虛擬化技術介紹

吳庭育

tyw429@gmail.com

Outline

- 虛擬化技術介紹
- 虛擬化方法與架構
- 虛擬化技術原理
- 虛擬化技術的產品

虛擬化技術介紹

虛擬化技術介紹

- 虛擬化技術漸漸普遍化，從伺服器到桌上型電腦都呈現急切需求導入其虛擬化之應用，因此相信大家都不會懷疑虛擬化技術的可用性和研究其技術的必要性。
- 一般而言，虛擬化就是把實體資源轉變為邏輯上可以管理的資源，以打破實體結構間的不可切割的障礙。
- 虛擬化技術本質就是一種資源管理技術，它將硬體、軟體、存儲、網絡等硬體設備分離開來,讓使用者能更合理、更充分的控制與管理這些資源。

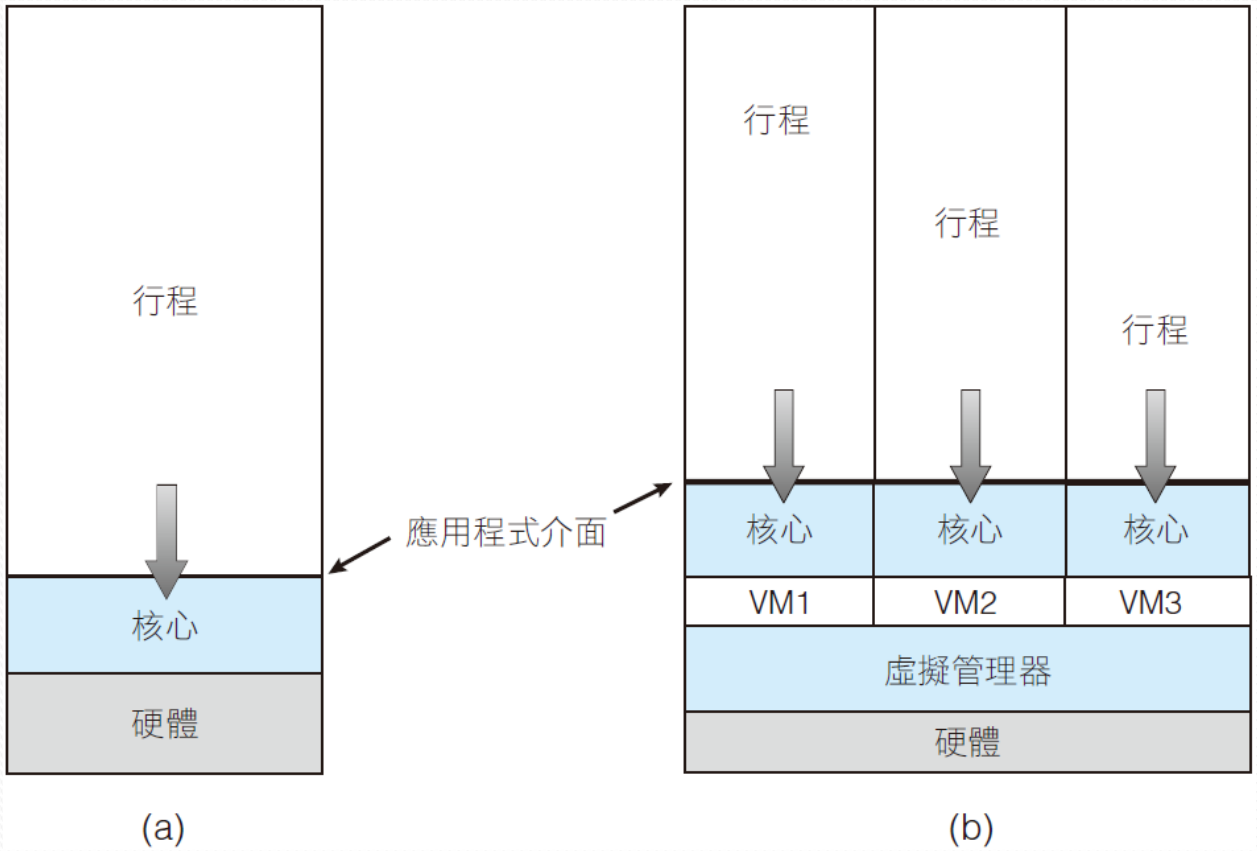
概　要

- 虛擬機的基本想法就是使用一部電腦中的硬體 ，來執行幾個不同的作業環境
  - 硬體：CPU、記憶體、磁碟裝置、網路介面卡等
- 虛擬機建置涉及多種元件
  - 在底層為**主機**（host），底層的硬體系統運行著虛擬機
  - **虛擬機管理器**（virtual machine manager, VMM）建立
  - 也稱為**虛擬機管理程式**（hypervisor）
- 每個**客戶**（guest）行程提供主機的虛擬副本
  - 通常，客戶行程實際上是一個作業系統

何謂虛擬化技術

- 將原本運行在真實環境上的電腦系統或文件，運行在虛擬的環境中
- 解除上下兩層間原本存在的耦合關係，使得上層運行不依賴於下層的具體實現
- 電腦系統從下至上可分為
  - 底層硬體資源
  - 作業系統
  - 應用程式撰寫介面
  - 應用程式

系統模組



虛擬化技術術語介紹

- 寄宿機**(Host)**
  - 即虛擬機管理程序所在的主機系統
- 客戶機**(Guest)**
  - 即運行在虛擬化管理器之上的虛擬機系統
- **VMM (Virtual Machine Monitor)**
  - 虛擬機監視器可以監視虛擬機的運作
- **Hypervisor**
  - 虛擬機管理程序(算是高階VMM)

虛擬化方法與架構

虛擬化方法

- 基本上虛擬化解決方案是要進行將實體機器虛擬化，這台機器可能直接支持虛擬化，也可能不會直接支持虛擬化。

- 若硬體不會直接支持虛擬化，那麼就需要使用虛擬化管理程序層的支持虛擬化。虛擬機管理程序稱為VMM或Hypervisor，可以看作是平台上硬體及作業系統的抽象化。

