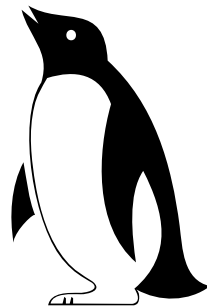


Experiences using VM on System z, the Infrastructure of Oracle

David Kreuter
VM Resources Ltd

Agenda

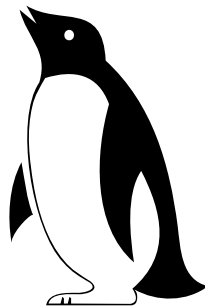
- User installation experiences with Oracle on Linux on System z using z/VM
- Description of installation configurations
- Lessons learned and best avoided
- Why we use z/VM and Linux on System z for Oracle
- Production results



Client profiles

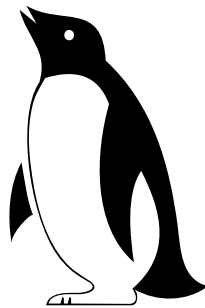
1. Large government service bureau
2. Police force
3. Software as a Service company

All with Oracle servers!



Client profiles: Government Service Bureau

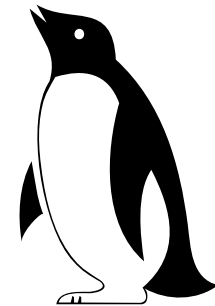
- IT service provider for many government offices (125)
 - *Going back several years:*
 - Exiting mainframe shop
 - 5 z890 + 1 z800 + 1 G5 on the floor on 3 sites
 - 1 z9/EC dedicated to Linux on z/VM
 - 450+ physical servers (750+ logical) (HP, SUN, pSeries, ...)



Government Service Bureau: History

- **Initial needs :**
 - Must solve many issues with the intermediate platform
 - Many operating systems
 - Many versions
 - Unsupported software
 - Unsatisfactory DR
 - Fast growing (unprecedented growth)
 - Understaffed
 - Need a flexible solution with rapid deployment
- **Mainframe is a stable and mature environment**
 - Staff is available and at early stages of their careers
 - Solid and well controlled DR process (MVS-like)

*Why choose Oracle on
Linux on System z?*



Government Service Bureau: Timeline

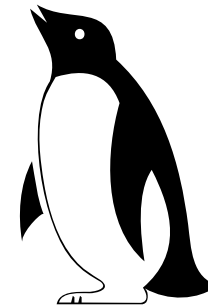
- **In 2003, beginning of the proof of concept (end-to-end)**
 - A mainframe was available (z800)
 - Installation of the software (z/VM, Novell SUSE Linux Enterprise Server (SLES), Oracle/DB, WAS, TAM and LDAP)
 - The objective was to prove the functionality and stability of the solution, plus the transportability of the tested applications

... performance characteristics were secondary
- **Spring 2005, first version of the business case which demonstrates benefits (financial, intangibles and human resources)**
- **Obtained approval from the board of directors**
- **Autumn 2005, installation of the new mainframe (z9 EC)**
- **Start of the first phase of the project**
 - Oracle/DB migration
- **Today in 2010 running 150 Linux Oracle Servers (and other servers running WAS, LDAP, TAM).**

Government Service Bureau: Original Configuration

- 1 z9 EC mainframe with 5 IFLs (~ 3000 mips)
- 5 LPARs
 - Oracle/DB
 - WAS
 - TAM & LDAP
 - Service Zone
 - Lab Zone
- 40 internal networks
- Software
 - Novell SLES (versions 8 & 9)
 - z/VM v.5.2 +
 - Oracle/DB (versions 9i & 10g)
 - Velocity Software Performance Tools
 - CA products (Automation, Scheduler)

Reduced Oracle licenses from 40 to 4!



Government Service Bureau: Original Configuration

- 1 z9 EC mainframe with 5 IFLs (~ 3000 mips)
- 5 LPARs
 - Oracle/DB in LPAR with 3 IFLs
 - WAS
 - TAM & LDAP
 - Service Zone
 - Lab Zone
- 40 internal networks
- Software
 - Novell SLES (versions 8 & 9)
 - z/VM v.5.2 +
 - Oracle/DB (versions 9i & 10g)
 - Velocity Software Performance Tools
 - CA products (Automation, Scheduler)

2007



Government Service Bureau: Current Configuration

- 1 z10 BC mainframe with 4 IFLs (~ 3000 mips)
- 5 LPARs
 - Oracle/DB in LPAR with 3 IFLs
 - WAS
 - Service Zone
 - Lab Zone
- Over 40 different networks
- Software
 - z/VM v.5.4 +
 - SLES8 SP4 Oracle 9i (holdouts/stragglers)
 - SLES9 SP3 Oracle 10gR1
 - SLES10 SP1 Oracle 10gR2
 - Velocity Software Performance Tools
 - CA products (Automation, Scheduler)

2010



TRANSZAP Customer profile

“The Complete e-Procure-to-Pay System”

- *Serving:*
 - 4,200 Companies
 - 44,000 Users
- *Processing:*
 - \$80 billion in transaction detail yearly



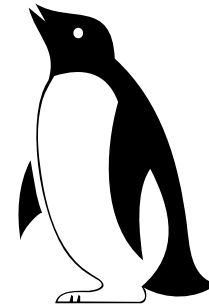
SaaS-based (Software as a Service)

TRANSZAP

Why choose Oracle on Linux on System z?

