>00064712/00064699</th>
</tr>
</thead>
<tbody>
<tr>
<td>OAK0788</td>
<td>Q1</td>
<td>PS</td>
<td>00064782/00064769</td>
<td>LINUX02</td>
<td>Q3</td>
<td>PS</td>
<td>00064873/00064860</td>
</tr>
<tr>
<td>CSECP1</td>
<td>Q2</td>
<td>PS</td>
<td>00097299/00097286</td>
<td>MSC01</td>
<td>Q2</td>
<td>PS</td>
<td>00064611/00064598</td>
</tr>
<tr>
<td>DTCVSW1</td>
<td>Q0</td>
<td>PS</td>
<td>00002692/00001752</td>
<td>SUBP1</td>
<td>Q1</td>
<td>PS</td>
<td>00130727/00130693</td>
</tr>
<tr>
<td>XFRD1</td>
<td>Q3</td>
<td>PS</td>
<td>00064823/00064810</td>
<td>LINUX01</td>
<td>Q1</td>
<td>PS</td>
<td>00091366/00091332</td>
</tr>
<tr>
<td>XFRTP1</td>
<td>Q3</td>
<td>PS</td>
<td>00097581/00097568</td>
<td>CNXT1</td>
<td>Q3</td>
<td>PS</td>
<td>00064613/00064600</td>
</tr>
<tr>
<td>CNXD1</td>
<td>Q3</td>
<td>PS</td>
<td>00064597/00064584</td>
<td>CNXM1</td>
<td>Q3</td>
<td>PS</td>
<td>00064665/00064652</td>
</tr>
<tr>
<td>IMG01</td>
<td>Q3</td>
<td>PS</td>
<td>00098304/00098291</td>
<td>TCPIP</td>
<td>Q0</td>
<td>PS</td>
<td>00003330/00002815</td>
</tr>
</tbody>
</table>

- See: IBM z/VM Performance (IBM SC24-5999) available on-line in PDF format:
  http://www.vm.ibm.com/perf/docs/
A Bit of Tuning

- Monitoring CPU Usage
  - Linux top command *(virtual)*
    
    ```
    top - 09:17:56 up 3 days, 20:13, 3 users, load average: 0.00, 0.13, 0.10
    Tasks: 73 total, 1 running, 72 sleeping, 0 stopped, 0 zombie
    Cpu(s): 28.7%us, 2.0%sy, 0.0%ni, 60.7%id, 7.0%wa, 0.7%hi, 0.3%si, 0.7%st
    Mem: 377000k total, 373176k used, 3824k free, 8704k buffers
    Swap: 0k total, 0k used, 0k free, 162608k cached
    ```
  - z/VM Indicate *(real)*
    
    ```
    CP IND
    AVGPROC-055% 02
    PROC 0000-056%  PROC 0001-055%
    ```
  - You must prorate your Linux CPU Usage stats with the real world.
    - If Linux says 50% use and z/VM says 10% then Linux is really using around 5%

- Minidisk Cache
  
  ```
  CP IND
  AVGPROC-054% 02
  XSTORE-000000/SEC MIGRATE-0000/SEC
  MDC READS-000001/SEC WRITES-000001/SEC HIT RATIO-046%
  PAGING-0/SEC STEAL-000%
  ```
  - Your **goal** is to have High Reads and a High Hit Ratio
  - High hit ratio with low reads is not a good use of the cache
A Bit of Tuning

- The Linux I/O Scheduler
  - Take a look at the zipl.conf file...
  - The `elevator` parameter can be: `as` | `deadline` | `cfq` | `noop`

```
[defaultboot]
default = ipl

[ipl]
traget = /boot/zipl
image = /boot/image
ramdisk = /boot/initrd
parameters = "dasd=200-20f, fixedbufferes root=/dev/dasda1 selinux=0 TERM=dumb elevator=cfg"
```

