淺談雲端運算的新趨勢

The Trend of Cloud Computing

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

Powered by DRBL

What is Cloud Computing?
何謂雲端運算？請用一句話說明！

Anytime 隨時

Anywhere 隨地

With Any Devices 使用任何裝置

Accessing Services 存取各種服務

More definition?
其他定義請參考：
NIST Notional Definition of Cloud Computing

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

Evolution of Cloud Services

雲端服務只是軟體演化史的必然趨勢

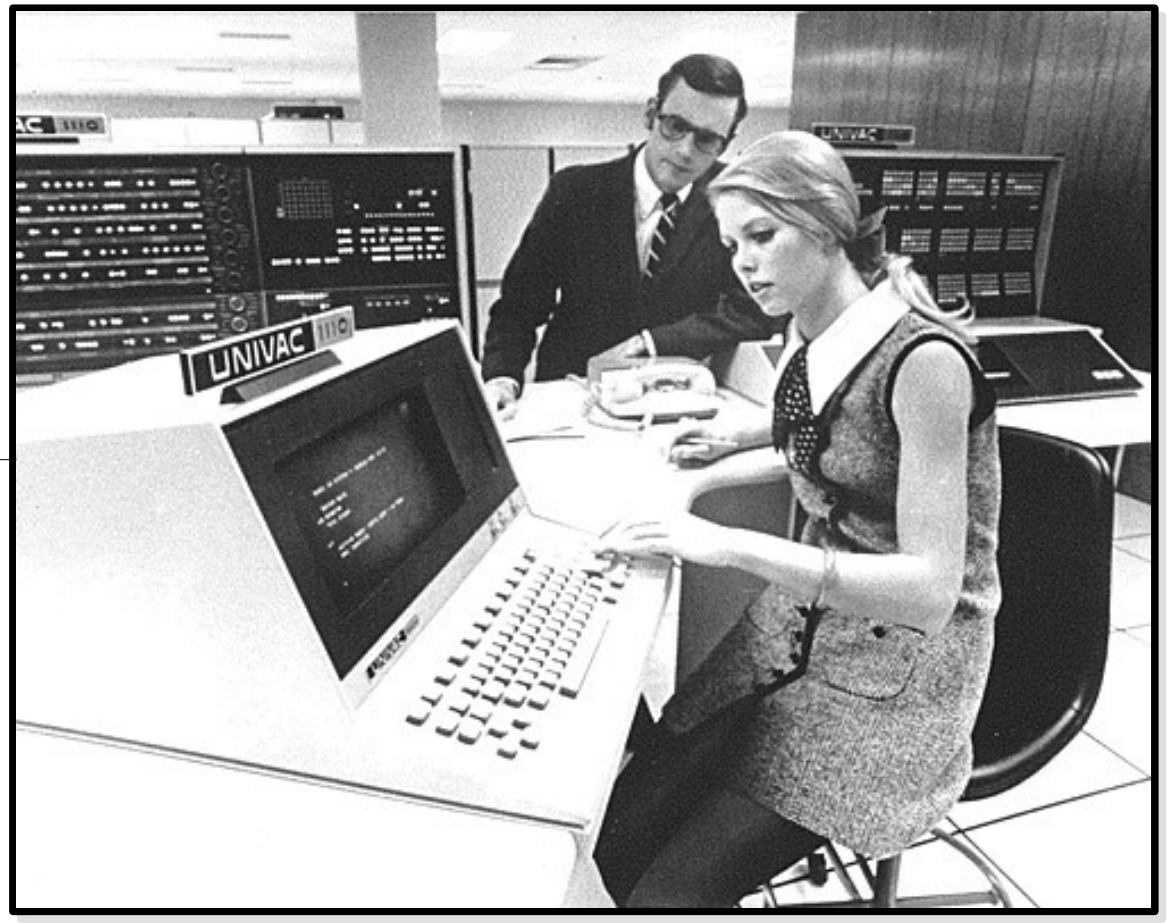
數位化 → e-Taiwan 成果系列報導 → 行動台灣 應用推動計畫 → 發展優質網路社會 (U-Taiwan) 計畫

實體	單機版 個人使用	網路版 多人共享	行動版 隨時存取
信箱	E-Mail	Web Mail	Mobile Mail
電視	電視盒	Web TV	Mobile TV
打字機	Office	Google Docs	M-Office
電話	數位電話	Skype	Flash Wengo
佈告欄	電子佈告欄	部落格	微網誌

When did the Cloud come ?!

這朵雲幾時飄過來的 ?!

Brief History of Computing (1/5)



Source: http://pinedakrch.files.wordpress.com/2007/07/

1960 PDP-1

.

.

.

1965 PDP-7

.

.

.

