

Course Information 課程資訊



- 講師介紹：
 - 國網中心 王耀聰 副研究員 / 交大電控碩士
 - jazz@nchc.org.tw
- 所有投影片、參考資料與操作步驟均在網路上
 - 由於雲端資訊變動太快，愛護地球，請減少不必要之講義列印。
- 礙於缺乏實機操作環境，故以影片展示與單機操作為主
 - 若有興趣實機操作，請參考國網中心雲端運算課程錄影
 - <http://trac.nchc.org.tw/cloud>
 - <http://www.classcloud.org/media>
 - <http://www.screentoaster.com/user?username=jazzwang>
- 若需要實驗環境，可至國網中心雲端運算實驗叢集申請帳號
 - <http://hadoop.nchc.org.tw>
- Hadoop 相關問題討論：
 - <http://forum.hadoop.tw>



雲端運算的關鍵技術概述

Part 1 : Overview Core Technologies of Cloud Computing

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



Powered by DRBL



什麼是雲端運算啊？可以個簡單的定義嗎？

What is Cloud Computing ?

雲端運算怎麼聽起來要買一些新硬體、新軟體啊？

Is it about buying NEW Hardware and Software?



雲端運算可能只是拿來振興經濟的幌子吧？

Is it a trap to another bubble economy ?



我聽你們在那裡講五四三.....

Cloud Computing is as simple as 5..4..3..2..1...



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

**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**

Community Cloud

社群雲端

Academia 學術為主



私有雲端

Private Cloud

3 Service Models of Cloud Computing

雲端運算的三種服務模式

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 R&D directions : Cloud or Device

兩大研究方向：你該選「雲」還是「端」？



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

Cloud Computing =~ Network Computing

雲端運算 =~ 網路運算

Key spirit of Cloud ~

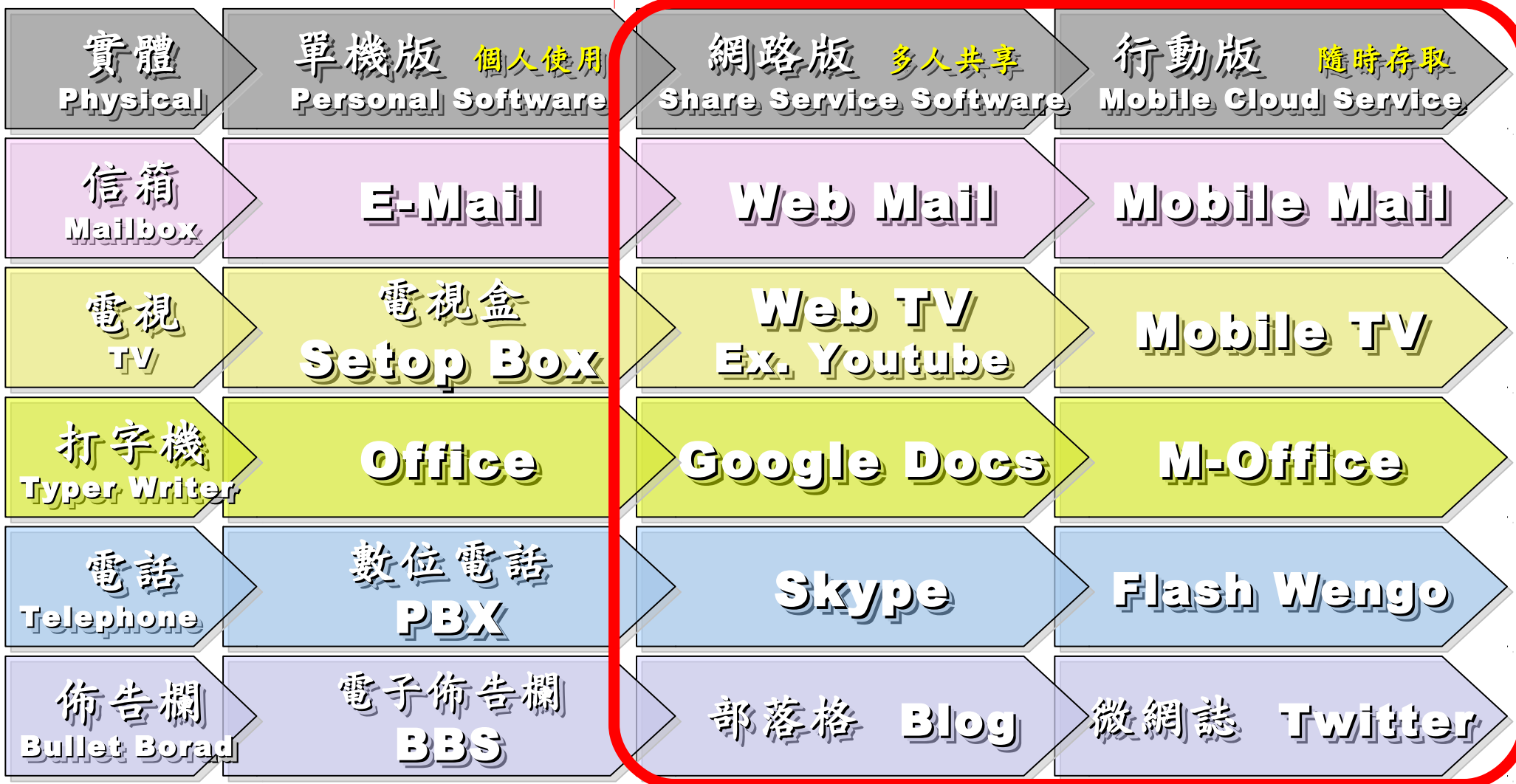
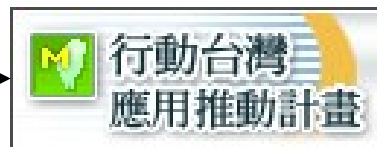
形成服務才是重點！！

