

Introduction to **Virtualization & Xen**

國網中心-格網技術組
Rock(郭文傑), Rider(涂哲原)
rock@nchc.org.tw, rider@nchc.org.tw

為何要討論

Virtualization?

為何要討論 Xen ?

<http://dev.yesky.com/59/11048559.shtml>

我們將會討論到

1. 虛擬化是啥
2. 虛擬化有哪些 類型
3. Xen 的設計哲學&架構

虛擬是什麼？

Virtualization

國語小字典

虛_{ㄊㄩˊ}

- (1)假的、不真實的。如：「虛偽」、「虛假」、「虛構」、「虛幻」
- (2)衰弱。如：「虛弱」、「虛脫」、「她的身子很虛。」
- (3)白白的、徒然的。如：「虛度」、「不虛此行」、「虛有其表」
- (4)不自滿、不驕傲。如：「謙虛」、「虛心求教」
- (5)有所愧疚而心中膽怯。如：「心虛」
- (6)空。如：「空虛」、「趁虛而入」
- (7)草率、敷衍。如：「虛應故事」

擬_{ㄋㄧˇ}

- (1)計劃、打算。如：「擬購」、「擬辦」、「擬赴約」
- (2)起草、編寫。如：「草擬」、「如擬」、「擬稿」
- (3)模仿、仿效。如：「模擬」、「擬人」

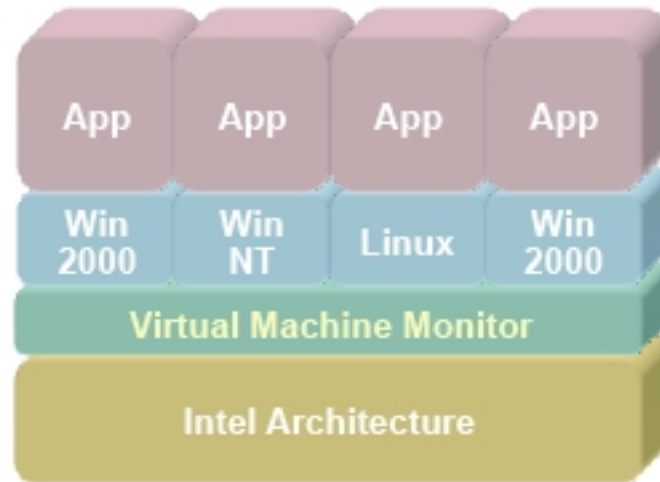
以 IT 界來說

A framework or methodology of **dividing the resources**
of a computer hardware into multiple execution
environments,

by applying one or more concepts or technologies such as hardware and software partitioning, time sharing, partial or complete machine simulation, emulation, quality of service, and many others.

比較實務的說法

A thin **software layer** that sits between Intel **hardware** and the **operating system**—virtualizing and managing all hardware resources



虛擬化的歷史

1959 - Christopher Strachey, Time Sharing in Large Fast Computers

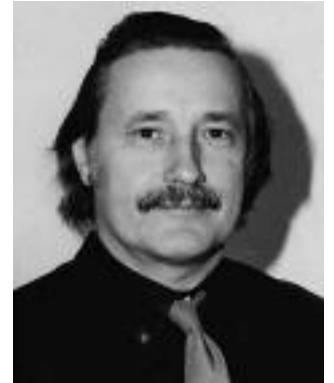
1965 - IBM 7044

1999 - VMware for x86 的VMM

2003 - 微軟收購connectix, 並推出virtual server
Xen 首次發表 Para-Virtualization

2006 - Amazon EC2 (Xen Hypervisor)
Intel & AMD Hardware Virtualization
KVM (Kernel-based Virtual machine)發表

2008 - MS Hyper-V



虛擬化與Linux的演進



2005 - VmWare runs on Linux

2007.04 - Kernel Virtual Machine Merged into Linux 2.6.20

2007.10 - Xen Merged into Linux 2.6.23

2009 - Red Hat 宣佈從Xen改為支援KVM(到2014就只能運行KVM)

虛擬化有哪些 Type

有哪些東西可以被虛擬呢？



Virtualization 的類型

Server/Platform Virtualization

(ex: VMware, OpenVZ, Xen...)

Storage Virtualization

(ex: LVM, IBM SAN Volume Controller...)

Network Virtualization

(ex: VPN...)

Application Virtualization

(ex: JVM, VMware ThinApp...)

