

Better HPC Better Living
為更好的生活努力

運用自由軟體快速打造私有雲端

Building Your Private Cloud Rapidly using Open Source

王耀聰 副研究員

jazz@nchc.org.tw

財團法人國家實驗研究院
國家高速網路與計算中心



國家實驗研究院
National Applied Research Laboratories

國家高速網路與計算中心
National Center for High-Performance Computing

WHO AM I? 這傢伙是誰啊? JAZZ?

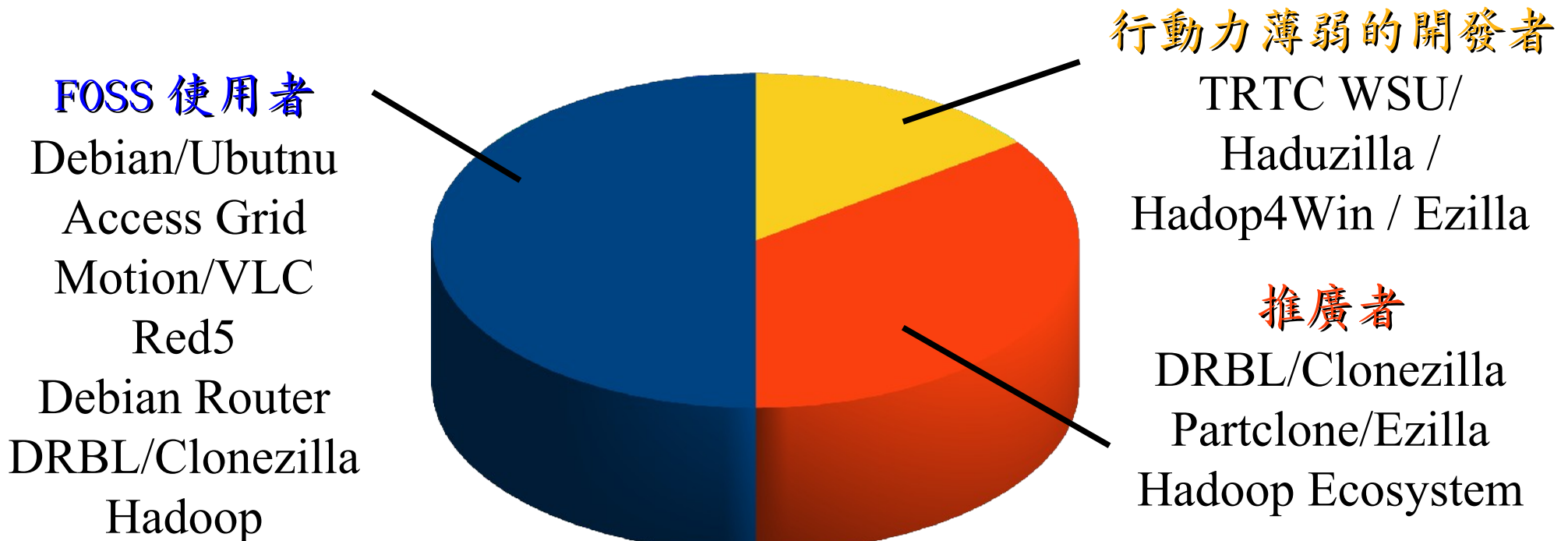
- 講者介紹：

- 國網中心 王耀聰 副研究員 / 交大電控八九級碩士
- jazz@nchc.org.tw



- 所有投影片、參考資料與操作步驟均在網路上

- <http://trac.nchc.org.tw/cloud>
- 由於雲端資訊變動太快，愛護地球，請減少不必要之列印。





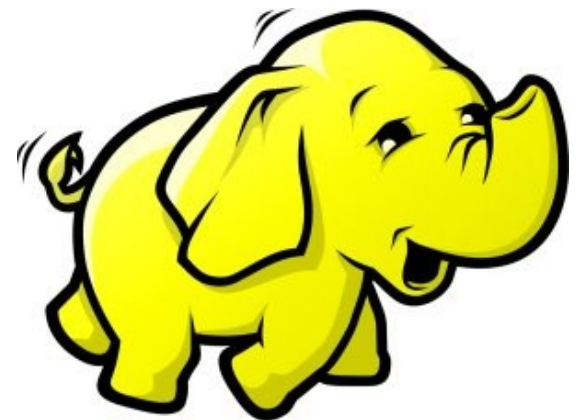
雲端趨勢、商業模式及三大關鍵技術

Three Key Technologies of Cloud Computing

Jazz Wang

Yao-Tsung Wang

jazz@nchc.org.tw



TRENDS

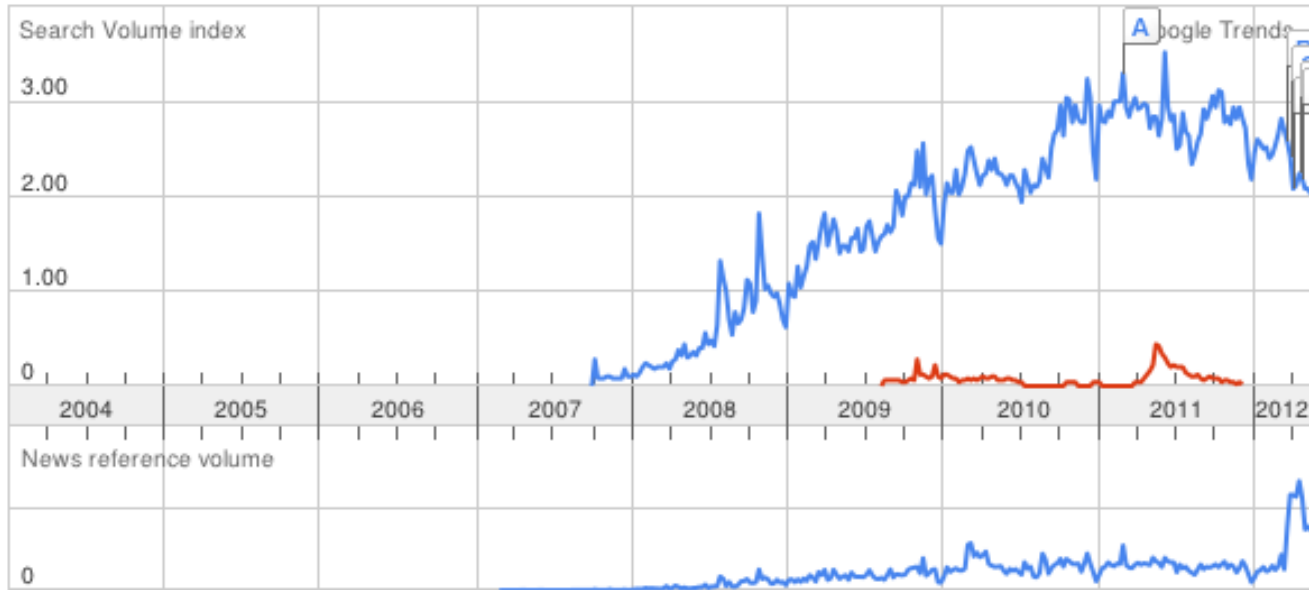
趨勢

The Trends of Cloud Computing

雲端運算的趨勢觀察：搜尋關鍵字熱門度

cloud computing

1.00 雲端運算 0.02



Rank by cloud computing

Regions

1. [India](#)
2. [Singapore](#)
3. [Hong Kong](#)
4. [Ireland](#)
5. [South Africa](#)
6. [South Korea](#)
7. [Philippines](#)
8. [Malaysia](#)
9. [Taiwan](#)
10. [United States](#)

Cities

1. Hyderabad, India
2. Bangalore, India
3. Pune, India
4. Chennai, India
5. Mumbai, India
6. New Delhi, India
7. Singapore, Singapore
8. Hong Kong, Hong Kong
9. Dublin, Ireland
10. San Francisco, CA, USA

Google Trends

觀察全球對雲端運算
這個字眼的熱門搜尋
趨勢～

約莫 2007 年開始炒
2011 年到達頂峰
2012 年開始退燒

<http://www.google.com/trends>

The Trends of Cloud Computing

雲端運算的趨勢觀察：企業重視度

CIO technologies	Ranking of technologies CIOs selected as one of their top 3 priorities in 2012			
	2012	2011	2010	2009
Analytics and business intelligence	1	5	5	1
Mobile technologies	2	3	6	12
Cloud computing (SaaS, IaaS, PaaS)	3	1	2	16
Collaboration technologies (workflow)	4	8	11	5
Virtualization	5	2	1	3
Legacy modernization	6	7	15	4
IT management	7	4	10	*
Customer relationship management	8	18	*	*
ERP applications	9	13	14	2
Security	10	12	9	8
Social media/Web 2.0	11	10	3	15

*Not an option that year

參考來源：<http://www.gartner.com/DisplayDocument?id=1524714>

Gartner CIO Agenda 2012

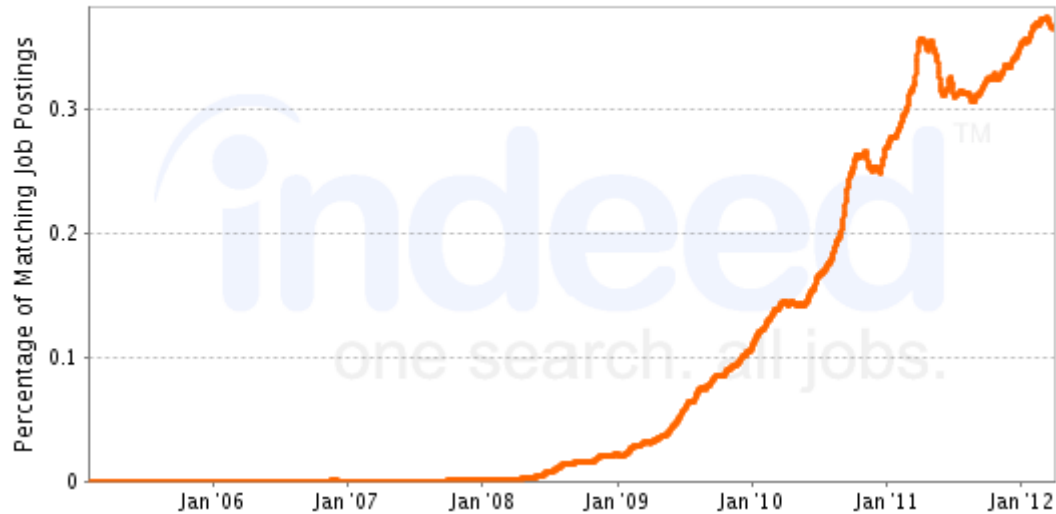
觀察全球資訊長對雲端運算的重視程度

The Trends of Cloud Computing

雲端運算的趨勢觀察：就業市場需求

Job Trends from Indeed.com

— cloud computing



indeed 與 simplyhired

觀察全美對雲端運算一詞的
職缺需求趨勢～

橫軸 = 時間

縱軸 = 佔全部職缺的百分比



約莫 2008 年開始變多

2011 年到達頂峰

2012 年開始退燒

<http://www.indeed.com/jobtrends>

<http://www.simplyhired.com/a/jobtrends>

WHAT

定義



什麼是雲端運算啊？

What is Cloud Computing ?

雲端概念

雲端運算不是一項全新技術，
而是一項概念。
雲端的意義不在技術，
而在商業模式的改變。

<http://www.youtube.com/watch?v=bJLSAcU6O3U>

<http://www.youtube.com/watch?v=VIMtd3nfPqc>

當紅「雲端運算」你瞭解了嗎？
雲端產業 8分鐘就上手

Paradigm Shift of Cloud Business Model !!

雲端商業模式的典範轉移

軟體從買斷變成租賃

Office 2007 →
Google Docs / Office 365

硬體從買斷變成租賃

PC / Server →
Hosting / Colocation →
Amazon EC2 / S3

軟體從單機變成行動

Outlook → Webmail →
Mail Web Apps →
Mail Mobile Apps

硬體從固定變成行動

PC / Server →
NB / Tablet →
Pad / Mobile

National Definition of Cloud Computing

美國國家標準局 NIST 給雲端運算所下的定義

5 Characteristics

五大基礎特徵

4 Deployment Models

四個佈署模型

3 Service Models

三個服務模式

1. On-demand self-service

隨需自助服務

2. Broad network access

隨時隨地用任何網路裝置存取

3. Resource pooling

多人共享資源池

4. Rapid elasticity

快速重新佈署靈活度

5. Measured Service

可被監控與量測的服務

4 Deployment Models of Cloud Computing

雲端運算的四種佈署模型

Public Cloud
公用雲端

Target Market
is **S.M.B.**
主要客戶為
中小企業

Community Cloud
社群雲端

Academia **學術**為主



Dynamic Resource Provisioning
between public and private cloud
私有雲端動態根據計算需求
調用公用雲端的資源

Hybrid
Cloud

以**大型企業**
為主要客戶
Enterprise is
key market



私有雲端
Private Cloud

3 Service Models of Cloud Computing

雲端運算的三種服務模式 (市場區隔)

IaaS

Infrastructure as a Service

架構即服務

PaaS

Platform as a Service

平台即服務

SaaS

Software as a Service

軟體即服務



2 perspectives : Services vs Technologies

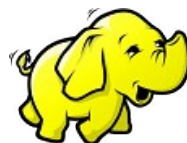
您想聽的是「雲端服務」還是「雲端技術」？

Google YouTube e W



雲端服務

Microsoft



雲端技術



Cloud computing hype spurs confusion, Gartner says
<http://www.computerworld.com/s/article/print/9115904>

淺談雲端運算 (Cloud Computing)

http://www.cc.ntu.edu.tw/chinese/epaper/0008/20090320_8008.htm

One key spirit of Cloud Computing

用一句話說明雲端運算！服務才是王道！

Anytime 隨時

Anywhere 隨地

With Any Devices 使用任何裝置

Accessing Services 存取各種服務

Key spirit of Cloud ~

形成服務才是重點！！

Everything as a Service !!

Cloud Computing =~ Network Computing
雲端運算 =~ 網路運算

WHEN

何時

The Wisdom of Clouds (Crowds)

雲端序曲：雲端的智慧始終來自於群眾的智慧

2006年8月9日

Google 執行長施密特 (Eric Schmidt) 於 SES'06 會議中首次使用「雲端運算 (Cloud Computing) 」來形容無所不在的網路服務

2006年8月24日

Amazon 以 Elastic Compute Cloud 命名其虛擬運算資源服務

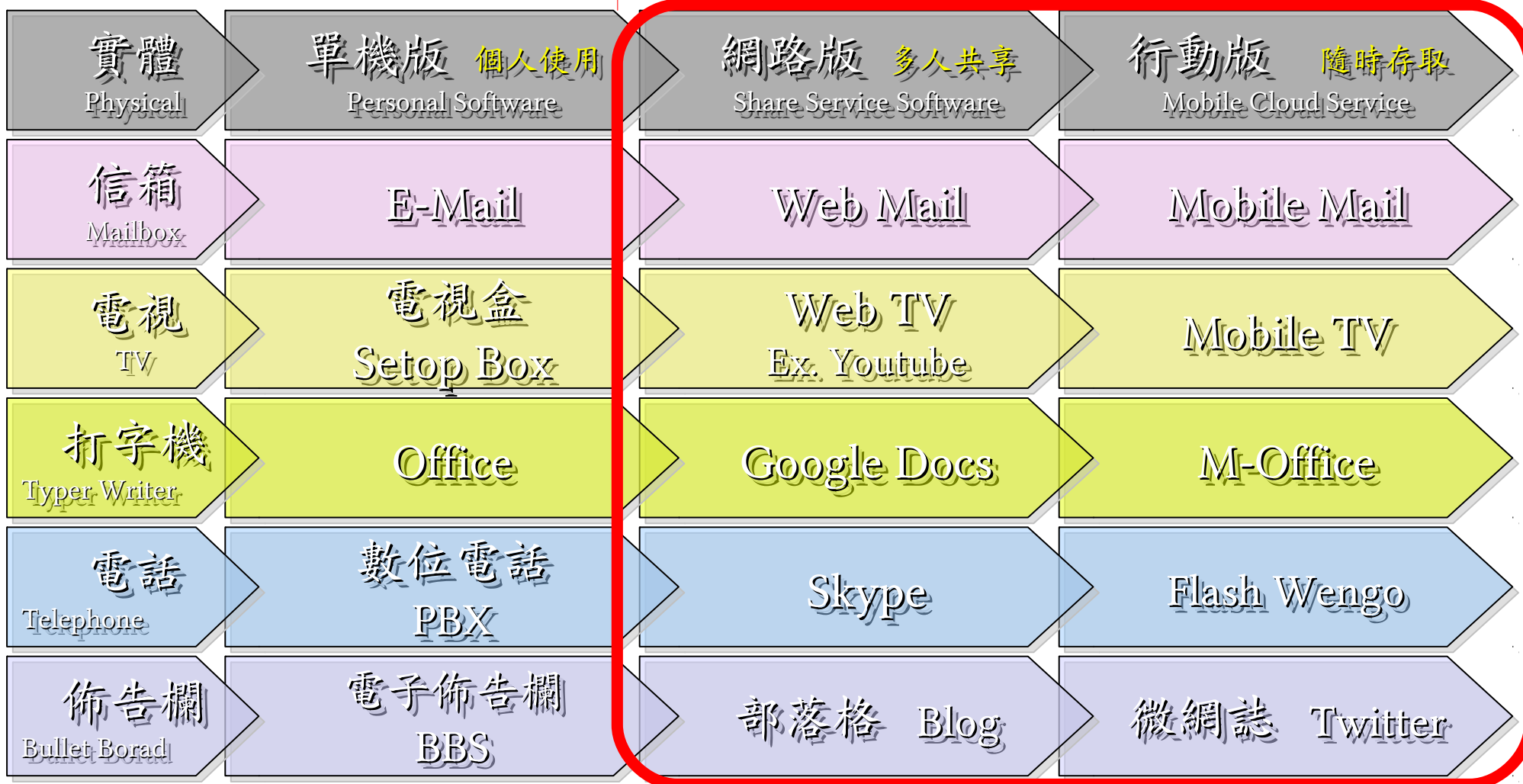
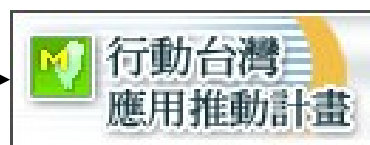
THE WISDOM OF CLOUDS

What you need to know
about cloud computing



Evolution of Cloud Services

雲端服務只是軟體演化史的必然趨勢



Why are they named by “SMART” ?!
智慧打哪兒來？！

Smart Phone
智慧手機

Smart Car
智慧車輛

Smart Grid
智慧電網

SMART ?
哪裡長
智慧了？

Smart City
智慧城市

Smart Home
智慧家庭

Smart Meter
智慧電錶

資料 Data

知識 Knowledge

智慧 Wisdom

Can Machine understand You? 讓機器更懂你?