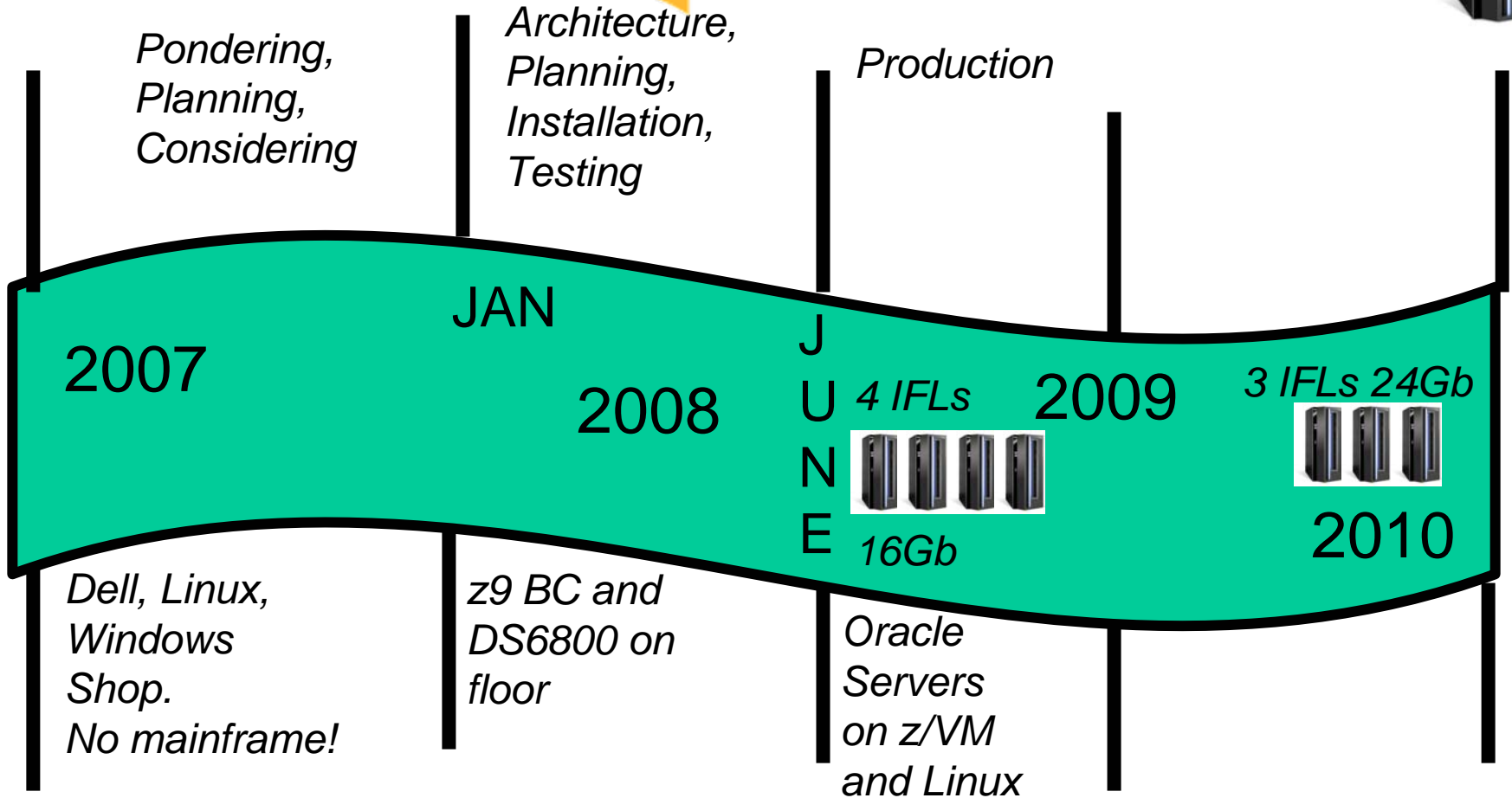
System z for the new mainframe customer

- Windows, Dell and Linux customer.
- Handle rapid company growth.
- Avoid server sprawl.
- Control costs with virtualization on System z.
- Applications are readily available: Oracle, WebSphere, Open source



TRANSZAP

System z timeline



TRANSZAP Lesson learned

Reduced from 4 to 3 IFLs



- *Workload reduction?*
- *Workload redeployment to other platforms?*
- *Decrease in transaction rates?*
- *Decrease in database size?*
- No! Workload, transactions and database size increased
- Memory increased to accommodate new workload
- The application was tuned!

TRANSZAP

z9 BC model R07-A01



3 IFLs

24Gb

DS6800



15 Tb

TS3400 Tape Library



TRANSZAP

z9 BC model R07-A01



3 IFLs

24Gb

z/VM 5.4

++ tools: systems management, automation, deployment, and monitoring

Linux Suse SLES 10

Oracle 10G EE

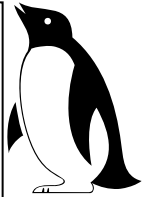
TRANSZAP

Production Challenges Early in Life

- **Pain point: unacceptable performance impact due to migration from Oracle 9i to 10g.**
 - And a platform change.
- **Significant issues with Oracle 10g configuration and poor performing SQL queries.**
- **System z9 sizing challenges:**
 - *2007 evaluation: 2 IFLs for all system workloads (application and DBMS).*
 - *But: 100% growth in database workload from 2007 to 2008.*
- **Severity 1 event on first day in production with both IBM and Oracle.**
 - *IBM CritSit team worked together to resolve issues (3.5 days)*
 - *Temporary capacity upgrade (CUoD) from 2 to 6 IFLs (now using 4).*
 - *Ongoing monitoring of situation and analysis of Oracle performance.*

TRANSZAP

*Today: Linux virtual machines
Servers named after minerals*



Peridot: 1 TB of data



Borax: 1 TB of data



Galena: 1 TB of data



Citrine: 500Gb



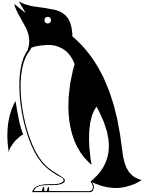
Barite: 400Gb



Cuprite: 3Tb (NFS server)



Garnet: 400Gb

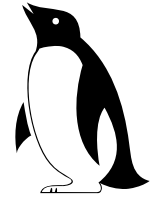


Best practices

- Use a performance monitor
 - The IBM Performance Toolkit
 - Generating daily csv format files used for analysis and reporting
- Using DIRMAINT for directory and storage management
- Use the z/VM wrapper: CMS tools.
- Using small locally written automation and remote control tools