Desktop Virtualization

Computer Cluster & Grid Computing

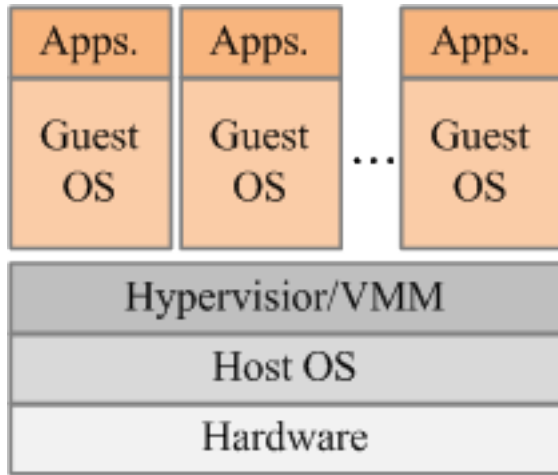
接下來將深入探討

Server Virtualization

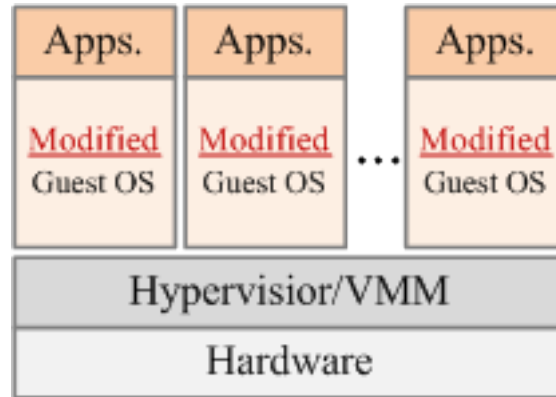
Term

- VMM(Virtual Machine Monitor) = Hypervisor
- Host = domain0 = dom0
- Guest = domainU = appliance = domU = instances = VM
- HVM = Hardware Virtual Machine
 - VT-x = Intel x86 VT-enabled processor
 - VT-i = Intel IA-64 ...
 - SVM = AMD Pacifica-enabled ...