1969 1st Unix

Mainframe
Super
Computer

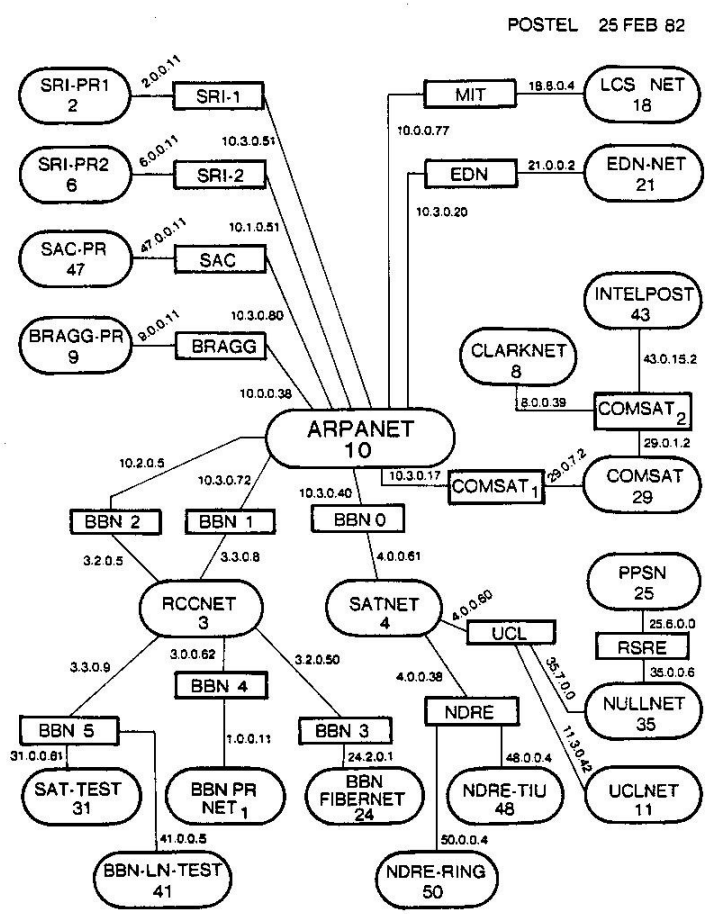
1977 Apple II

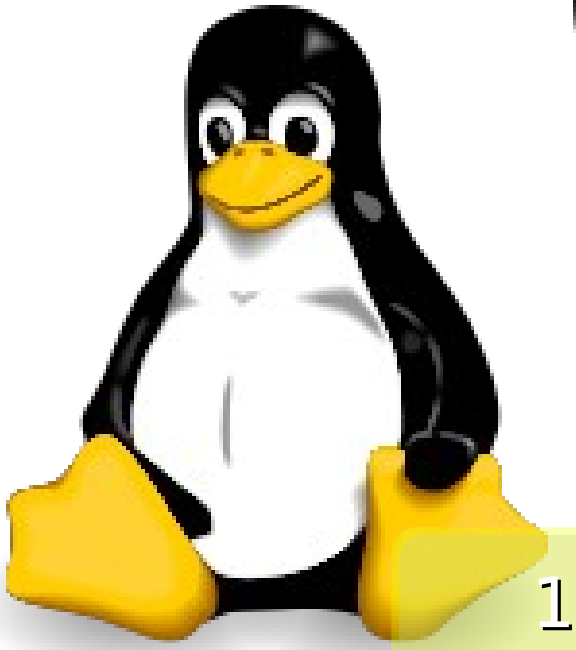
1981 IBM 1st PC 5150

Back to Year 1970s ...

1982 TCP/IP

1983 GNU





1991 Linux

Back to Year 1980s ...

Brief History of Computing (2/5)