Everything as a Service !!

Evolution of Cloud Services

雲端服務只是軟體演化史的必然趨勢

數位化



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



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

用任何連網裝置
都可以存取資料

租賃取代買斷
動態隨需付費

降低
經營成本

Cost Down

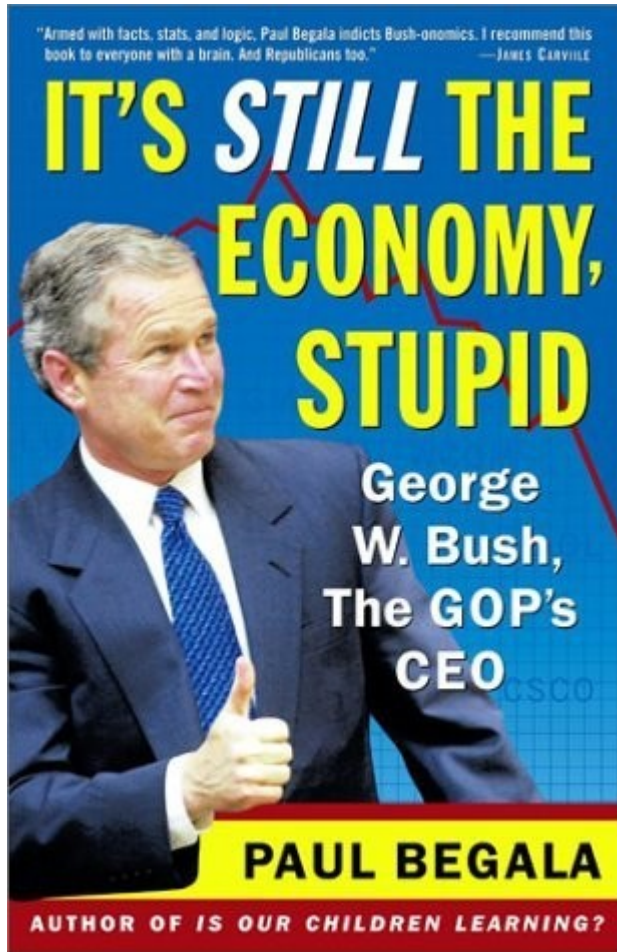
資料往雲擺
減少資料傳輸

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"**)