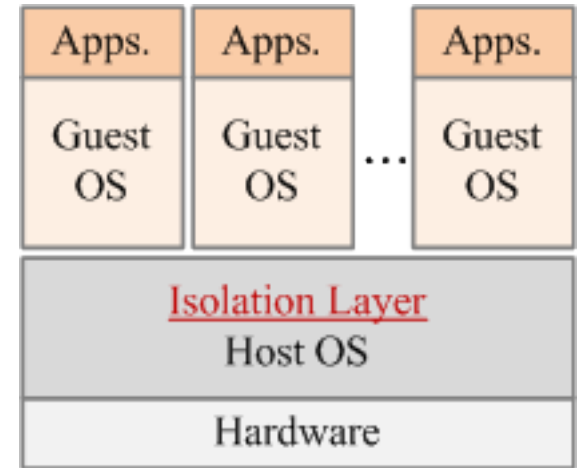
Server Virtualization 的技術有



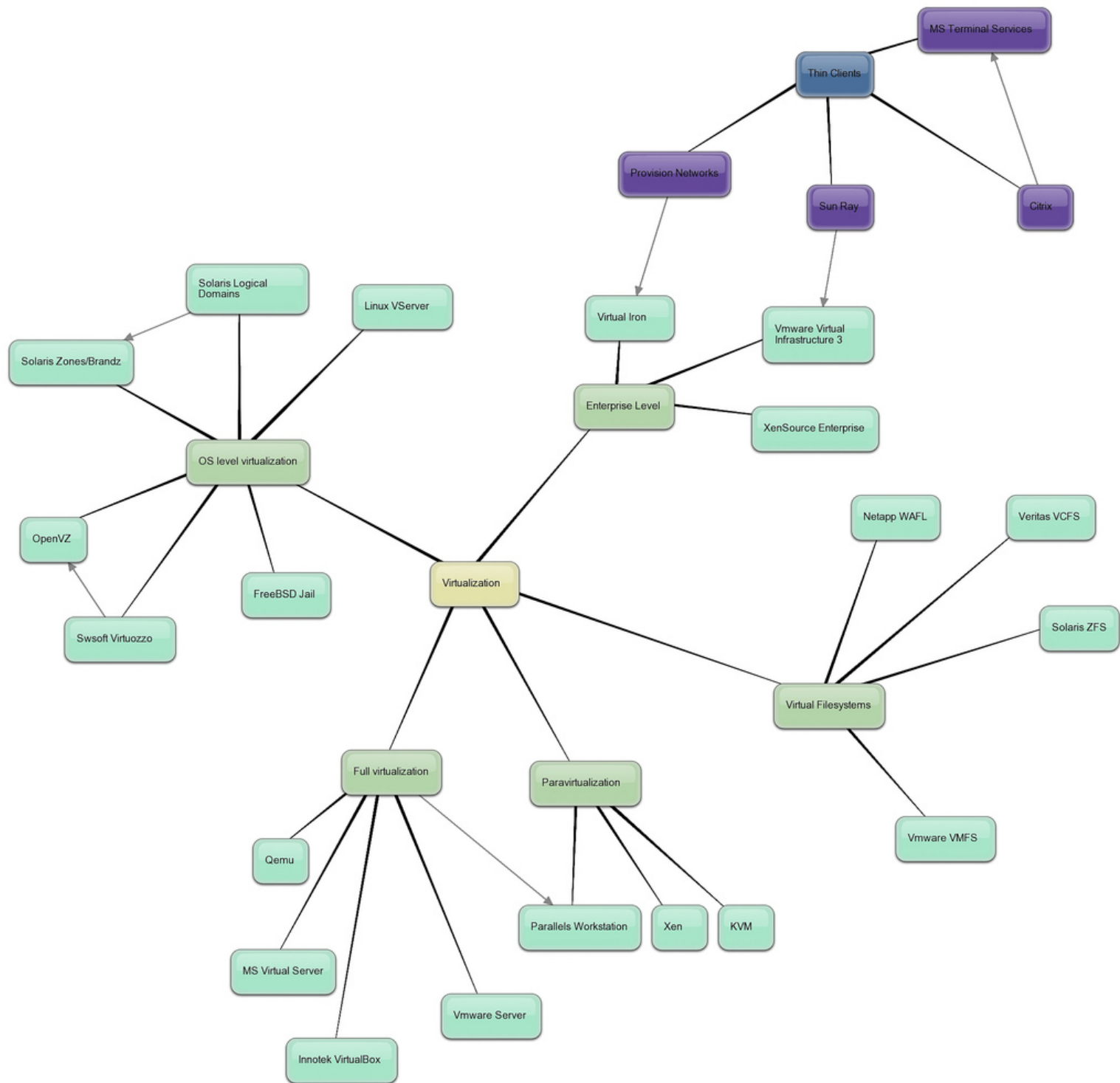
**Full
Virtualization**



Para-Virtualization



**OS-level
Virtualization**



source: <http://www.flickr.com/photos/ibegtin/469821424/sizes/l/in/photostream/>

而今天的主角

Xen

是 **Para-Virtualization** 流派

Xen

是一個 Open source 的
Virtual Machine Monitor/Hypervisor，
目前可支援 x86, x86_64, IA64, ARM 等
CPU架構上的虛擬化

設計想法

避免虛擬化的效能損耗

20~30%↓ -> 2~8%↓ (98% CPU Speed)

1. no modify guest application
2. support multi-app OS
3. high performance & strong resource isolation
4. hiding effect of virtual resource -> correctness & performance

Para-Virtualization

設計實作

- 主機跟虛擬機器的 Kernel 必須能夠支援 Xen
- Xen patch Kernel:
 1. 記憶體管理
 2. CPU管理
 3. 設備 I/O

若是對細節設計有興趣的可以參考這篇 Paper
[Xen and the art of virtualization](#)

PV ***x86 Interface***

- **Memory-Management**

- Segmentation
- Paging

- **CPU**

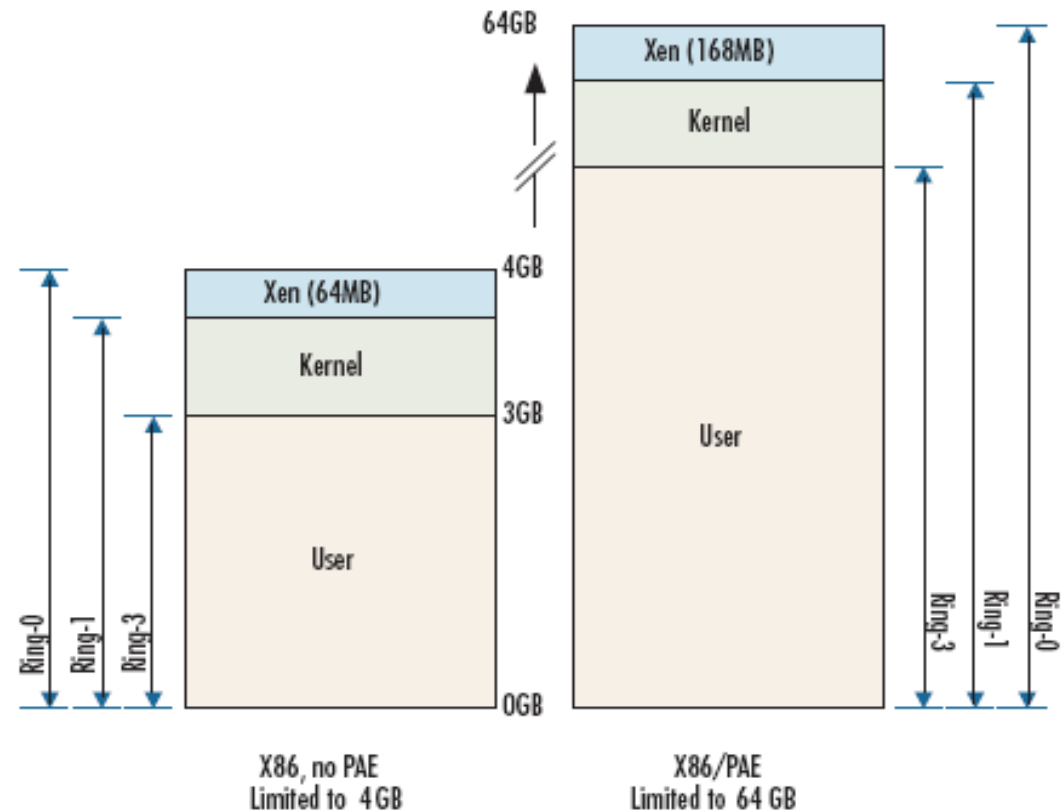
- Protection
- Exceptions
- System Calls
- Interrupt
- Time