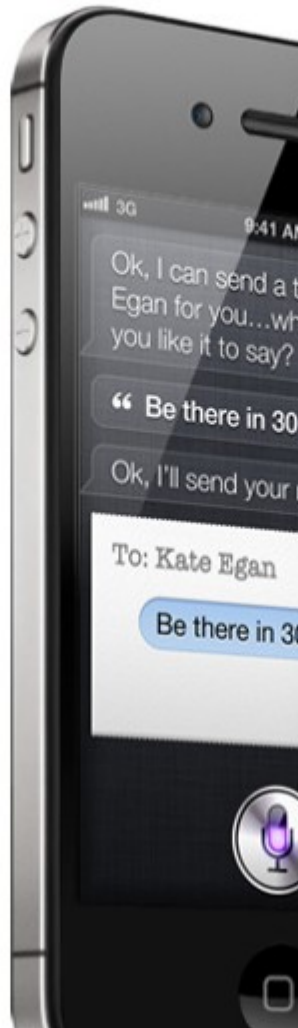
iPhone

Features Built-in Apps



Siri. Beta
Your wish is
its command.

Siri on iPhone 4S lets you use your voice to send messages, schedule meetings, place phone calls, and more. Ask Siri to do things just by talking the way you talk. Siri understands what you say, knows what you mean, and even talks back. Siri is so easy to use and does so much, you'll keep finding more and more ways to use it.



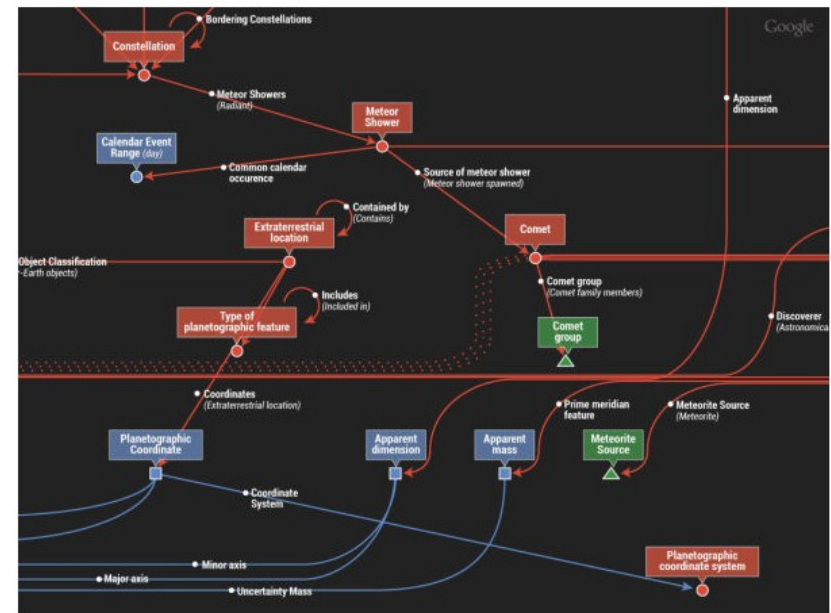
Google將發展「人工智慧」 永久改變搜尋引擎

2012年02月15日 00:11

點評：超級阿斯拉，衝啊！（阿斯拉：好的，準人！）

記者黃郁棋 / 綜合報導

每個人都再猜，下一波網路革命是什麼？每個人都再猜，未來的世界會如何運作？Google的資深副總Amit Singhai透露了一點訊息。「Google正努力從『單字』層面進展到『意義』層面，未來搜尋引擎提供的不只是關鍵字搜尋，搜尋引擎甚至會『明白』你到底要什麼。」



▲ Google未來將會朝「人工智慧」前進。(圖 / 取自mashable.com)

WHY

爲何

Key Driving Forces of Cloud Computing

雲端運算的關鍵驅動力

端

隨需
行動服務

Mobile Service

個人

用任何連網裝置
都可以存取資料

企業

租賃取代買斷
動態隨需付費

降低
經營成本

Cost Down

全球

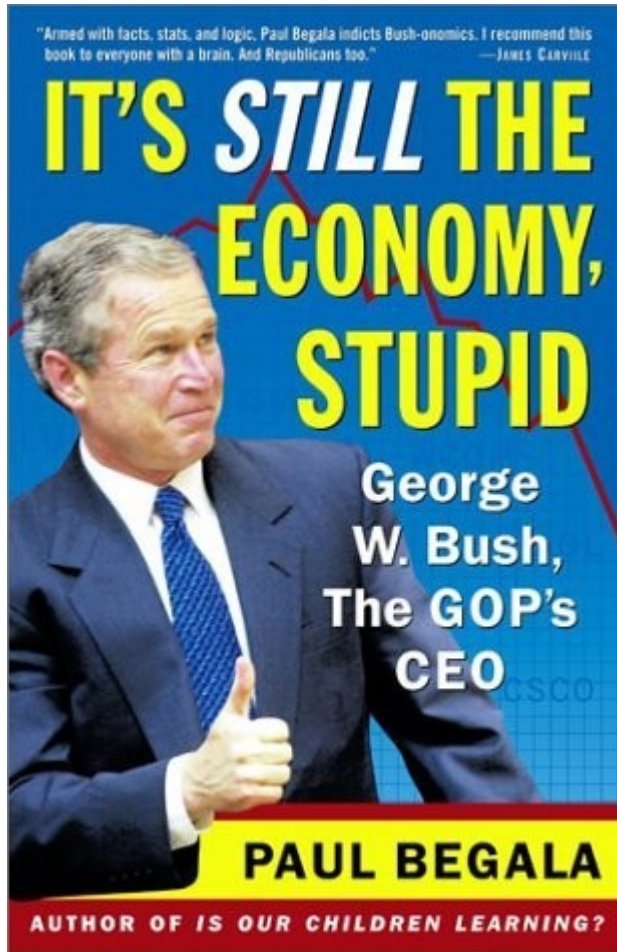
資料往雲擺
減少資料傳輸

因應
資料爆炸

Data Explore

雲

IT'S THE DATA, STUPID!



「笨蛋！重點在經濟」

("It's the economy, stupid")

卡維爾 (James Carville) 自創這句標語，
促使柯林頓當上美國第 42 屆總統。

- 1992 年

「笨蛋！重點還是在經濟」

("It's STILL the economy, stupid")

卻讓小布希嘲笑是幼稚的總統。

- 2002 年

雲端時代，谷歌會說：「笨蛋！重點在資料」

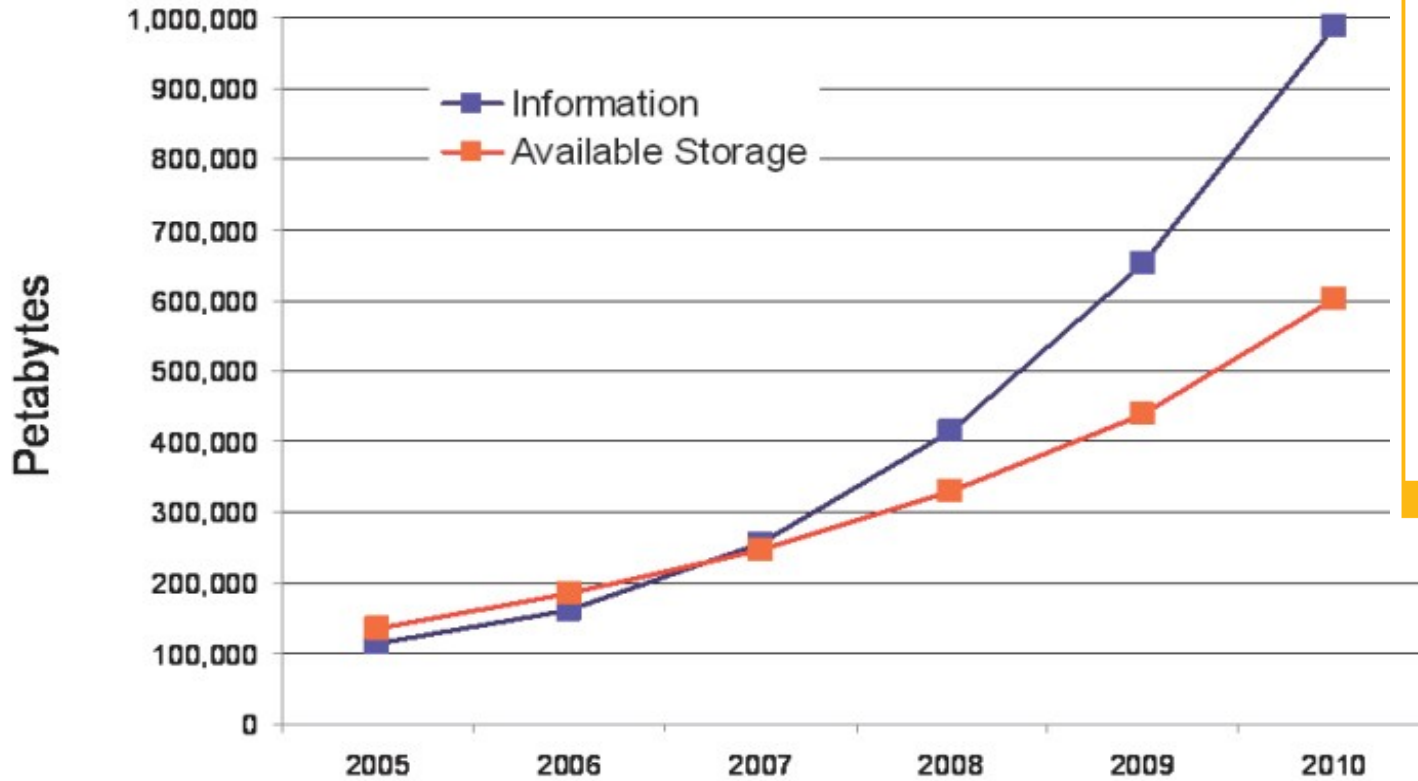
("It's the data, stupid")

**誰掌握了你的資料，就有機會掌握你的荷包
想想看，電腦、手機掉了，您心疼的是甚麼呢？**

- 2007 年

Data Explosion!! 始於 2007 的「資料大爆炸」時代

Information Versus Available Storage

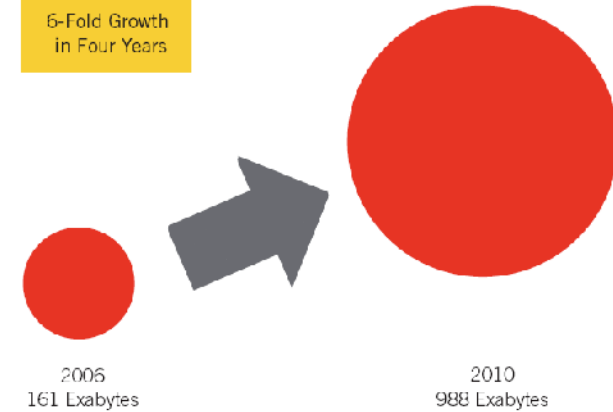


Source: IDC, 2007

Figure 1

Information Created, Captured and Replicated

6-Fold Growth
in Four Years



Source: IDC, 2007

2007 年，IDC 預估
2010 年會成長**六倍**！
(相較 2006 年)

2006 161 EB
2010 988 EB (預測)

出處：The Expanding Digital Universe,
A Forecast of Worldwide Information Growth Through 2010,
March 2007, An IDC White Paper - sponsored by EMC
<http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>

Data expanded 1.6x each year !! 每年約略 1.6 倍

THE
2011
IDC

DIGITAL UNIVERSE

STUDY

sponsored by EMC

1.8

ZETTABYTES

(1.8 trillion gigabytes) of information
will be created and replicated in 2011-
and growing fast (it has grown by a factor of 9 in just five years)

追蹤歷年的 IDC 數據：

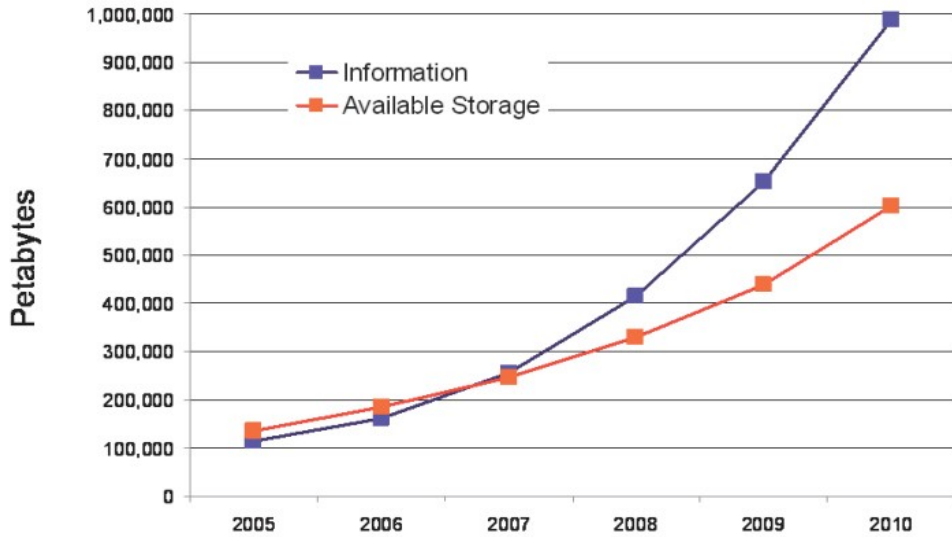
2006	161	EB	
2007	281	EB	
2008	487	EB	
2009	800	EB	(0.8 ZB)
2010	988	EB	(預測)
2010	1200	EB	(1.2 ZB)
2011	1773	EB	(預測)
2011	1800	EB	(1.8 ZB)

景氣差而成長趨緩？
或受新技術抑制？

出處：[Extracting Value from Chaos](#),
June 2011, An IDC White Paper - sponsored by EMC

<http://www.emc.com/collateral/about/news/idc-emc-digital-universe-2011-infographic.pdf>