Best practices: performance analysis



Feb CPU Usage: 3 IFLs

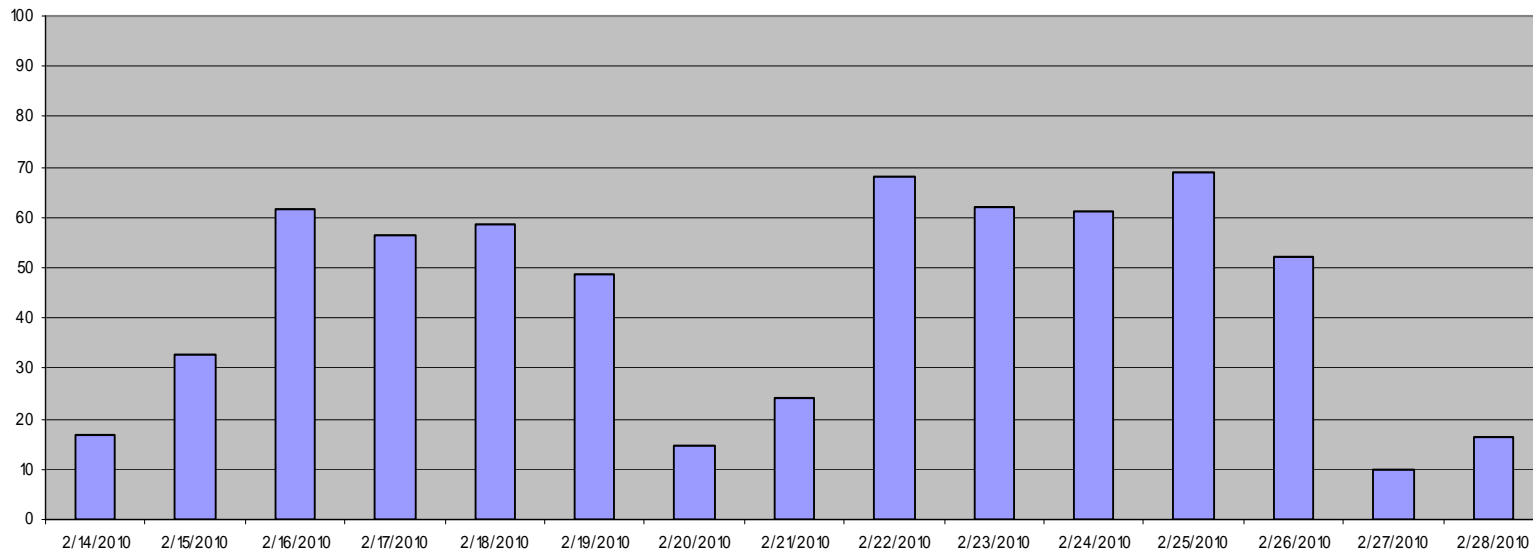
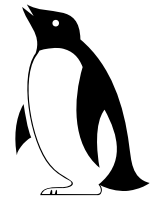


Chart generated from CSV data derived from Performance Toolkit data. Use the CP and CMS wrapper tools well!



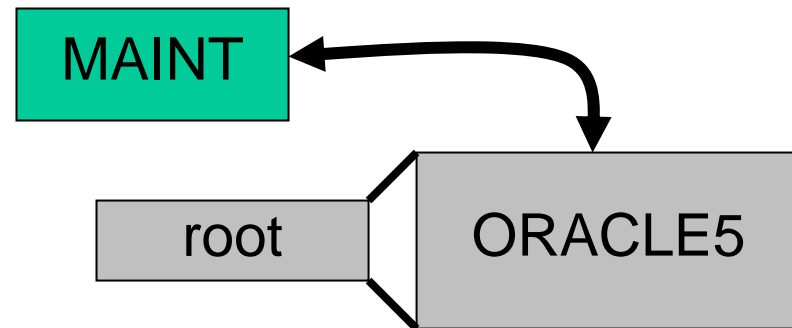
Best practices: locally written tools



```
query secuser ORACLE5
      Secondary
Userid   Userid   Status
ORACLE5 MAINT   Logged on
Ready; T=0.01/0.01 08:47:51
type gal ena exec
```

```
:
gal ena pwd
ORACLE5: pwd
ORACLE5: /root
ORACLE5: gal ena: ~ #
```

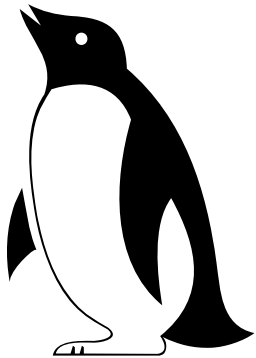
```
/**/
'CP SET SECUSER ORACLE5 *'
Address 'COMMAND'
parse arg cmd
'CP SEND ORACLE5' cmd
:
```



Best practice: Use CP and CMS tools provided from the z/VM wrapper

TRANSZAP

System z Plans

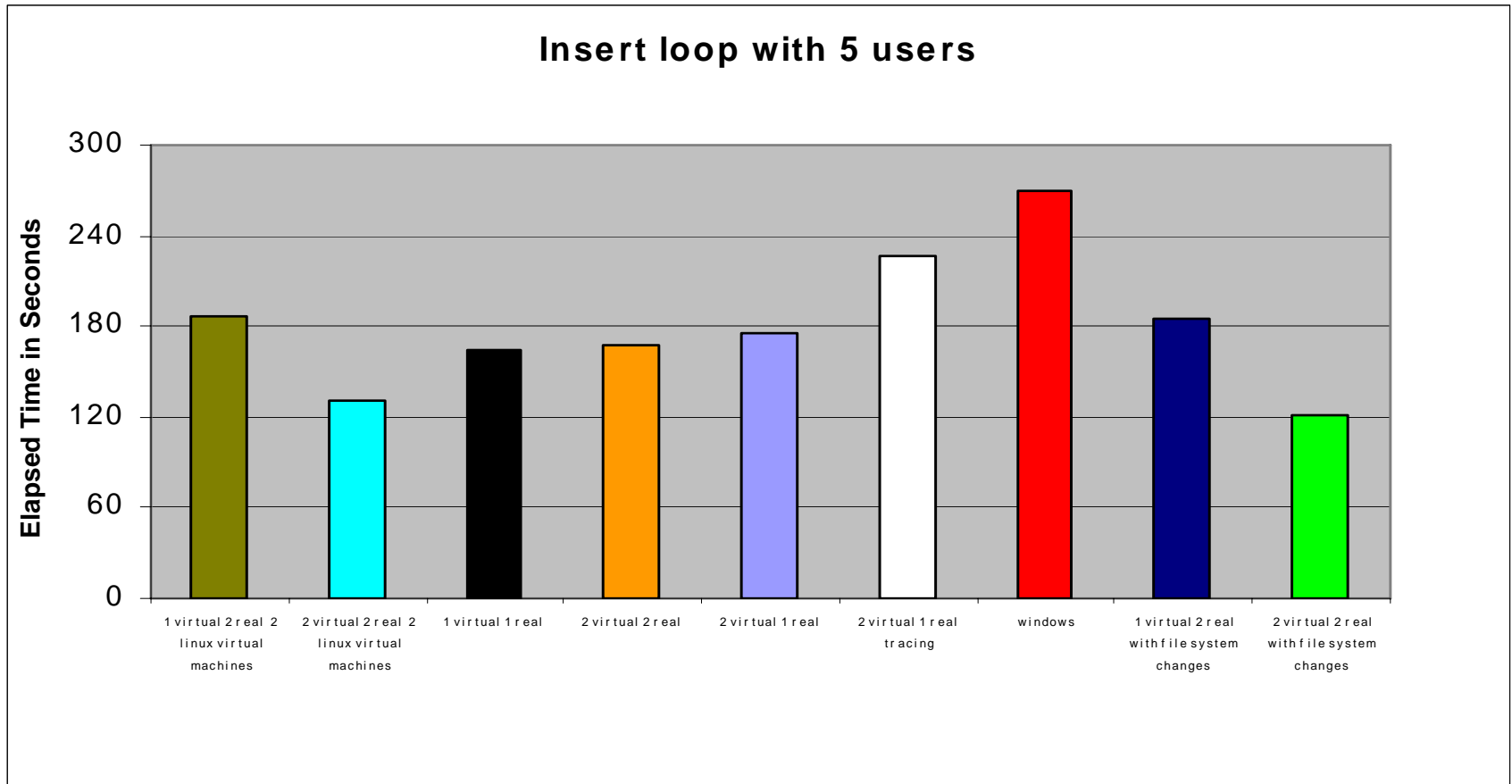


- Additional IFLs and memory
- Horizontal and vertical growth
- More Oracle servers
- More disk space
- Open source: Apache and Tomcat

