

高效能運算虛擬化應用

Machine Learning / Deep Learning

資深技術顧問

羅元佃, Louis Lo

December 22, 2017

vmware®

© 2017 VMware Inc. All rights reserved.

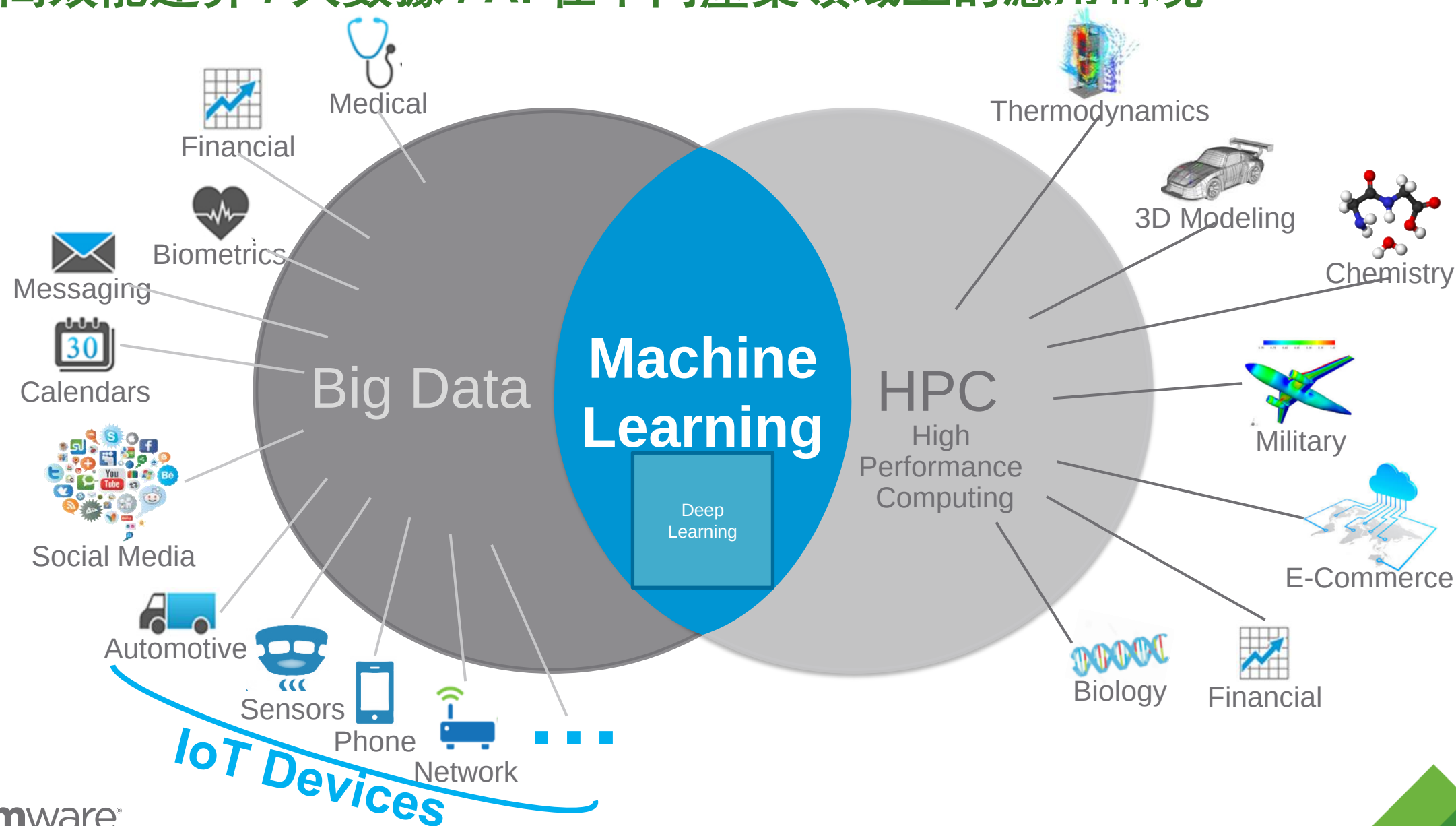
議程

-
- 1 高效能運算虛擬化
 - 2 Deep Learning 實測案例
 - 3 VMware 軟體定義資料中心
-

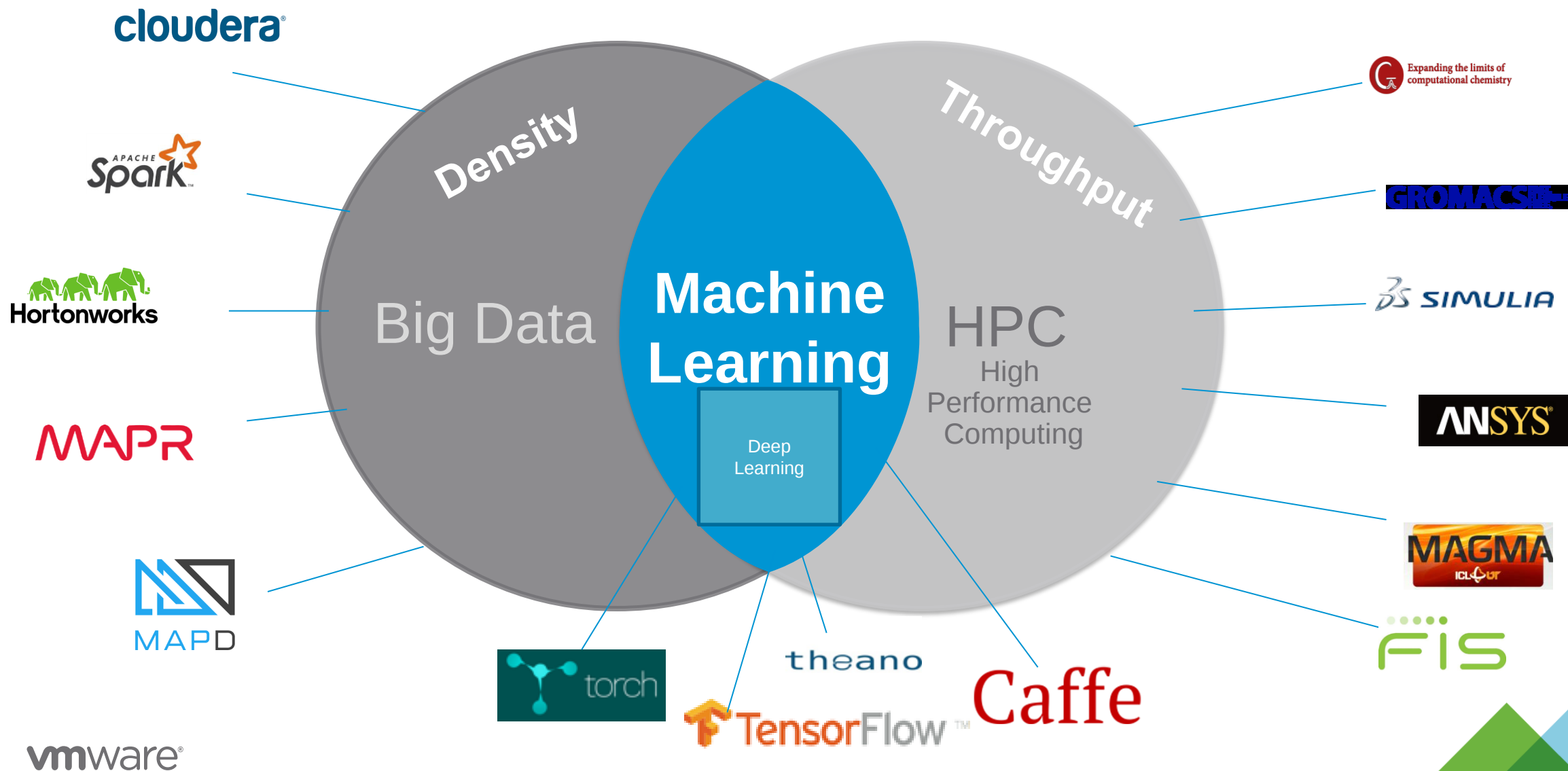
議程

-
- 1 高效能運算虛擬化
 - 2 Deep Learning 實測案例
 - 3 VMware 軟體定義資料中心
-

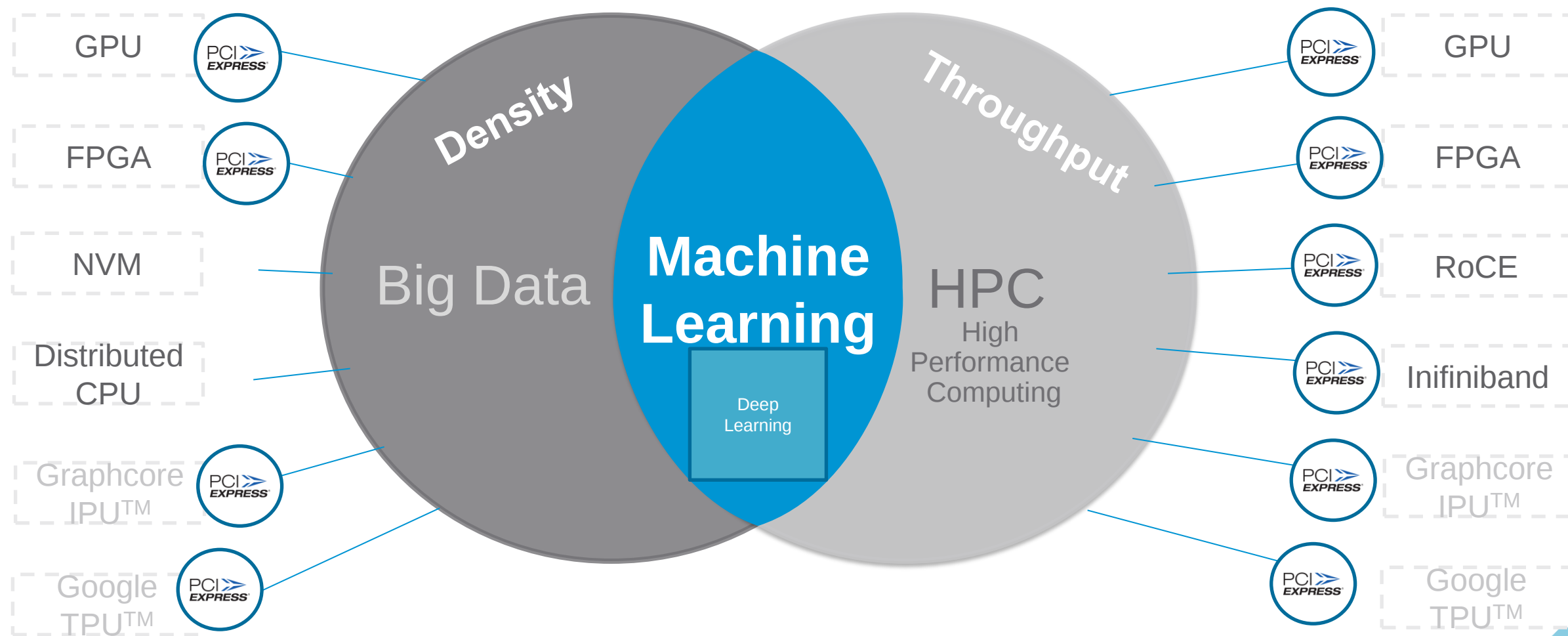
高效能運算 / 大數據 / AI 在不同產業領域上的應用情境



高效能運算 / 大數據 / AI 在不同應用情境所使用的應用工具



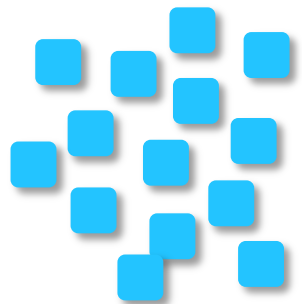
高效能運算 / 大數據 / AI 在不同產業領域上相關應用所需使用的資源



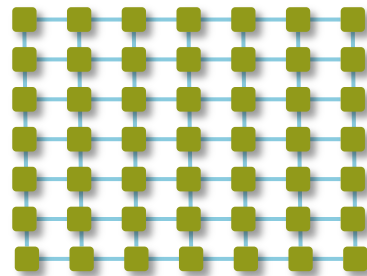
HPC Workloads

- 科學/研究型計算場景
- 通常需要更多浮點運算
- 通常需要高效IO
- 常見分散/平行運算架構

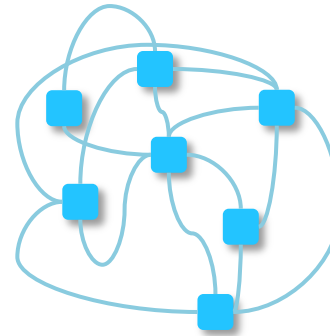
- Mechanical Design / Drafting
- Chemical Engineering
- Economics/Financial
- Weather
- Electronic Design Automation (EDA)
- Geosciences
- Defense
- Computer-Aided Engineering (CAE)
- Bioscience
- Government Lab
- University/Academic