Information Versus Available Storage



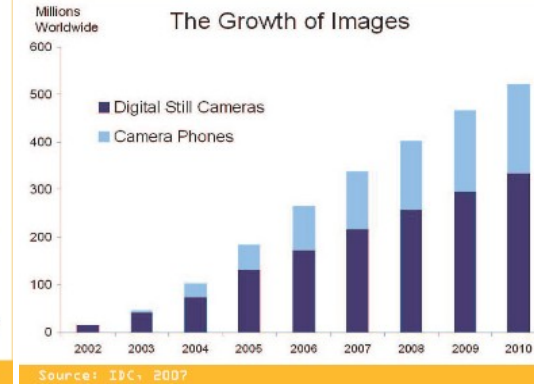
Source: IDC, 2007

2007 Data Explore

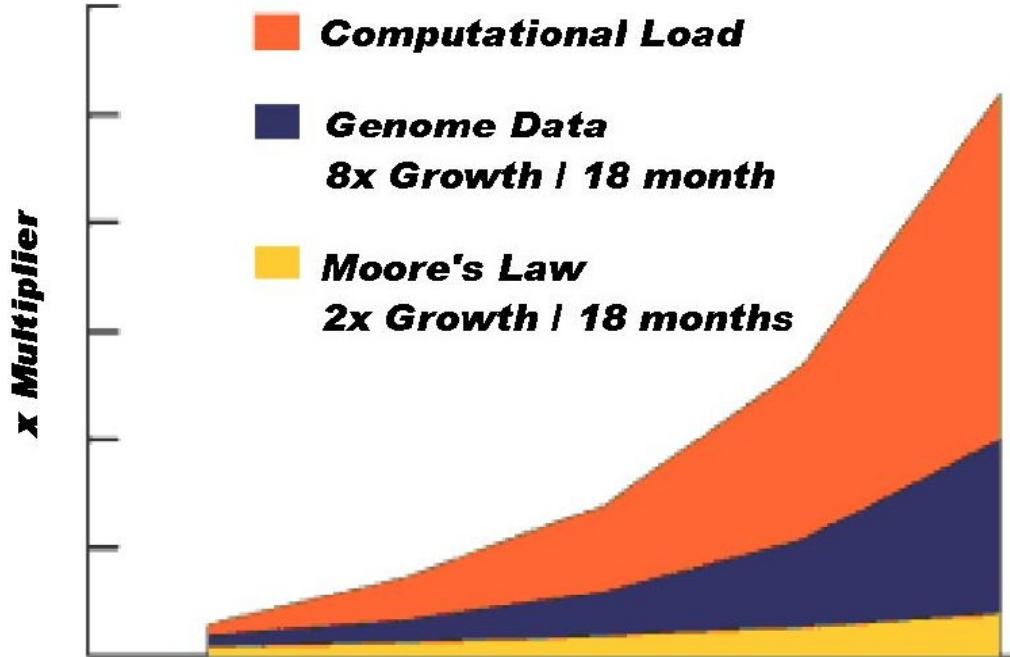
- Top 1 : Human Genomics – 7000 PB / Year
- Top 2 : Digital Photos – 1000 PB+ / Year
- Top 3 : E-mail (no Spam) – 300 PB+ / Year



Source: IDC, 2007



Source: IDC, 2007



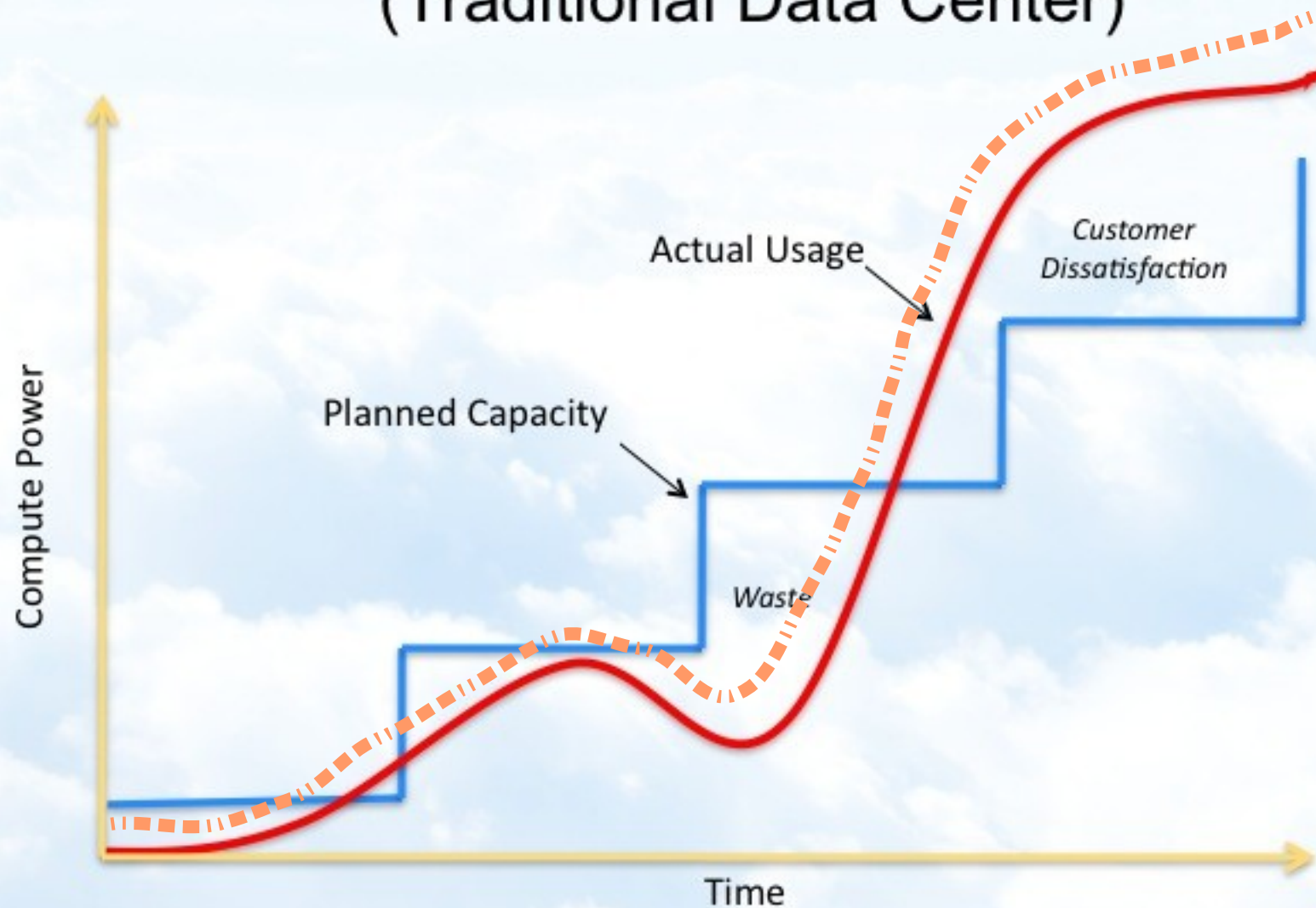
Particle Physics Large Hadron Collider (15PB)	Human Genomics (7000PB) 1GB / person 200PB+ captured 200% CAGR	World Wide Web (~1PB)	Wikipedia (10GB) 100% CAGR
Annual Email Traffic, no spam (300PB+)	Internet Archive (1PB+)	Estimated On-line RAM in Google (8PB)	Personal Digital Photos (1000PB+) 100% CAGR
200 of London's Traffic Cams (8TB/day)	2004 Walmart Transaction DB (500TB)	Typical Oil Company (350TB+)	Merck Bio Research DB (1.5TB/qtr)
UPMC Hospitals Imaging Data (500TB/yr)	MIT Babytalk Speech Experiment (1.4PB)	Terashake Earthquake Model of LA Basin (1PB)	One Day of Instant Messaging in 2002 (750GB)
Total digital data to be created this year 270,000PB (IDC)			

Phillip B. Gibbons, Data-Intensive Computing Symposium

Source: <http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>

Source: http://lib.stanford.edu/files/see_pasig_dic.pdf

Capacity vs. Usage (Traditional Data Center)



Source : http://awsmedia.s3.amazonaws.com/chart01_traditional_720x540.jpg

WHO

誰在賣

Who are the Public Cloud Service Providers ?

公用雲端關鍵提供者

Public Cloud
公用雲端



Target Market
is **S.M.B.**
主要客戶為
中小企業

Hybrid
Cloud

以大型企業
為主要客戶
Enterprise is
key market

Community Cloud
社群雲端

Academia **學術**為主



私有雲端
Private Cloud

IaaS Key Player:

Amazon 亞馬遜網路書店



- Amazon Web Service (AWS)
- 虛擬伺服器：**Amazon EC2**
 - Small (Default) \$0.085 per hour(L) - \$0.12 per hour(W)
 - All Data Transfer \$0.15 per GB
- 儲存服務：**Amazon S3**
 - \$0.15 per GB – first 50 TB / month of storage used
 - \$0.15 per GB – all data transfer in
 - \$0.01 per 1,000 PUT, COPY, POST, or LIST requests
- 觀念：**Paying for What You Use**

參考來源：
<http://eblog.cisnet.org.tw/post/Cloud-Computing.aspx>
<http://aws.amazon.com/ec2/pricing/>
<http://aws.typepad.com/aws/2010/02/aws-data-transfer-prices-reduced.html>
<http://aws.amazon.com/s3/#pricing>

Amazon Web Services 亞馬遜的網頁服務

E-Commerce

Fulfillment Web Service (FWS)

Database

SimpleDB, Rational DB Service

Compute

EC2, Elastic MapReduce (EMR)

Network

VPC, Elastic Load Balance

Storage

S3, Elastic Block Storage (EBS)

較外顯的基本服務

Support

AWS Premium Support

Payment / Billing

Amazon DevPay

Messaging

Simple Queue / Notification Service

Web Traffic

Alexa Web Info / Top Sites

Content Delivery

Amazon CloudFront

較內隱的流程細節

- Google App Engine (GAE)
- 讓開發者可自行建立網路應用程式於 Google 平台之上。
- 提供：
 - 500MB of storage
 - up to 5 million page views a month
 - 10 applications per developer account
- 限制：
 - 程式設計語言只能用 Python 或 Java
- 計費標準：
 - 連出頻寬 \$0.12 美元/GB, 連入頻寬 \$0.10 美元/GB
 - CPU 時間 \$0.10 美元/時
 - 儲存的資料 \$0.15 美元/GB-每月
 - 電子郵件收件者 \$0.0001 美元/每個收件者



PaaS / SaaS Key Player :

Google 谷歌



Gmail / 電子郵件



Contact / 通訊錄



Postini / 通訊安全



Calendar / 行事曆



Talk / 即時通



Group / 網上論壇



Doc / 文件



Video / 影音



Sites / 協作平台



Mobile / 行動使用Apps



Admin / 管理後台

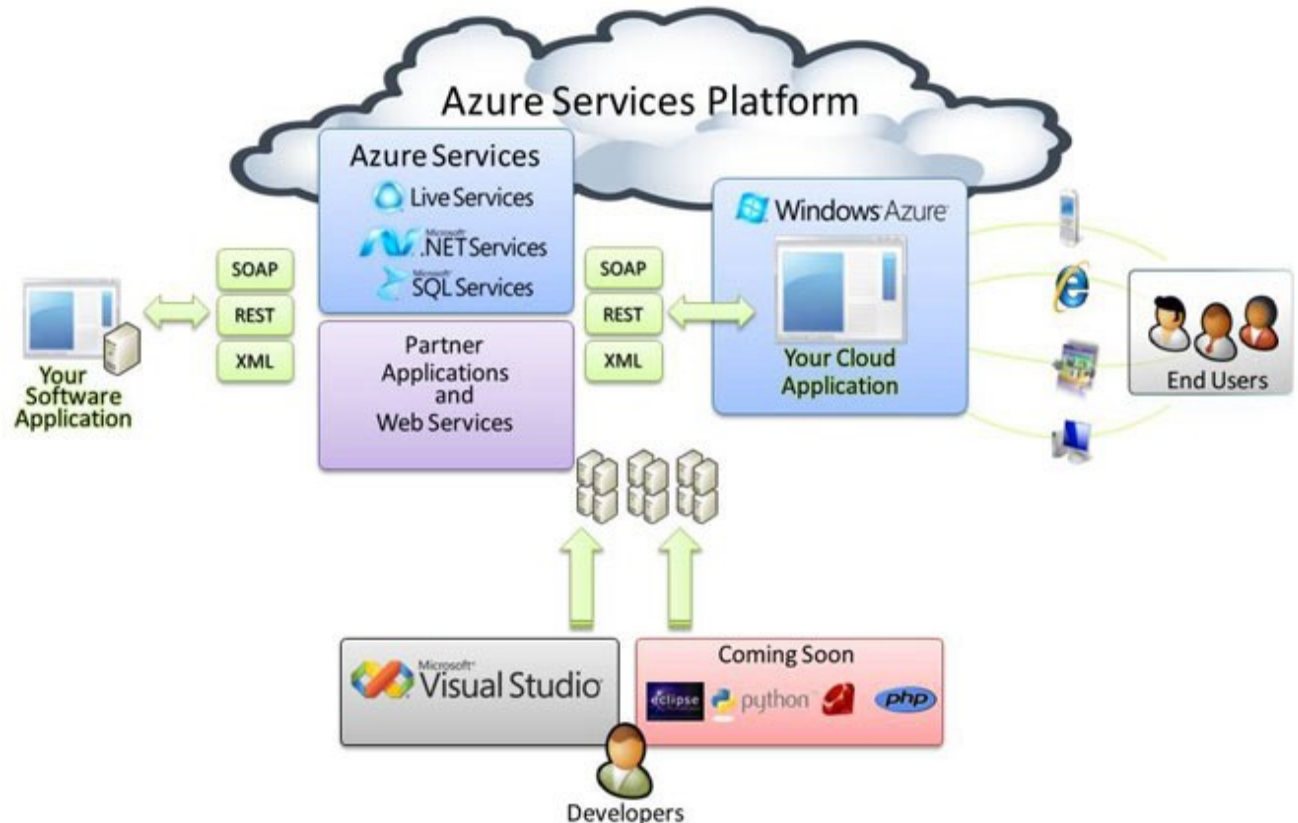
- Google Apps
- Google Apps for Business
- Google Apps for Government
- Google Apps for Non-profits
- Google Apps for ISPs

如果無法掌握雲端技術，至少該學會使用雲端服務！

XaaS Key Player :

Microsoft 微軟

- Microsoft Azure 是一套雲端服務作業系統。
- 作為 Azure 服務平台的開發、服務代管及服務管理環境。
- 服務種類：
 - .Net services
 - SQL services
 - Live services



XaaS Key Player :

Microsoft 微軟

Microsoft Cloud Computing 全貌

Private

Public

Microsoft SharePoint Server
Microsoft Exchange
Microsoft Dynamics

Software as a Service (SaaS)

Microsoft Online Services

Microsoft SharePoint Services

Microsoft Office Live

Microsoft SQL Server
Microsoft .NET

Platform as a Service (PaaS)

Windows Azure

SQL Services

Windows Azure platform
AppFabric

Microsoft System Center
Windows Server

Infrastructure as a Service (IaaS)

Windows Azure

Microsoft System Center

Windows Server

Microsoft Dynamic Data Center Toolkit For Enterprises

Microsoft Dynamic Data Center Toolkit For Hosters

Dallas
→ DaaS

Azure
AppFabric
→ PaaS
(類似 GAE)

SQL Azure
→ PaaS
(雲端 SQL)

Windows Azure
→ PaaS
(類似 EC2)

Hyper-V
→ IaaS
(虛擬化)

IT as a Service

HOW

怎麼做

New IT Architecture toward Cloud Computing !!