- CFQ – (Completely Fair Queuing) ✓
- Deadline – attempts to reduce I/O latency ✓
- NOOP – Basically, a FIFO queue
- AS – Anticipatory – delays requests to gather & re-order to improve disk seeks
A Quick Look at Security

- Security
  - Everyone’s Linux is different with respect to configurations
  - Out-of-the-box – the SuSE 9/10 is quite secure
    - A couple of years ago – everything was open... Now it’s locked up tight!
  - Don’t rely on just one Security scanning package
    - There are overlaps, but one package might provide tests that another does not
    - Always be checking for revisions on your scan packages
    - Always be looking at new packages
    - We tried: Nessus – http://www.nessus.org
  - Anti-Virus – not just for Windows machines!
    - Limited number of vendors
      - Remember what you’re running – 64 bit & s/390!
A Quick Look at Security

- At the z/VM / zSeries level we can have CCA
  - IBM Cryptographic Common Architecture
    - PCICA & PCICC (PCI Cryptographic Accelerator & PCI Cryptographic Coprocessor)
    - CEX2A & CEX2C (Cryptographic Express2 Accelerator & Cryptographic Express2 Coprocessor)
    - Key management
    - Data Integrity
      - Message Authentication Code (MAC)
      - Modification Detection Code (MDC)
      - Digital Signatures
    - Encryption/Description at all levels of the Network protocol stack
    - PIN authentication
A Quick Look at Security

- For Linux a zSeries Crypto driver is required
  - Provides improvements in performance for SSL
  - May be shared with numerous Linux guests

- Reference: zSeries Crypto Guide Update (Redbook)

- You don’t have a Crypto Card??
  - OpenSSL can help you... check out their web site: [http://www.openssl.org](http://www.openssl.org)
    - Some distributions already have it packaged
A Quick Look at Security

- Linux LDAP & PAM support allows for easy integration to your security

- PAM = Pluggable Authentication Modules for Linux
  - Lot’s of support across Linux implementations
    - Caldera
    - FreeBSD
    - SuSE
    - Red Hat
    - Others...

- RACF
- Microsoft Active Directory
- LDAP
- Kerberos
The Good & Bad

- The so-so –
  - After initial play period...
    - We needed training for VM & Linux
    - We needed a Linux Admin
    - We needed “real tools”
      - VM Indicate & Linux top are fine, but we need to look at the whole environment
      - Backup utilities – you can quickly get tired of RPM’ing files off to a SAN

- The Bad –
  - We had a runaway guest that was difficult to track down
  - An MVS Utility wiped out one of the Linux system volumes
  - We’ve had two unplanned power outages!
  - Once proven – we had to follow procedures & policies! 😞
The Good & Bad

The Good –

- VM has never crashed

- Linux (guests) have never crashed

- We have not lost any data (due to VM or Linux), not power outages!

- Even without personnel (VM & Linux) we’ve been able to deploy product systems successfully

- Mainframe procedures fit nicely in this environment

- Once proven – we had to follow procedures & policies! 😊
# Some Costs

<table>
<thead>
<tr>
<th>Item</th>
<th>What installations paid in 2003</th>
<th>Today</th>
</tr>
</thead>
<tbody>
<tr>
<td>One IFL Engine</td>
<td>$200,000 (~250 mips)</td>
<td>$95k - $125k (~480 mips)</td>
</tr>
<tr>
<td>z/VM initial Fee (and yearly support)</td>
<td>$45,000 ($11,000)</td>
<td>Ask IBM</td>
</tr>
<tr>
<td>SuSE Linux (Annual Fee)</td>
<td>$11,500</td>
<td>$26,000 /2 IFLs</td>
</tr>
</tbody>
</table>

- Today – IFL’s are around ¼ the cost of a standard CP configured for z/OS
  - Z9 EC Machines are ~ $125/IFL
  - Z9 BC Machines are ~ $95/IFL
  - Each IFL comes with Ginsu knifes!
Where We’re at Now…

- In July, after proving that Linux works and is very stable...

