LE CENTRE DE SERVICES PARTAGÉS

Building a strong z/VM and Linux on the mainframe architecture

SHARE User Events

Tampa , Florida

Session 9231

February 14th, 2007

In collaboration with

IBM Canada LTD

VM Resources LTD

Centre
de services partagés
Québec

Table Of Contents

- Client Context
  - The DGTIC
  - Environment
- Architecture
  - Guaranteed isolation of multiple clients
    - Taking advantage of the System z, LPARs
    - Networking within the box
  - Security and data integrity
    - RACF, Hardening
    - Networking with the real world
- Best practices
  - Networking
  - System
  - Lessons learned

Centre
de services partagés
Québec

Client context

... DGTIC ...

Centre
de services partagés
Québec

Client context
The DGTIC

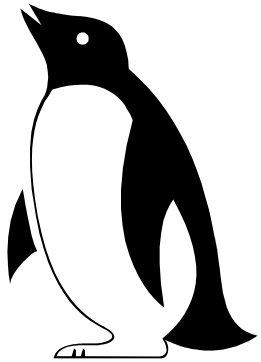
- IT service provider for many Québec government offices (125)
  - Already a mainframe shop
  - 5 z/890 + 1 z/800 + 1 G5 on the floor on 3 sites
  - 1 z9-EC dedicated to Linux on z/VM
  - 450+ physical servers (750+ logical) (HP, SUN, pSeries, …)

- DGTIC orientations :
  - Promote the mainframe environment
  - z/VM is the prime choice for future projects
  - Server consolidation is a priority
  - This project is in line with the new « online government » policy

Centre
de services partagés
Québec

Client context
Environment

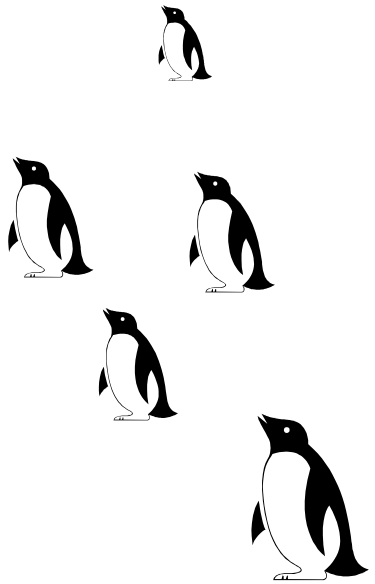
- 1 z9-EC mainframe with 5 IFLs (~ 3000 mips)
- 5 LPARs
  - Oracle/DB
  - WAS
  - TAM & LDAP
  - Service Zone
  - Lab Zone
- 40 internal networks
- Software
  - SuSE Linux (versions 8 & 9)
  - z/VM v.5.2 +
  - Oracle/DB (versions 9i & 10g)
  - Velocity Software Performance Tools
  - CA products (Automation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



Centre
de services partagés
Québec

Client context
Environment

- Oracle/DB – Migration Project Status
  - Golden images
  - 165 Oracle instances with 125 Linux virtual machines
  - Growth of over 100 new instances planned per year for the next few years
  - 25 instances in production as part of the government portal
  - For the first migrations (~ 60), on average
    - 1 migration per day (20-25 databases per month)
- Our current challenge is to synchronize the migrations with date restrictions imposed by our external clients



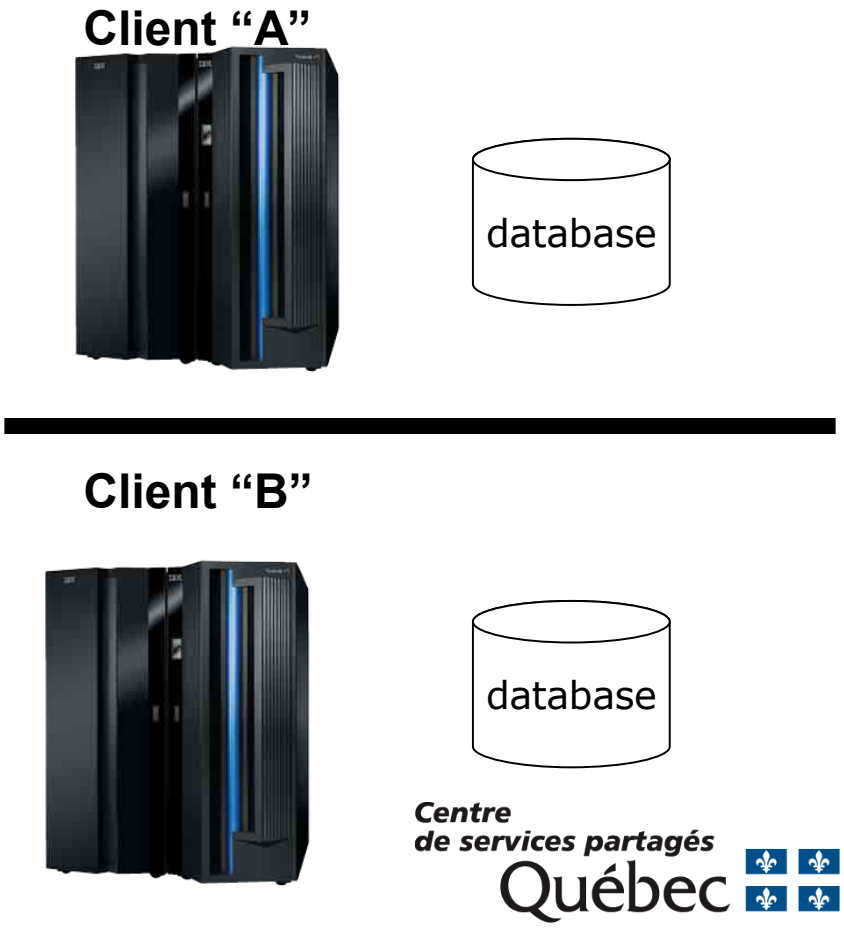
Centre de services partagés Québec

Architecture

Centre
de services partagés
Québec

Architecture
Guaranteed isolation of multiple clients

- The DGTIC serves the needs of over 125 clients, some large, some small.
- Clients must have their applications and data separated from each other.
- The challenge is how to do this in a centralized shared environment.
- System z hardware with z/VM leads the way!



Architecture
Taking advantage of System z hardware and z/VM

- Old school IBM mainframe virtualization
  - *35+ years of storage and CPU sharing with integrity and isolation*
- LPARs
  - *Hardware partitioning*
- Networks within the box
  - *Hipersockets, guest lans, and Vswitch*
- Securing resources with RACF

1970's

1980's

2000's

Centre de services partagés Québec

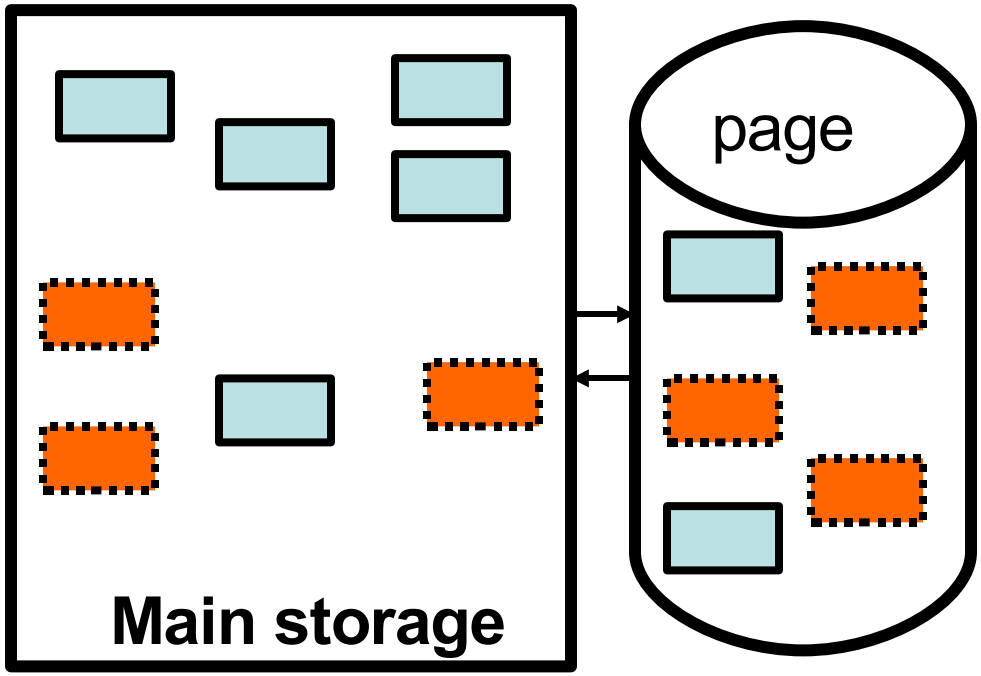
Architecture
Old school IBM mainframe virtualization
Memory