- **Device I/O**

- Network, Disk, etc...

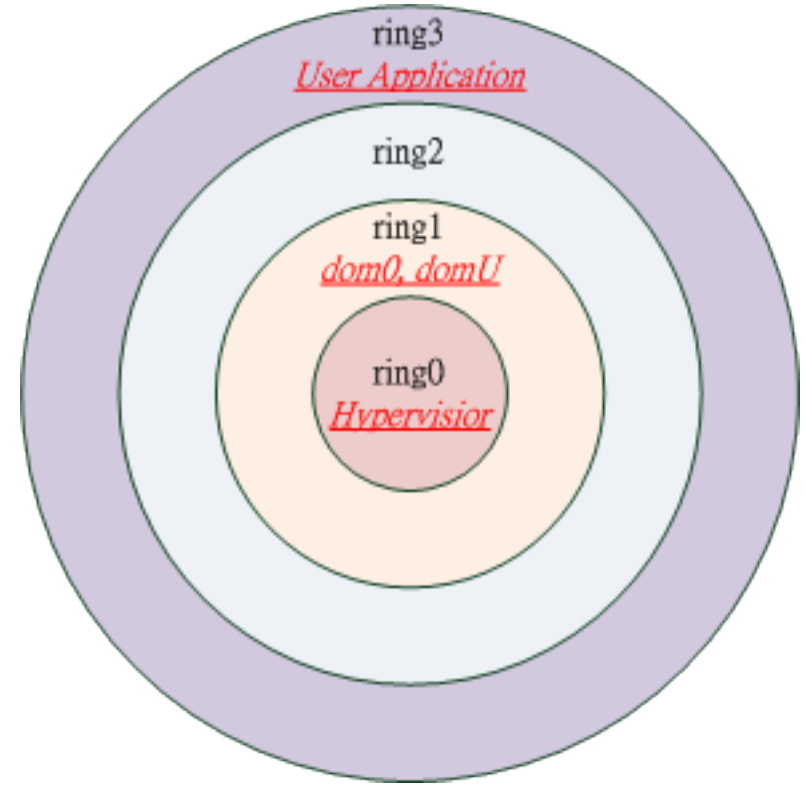
PV Memory Management

- Virtualizing memory is difficult part
- Performance & Isolation(safety)
 - Guest direct read page table, only Xen can update
 - discontinuous page