虛擬化架構

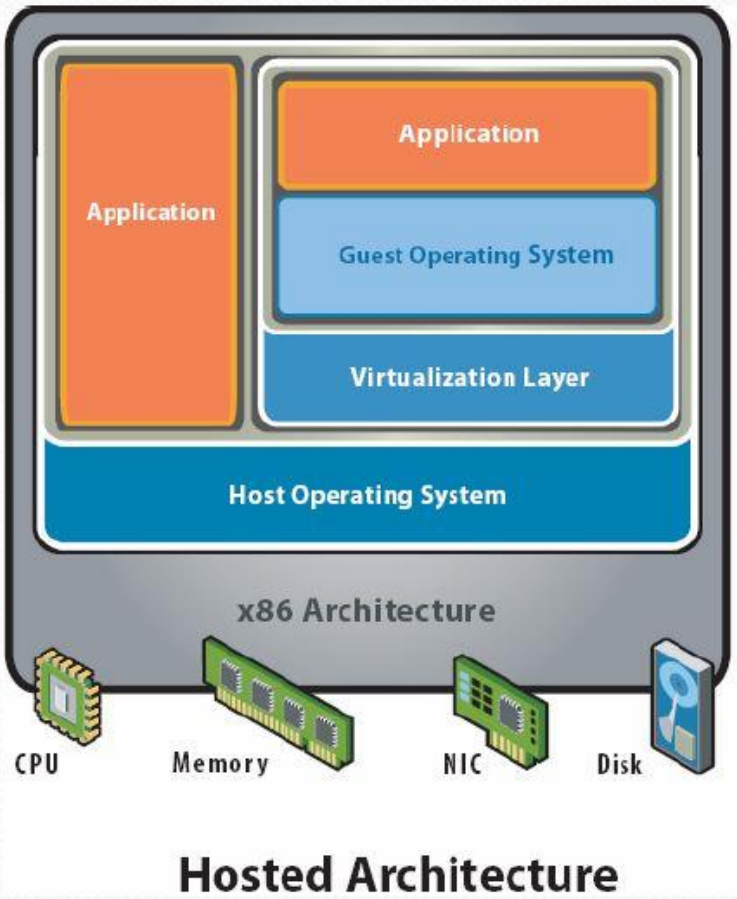
- 在某些情況中，這個虛擬機管理程序(VMM)等同於一個作業系統，可以稱為主機作業系統(Host OS)。

- 架在虛擬機管理程序之上是虛擬機（VM），其內部運行的是客戶機作業系統(Guest OS) 。

- 這些VM都是一些相互隔離的客戶機作業系統，它們將底層硬體平台視為自己所擁有。但是實際上，是虛擬機管理程序為它們製造了這種假象(虛擬化)。

寄宿架構

- 寄宿架構：採用模擬軟體技術模擬出計算機硬體和軟體。模擬層與作業系統對話，而作業系統與計算機硬體對話。在模擬層上虛擬機器安裝的作業系統並不知道自己是被安裝在模擬環境下的，你可以按照一般的方法安裝該作業系統。這種虛擬化需要付出很大的性能代價。(Hosted Architecture)



Hosted Architecture

模擬器 (emulator)

- 模擬器（emulator）允許針對一種硬體環境撰寫的應用程式可在不同的硬體環境上運行(例如不同類型的 CPU)
- **應用程式容量**（application containment）不完全是虛擬化，而是經由將應用程式與作業系統隔離，提供類似於虛擬化的功能
  - Oracle Solaris Zones、BSD Jails 和 IBM AIX WPAR “包含”應用程式，使它們更加安全和可管理

歷　史

- 虛擬機於 1972 年首次在 IBM 大型主機上商業化出現
- 虛擬機建立後，使用者可以執行任何作業系統
  - 真實度 (fidelity)：VMM 為程式提供與原始電腦基本相同的環境
  - 效能 (performance)：在該環境中運行的程式僅顯示輕微的效能下降
  - 安全 (safety)：VMM 完全控制系統資源
- 到 1990 年代後期，Intel 80x86 CPU 在功能上已經變得普遍、快速和豐富
  - 開發人員展開多種努力來實現該平台上的虛擬化
  - Xen 和 VMware 都開發相關技術

優點和特點

- 主機系統可以被虛擬機保護，就像虛擬機保護彼此一樣
  - 因為每個虛擬機幾乎與其它所有虛擬機完全隔離，幾乎沒有保護問題
- 隔離的潛在缺點是，它能夠阻止共享資源
- 客戶可以**恢復**（resume）原本的狀態，就像在原始位置上一樣機器，建立一個**複製體**（clone）
  - 快照記錄時間點，客戶如有必要可以將其恢復到該時間點
    - 例如我們做了更改但已不再需要

優點和特點

- 虛擬機在生產資料中心使用中的主要優勢是系統**合併**(consolidation)
  - 涉及採用兩個或多個單獨的系統併在一個系統上的虛擬機中運行它們
- **模板化**（templating），為其中一個標準虛擬機映像，包括一個已安裝和配置的客戶作業系統和應用程式，已保存並用做多個正在運行的 VM 的來源
- **即時遷移**（live migration）功能，該功能可以將客戶作業系統從一台實體伺服器轉移到另一台實體伺服器

優點和特點

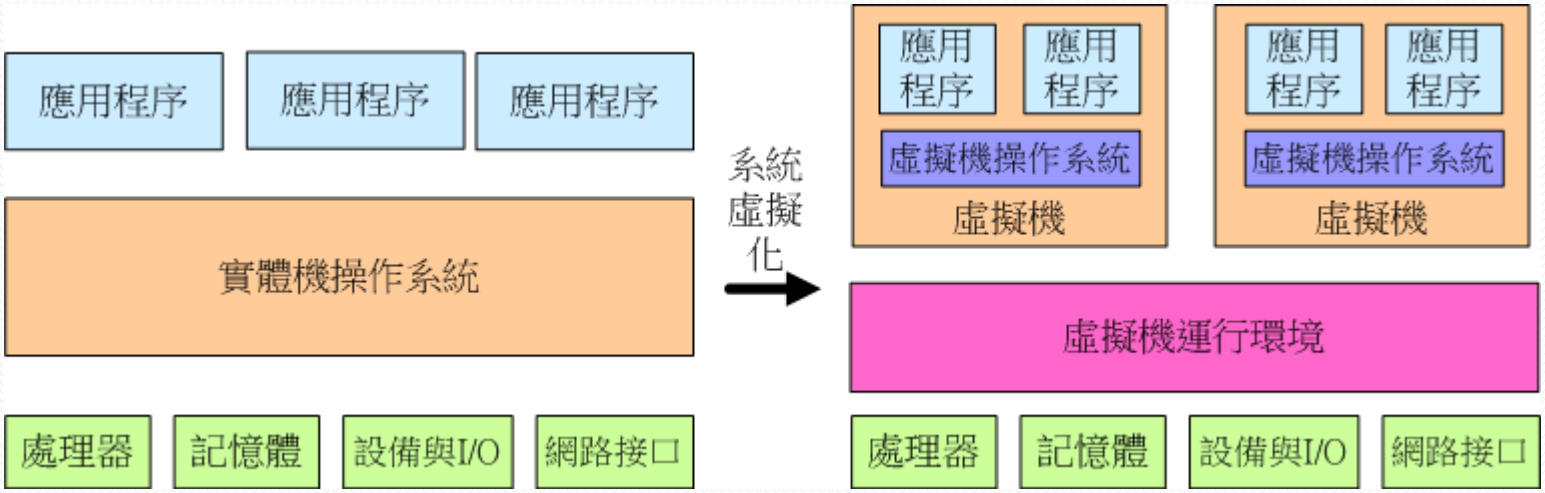
- 虛擬化為電腦設備實現、管理和監視方面的許多其它進步奠定基礎
  - 例如，透過虛擬化讓**雲端運算**（cloud computing）成為可能，在虛擬化中，使用網際網路技術將像是 CPU、記憶體和 I/O 之類的資源作為服務提供給客戶
  - 藉由使用 API 程式，可以告訴雲端運算裝置建立數千個 VM，這些 VM 都執行於特定的客戶作業系統和應用程式中，其它人可以經由網際網路存取這些 VM

系統虛擬化(1/5)