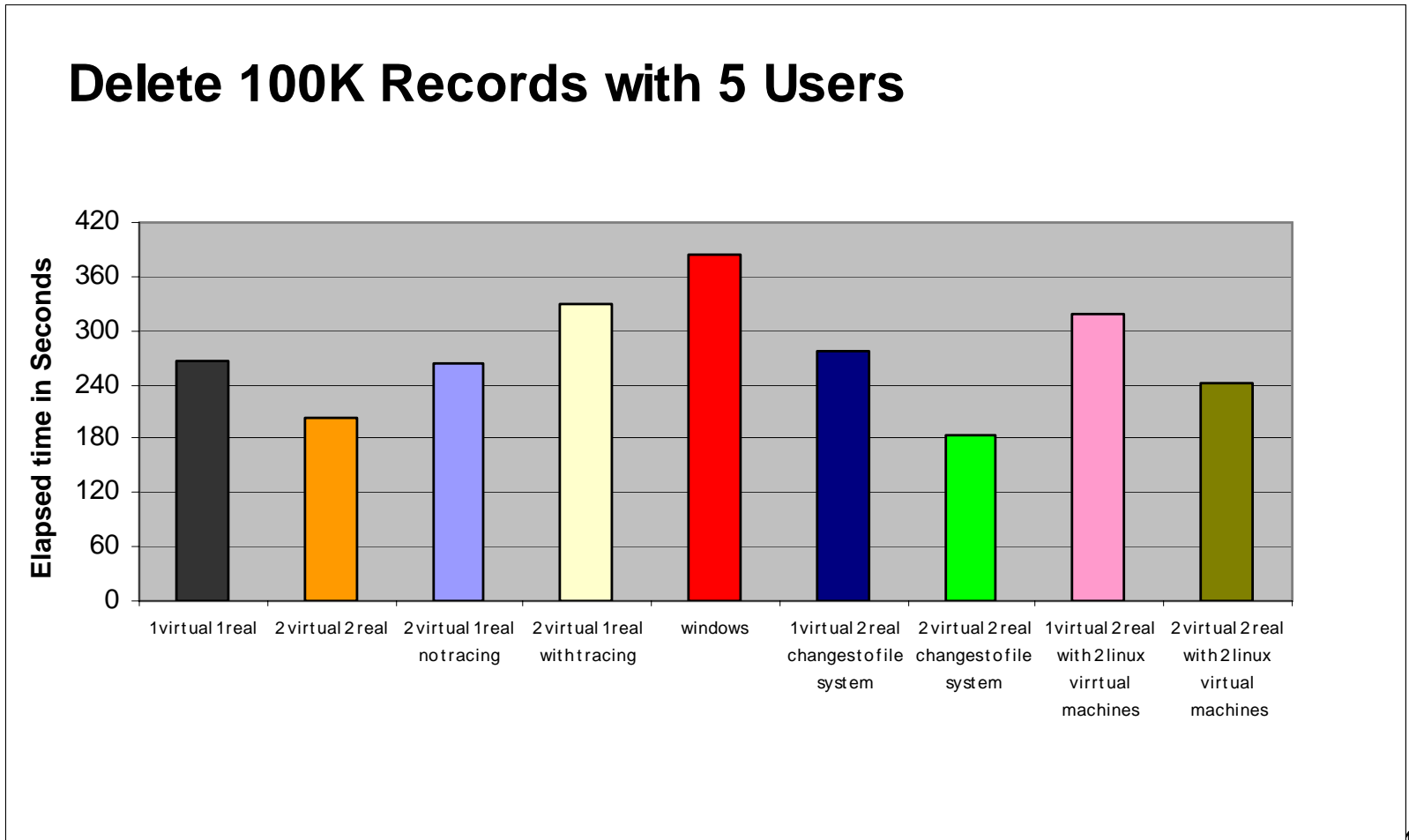
Client Profile: Major Police Force

- Will deploy Oracle on z10BC 2 IFL machine with z/VM, SUSE SLES Linux, and Oracle in 2010
 - Releases??/???
- Completed study in early 2009 with sample scripts executing in Windows compared to System z.
 - Performed on z/890 with two IFLs
 - Windows machine was a 4 way.
- Most scripts performed better on System z
 - Exception was a long running script.
 - *in production z/10 will handle CPU intensive work better than the z/890.*
 - System z performed better than windows on insert and delete loop tests with multiple users.