  - z/9BC (business computer) W02
    - 2 z/OS CP's
    - 2 IFL's
  - 32 GIG memory (*distribution*):
    - 13GB Production z/OS
    - 12GB z/VM Production
    - 4GB z/VM Production Expanded Storage
    - 2GB z/VM Test Central - New LPAR
    - 1GB z/VM Test Expanded Storage

- Remember – we run a lot of mainframe/legacy code.
  - Your shop might be the same... or not
  - We’re putting the interface near the data and business logic
    - That might be a server or it might be the mainframe
What’s Next for Us?

- Look for projects with a high probability of success
  - Pick other existing *small* applications
    - For throughput – make use of HiperSockets
      - DB/2 Connect Gateway
      - Other applications with high data rates
    - Anything that might run better in Java or that requires MVS data
      - Convert .NET to Java if necessary
  - Keep everyone aware of Linux *(developers up through management)*
  - Mono under Linux

- Keep it secure – just because it’s Linux doesn’t mean it’s hack proof.
  - Look at VLANs
  - Virtual Firewalls

- WebSphere 6.x migration from Servers (Intel) to Linux
  - Use Network Deployment feature to migrate between platforms and across data centers
Summary – Section 1

- You now have a very brief overview of z/VM and basic virtualization.
- The mainframe is a great server – not just a big Cobol compiler.
- Questions to ask yourself...
  - Do you have a mainframe?... If not, thanks for attending... come back for Part 2!
  - Do you have any extra capacity?
  - Do you have Mainframe Systems Programmer willing to have some fun?
  - Can they get hold of a demo copy of z/VM *(ask for version 5.1 or 5.2)*
  - If your testing looks good – Do you have applications that would benefit?
    - Multiple WebSpheres or DB/2 products are great candidates!
    - Don’t use Open Source as a cost benefit – it’s free on Intel too...
  - Begin to sell it (one-on-one) to architects and management
  - Get IBM in to help you – encourage the soft sell approach...
  - If you get approval
    - Get z/VM training
    - Get a qualified Linux support person
    - Get your procedures in place – it’s a real system now!
Section 2 – Introduction

During Part 2 we’ll be talking about...

- Virtualization – more general concepts
- Virtualization with Workstations and Servers
- Deploying Linux in a Virtualized server environment
- Experiences with Introducing Open Source to your company
- What’s in the future for Linux, Virtualization, and Open Source
More Virtualization

- You don’t need a mainframe to use virtualization.
  - You do need Memory!
  - 2+ CPUs also help
  - Newer technology should also help
    - Timings & Tests to be done...

- You don’t need lots of Windows or Linux images to do virtualization.
  - There are benefits beyond server consolidation
    - More later...
More Virtualization

- For the companies without mainframes there are alternatives...
  - **VMWare** – [http://www.vmware.com](http://www.vmware.com)
  - **Parallels** – [http://www.parallels.com](http://www.parallels.com)
  - **Virtuozzo** – [http://www.virtuozzo.com](http://www.virtuozzo.com)
  - **QEMU** – [http://fabrice.bellard.free.fr/qemu/](http://fabrice.bellard.free.fr/qemu/)
  - **Virtual Server / Virtual PC** – [http://www.microsoft.com](http://www.microsoft.com)
  - You can use your existing Intel servers, Sun Midrange to high end, Blade systems, ...
  - You can mix Windows Server, Linux, Windows XP – all on a single box
  - Backups and Disaster Recovery benefit greatly
  - Hardware upgrades/switch-outs are a greatly simplified
  - Individual OS’s can be transferred to other Virtual servers
More Virtualization

- Who can it help?
  - Server Team –
    - Save Space, power, money, ...
  - Developers
    - Desktop versions of available or setup private development areas on servers

- The Great Give Away... July 2006
  - Microsoft and VMWare give away some good stuff...
  - Not Enterprise level code, but still very good for a small company or developer
    - The free ones require a base operating system (Windows/Linux)

- Devices – the basic stuff
  - Much of the virtual equipment the guest sees are emulations of the real thing
  - Memory
  - Video
  - LAN Cards
  - Hard Drives
Virtualization – The Claims

What vendors are claiming as benefits...