Throughput
Jobs



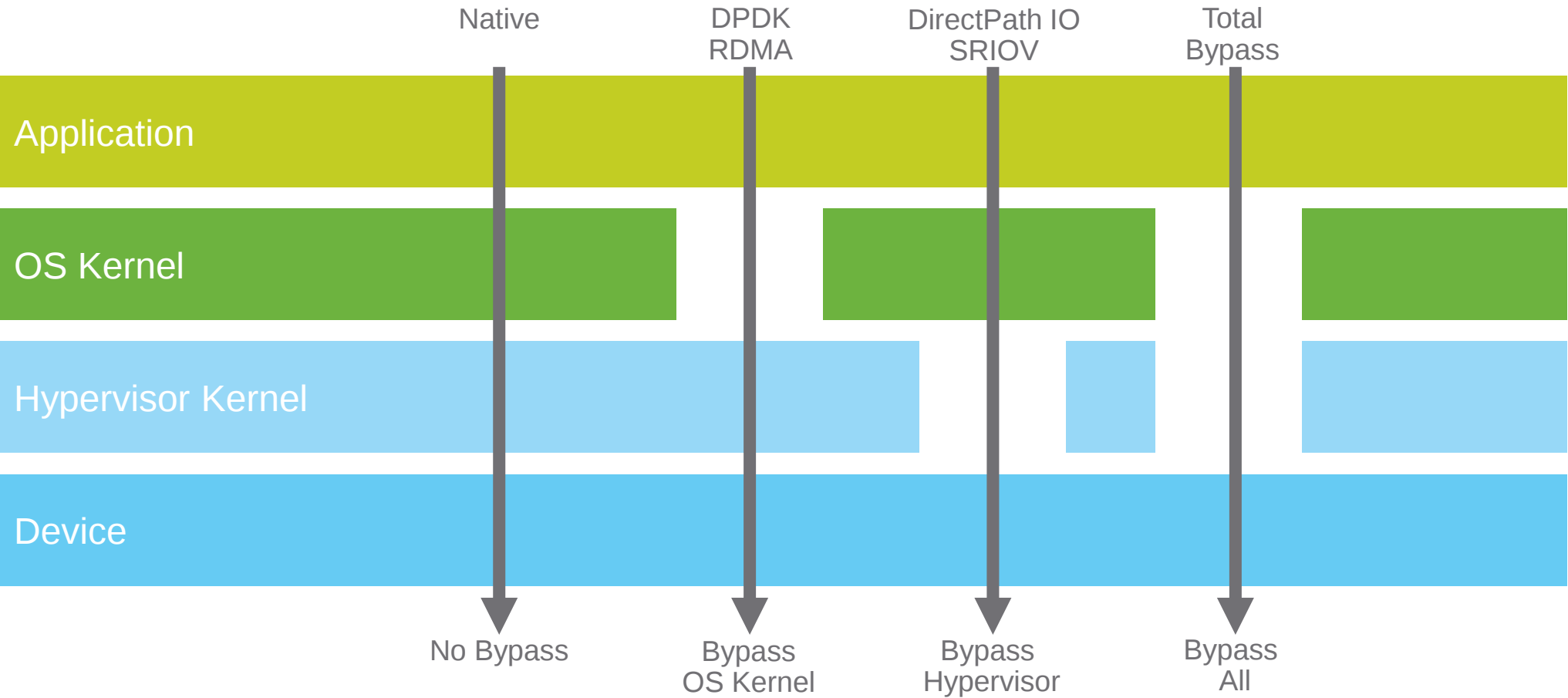
HPC
Cluster



MPI
Jobs

Message
Passing
Interface

超高效能關鍵：直接硬體存取



虛擬機能夠因應不同的應用，配置不同資源符合不同的情境使用


Simulation
Engineer


Machine
Learning
Expert


Data
Scientist


Developer


App
Admin



- ✓ Websites
- ✓ Dev & Test, Staging
- ✓ LoB Applications
- ✓ Small to mid-size DB
- ✓ SAP, Oracle, SharePoint
- ✓ Cluster computing
- ✓ Enterprise applications

vCPU

Mem

vNIC

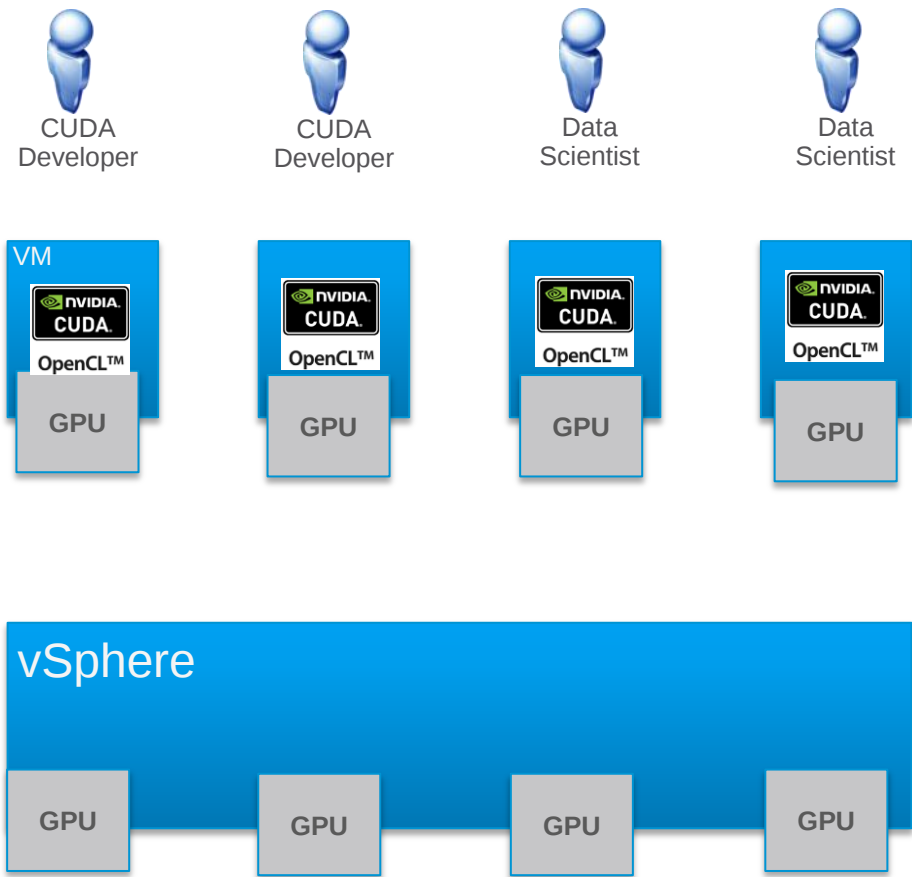
VMFS
or
VSAN


VI Admin

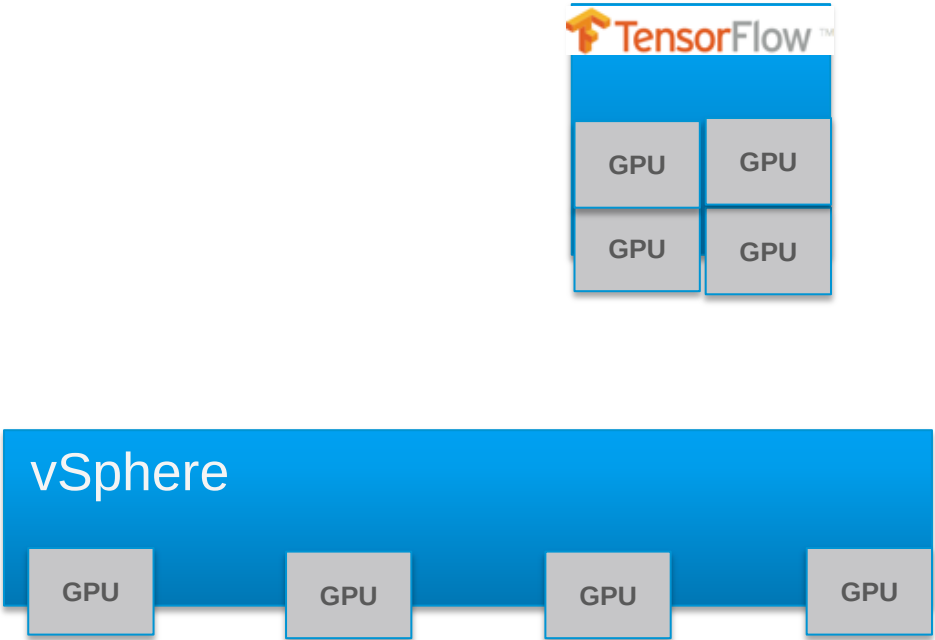
於 ML/DL 應用方式一

vSphere 運用 DirectPath IO 使用 GPU 資源

1:1

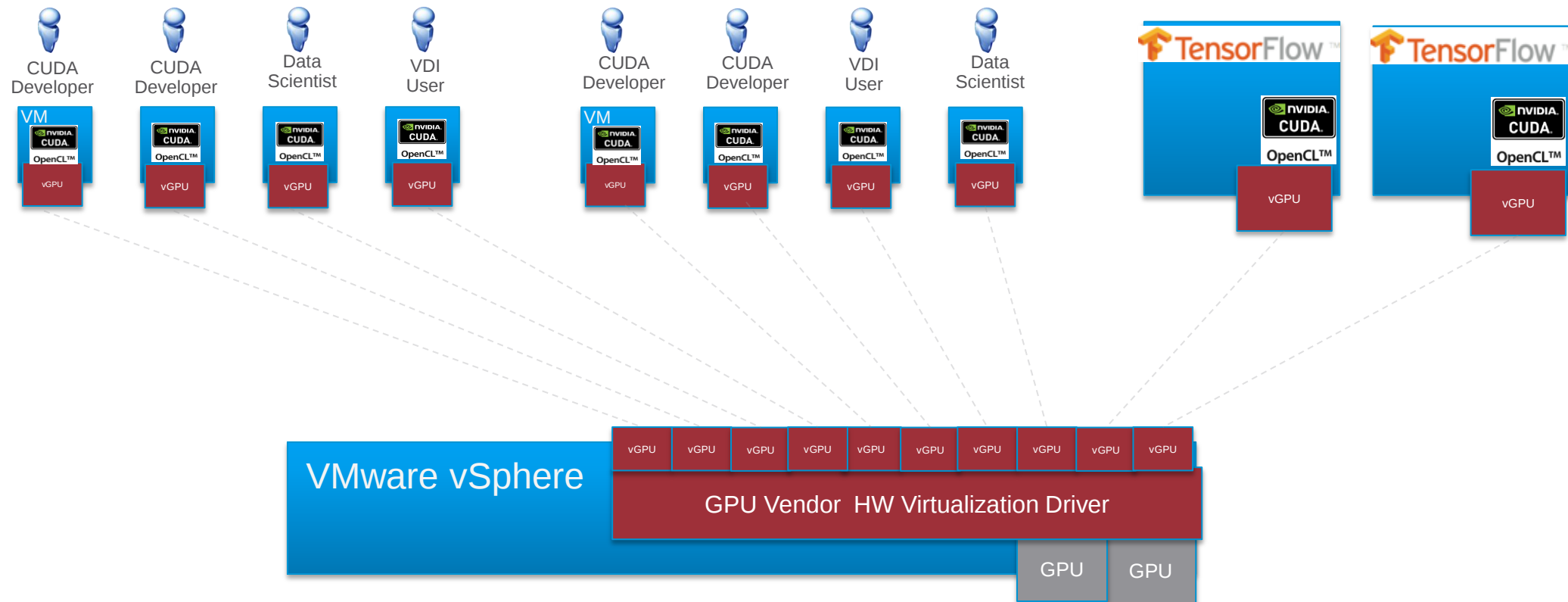


X:1



於 ML/DL 應用方式二

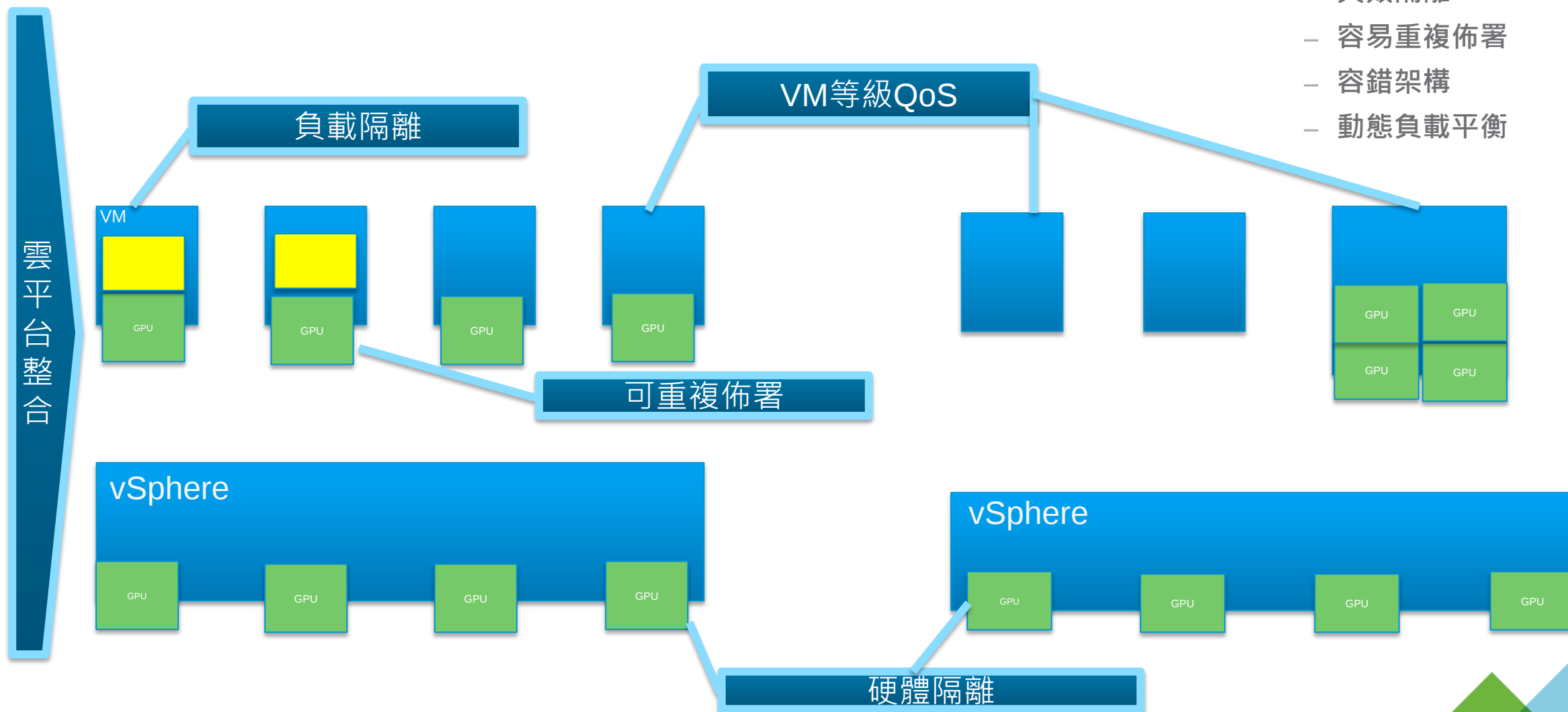
vSphere 使用 Pass-through 方式 – SRIOV or NVIDIA GRID