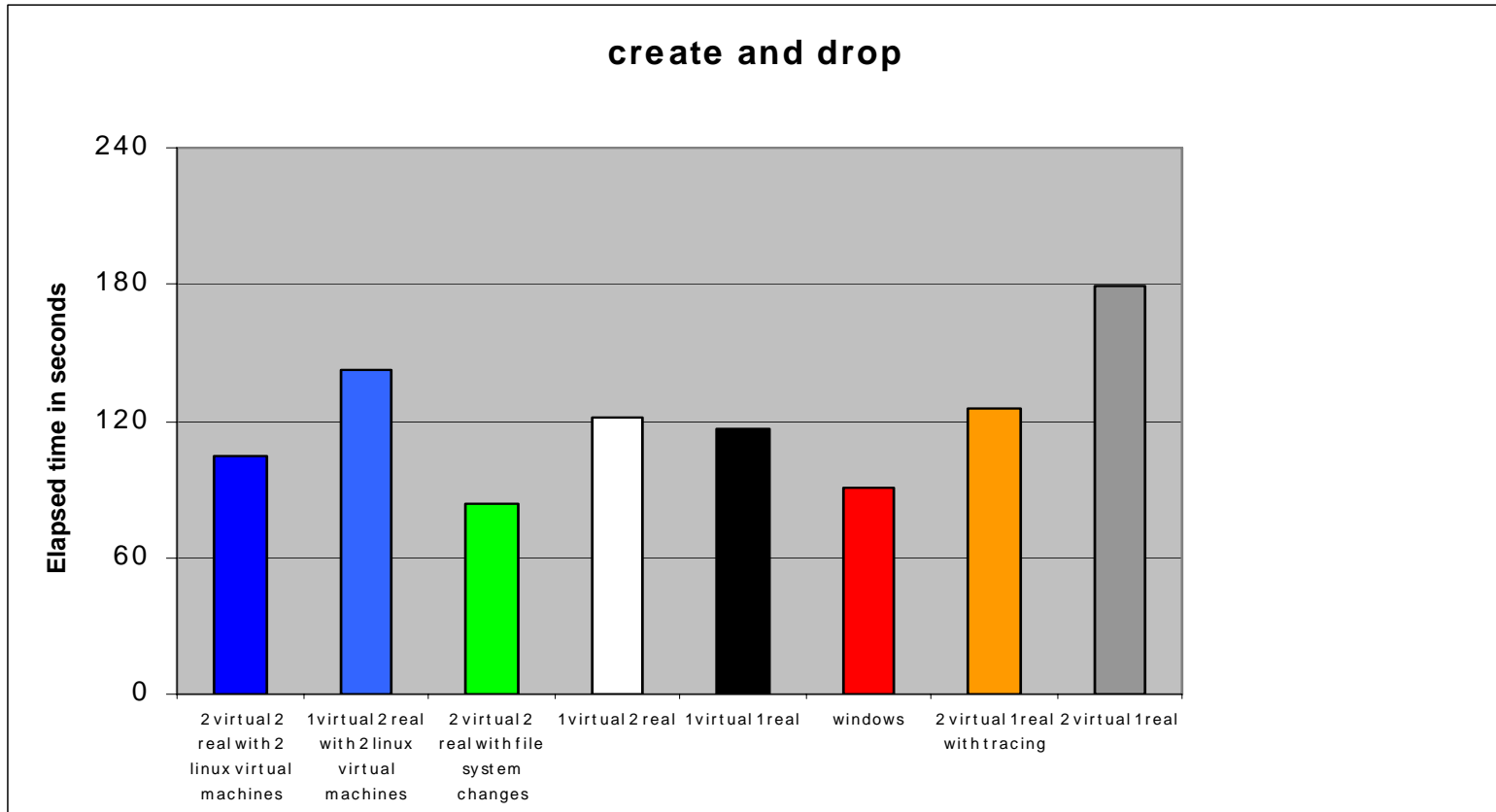
Client Profile: Major Police Force: Comparison Report of Record Insert



Client Profile: Major Police Force: Comparison Report of Record Delete



Client Profile: Major Police Force: Comparison Report of Record Create and Drop



Best Practices: Memory and Swapping

- Calibrate Linux virtual machine size so it consists of SGA adequate for workload and additional memory for kernel
 - Do not size virtual machine too high - wasting precious resource
- Define multiple swap disks on virtual disk space:
 - Cascading priority
- Set sysctl swappiness value low – let VM do paging!
`sysctl -a | grep vm.`
`vm.swappiness = 15`
- Set swap size to be around 50% of virtual machine size

```

INTERIM period, select current or average data)
Data Userid      ....   Size
>System<
ORACLE3          .....   512M
ORACLE4          .....  1024M
    
```

Both of these machines were reduced by 50% when moved from physical to virtual server

Perfkit data from screen "22" – user paging load

Linux Swap Usage

- Linux Swap Space – Purpose
 - What is Swap? Storage space auxiliary (on disk) to hold system memory
 - Holds system memory pages that are not currently needed by the processes that are running
 - Physical memory + swap space memory = virtual storage size
 - Allows processes larger than the physical storage to be executed



Best practice: Linux Swap Usage

Hints and Tips: Priorities

- Linux Swap Space – Swap priorities
 - Different swap devices can be assigned different priorities
 - Priorities are values between 0 and 32767 Higher numbers indicate higher priority Linux uses the higher priority swap devices first
- Define two swap devices –
 - A smaller swap area with a higher priority on a fast “device”, e.g., VDISK or DCSS
 - A larger swap area with a lower priority on a slower “device”, e.g., a real disk
 - Monitor swap area usage

Linux Swap Usage Hints and Tips: Swapiness

- *Try the SWAPINESS setting at your shop, monitor, adjust for your purposes. This is not a set it and forget it value.*
- Good results with setting swapiness to 0 or low value (< 20).
 - When Linux machine sizes properly planned avoids double paging (i.e. let z/VM do the paging).
 - Avoidance of double paging

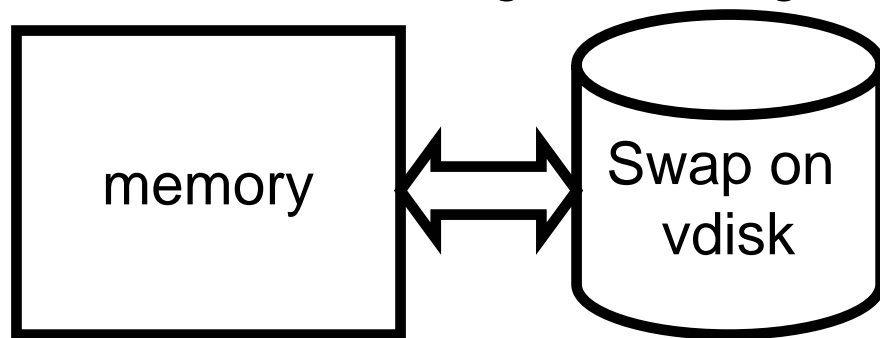
- **SWAPINESS:**

From root: echo 0 /proc/sys/vm/swapi ness

- When swapiness = happiness:
- *Change the vm. swapi ness parameter in /etc/sysctl . conf*

Oracle on Linux on System z: Hints and Tips: Virtual machine sizes

- Must plan storage sizes to work in system z!
 - Usually lower than storage size of same workload on other platforms
 - *Since Linux is an all in use all the time memory system, memory will be quite happily used for file and buffer caching. This might be self defeating in a heavily shared environment like VM. Adjust storage sizes so that workloads do little or no swapping at average utilizations.*
 - Get the a product to measure your system!
- Swap sizes usually less than 50% of assigned storage values.