- Server consolidation
- Easier Test & Development – no wasted servers
- Hardware migrations & upgrades are easier
- Business Continuity – easier failover & Disaster Recovery
- Capacity Planning – can increase & decrease somewhat dynamically
- Load Balancing – Virtual images can be moved server to server
  - In some cases while OS is up and active (VMWare)
- Faster – not really virtualization, but new chips containing VT are faster

Is it true?

- Different sites might have different perspectives
More Virtualization

- What do we look for in a Virtualization Product?
  - Disk Management
  - Good Memory use
  - Speed
  - Devices (networking & other I/O)
- With limited staff — **what we really look for & why...**
  - Easy to Install
  - Easy to Administer
  - Easy to understand
  - Easy to maintain
- What you might already be knowledgeable about:
  - DosEMU
  - Wine
  - Bochs (PC Emulator)
  - QEMU
  - V86
Hardware – The Small Systems

- What are Intel & AMD doing these days?

- **AMD-V**
  - Athlon 64™, AMD Turion™, and AMD Opteron™

- **Intel – VT** (*VMX – Virtual Machine Extensions*)
  - VT-x – IA32 Virtualization enhancements
    - 32-bit instruction set – see Intel site for *very* exact details
  - VT-i IFP / IA64 virtualization enhancements
    - 64-bit instruction set

- **The I/O chips ..........**
  - VT-d – Directed I/O
  - AMD IOMMU
Hardware – The Small Systems

- The GOAL

  - Performance
    - CPU
    - Memory
    - I/O

  - Get the work out of the VMM and down into the Chip
Terms

Some of the lingo –

- Virtualization – An abstraction of a hardware platform that can be the same or different from the underlying physical hardware, typically sharing the “real” resources among several operating systems that may or may not be similar.

- Emulation – is not virtualization – although virtualization makes use of emulation

- Multitasking – one operating system with numerous tasks running in parallel

- Hyperthreading – One CPU appears or simulates two CPUs

- Paravirtualization – is a technique that presents a software abstraction interface to guests that is similar to, but not identical to, that of the underlying hardware. This requires the guest OS be modified, but guest applications run normally.

- VMM –
  - Guests run with the aid of a VMM and a Host OS
  - Pretty much everything goes through the VMM (managed by Host) to execute
Hardware & Virtualization

- Some basics on the new Virtualization for Intel & AMD
  - Remember when there were 8 security levels / rings?
    - Probably not – but there were – Multics
  
  - As Desktops were developed we reduced that to 4 levels
    - Although most systems (even today) only use ring 0 & Ring 3
      - Privileged & non-Privileged modes
    - OS/2 used a few more (3)

  - New Virtualization introduces, what some call Ring -1
    - A new set of instructions specifically for Virtualization
      - Both AMD and Intel
    - The new chips require a VMM/Hypervisor to perform the calls
Hardware & Terms

- A few of the models we’re talking about today..
  - Some absolutely require a Host system
  - Others require some of the Host
  - z/VM is the Host and the Hypervisor

- With x86 types – the more passed though to the CPU the better
  - But, that’s only been available for the last year or so

Diagram:

- Host Apps
- Guest A
- Guest B
- VMM
- Host OS
- Hardware

- Guest A
- Guest B
- Guest C
- VMM
- Host OS
- Hardware

- Guest A
- Guest B
- Guest C
- Hypervisor
- Hardware

- Guest A
- Guest B
- Guest C
- Guest D
- Hypervisor/OS
- Hardware

Models:
- VMware Workstation
- MS Virtual PC 2004
- DOS, JVM, .NET
- Xen
- z/VM
- ESX Server
XEN & VMWare

- Both XEN & VMWare have free offerings
  - VMWare is commercial, but some versions are free to use
  - XEN is Open Source & is being packaged with some Operating Systems

- Is one really better – what are you paying for or not?