高效能運算進行虛擬化達到的效益

Virtual Machines 提供:

- 同質性平台
- 多租戶資料保護
- 失效隔離
- 容易重複佈署
- 容錯架構
- 動態負載平衡



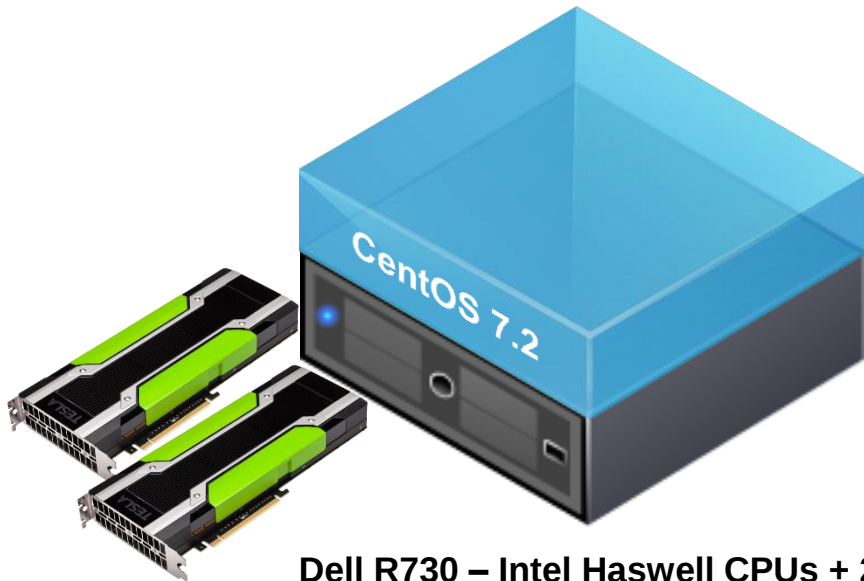
議程

-
- 1 高效能運算虛擬化
 - 2 Deep Learning 實測案例
 - 3 VMware 軟體定義資料中心
-

實測環境比較 - Native GPU vs Virtual GPU Comparison

Native Configuration

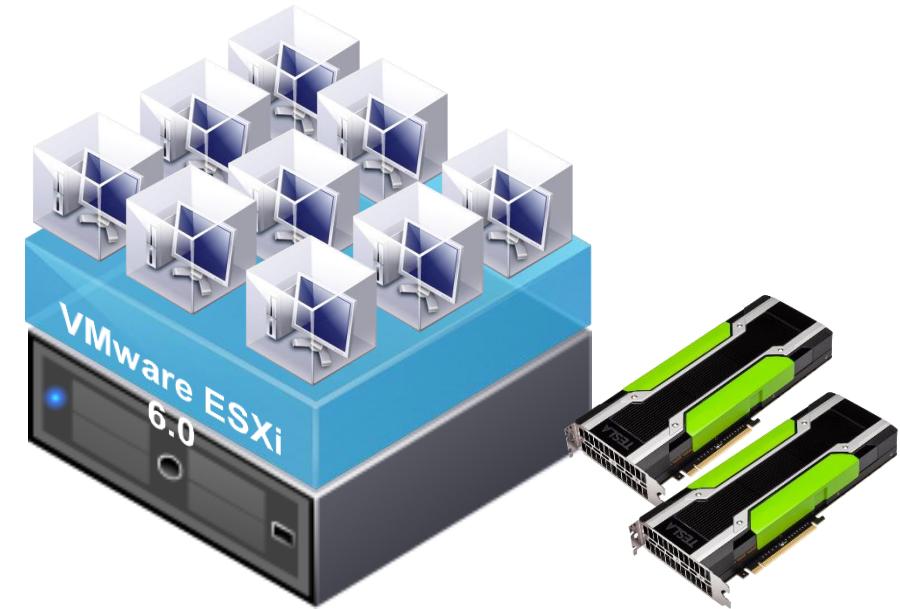
- TensorFlow 0.10
- CentOS 7.2



Dell R730 – Intel Haswell CPUs + 2 x NVidia GRID M60
24 cores (2 x 12-core socket) E5-2680 V3
768 GB RAM

Virtual Configuration

- TensorFlow 0.10
- CentOS 7.2
- ESX 6.X



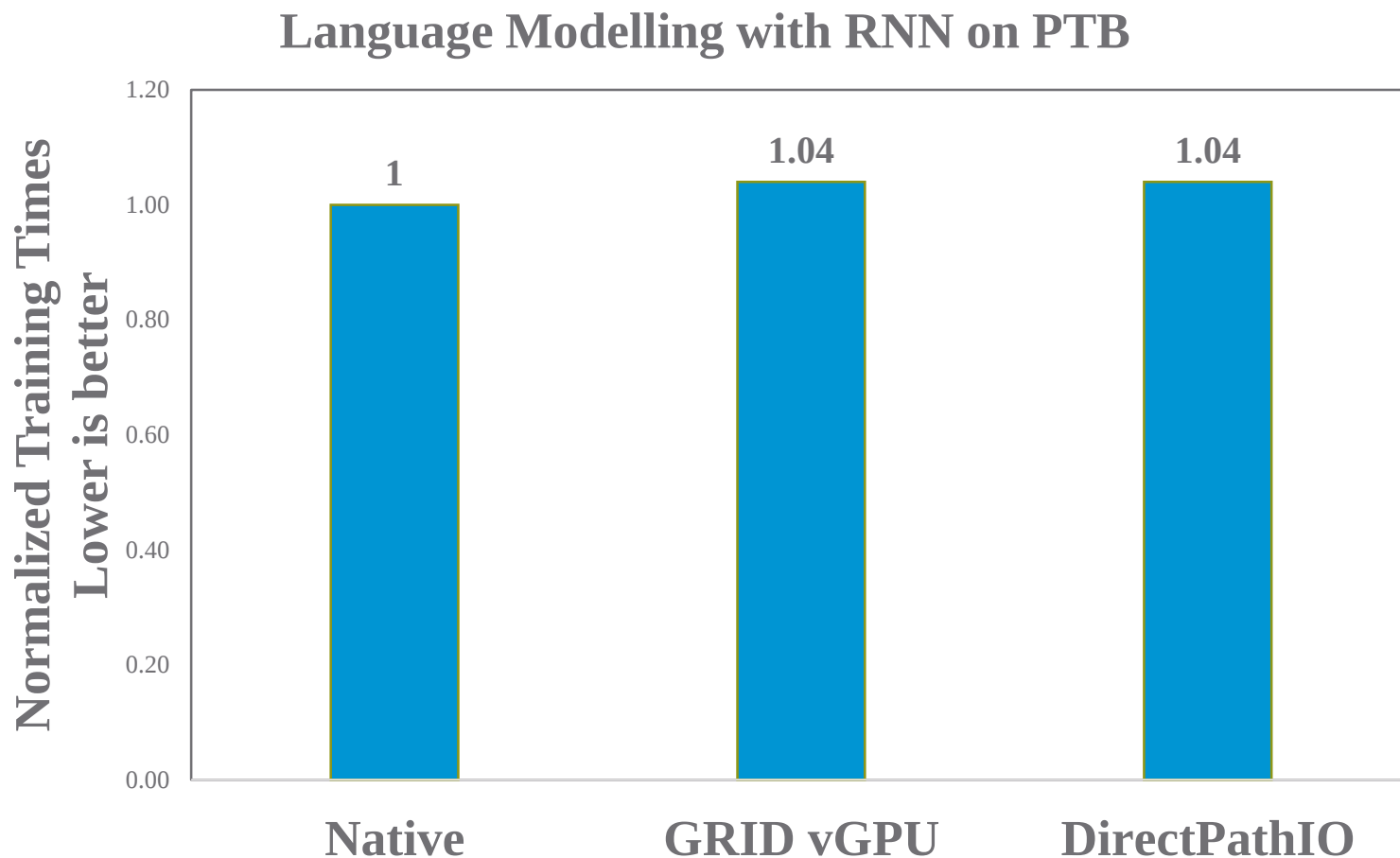
Dell R730 – Intel Haswell CPUs + 2 x NVidia GRID M60
24 cores (2 x 12-core socket) E5-2680 V3
768 GB GB RAM

測試過程與驗證方式 - Native GPU vs Virtualized GPU

- **Workload**
 - Complex Language Modelling
 - Given history of words, predicts next word
- **Neural Network Type: Recurrent Neural Network**
 - Large Model
 - 1500 LSTM units /layer
 - Penn Tree Bank (PTB) Database:
 - 929K training words
 - 73K validation words
 - 82K test words
 - 10K vocabulary

實測結果: Training Times on native GPU vs virtualized GPU

4% of overhead for both GRID vGPU & DirectPath I/O compared to native GPU



硬體映射高階加速卡配置方式

VMDirectPath I/O for High-End Devices on vSphere 6.5u1

- Special settings needed for large-BAR PCI devices (16GB+)

- Examples: nVidia P100, nVidia K80, nVidia K40m, Intel Xeon Phi, FPGA

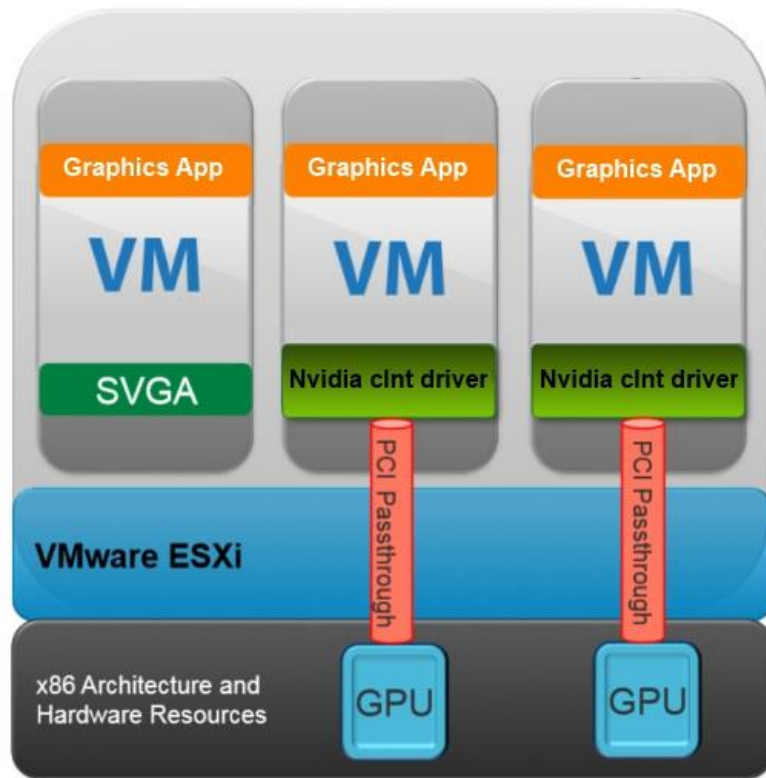
```
2017-03-07T07:40:38.467Z| vmx| I125: PCIPassthru: Device 0000:09:00.0 barIndex 0 type 2 realaddr 0xc6000000 size 16777216 flags 12
2017-03-07T07:40:38.467Z| vmx| I125: PCIPassthru: Device 0000:09:00.0 barIndex 1 type 3 realaddr 0x3b80000000 size 17179869184 flags 12
2017-03-07T07:40:38.467Z| vmx| I125: PCIPassthru: Device 0000:09:00.0 barIndex 3 type 3 realaddr 0x3bc0000000 size 33554432 flags 12
```

- BIOS: Enable “above 4GB decoding” or similar, depending on vendor
- Enable EFI boot and perform EFI guest OS installation
- VMX file entries:
 - `pciPassthru.use64bitMMIO= "TRUE"`
 - `pciPassthru.64bitMMIOSizeGB = "NN"`
 - `NN = ROUND-UP-TO-NEXT-POWER-OF-TWO( (# high-end passthrough devices) * 16 )`
 - Two-device example: `NN = ROUND-UP-TO-NEXT-POWER-OF-TWO(32) = 64`
- Enable Passthrough devices in ESX and VM, as usual. For details and further help:

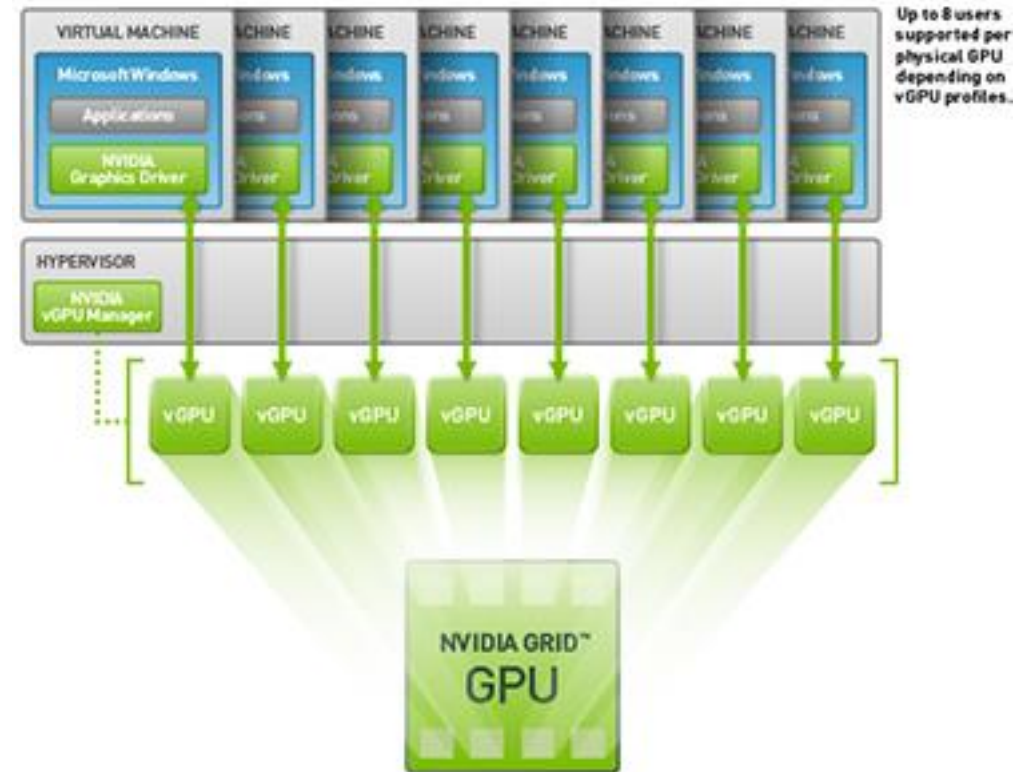
<https://octo.vmware.com/enable-compute-accelerators-vsphere-6-5-machine-learning-hpc-workloads/>

GPU加速方式

vDGA / VMDirectPath IO / PCI Passthrough



vGPU

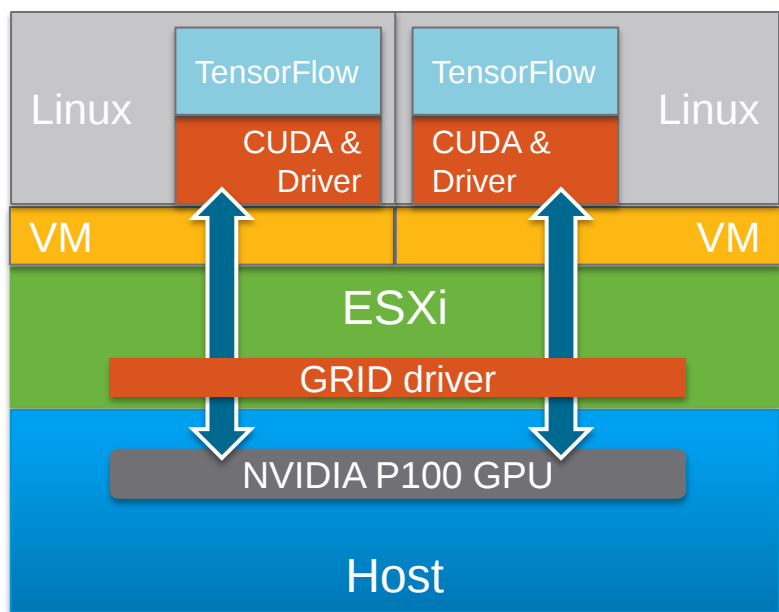


NVIDIA 目前對於 vSphere 的相容性認證

Package	Hypervisor/BM OS	SW Product Deployment	HW Supported ¹	Guest OS Support
GRID for vSphere 6.5	ESXi 6.5	vGPU, vDGA, vSGA ³	M6, M10, M60, P4, P6, P40, P100 ⁴ Since 5.1: P100 12GB⁴	Win 7 (32/64-bit), Win8/8.1 (32/64-bit), Win 10 (32/64-bit)
				WinServer 2008 R2, WinServer 2012 R2, WinServer 2016
				RHEL 6.6/7.x (64-bit), Ubuntu 14.04 LTS/16.04 LTS (64-bit)
GRID for vSphere 6.0	ESXi 6.0 ⁵	vGPU, vDGA, vSGA ³	M6, M10, M60, P4, P6, P40, P100 ⁴ Since 5.1: P100 12GB⁴	Win 7 (32/64-bit), Win8/8.1 (32/64-bit), Win 10 (32/64-bit)
				WinServer 2008 R2, WinServer 2012 R2, WinServer 2016
				RHEL 6.6/7.x (64-bit), Ubuntu 14.04 LTS/16.04 LTS (64-bit)

<http://docs.nvidia.com/grid/5.0/product-support-matrix/index.html>

高階加速卡除直接映射外，也支援GPU共享



- TensorFlow RNN
- SuperMicro dual 12-core system
- 16GB NVIDIA P100 GPU
- Two VMs, each with an 8Q GPU profile
- NVIDIA GRID 5.0
- ESXi 6.5

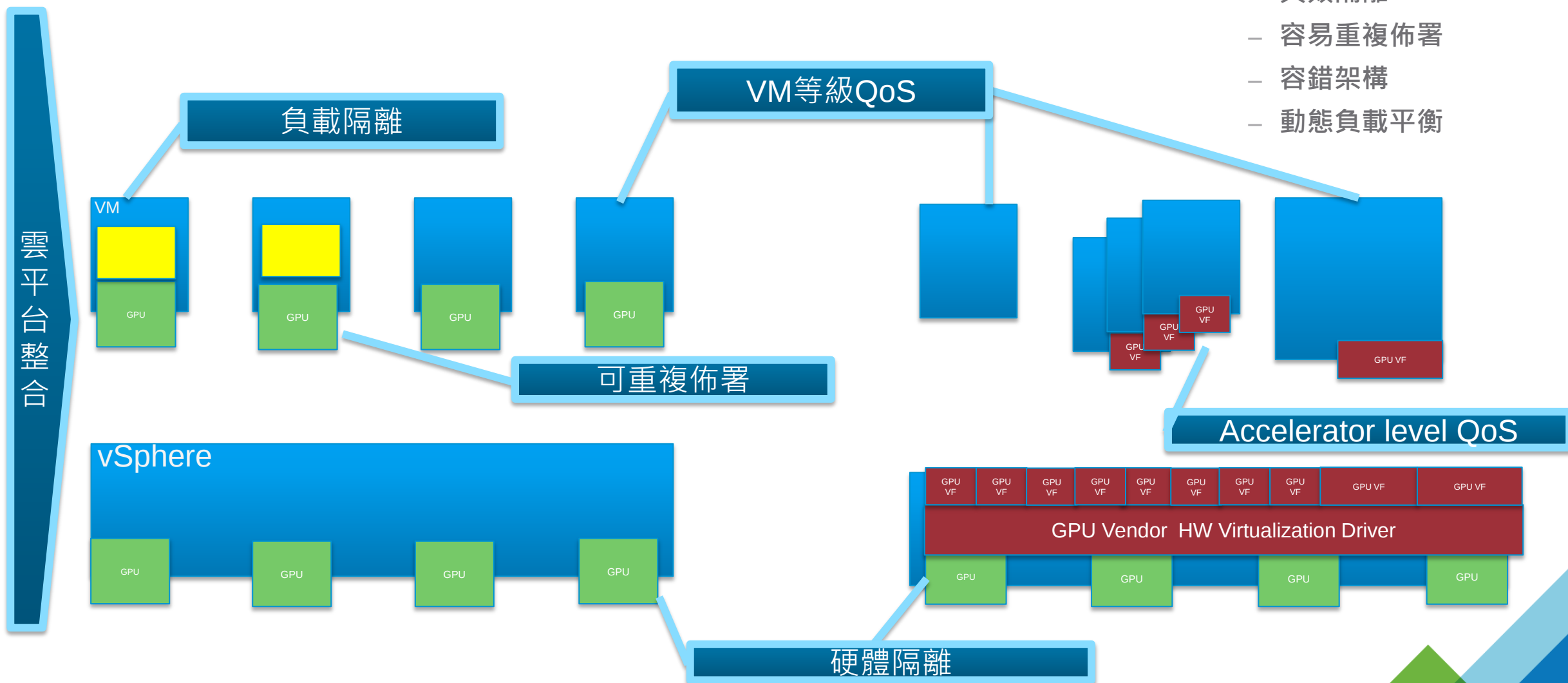
Scheduling policies:

- Best effort
- Equal share
- Fixed share

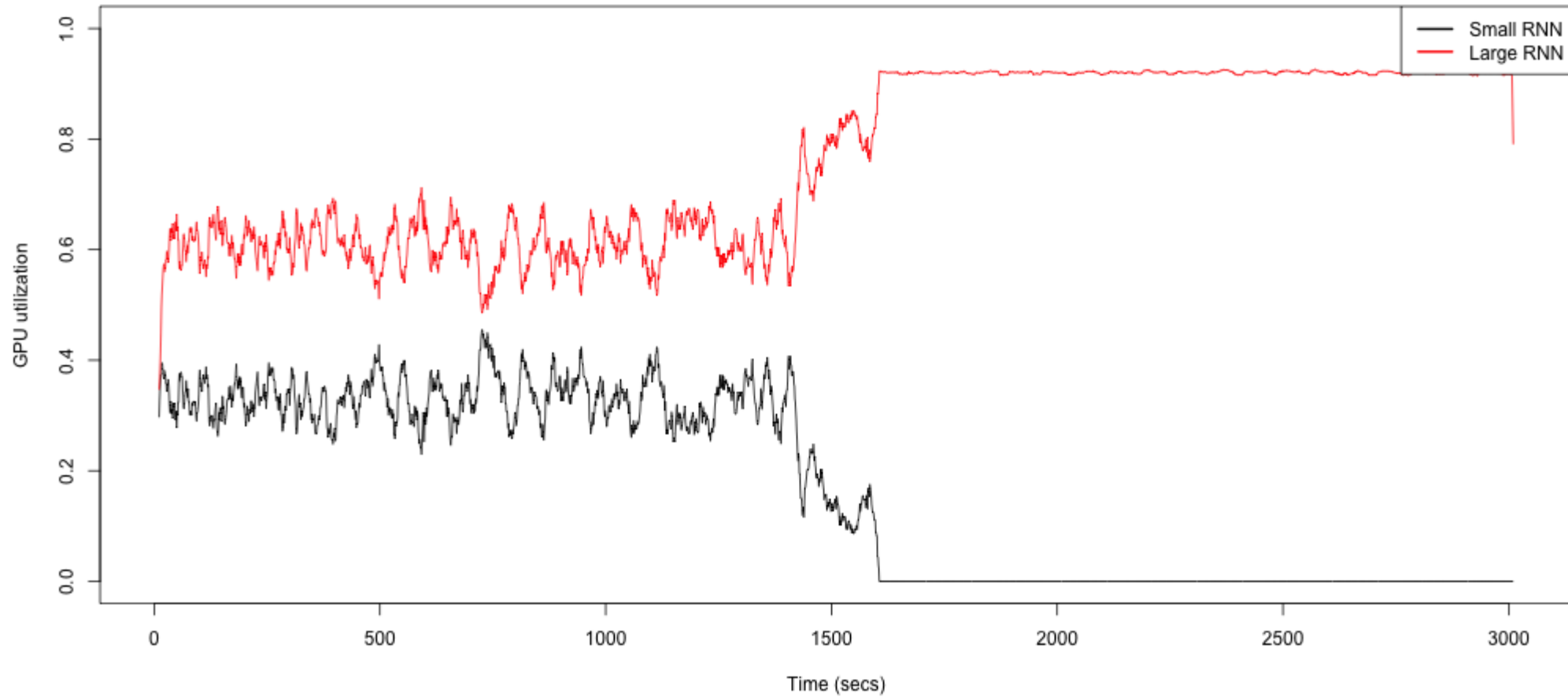
GPU 虛擬化效益

Virtual Machines 提供:

- 同質性平台
- 多租戶資料保護
- 失效隔離
- 容易重複佈署
- 容錯架構
- 動態負載平衡



GPU共享效能實測



Single P100, two 8Q VMs, Best effort scheduler

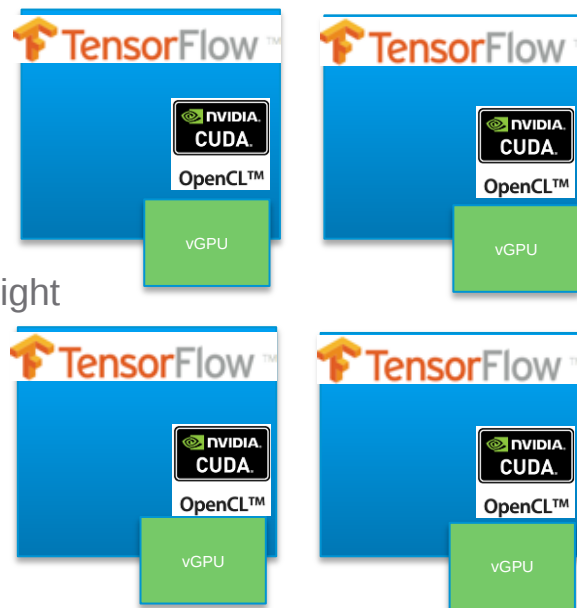
GPU虛擬化效益-GPU共享，可依據時間調配資源分配



Develop/VDI/
Inference by day

ML Training by night

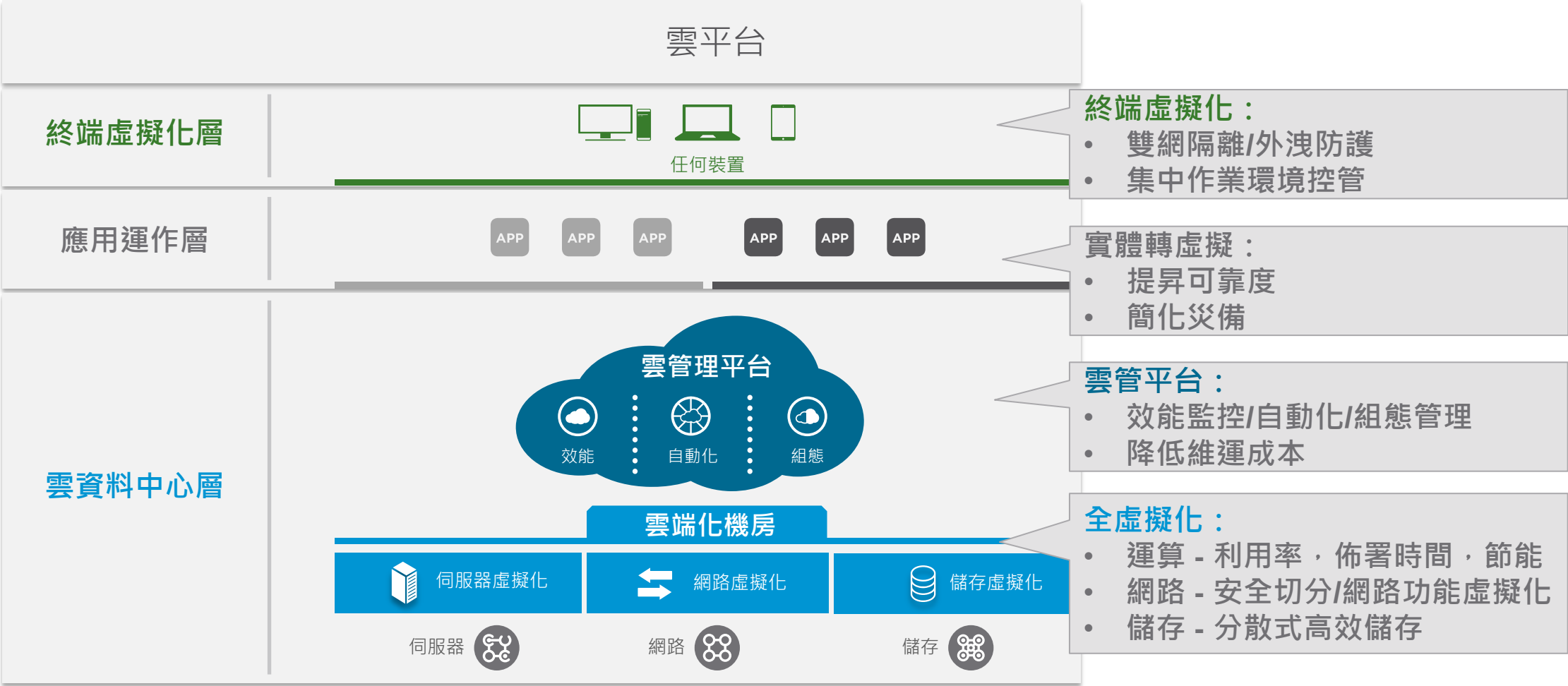
Same Infrastructure



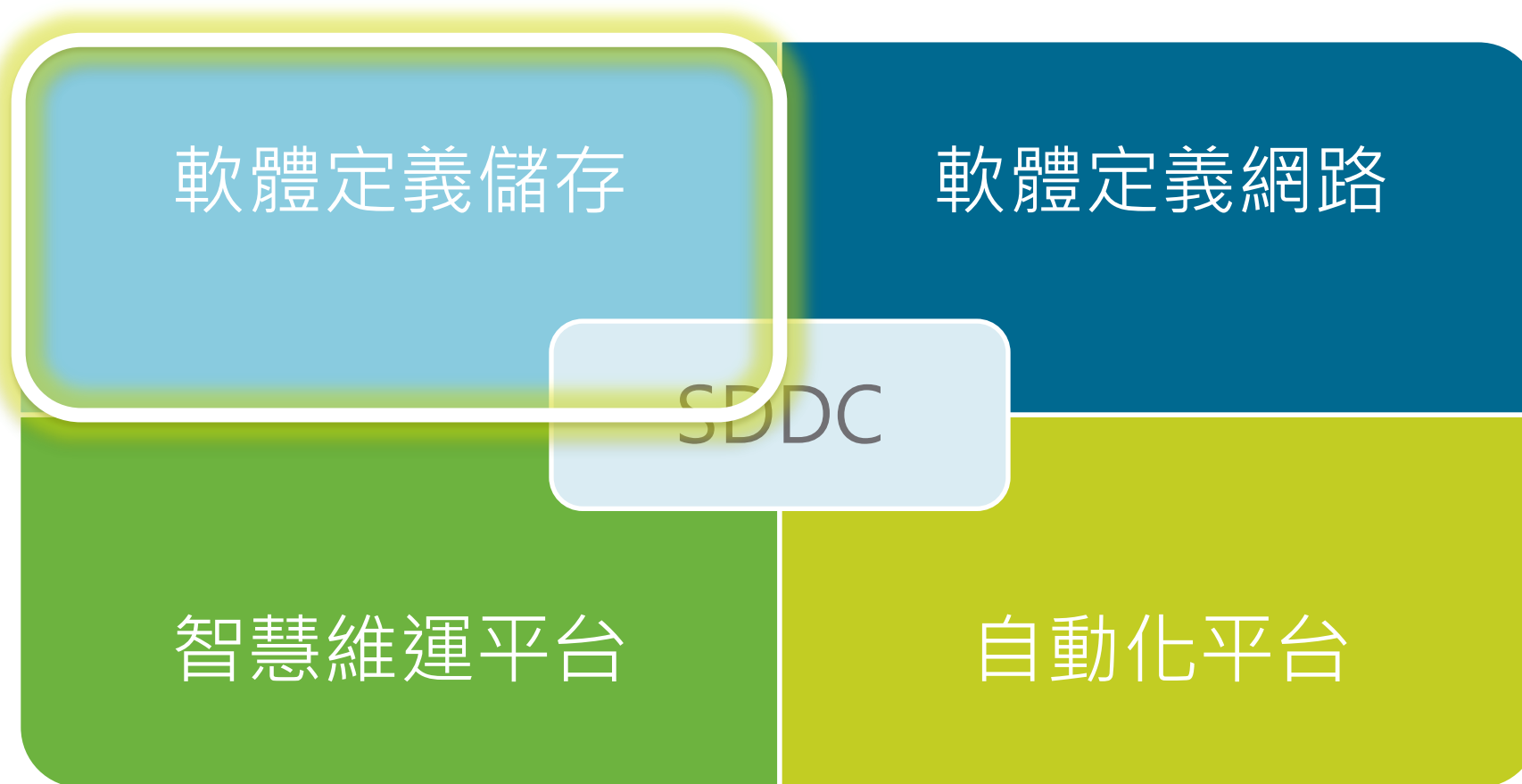
議程

-
- 1 高效能運算虛擬化
 - 2 Deep Learning 實測案例
 - 3 VMware 軟體定義資料中心
-

VMware SDDC – 整體雲端框架

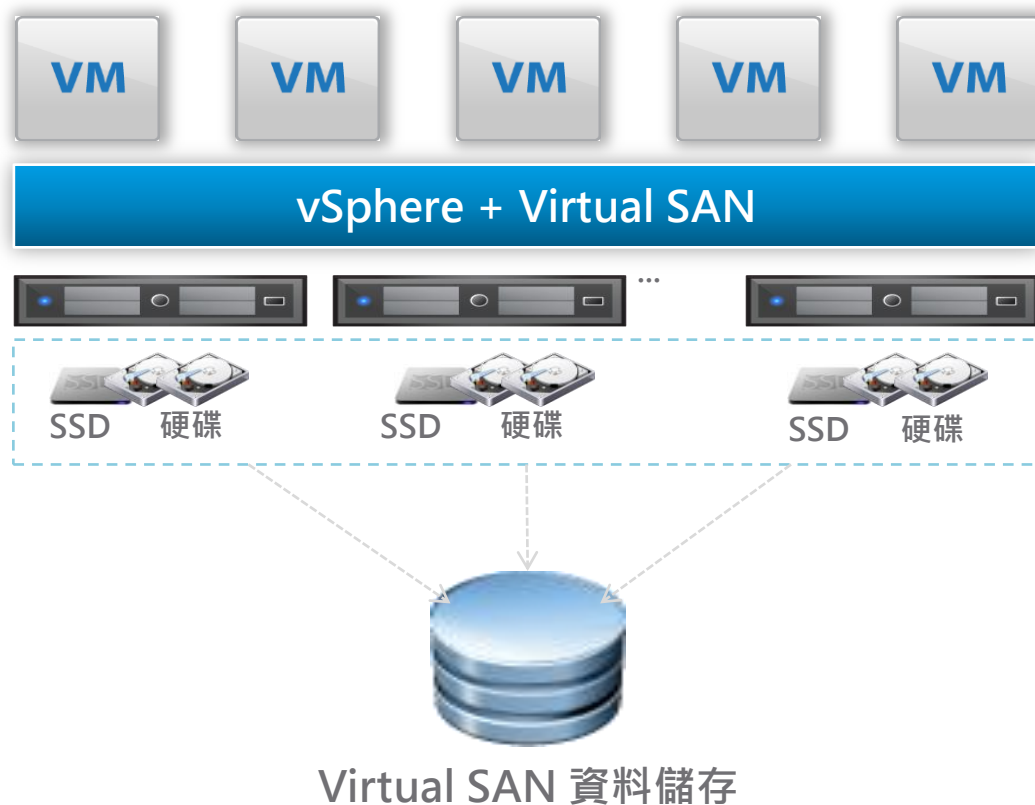


新一代資料中心四大主軸



分散式高效儲存 – VMware vSAN

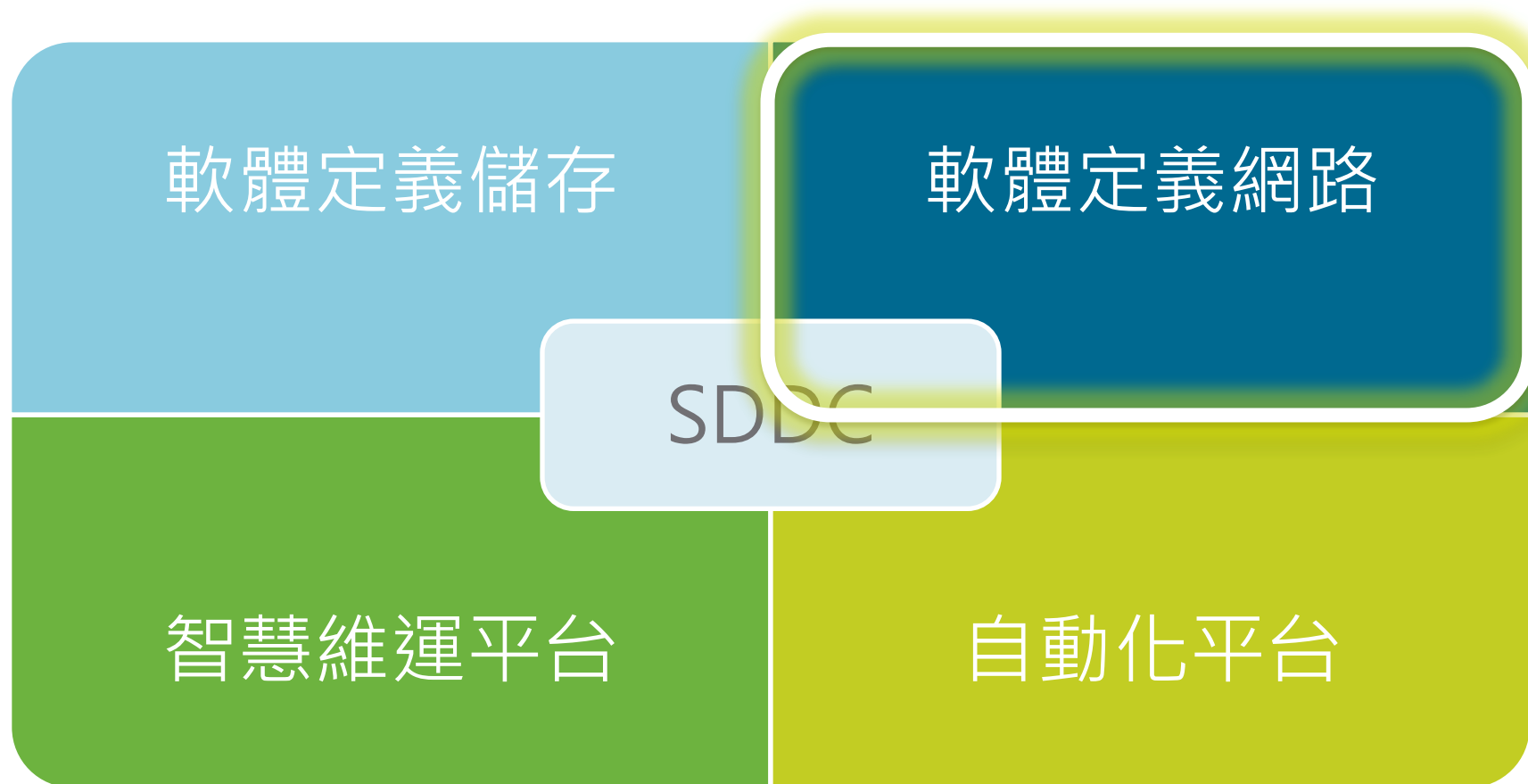
vSAN 內嵌在 vSphere Kernel，將 I/O 資料路徑縮至最短



高科技製造業導入現況

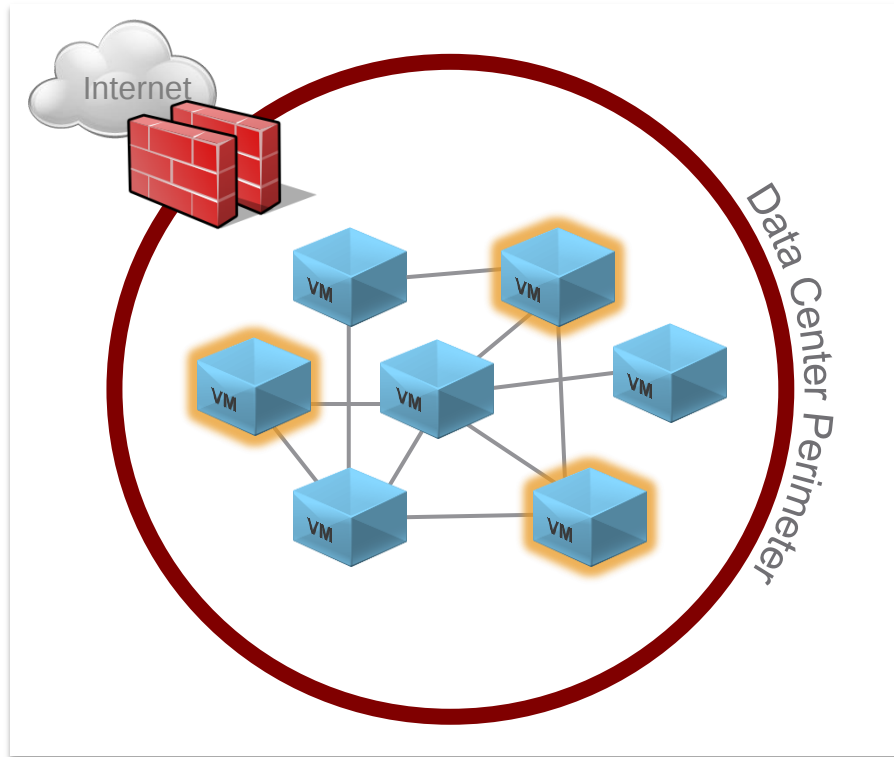
UMC(產線機台)
MTK (研發單位VDI)
群創(產線機台)
TSMC(產線機台)