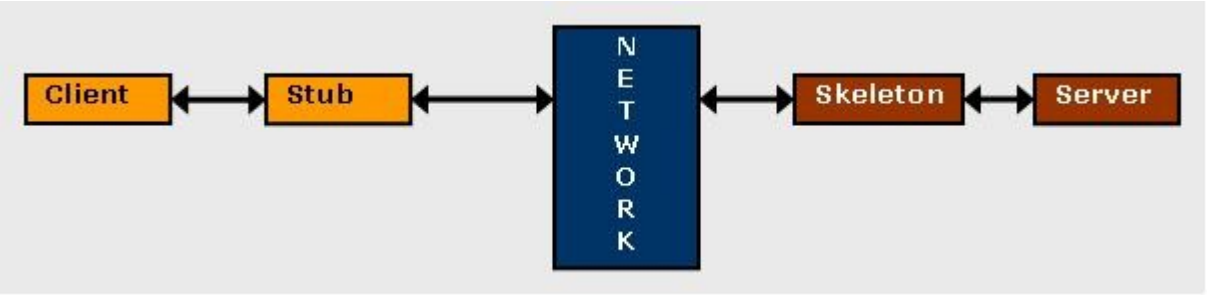
Source: http://www.nchc.org.tw

Mainframe
Super
Computer

PC / Linux
Cluster
Parallel

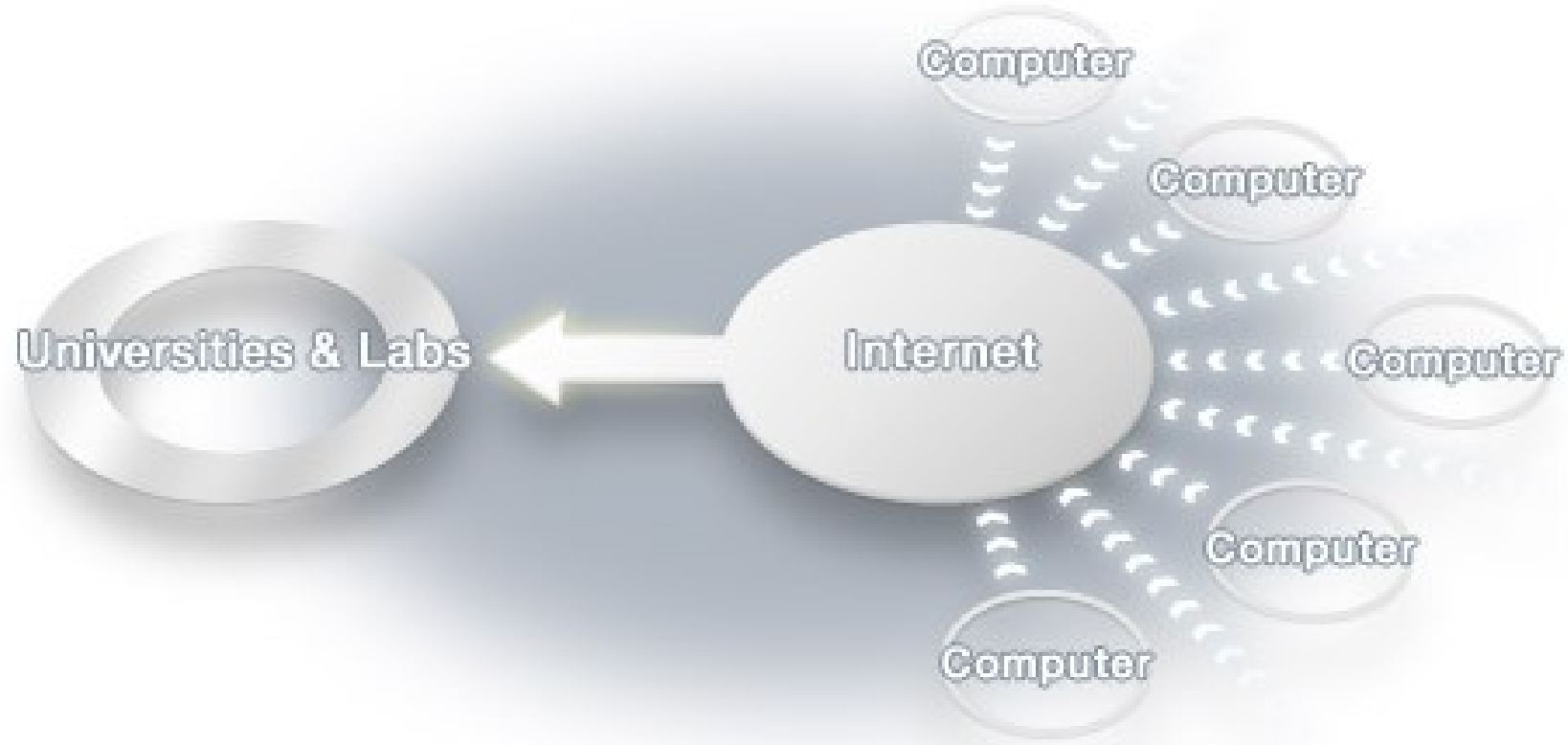
1990 World Wide Web
by CERN
...
...
1993 Web Browser
Mosaic by NCSA

1991 CORBA
...
Java RMI
Microsoft DCOM
...
Distributed Objects



Back to Year 1990s ...

Brief History of Computing (3/5)



Source: http://www.scei.co.jp/folding/en/dc.html

Mainframe
Super
Computer

PC / Linux
Cluster
Parallel

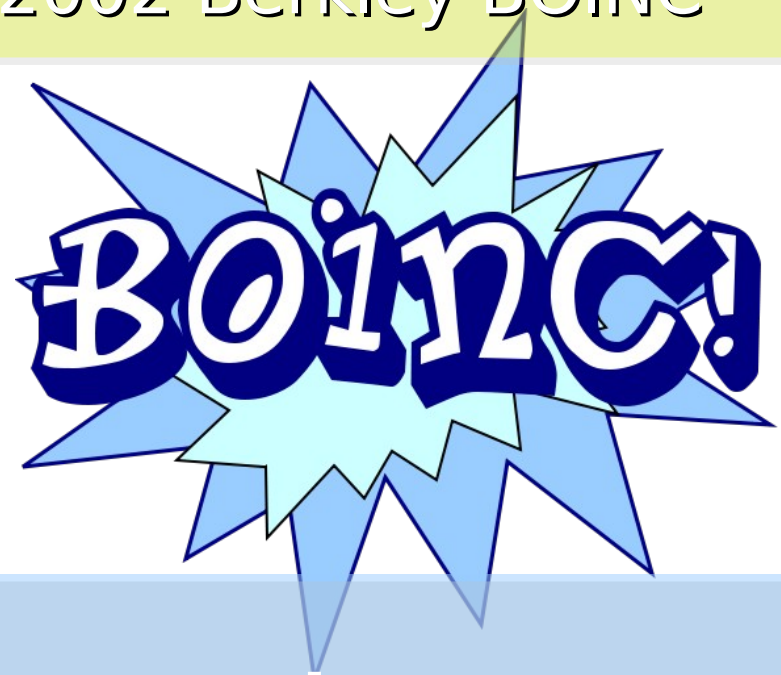
Internet
Distributed
Computing

1997 Volunteer Computing
1999 SETI@HOME

2003 Globus Toolkit 2

2002 Berkley BOINC

2004 EGEE gLite



Back to Year 2000s ...