VMA pages →

VMB pages →

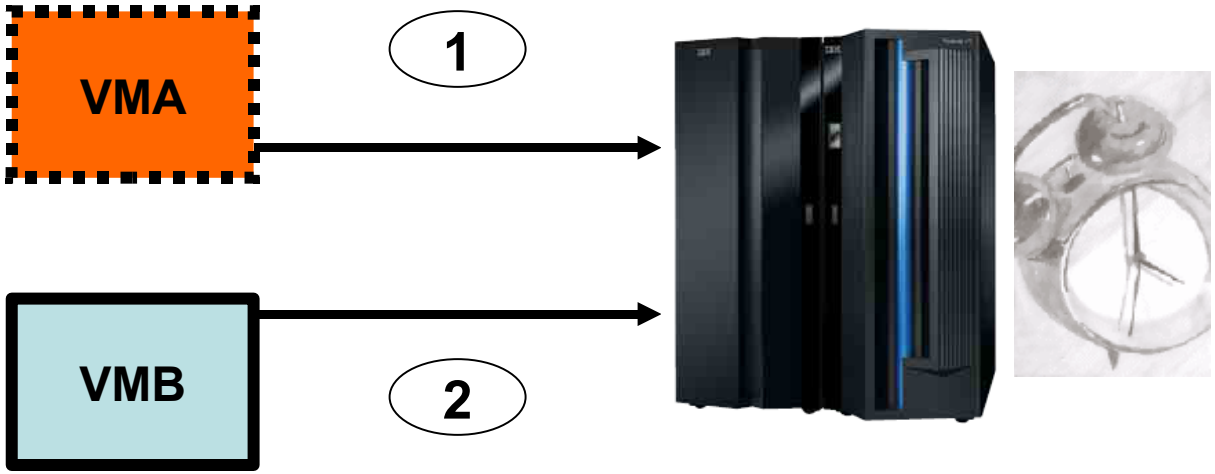
Virtual storage isolates the memory of address spaces (virtual machine memory). Memory is isolated in both main and auxiliary storage devices, and controlled by CP.

Two virtual machines, VMA and VMB are shown.



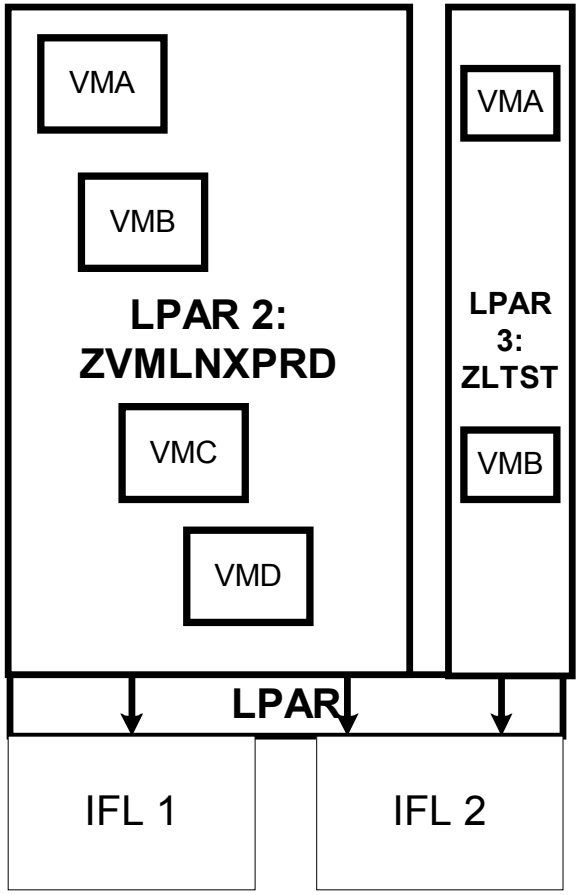
Centre de services partagés Québec

Architecture
Old school IBM mainframe virtualization CPU



The CPU(s) are time shared. The CP scheduler and dispatcher subsystem organize the virtual machine work. Control of the CPU is given for milliseconds. In the example VMA gets cycles before VMB.

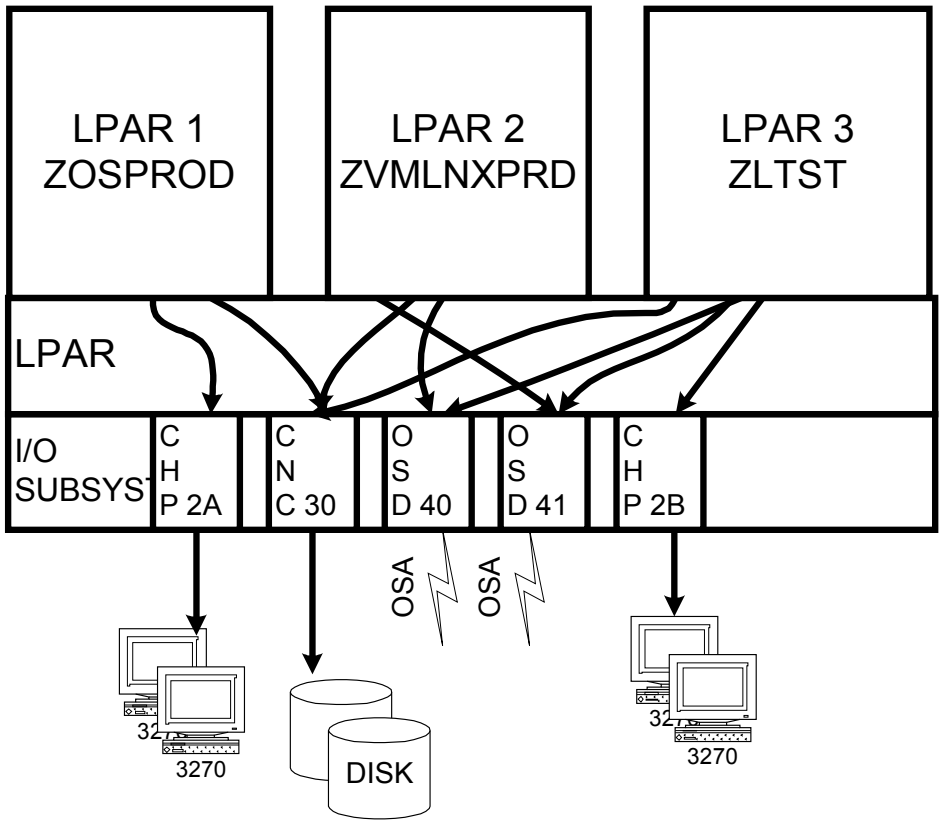
Architecture
LPARs



LPAR and z/VM provide the most virtualization of any platform. Logical partitioning of the physical resources with PR/SM (up to 30 LPs) and software partitioning with z/VM (thousands of virtual machines).

Centre
de services partagés
Québec

Architecture
Logical Partitioning



- In the System z platform the hardware resources are logically distributed among multiple control programs.
- Each instance runs simultaneously and independently.
- The set of resources available to each control program is called a logical partition (LP). Resources are logical CPUs, storage, and I/O.