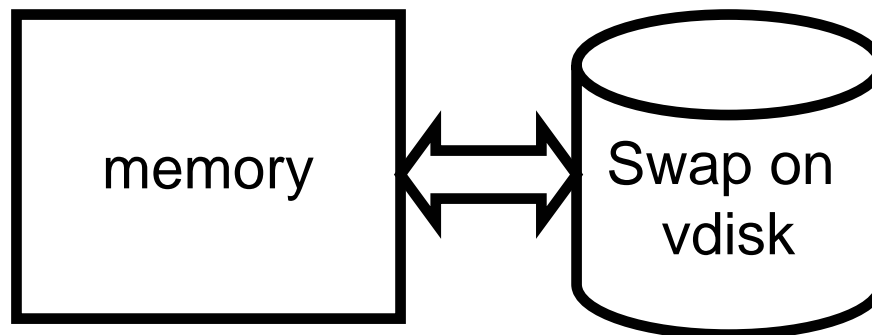


Linux Swap Usage Hints and Tips: Performance Reports on Vdisk

Interval 14:00:18-14:45:18, on 2010/03/16 (INTERIM period, select current or average data)

Owning Userid	Virtual Devno	Disk Blocks	Links	Data Space, Rate per Sec.							Nr of Pages			
				I/Os	Stl	Rd	Wr	X-rd	X-wt	X-mg	Res	Lock	XST	DASD
>System<	1181k	1	.044	.030	.068	.000	.000	.030	.000	32683	0	238	7234
ORACLE3	0160	200000	1	.000	.000	.000	.000	.000	.000	.000	0	0	0	237
ORACLE3	0161	1572k	1	.057	.000	.045	.000	.000	.000	.000	2311	0	0	190k
ORACLE4	0160	200000	1	.005	.000	.076	.000	.000	.000	.000	3597	0	0	24954
ORACLE4	0161	1572k	1	.006	.000	.042	.000	.000	.000	.000	3539	0	0	196k

Perfkit data from screen "B" – virtual disks in storage

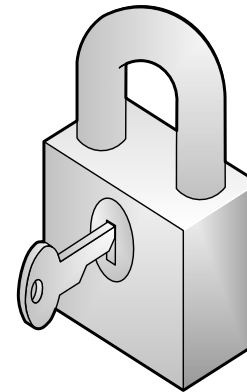


So ... What do we know now? What have we learned?

- Oracle is a natural application for Linux on System z.
- Vertical and horizontal server growth.
- Must calibrate virtual machine storage size
 - Almost always smaller then in distributed environment
- Plan for swapping but avoid in most cases.
- Must use a performance monitor:
 - You paid for it must know how the resources are being utilized
- Must keep z/VM and Linux Secure
 - Linux security is ... Linux security
 - Secure the z/VM environment

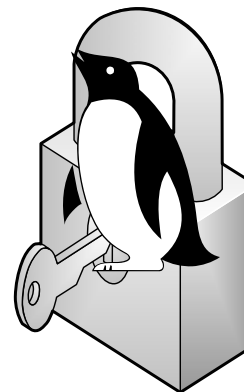
RACF

- In z/VM RACF provides security, access control, audit capabilities, and administrative control.
- Runs in a disconnected service virtual machine.
- Protects logon passwords, minidisk password, spool devices, and guest lans.



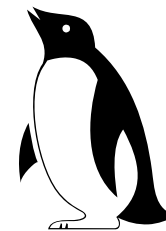
RACF and Linux virtual machines

- Standard protection mechanisms for userid password and minidisks.
- Additional protection available for guest lan membership.



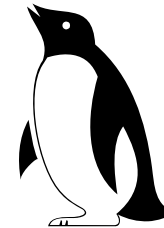
Intrusion Detection

- Intrusions or attacks can happen on your system.
- Securing networks through firewalls helps.
- There are intrusion detection methods available for Linux:
 - Network port scanning
 - Auditing of file systems, system logs, system resources
 - Honey pot traps
- While automated still requires sysadmin attention.

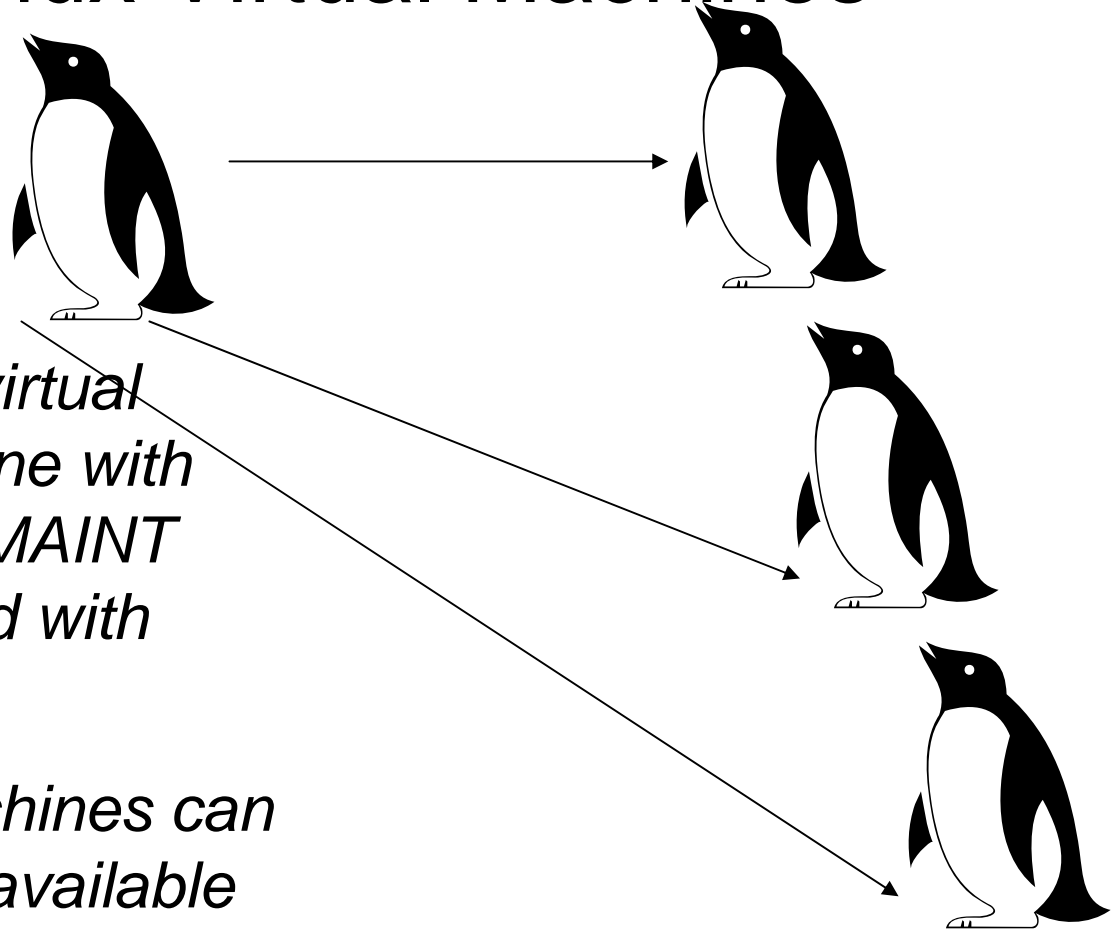


Hardening

- Customizing linux to make it more secure.
- Different levels of hardening depending on requirements of server.
- Harden:
 - Password authentication
 - Remove unneeded userids
 - Remove unused network services
 - Control access to files