用雲掌握資料，加以分析，形成智能給端用



雲

資料中心
提供服務

雲端設計新思維：端的智能來自於雲的服務

Devices share the wisdom of Cloud

端



各類裝置
存取服務

4. 下載最佳路徑

1. 上傳行進軌跡

6. 上傳最佳路徑執行狀態
(判斷是否需要修正)

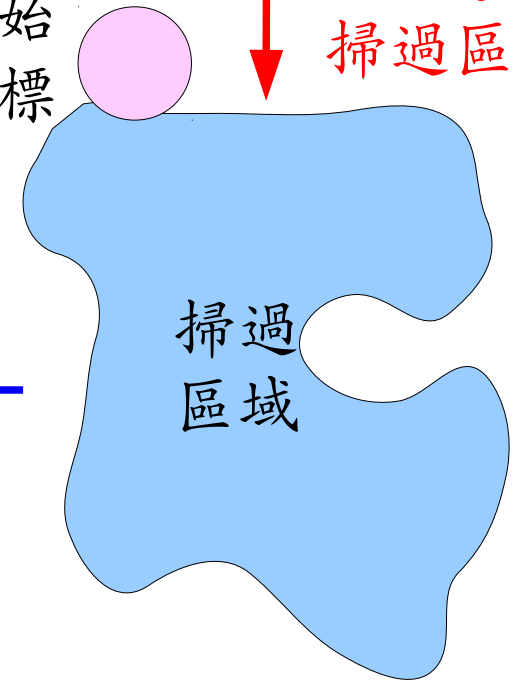
3. 運算
最佳路徑

2. 運算
掃過區域

5. 執行
最佳路徑

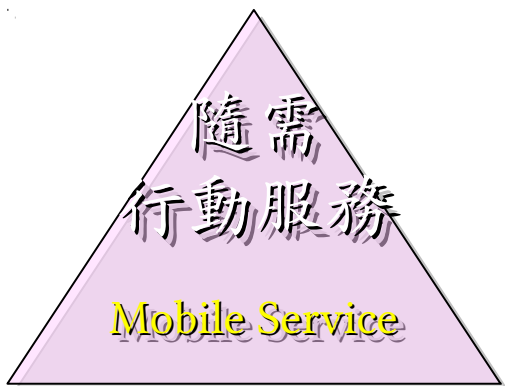


起始
座標



曾經有人問道：我是學電機的，那電機與控制跟雲端有甚麼關聯性？

Three Key Technologies !! 三種服務模式 vs. 三大關鍵技術

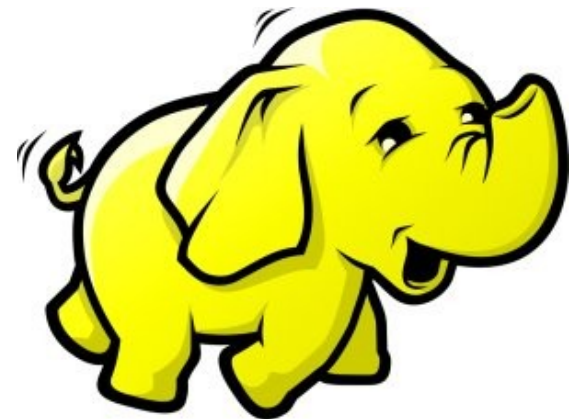




自由軟體在雲端世代中的角色與貢獻

Roles and contributions of Open Source in Cloud Computing Era

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw



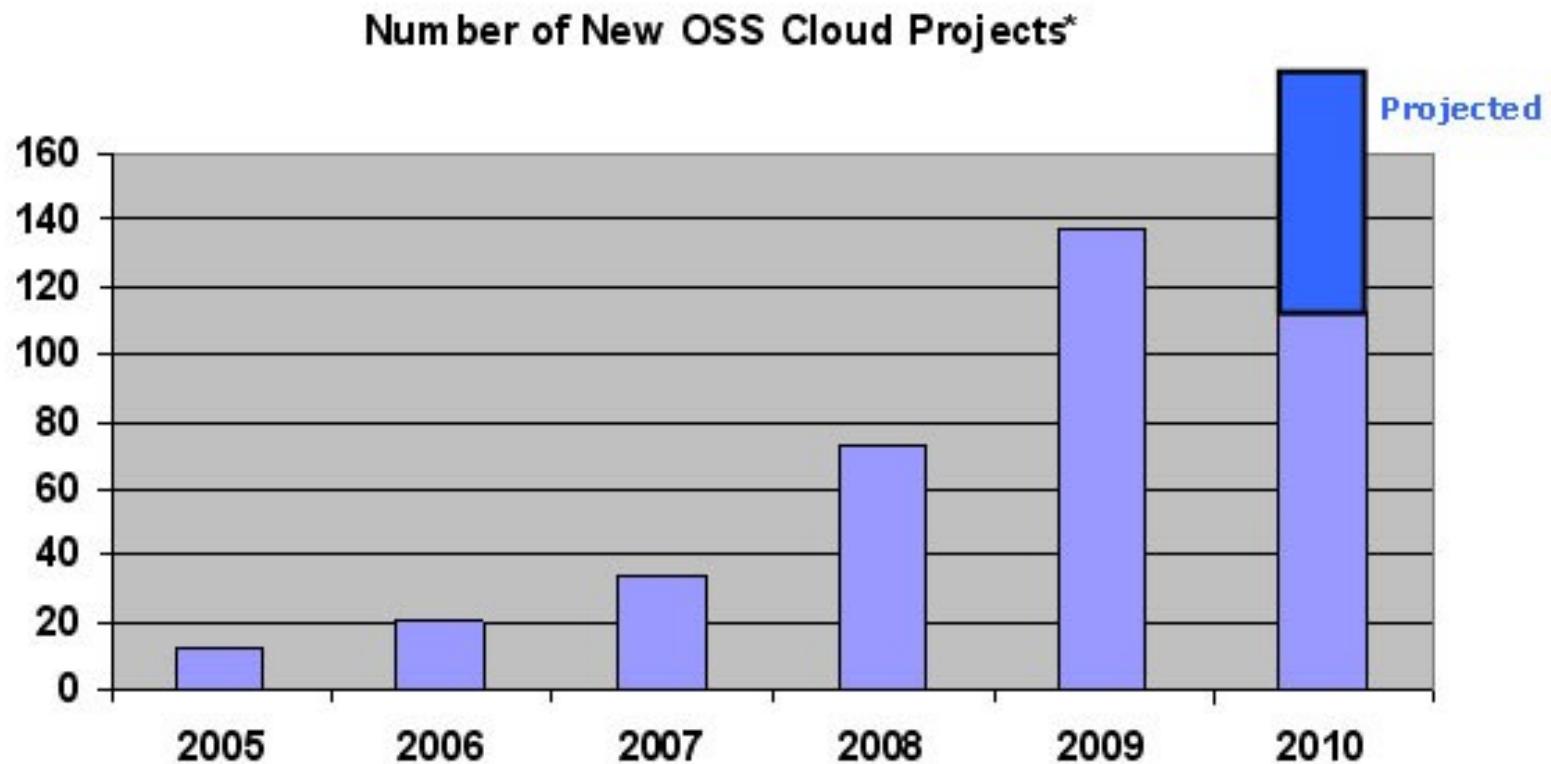
Rapid Growth of Open Source Projects for Cloud

雲端相關的自由軟體專案於過去三年急速成長

Open Source Projects for Cloud on Rise, According to Black Duck Software Analysis

根據 Black Duck 的分析，近年來雲端相關的自由軟體專案有大幅度的成長

Open Source Projects for Cloud on Rise



* Projects specifically referencing cloud. Many OSS projects are suitable for cloud but don't specify it.

圖片來源：<http://www.blackducksoftware.com/oss/projects/>

Three Roles and Contribution of Open Source for Cloud 自由軟體在雲端科技發展中的三大角色與貢獻

Role #1: The foundation of Cloud Services !

角色一：自由軟體是雲端服務的幕後技術基石

Contribution #1: Reduce IT Cost 貢獻一：節省軟體成本

Role #2: Interoperability Standard between Clouds !

角色二：跨越不同雲端服務的互通標準！

Contribution #2: Less Vendor Lock-In 貢獻二：減少技術壟斷

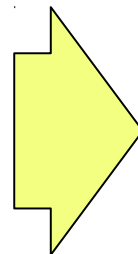
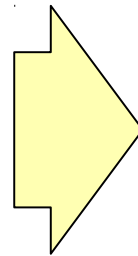
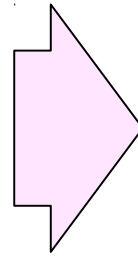
Role #3: Open Platform for Open Education !

角色三：雲端人才培育需要開放平台！

Contribution #3: Open Ladder to Cloud 貢獻三：搭建登雲天梯

Role #1: The foundation of Cloud Services !

角色一：自由軟體是雲端服務的幕後技術基石



Reference Cloud Architecture

雲端運算的參考架構

應用軟體 Application

Social Computing, Enterprise, ISV, ...

程式語言 Programming

Web 2.0, Mashups, Workflows, ...

控制管理 Control

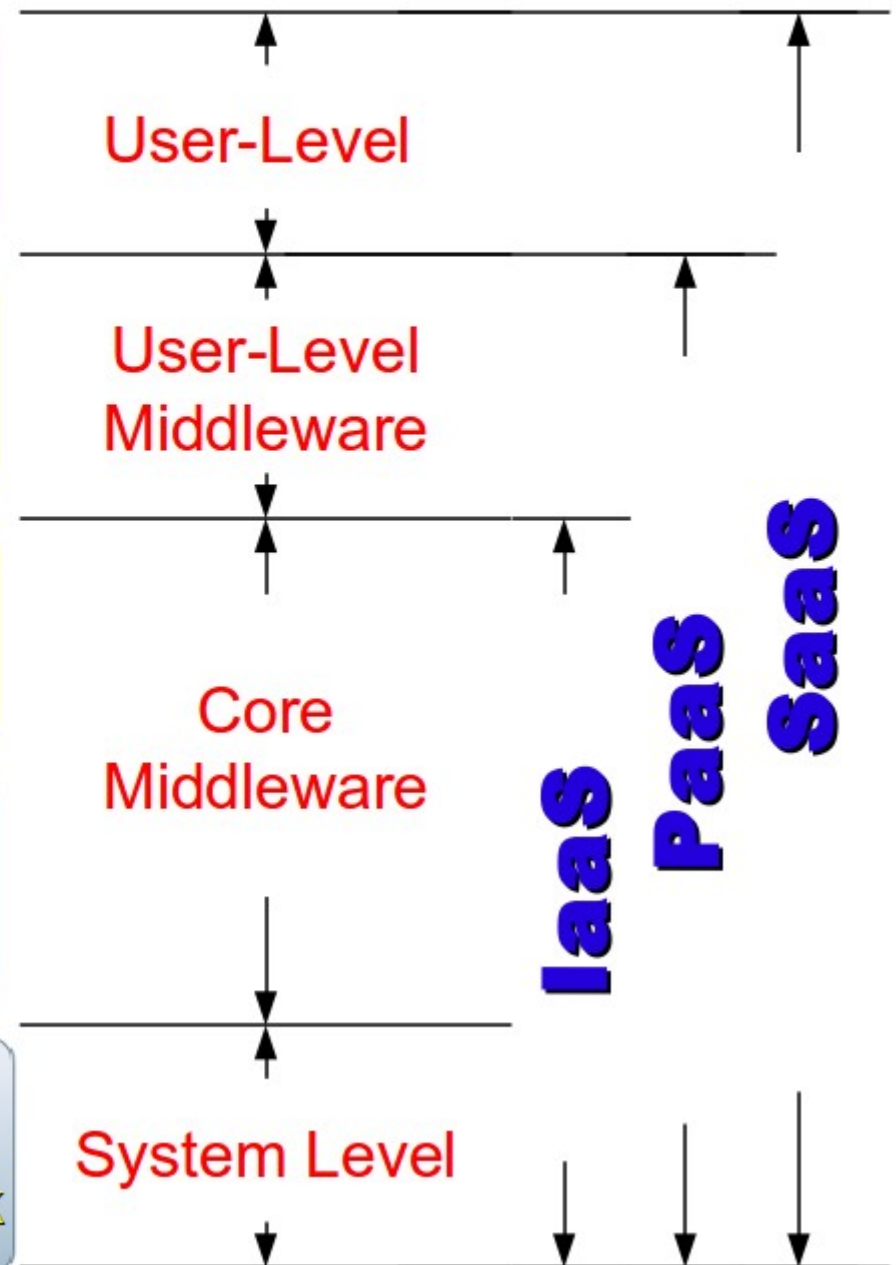
Qos Negotiation, Admission Control, Pricing, SLA Management, Metering...

虛擬化 Virtualization

VM, VM management and Deployment

硬體設施 Hardware

Infrastructure: Computer, Storage, Network



Building IaaS with Open Source

用自由軟體打造 IaaS 服務

應用軟體 Application
Social Computing, Enterprise, ISV, ...

eyeOS, Nutch, ICAS,
X-RIME, ...

程式語言 Programming
Web 2.0, Mashups, Workflows, ...

Hadoop (MapReduce),
Sector/Sphere, AppScale

控制管理 Control
Qos Negotiation, Admission Control,
Pricing, SLA Management, Metering, ...

OpenNebula, Enomaly,
Eucalyptus, OpenQRM, ...

虛擬化 Virtualization
VM, VM management and Deployment

Xen, KVM, VirtualBox,
QEMU, OpenVZ, ...

硬體設施 Hardware
Infrastructure: Computer, Storage,
Network

What is Virtualization ??

虛擬化技術有哪些呢??

