Millennium Systems on the Cloud:
Experience on System Migration from SUN V490 to VMWare Platform

Li Yiu On, Systems Section, Library
Tsang Jimmy, Office of Information Technology
Hong Kong Baptist University

Outline

A. Introduction – What is Cloud Computing

B. Building Cloud Computing Infrastructure

C. Millennium Migration to New Cloud

D. Conclusion

A. Introduction

1. What is Cloud Computing

- In June 2011, a study conducted by VersionOne found that 91% of senior IT professionals actually don't know what cloud computing is [1]

- Cloud computing may be defined as a new technology that allows a computer system running on a virtual machine (VM, or virtual server) which can be access on internet

- Traditional Physical Server-Based Model vs Cloud Computing Model

[1] Cloud Computing, http://www.wikinvest.com/concept/Cloud_Computing

2a. Traditional Sever-based Model

- Millennium system consists of two main components:
  - Hardware including
    - a dedicated physical server -- entire server is dedicated to running the Millennium software.
    - tape drive, UPS (Uninterrupted Power Supply)
  - Software (computer program developed by III)

2b. Traditional Sever-based Model

- We need to **buy** a new server for running Millennium system (software) every 3 to 5 years
  - A hardware, just like a human physical body, is getting old every day.
  - It is not secure to run a system on a obsolete server
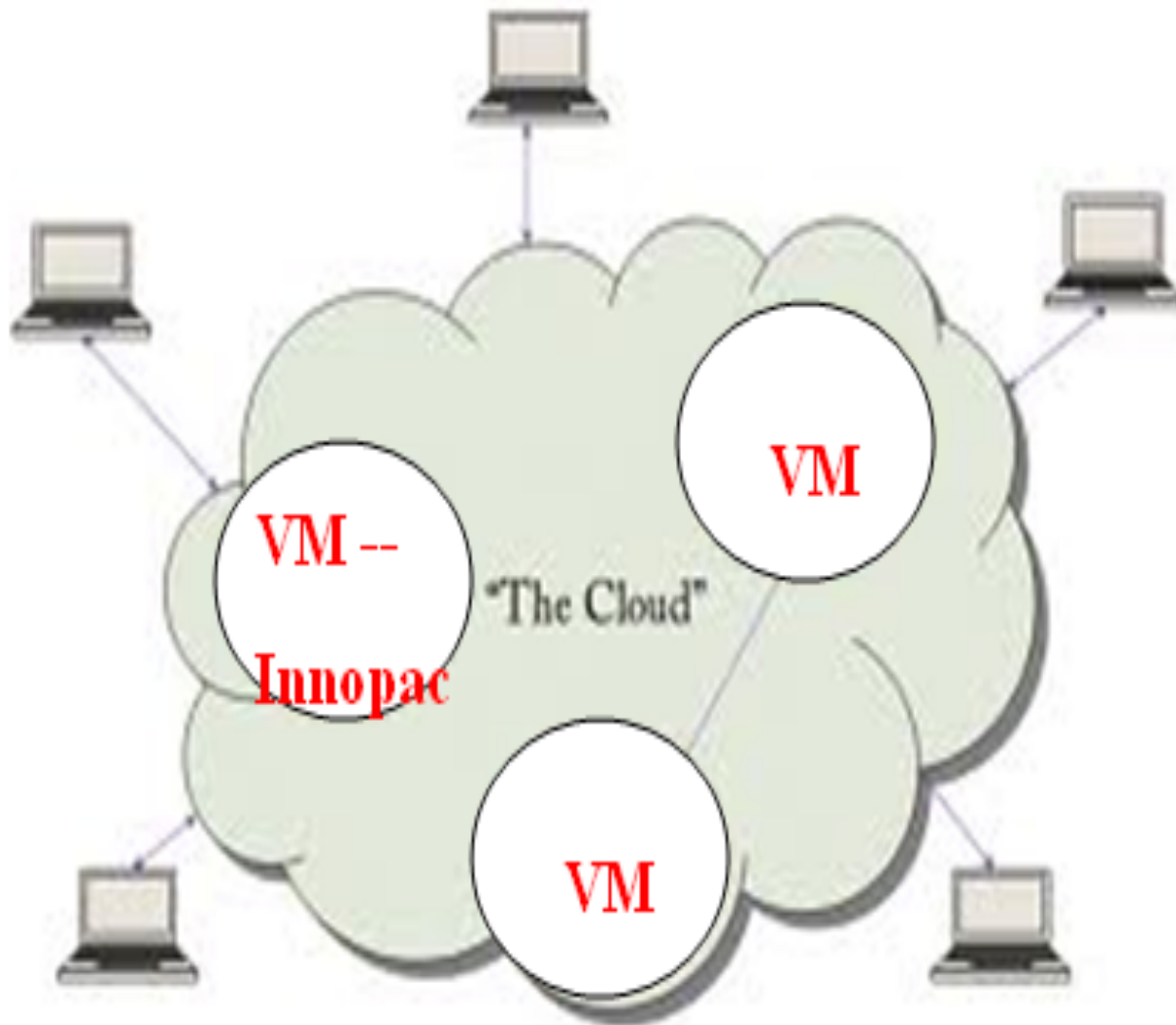
3. Heavy Hardware Costs