新一代資料中心四大主軸

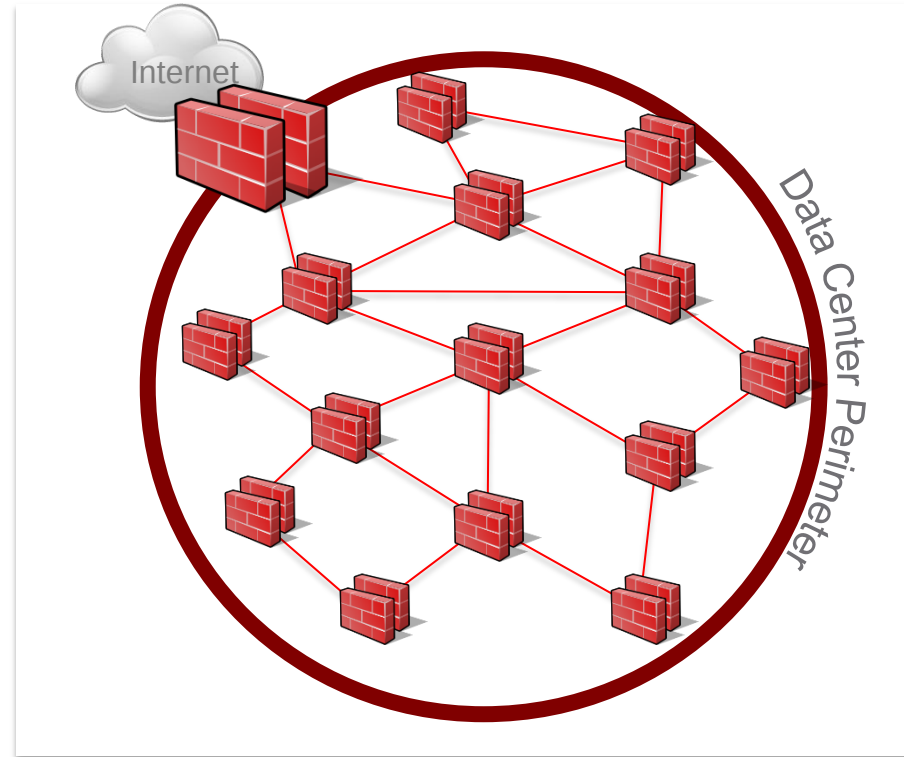


傳統架構下，安全防護僅著重邊界防護。但現今環境內動態建立的業務、各個虛擬機器都要求獨立的保護

傳統邊界防護：2~4組防火牆，管理數十至數百台實體或虛擬機器

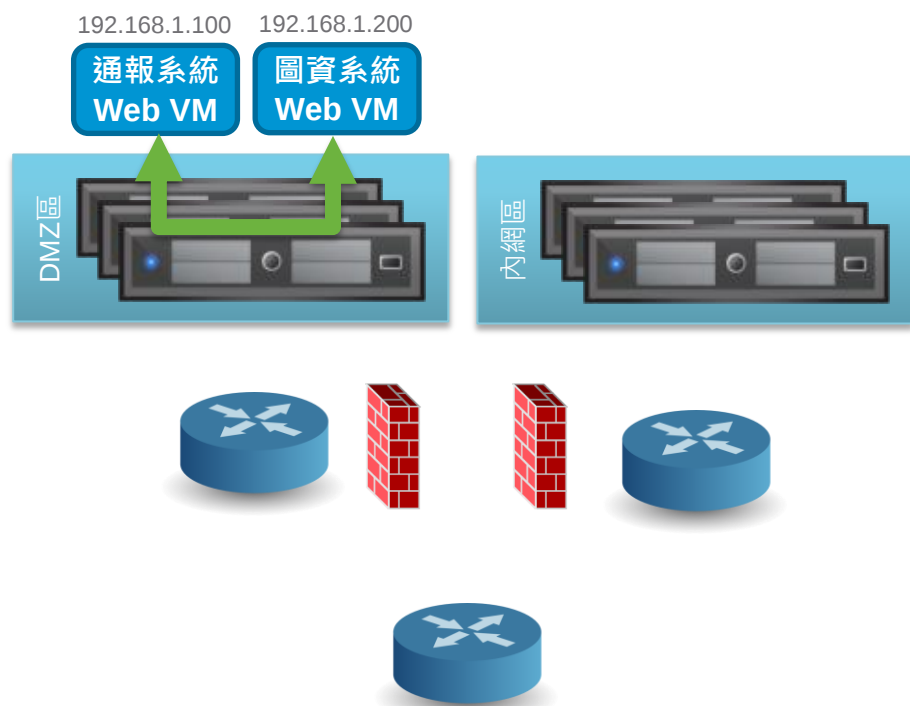


現今的需求：快速增長的業務，要求直接針對業務甚至細至VM層的防護

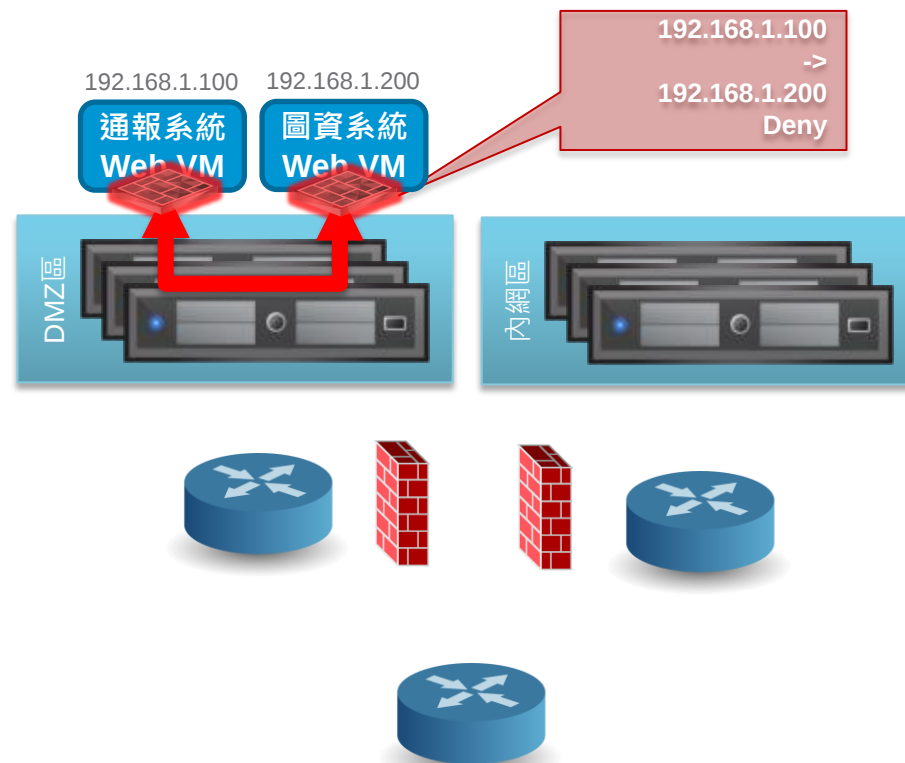


微切分防火牆功能可對同區流量進行過濾

傳統防火牆無法支援同區/同網段流量過濾

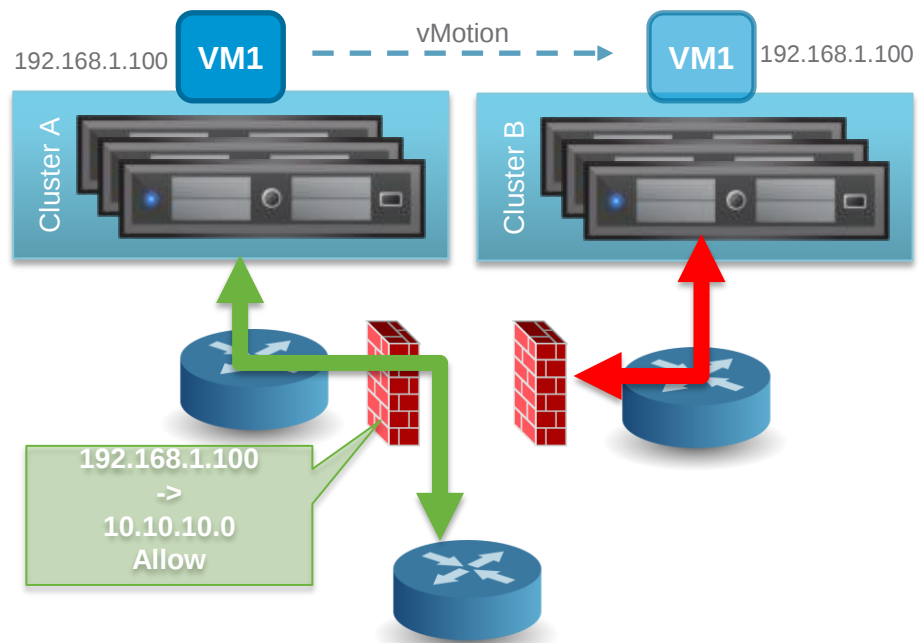


分散式防火牆可針對任意VM間過濾流量

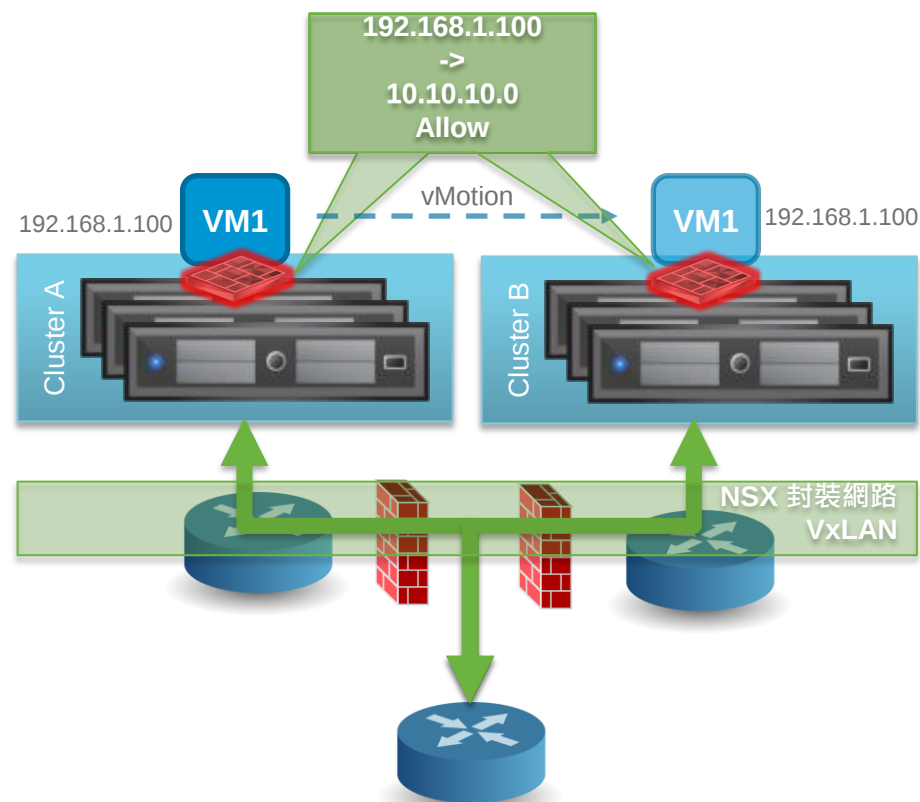


安全政策隨虛擬機移動而移動

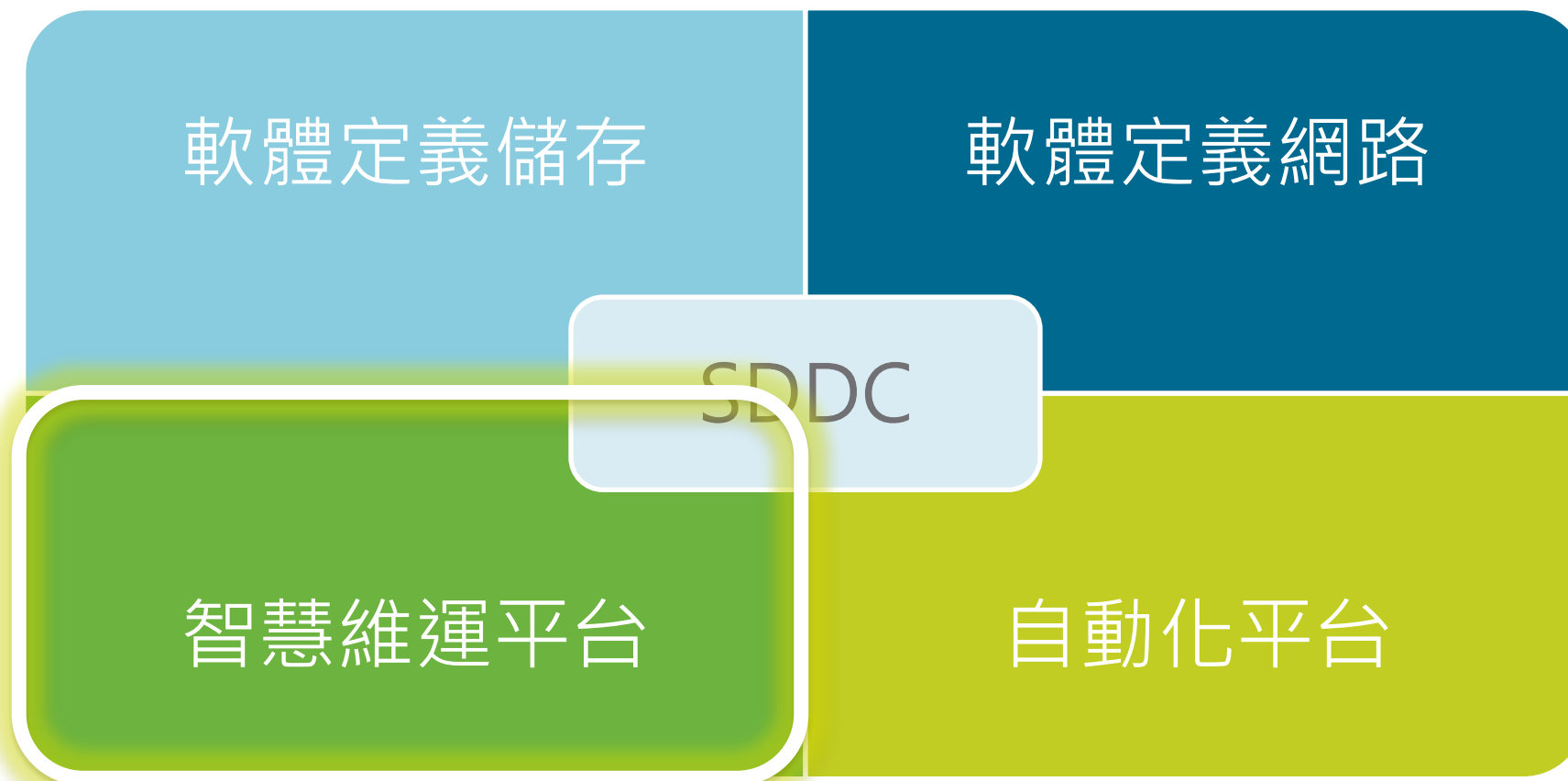
傳統防火牆無法支援虛擬機跨區域移轉



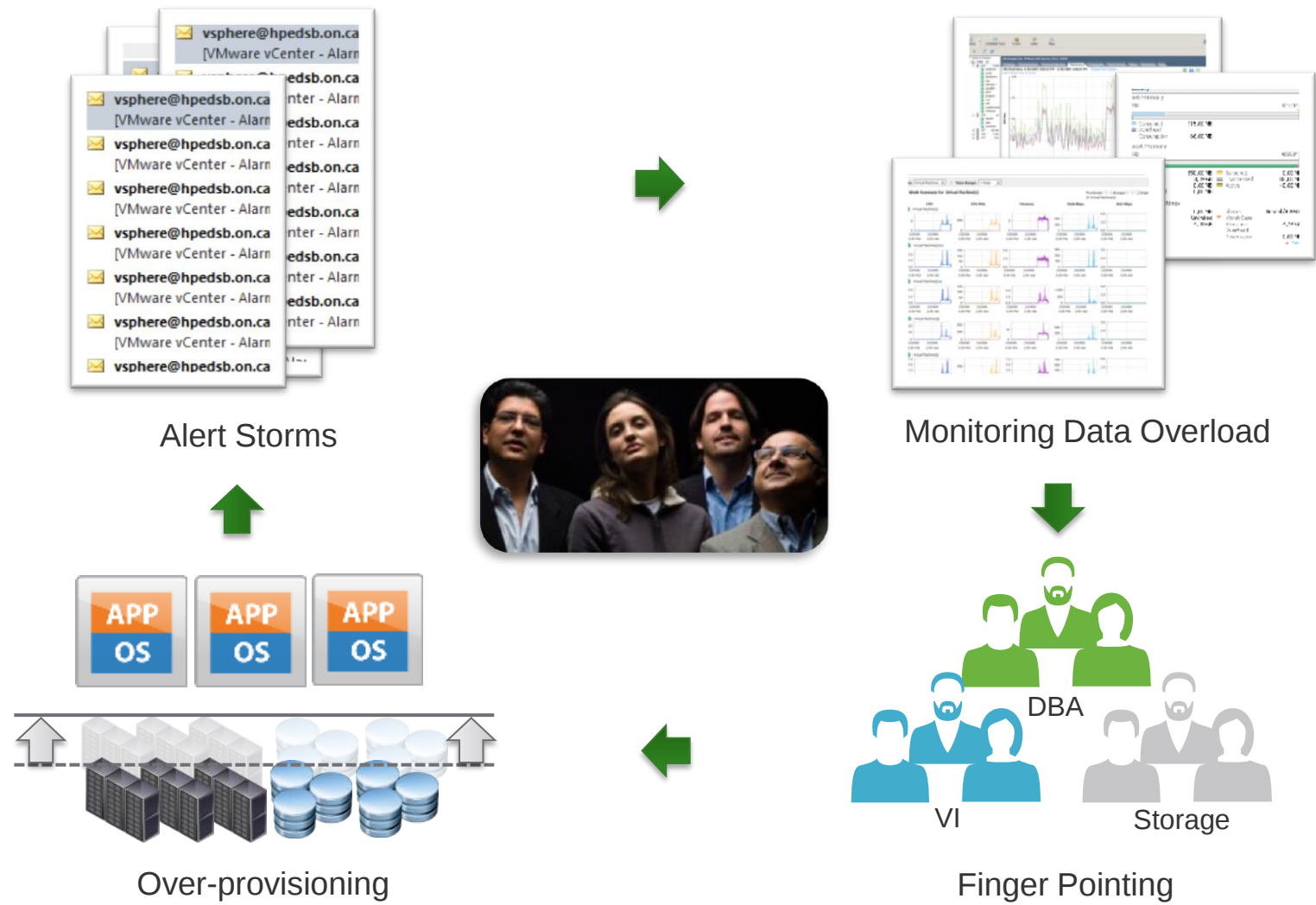
分散式防火牆綁定所有VM網路口，即便進行vMotion仍持續有效



新一代資料中心四大主軸



作業管理現況與難題

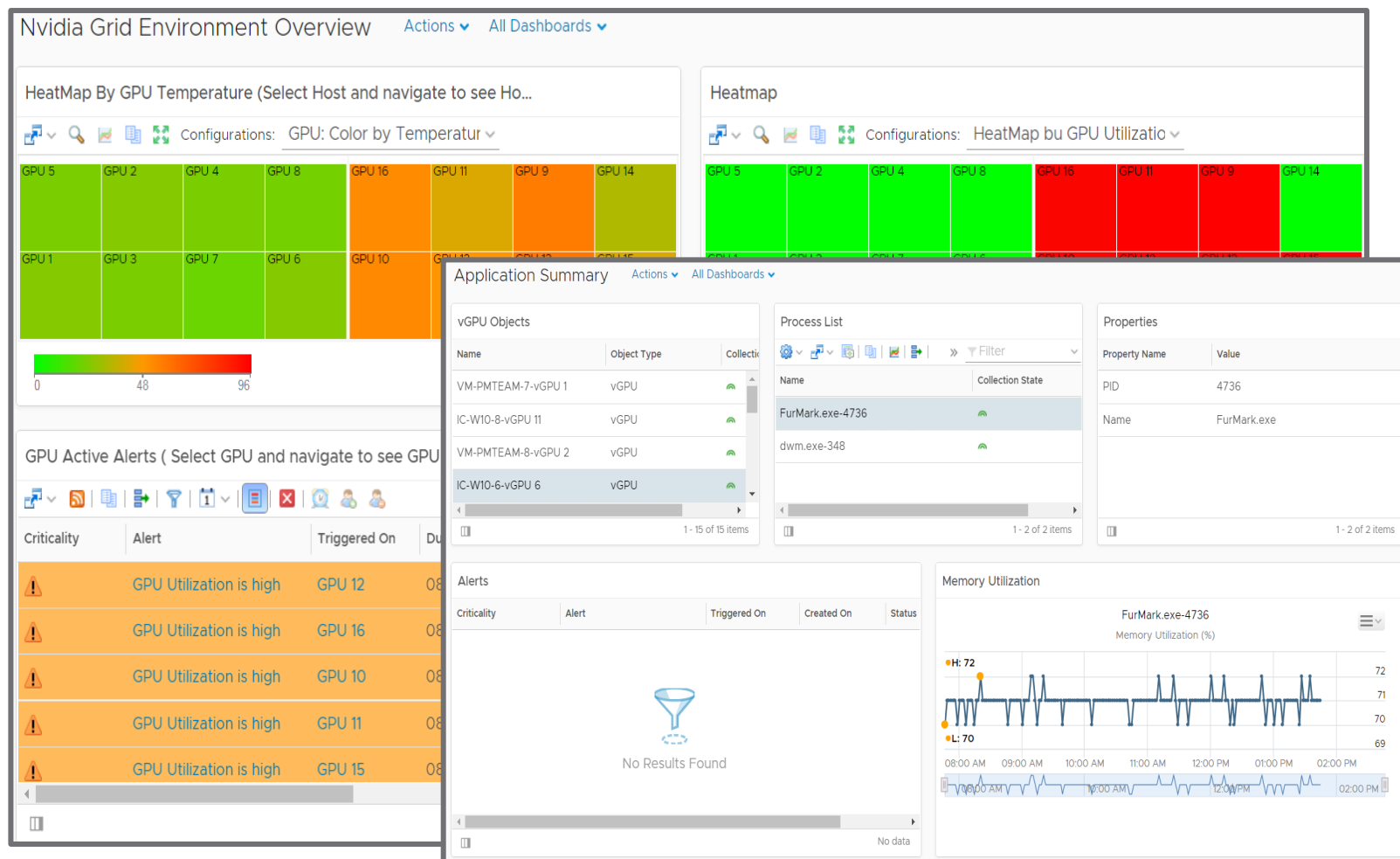


vRealize Operations 協助運營轉換 “被動” 的角色

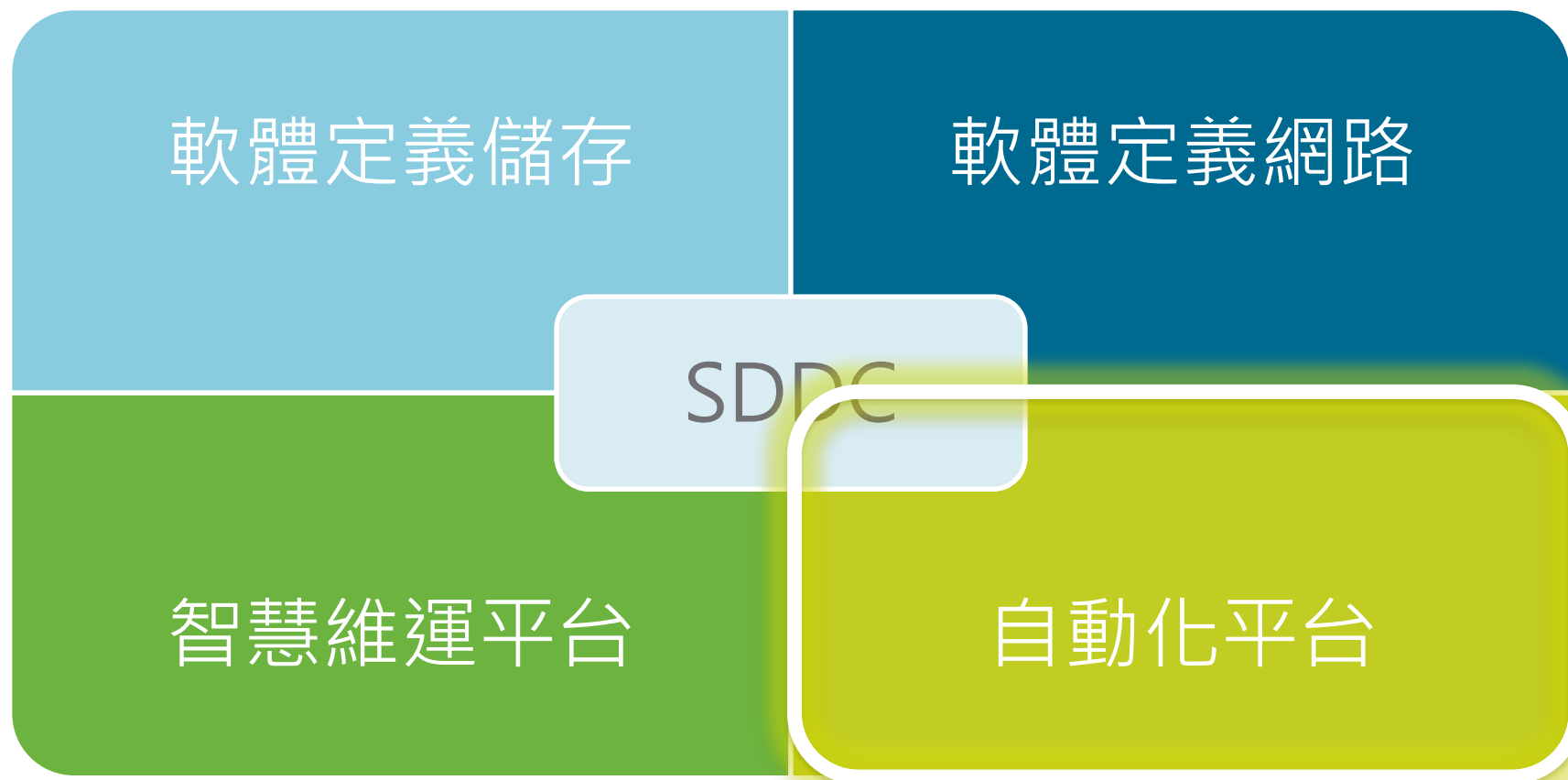


GPU虛擬化效益-雲平台整合，效能監控

- 使用虛擬化平台運行GPU運算可直接與雲平台整合
 - GRID 環境儀表板
 - vGPU 效能監控
 - 應用層次監控可下探至GPU狀態



新一代資料中心四大主軸



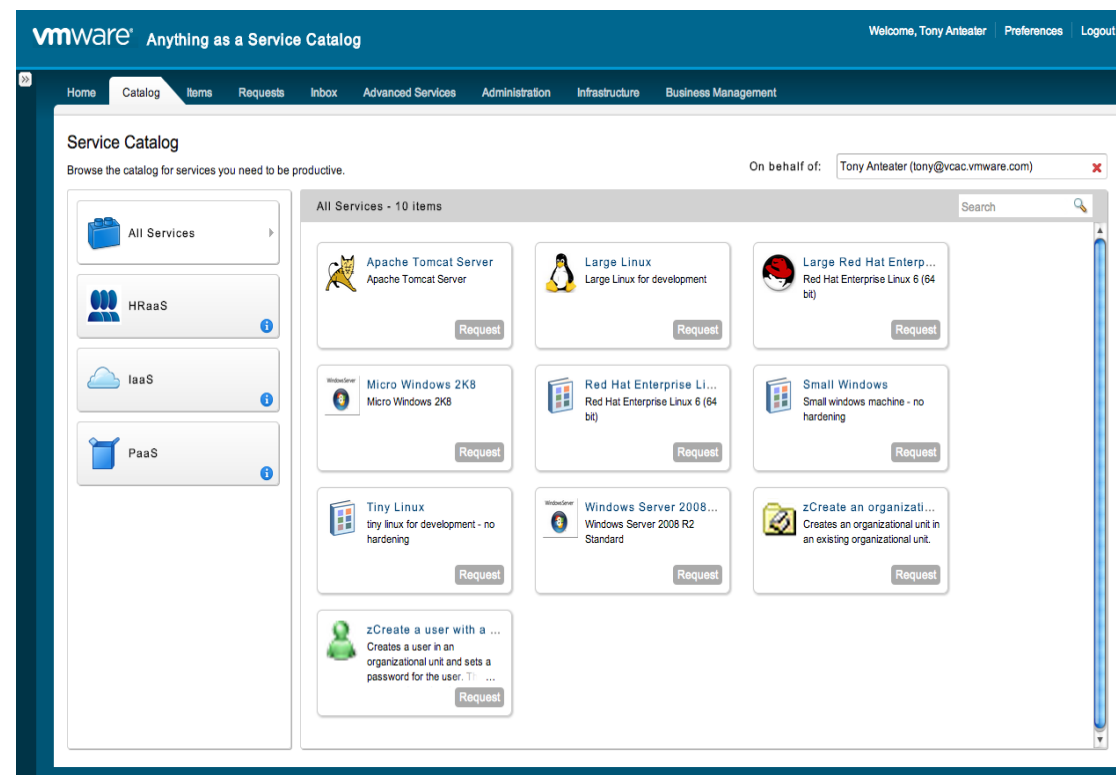
雲端自動化平台 vRealize Automation

重點功能

- 單一Web入口自助服務
- 彈性服務藍圖
- 結合核准流程
- 跨平台整合
- 整合虛擬網路自動部署功能

方案效益

- 以彈性的自動化解決方案擴充現有的基礎架構、程序及環境
- 佈建標準化與流程化
- 自助服務與快速部署提升服務滿意度
- IT 即是服務



落地自動化IT

Home

Catalog

Items

Requests

Inbox

Design

Administration

Infrastructure

Containers

Business Management

