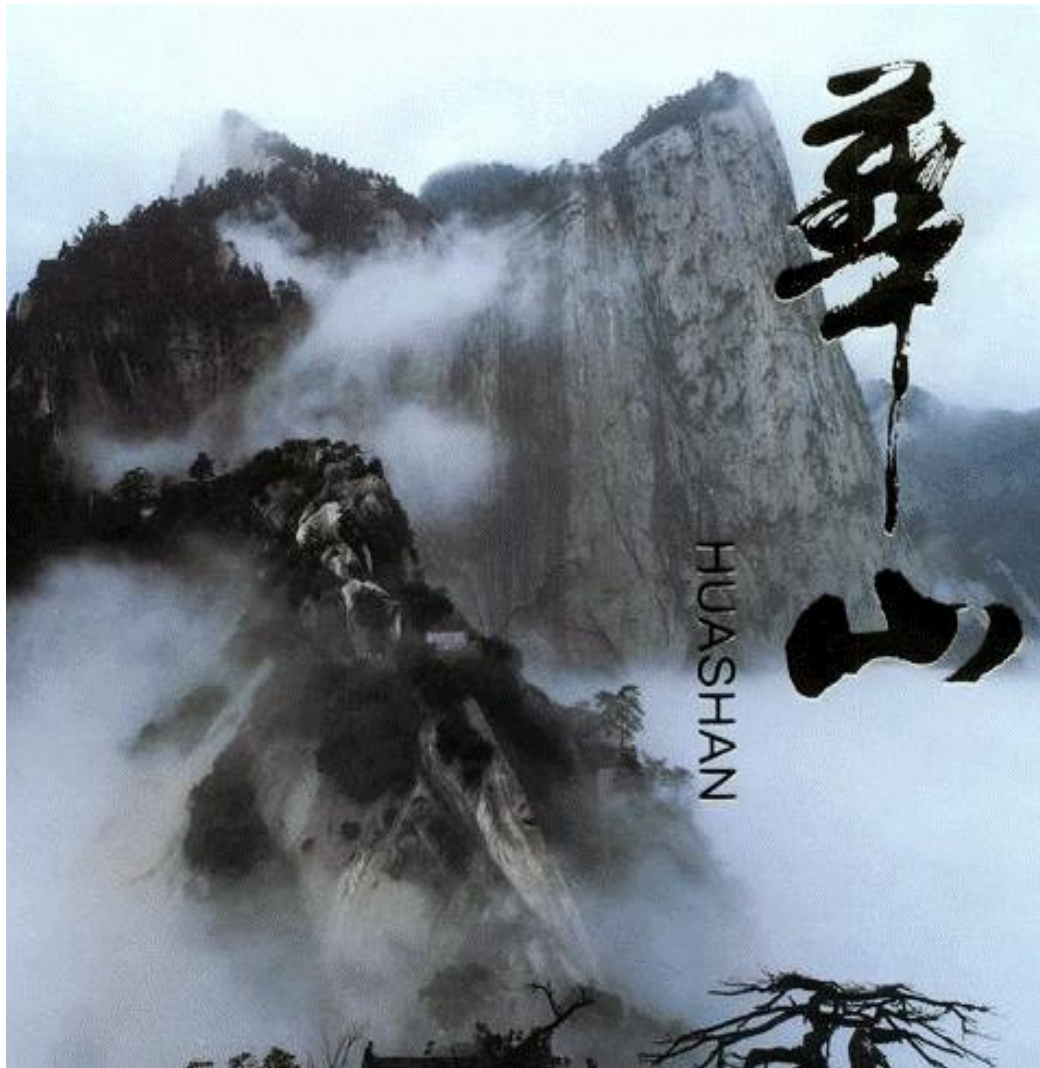


Introduction to **Cloud Computing**

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



只有天上在
更無山與齊
舉頭紅日近
回首白雲低

宋 寇準(華山)

我們將討論到

1. 雲端運算是啥？
2. 雲端運算的類型有哪些？
3. 如何建置自己的雲端？

Distributed Computing

Utility Computing

Virtualization

SOA

Web

SaaS
(Software as a Services)

Cloud

Computing?