- **New Hardware Cost is High**: To change a server regularly means we need extra money to acquire a new server to replace the obsolete one
  - a dedicated high end server may cost over HK $200,000
- **Maintenance Costs is also High**:
  - Disaster recovery (usually require to acquire an additional server for backup)
  - Daily backup: labor cost to change tapes every day

4a. Virtualization

- We only want to use the Millennium system (*the software*) but do not want to buy/maintain the server, how can we do that?
- Cloud computing allows systems librarians to maintain a integrated library system like Millennium *without* installing a hardware (i.e. physical server)
- Virtualization -- Instead of requiring a separate physical server for each system, dozens of virtual servers (Virtual Machines) can co-reside on the same computer.

- **HKBU Millennium System running on a dedicated physical server SUN Fire V490 before migrating to a cloud platform on Oct 2011**
- **Last photo**
  - with a physical body
  - exact location

- Millennium System on the cloud
  - without a physical body
  - without an exact physical location

5a. Advantages of putting Millennium on the Cloud

- As hardware is virtualized (VM) in the new cloud computing technology, one of the immediate benefits are money saving on hardware investment:
  - End of endless hardware purchases and upgrades
  - No need to purchase dedicated UPS (Uninterrupted Power Supply)
  - No need to purchase additional backup server for disaster recovery

5b. Advantages of putting Millennium on the Cloud

- **Save Energy**
  - Run ONE physical server instead of several to host multiple VMs
  - Less air conditioning usage to keep the environment intact
  - Less power backup resources required

5c. Advantages of putting Millennium on the Cloud

- Save Time
  - Reduced IT overhead: with fewer servers, you can spend less time on the manual tasks required for server maintenance
  - Safer, faster backups and restore
  - Ability to take live snapshots while OS is running
  - Can schedule snapshots of OS
  - Less time to recover the whole OS
  - It's also much faster to deploy a virtual machine than it is to deploy a new physical server.

5c. Advantages of putting Millennium on the Cloud

- ● More Secure
  - – Running Millennium on a standalone physical server is, indeed, without any fault tolerance
  - – Even a minor hardware fault may take days to recover the services (we did have a terrible experience before, our Millennium system was down several days simply because of a server hard disk problem)
  - – In case of serious disaster, the process to recover from backup could take week, as well as, losing data since last backup

5c. Advantages of putting Millennium on the Cloud

- **More Secure (con't)**
  - Running Millennium on a VM can
  - Decrease downtime and improve reliability with business continuity and disaster recovery
  - Proactively migrate virtual machines away from failing or under performing servers

B. Building Cloud Computing Infrastructure

1a. Two Models to build Cloud Computing Platform

- In Oct 2010, the maintenance contract for the SUN Fire V490, which our Millennium system was running for over 5 years, would be expired

- Given the benefits of cloud computing mentioned above, it did not justify to purchase a new standalone physical server for replacement

- All were agreed to go for cloud computing

- Questions: who should be responsible for implementing and maintaining cloud computing?

1b. Two Models to build Cloud Computing Platform

- Two models:

  A. Library Systems Section Do-it-Ourselves

  B. Outsourcing
    - Commercial
    - HKBU ITO (Office of Information Technology)

2a. Cloud Computing Costs

- Many programs are available in the market to transform an IT infrastructure into a cloud. One of the sophisticated programs is VMWare vSphere

- In HKBU Library, Systems Section have already build a VMWare platform concurrently hosting 9 virtual machines of which

  - over 24 library databases (including full-text, image and video streaming databases),

  - numerous commercial library system packages (e.g. ILLiad) and

  - in-house computer developed programs since 2009

2b. Cloud Computing Costs

- We would say it is not difficult and cost-affordable to use VMWare to build our own private cloud platform inside the Library, but it must be emphasized that this platform only aims to host comparatively the non-critical library systems

- For the provision of hosting services to a large and critical library system like Millennium is another story

2c. Cloud Computing Costs

Use VMware's words:

"VMware vSphere is comprised of a ***number of features*** that transform industry standard hardware into a shared, mainframe-like resilient environment with built-in service level controls for all applications."*

- "A number of features" include:

*http://www.vmware.com/products/vsphere/mid-size-and-enterprise-business/overview.html

VMware vSphere

Application Services

Availability
vMotion
Storage vMotion
HA
Fault Tolerance
Data Recovery

Security
vShield Zones
VMSafe

Scalablity
DRS
Hot Add

Infrastructure Services

Compute
ESXi
DRS and DPM
Memory
Overcommit
Auto Deploy

Storage
VMFS
Thin Provisioning
Storage I/O Control
Storage DRS
Profile-Driven Storage