Brief History of Computing (4/5)

Source: http://gridcafe.web.cern.ch/gridcafe/whatisgrid/whatis.html

Mainframe
Super Computer

PC / Linux
Cluster Parallel

Internet
Distributed Computing

Virtual Org.
Grid Computing

2001 Autonomic Computing
IBM

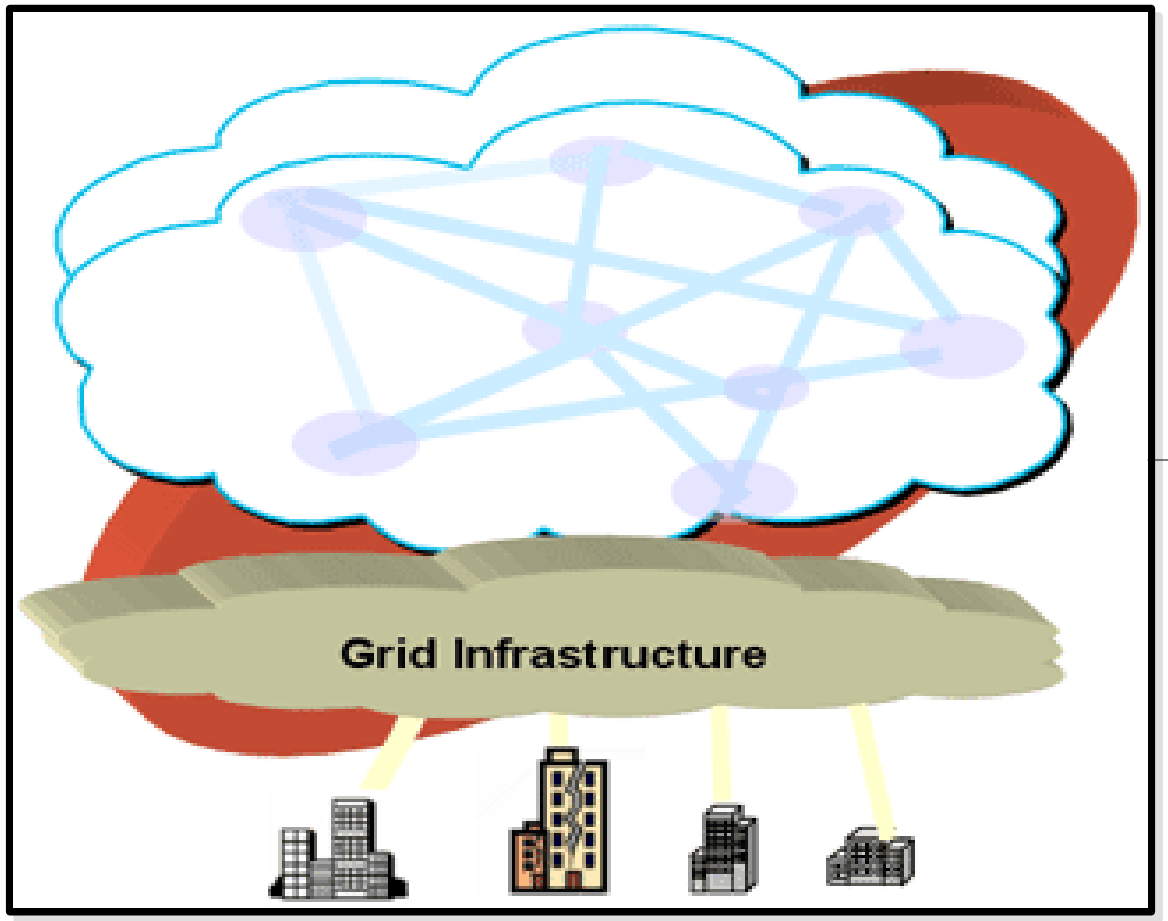
2006 Apache Hadoop

2005 Utility Computing
Amazon EC2 / S3

2007 Cloud Computing
Google + IBM

Back to Year 2007 ...