Application Virtualization 應用程式虛擬化

Desktop Virtualization
Client Virtualization 桌面虛擬化

Presentation Virtualization 顯示虛擬化

OS-level Virtualization 作業系統虛擬化

Network Virtualization 網路虛擬化

Storage Virtualization 儲存虛擬化

資料庫虛擬化

資料虛擬化

Virtualization

Data

Data Virtualization

大家對虛擬化的
狹義認知

Open Source for Virtualization 虛擬化技術對應的自由軟體

Application Virtualization

Ex. VMWare ThinApp

Desktop
Client

建置虛擬化叢集不可只
注重作業系統虛擬化，
須同時具備底層支援的
網路與儲存虛擬化！！

Redhat SPICE

Presentation

VNC, FreeNX

OS-level Virtualization

Xen, KVM, OpenVZ

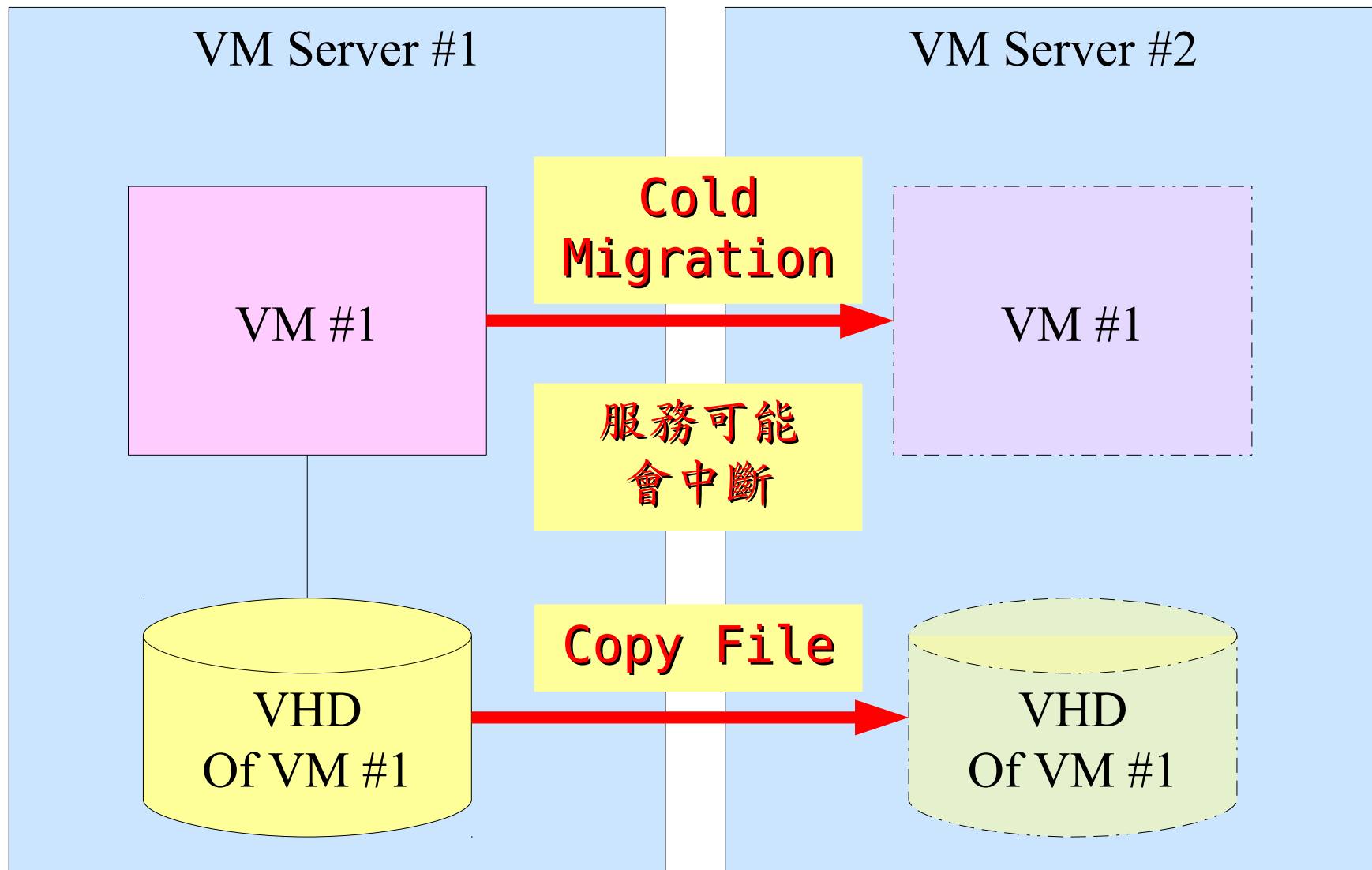
Network Virtualization

OpenFlow vSwitch

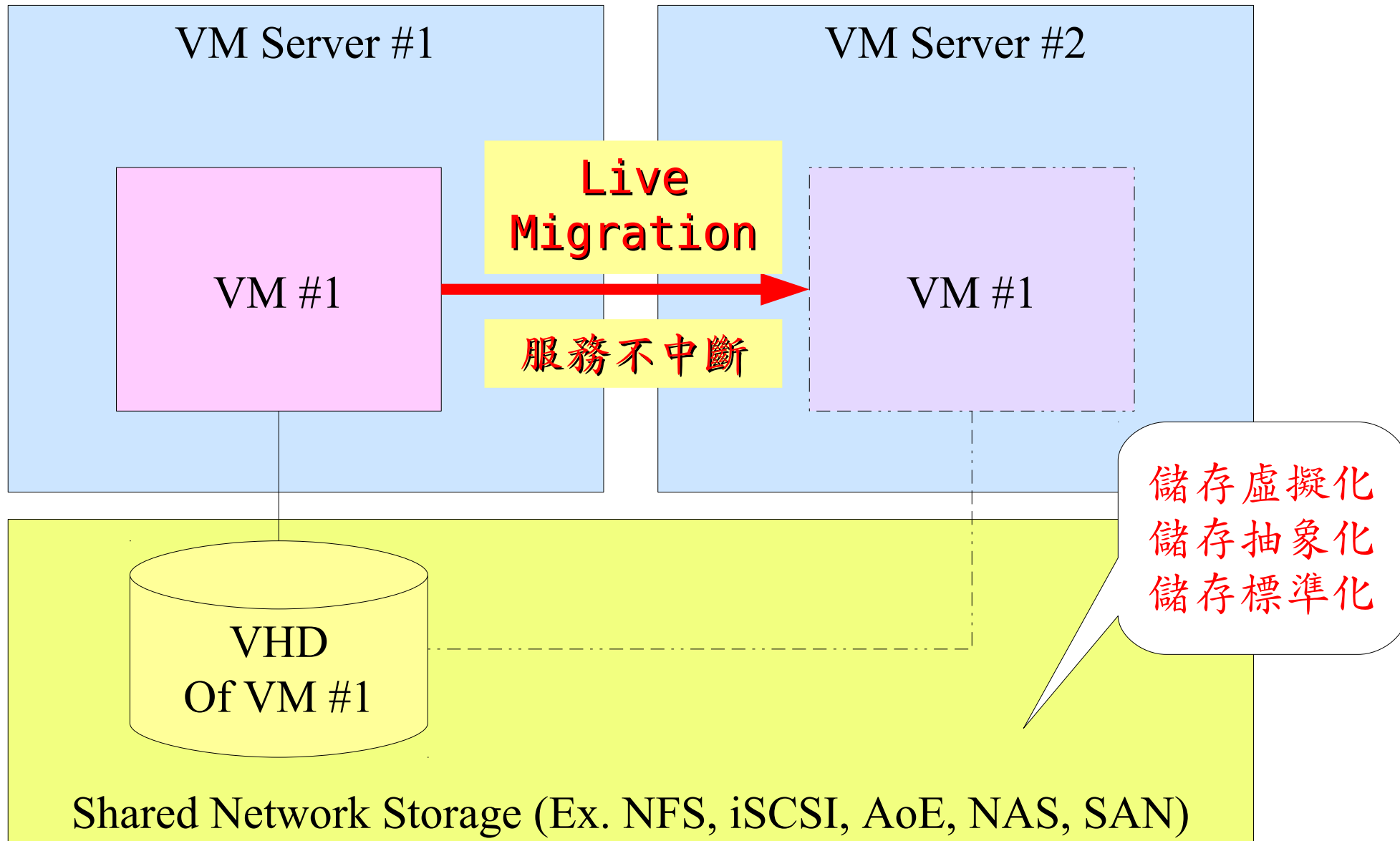
Storage Virtualization

Lessfs, SDFS

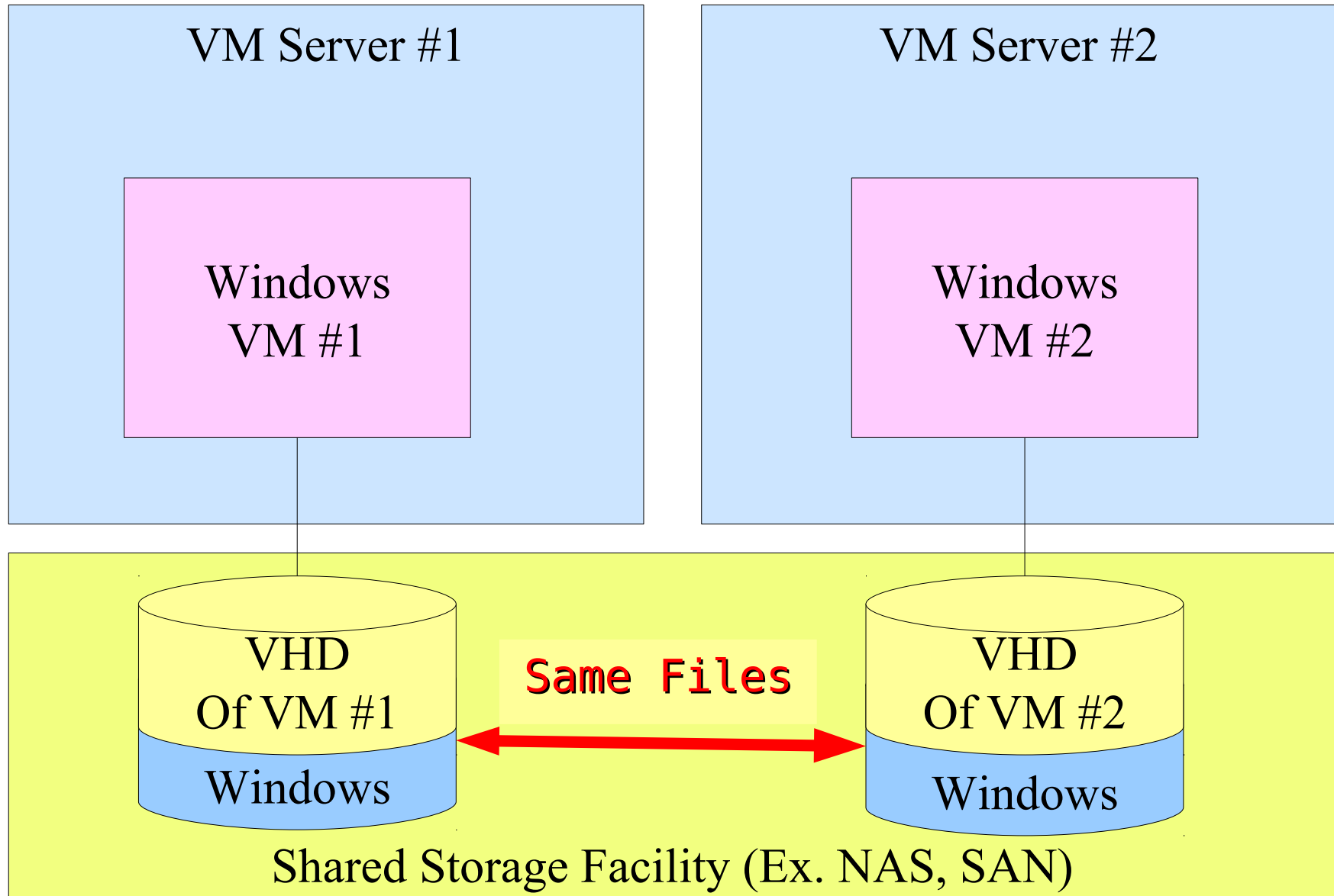
Why Storage Virtualization !! 為何需要儲存虛擬化？



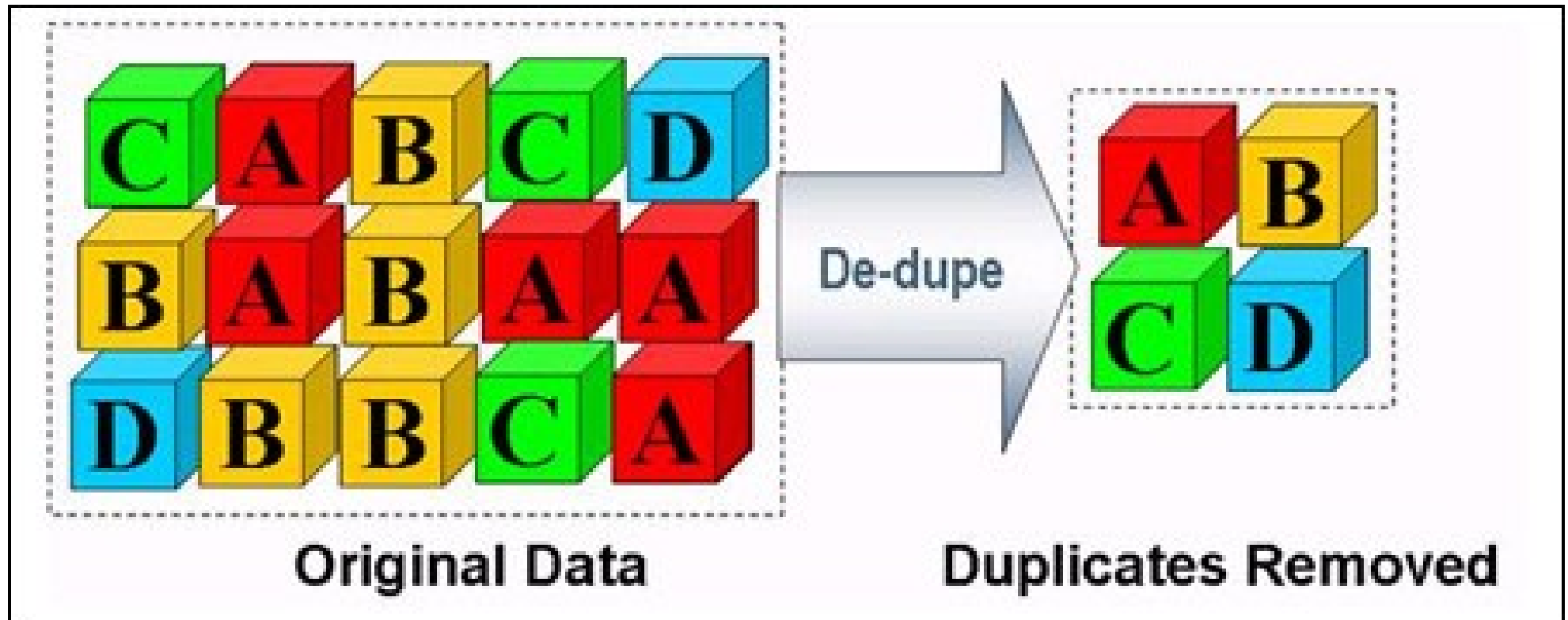
Why Storage Virtualization !! 為何需要儲存虛擬化？



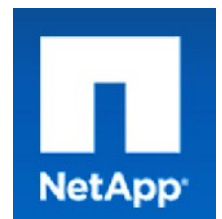
Why Storage Virtualization !! 為何需要儲存虛擬化？



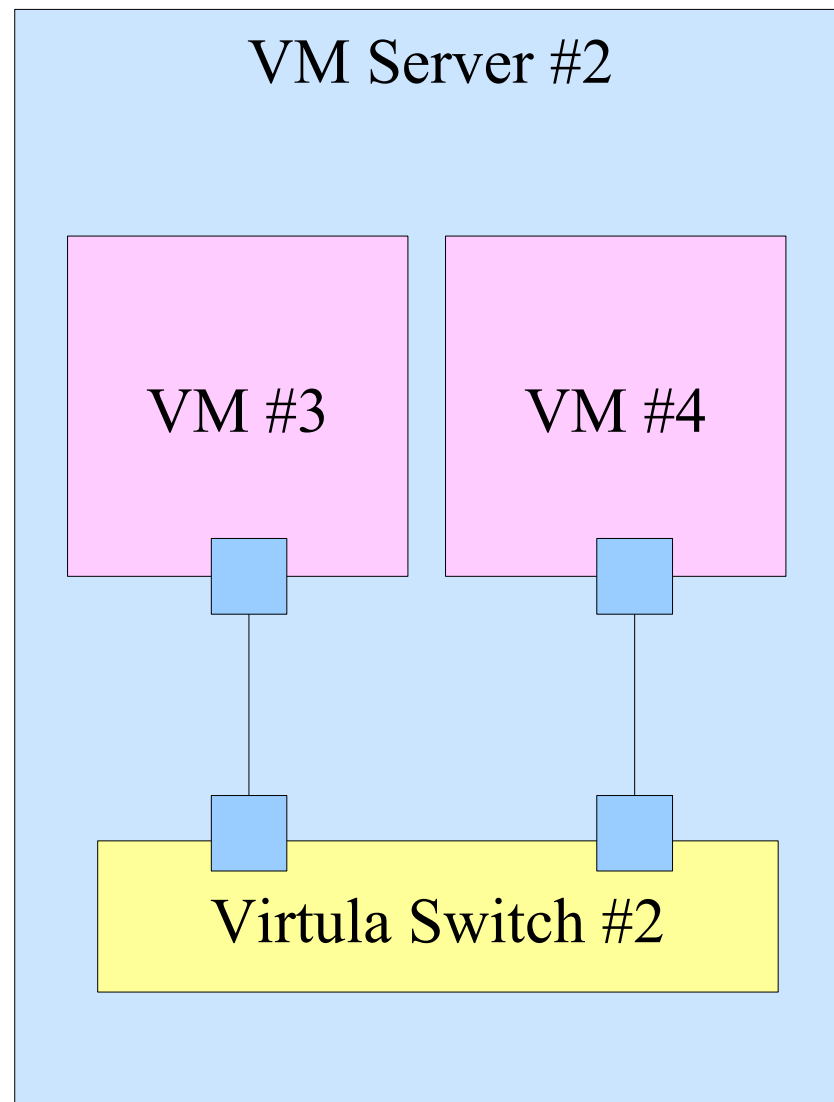
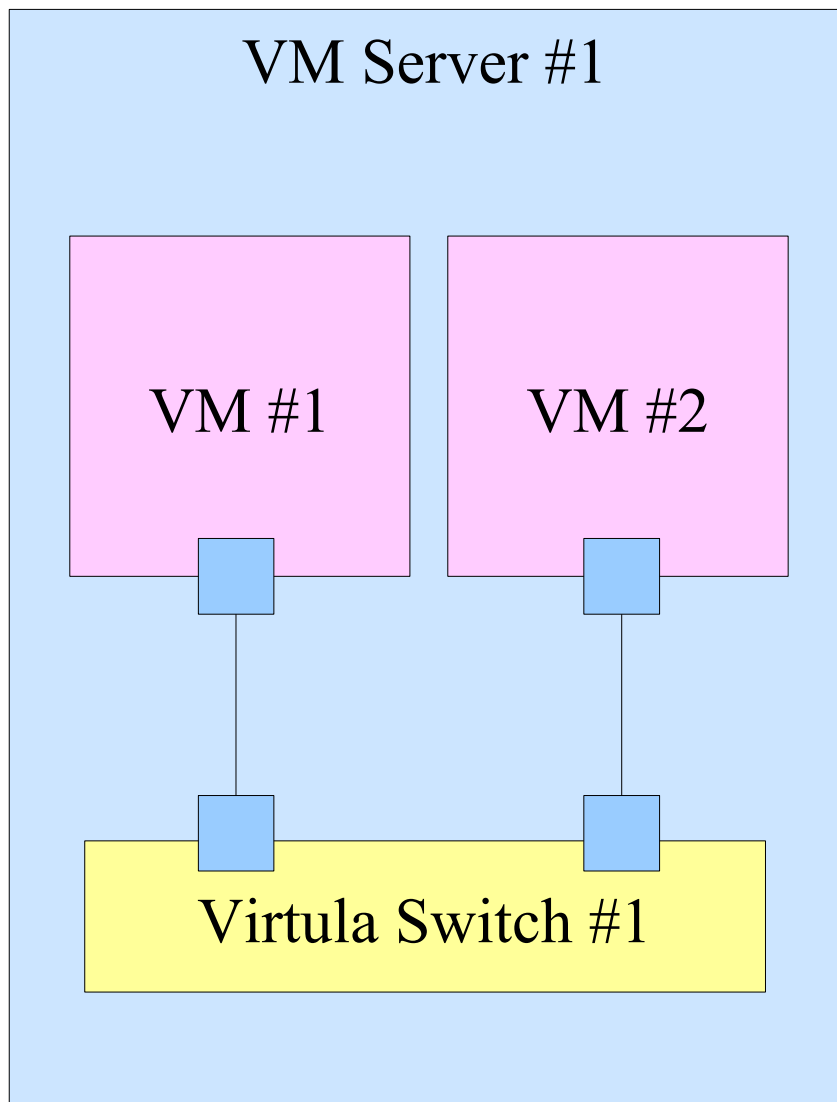
Deduplication? 去除重複儲存的資料?