Network
Distributed Switch
Network I/O Control

2d. Cloud Computing Costs

- ***Basic features of VMware***
  - vSphere ESXi
  - vCentre
- ***Due to limited resources, we do not turn on***
  - vSphere Fault Tolerance (FT)
  - vSphere motion
  - vSphere DRS
  - vSphere High Availability (HA)
  - vSphere Fault Tolerance (FT)
  - vSphere Distributed Switch
  - NEW: vSphere Storage DRS
  - NEW: vSphere Profile-Driven Storage

2f. Cloud Computing Costs

- Implementation and operation cost of VMware at full scale is *Very High*
    - Initial hardware/software acquisition and setup cost, as well as the re-current maintenance expenditure are very expensive
    - A full team of systems expertise and personnel needed to master such complex infrastructure, also additional man power, and training; and provision for staff turn-over

3. Core business of Library

- Clearly, IT infrastructure is not the core business of the Library

- With such a small 3 team members at the Systems Section, we do not have adequate human and hardware resources to do it ourselves

- Hardware, software and staff costs are too high for the Library to build such an advance platform alone

- Thus, the Library Systems Section Do-it-Ourselves is not considered

4. Disadvantages of Commercial Outsourcing

- Privacy and Security Concerns
- It means transmitting sensitive data to a third-party commercial vendor over internet
  - Patron records containing personal address, phone number, and library expiry date
  - Patron check-out records and borrowing history
  - Book order records containing book price information
- Too risky → not suggested

5a. Cloud Infrastructure maintained by ITO

- In 2010, ITO create a new cloud infrastructure which aims to provide easy, scalable access to computing resources and IT services for all BU departments/offices

- ITO Cloud Computing → ITO is like the landlord of a big factory building where have many well-equipped offices/rooms for leasing

- All BU Departments/offices may pay a annual subscription fee to ITO for renting the cloud computing services

5b. Cloud Infrastructure maintained by ITO

- Best model for Library: Library Millennium becomes a tenant of ITO cloud computing
- Do not have Privacy and Security Concerns of transmitting data outside campus
- Systems Section are longer required to take care of hardware (UPS and data backup) and discovery recovery
- Of course, No free lunch. The Library still has paid a subscription fee to ITO, but the paid amount should be the *lowest*, comparing with
  - to buy a new dedicated physical server, or
  - to implement and maintain a private cloud by the Library, or
  - to outsource a commercial cloud company

C. Millennium Migration to New Cloud

1. Configuration of Millennium Virtual Server

- Resources (Virtualization)
  - 4 CPU(s)
  - 16 GB Memory
  - 150 GB Disk Space

2. New OS

- SUN V490 OS is Unix (Solaris)
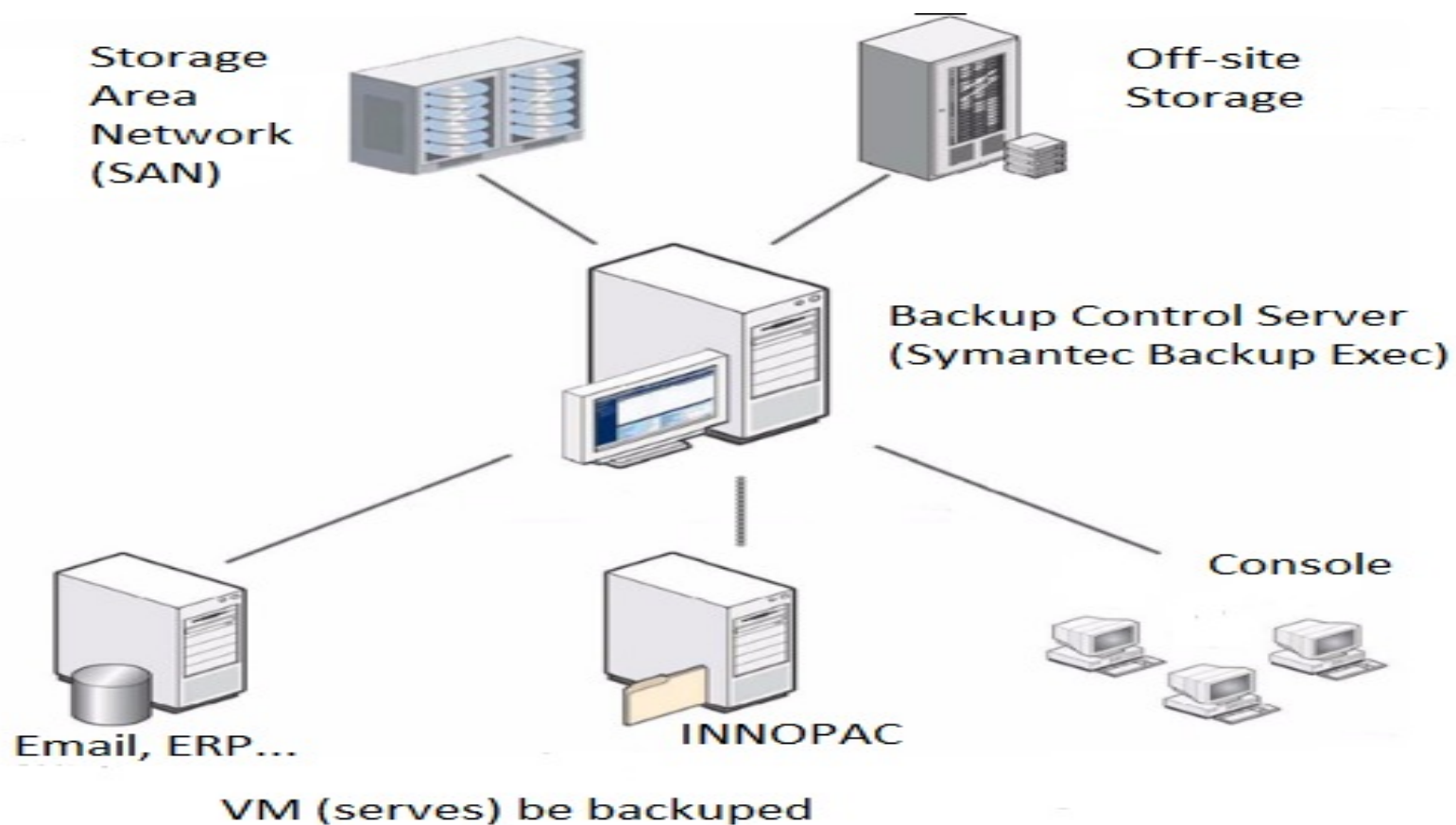- To run on a virtual server, OS has been changed to "Red Hat Enterprise Linux Server release 5.7 (Tikanga)"