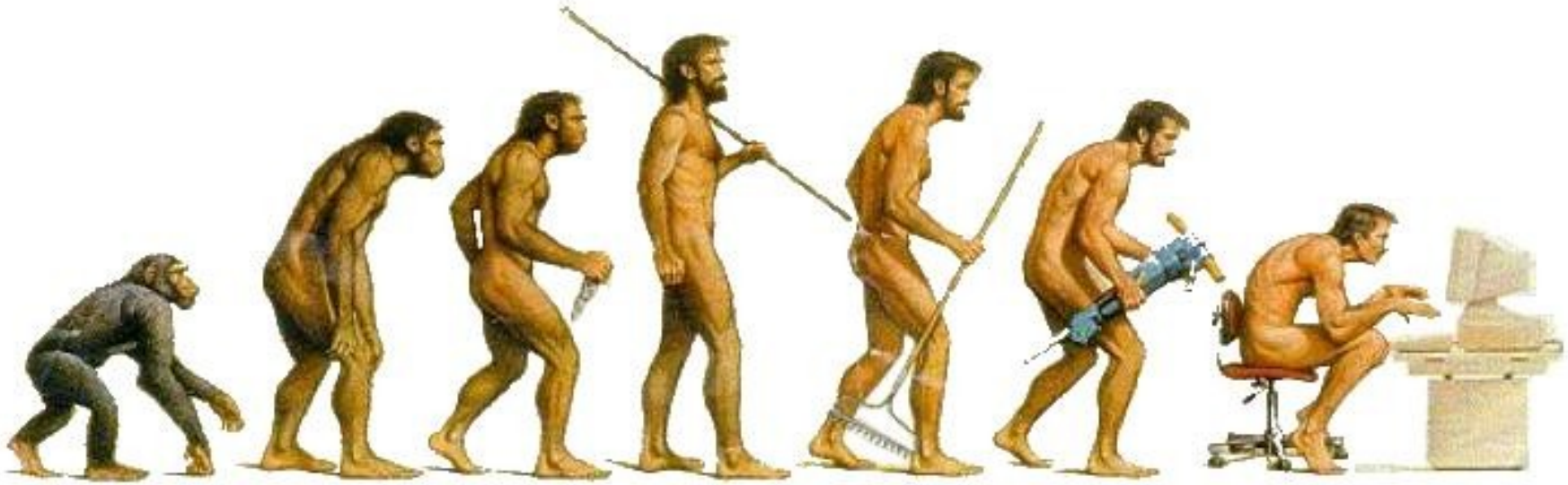
Brief History of Computing (5/5)



Source: http://mmdays.com/2008/02/14/cloud-computing/

nframe per nputer	**PC / Linux** Cluster Parallel	**Internet** Distributed Computing	**Virtual Org.** Grid Computing	**Data Explode** Cloud Computing

Evolution



(or is it?)

What can we learn from the past ?!

在這漫長的演化中，我們到底學到些什麼 ?!

Source: http://cyberpingui.free.fr/humour/evolution-white.jpg

Lesson #1: One cluster can't fit all !
教訓一：叢集的單一設定無法滿足所有需求！

Answer #1: Virtual Cluster 新服務：虛擬化叢集

Lesson #2: Grid for Heterogeneous Enterprise !
教訓二：格網運算該用在異業結盟的資源共享！

Answer #2: Peak Usage Time 尖峰用量發生時間點

Lesson #3: Extra cost to move data to Grid !
教訓三：資料搬運的網路與時間成本！

Answer #3: Total Cost of Ownership 總擁有成本

This is why Cloud Computing matters ?!
這就是爲什麼雲端運算變得熱門?!

Trend #1: Data are moving to the Cloud
趨勢一：資料開始回歸集中管理

Access data anywhere anytime 爲了隨時存取

Reduce the risk of data lost 降低資料遺失風險

Reduce data transfer cost 減少資料傳輸成本

Enhance team collaboration 促進團隊協同合作

How to store huge data ?!

如何儲存大量資料呢 ?!

Trend #2: Web become default Platform!
趨勢二：網頁變成預設開發平台

Open Standard 網頁是開放標準

Open Implementation 實作不受壟斷

Cross Platform 瀏覽器成爲跨平台載具

Web Application 網頁程式設計成爲顯學

Browser difference become entry barrier ?!

瀏覽器的差異造成新的技術門檻 ?!

Trend #3: HPC become a new industry
趨勢三：高速計算已悄悄變成新興產業

Parallel Computing 平行運算的技能

Distributed Computing 分散運算的技能

Multi-Core Programming 多核心程式設計

Processing Big Data 處理大資料的技能

Education and Training are needed !!

爲了讓這些技能與產業接軌，亟需教育訓練！！

該使用別人打造的雲端，還是自己打造專屬雲端呢？

Types of Cloud Computing
雲端運算的三種型態

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

amazon web services™ Microsoft Google

Public Cloud
公用雲端

Target Market
is S.M.B.
主要客户爲
中小企業

Hybrid
Cloud

以大型企業
爲主要客户
Enterprise is
key market

IBM

私有雲端
Private Cloud

Types of Cloud Service Provider
雲端服務的市場區隔

SaaS
Software as a Service
軟體即服務

PaaS
Platform as a Service
平台即服務

IaaS
Infrastructure as a Service
架構即服務

Everything as a Service 啥米鬼都是一種服務

AaaS	Architecture as a Service
BaaS	Business as a Service
CaaS	Computing as a Service
DaaS	Data as a Service
DBaaS	Database as a Service
EaaS	Ethernet as a Service
FaaS	Frameworks as a Service
GaaS	Globalization or Governance as a Service
HaaS	Hardware as a Service
IMaaS	Information as a Service
IaaS	**Infrastructure or Integration as a Service**
IDaaS	Identity as a Service
LaaS	Lending as a Service
MaaS	Mashups as a Service
OaaS	Organization or Operations as a Service
SaaS	**Software or Storage as a Service**
PaaS	**Platform as a Service**
TaaS	Technology or Testing as a Service
VaaS	Voice as a Service