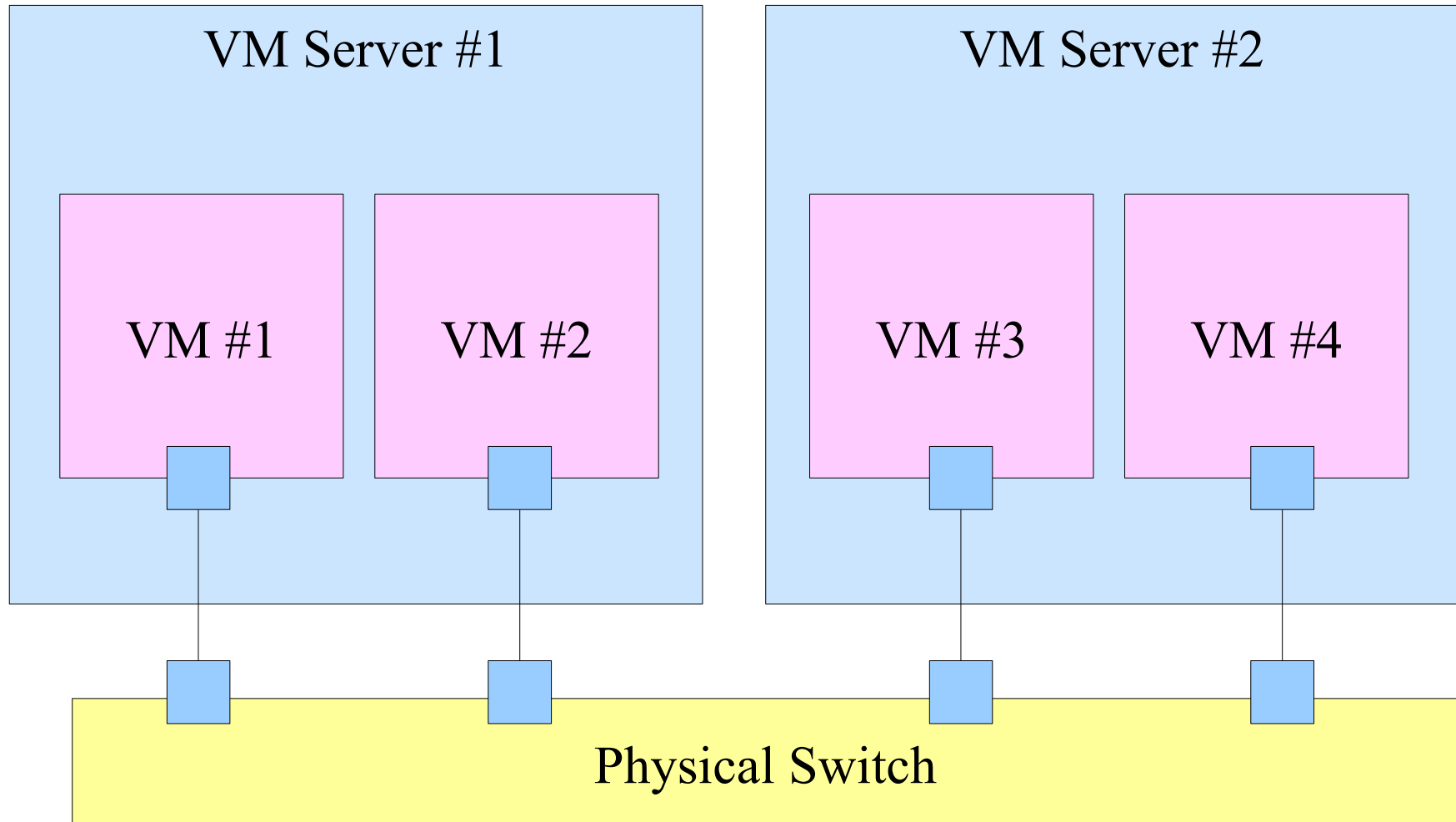
- 資料整合為跨單位整合的第一步 !!
- 商業硬體方案：EMC、NetApp
- 自由軟體方案：
 - ZFS、Lessfs、SDFS...



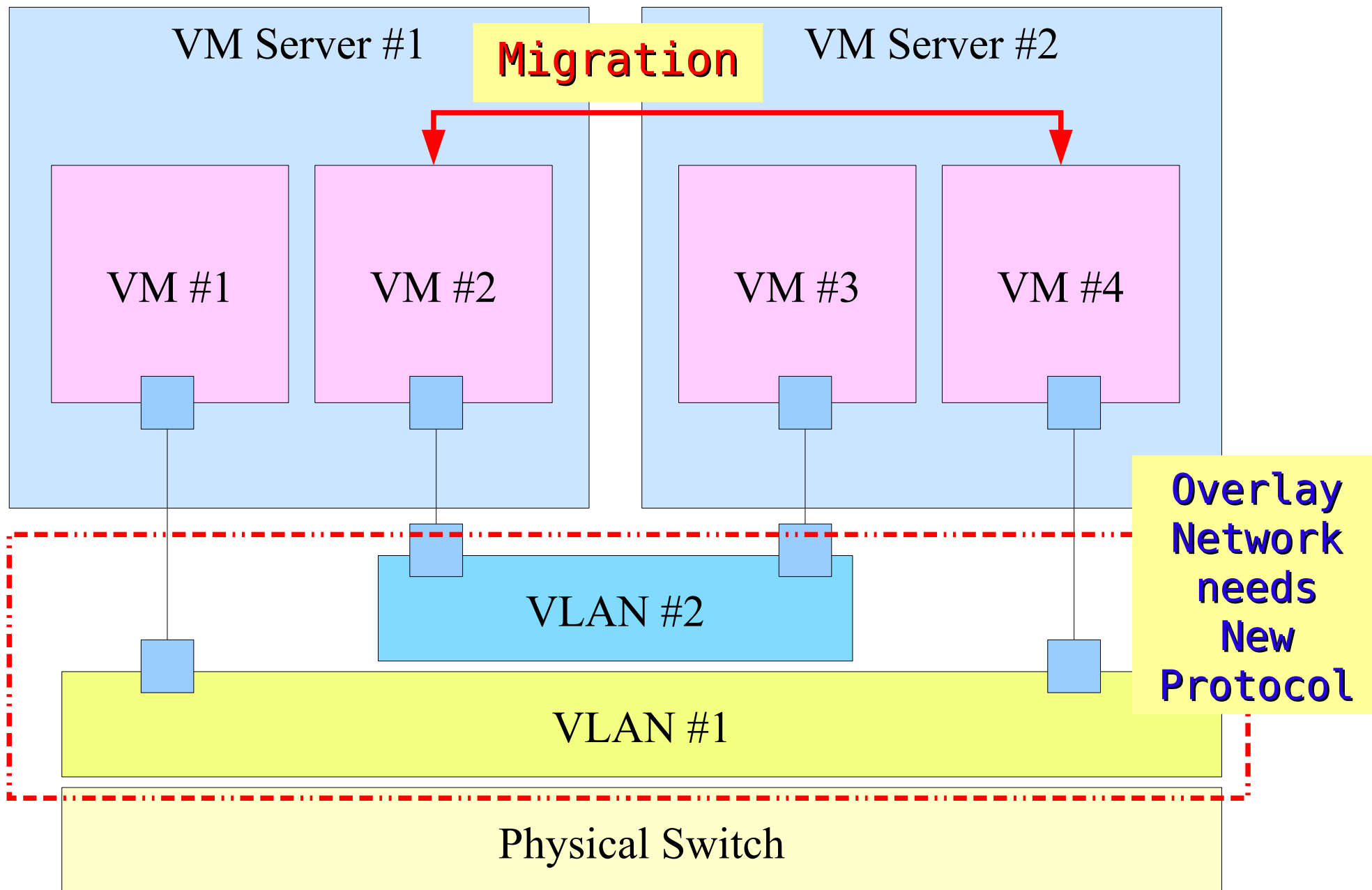
Why Network Virtualization !! 為何需要網路虛擬化？



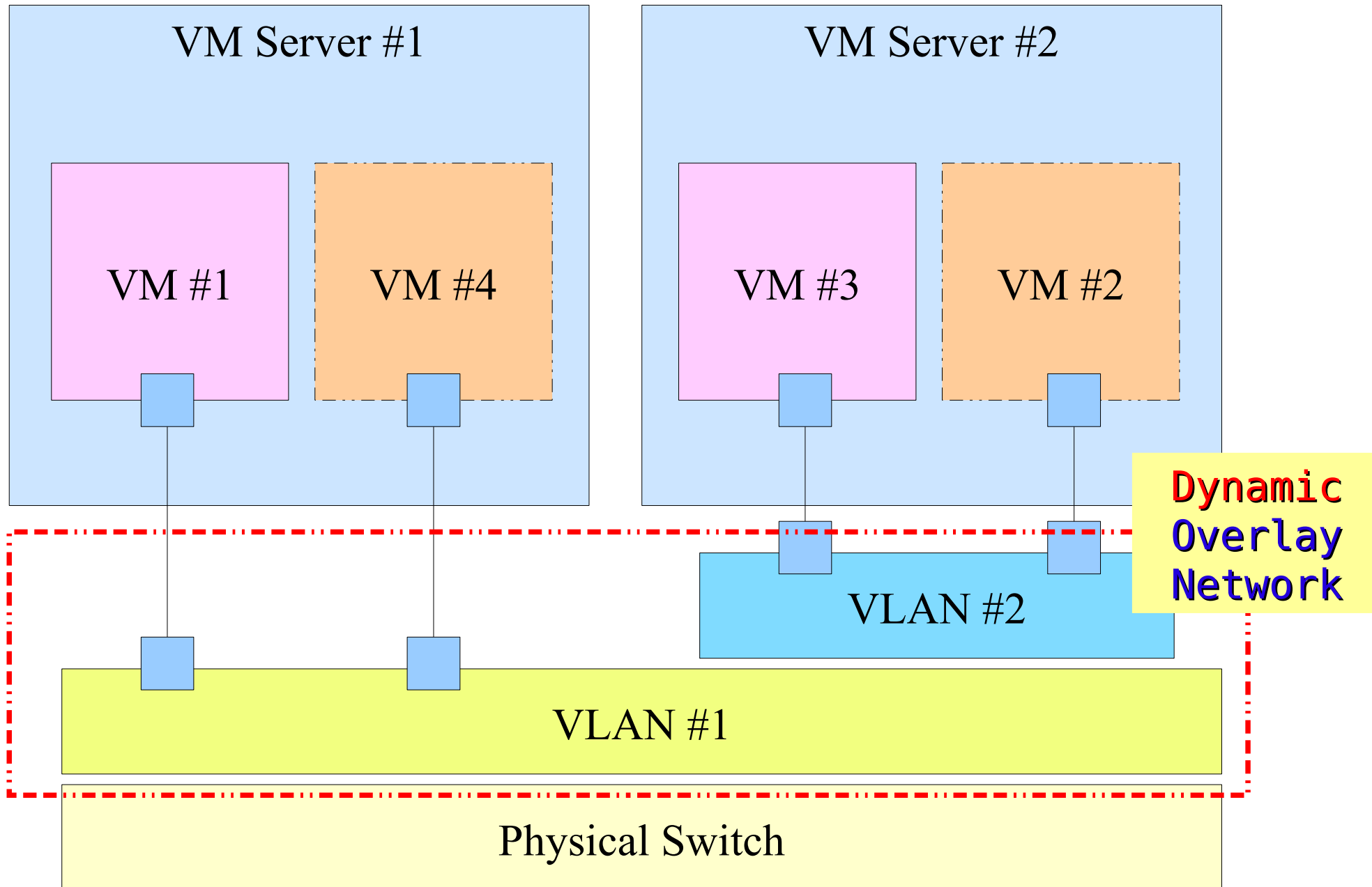
Why Network Virtualization !! 為何需要網路虛擬化？



Why Network Virtualization !! 為何需要網路虛擬化？

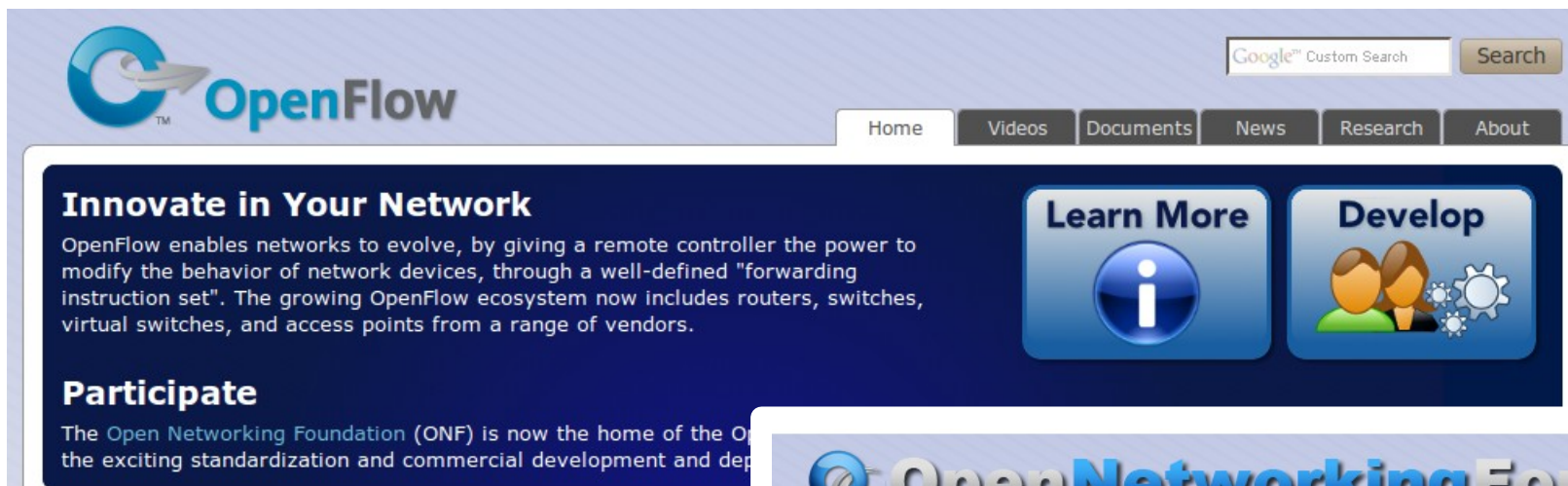


Why Network Virtualization !! 為何需要網路虛擬化？



OpenFlow Network Virtualization 網路虛擬化標準

- Open Networking Foundation 為非營利組織，致力推廣「軟體定義網路 (Software-Defined Networking; SDN)」的新網路方法
- SDN 包含幾個相當重要的元素，其中關鍵為 OpenFlow，其能為服務供應商提供更精確的網路基礎設施管理能力



The screenshot shows the OpenFlow website homepage. At the top left is the OpenFlow logo, which consists of a blue circle with a white arrow pointing right, followed by the text "OpenFlow". To the right of the logo is a search bar with the text "Google™ Custom Search" and a "Search" button. Below the search bar is a navigation menu with buttons for "Home", "Videos", "Documents", "News", "Research", and "About". The main content area features a dark blue banner with the text "Innovate in Your Network" and a paragraph describing OpenFlow's capabilities. To the right of this banner are two buttons: "Learn More" with an information icon and "Develop" with an icon of two people and gears.



The screenshot shows the Open Networking Foundation website homepage. At the top is the Open Networking Foundation logo, which consists of a blue circle with a white arrow pointing right, followed by the text "OpenNetworkingFoundation". Below the logo is a dark blue banner with the text "Welcome to the Open Networking Foundation!" and a paragraph describing the foundation's mission. To the right of this banner is a photograph of network cables plugged into a switch.