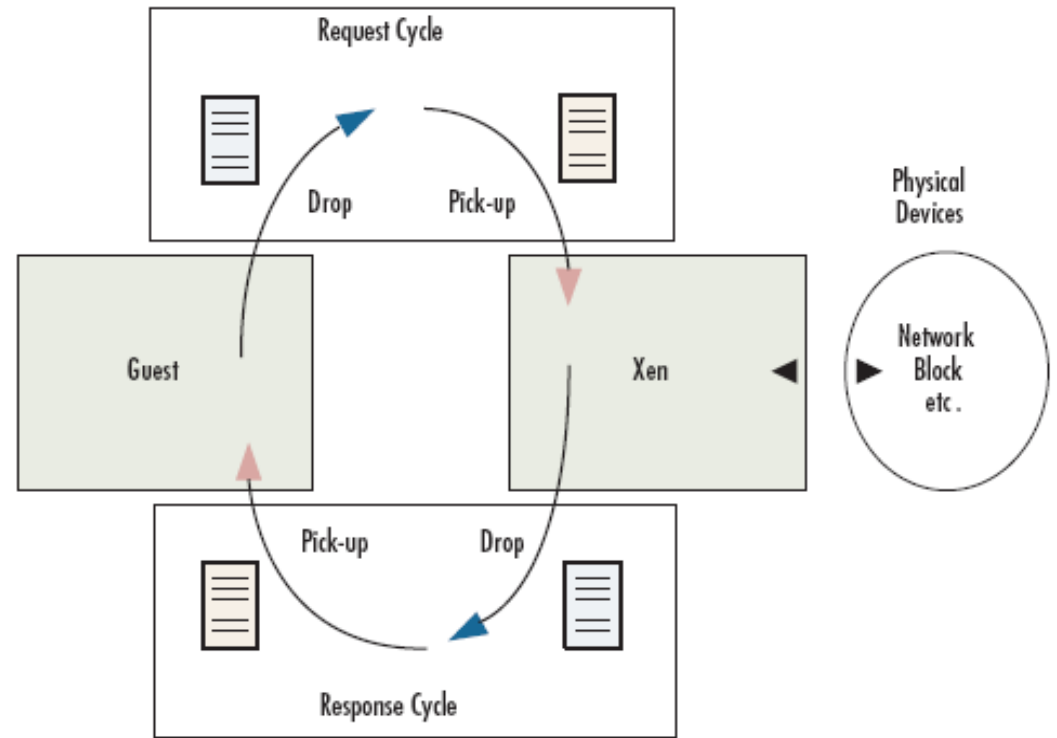


PV CPU

- X86 ring
 - ring0 -> OS
 - ring3 -> Application
- X86 protection ring in Xen
 - ring0 -> Hypervisor
 - ring1 -> dom0, domU
 - ring3 -> Applications



PV Device I/O



- **Virtual Device**

- **Data Transfer**

- **Asynchronous I/O rings**

- **Interrupt & Notifications**

- **Event Mechanism**

Xen Kernel Option

檔案(E) 編輯(E) 顯示(V) 終端機(T) 分頁(T) 求助(H)

.config - Linux Kernel v2.6.22.9 Configuration

XEN

Arrow keys navigate the menu. <Enter> selects submenus —>. Highlighted letters are hotkeys. Pressing <Y> includes, <N> excludes, <M> modularizes features. Press <Esc><Esc> to exit, <?> for Help, </> for search. Legend: [*] built-in []

[*] Privileged Guest (domain 0)

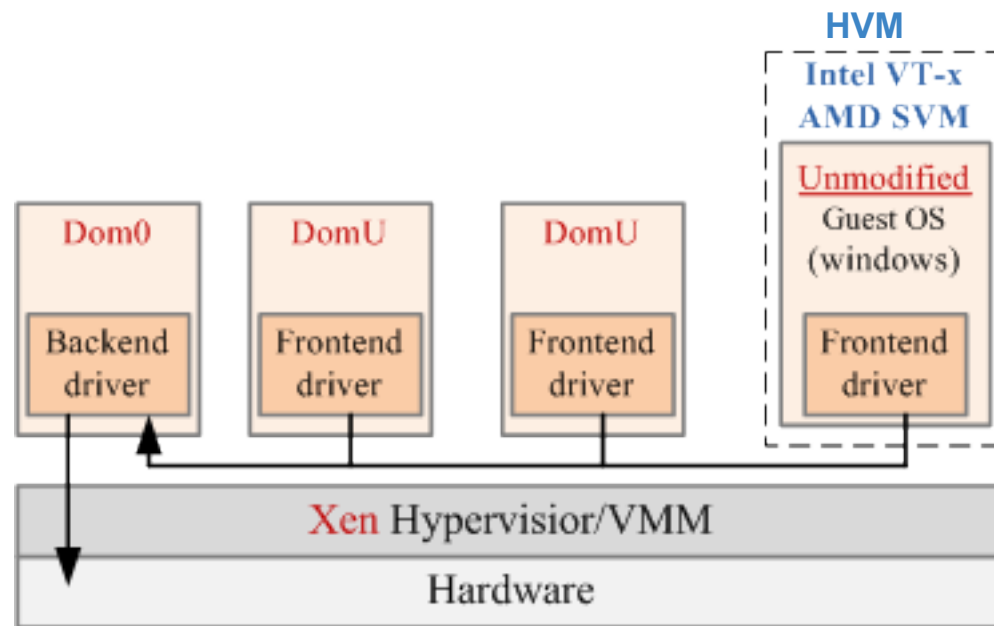
- <*> Backend driver support
 - <*> Block-device backend driver
 - <*> Block-device tap backend driver
 - <*> Network-device backend driver
 - [] Pipelined transmitter (DANGEROUS)
 - <*> Network-device loopback driver
 - <*> PCI-device backend driver
 - PCI Backend Mode (Virtual PCI) —>
 - [*] PCI Backend Debugging
 - <M> TPM-device backend driver
- <*> Block-device frontend driver
- <*> Network-device frontend driver
- <M> Framebuffer-device frontend driver
- <M> Keyboard-device frontend driver
- [*] Xen virtual console
- [*] Scrub memory before freeing it to Xen
- [*] Disable serial port drivers
- <*> Export Xen attributes in sysfs
 - Xen version compatibility (3.0.4 and later) —>