- 系統虛擬化
  - 使用虛擬化軟體在一台實體主機上虛擬出一台或多台虛擬機
  - 虛擬機是指使用系統虛擬化技術，運行在隔離環境中、具有完整硬體功能的邏輯電腦系統，包括客戶(作業系統(Guest OS)和其中的應用程式
  - 系統虛擬化中，多個作業系統可以互不干擾，運行於同一台實體主機上，重複使用實體主機資源

系統虛擬化 (2/5)

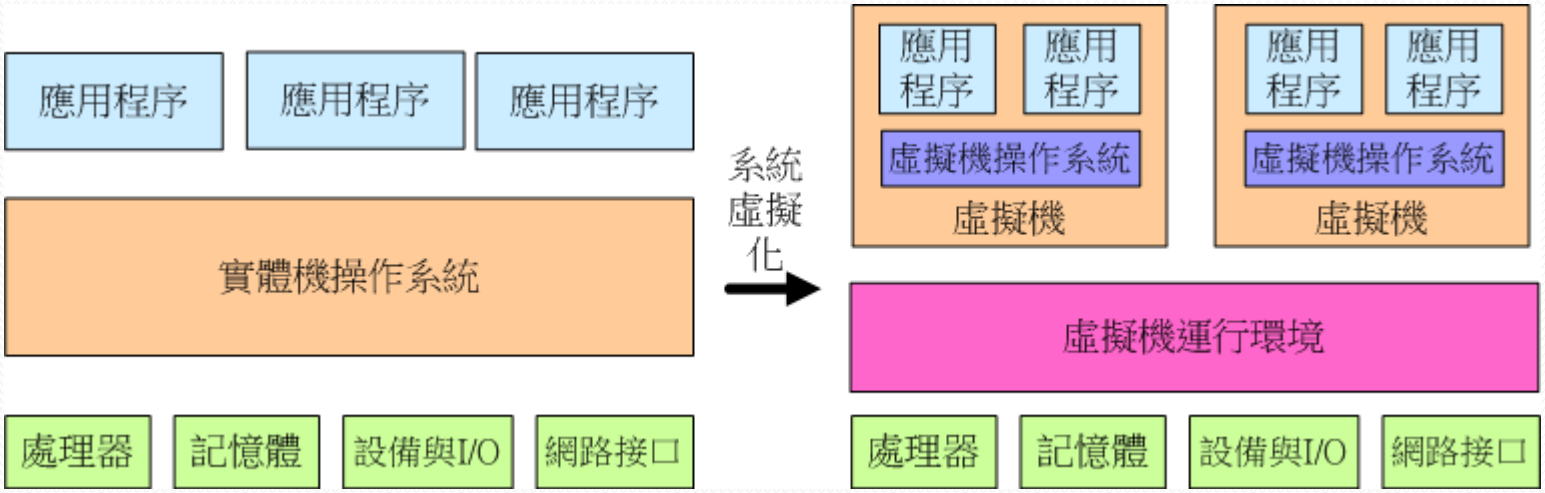
- 系統虛擬化中虛擬運行環境都需要為在上面運行的虛擬機提供一套虛擬的硬體環境
- 包括虛擬的處理器、記憶體、設備與I/O及網路介面



圖片來源:雲端策略 陳瑩

系統虛擬化 (2/5)

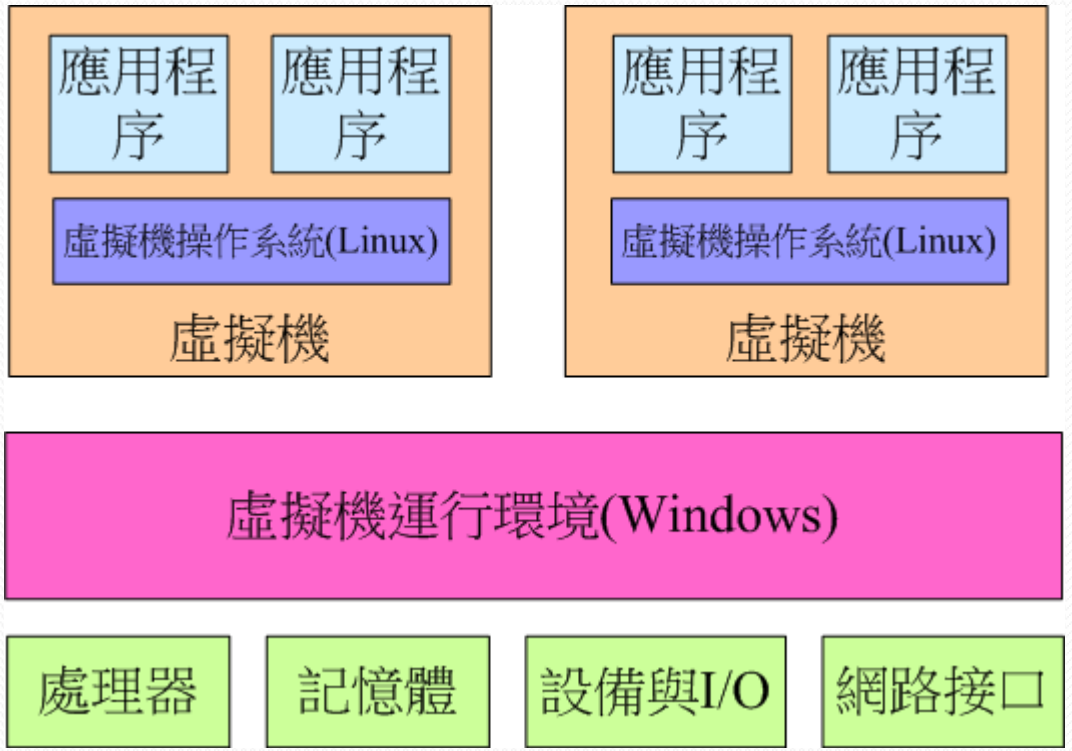
- 系統虛擬化中虛擬運行環境都需要為在上面運行的虛擬機提供一套虛擬的硬體環境
- 包括虛擬的處理器、記憶體、設備與I/O及網路介面



圖片來源:雲端策略 陳瑩

系統虛擬化(3/5)

- 應用最普遍的是運行與本機作業系統不相容的應用程式
  - 例如，一個用戶使用的是Windows系統的個人電腦，可以在個人電腦上虛擬出一個虛擬機，並在上面安裝Linux作業系統，以使用所需的應用程式