Centre de services partagés Québec

Architecture
Networks within the box

- A variety of choices for networks that keep the clients isolated
  - **OSA devices**

  Traditional connectivity from mainframe to physical switches
  - **Hipersockets**

  Inter and Intra LPAR connectivity
  - **Guest lans**

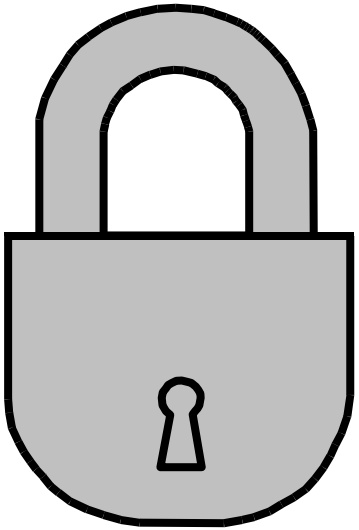
  Connect virtual machines on virtual networks within an LPAR
  - **Vswitches**

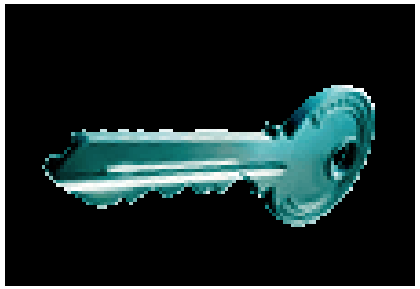
  *Connect **guest lans** to physical switches using **OSA devices***

Centre
de services partagés
Québec

Architecture
RACF Serves and Protects

- RACF ensures isolation as it provides security for these protected resources and events in z/VM:
  - Logon
  - Link
  - Vswitch
  - Vlan
  - Shared File system
  - VM FTP





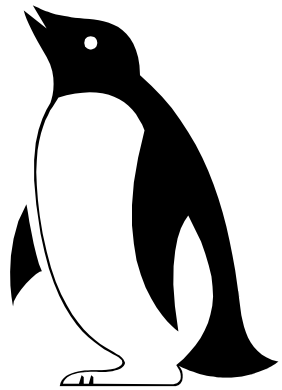
Centre
de services partagés
Québec

Architecture
We harden our Linux on the mainframe servers

- The Linux golden images are hardened, tested and certified by an independent team before allowing the image to be cloned.

Hardening tasks:

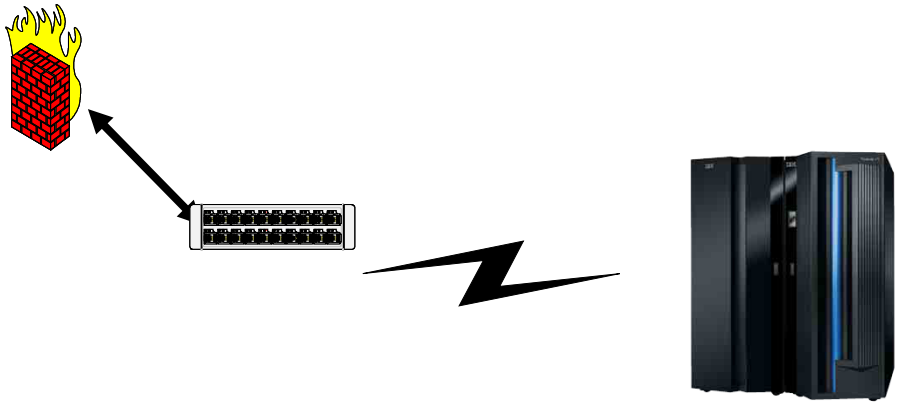
- Removing unneeded login accounts
- Removing many supplied services such as FTP, Telnet, and NFS.
- Sifting through the startup /etc/rc.d tasks and removing unneeded tasks.
- Using PAM authentication with strong password practices.
- Using Tripwire to inventory software and for file anomaly detection.
- Ethical hacking done on a regular basis for penetration testing and cracking.
    - Certified by an independent team.



Centre
de services partagés
Québec

Architecture
Networking with the real world

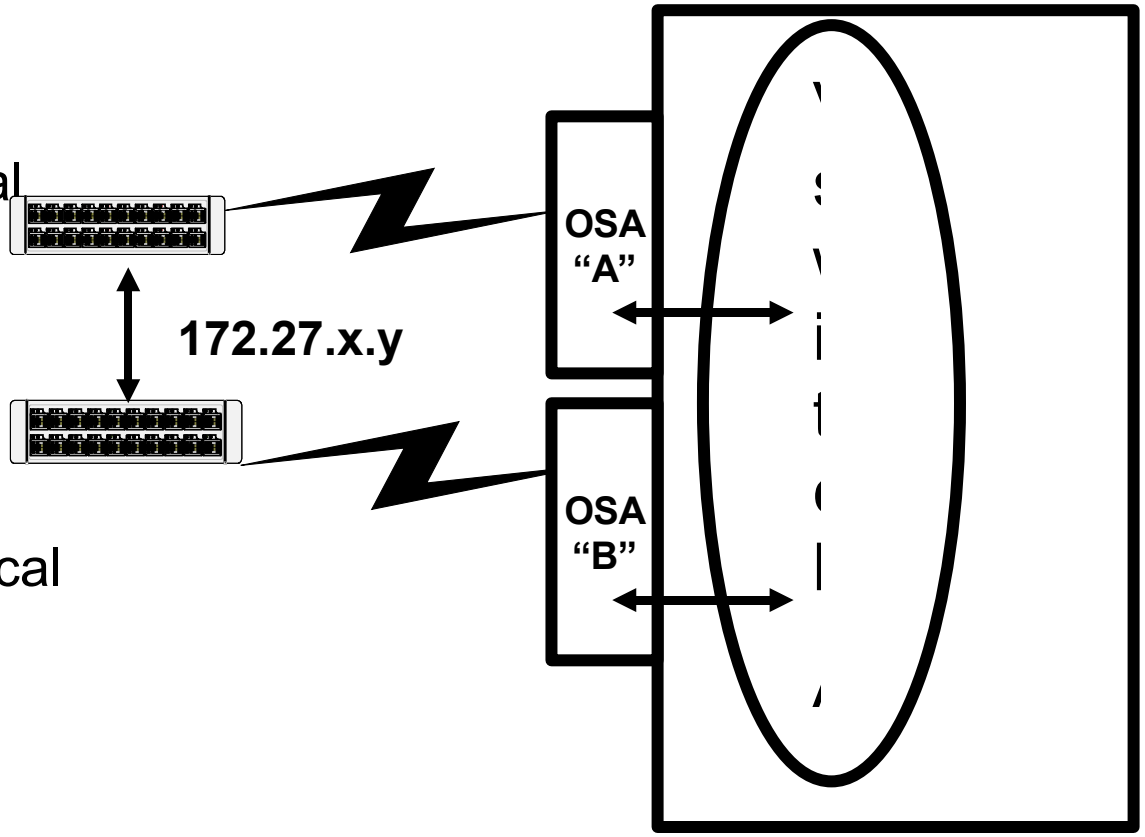
- We use a lot of Vswitch networks.
  - Over 40 ….
- Vswitch connects to OSA port as conforms to the physical network topology.
- Redundancy provided only for production networks.
  - Handled within Vswitch connecting to multiple unique OSA ports.
  - Does not require VIPA
- Some OSA ports shared across zones in multiple LPARs.
- Firewalling done downstream from the mainframe.



Centre
de services partagés
Québec

Architecture
Production Network Redundancy

- 15 production networks have redundancy with dual OSA ports.
  - *All others (25+) do not have redundant networking*
- Managed by Vswitch.
- Connect to different physical switches.
- Switches are bridged.