Cloning Linux Virtual Machines



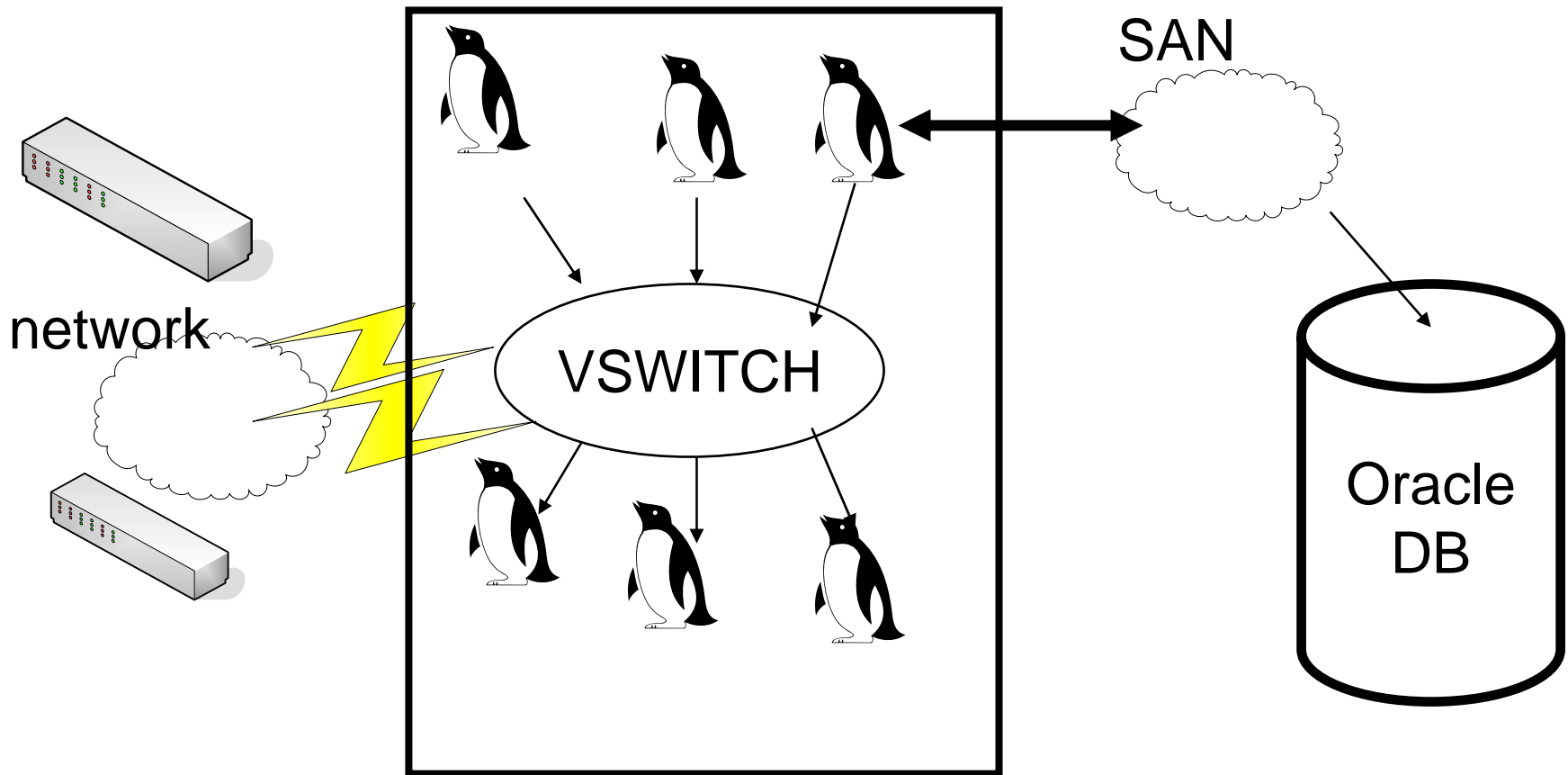
Cloning Linux virtual machines is done with CP, CMS, DIRMAINT commands, and with Linux tools.

New Linux machines can be cloned and available on your networks within minutes.