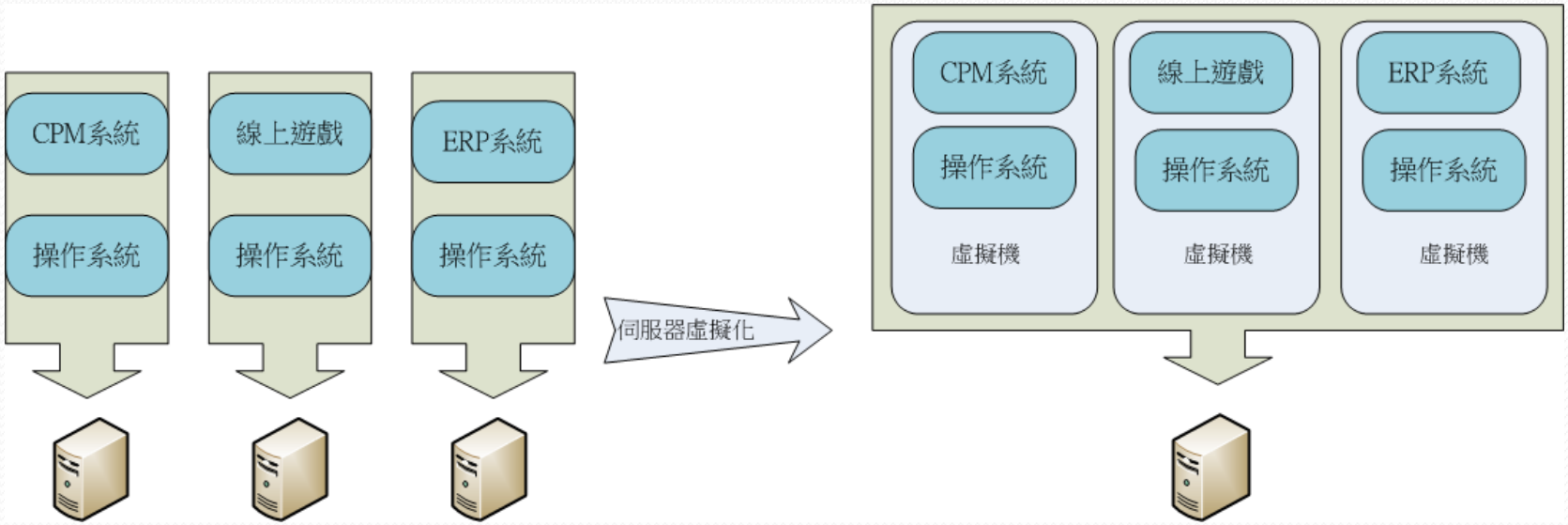
圖片來源:雲端策略 陳瑩

系統虛擬化(4/5)

- 系統虛擬化更大的價值在於伺服器虛擬化
- 伺服器基本只運行著一個應用服務，導致伺服器利用率低下
- 優點為
  - 提高伺服器的利用率
  - 減少機器數量
  - 降低營運成本
  - 節省實體儲存空間及電能

系統虛擬化的應用–伺服器虛擬化

- 將系統虛擬化技術應用在伺服器上，將一個實體伺服器虛擬成許多個伺服器來使用



圖片來源:雲端策略 陳瑩

伺服器虛擬化的典型實現

- 虛擬化軟體的不同實現方式，產生了兩種虛擬化軟體的專用術語
  1. 虛擬機監視器(Virtual machine monitor , VMM)
     - 負責對虛擬機提供硬體資源抽像化，為客系統提供運行環境
  2. 虛擬化平台(Hypervisor)
     - 負責虛擬機的託管和管理。
     - 直接運行在硬體上，因此運行直接受底層體系架構的規範

寄宿虛擬化(Hosted Architecture)

- 虛擬機監視器是運行在宿主作業系統(Host OS)上的應用程序，利用宿主作業系統的功能來實現硬體資源的抽象和虛擬機的管理
  - 需透過宿主作業系統來完成，效能通常較低
  - 容易實現

圖片來源:雲端策略 陳瑩

原生虛擬化(Bare Metal Architecture)

- 虛擬機運行在虛擬化平台上，虛擬化平台提供指令集和設備介面，以提供對虛擬機的支援
- 有較好的效能，但實作更為複雜

原生虛擬化

圖片來源:雲端策略 陳瑩

伺服器虛擬化之技術優勢

- 伺服器虛擬化技術具有以下優勢
  - 降低營運成本
  - 提高應用相容性
  - 加速應用佈署
  - 提高服務可用性
  - 提升資源利用率
  - 動態調度資源
  - 降低能源消耗

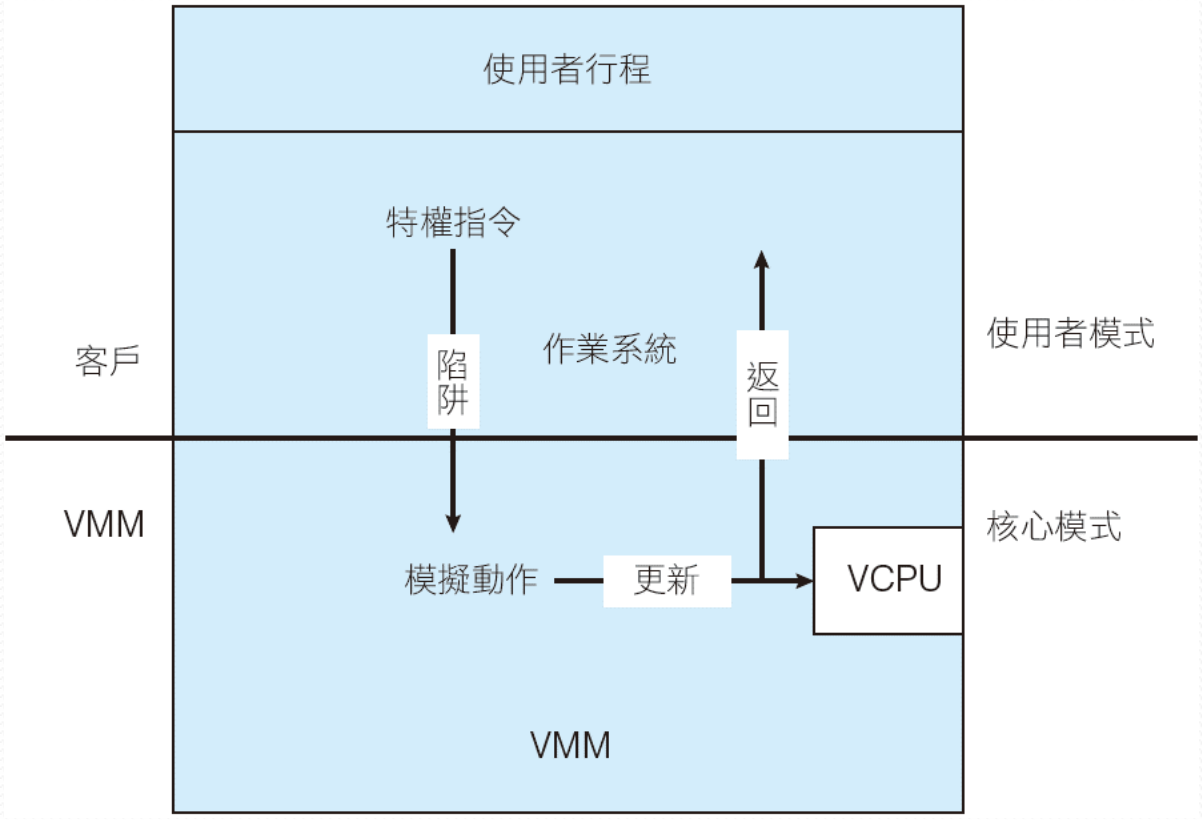
陷阱和模擬

- 在典型的雙模式系統上，虛擬機客戶只能在使用者模式下執行
  - 除非提供額外的硬體支援
- 虛擬機也必須具有兩種模式，因此我們必須具有虛擬使用者模式和虛擬核心模式，兩者均在實體的使用者模式下運行
  - 在真實電腦上導致從使用者模式轉換到核心模式的那些動作
    - 例如系統呼叫、中斷或嘗試執行特權指令
  - 必須在虛擬機中從虛擬使用者模式轉換為虛擬核心模式