- Again – what is your company’s attitude toward Open Source tools?

- As Chip makers advance, the two are becoming similar...
XEN & VMWare

- For awhile (and still) we have quite a bit of emulation going on behind the scenes.
- With new chip-sets the virtualization is being pushed down to the chip.

Diagram showing the relationship between CPU/Memory, Storage, Virtual Disk Images, Networking Disk, I/O Devices, and layers of virtualization including native (bare hardware) virtualization, software virtualization, and paravirtualization.
XEN & VMWare

Questions – Commercial or Open Source

➤ Does it do what you need it to do?

➤ It might run the demo okay – does it run your application?

➤ How much time & effort are required?
   • To install?
   • To maintain?

➤ Is it documented well – or do you need to pay a support person??

➤ If it’s critical code – can you maintain it if the vendor/Open Source group disappears?

➤ Even commercial packages don’t do it all...
Not everyone has a mainframe with extra cycles...

- One of the solutions for the smaller platforms is VMWare

- Main offerings
  - VMPlayer
  - Workstation
  - Server & ESX
  - Tools (ESX) – Virtual Center, SMP, VMFS
    - P2V Assistant – convert physical Windows® systems to virtual images

- Is it like a z/VM LPAR?
VMWare

- The easy way... VMPlayer

- Free

- Dozens of pre-configured Appliances – some are interesting
  - Demo a couple of Appliances

- Make your own images with Workstation or Server
  - Check your licensing on the Guest systems.
  - You still must pay for each MS Windows product (some are by # of CPUs)
  - This is where z/VM begins to look nice... if you’re running Linux with commercial products
VMWare®

- **Remember licensing!!!!**
  - If you’re running 3 extra Windows® images – **you’ll need a license for each**
    - Don’t forget – your host (if it’s Windows®, counts as a license too)
  
  - If you’re running 3 extra FreeBSD® ... GNU Public License

- **Devices**
  - Depending on the product, some require specific drivers for your guest
    - See product notes on emulating SCSI adapters
    - Virtual or Real??

- **Networking**
  - It launches it own virtual Switches and Firewall
  - Some issues:
    - Virtual Terminal Servers & Virtual Web Servers on same ESX Server
    - Not a good selection for heavy I/O servers (*Large Mail Servers, Databases, etc.*)
XEN

- We’ve been talking about Guests –
  - in XEN terminology a guest is referred to as a Domain
  - Only the 2.6 Kernel is supported – old stuff might exist, but not worth using.

- As with other Virtualization techniques, XEN supports
  - Different Operating Systems, including Windows
  - Multiple domains running concurrently

- A very tiny hypervisor is installed.
  - Runs in Ring 0 (privileged)
  - Controls the real system resources
  - Domains use Rings 1 – 3 (32-bit) and Ring 3 only for 64-bit

- Source is available – check it out.
  - Embedded Assembler – get out your reference book!
**XEN**

- Rather than virtualizing the underlying hardware, modified guest OS’s work with the base system
  - Specifically, hardware/driver code is modified
  - Things are changing with the new Intel & AMD chips

- Domain 0 is the “Master” domain that manages the Control Interface
  - This allows management of CPU, Memory, and other I/O devices

- Networking – LVS Linux Virtual Server
  - Distributed in the 2.6.10 kernel
  - Patches available for earlier versions

- XEN is generally an answer to Virtualization on non-mainframe platforms
  - z/VM already provides some of the best virtualization available
Disk Image Options

- Lots of interesting options
  - Local or remote physical drives
  - Easy movement of entire guests to other VM Hosts
  - Some products provide conversion utilities between each other
    - Example...