redhat

Red Hat Enterprise Linux 5

3a. New Backup Method

- As the virtual server does not provide a directly attached tape drive device, Millennium requires us to acquire a new product, Enterprise Backup API, for performing daily or full backup

- Every late night, the Millennium data snapshot will be taken through the "Enterprise Backup API". Then the data snapshot will be archived on the ITO backup infrastructure through network

- As backup data is stored on network, we do NOT need to change backup tape every day → save a lot of labor works

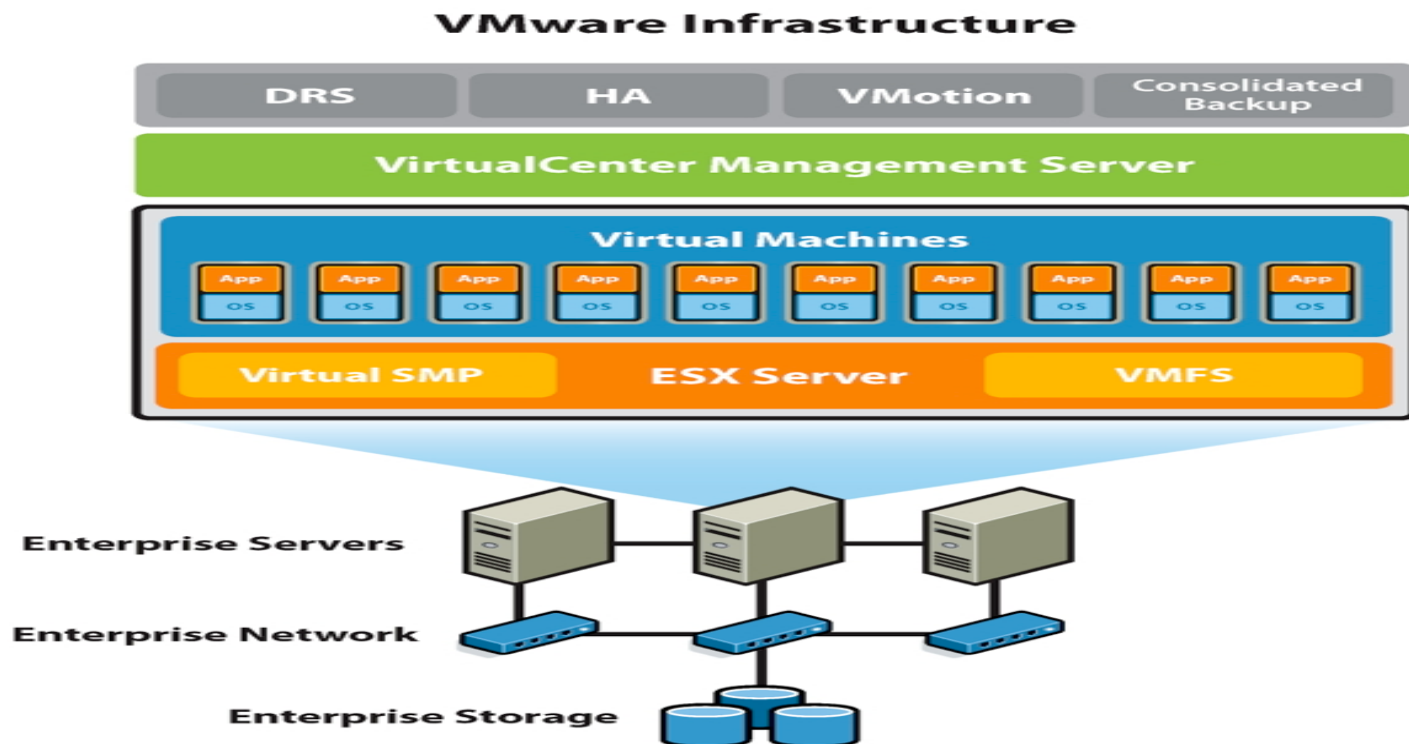- Average full backup time is reduced from 1 hour to 15 minutes