卻讓小布希被嘲笑是幼稚的總統。

- **2002** 年

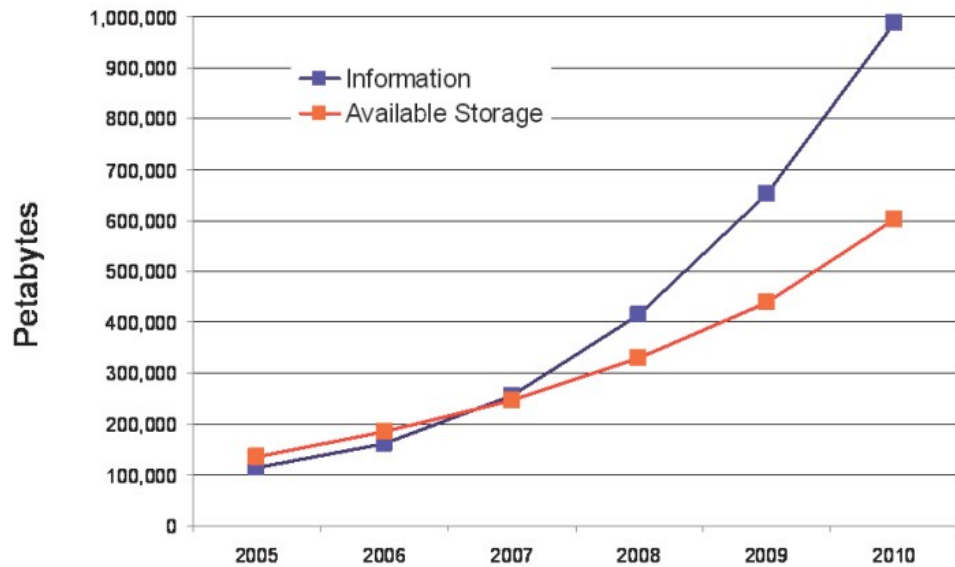
雲端時代，谷歌會說：「笨蛋！重點在資料」

(**"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？

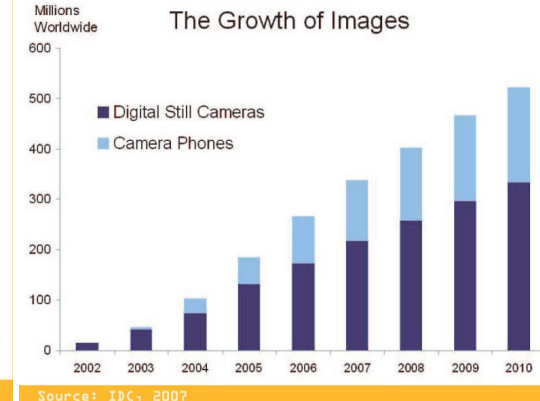
- **2007** 年

Information Versus Available Storage



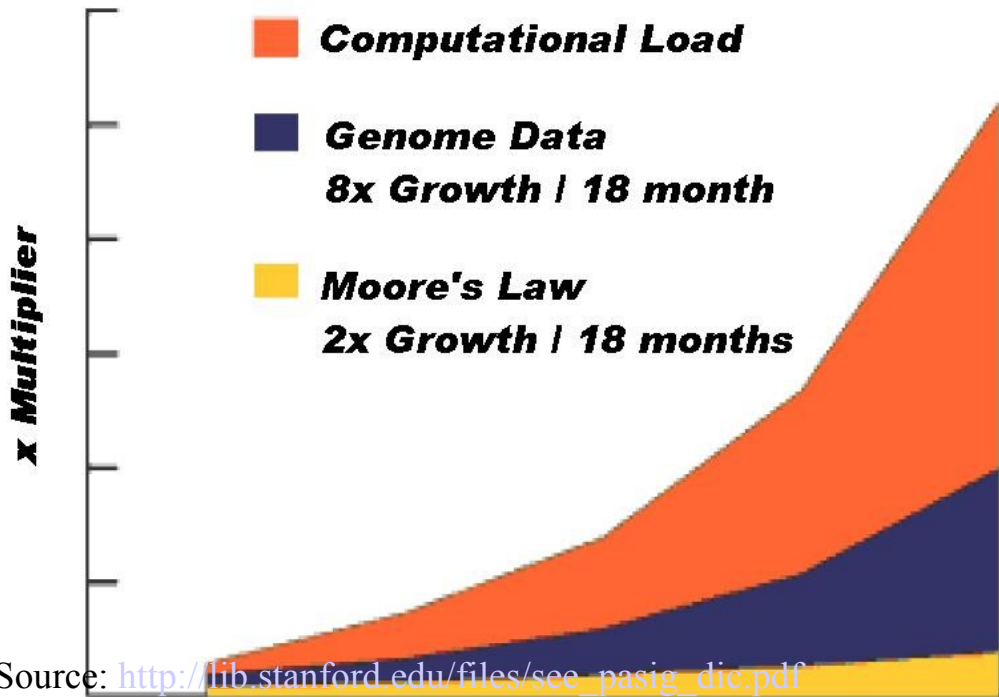
2007 Data Explore

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



Source: <http://www.emc.com/collateral/analyst-reports/expanding-digital-idc-white-paper.pdf>
 Source: IDC, 2007