Centre de services partagés Québec

Best Practices

Centre de services partagés Québec

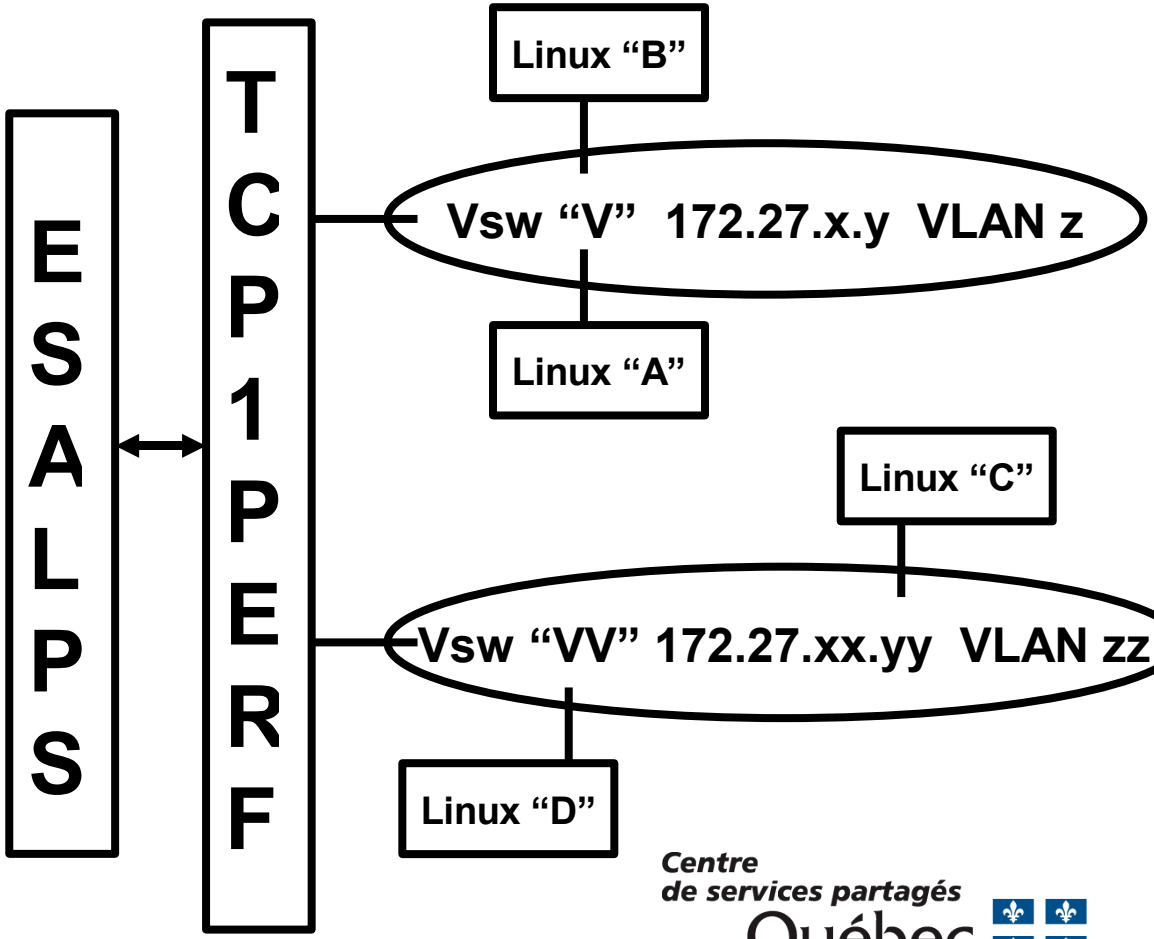
Best practices
At the DGTIC

- In our project we planned to utilize best practices for systems and network management.
- Examples of in use best practices:
  - Networking:
    - Performance data collection using private Vswitches
    - Manage multiple networks from a single TCPMAINT
  - Systems:
    - Golden images (z/VM & Linux)
    - Cloning engine
    - Sharing resources the DGTIC way
    - Sharing resources the IBM way

Centre
de services partagés
Québec

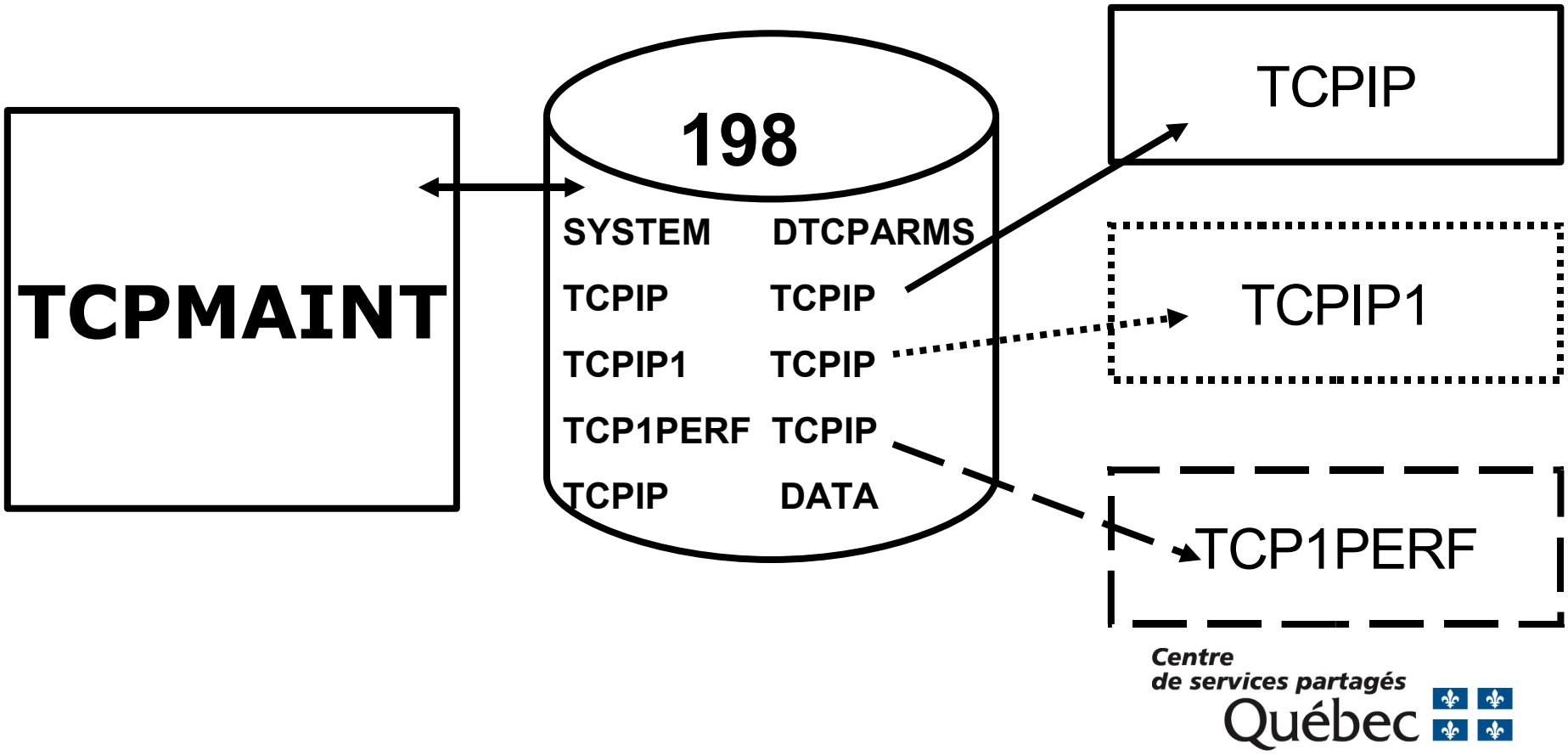
Best Practices
Performance data collection using private Vswitches

- A TCPIP stack with multiple guest lans and vlans collects data for the Velocity SNMP data collection.
- The Vswitches are defined without real devices.
- Membership in the Vswitch and vlan is RACF protected.