Grid Computing

IaaS
(Infrastructure as a Services)

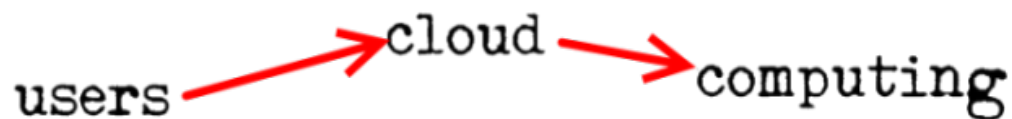
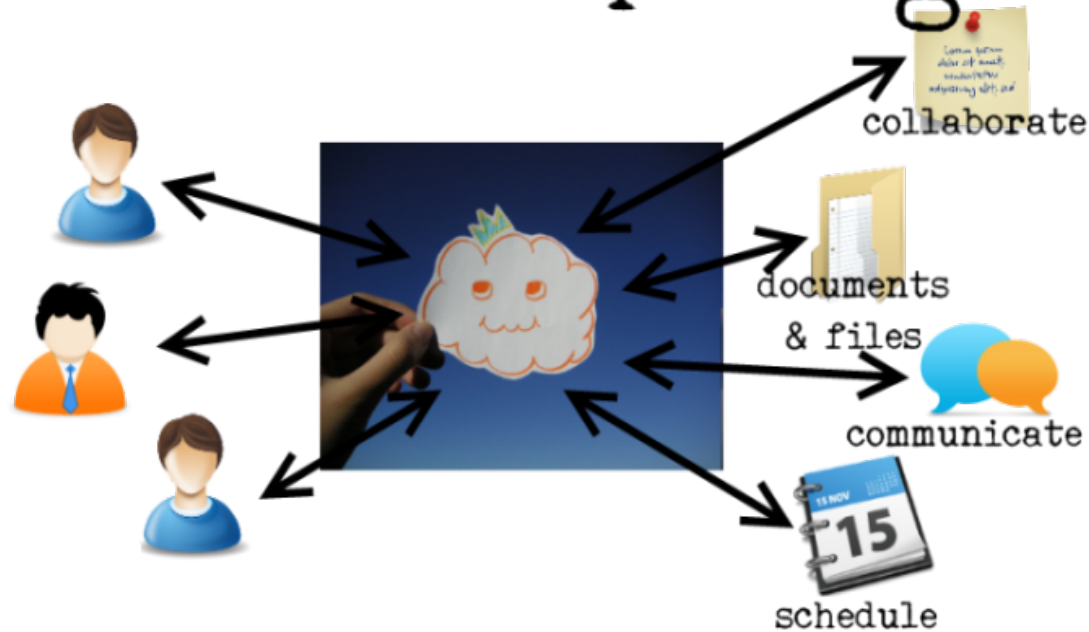
PaaS
(Platform as a Services)

Storage



吾道(雲端運算)一以貫之
孔夫子

cloud computing



凡能透過 **Internet**
得到的 **Services**

(差別只在於規模和類型)

任何一個國家都有自己的歷史，只有了解這些**歷史**，
才能理解我們今天面臨的問題。危險、挑戰和機會。

邱吉爾



What's **history** of
Cloud Computing



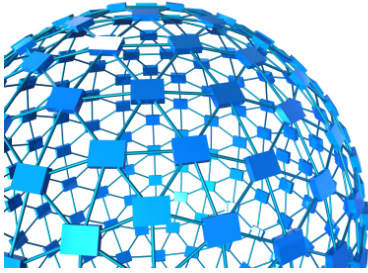
Data Explode Cloud computing

- 2005 Utility computing (Amazon EC2/S3)
- 2006 Apache Hadoop
- 2007 Cloud computing (Google + IBM)



Virtual Org. Grid computing

- 2002 Berkley Bonic
- 2003 Globus toolkit
- 2004 EGEE gLite



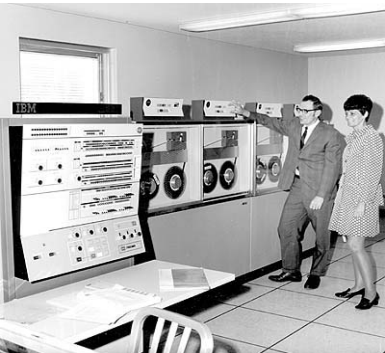
Internet Distributed computing

- 1990 網路對公眾開放
- 1991 HTML, http, WWW
- 1993 Mosaic (web browser)
- 1996 Internet



PC/Linux Clustering computing

- 1975 Apple I
- 1981 IBM PC
- 1982 TCP/IP
- 1985 Windows
- 1991 Linux



Mainframe Super computing

- 1946 Von Neuman
- 1964 IBM System/360
- 1969 Unix
- 1972 C

History tell us...