- Categories as defined by VMWare
  - Virtual Hard Drive
  - Physical
  - Dynamic Expanding
  - Pre-allocated
  - Independent Persistent & non-persistent
  - Persistent & Non-persistent
  - Undoable
  - Append
  - Differencing (*Microsoft*)
Disk Image Options

- Virtual Hard Drive – not a real physical disk, but an image of one.
  - Appears as a large file or group of files (extensions like .vhd & .vmdk)
  - Can be on SANs or local hard drives

- Dynamically Expanding
  - You start with a small footprint and allow it to grow as required
  - Nice, if you don’t know how big a disk might become
  - Depending on its growth, you may encounter poor performance & real fragmentation

- Pre-allocated
  - The drive size is defined at the start and stays the same size
  - Does not grow.
  - Better performance than dynamic, since you don’t have to enlarge the drive
Disk Image Options

- Physical/Raw – a real hard drive
  - This is a normal hard drive that has been attached to the VMWare Guest.
  - Best option for high performance
    - No need to map to a Virtual Image

- 😊 Undoable Image –
  - The primary Disk image is not written to during the session
  - All changes are logged to a temporary disk image area
  - At the end of a session you can commit all changes or discard them
  - Great for software/DB Testing & First time software installations
    - For Testing, you always start with a “known” environment
    - For installations, you can mess it up numerous times without rebuilding the system

- Persistent Independent Images
  - Normal processing mode – as you write data it is retained on the image
Disk Image Options

- Non-persistent Independent Images
  - Upon reboot of the guest, this drive reverts to its original state
  - Nothing is actually written to the disk image

- Locking Files – enforces data consistency between guests that share a drive
Virtual Networking

- Typical networking options

  - Host Only
  - Bridged – Your virtual guest will look like a normal server on the regular network
  - NAT – Guest IP addresses are hidden behind a gateway
Demos of Virtualization

- Demo of Bugzilla with VMPlayer
  - 42 Mb VMWare image.
  - 3 IDE drive images (system, user, swap)
  - Automatic database configuration
  - Bugzilla in 5 minutes

- Demo of SuSE 10 Enterprise Desktop with Mono
  - A preconfigured virtual system with Open Source Mono

- Demo of XEN **under** Linux, **under** VMWare Workstation, **under** Windows XP
  - A little overkill – but just to show you that you can have multiple layers
VMPlayer

- The Mono demo *(example)* –
  - VMPlayer is free
  - A very convenient way to send a demo or preconfigured system
  - Comes as a suspended VMWare appliance
You can query the defined devices

Reconfigure it...

You can query the defined devices:

- Memory: 256 MB
- Hard Disk (SCSI 0:0)
- CD-ROM (IDE 1:0)
- Floppy
- Ethernet
- USB Controller
- Audio
- Virtual Processors: 1

You can change settings such as:

- Disk file
  - Name: SLED 10.vmdk
- Capacity
  - Current size: 2.62 GB
  - System size: 211 MB
  - Maximum size: 4.00 GB
- Disk information
  - Disk space is not preallocated for this virtual disk.
  - Virtual disk contents are stored in one or more files up to 3 GB each.

Reconfiguring the settings:

- Advanced
  - Virtual device mode
    - SCSI 0:0
  - Mode
    - Independent
    - Nonpersistent

Changes to this disk are discarded when you power off or restore a snapshot.
VMPlayer Example

- Device settings allow alteration of
  - Networking

<table>
<thead>
<tr>
<th>Device</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>256 MB</td>
</tr>
<tr>
<td>Hard Disk (SCSI 0:0)</td>
<td>Auto detect</td>
</tr>
<tr>
<td>CD-ROM (IDE 1:0)</td>
<td>Using drive (dev/rd)</td>
</tr>
<tr>
<td>Floppy</td>
<td></td>
</tr>
<tr>
<td>Ethernet</td>
<td>Bridged</td>
</tr>
<tr>
<td>USB Controller</td>
<td>Present</td>
</tr>
<tr>
<td>Audio</td>
<td>Default adapter</td>
</tr>
<tr>
<td>Virtual Processors</td>
<td>1</td>
</tr>
</tbody>
</table>