Centre de services partagés Québec

Best Practices
Administering multiple z/VM TCPIP machines from a single TCPMAINT

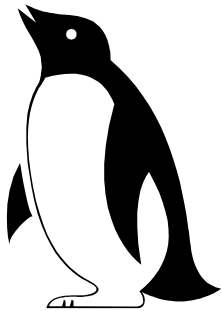


Centre de services partagés Québec

Best Practices
Golden images : z/VM & Linux

- Our z/VM golden image:
  - z/VM 5.2.0 RSU 0602+ reach ahead service
  - All production mdisks on one volume per system
  - Goal is to service from one system
  - One flavor

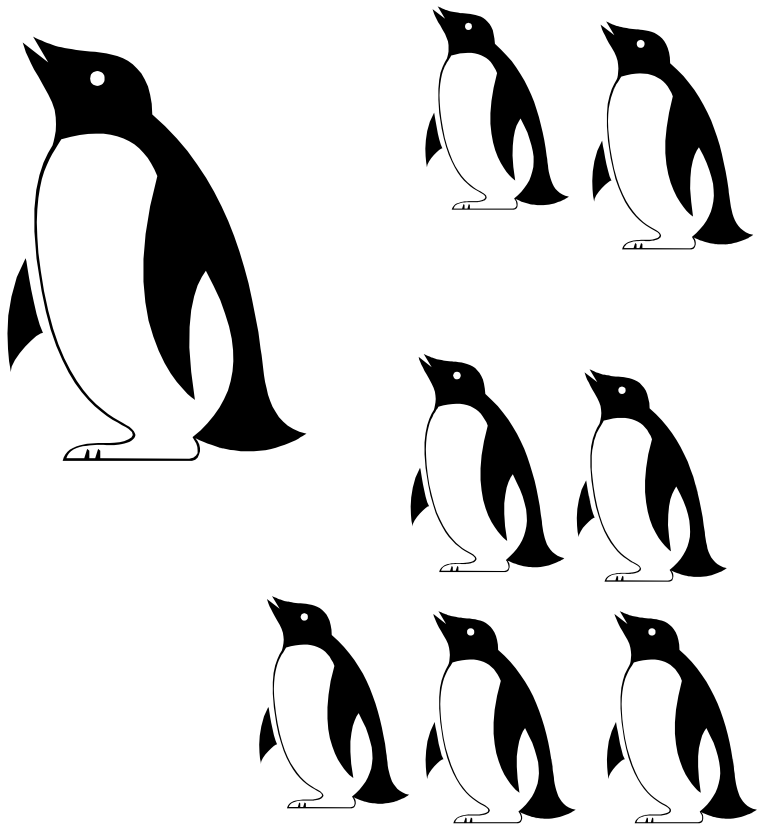
  Our Linux on the mainframe golden images:
  - SuSE SLES 8 or 9 (evaluating v.10)
  - Service pack 2 or 3
  - Hardened
  - One application flavor (Oracle or WAS or TAM/LDAP)
  - Input to the cloner
- Both are rigorously tested and certified



Centre
de services partagés
Québec

Best Practices
The Linux Golden Image
"*install once and clone often*"

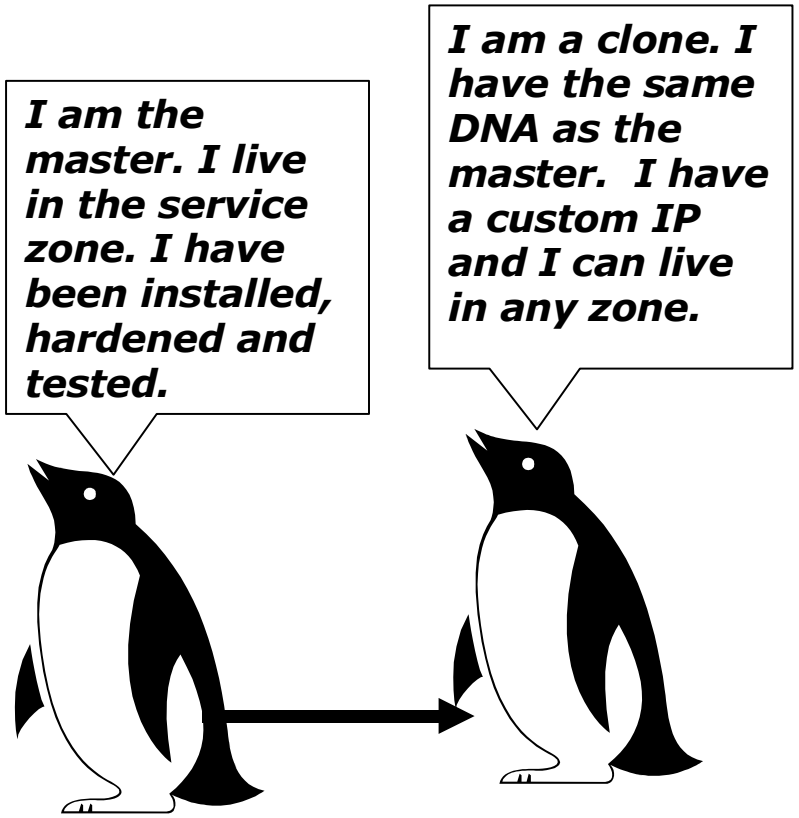
- The golden image is really black and white and waddles on ice but not until:
  - Installed
  - Serviced
  - Hardened
  - Tested by various groups
  - Passes security penetration tests and certification
- There are a few masters and many many many clones!



Centre
de services partagés
Québec

Best Practice
Our cloner : Overview

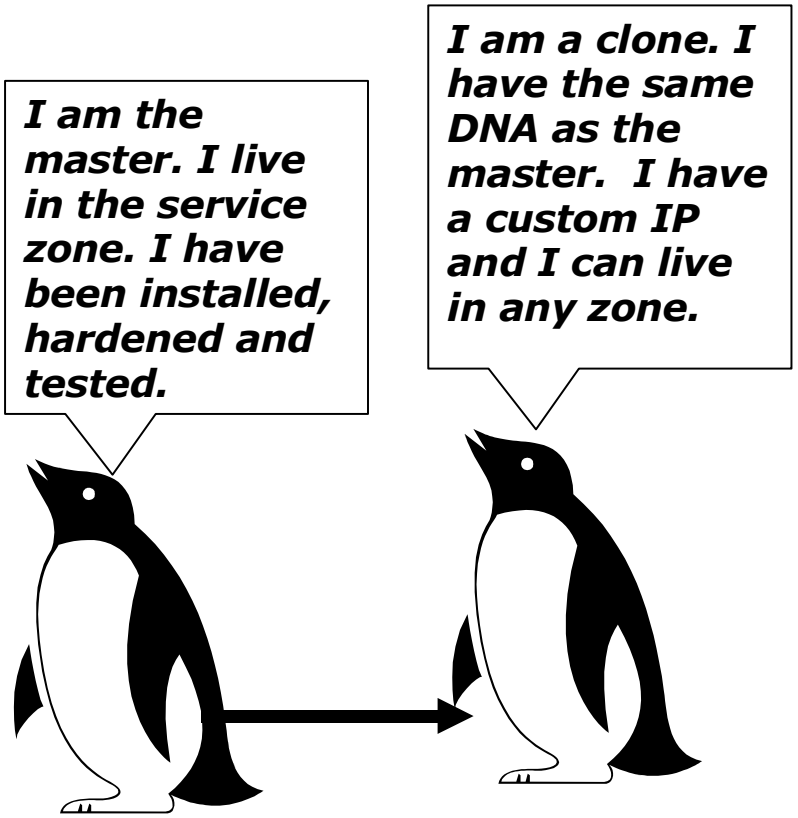
- Hand crafted
- Pride of ownership
- Not a disk copier
- Intelligent decisions:
  - Choice of Linux
  - Choice of application
  - System and application position
  - Vswitch membership
  - Vlan membership
  - IP address
  - Data replicated
  - Strong passwords