Customer-Oriented
客户導向

引用自：
https://www.ibm.com/developerworks/mydeveloperworks/blogs/sbose/entry/gathering_clouds_of_xaas

Public Cloud #1: Amazon 亞馬遜網路書店

- Amazon Web Service (AWS)
- 虛擬伺服器：Amazon EC2
  - Small (Default) $0.10 per hour $0.125 per hour
  - All Data Transfer $0.10 per GB
- 儲存服務：Amazon S3
  - $0.150 per GB – first 50 TB / month of storage used
  - $0.100 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.cisanet.org.tw/post/Cloud-Computing.aspx

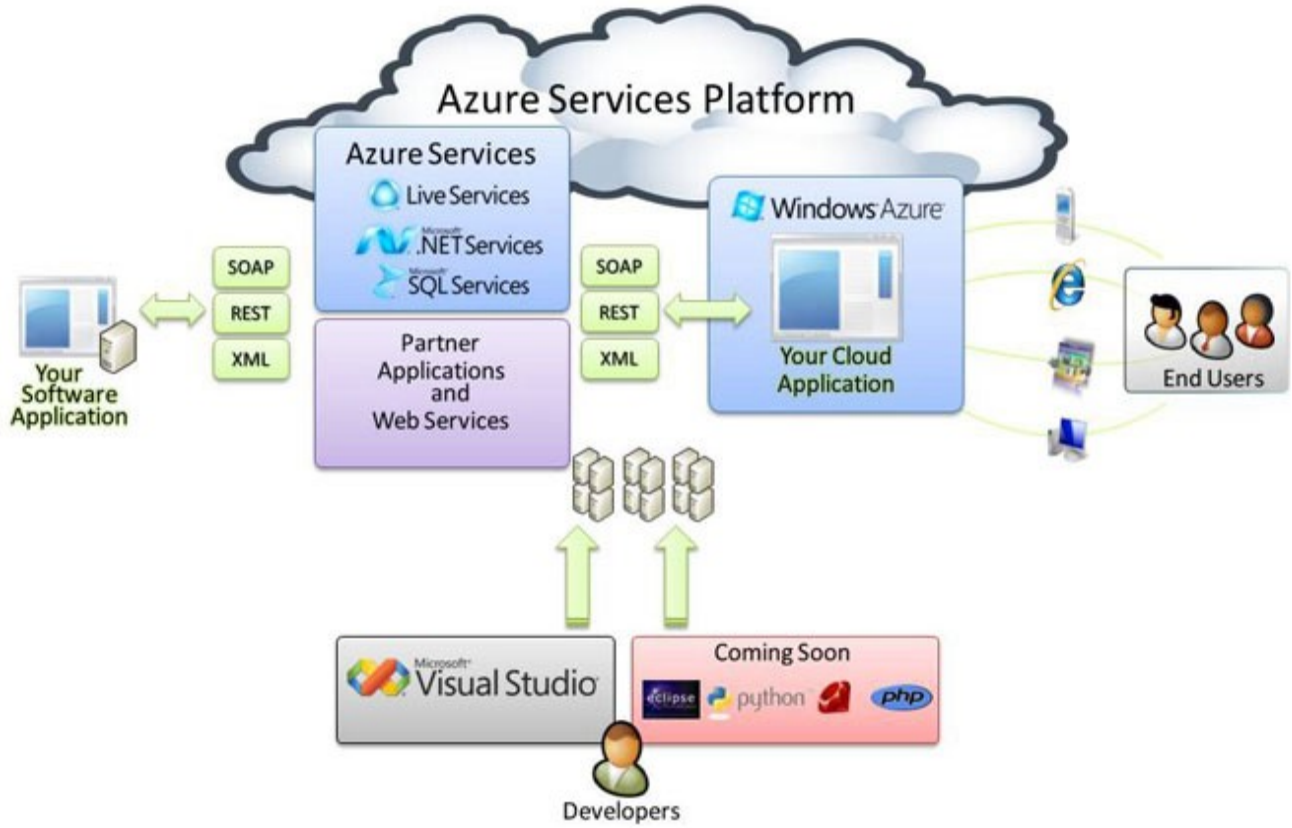
Public Cloud #2: Google 谷歌

- Google App Engine (GAE)
- 讓開發者可自行建立網路應用程式於Google平台中。
  - 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: Python、Java

參考來源：http://code.google.com/intl/zh-TW/appengine/

Public Cloud #3: Microsoft 微軟

- Microsoft Azure 是一套雲端服務作業系統。
- 作爲 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