趨勢一

資料重回集中管理

(1.隨時存取 2.降低資料遺失風險 3.減少資料傳輸成本 4.促進協同作業)

趨勢二

高速計算的興盛

(1.平行計算的技能 2.分散計算的技能 3.多核心程式設計 4.處理大量資料的能力)

趨勢三

網頁成為預設開發平台

(1.網頁是開放標準 2.瀏覽器成為跨平台的工具 3.網頁程式設計成為顯學)

地球萬事萬物，在天堂都有個理想版本，
其重要性不在於他們是否存在，
而在我們無暇的追求

柏拉圖



Give me a perfect
definition
to Cloud Computing

Wikipedia -

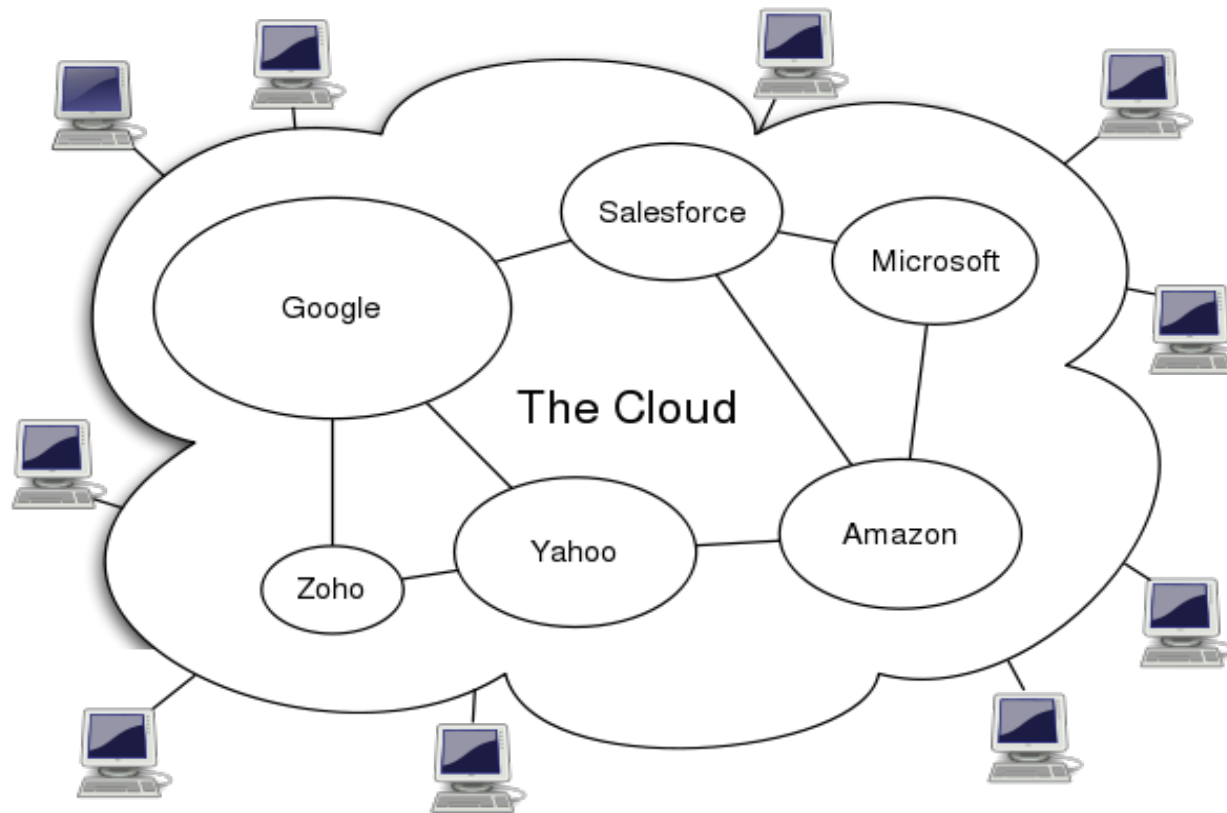
雲端運算是一種運算方式，
運算資源是動態擴充而且虛擬化的，
並透過網際網路以服務的形式來提供

使用者不需要了解「雲端」中基礎設施的細節，
不必具有相應的專業知識，也無需直接進行控制

Cloud computing is a style of computing in which dynamically [scalable](#) and often [virtualised](#) resources are provided [as a service](#) over the Internet ~ Wikipedia