陷阱和模擬

- 陷阱和模擬 (trap-and-emulate) 有了特權指令，時間就成了問題
- 例如 IBM VM 允許虛擬機的常用指令直接在硬體上執行
  - 只有特權指令 (主要是 I/O 所需) 才必須被模擬，因此執行起來會更慢
  - 通常隨著硬體的發展，陷阱和模擬功能的效能得到改善

實作陷阱及模擬虛擬化

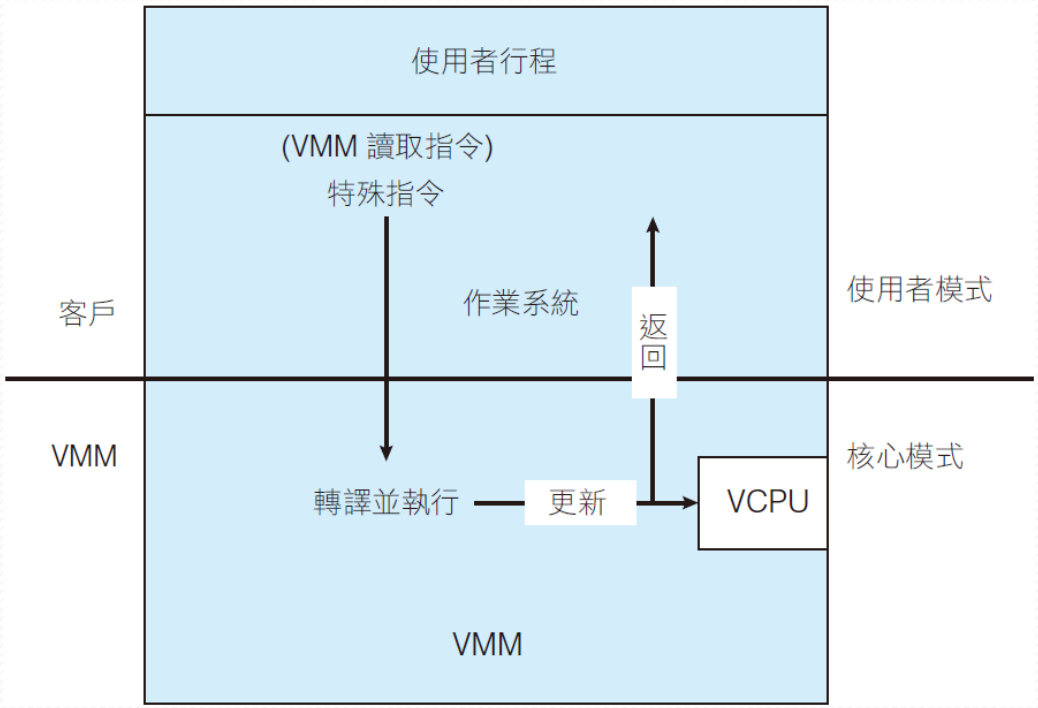


二進制轉譯

- 某些 CPU 不能完全區分特權和非特權指令
  - 對於虛擬化實現者而言，Intel x86 CPU 系列就是其中之一
    1. 如果客戶 VCPU 處於使用者模式，則客戶虛擬機可以在實體 CPU 上區域執行其指令
    2. 如果客戶 VCPU 處於核心模式，則客戶相信其正在核心模式下執行
       - VMM 透過客戶的程式計數器讀取客戶將要執行的後續幾個指令，來檢查客戶在虛擬核心模式下執行的每個指令
       - 除特殊指令外

二進制轉譯

- 二進制轉譯
  - 透過轉譯 VMM 中的程式碼，該程式碼自然地動態讀取從客戶來的二進制指令，按照需求生成二進制程式碼，並且在原始碼的地方執行



二進制轉譯

- 一個通用方法為陷阱和模擬和二進制轉譯一起使用，是使用**巢狀分頁表**（nested page tables, NPT）
  - 每個客戶作業系統維護一個或多個分頁表，以將虛擬記憶體轉換為實體記憶體
  - VMM 維護 NPT 來表示客戶的分頁表狀態，就像它建立 VCPU 時表示客戶的 CPU 狀態
  - VMM 知道客戶何時嘗試更改其分頁表，並在 NPT 中進行相同的更改
  - 當客戶在 CPU 上時，VMM 會將指針放置到適當的位置

二進制轉譯

- NPT 進入相應的 CPU 暫存器，以使該表成為活動分頁表如果客戶需要修改分頁表（例如分頁填寫錯誤）
  - 則該操作必須被 VMM 攔截，並適當對巢狀分頁表和系統分頁表做更改
- 使用 NPT 可能會導致 TLB 遺失率增加，並且需要解決許多其它複雜問題才能實現合理的效能

二進制轉譯

- NPT 進入相應的 CPU 暫存器，以使該表成為活動分頁表如果客戶需要修改分頁表（例如分頁填寫錯誤）
  - 則該操作必須被 VMM 攔截，並適當對巢狀分頁表和系統分頁表做更改
- 使用 NPT 可能會導致 TLB 遺失率增加，並且需要解決許多其它複雜問題才能實現合理的效能