3b. ITO backup infrastructure

Storage Area Network (SAN)

Off-site Storage

Backup Control Server (Symantec Backup Exec)

Console

Email, ERP...

INNOPAC

VM (serves) be backuped

4a. ITO VM Infrastructure(1)

- Built upon VMware technology

VMware Infrastructure

4a. ITO VM Infrastructure(1)

- ITO virtual server infrastructure is built upon
    1. VMware technology
    2. High reliable storage system (SAN – Storage Area Network)
   3. Remote recovery site

4a. ITO VM Infrastructure(1)

- Advantages: A lot as mentioned above, such as save hardware, staffing, space, energy, standardization…

- Most important features appeared as a critical system for the Millennium,
  1. HA (High Availability)
  2. DR (Disaster Recovery)
  3. Consolidated Backup

4b. ITO VM Infrastructure(2)

- High Availability:
  1. The Millennium system image is on the fault-tolerate SAN storage
  2. As the hosting server crash, the Millennium service can be resumed on other server

4b. ITO VM Infrastructure(2)

- HA (High Availability)
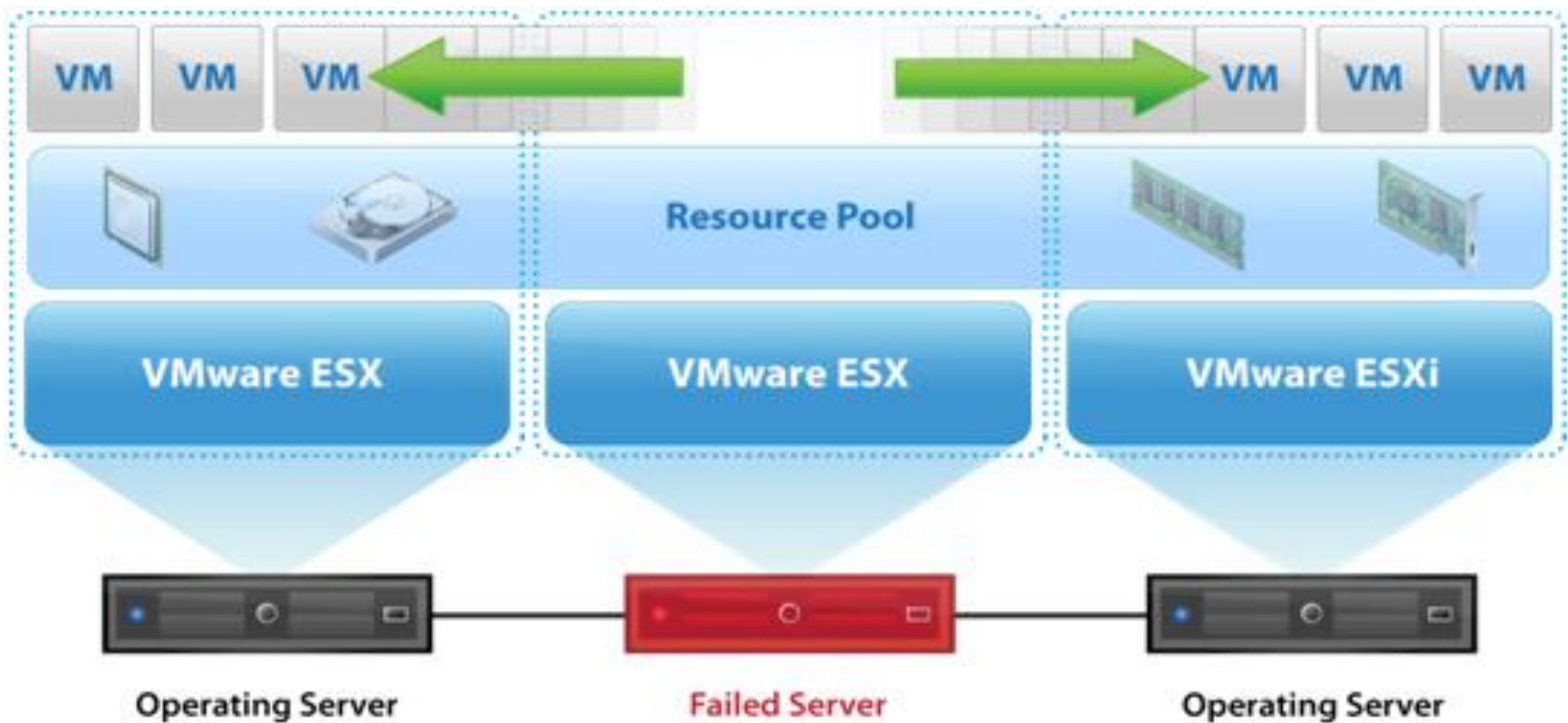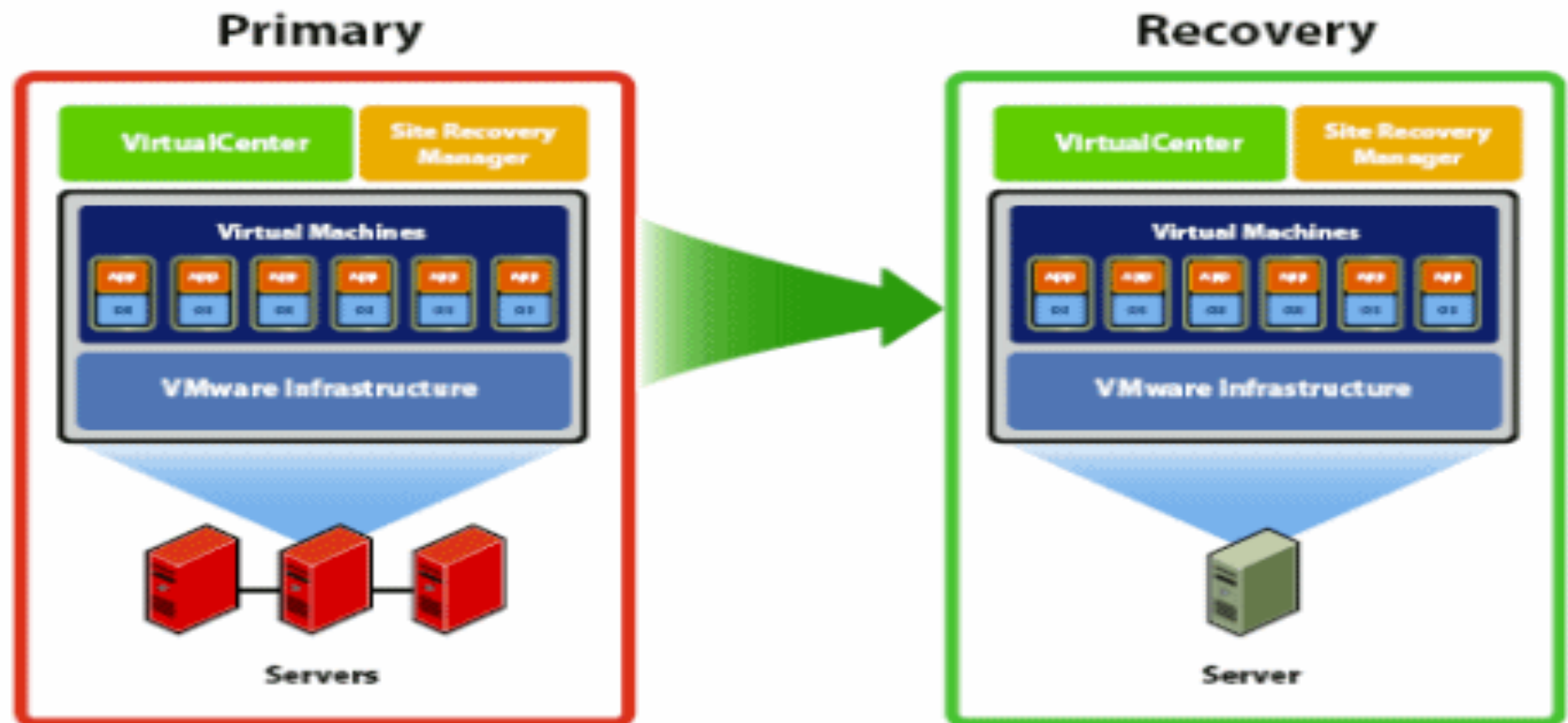
4c. ITO VM Infrastructure(3)

DR (Disaster Recovery)