參考來源：http://tech.cipper.com/index.php/archives/332

Reference Cloud Architecture
雲端運算的參考架構

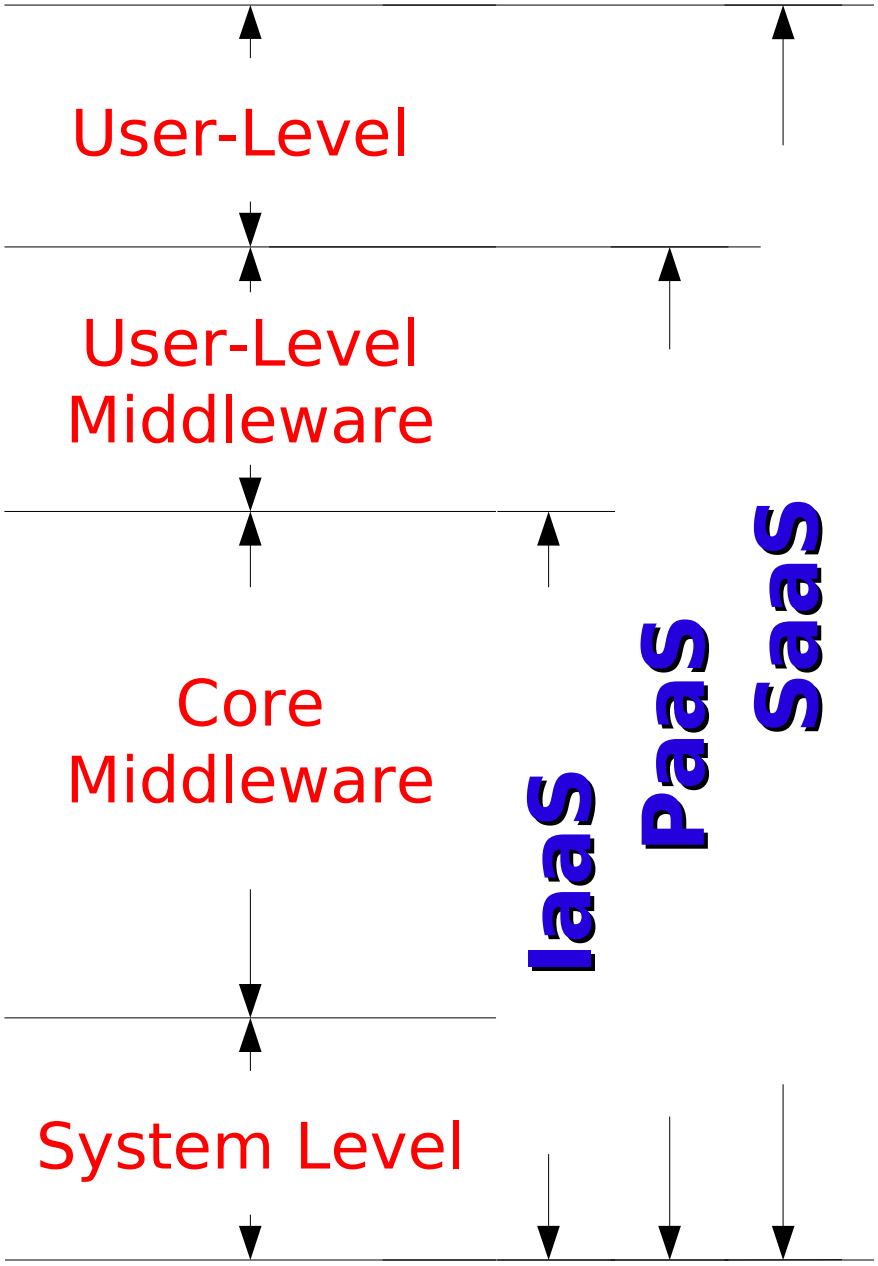
應用
Social Computing, Enterprise, ISV,...

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

控制
Qos Neqotiation, Ddmission Control, Pricing, SLA Management, Metering...

虛擬化
VM, VM management and Deployment

硬體設施
Infrastructure: Computer, Storage, Network



Open Source for Private Cloud
建構私有雲端運算架構的自由軟體

應用
Social Computing, Enterprise, ISV,...

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

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

Hadoop (MapReduce), Sector/Sphere, AppScale

控制
Qos Neqotiation, Ddmission Control, Pricing, SLA Management, Metering...

OpenNebula, Enomaly, Eucalyptus , OpenQRM, ...

虛擬化
VM, VM management and Deployment

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

Open Cloud #1: Eucalyptus

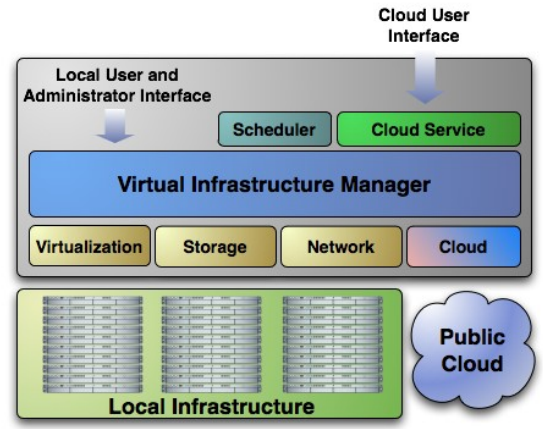
- http://open.eucalyptus.com/
- 原是加州大學聖塔芭芭拉分校(UCSB)的研究專案
- 目前已轉由Eucalyptus System這間公司負責維護
- 創立目的是讓使用者可以**打造自己的EC2**
- 特色是相容於 Amazon EC2 既有的用户端介面
- 優勢是Ubuntu 9.04 已經收錄 Eucalyptus 的套件
- Ubuntu Enterprise Cloud powered by Eucalyptus in 9.(
- 目前有提供 Eucalyptus 的官方測試平台供註冊帳號
- 缺點：目前仍有部分操作需透過指令模式