Centre de services partagés Québec

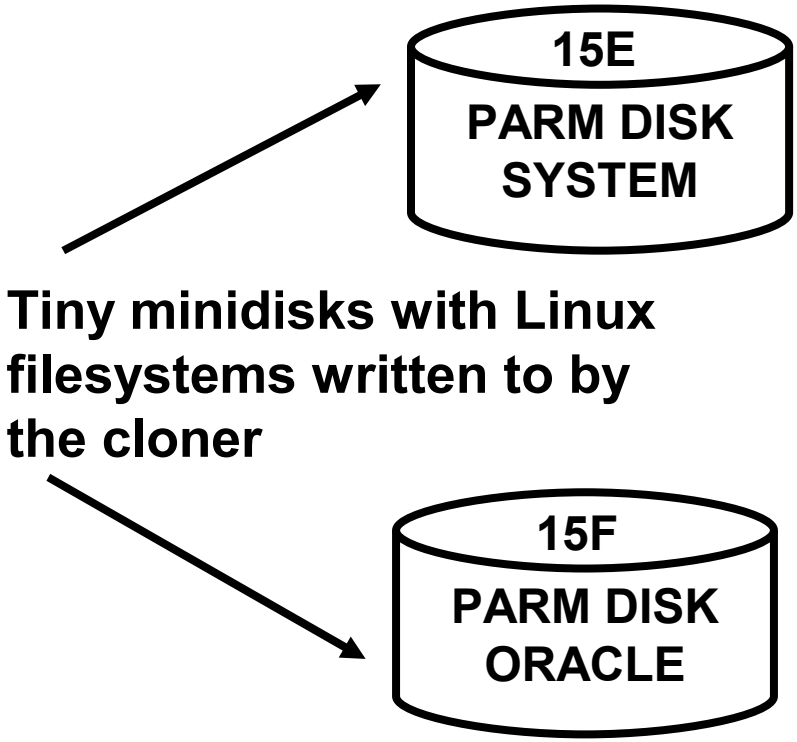
Best Practices
Our cloner: Coding and interfaces

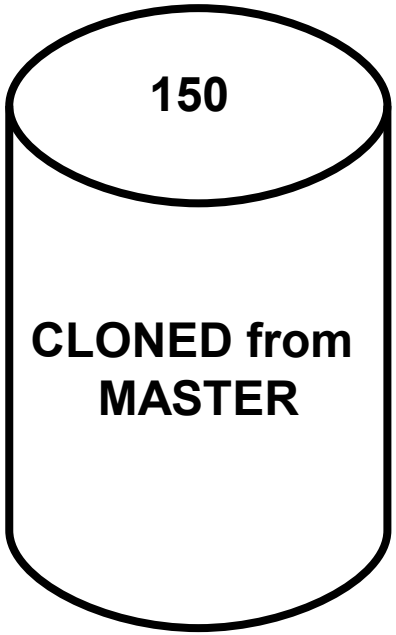
- Coded in REXX and PIPELINEs.
- Interfaces to DIRMAINT and RACF.
- Inputs include which system, application, storage size, etc.
- Interfaces with 3270.
- Can clone only from service zone to any other zone.