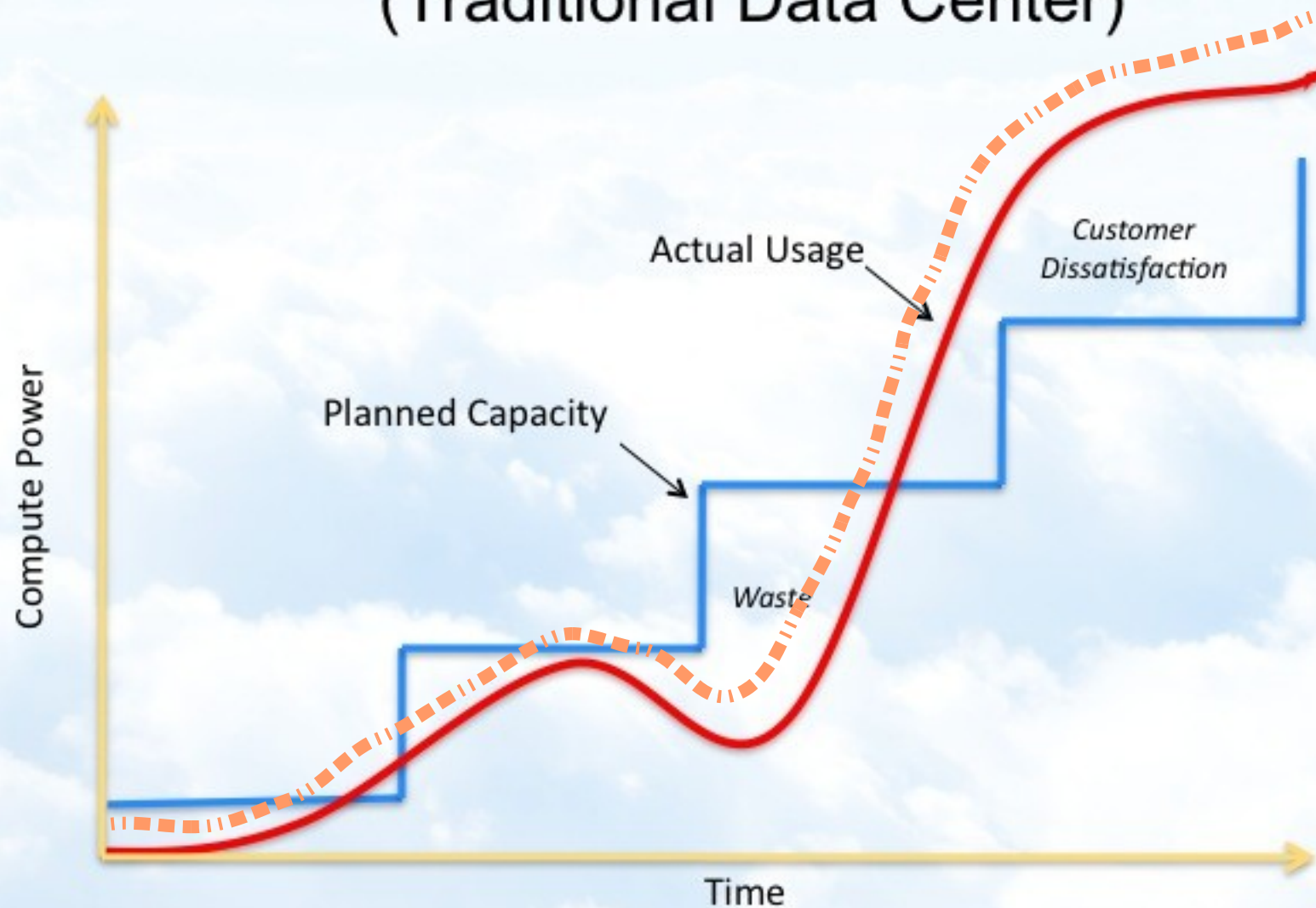
Source: IDC, 2007



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

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)			

Capacity vs. Usage (Traditional Data Center)



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