Service Catalog

Browse the catalog for services you need.

All Services

AWS 公有雲

Linux 作業系統

NSX 虛擬網路專區

Windows 作業系統

台北資料中心

On behalf of: colin@sysage.com

Welcome, colin. | Preferences | Help | Logout

vmware vRealize Automation

Edit Blueprint: 3Tier-Service

Categories

Machine Types 11

Software Components 1

Blueprints 4

Network & Security 8

XaaS 1

Containers 2

Configuration Management 1

Other Components 0

Search

Amazon EC2 Machine

Azure Machine

Generic Virtual Machine

Hyper-V (SCVMM) Machine

Hyper-V (Standalone) Machine

KVM (RHEV) Machine

OpenStack Machine

vCloud Air Machine

vCloud Director Machine

vSphere (vCenter) Machine

XenServer Machine

Design Canvas

WLB_VIP

NET24

WEB_pool

APP_pool

DBS_pool

vRASGWeb

WEBVM

vSphere (vCenter...)

vRASGAP

APPVM

vSphere (vCenter...)

vRASGDB

DBSVM

vSphere (vCenter...)

Save | Finish | Cancel

GPU虛擬化效益-雲平台整合，自動化隨選申請AI服務

vmware vRealize Automation

歡迎使用 · anson · 喜好設定 | 說明 | 登出

首頁 目錄 項目 申請 收件匣 設計 管理 基礎結構 容器 Business Management

服務目錄

瀏覽目錄以尋找所需的服務。

代表: anson@sysage.com

業務群組: 全部 搜尋

所有服務

- AI 服務
- AWS 公有雲
- Container 容器服務
- DB 資料庫服務
- Linux 作業系統