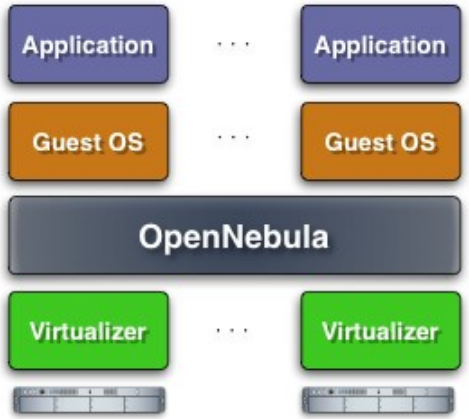
關於 Eucalyptus 的更多資訊，請參考
http://trac.nchc.org.tw/grid/wiki/Eucalyptus

Open Cloud #2: OpenNebula

- http://www.opennebula.org
- 由歐洲研究學會(European Union FP7)贊助
- 將實體叢集轉換成具管理彈性的虛擬基礎設備
- 可管理虛擬叢集的狀態、排程、遷徙(migration)
- 優勢是Ubuntu 9.04 已經收錄 OpenNebula 的套件
- 缺點：需下指令來進行虛擬機器的遷徙(migration)。

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





Open Cloud #3: Hadoop

- http://hadoop.apache.org
- Hadoop 是 Apache Top Level 開發專案
- 目前主要由 Yahoo! 資助、開發與運用
- 創始者是Doug Cutting，參考Google Filesystem，以Java開發，提供HDFS與MapReduce API。
- 2006年使用在Yahoo內部服務中
- 已佈署於上千個節點。
- 處理Petabyte等級資料量。
- Facebook、Last.fm、Joost … 等
- 著名網路服務均有採用Hadoop。

Open Cloud #4: Sector / Sphere

- http://sector.sourceforge.net/
- 由美國資料探勘中心(National Center for Data Mining)研發的自由軟體專案。
- 採用C/C++語言撰寫，因此效能較 Hadoop 更好。
- 提供「類似」Google File System與MapReduce的機制
- 基於UDT高效率網路協定來加速資料傳輸效率
- Open Cloud Consortium的Open Cloud Testbed，有提供測試環境，並開發了MalStone效能評比軟體。

Sector-Sphere

National Center for Data Mining
University of Illinois at Chicago
UIC

open data
Open Data Group
http://www.opendatagroup.com/

NCHC Cloud Computing Research Group
團隊小檔案：國網中心雲端運算研究小組

- 主要研究雲端運算的基礎架構組成元件
- http://trac.nchc.org.tw/cloud,
  http://trac.nchc.org.tw/grid
- 團隊成員：6名
  - —王耀聰— drbl-xen / drbl-hadoop (~6 Years) 架構
  - —陳威宇— Hadoop / NutchEz / ICAS (~3 Years) 應用
  - —郭文傑— Xen / OpenNebula / Eucalyptus (~3 Years) 元件
  - —涂哲源— Xen GPU / OpenMP / VirtualGL (~3 Years) 元件
  - —鄭宗碩— Google App Engine (~2 Years) 新技術
  - —鄧偉華— AMQP / OpenID (~2 Years) 新技術
- 定位：
  - —研發快速佈建軟體，提供實驗平台服務，開辦訓練課程育才
- 獨特性：
  - —基於企鵝龍(DRBL)，可快速佈署雲端運算的叢集環境

更多相關的開放教材－生物叢集、GAE...

- 陽明生資所 97 年度暑期學分班 格網及平行運算（實驗課程） http://trac.nchc.org.tw/course/
- 陽明生資所 98 年度暑期學分班 格網及平行運算（實驗課程） http://bio.classcloud.org
- 雲端運算基礎課程（一）Hadoop 簡介、安裝與範例實作 http://www.classcloud.org/media/
- 「Ruby on Rails 初學」電子書 by 鄭立竺 http://nchcrails.blogspot.com
- Google App Engine 電子書 by 鄭宗碩 http://nchc-gae.blogspot.com/
- More to come .....

What we learn today ?

WHAT	隨時隨地用任何裝置存取各種服務!! Accessing services with any device anytime anywhere!!
WHO	亞馬遜、谷歌、微軟等! 什麼都可以是服務 ~ Amazon, Google, Microsoft and more! Everything as a Service!
WHEN	雲端運算是2007年繼格網運算之後的新趨勢!! Cloud Computing become new trend since year 2007 !!
WHY	資料集中、虛擬化、異業資源共享 Data-intensive, Virtualization, Heterogeneous
HOW	採用自由軟體也能打造私有雲端 Hadoop, Sectore/Sphere, Eucalyptus, and more ....

Questions?

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

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

Powered by DRBL