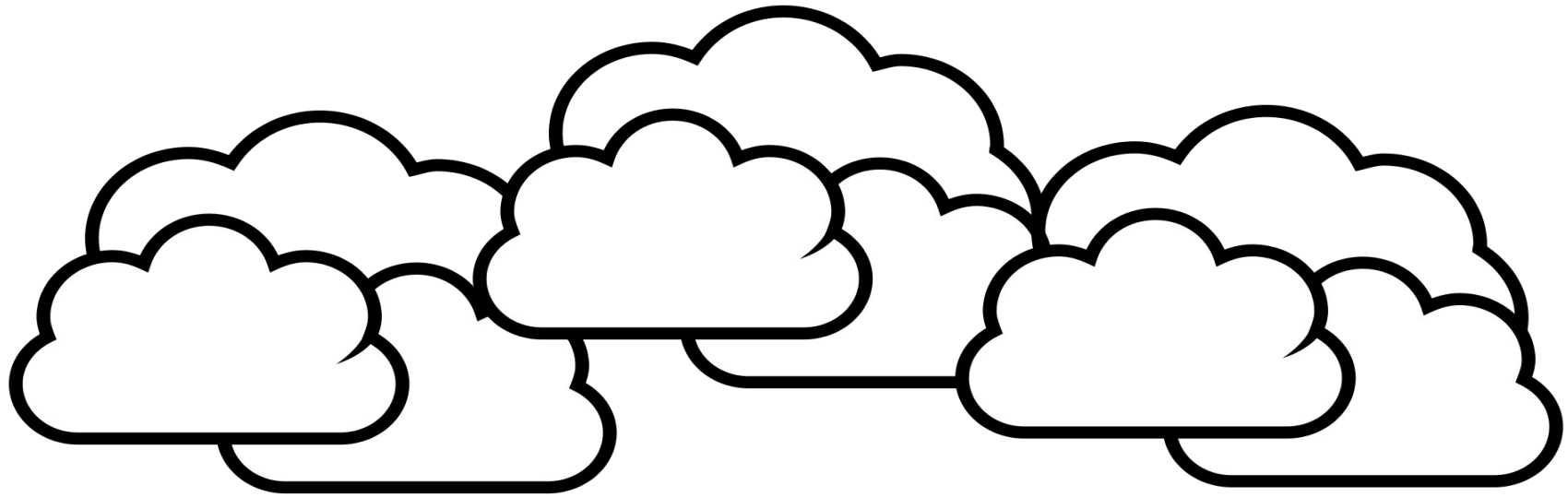
Answer #3: Total Cost of Ownership 總擁有成本

Cost Down is the Key Drive !!

降低營運成本才是企業導入雲端運算的關鍵考量！！

Data is the source of Wisdom !!

用雲掌握資料，加以分析，形成智能給端用



嵌入式的新思維：未來，**端**的智能來自於**雲**

Devices share the wisdom of Cloud



What are the trend of next 10 years ?