- The critical system images are replicated onto the remote DR site

- In case of disaster, the Millennium server can be resumed as soon as the DR site up, and loading the data from daily backup
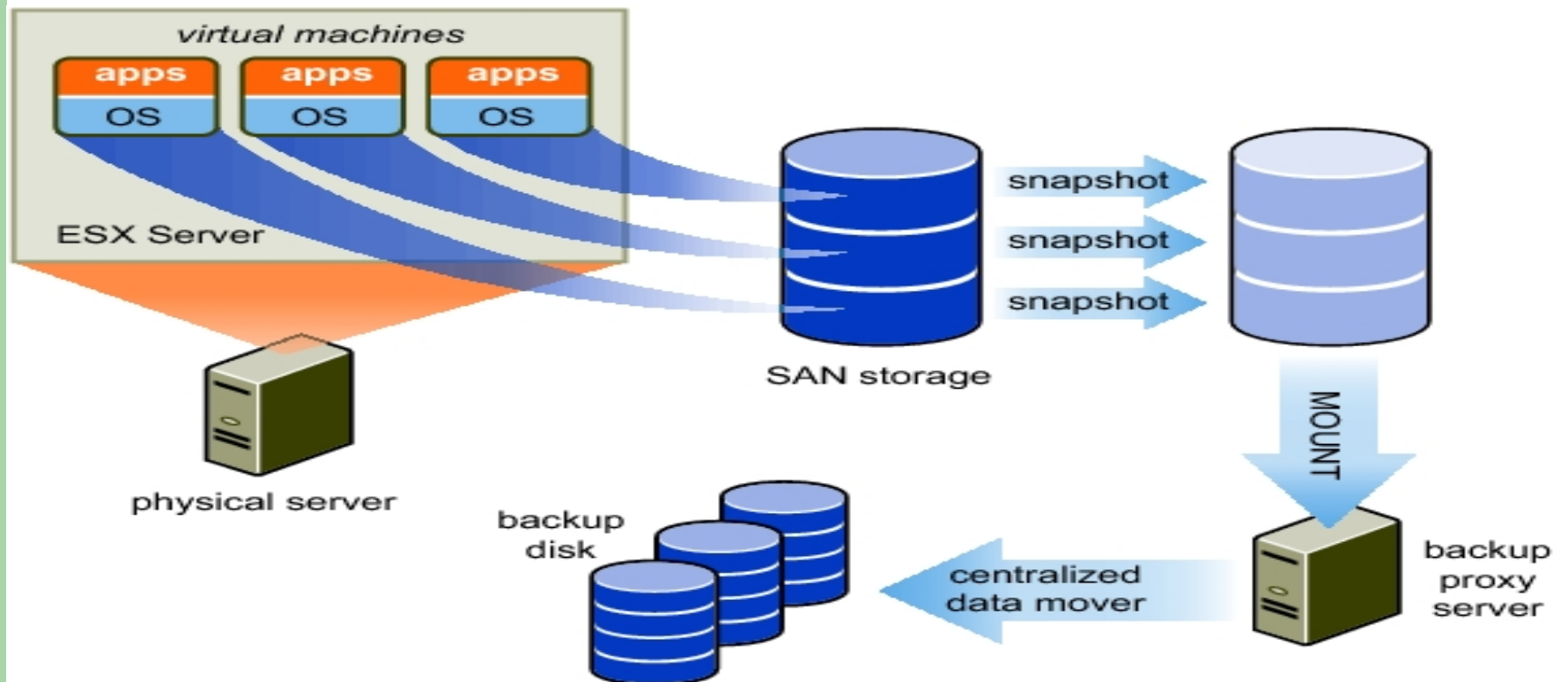
4c. ITO VM Infrastructure(3)

- DR (Disaster Recovery)

4d. ITO VM Infrastructure(4)

Consolidated Backup

- The system image snapshot is taken every night

- As necessary, it will only take minutes to restore the system backup to yesterday, or specified date state

4d. ITO VM Infrastructure(4)

- Consolidated Backup

5a. Migration Process

- The Millennium-VM Migration was conducted on 11 Oct., 2011
- Whole process was very smooth, and did not encounter any problems
- Special credits should be sent to
  - III Migration team
  - ITO VM Team, specially thanks to Jimmy

5b. Problems after Migration

- The performance of VM is very good, the overall speed of Millennium system are increased

- However, Millennium has shutdown *three times* suddenly on the first month after migration