較實務的定義 -

雲端運算描繪了如何透過網際網路從遠端的大量資料中心，得到軟體、儲存或是運算的服務



那雲端運算可分哪些類呢？



一款米養百樣人
台灣俗諺

依據 **Services Layer** 分類

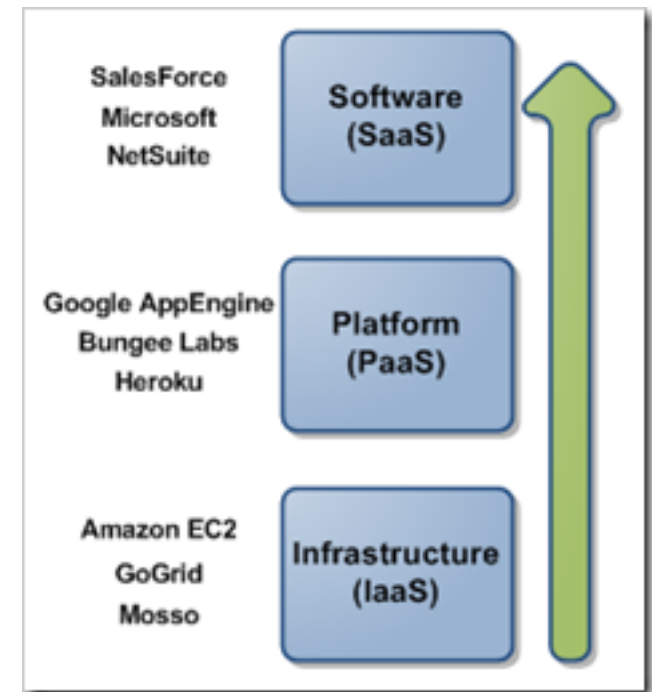
軟體及服務 – **SaaS**

位於最高層，將完整的應用做為服務隨需提供
平台及服務 – **PaaS**

位於中層，透過網路將基礎的儲存運算功能作為
標準化的服務來提供

基礎架構及服務 – **IaaS**

位於最底層，透過網路將基礎的儲存運算功能作為
標準化的服務來提供



依據 **Deployment** 分類

公有雲 - **Public Cloud**

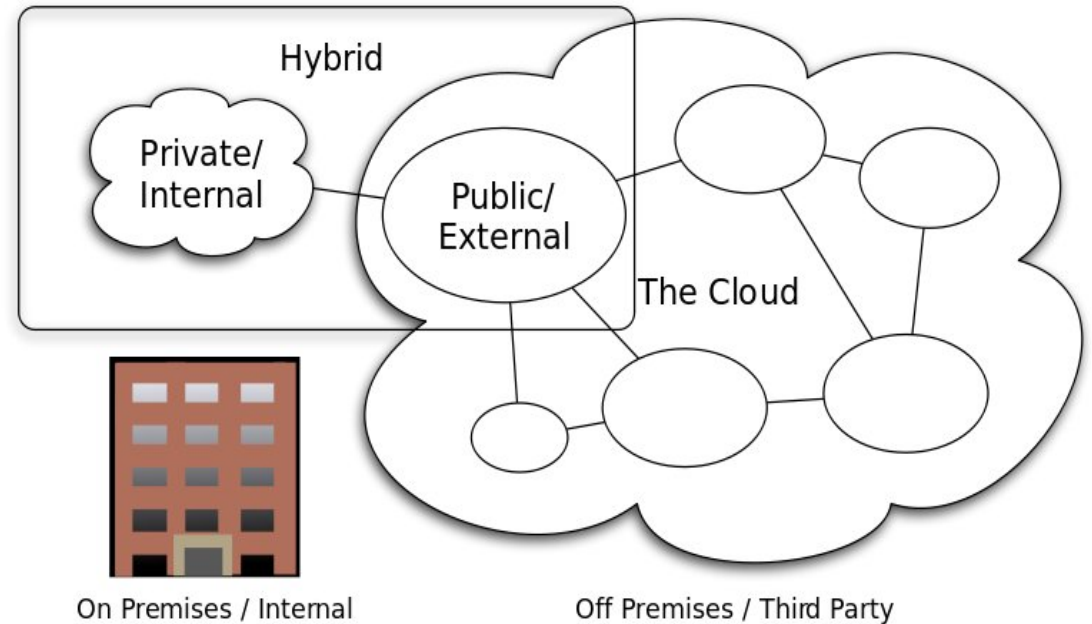
來自於協力廠商，可讓不同的客戶再其伺服器上運作(在雲端內廠商擁有的伺服器、儲存系統和其他基礎架構上混同執行)

私有雲 - **Private Cloud**

單一客戶所有，客戶可控制各應用的執行位置(客戶擁有server, storage, network)