- Network Security
- Windows 作業系統
- 台北資料中心

所有服務 (25)

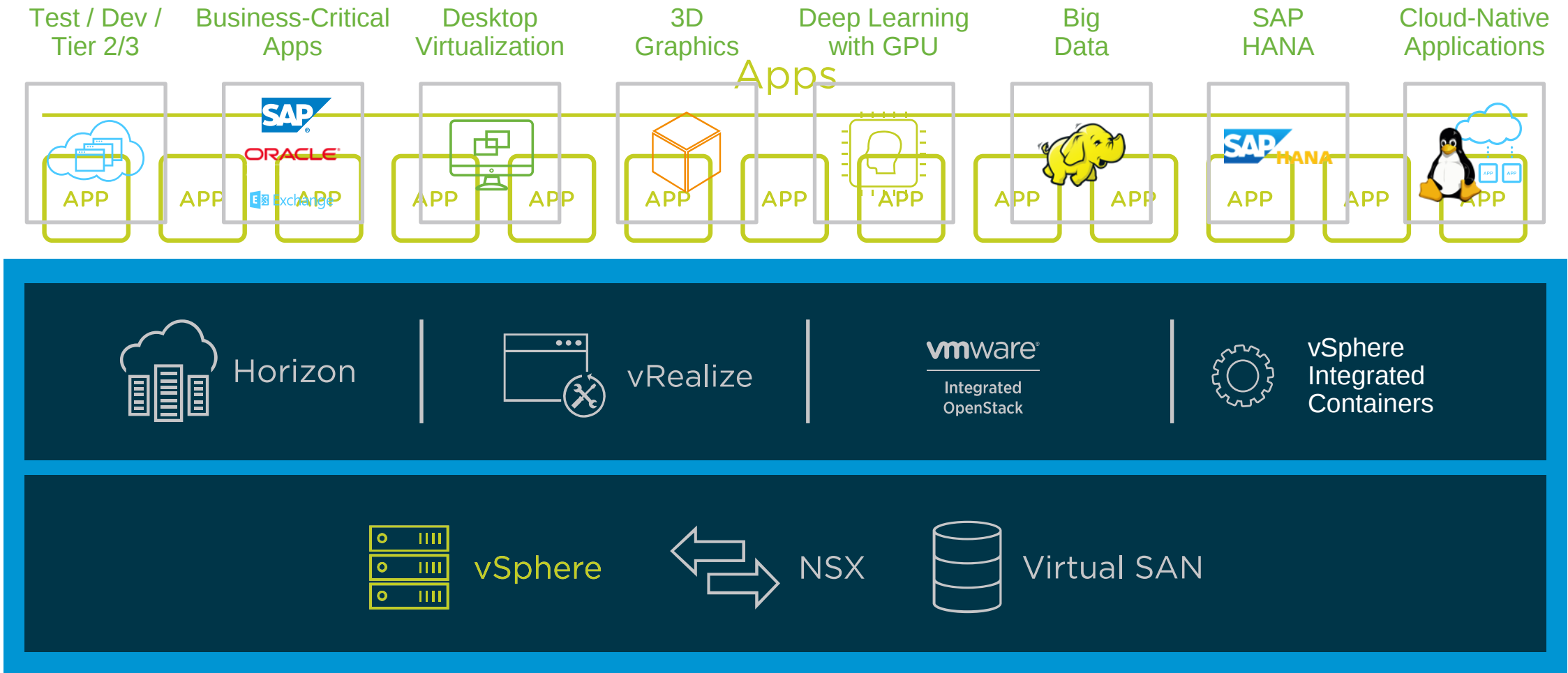
3Tier-Service BU1-vRA 展示 3Tier Service 安裝 Web+AP+DB 申請	3Tier-Service AI BU1-vRA AI 服務 申請	3Tier-Service AI-2 BU1-vRA AI 2 服務 申請	Add Security Policy (Con... BU1-vRA Create security policy between consumer and ... 申請
AWS-EC2-Ubuntu BU2-AWS 申請	AWS-EC2-Win2008 BU2-AWS 申請	Delete Security Policy (C... BU1-vRA Deleting security policy between consumer and ... 申請	Docker - CoreOS BU1-vRA Sample template for provisioning Docker ho ... 申請
Docker - PhotonOS BU1-vRA Sample template for provisioning Docker ho ... 申請	FreeBSD BU1-vRA 申請	JBoss BU1-vRA 申請	PostgreSQL BU1-vRA 申請
RHEL62 BU1-vRA 申請	SUSE BU1-vRA 展示 SUSE 12 安裝 申請	Tomcat BU1-vRA 申請	Update Access List Secu... BU1-vRA Update Access List Rules (security filters) 申請
W2012-AP1 BU1-vRA with Agent 申請	W2012-IISv8.5 BU1-vRA 展示 IIS v8.5 安裝 申請	W2012-noAdmin BU1-vRA 展示 1.Rename Administrator to os_ho ... 申請	W2012r2-input BU1-vRA 展示 自訂 Hostname & IP- 170 申請

Copyright © 2012-2017 VMware, Inc. 版權所有。

版本 7.3.0 (組建編號 5493918) | 隱私權政策 | 連絡我們

AI服務目錄

VMware SDDC 提供符合各種企業應用需求平台



Thank You

llo@vmware.com