Network adapter

- Device status:
  - Connected
  - Connect at power on

- Network connection:
  - Bridged: Connected directly to the physical network
  - NAT: Used to share the host's IP address
  - Host-only: A private network shared with the host
  - Custom: Specific virtual network

- Virtual NIC 0 (default Bridged)
More on Open Source

- Just the term seems to un-nerve people

- Open Source was a harder sell than z/VM on the mainframe
  - $0 :: $???

  - z/VM, Mainframe, IBM carries a *positive* message in most mainframe shops

  - Management/Leads want to know – who supports me when I have a problem?
    - A valid question!

  - If you use an open source component..
    - Did you read the License??  **Related:** Donald Smith - *Open Source Licensing*
    - Did you know you might have to make your application Open?
    - Know what you’re trying to bring in-house
      - Is it secure?
      - Bring your Security group in on this early
    - By bringing in the Security group you may have a very good ally

- Joe Oak — Introducing Linux to the Enterprise (Parts 1 and 2)
More on Open Source

- Having a hard time convincing people?

  - Are you using JAVA? .NET?
    - Yes – there’s your foot in the door!
    - No matter if you’re JAVA or .NET Based – there’s Open Source available

  - Remember – you must EDUCATE people around you!
    - Knowledge is power, but not if you keep it to yourself.

- You’re probably using some and don’t realize it.
More on Open Source

- Are the applications/utilities/languages you have stable?
  - If yes, spread the news!
  - If no, maybe you should look for a more stable Open Source package??

- What are some good starters?
  - JAVA
  - Linux
  - ANT & nANT
  - subVersion
  - Tomcat / Apache / Axis
  - mySQL
  - Log4j & Log4Net
  - Way too many to list...
    - http://www.apache.org/
    - http://sourceforge.net
    - http://csharp-source.net/
More on Open Source

As you bring in potential Open source products

- Have a review process with other departments
  - Do other areas need this? Let them help sell it.
  - Have show & tell lunch times – *again, spread the news!*
  - As previously mentioned, get an Architecture group together
    - Have your CIO & IT Management back it

- Don’t go overboard
  - Look for sure successes – *start small*
  - Just a couple of packages to start.
    - People get scared if they see *too much change!*
  - Remember this is a *SLOW* process
    - Let people get used to new ideas and tools
    - By going slow, standards can be implemented in an easier fashion
More on Open Source

- Let others carry the ball for you – *if they want to*
  - If it’s a net benefit for all – there might be other enthused people to help you

- Use highly stable vendor products & keep the environment stable
  - Offer SLA to projects/customers
  - Ask other CSS attendees about their experiences
Experiences...

So, who owns and manages:

- The new Linux environments?
  - zSeries, Server based, or both?
    - Mainframe support group
    - Server support group
  - Who has qualified people?
  - What/who defines what support people might be needed?

- Open Source products?
  - Product support should fit nicely with current product support
  - What do you do for products under Linux?
  - What special qualifications should be looked for?
Experiences...

As always – get people talking together to **SHARE**:

- Knowledge
- Talents
- Experience
Experiences…

- With the new environment and products

- What type of Quality Control do you currently have?

- Will it *(Quality Control)* work for the new environment?

- Mainframes traditionally have well documented procedures
  - Are they flexible enough for a *very* flexible environment (Linux)?

- Have all departments signed off on this?
  - Audit
  - Legal – watch those Open Source licenses!
  - Security
  - Architecture
Experiences…

➢ Do you have a Disaster Recovery policy?
  • How will Linux and new products be rolled into the existing DR plan?
  • Do you have the proper products available to perform
    ✓ Backups
    ✓ File & Data Synchronization
    ✓ Timed/Scheduled events – Cron on numerous servers can become a management problem
  • Do your products (vendor/Open source) run on a 64bit s/390 system?
    ✓ Good reason to use Java
    ✓ Standard C is good if you have all supporting libraries

➢ Do you have any single points of failure?
  • Maybe your Power?
  • Maybe your network?
  • Maybe your SAN??
What’s Next for Us?