混合雲 - **Hybrid Cloud**

兩者的混合，通常攸關公司核心或機密的應用在私有雲運作，公有雲則為一般事物運作



Cloud Computing Types

小王子：「建立關係」

狐狸說：「不錯。對我來說，你只不過是個小孩，跟其他成千成萬的小孩沒有分別，我不需要你，你也一樣不需要我。我對於你也只不過是一隻狐狸，跟成千成萬其他的狐狸一模一樣。

但是，假如你馴養我，我們就彼此互相需要。

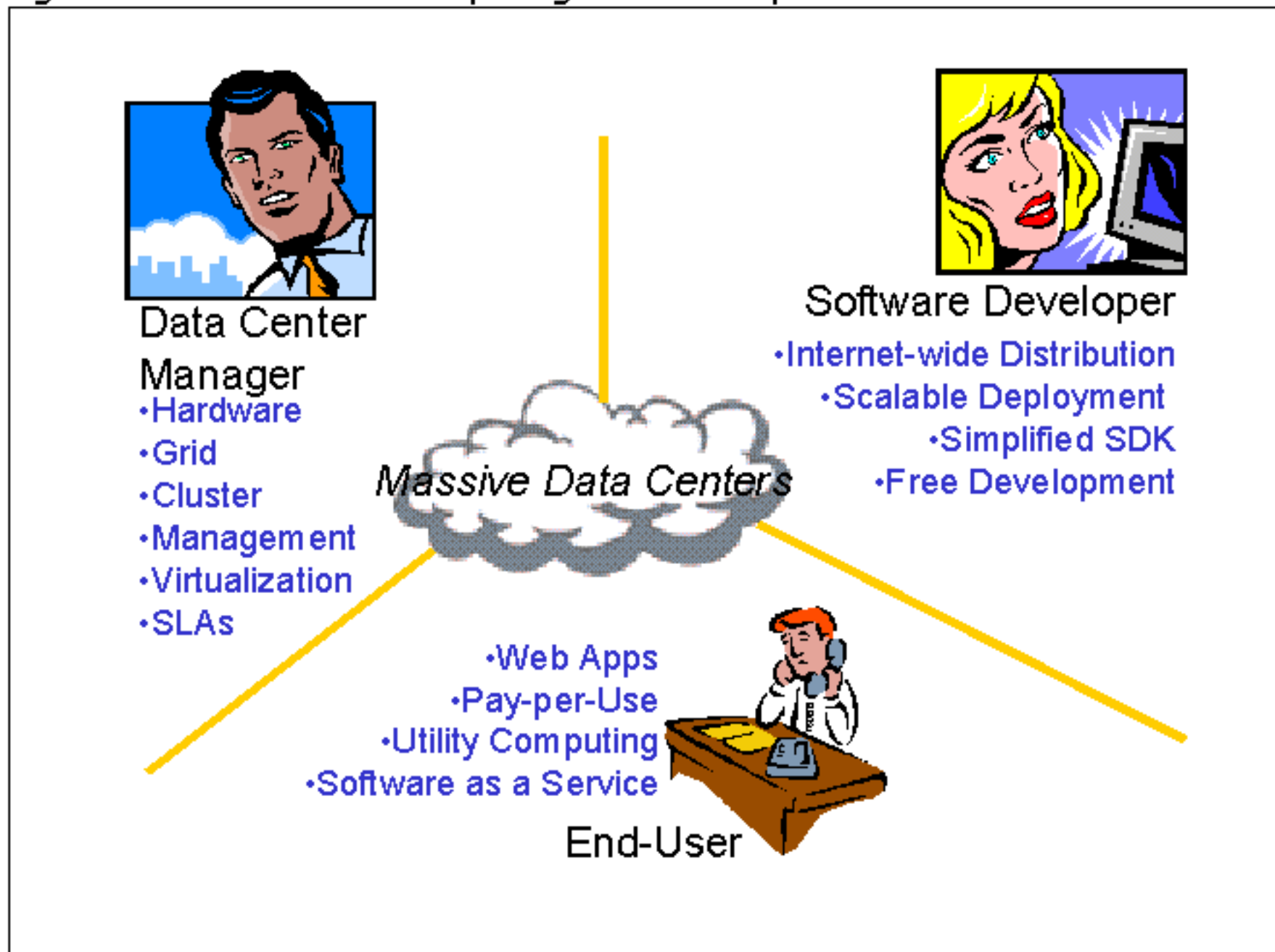
你對於我將是世界上唯一的，我對於你也將是世界上唯一的...