<http://www.openflow.org/>

NIST Mapping of Cloud Technologies

美國國家標準局的定義主要鎖定虛擬化技術

5. Measured Service
可被監控與量測的服務

Monitoring / AAA
狀態監控與認證收費機制

1. On-demand self-service
隨需自助服務

VM Management Tool
虛擬機器管理平台

2. Broad network access
隨時隨地用任何網路裝置存取

OS-level Virtualization
作業系統虛擬化

3. Resource pooling
多人共享資源池

Network Virtualization
網路虛擬化

4. Rapid elasticity
快速重新佈署靈活度

Storage Virtualization
儲存虛擬化

NIST Mapping of Open Source Cloud 美國國家標準局的定義對應自由軟體技術

Monitoring / AAA
狀態監控與認證收費機制

Ganglia, Nagios
/ OpenID, SAML

VM Management Tool
虛擬機器管理平台

OpenNebula, Eucalyptus

OS-level Virtualization
作業系統虛擬化

Xen, KVM, OpenVZ

Network Virtualization
網路虛擬化

OpenFlow vSwitch

Storage Virtualization
儲存虛擬化

Lessfs, SDFS, ZFS

- 原是加州大學聖塔芭芭拉分校 (UCSB) 的研究專案
- **It was a research project of UCSB, USA**
- 目前已轉由 Eucalyptus System 這間公司負責維護
- **Now Eucalyptus System provide technical supports.**
- 創立目的是讓使用者可以**打造自己的 EC2**
- **It designed to help user to build their own Amazon EC2**
- 特色是相容於 Amazon EC2 既有的用戶端介面
- **Its feature is compatible with existing EC2 client.**
- 優勢是 Ubuntu 9.04 已經收錄 Eucalyptus 的套件
- **Ubuntu Enterprise Cloud powered by Eucalyptus in 9.04**
- 目前有提供 Eucalyptus 的官方測試平台供註冊帳號
- **You can register trail account at <http://open.eucalyptus.com/>**
- 缺點：目前仍有部分操作需透過指令模式
- **Cons : you might need to type commands in some case**

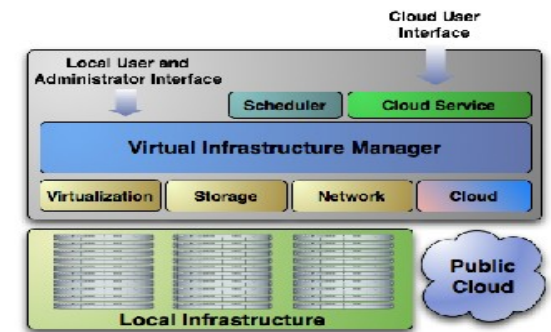


Eucalyptus

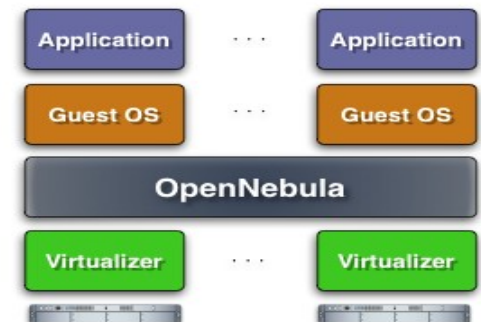
VMM Tool #2: OpenNebula

- <http://www.opennebula.org>
- 由歐洲研究學會 (European Union FP7) 贊助
- **Sponsor by European Union FP7**
- 將實體叢集轉換成具管理彈性的虛擬基礎設備
- Turn Physical Cluster into Virtual Cluster
- 可管理**虛擬叢集**的狀態、排程、遷徙 (migration)
- **manage status, scheduling and migration of virtual cluster**
- [Ubuntu 9.04 provide package of opennebula](#)
- 缺點：需下指令來進行虛擬機器的遷徙 (migration) 。
- **Cons** : You need to type commands to check or migration

OpenNebula.org



關於 OpenNebula 的更多資訊，
請參考 <http://trac.nchc.org.tw/grid/wiki/OpenNEbula>



NIST Mapping of Open Source Cloud 美國國家標準局的定義對應自由軟體技術

Monitoring / AAA
狀態監控與認證收費機制

Ganglia, Nagios
/ OpenID, SAML

VM Management Tool
虛擬機器管理平台

OpenNebula, Eucalyptus

OS-level Virtualization
作業系統虛擬化

Xen, KVM, OpenVZ

Network Virtualization
網路虛擬化

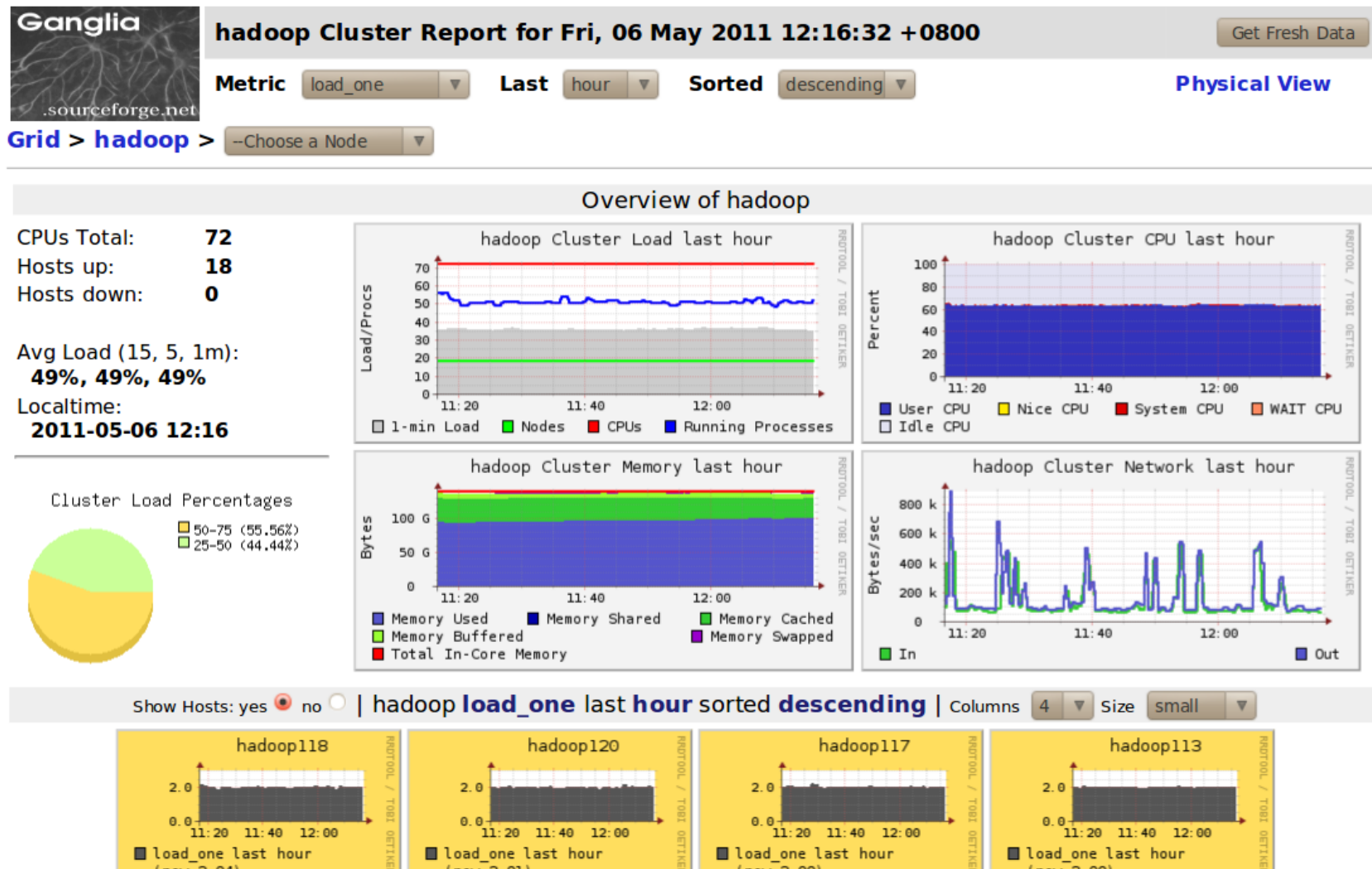
OpenFlow vSwitch

Storage Virtualization
儲存虛擬化

Lessfs, SDFS, ZFS

系統狀態監控 Ganglia

- 採用自由軟體 Ganglia 來蒐集電腦叢集的負載狀態
- <http://ganglia.sourceforge.net/>



日本政府已開始接受 OpenID

2010年3月9日 Nat Sakimura 發表「Government of Japan started accepting OpenID」。

日本經濟產業省開設了一個名為「IdeaBox」的網站，可以使用 mixi、Yahoo!、Livedoor 和 Google 等的帳號來登錄。IdeaBox 是經濟產業省用來向國民廣泛徵求意見，以促進國民間政策的討論的網站。在該網站上，人們可參與出謀劃策和討論，並進行表決。

經濟產業省是想基於網際網路開設一個向公眾開放的委員會。一個類似網站在去年10月到11月間運作並收集了1700多條政策建議。IdeaBox 2月23日開放的這個版本，它還有其它各種社交網站的組件，例如可以直接從這個網站發送 tweet 到 twitter、收藏到 delicious 和 hatena 書籤等。這個版本將運行到3月15日。

網站地址：<http://open-meti.go.jp/>



Building PaaS with Open Source

用自由軟體打造 PaaS 雲端服務

應用軟體 Application
Social Computing, Enterprise, ISV, ...

eyeOS, Nutch, ICAS,
X-RIME, ...

程式語言 Programming
Web 2.0, Mashups, Workflows, ...

**Hadoop (MapReduce),
Sector/Sphere, AppScale**

控制管理 Control
Qos Negotiation, Ddmission Control,
Pricing, SLA Management, Metering...

OpenNebula, Enomaly,
Eucalyptus , OpenQRM, ...

虛擬化 Virtualization
VM, VM management and Deployment

Xen, KVM, VirtualBox,
QEMU, OpenVZ, ...

硬體設施 Hardware
Infrastructure: Computer, Storage,
Network

Hadoop

- <http://hadoop.apache.org>
- Hadoop 是 Apache Top Level 開發專案
- **Hadoop is Apache Top Level Project**
- 目前主要由 Yahoo! 資助、開發與運用
- **Major sponsor is Yahoo!**
- 創始者是 Doug Cutting，參考 Google Filesystem
- **Developed by Doug Cutting, Reference from Google Filesystem**
- 以 Java 開發，提供 HDFS 與 MapReduce API。
- **Written by Java, it provides HDFS and MapReduce API**
- 2006 年使用在 Yahoo 內部服務中
- **Used in Yahoo since year 2006**
- 已佈署於上千個節點。
- **It had been deploy to 4000+ nodes in Yahoo**
- 處理 Petabyte 等級資料量。
- **Design to process dataset in Petabyte**



**Facebook, Last.fm,
Joost, Twitter
are also powered
by Hadoop**

Sector / Sphere

- <http://sector.sourceforge.net/>
- 由美國資料探勘中心研發的自由軟體專案。
- **Developed by National Center for Data Mining, USA**
- 採用 C/C++ 語言撰寫，因此效能較 Hadoop 更好。
- **Written by C/C++, so performance is better than Hadoop**
- 提供「類似」Google File System 與 MapReduce 的機制
- **Provide file system similar to Google File System and MapReduce API**
- 基於 [UDT 高效率網路協定](#) 來加速資料傳輸效率
- **Based on UDT which enhance the network performance**
- [Open Cloud Testbed](#) 有提供測試環境，並開發 [Ma1Stone 效能評比軟體](#)
- **Open Cloud Consortium provide Open Cloud Testbed and develop Ma1Stone toolkit for benchmark**



National Center for Data Mining
University of Illinois at Chicago



Open Data Group
<http://www.opendatagroup.com/>

Building SaaS with Open Source

用自由軟體打造 SaaS 雲端服務

應用軟體 Application

Social Computing, Enterprise, ISV, ...

eyeOS, Nutch, ICAS,
X-RIME, ...

程式語言 Programming

Web 2.0, Mashups, Workflows, ...

Hadoop (MapReduce),
Sector/Sphere, AppScale

控制管理 Control

Qos Negotiation, Admission Control,
Pricing, SLA Management, Metering...

OpenNebula, Enomaly,
Eucalyptus, OpenQRM, ...

虛擬化 Virtualization

VM, VM management and Deployment

Xen, KVM, VirtualBox,
QEMU, OpenVZ, ...

硬體設施 Hardware

Infrastructure: Computer, Storage,
Network

快速佈署內網搜尋引擎的軟體工具 Crawlzilla

- 雲端運算在大型資料分析、資料探勘方面的示範應用
- 可協助企業或校園建置內網的網站、檔案搜尋
- 支援叢集式安裝、中文分詞，並提供網頁管理介面
- 更多資訊詳見專案網站：<http://code.google.com/p/crawlzilla>

Web UI (Crawlzilla Website + Search Engine)

JSP + Servlet +
JavaBean

Nutch

Lucene

Crawlzilla System Management

Tomcat

Hadoop

PC1 (Master)

PC2

PC3

PC4

CrawlZilla 網頁管理介面



HOME

Crawl

資料庫管理

系統狀態

使用者設定

登出系統

管理介面功能介紹

* 抓取網頁設定：透過網頁介面建立搜尋引擎索引

* 資料庫管理：查詢、刪除已爬取的索引資料庫

* 系統狀態：瀏覽目前系統狀態

* 管理員設定：包含個人化與語言設定

* 修改密碼

搜尋引擎快速連結

CrawlZilla 搜尋引擎範例

3@books

3@businessweekly

3@yahoo

系統功能

修改管理員密碼

相關資源

CrawlZilla@GoogleCode



Role #2: Interoperability Standard between Clouds ! 角色二：跨越不同雲端服務的互通標準！

IBM、思科（Cisco Systems）等 20 多家科技公司於 2009 年 3 月 30 日發佈「開放雲端宣言」（Open Cloud Manifesto），希望為雲端運算建立標準，確保業者間的相容性，讓客戶更容易轉換平台。

「客戶使用雲端運算的經驗愈多，就愈明白相容性有多重要。他們不想被綁死在單一業者。」
— IBM 開放標準部門副總裁諾斯沃西（Karla Norsworthy）



2009年四月28日，Open Grid Forum 宣佈成立 Open Cloud Computing Interface (OCCI) 工作小組，以協助制定雲端服務之間的互通國際標準。



ABOUT OGF RESOURCE CENTER EVENTS DOCUMENTS AREAS/GROUPS
MEMBERS **NEWS** STANDARDS CONTACT US SITE MAP GRIDFORGE

OGF ANNOUNCES OPEN CLOUD COMPUTING INTERFACE (OCCI) WORKING GROUP

OGF Officially Launches Working Group To Create An API For Cloud Infrastructure Services

CHICAGO, 28 April 2009 The Open Grid Forum (OGF) has officially launched the Open Cloud Computing Interface Working Group (OCCI-WG). Its aim is the rapid development of a clean, open API for cloud infrastructure delivered on-demand. Infrastructure as a Service (IaaS) is one of three primary segments of the emerging cloud industry in which compute, storage and network resources are provided as a service. The API will be based on existing service provider functionality, and will cover management of the entire life-cycle of virtual machines and other similar workloads (including virtual private servers and physical dedicated servers). It will as much as possible complement and build on existing work such as the Open Virtualisation Format (OVF), as well as serving as an integration point for others standards efforts including DMTF and SNIA.

Craig Lee, OGF's President, stated: "OCCI is an extremely important development that will define a much needed, open standard API in the cloud space. This is, however, just the first step of many. We are committed to working with all stakeholders and organizations to move the ball forward. As always, OGF strives to facilitate productive technical environments for work across the spectrum of distributed computing technologies."

Existing specifications for cloud infrastructure services (including the Amazon EC2 API and Sun Cloud APIs) are provided by single vendors and OCCI will be the first multi-vendor initiative to deliver a royalty-free, open standard API. This will improve cloud computing interoperability and increase competition in the space, facilitating the emergence of a marketplace of services, which will in turn increase customer choice and decrease the risk of vendor lock-in. OCCI will also be important enabler for the creation of hybrid cloud architectures that bridge multiple data centers and cloud services, through adapters and tooling that support the open API. This and related open standards will lower the cost of migration to and from public clouds, delivering automated management of peak capacity, minimizing outages and reducing costs.