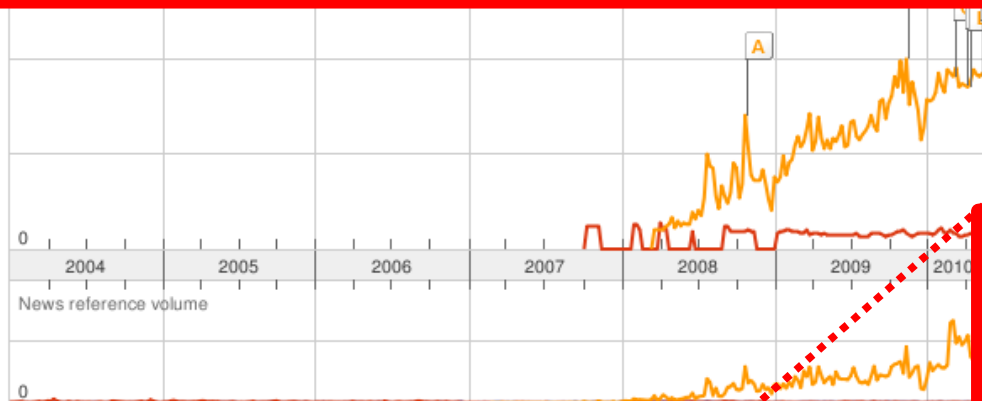
什麼是下個十年的熱門工作技能？

● distributed computin... ● grid computing ● cloud computing

[Sign in](#) to see and export additional Tren

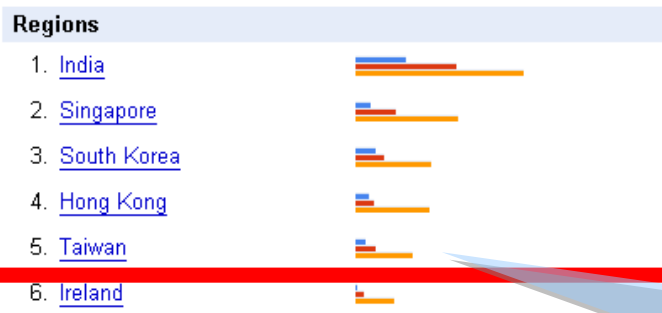
All regions All years

Search Volume index

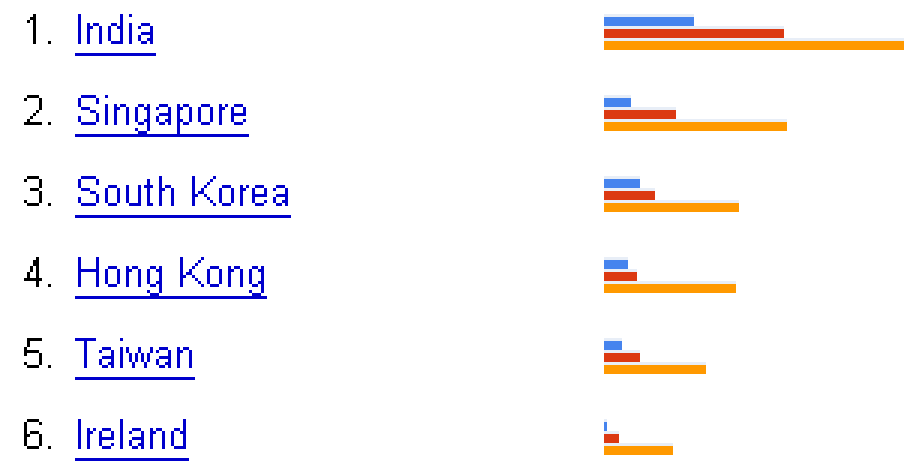


- [" cloud" computing space](#)
Canada.com - Oct 26 2008
- [Microsoft's cloud computing system is growing up](#)
Philadelphia Inquirer - Nov 17 2009
- [Google looks to be 'cloud-computing' rainmaker for other online business services](#)
Winnipeg Free Press - Mar 10 2010

Rank by cloud computing



Regions



似乎亞洲國家特別熱愛雲端?! Too Hot in Asia ?!