小王子 - 聖修伯里



我們與雲端運算的關係是啥？

Figure 1 – Three Cloud Computing User Participants



Source: Martin Hingley, June 2009

雲端運算的服務有啥呢？

Infrastructure Services

Storage

- Amazon S3
- Amazon EBS
- CTERA Portal
- Mosso Cloud Files
- Nirvanix

Compute

- Amazon EC2
- Serve Path GoGrid
- Elastra
- Mosso Cloud Servers
- Joyent Accelerators
- AppNexus
- Flexiscale
- ElasticHosts
- Hosting.com CloudNine
- Terremark
- GridLayer
- ITRICITY
- LayeredTech

Services Management

- RightScale
- enStratus
- Scalr
- CohesiveFT
- Kaavo
- CloudStatus
- Ylastic
- Dynect
- CloudFoundry
- NewRelic
- Cloud42

Cloud Software

Data

- 10Gen MongoDB
- Oracle Coherence
- Gemstone Gemfire
- Apache CouchDb
- Apache HBase
- Hypertable
- TerraCotta
- Tokyo Cabinet
- Cassandra
- memcached

Appliances

- PingIdentity
- Symplified
- rPath
- Vordel

Compute

- Globus Toolkit
- Xeround
- Beowulf
- Sun Grid Engine
- Hadoop
- OpenCloud
- Gigaspace
- DataSynapse
- Xeround

File Storage

- EMC Atmos
- ParaScale
- Zmanda
- CTERA

Cloud Management

- 3Tera App Logic
- OpenNebula
- Open.ControlTier
- Enomaly Enomalism
- Altor Networks
- VMware vSphere
- OnPathTech
- CohesiveFT VPN Cubed
- Hyperic
- Eucalyptus
- Reductive Lbs Puppet
- OpenQRM
- Appistry

CLOUD TAXONOMY

Platform Services

General Purpose

- Force.com
- Etelos
- LongJump
- AppJet
- Rollbase
- Bungee Labs Connect
- Google App Engine
- Engine Yard
- Caspio
- Qrimp
- MS Azure Services Platform
- Mosso Cloud Sites

Business Intelligence

- Aster DB
- Quantivo
- Cloud9 Analytics
- Blink Logic
- K2 Analytics
- LogiXML
- Oco
- Panorama
- PivotLink
- Sterna
- ColdLight Neuron
- Infobright
- Vertica

Integration

- Amazon SQS
- MuleSource Mule OnDemand
- Boomi
- SnapLogic
- OpSource Connect
- Cast Iron
- Microsoft BizTalk Services
- gnip
- SnapLogic SaaS Solution Packs
- Appian Anywhere
- HubSpan
- Informatica On-Demand

Development & Testing

- Keynote Systems
- Mercury
- SOASTA
- SkyTap
- Aptana
- LoadStorm
- Collabnet
- Dynamsoft