硬體協助

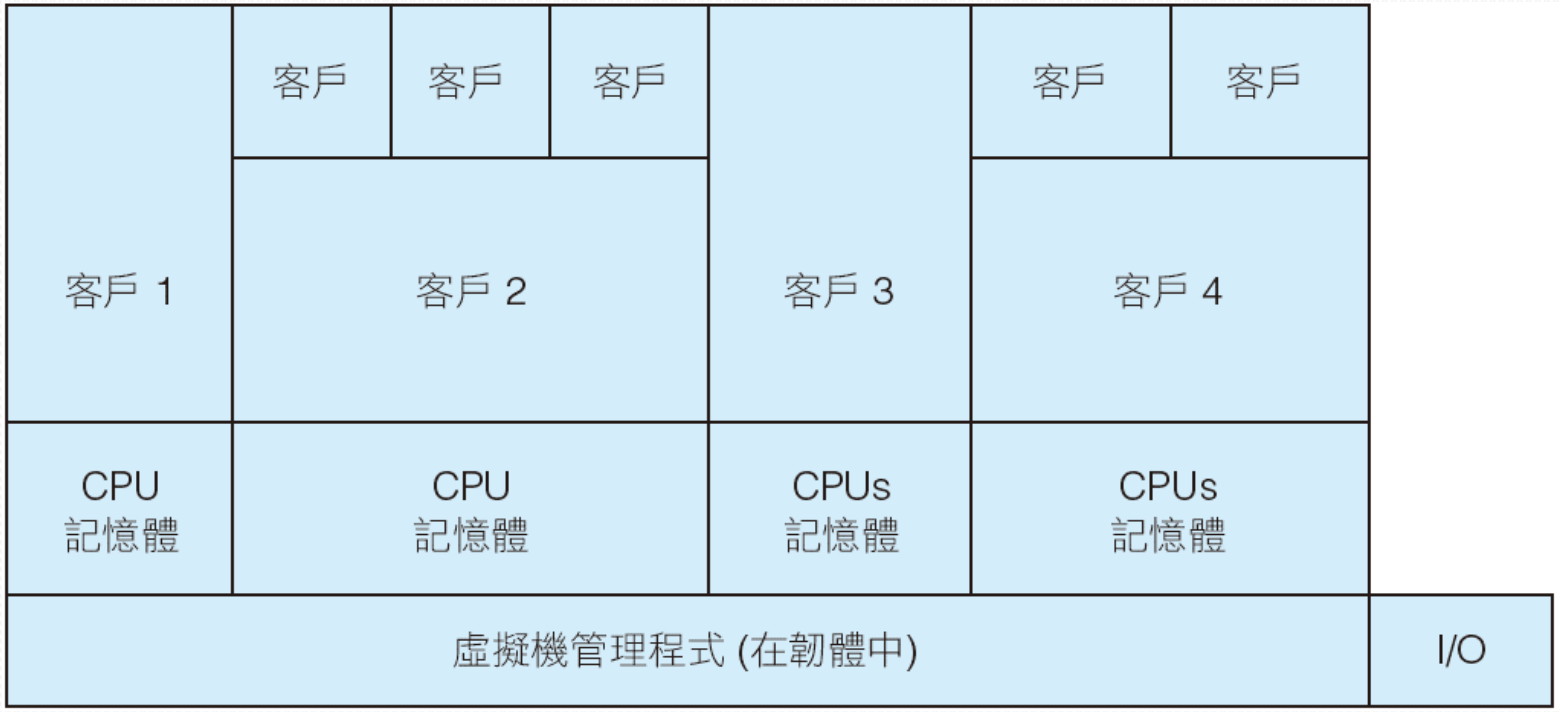
- 沒有某種程度的硬體支援，虛擬化的實現是不可能
  - 硬體支援越多，虛擬機的穩定性跟功能就越豐富
  - 在 Intel x86 CPU 家族裡，Intel 從 2005 年開始增加虛擬化的支援（VTX 指令），現在我們不再需要二進制轉譯
- 主要的通用 CPU 現在都為虛擬化提供擴展的硬體支援
  - 例如，AMD 虛擬化技術（AMD-V）從 2006 年開始出現在幾種 AMD 處理器中
  - 它定義兩個新的執作模式(主機和客戶)，從雙模式轉移到多模式處理器

硬體協助

- VMM 可以啟用主機模式，定義每個客戶虛擬機的特徵，然後將系統切換到使用者模式，從而將系統的控制權傳遞給在虛擬機中執行的客戶作業系統
- 在使用者模式下，虛擬化作業系統認為它正在本機硬體上運行，並且可以查看客戶主機定義中包含的任何裝置

- 兩者都提供客戶 VCPU 狀態資料結構，以在客戶作業系統內容轉換期間自動載入和保存客戶 CPU 狀態

類型 0 虛擬機管理程式


客戶
客戶
客戶
客戶
客戶
客戶 1
客戶 2
客戶 3
客戶 4
CPU
記憶體
CPU
記憶體
CPUs
記憶體
CPUs
記憶體
虛擬機管理程式 (在韌體中)
I/O

虛擬化和作業系統元件

虛擬化和作業系統元件

- 將深入研究虛擬化的作業系統方面
  - 包括 VMM 如何提供核心作業系統功能
    - 如排班、I/O 和記憶體管理
- VMM 如何排班 CPU 在客戶作業系統，認為它們具有專用 CPU 時使用？
- 當許多客戶需要大量記憶體時，記憶體管理如何工作？

關鍵特性與核心技術

- 伺服器虛擬化需要具備關鍵特性
  - 多實例
  - 隔離性
  - 封裝性
  - 高效能

- 伺服器虛擬化核心技術
  - CPU虛擬化
  - 記憶體虛擬化
  - 設備與I/O虛擬化
  - 即時搬移技術

CPU虛擬化 (1/2)

- CPU虛擬化
  - 把實體CPU抽象成虛擬，每個客系統可同時使用多個虛擬CPU
  - 各個客戶作業系統之間CPU運作互相隔離、互不影響

- x86架構的作業系統一般是被設計用於實體主機上，假設其擁有完整的底層實體硬體，於是在x86系統結構中要實現虛擬化，要在客系統層下加入虛擬化層，來實現實體資源的共用。

CPU虛擬化(2/2)

- 在x86體系中，CPU有四個運行級別，分別為Ring0~3，其中Ring0具有最高權限，而一般應用程式是運作在Ring3。

- 為了解決x86體系的虛擬化問題，業界提出了全虛擬化(Full-Virtualization)和半虛擬化(Half-Virtualization)兩種軟體方案。

全虛擬化(Full-Virtualization)

- 全虛擬化採用二進位碼動態組譯技術(Dynamic Binary Translation)解決特權指令的問題
- 虛擬機監視器會將指令動態轉換成相同功能的指令序再執行
- 優點
  - 程式碼的轉換動作動態完成，不用修改客戶作業系統，可以支援各種不同的作業系統
- 缺點
  - 動態轉換需要消耗一定的效能

半虛擬化(Half-Virtualization)

- 半虛擬化透過修改客戶作業系統來解決虛擬機特權指令的問題
- 在虛擬化平台上的客戶作業系統需修改作業系統，將所有敏感指令替換為對底層虛擬化平台的“超級呼叫(Hypercall)”，虛擬化平台也為這些敏感指令的特權指令提供呼叫介面

  - 在半虛擬化中，客戶作業系統和虛擬化平台必須要相容

CPU 排班

- 具有虛擬化功能的系統，甚至是單 CPU 系統，通常也像多處理器系統一樣工作
  - 虛擬化軟體向系統上執行的每個虛擬機提供一個或多個虛擬 CPU
  - 然後在虛擬機之間安排實體 CPU 的使用
- 執行緒可以是 VMM 執行緒或客戶執行緒。在建立時為客戶虛擬機配置一定數量的虛擬 CPU
- 以從使用者中竊取週期，但是此操作的影響相對較小。**過量使用**（overcommitment）的情況更加困難

CPU 排班

- 具有虛擬化功能的系統，甚至是單 CPU 系統，通常也像多處理器系統一樣工作
  - 虛擬化軟體向系統上執行的每個虛擬機提供一個或多個虛擬 CPU
  - 然後在虛擬機之間安排實體 CPU 的使用
- 執行緒可以是 VMM 執行緒或客戶執行緒。在建立時為客戶虛擬機配置一定數量的虛擬 CPU
- 以從使用者中竊取週期，但是此操作的影響相對較小。**過量使用**（overcommitment）的情況更加困難

CPU 排班

- 作業系統嘗試將 100毫秒分配給每個時間片段，以便為使用者提供合理的回應時間
  - 虛擬機內，此作業系統僅接收虛擬化系統為其提供的 CPU 資源
  - 一個 100 毫秒的時間片段可能會花費超過 100 毫秒的虛擬 CPU 時間
  - 時間片段可能需要 1 秒或更長時間，具體取決於系統的繁忙程度
    - 從而導致登入該虛擬機的使用者的回應時間非常短，對即時作業系統的影響可能更加嚴重