5c. Millennium Shutdown Error Message
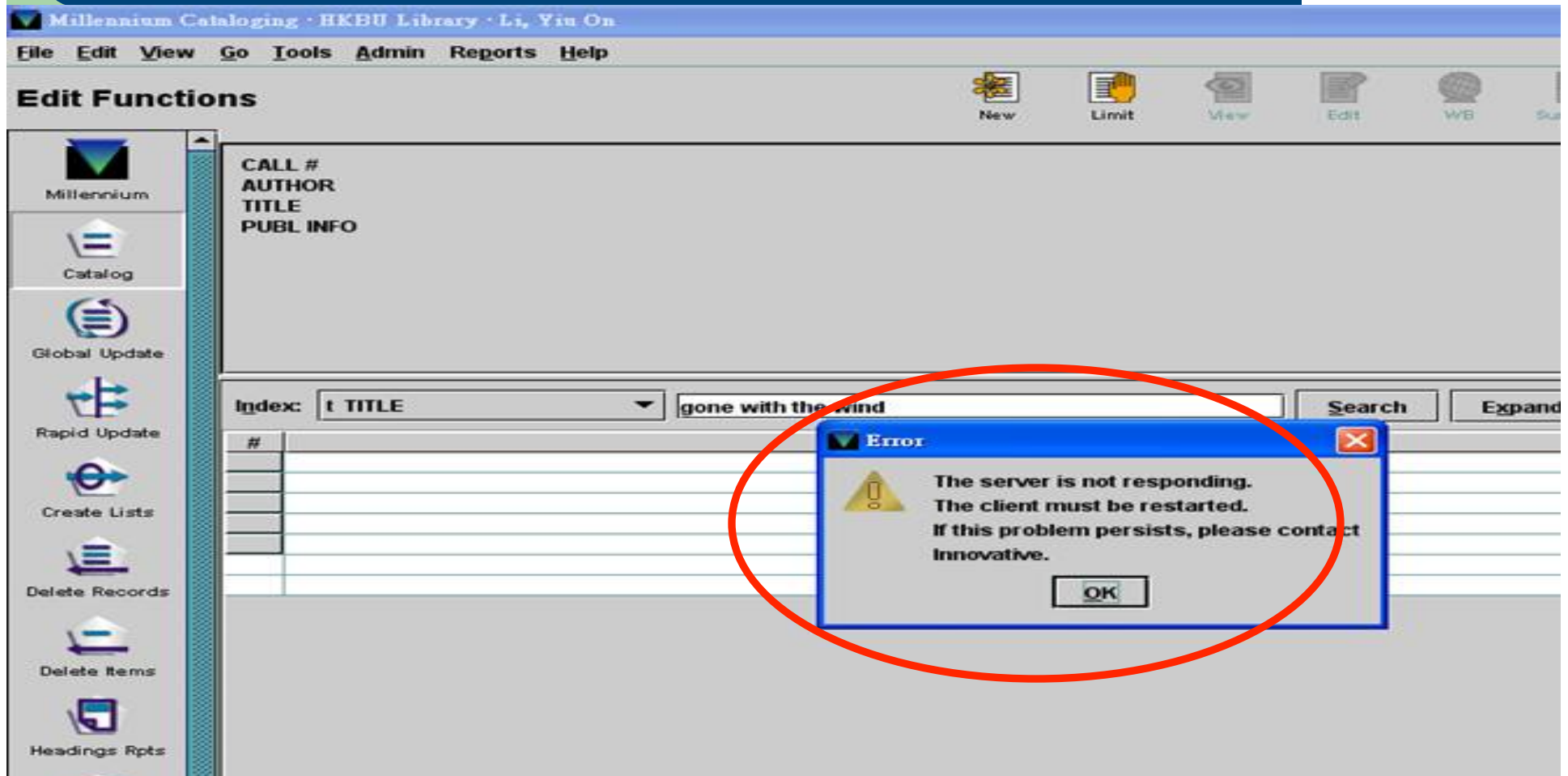
5c. III Helpdesk (1)

- All the three system shutdown incidents occurred on our HK day time, (i.e. night time at U.S.)

- We phoned the III Helpdesk immediately, but the response that we got was not very helpful –

  "the specialists/experts/mangers are not on duty at night, and problems should be resolved when they are back on tomorrow morning"

5c. III Helpdesk (2)

- The specialists/experts/mangers did help us to resolve the problem on the next morning, but we had to suffer an almost 24 hours downtime
- → We need to discuss the Helpdesk Service Level with the Innovative's representatives at the Discussion Forum later

5c. III Helpdesk (3)

- The report on detecting the causes of the persisting problems is

One thing we noticed was the ownership and group for /iiidb/software/shared/ were owned by "root". They should both be owned by "iii". We have since changed the ownership/group to "iii" which is the correct ownership values. One mystery is why they were changed. How did they change? At the OS level are you aware of anything different which might contribute to this behavior? Are you mounting or unmounting filesystems at various times?? Anything with the way you do backups? Something which might cause the ownership for this directory to change?

5c. III Helpdesk (4)

- It seems the problems were caused by the Enterprise Backup API
- We urge the Helpdesk to do a thorough investigation in order to prevent the same problem to be occurred on the other VM migration libraries

D. Conclusion

1. Win-Win Strategy

- We can use a Win-win Strategy to describe the Millennium-VM Migration project
    - Library: say good-bye to the Traditional Physical Server-Based Model
    - ITO: find a new tenant

2. Smooth Immigration

- Migration is quiet smooth except the unpleasant system shutdown experiences (*3 times!!!*)

3. Milk-Cow Model

- As a systems librarian, I am highly delighted that, with the support of ITO, we are can run Millennium on Cloud (VM).

- We are finally able to get rid of hardware stuff. Let us use a analog to end our presentation

    "I want my daughter to drink milk every morning, but I do not want/need to keep the cow"