Database

- Google BigTable
- Amazon SimpleDB
- FathomDB
- Microsoft SDS

Software Services

Billing

- Aria Systems
- eVapt
- OpSource
- Redi2
- Zuora

Financials

- Concur
- Xero
- Workday
- Beam4d

Legal

- DirectLaw
- Advologix
- Fios
- Sertifi

Sales

- Xactly
- LucidEra
- StreetSmarts
- Success Metrics

Desktop Productivity

- Zoho
- IBM Lotus Live
- Google Apps
- Desktoptwo
- Parallels
- ClusterSeven

Human Resources

- Taleo
- Workday
- iCIMS

Content Management

- Clickability
- SpringCM
- CrownPoint

Backup & Recovery

- JungleDisk
- Mozy
- Zmanda Cloud Backup
- OpenRSM
- Syncplicity

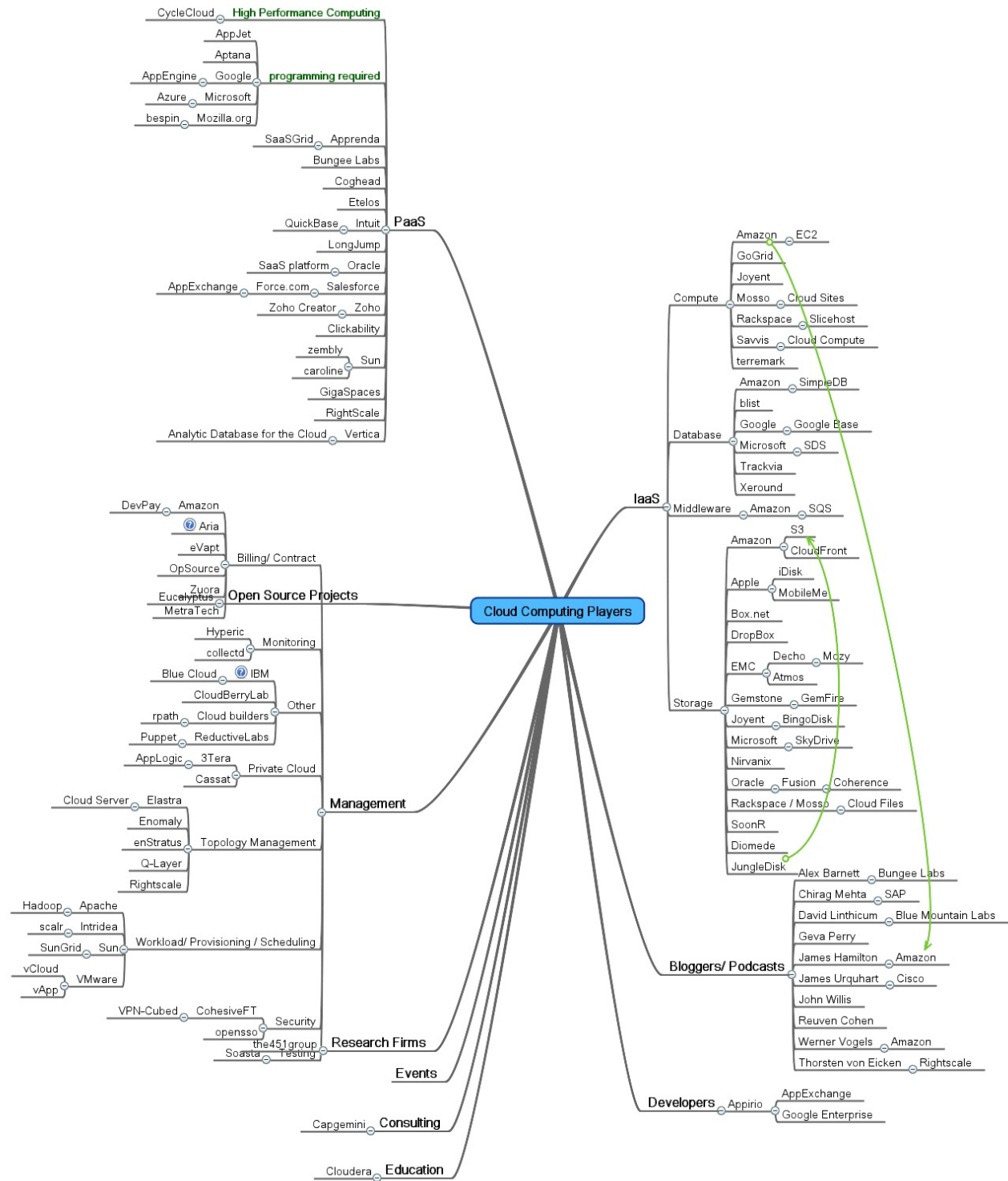
CRM

- NetSuite
- Parature
- Responsys
- Rightnow
- Salesforce.com
- LiveOps
- MSDynamics
- Oracle On Demand

Document Management

- NetDocuments
- Questys
- DocLanding
- Aconex
- Xyθος
- Knowledge TreeLive
- SpringCM








source: http://www.cloudiquity.com/wp-content/uploads/2009/02/cloud_computing_players.jpg

Let's see some Examples!

- [Amazon Web Services](#)
- [Slaesforce.com](#)
- [iCloud](#)
- [Google Documents](#)
- [Google Mail](#)
- [Google App Engine](#)
- [Windows Azure](#)

Comparing Google、Amazon、Microsoft

	類型	功能	費用	開發工具
Amazon EC2 	IaaS	使用者可使用AMI或自行上傳的Xen VM Image 來產生所需的 VM, 透過API 來控制和管理VM (可與S3、DB2、Simple Queue Services互動)	<u>付費</u> (根據VM 的能力來收費)	不拘(Linux 可支援的任何語言)
Google App Engine 	PaaS	可將撰寫好的 web 應用程式部署至 GAE (可與其他的Google服務整合, 亦可透過API來使用 Google DB/BigTable)	<u>免費/ 加值付費</u>	Python, Django, Java
Microsoft Azure 	PaaS	使用者可將開發好的應用程式佈署至微軟的data center, 透過 Windows 2008 VM 來執行應用程式 (Azure 提供 application storage 和 SQL services)	<u>付費</u> (2010/2/1 開始收費)	整合Visual Studio, 支援任何.NET語言

Top10 值得關注的 雲端服務供應商



Amazon - Amazon Web Services, 共有六項服務 (EC2, S3...)