CPU 排班

- 為了解決這個問題，VMM 使應用程式可用於每種類型的系統管理員可以安裝到使用者中的作業系統
  - 這個應用程式可以校正時間偏差，並可以具有其它功能

記憶體虛擬化

- 記憶體虛擬化
  - 虛擬機監視器要能夠管理實體主機上的記憶體，並按每個虛擬機的需求來劃分實體記憶體，同時保持虛擬機對記憶體存放的互相隔離

- 為了在實體伺服器上能夠運行多個虛擬機，虛擬機監視器必須具備虛擬機記憶體管理單元
- 由於新增了一個虛擬機記憶體管理單元，在記憶體管理上與傳統的記憶體管理不同

記憶體管理

- VMM 的使用者不會失去對有效使用記憶體的額外需求
  - VMM 的使用者會採取廣泛措施來確保記憶體的最佳使用
- 例如 VMware ESX 使用多種記憶體管理方法
  - 在進行記憶體最佳化之前，VMM 必須確定每個客戶應使用多少實際記憶體
  - VMM 首先評估每個客戶的最大記憶體大小

記憶體管理

- 每個客戶計算目標實際記憶體分配，然後使用下面列出的三種低階機制從客戶中回收記憶體：
  1. 客戶認為它透過其分頁表管理來控制記憶體分配，而實際上 VMM 維護一個巢狀分頁表，該表將客戶分頁表轉換為實際分頁表
     - VMM 可以使用此額外的間接層級來最佳化客戶對記憶體的使用，而無須客戶的知識或幫助
     - 一種方法是提供雙分頁排班，在此 VMM 具有自己的分頁替換演算法，並將分頁載入客戶認為是實體記憶體的備份儲存中
     - VMM 對客戶的記憶體存取模式的瞭解要少於客戶，因此其分頁效率較低，從而導致效能問題
     - 當其它方法不可用或沒有提供足夠的可用記憶體時 VMM 會使用此方法，但這不是首選方法

記憶體管理

2. 常見的解決方案是讓 VMM 在每個客戶中安裝偽 VMM 控制的裝置驅動程式或核心模組
   - **偽裝置驅動程式**（pseudo-device driver）使用裝置驅動程式介面，在核心中似乎是裝置驅動程式，但實際上不控制裝置
   - 相反地，這是一種簡單在不直接修改核心的情況下，增加核心模式程式碼的方法
   - 氣球記憶體管理器（balloon memory manager）與 VMM 通信
   - VMM 將告知客戶內部的氣球行程取消固定，並釋放部份或全部記憶體，從而允許客戶使用更多分頁

記憶體管理

2. 常見的解決方案是讓 VMM 在每個客戶中安裝偽 VMM 控制的裝置驅動程式或核心模組
   - **偽裝置驅動程式**（pseudo-device driver）使用裝置驅動程式介面，在核心中似乎是裝置驅動程式，但實際上不控制裝置
   - 相反地，這是一種簡單在不直接修改核心的情況下，增加核心模式程式碼的方法
   - 氣球記憶體管理器（balloon memory manager）與 VMM 通信
   - VMM 將告知客戶內部的氣球行程取消固定，並釋放部份或全部記憶體，從而允許客戶使用更多分頁

記憶體管理

3. 降低記憶體壓力的另一種常用方法是，讓 VMM 確定同一分頁是否已載入多次
   - 在這種情況下，VMM 會將分頁的副本數減少為一個
   - 並將該分頁的其它使用者映射到該使用者複製

記憶體管理

3. 降低記憶體壓力的另一種常用方法是，讓 VMM 確定同一分頁是否已載入多次
   - 在這種情況下，VMM 會將分頁的副本數減少為一個
   - 並將該分頁的其它使用者映射到該使用者複製

設備與I/O虛擬化

- 實體主機的設備虛擬化
  - 虛擬化平台將這些設備標準化為一系列虛擬設備，為虛擬機提供一個可以使用的虛擬設備集合

- 經過虛擬化的設備並不一定與實體設備的型號、組態、參數等完全相符。
- 虛擬化設備能有效的模擬實體設備的動作，將虛擬機設備操作組譯給實體設備，並將實體設備的組譯結果送回虛擬機

I/O

- 在 I/O 範圍中，虛擬機管理程式具有一定的彈性，因此不必擔心它們如何為客戶重現底層硬體
  - 因為在 I/O 裝置的廣泛變化中，作業系統用於處理具有各種靈活的 I/O 機制
    - 例如作業系統的裝置驅動程式機制，不論 I/O 裝置如何，都能提供作業系統統一介面
  - 裝置驅動程式介面的目的在於，讓第三方硬體製造商提供連接他們的裝置驅動程式到作業系統
  - 通常，裝置驅動程式可以動態載入和卸載
  - 虛擬化利用這種內置的靈活性，向客戶作業系統提供特定的虛擬化裝置

I/O

- 客戶可以經由 IP 位址“直接”連接到網路［這稱為**橋接**（bridging)]，或者 VMM 可以提供網路位址轉換（network address translation, NAT）位址
- NAT 位址對於執行客戶的伺服器和 VMM 是區域的，且提供更廣泛的網路和客戶之間的路由
- VMM 也提供防火牆，以保護系統內客戶之間的連接及客戶和外部系統

儲存管理

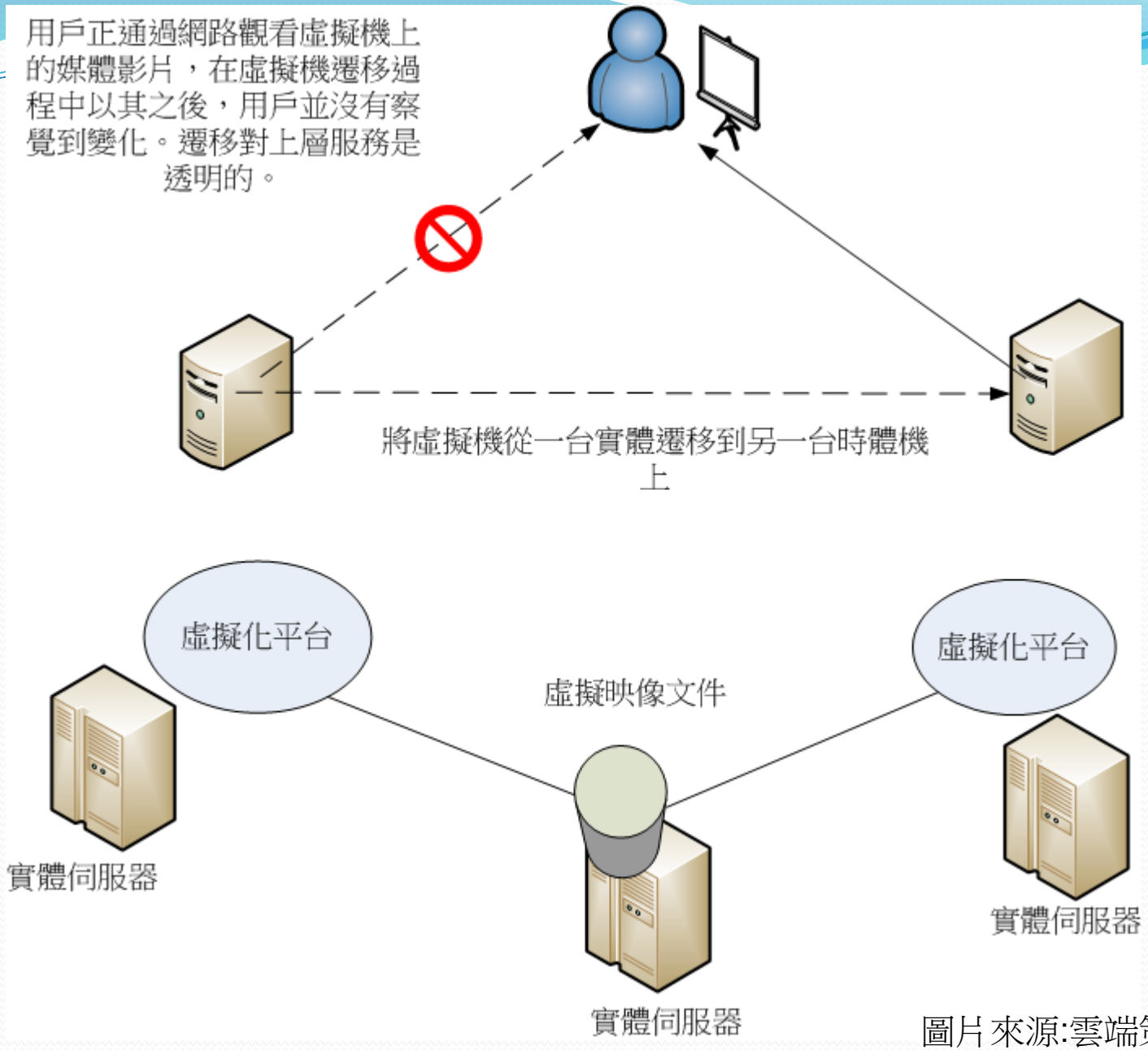
- 類型 1 虛擬機管理程式儲存客戶 root 磁碟（跟配置資訊）在一個或多個 VMM 提供的檔案系統上
- 類型 2 虛擬機管理程式將相同的資訊儲存在主機作業系統的檔案系統中
- 本質上，**磁碟映像**（disk image）包含所有 root 磁區的內容，被包含在 VMM 的一個檔案中