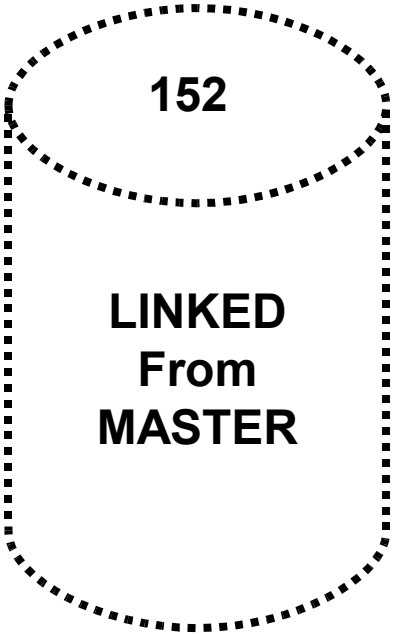


Centre de services partagés Québec

Best Practices
Disk layout for the Cloned Linuxen

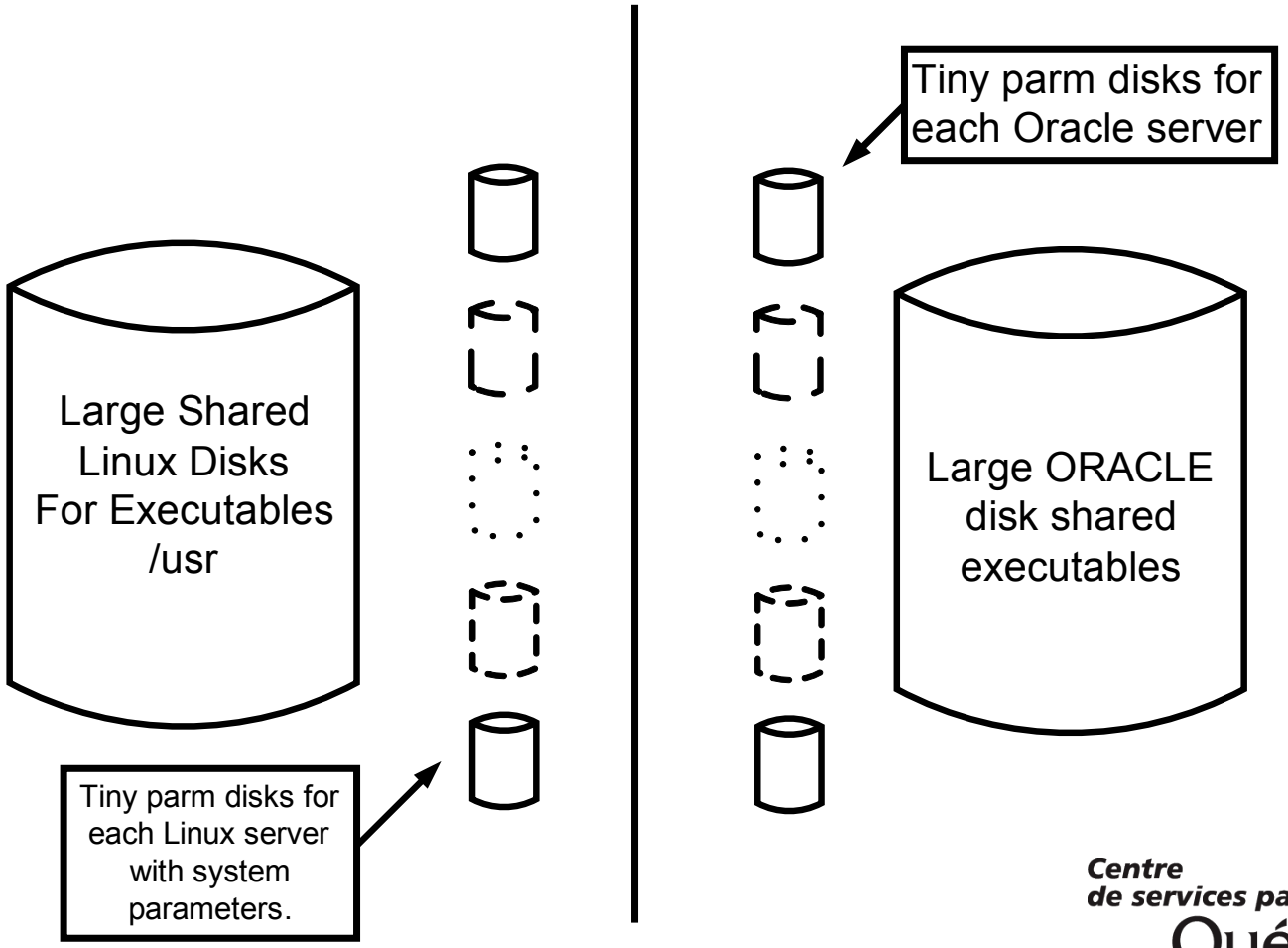






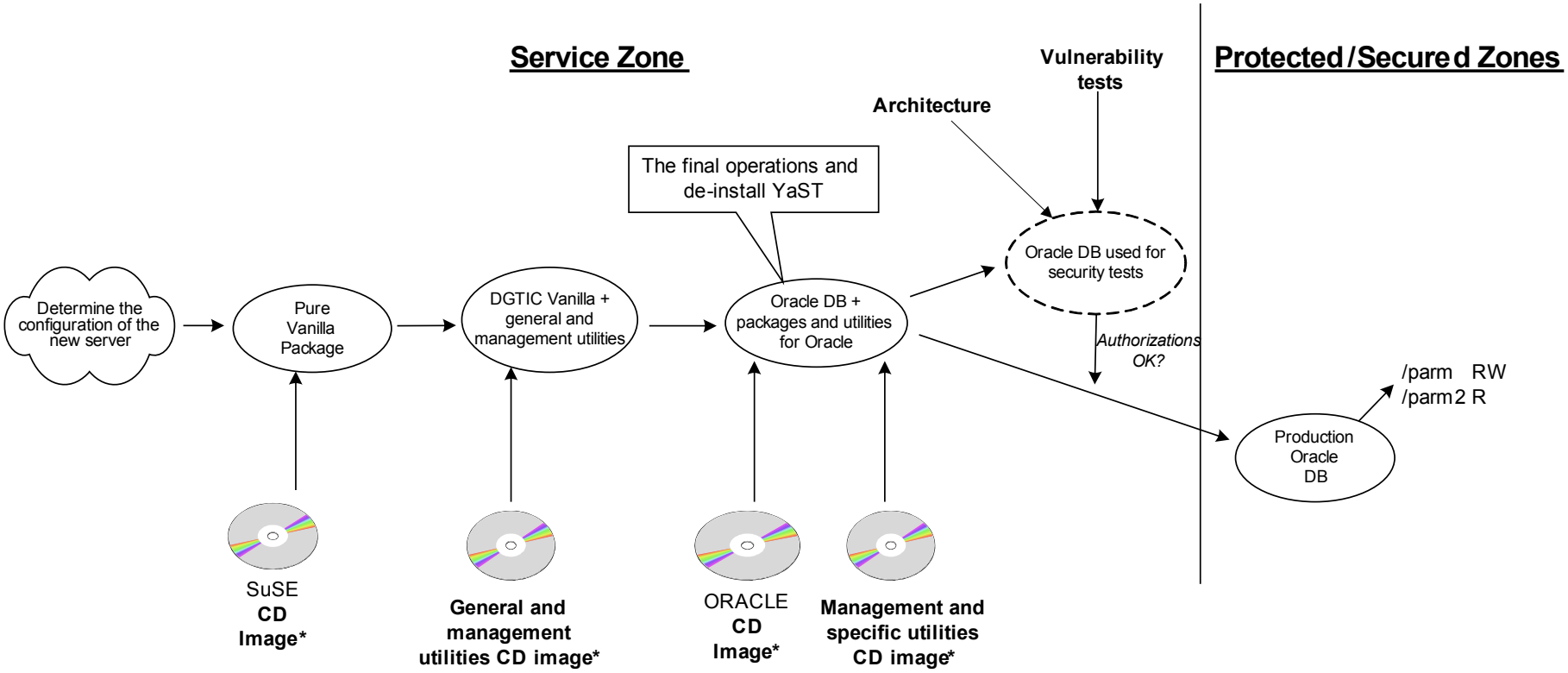
Centre de services partagés Québec

Best Practices
The big picture of the cloning



Centre de services partagés Québec

Best Practices
The cloner at the DGTIC

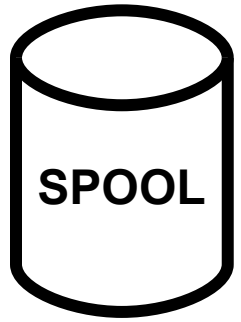


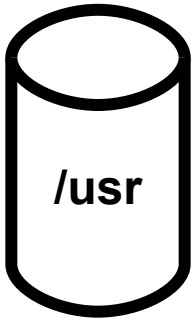
* : These images reside on a virtual Linux server in the service zone for access via FTP. This server contains software libraries.

Centre de services partagés Québec

Best Practices
Resource sharing at the DGTIC

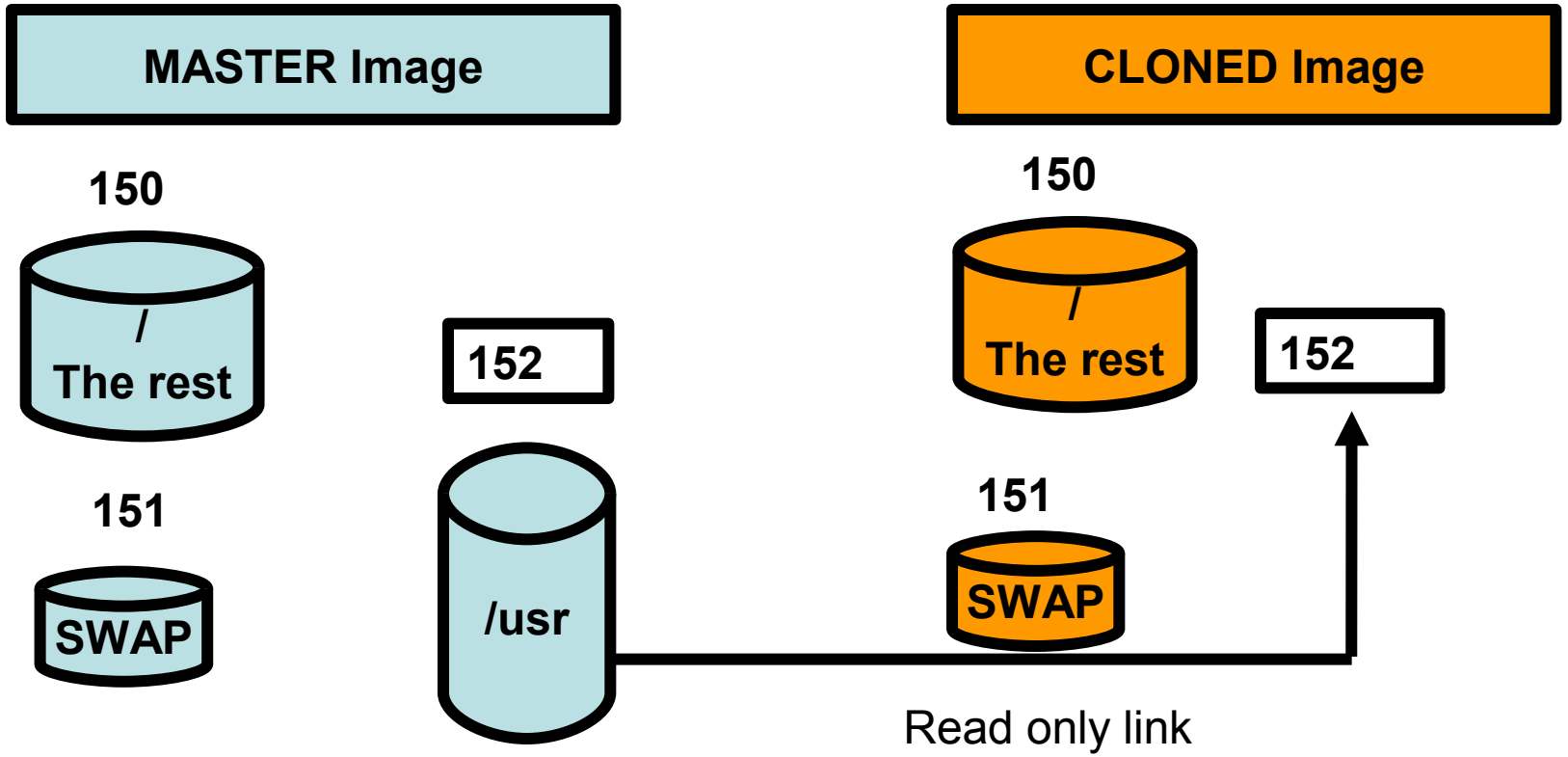
- IBM way (old school – 35+ years)
  - CPU
  - Memory
  - Minidisk i/o
  - Spooling
- DGTIC sharing:
  - Linux file systems (/usr)
  - Heavy usage of Vswitch
    - Lots of guest lans
    - Many OSA ports