AT&T - Synaptic Hosting

Enomaly - Elastic Computing Platform/ECP

Google - Google App Engine, Google Apps(Gmail, Gdocs...)

GoGrid - Cloud hosting

Microsoft - Azure

NetSuite - SAP ERP, CRM

Rackspace - Rackspace Cloud/Mosso, 提供三項服務(Cloud sites, Cloud Files, Cloud Servers)

RightScale - Right Cloud management platform

Salesforce.com - CRM

[資料來源: Jon Brodkin](#)

接下來 **Let's Case Study**



Case Study (Amazon EC2 & S3 Cloud services)

The New York Times & TimeSelect

- Scenario

- 1851 ~ 1922 articles
- 4 TB TIFF -> pdf

- Solution

- Amazon S3 (Upload TIFF)
- Amazon EC2 (100 EC2 instance)
- 1.5 TB pdf

- Cost

- USD \$3000
- 24 hours



[The New York Times — April 4, 1912](#)
[Link1](#) [Link2](#)

是否有
Open Source
的雲端運算建置工具呢？

IaaS 的 Open Source Tools



Eucalyptus
Systems

可建置類似Amazon EC2的架構

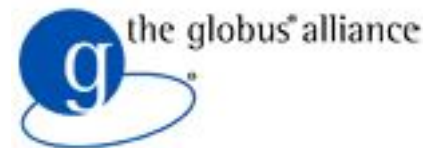
可動態分配和管理Public & Private Cloud

OpenNebula



將Data center轉換成雲端運算服務

Globus Nimbus 將叢集轉換成 IaaS Cloud



PaaS 的 Open Source Tools



可撰寫和處理巨量資料的平行運算平台

一套雲端的網路桌面環境



提供免費的 Hadoop Linux

可建置類似 Google App Engine 的平台



雲端運算與本課的關係

希望這門課可以獲得

1. 虛擬化的技術

Xen

2. 如何佈署你的私有雲

Eucalyptus, OpenNEbula

精彩回顧



- 了解了雲端運算的內涵
- 不同雲端計算的類型
- 各企業提供的雲端服務

討論時間



Appdenix

Cloud vs. Grid

Grid

- Massive Parallelization & Running large-scale jobs
- Concept of VO
- Focus on resource sharing & interoperability
- Standardization of protocols
- Make use of idle resources (utility-based)
- Scientific & engineering applications

Cloud

- Auto-provisioning of computing environment
- Intensive use of virtualization technology
- Enable the simplification of complex services: "Commodity Grid computing"
- Typically supported by data center

雲端運算vs.網格運算

	雲端運算	網格運算
主要推動者	資訊供應商 (如Google、Yahoo、IBM、Amazon等)	學術機構 (如歐洲粒子研究中心CERN、中研院、國家高速網路與計算中心)
標準化程度	無標準化，各家採用的技術架構也不同。	有標準化的協定和信任機制
開源幅度	部分開源，目前有開源Hadoop框架，但Google GFS和資料庫系統BigTable則未開源。	完全開源
網域限制	企業內部網域	可跨企業、跨管理網域
單一運算叢集可支援的硬體	相同標準規格的個人電腦 (如x86處理器、硬碟、4GB 記憶體、Linux等)	可混合異質性伺服器 (不同處理器、不同作業系統、不同編譯器版本等)
擅長處理的資料特性	單次運算資料量小 (可於單臺個人電腦上執行)，但需要重複大量處理次數的應用。	單次運算資料量大的應用。例如單筆數GB的衛星訊號分析。

資料來源：iThome整理，2008年6月

Powers Cloud Computing in Google

Distributed storage -

Google File System/GFS

Distributed semi-structure data system -

Big Table

Distributed data processing system -

MapReduce

Enterprise Computing vs Scientific Computing

Enterprise Computing	Scientific Computing
<p data-bbox="472 427 783 480">Application</p> <p data-bbox="197 491 559 762">Web/Mail server Database ERP CRM Backup</p>	<p data-bbox="1353 427 1664 480">Application</p> <p data-bbox="1076 491 1308 646">HPC CAD/CAE Simulation</p>
<p data-bbox="355 805 921 858">Usage Characteristic</p> <p data-bbox="197 869 517 1024">Small Job 20% loading (Virtualization)</p>	<p data-bbox="1227 805 1793 858">Usage Characteristic</p> <p data-bbox="1076 869 1553 976">Big Job High Loading (phase)</p>