<select>

< Exit >

< Help >

根據此想法設計出的架構為

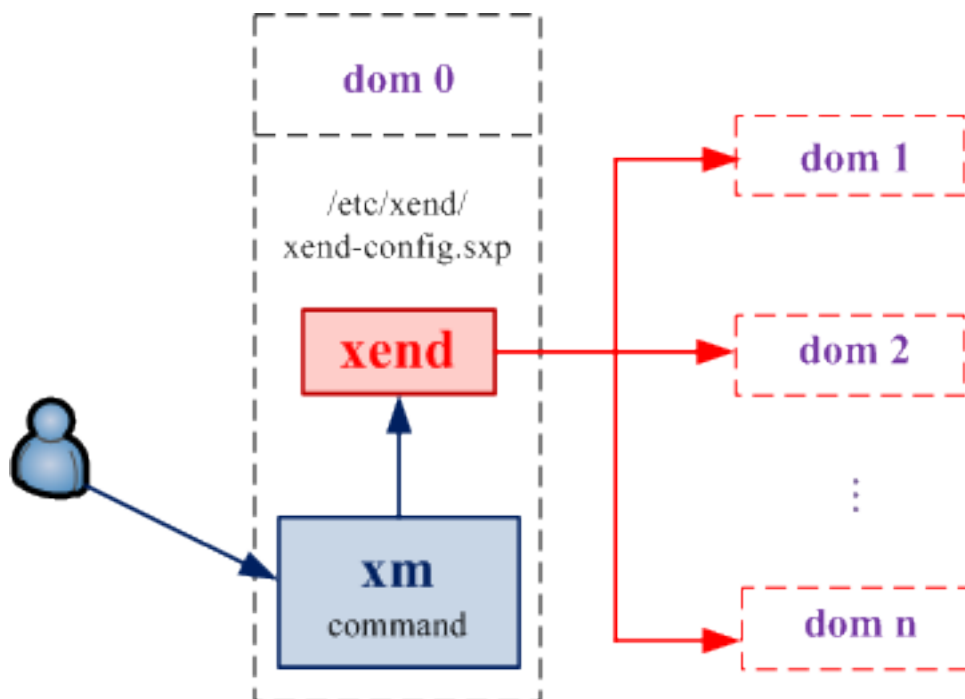


Dom 0 / Host - 開機後就會啟動的 domain 0

Dom U / Guest - Virtual machine (需透過 dom0 來控制硬體資源)

HVM - Xen可藉由 CPU 指令集(Intel VT-X, AMD SVM)的支援來做到 Full Virtualization

如何運作呢?