- Open Source is gaining ground – a little here – a little there – it adds up!
  - Problems with Here & There
    - No coordination of efforts – different departments doing different things
    - Different versions
    - Different setups
    - Time consuming to get everyone back in the same ball park
      - Some groups become very entrenched in their own infrastructure
  - This can happen when support people and support departments don’t change...
    - They might be overwhelmed or understaffed
    - They might not understand the technology
    - Fear of losing control
  - To help alleviate this:
    - Form an architecture group to keep your senior developers and architects informed
    - Share ideas, product test, demos, bring in vendors for presentations
    - Have a leader who has been given the authority to prioritize changes
Summary

For us...

- Will Open Source be accepted and be successful?
- Will z/VM & Linux be accepted and be successful?
- Where will Virtualization be heading?
  - Does it have a home with us?

Can any of these tools help you... as Developer, Support Person, or Company?

- Open Source
- Linux
- z/VM
- Virtualization

When dealing with change... **Patience!**

*You can learn many things from children. How much patience you have, for instance.*

- Franklin P. Jones
 References

- IBM – Getting Started with Linux on System Z9 and zSeries v5.2
- Linux/390 General Information
  - [http://linuxvm.org/]
- z/VM V5R2.0 PDF Library at IBM
  - [http://www-03.ibm.com/servers/eserver/zseries/zos/bkserv/zvmpdf/zvm52.html]
- z/Architecture
- z/VM Paging Tips
  - [http://www.vm.ibm.com/perf/tips/storconf.html]
- IBM Linux on zSeries Device Drivers, Features, and Commands (March 23, 2006)
- IBM z/VM and Linux on IBM System z: The Virtualization Cookbook
  - Michael MacIsaac and Jin Xiong
- Virtualization From the Desktop to the Enterprise, Apress
  - Chris Wolf & Erick M. Halter
References

- **Linux on IBM zSeries and S/390: High Availability for z/VM and Linux** (IBM Redbook)
  - Erich Amrehn, Ronald Annuss, Bernhard Dierberger, Richard Plasun, & U. Sager

- **XEN**
  - [http://www.cl.cam.ac.uk/Research/SRG/netos/xen/](http://www.cl.cam.ac.uk/Research/SRG/netos/xen/)

- **IBM Developer Works – Linux on system z**
  - Check-out Tuning Hints section
Trademarks

The following terms are trademarks of:

- Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
- z/OS, z/VM, zSeries, DirMaint, HiperSockets, IBM, S/390, Rexx, and OS/2 are trademarks or registered trademarks of International Business Machines Corporation in the United States.
- Microsoft, Windows, Windows XP trademarks of Microsoft Corporation in the United States, other countries, or both.
- Intel is a trademark of Intel Corporation in the United States, other countries, or both.
- Linux is a trademark of Linus Torvalds in the United States, other countries, or both.
- SuSE is a registered trademarks of Novell, Inc., in the United States and other countries.
- Red Hat is a registered trademarks of Red Hat, Inc.
- Xen and XenSource are either registered trademarks or trademarks of XenSource Inc. in the United States and/or other countries.
- VMware, GSX Server, ESX Server & VMotion are trademarks (the "Marks") of VMware, Inc.
THAT CONCLUDES MY TWO-HOUR PRESENTATION. ANY QUESTIONS?

DID YOU INTEND THE PRESENTATION TO BE INCOMPREHENSIBLE, OR DO YOU HAVE SOME SORT OF RARE "POWER-POINT" DISABILITY?

ARE THERE ANY QUESTIONS ABOUT THE CONTENT?

THERE WAS CONTENT?