儲存管理

- **實體到虛擬**（physical-to-virtual, P-to-V）的轉換，將讀取以下的硬碟區塊：
  - 實體系統的磁碟，並將它們儲存在 VMM 的系統上
    - 或 VMM 可以存取的共用儲存中的檔案
  - VMM 還提供從**虛擬到實體**（virtual-tophysical, V-to-P）的過程，可將客戶轉換為實體系統

即時搬移技術

- 即時搬移(Live Migration)技術
  - 將整個虛擬機器的運行狀態完整、快速的從原來所在的宿主機硬體平台搬移到新的宿主機硬體平台上
  - 搬移過程平順無縫，用戶不會察覺到任何差異
  - 最大的用途
    - 資料中心在伺服器做維護時，可將用戶轉移到別處繼續使用，而不用造成服務的停止


用戶正通過網路觀看虛擬機上的媒體影片，在虛擬機遷移過程中以其之後，用戶並沒有察覺到變化。遷移對上層服務是透明的。
將虛擬機從一台實體遷移到另一台時體機上
虛擬化平台
虛擬映像文件
虛擬化平台
實體伺服器
實體伺服器
實體伺服器

圖片來源:雲端策略 陳瑩

即時遷移

- 每個客戶和 VMM 之間的介面定義明確，並且 VMM 為客戶保留的有限狀態，使即時遷移成為可能
  - VMM 透過以下步驟遷移客戶：
    1. 來源 VMM 與目標 VMM 建立連接，並確認允許其發送客戶
    2. 目標透過建立新的 VCPU、新的巢狀分頁，來建立新的客戶表和其它狀態儲存
    3. 來源將所有唯讀記憶體分頁發送到目標
    4. 來源將所有唯讀分頁發送到目標，將它們標記為空

即時遷移

5. 來源重複步驟 4，因為在此步驟中某些分頁可能是客戶修改的，現在資料可能有問題，因此這些分頁需要再次發送，並再次標記為空
6. 當步驟 4 和 5 的週期變得非常短時，來源 VMM 凍結客戶作業系統，發送 VCPU 的最終狀態（其它狀態詳細資訊），以及最後的分頁，並告訴目標開始運行客戶虛擬機

- 一旦目標確認客戶正在執行，便終止客戶

虛擬化技術原理

虛擬化技術原理

- 有非常多有關於虛擬化名詞:

  Java VM、Dalvik VM、Apple Rosetta、Transmeta Crusoe、VMWare ESXi、VMWare Workstation、Xen、KVM、Parallels Desktop … etc.

- 以上名詞都是屬於虛擬機(Virtual Machine, VM), 差別在於不同工作性質，有些屬於Process VM，有些屬於System VM！

Process VM VS. System VM

- Process VM
  - 只虛擬一個Process 的行為，可能是為了支援舊架構上的應用程式
  - e.g., Apple Rosetta；也可能是為了跨平台上執行應用程式e.g., Java 。
- System VM
  - 虛擬的是一個完整的OS，需要更複雜的處理，
  - VMWare ESX Server，Hyper-V R2, Xen, KVM, and Citrix Xen Server。

虛擬化技術的產品

虛擬化技術的產品

- PC虛擬機
  - Virtual PC
  - Virtual Box
  - VMware Workstation
- 伺服器虛擬機
  - Xen (Xen Hypervisor)
  - KVM (Kernel-Based Virtual machine)
  - XenServer (Citrix XenServer)
  - VMware vSphere 4 (VMware ESX Server) (see another topic)
  - MS Hyper-V R2 (Windows Server 2008 R2) (see another topic)
  - OpenNebula (OpenNebula.org)
  - OpenStack (Nova-Swift-Glance-Keystone-Horizon)
  - Proxmox Virtual Environment (Proxmox VE)

Virtual PC

- 原本「Virtual PC」跟「VMWare」一樣，都是付費軟體，不過自從微軟將Virtual PC買下之後，就變成免費軟體了。
- 官方網站：
  - https://www.microsoft.com/zh-tw/download/details.aspx?id=3702

Virtual Box

- 軟體資訊
  - 軟體名稱：Virtual Box
  - 軟體版本：5.1
  - 軟體語言：中文（內建多國語言）
  - 軟體性質：免費軟體
  - 系統支援：Windows7/XP/2003/Vista/Mac OS
  - /Linux/Solaris
  - 官方網站：http://www.virtualbox.org
- 目前Virtual Box 可支援的客戶端系統(可安裝在Virtual Box裡面的作業系統)有十幾種，幾乎各種常見的x86作業系統都可以支援

VMware Workstation

- 經由VMware Workstation，可以發現PC或Notebook的真正功能和靈活性。通過一台PC 上同時運行多個作業系統，可將系統所需的硬體總成本降低50% 或更低，另外使用自動執行和任務優化，可以節省時間並提高執行效率。
- 官方網站：
  - http://www.vmware.com/tw/products/workstation.html
- VMware Workstation具有四項優點
  1. 代管舊版應用程式並克服平台遷移問題
  2. 在隔離的環境中配置和測試新軟體或修補程式
  3. 自動執行開發軟體和測試任務
  4. 在單台PC 上展示多層配置

動腦時間

- 請說明虛擬化技術的用途?

END!