OCCI will rely heavily on existing Internet standards and work already done by cloud providers, so as it is easy to both implement and consume. This will benefit software vendors and service consumers alike. For example software vendors who wish to deliver their applications as a service

圖片來源：<http://www.ogf.org/News/news.php?id=132>

數位時代 Beta 3.0 - Topics and Links新聞精選

全球廿五間科技公司將攜手合作，計畫促進更健全、更開放的雲端科技。

這項稱作OpenStack的合作計畫在今天公布，以IT代管服務公司Rackspace和美國太空總署NASA為首，將以Open Source（開程式碼）的方式，把雲端平台的技術分享到世界每一個角落。

有了OpenStack以後，任何一個想要提供雲端服務的單位，只要準備好伺服器 and 網路，就可以下載OpenStack的開放軟體，架設自己的雲端平台。目前這個計畫的網站提供了兩套軟體：OpenStack Compute用來設定和管理大量的伺服器，而OpenStack Storage則提供一套雲端儲存系統，只要使用多台一般市面上出售的伺服器，就可以儲存上千兆個位元組的資料。

Rackspace和NASA在這項計畫中，都有卓著的貢獻。Rackspace開放了自家公司產品「雲端檔案」和「雲端伺服器」的程式碼，這兩項產品是該公司雲端設備的基礎。NASA也將貢獻出他們用以架設「Nebula雲端平台」的技術。

NASA Nebula是目前全世界最強大的雲端平台之一。上週Microsoft Research所公布的最新火星影像，就是在NASA的Nebula平台上，可用Microsoft Research的影像，目前也存放在Nebula

藉由開程式碼的方式，增加不同雲端科技產品之間的互通性。

以Nebula平台的

參與這項計畫的公司還有Intel、Dell、Citrix等，而其他知名公司如Microsoft也在陸續洽談中。這個計畫的目標，是希望為雲端科技提供一個良好的開發環境，同時藉由開程式碼的方式，增加不同雲端科技產品之間的互通性。這樣一來，雲端科技一旦有了新的突破，不同的產品都能夠同時受惠。

Open source software for building private and public clouds.

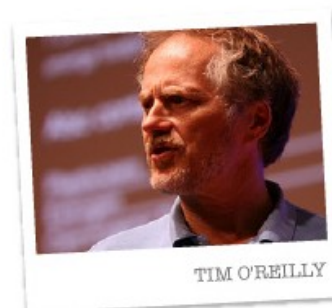
Software

OpenStack Software delivers a massively scalable cloud operating system. The three major components are:

[Compute](#)[Details & Download](#) ▶

Community

Join our global community of technologists, developers, researchers, corporations and cloud computing experts.

1004**77**

"Rackspace and NASA are taking an amazing step towards my vision of an future."

REILLY.COM



OpenStack新版開源軟體Bexar，將在Ubuntu上實現雲端技術

撰文者：戴佳慧

發表日期：2011-02-08

開放原始碼計畫OpenStack上週發表新版雲端軟體Bexar，讓企業用戶能輕鬆在自家伺服器上搭建雲端平台，實現雲端虛擬技術。

OpenStack計畫由積極開發雲端科技的Rackspace和NASA所領導，致力於促進雲端平台的開放。目前OpenStack計畫包含了三個部份：OpenStack Compute雲端運算、Object Store雲端儲存、以及Glance影像傳遞服務。

這次代號Bexar的新版本發佈，是OpenStack自去年夏天成立以來推出的第二版雲端軟體，首度讓新加入的用戶能夠自由部署OpenStack虛擬平台。Bexar加入的新功能包括了IPv6支援、影像傳遞技術、以及Hyper-V和Xen等虛擬伺服器功能。

OpenStack同時宣佈了新夥伴Canonical、Cisco等四家業者的加入。Canonical承諾，會把OpenStack軟體整合進即將在四月發表的Ubuntu伺服器11.04版，協助推廣OpenStack開放雲端平台。

OpenStack也預告了下一個版本Cactus的發行計畫，強調將為大規模部署提供最佳化方案。OpenStack將於四月底在美國加州舉辦設計高峰會，邀請開發人員共同來擬定OpenStack今年下半年的產品規劃藍圖。

Role #3: Open Platform for Open Education !

角色三：雲端人才培育需要開放平台！

<http://www.captivating-content.com/Hidden%20Treasure.jpg>

第一波雲端運算
的成長主要帶動了
台灣硬體產業的
經濟復甦 ...

建再多的資料中心
也是要有**人**管



第二波雲端運算
的成長將決定於
人才培育速度能否
快速銜接產業需求

<http://www.sysadminaday.com/images/people/136-3697.JPG>

Supply Chain of Cloud Industry !! 雲端產業的供應鍊

應用軟體
供應商



端

行動裝置
共通平台



各類裝置
存取服務

軟體服務
供應商



雲

資料中心
機房維運



資料中心
提供服務

基本硬體
建設組件



Relation of different departments !! 各科系人才如何供給雲端產業

應用軟體
供應商



資工

資管

美術

其他

行動裝置
共通平台



電機

資工

軟體服務
供應商



資工

資管

統計

資料中心
機房維運



機械

土木

資管

基本硬體
建設組件

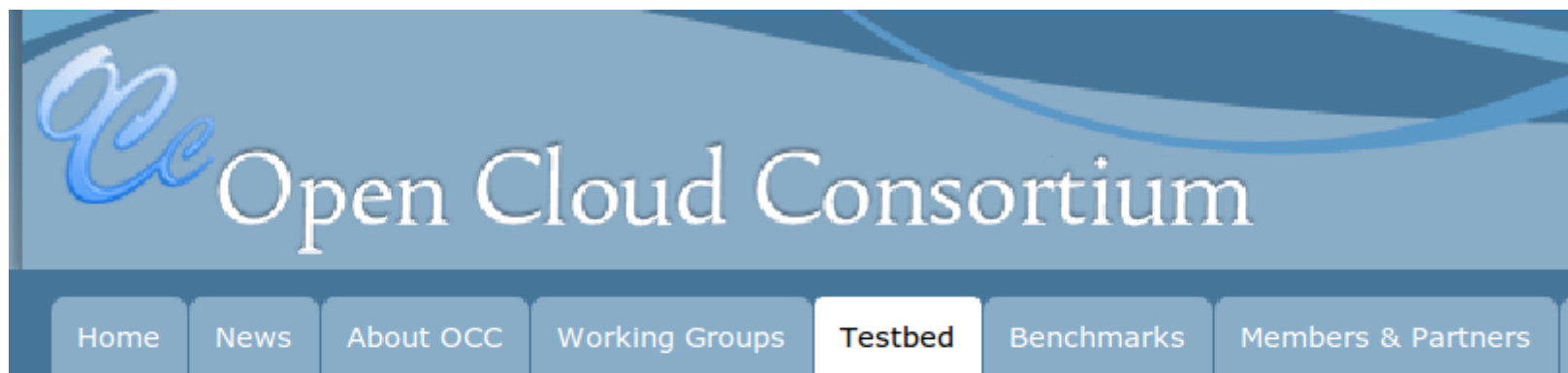


電機

資工

美國結合產學資源，成立 Open Cloud Consortium，
並建立 Open Cloud Testbed，作為人才培育與研究的開放平台。

會員包括：Aerospace、Booz Allen Hamilton、Cisco、Infoblox、MIT Lincoln Labs、NASA Ames Research Center、Northwestern University、Open Data Group、Raytheon、Sector Project、SIOS、University of Illinois at Chicago、Yahoo、Calit2、Johns Hopkins University、National Lambda Rail、University of Chicago



Testbed

The Open Cloud Consortium (OCC) operates a wide area testbed for cloud computing called the Open Cloud Testbed. The testbed is being rolled out in phases. We are currently in Phase 2 of the roll out.

As far as we know, the Open Cloud Testbed is the only wide area cloud that utilizes wide area 10 Gb/s networks. If you know of any other such clouds, please let us know.

國網中心 2009 年開放實驗叢集間接促成研究成果

- 國網中心於 2009 年四月啓用 Hadoop 實驗叢集，目前共 21 台 PC 服役中。
- 促成台大資工系資訊網路與多媒體研究所發表論文至 ACM Multimedia 2009
- 自 2009 年四月至 2010 年八月，雲端實驗叢集共註冊 937 人，服務 37 個學術單位、5 個研究單位、19 間業界公司、2 所醫院。累計於 14 個月執行 16,424 個 Job。
- 註冊排行前五大依序為交通大學、台灣大學、成功大學、中央大學與陽明大學

Canonical Image Selection and Efficient Image Graph Construction for Large-Scale Flickr Photos

Liang-Chi Hsieh, Kuan-Ting Chen, Chien-Hsing Chiang, Yi-Hsuan Yang, Guan-Long Wu
Chun-Sung Ferng, Hsiu-Wen Hsueh, Charng-Rurng Tsai, Winston H. Hsu
National Taiwan University, Taipei, Taiwan

viirya@gmail.com, ktchen@cmlab.csie.ntu.edu.tw, {pacifistboy, affige}@gmail.com, {b95109, b95108, b95057, b95093, winston}@csie.ntu.edu.tw

ABSTRACT

Efficient image search clustering is prominent for image search engines for exponentially growing photo collections. In this work, we propose an image search clustering approach which selects multiple canonical images from image search results and constructs image clusters in real time on an image sub-graph for the search results. The efficiency is achieved with the help of offline-computed image context graphs by distributed computing methods. Extending our prior works, we demonstrate the results of the proposed canonical image selection and preliminary outcomes of large-scale image graph construction in this proposal. We experiment in Flickr550 dataset, containing 540,321 Flickr photos.

Categories and Subject Descriptors: H.3.5 [Informa-

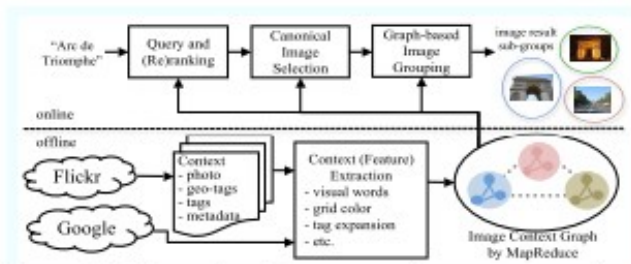
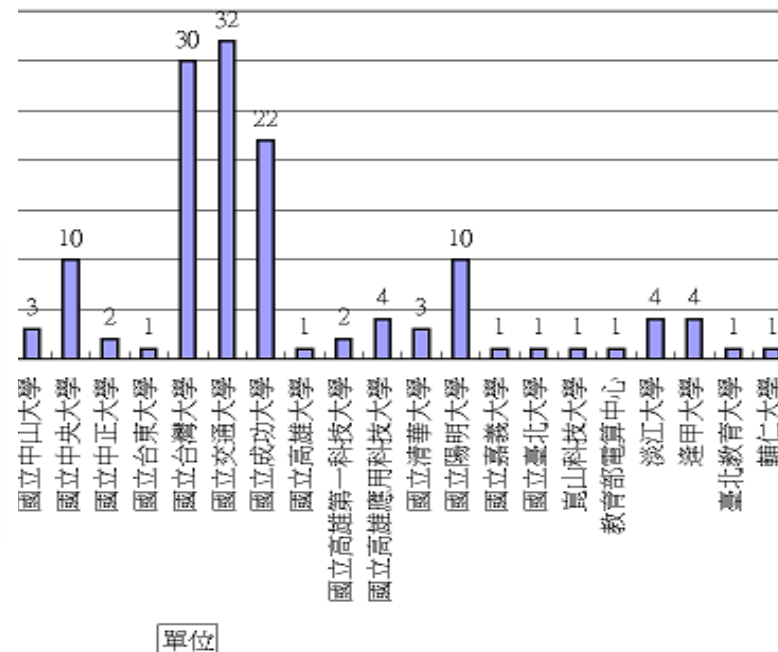


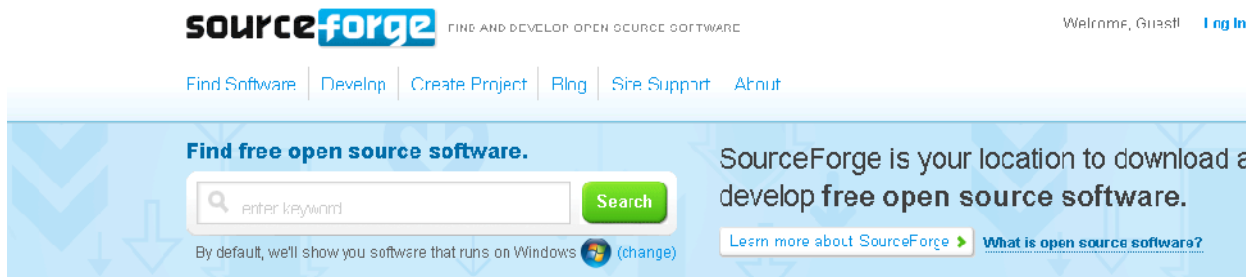
Figure 1: The system diagram: online query reranking, canonical image selection, and image result grouping are based on the image context graph, constructed offline by Hadoop MapReduce over rich context features in Flickr photos.



Conclusion: Open Source as a Strategy !!

未來自由軟體將成爲吸引注意力的一種策略

當免費、開放才能吸引使用者或開發者的注意力～
自由軟體就會變成一種行銷策略，將不得不開放！！



The screenshot shows the SourceForge homepage. At the top left is the SourceForge logo with the tagline "FIND AND DEVELOP OPEN SOURCE SOFTWARE". To the right, it says "Welcome, Guest!" and "Log In". Below the logo are navigation links: "Find Software", "Develop", "Create Project", "Blog", "Site Support", and "About". A main banner area contains the text "Find free open source software." and "SourceForge is your location to download and develop free open source software." There is a search bar with the placeholder "enter keyword" and a "Search" button. Below the search bar, it says "By default, we'll show you software that runs on Windows" with a globe icon and "(change)". There are also links for "Learn more about SourceForge" and "What is open source software?".

Browse by category

Clustering

Communications

What's Hot for Windows?



1

XBMC Media Center

XBMC Media Center is a free cross-platform media



85% recommend

Project Hosting on **Google** code

Search projects

Explore hosted projects:

[Python](#)

[JavaScript](#)

[Django](#)

[Web](#)

[CMS](#)

[Google](#)

[Java](#)

[AJAX](#)

[Rails](#)

[Plugin](#)

[Student](#)

[Android](#)

[CPlusPlus](#)

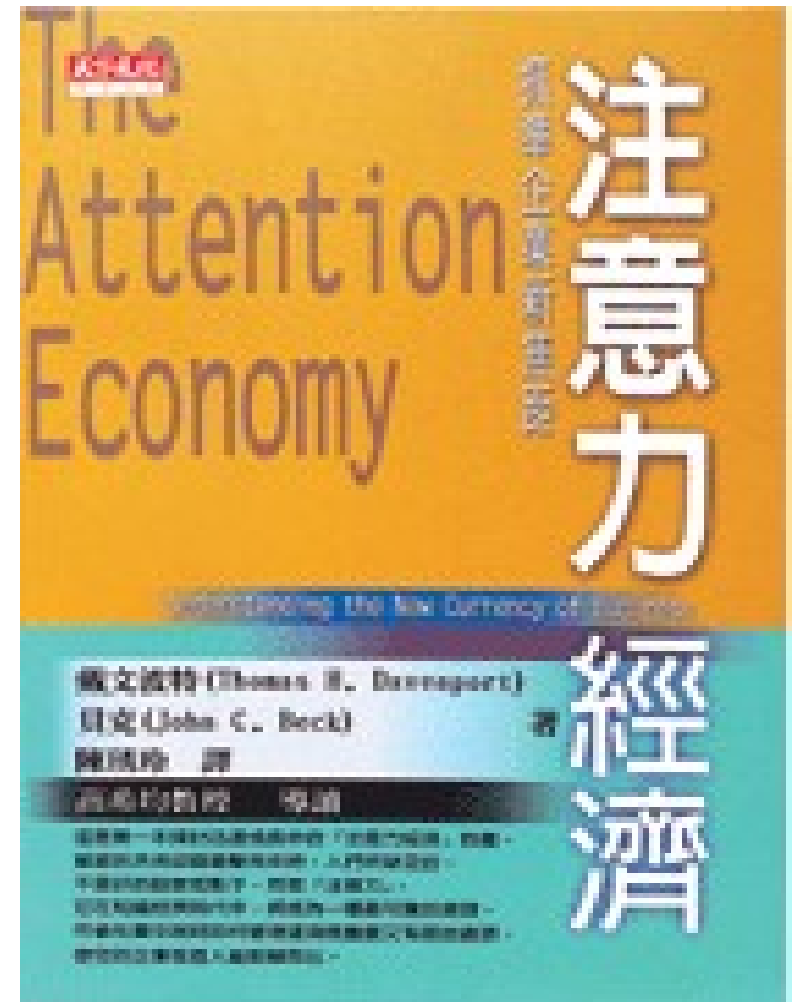
[MySQL](#)

[.Net](#)

[Game](#)

[Windows](#)

[AppEngine](#)





Questions?

Slides - <http://trac.nchc.org.tw/cloud>

Jazz Wang
Yao-Tsung Wang
jazz@nchc.org.tw