xend - Xen 的核心程序

xm - user 透過 xen manager 來管理 **xend** 和控制 **VM**



PLAY

PLAY
COMME des GARÇONS*

Single Machine

1. Xen-enabled Kernel Installation
2. Xen Image & tools Installation
3. Configuration
4. Create VM

HaHa...
It's a piece of cake



**But, The Challenge
is**

Build *Multiple VMs*
in *Multiple Hosts*

再多台機器上你必須想的更多&看的更遠

1. **Network Arch Design**
2. **Stroage Design**
3. **VM Central Management**

精彩回顧



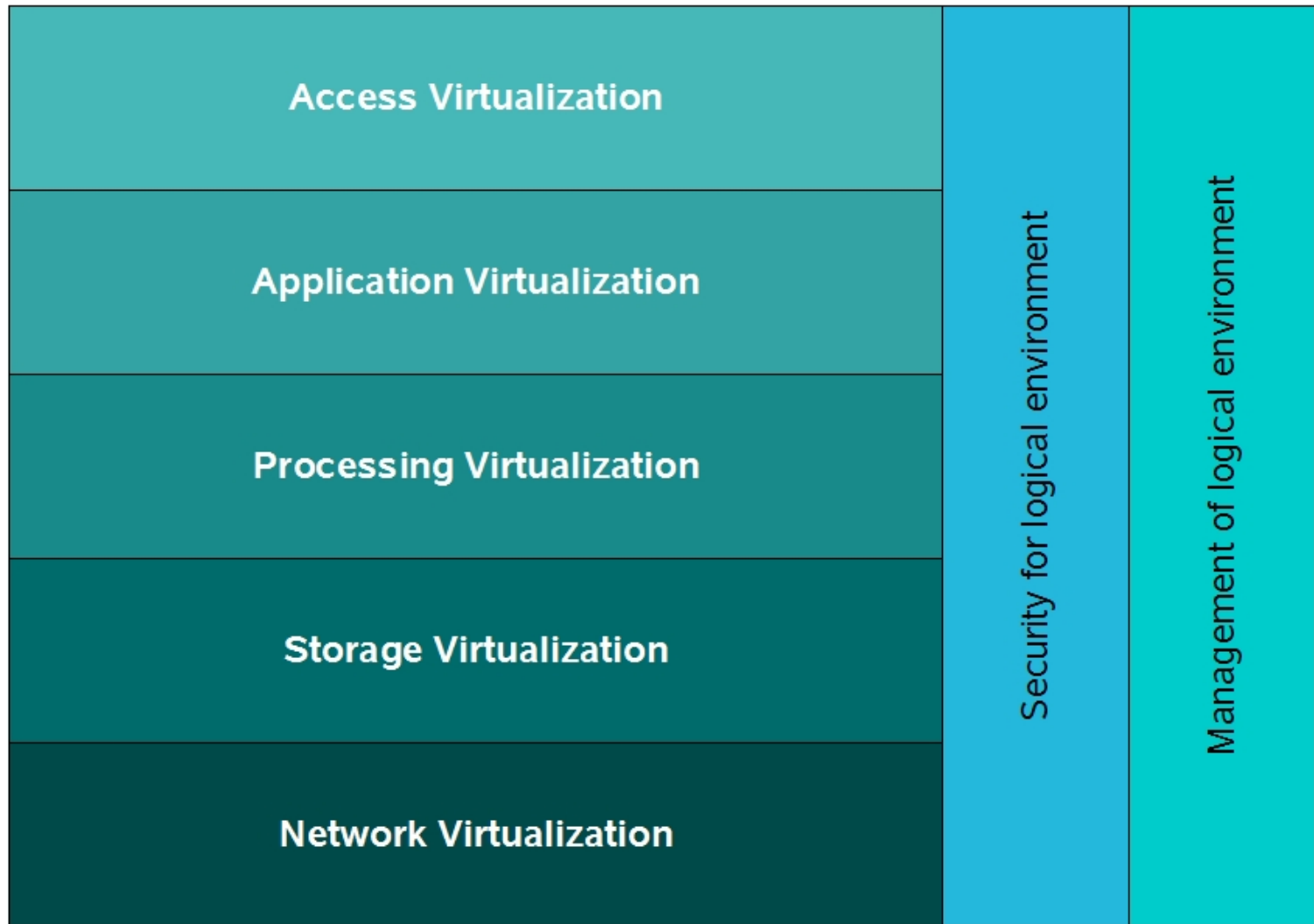