Centre
de services partagés
Québec

Best Practices
Sharing /usr with the master



Centre
de services partagés
Québec

Best Practices
Resource sharing
HiperSockets network on the z9-EC

HiperSockets 192.168.150.x

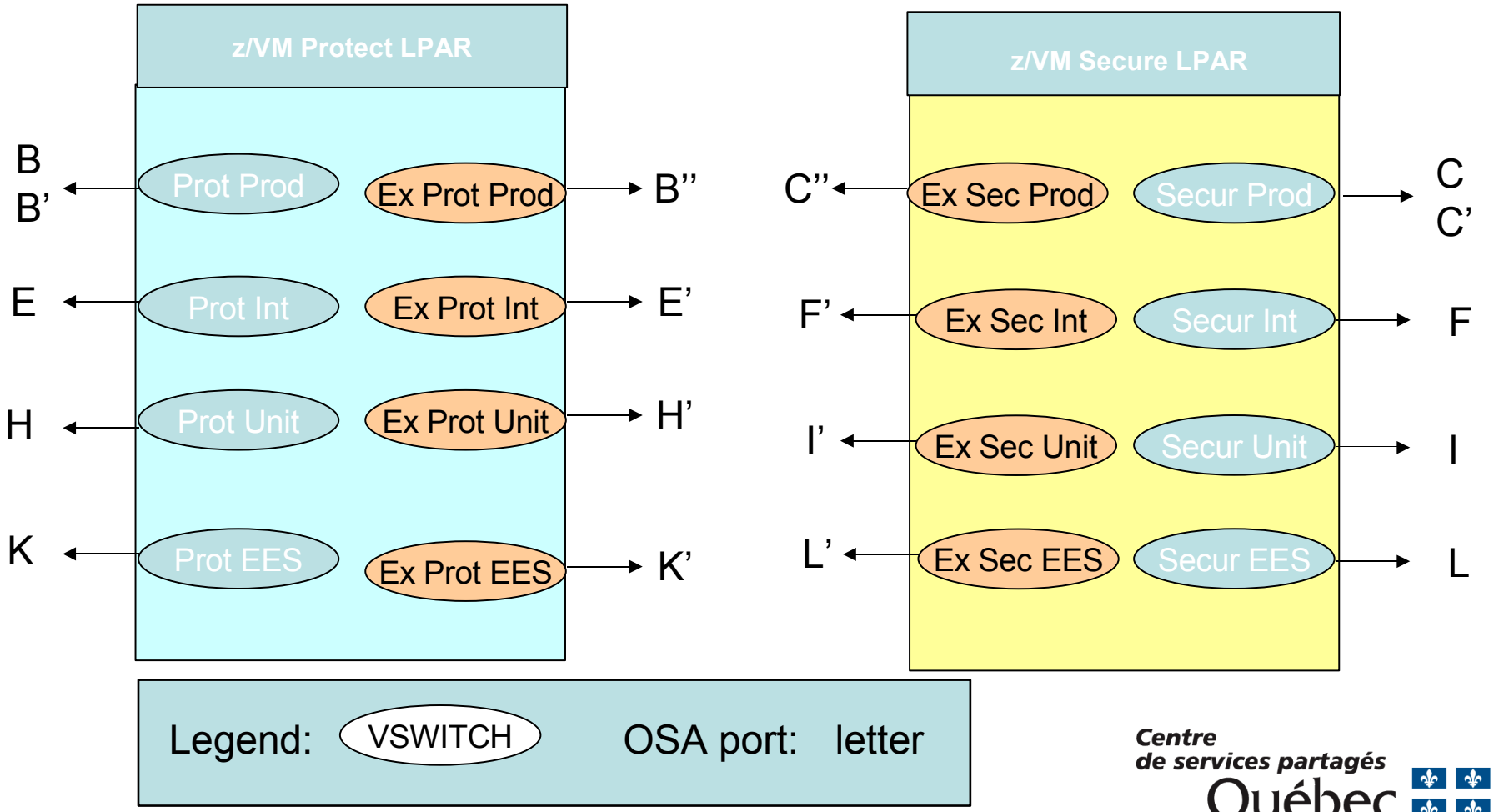
Chpid “BF”

telnet *rscs*	*telnet* *rscs*	*telnet* *rscs*	*telnet* *rscs*
z/VM protect LPAR	**z/VM secure LPAR**	**z/VM service LPAR**	**z/VM other LPAR**

- **Internal network only.**
- **Used for administrative purposes.**
- **Applications include the cloner, telnet, RSCS (file transfer and message queues).**
- **Secure memory-to-memory transfer.**

Centre de services partagés Québec

Best Practices
Resource sharing
Vswitch usage at the DGTIC



Centre de services partagés Québec

Best Practices
Lessons learned: Volume 1

- Acceptance of virtual servers quicker than expected.
  - Grew to 100+ Oracle servers ahead of plan.
- Fully tasked personnel (big shoulders):
  - Confirmed our expectation that 2 Linux administrators can support all virtual Linux servers.
    - *100:1 ratio of Linux virtual machines to administrator*
  - 2 z/VM systems programmers supporting 5 LPARs: (could support many more)
    - *New to z/VM*
    - *Mentored by consultant*
    - *z/VM support will be integrated into MVS group by year end 2007*
- Less than fully tasked personnel (arms and legs):
  - *Security administrator*
  - *Network programming*
  - *Storage*
  - *Automation*
  - *Performance*

Centre
de services partagés
Québec

Best Practices
Lessons learned: Volume 2

- Big win early win with successful disaster recovery.
- Administration and reporting on centralized servers is excellent.
- Lots of new documentation and procedures integral part of project.
- Lots of training required.
- Require a lot of z/VM paging space.
  - Core memory of 32 gigabytes, 2 gigabytes of expanded storage, and 72 gigabytes of DASD paging space.

Centre
de services partagés
Québec

Best Practices
Lessons learned: Volume 3

- Critical mass of servers required – use more than 1 Linux virtual machine for benchmark, POCs, and business case!
- Initially, project was done for the $ savings, now the important gains:
  1. The flexibility of the solution
  2. Disaster recovery
  3. $ savings
- You must have a sponsor. Our sponsor was the operations directorate for the mainframe business interested in solving DR issues.

Centre
de services partagés
Québec

Best Practices
Oracle Lessons Learned

- Mostly business as usual for the DBAs:
  - Use SSH client or “X” windows (no 3270 usage)
  - DBAs comment on rapid performance of I/O
  - DB loading faster than in other platforms.
- Benign ignorance of the virtual machine
  - Linux administration performed by Linux sysadmin.
  - z/VM administration performed by VM sysprogs.

- Rapid creation of new databases in virtual machines for testing, acceptance, and production.
- Initial install was difficult but once incorporated into cloning methods subsequent installs quick and easy.
- Almost all client needs satisfied with ORACLE cloned image (they don’t know).
  - ~ 2% require some sort of customizing.

Centre
de services partagés
Québec

Questions ?

For more information :

Karen-Ann Plourde
karen-ann.plourde@cspq.gouv.qc.ca

Jocelyn Hamel
jhamel@ca.ibm.com

David Kreuter
dkreuter@vm-resources.com

Centre
de services partagés
Québec

www.gouv.qc.ca