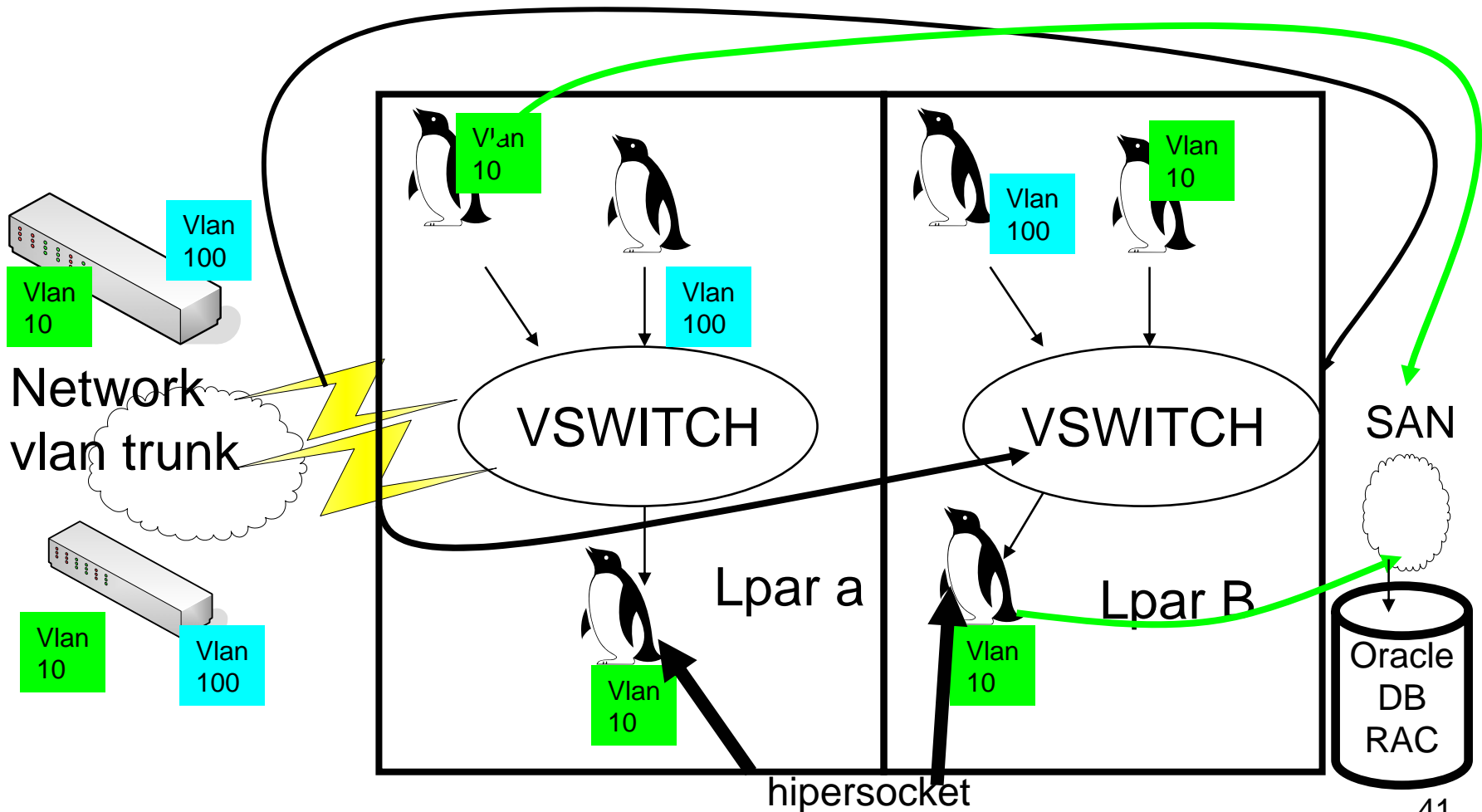
Lessons Learned: Best Avoided

- Vendor documentation and recommendations often not optimized for System z.
 - Creates confusion with personnel new to System z
- Overzealous Proof of Concept Deliverables
 - Don't attempt to test the everything in the house – just choose a room
- Executive scope creeping
 - Have a sponsoring manager interested in your success

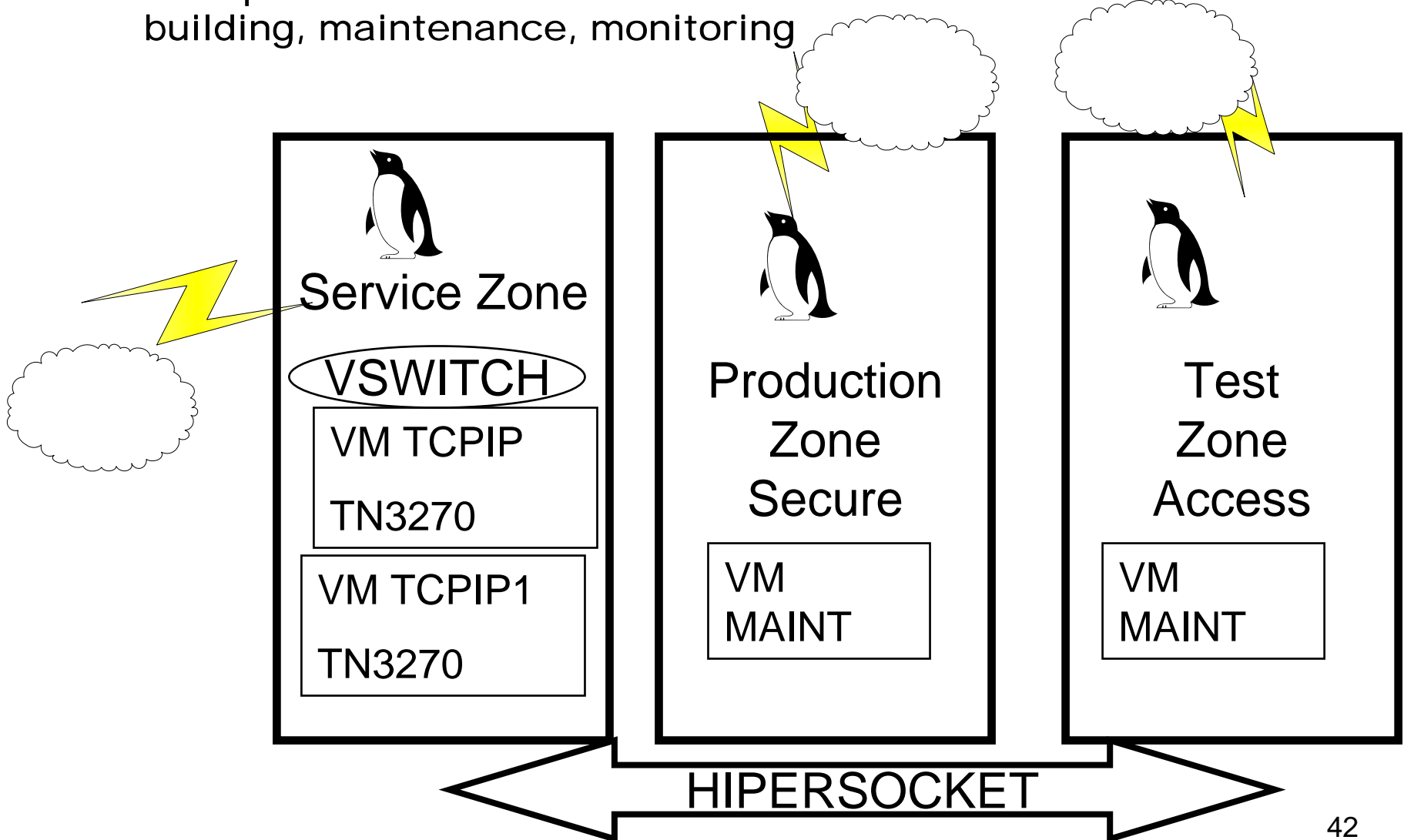
z/VM LPAR with Linux Oracle Servers.
Data is on FCP SAN. Vswitch network (built in redundancy)



z/VM LPARs with vswitches, vlans and hipersocket. Shared OSAs and ORACLE RAC on the SAN. Creates a nice maintenance window method.



Best practice: Service Zone LPAR for z/VM and Linux building, maintenance, monitoring



Oracle HA Scenarios: Multiple CECs, OSAs, LPARs, Hipersockets and Vswitches !