- 了解虛擬化的類型
- Xen 的架構
- Xen 的運作

討論時間

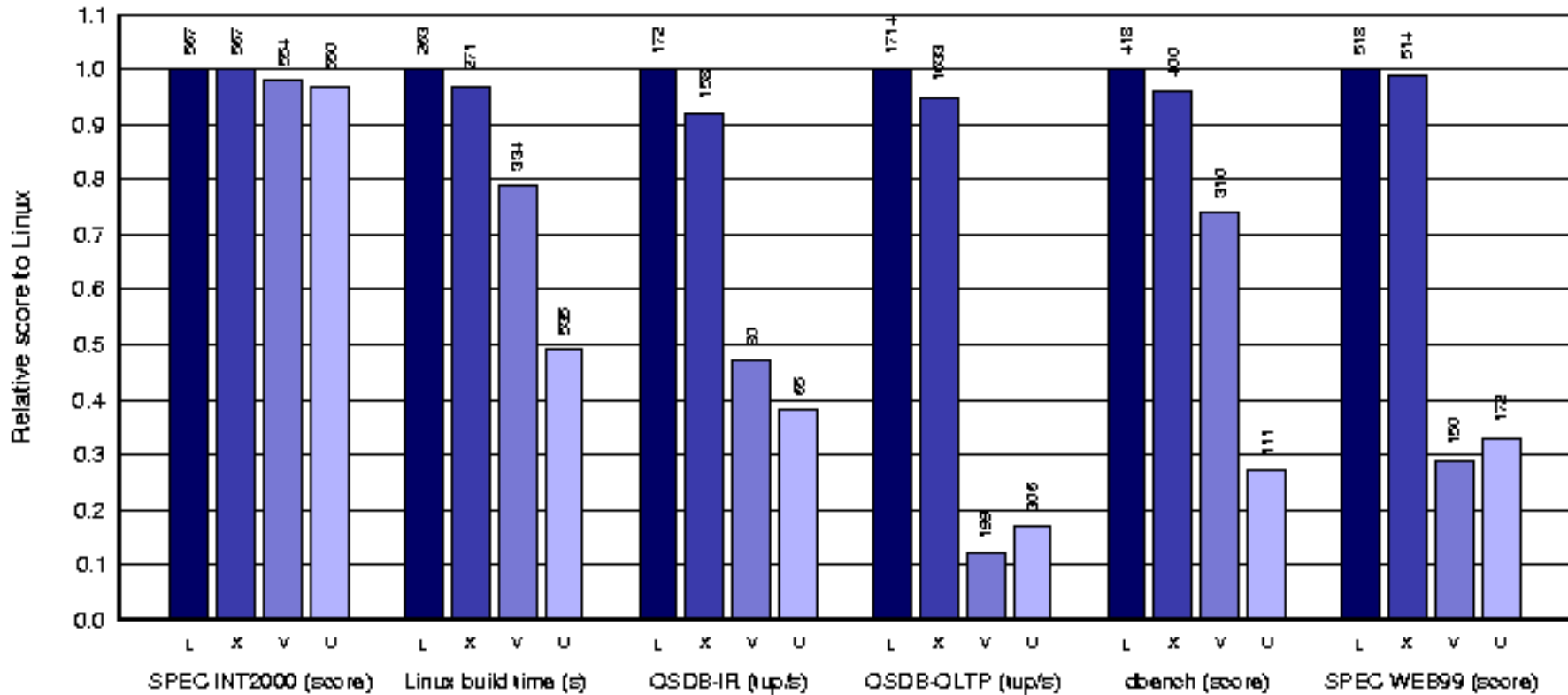


Appendix

Type of Virtualization



Performance (Hypervisor)



L -> native Linux

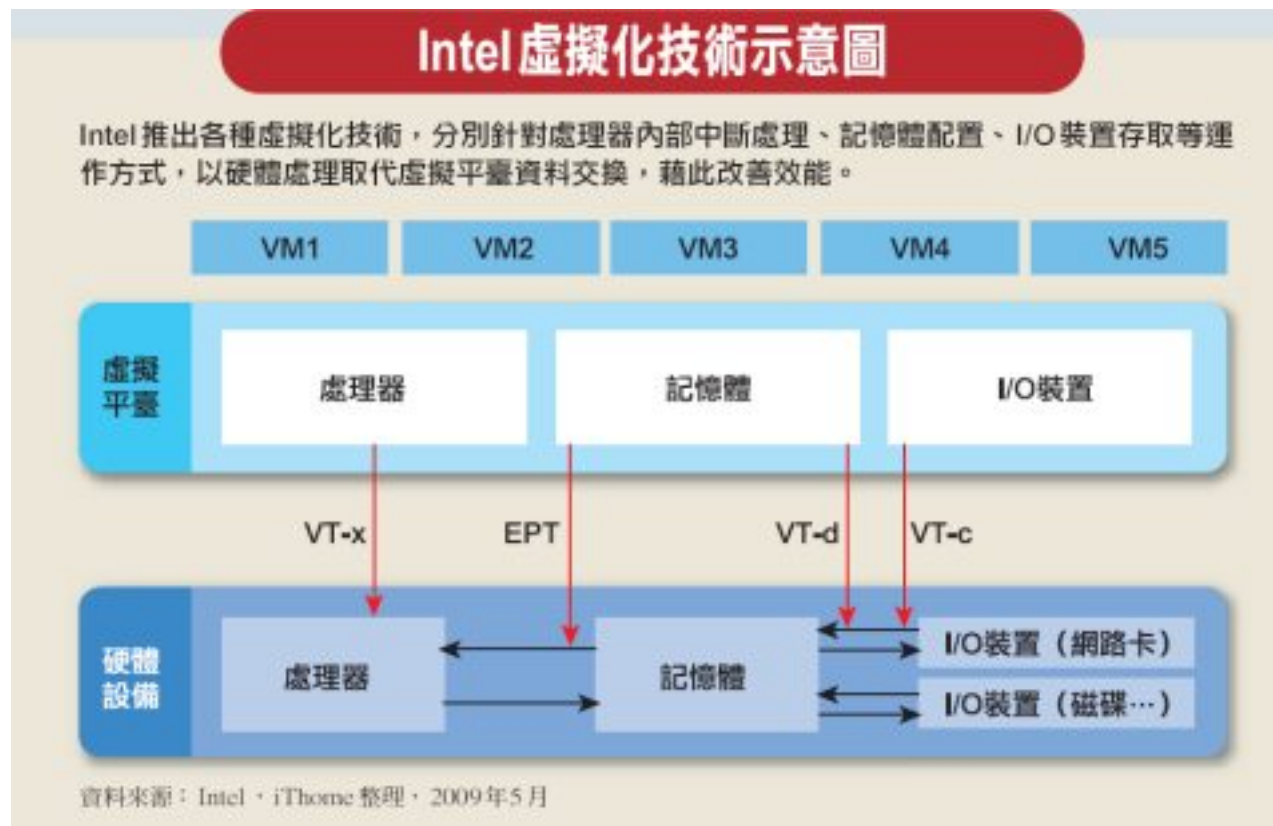
X -> Xen/Linux

V -> VMware Workstation 3

U -> User Mode Linux

Intel 的虛擬化簡介

- **VT-x:** VT FlexPriority、VT FlexMigration 與 VT Extended Page Tables(EPT)
- **VT-d:** 改善I/O效能
- **VT-c:** VMD 與 VMDc



<http://vaemon.com/article/633.html>