CIO 2010 : Virtualization, Cloud and Web 2.0

虛擬化技術是 2010 年 CIO 資訊長認為最重要的資訊投資

CIO strategic technologies reflect increased interest in “lighter-weight” solutions

CIO technologies

Ranking of technologies CIOs selected as one of their top 5 priorities in 2010

Ranking	2010		2009	2008	2007
Virtualization	1	↑	3	3	5
Cloud computing	2	↑	16	*	*
Web 2.0	3	↑	15	15	*
Networking, voice and data communications	4	↑	6	7	4
Business intelligence (BI)	5	↓	1	1	1
Mobile technologies	6	↑	12	12	11
Data/document management and storage	7	↑	10	9	9
Service-oriented applications and architecture	8	↑	9	10	7
Security technologies	9	↓	8	5	6
IT management	10		*	*	*
Enterprise applications	11	↓	2	2	2

* New question for that year

Source: *Gartner Executive Programs* : “ *Leading in Times of Transition: The 2010 CIO Agenda* ”

How hot is Cloud and Virtualization ?

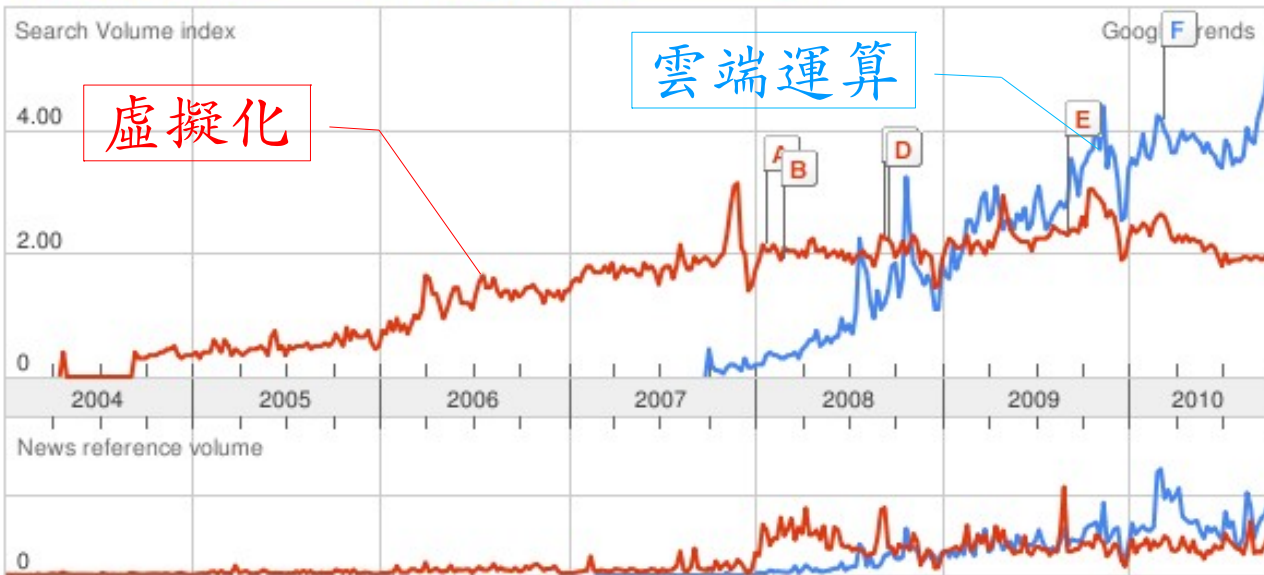
雲端運算與虛擬化有多熱門呢？

Searches [Websites](#)

All regions

Scale is based on the average worldwide traffic of [cloud computing](#) in all years. [Learn more](#)

cloud computing 1.00 virtualization 1.45



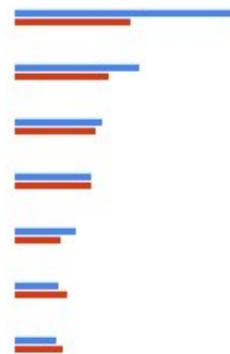
- A** [Microsoft joins 'virtualization' push](#)
Globe and Mail - Jan 22 2008
- B** [Microsoft enters virtualization with Windows](#)
Canada.com - Feb 27 2008
- C** [Microsoft Launches Virtualization Blitzkrieg](#)
CRN - Sep 8 2008
- D** [CA to Deliver Integrated Virtualization Manag Program](#)
WELT ONLINE - Sep 15 2008
- E** [Wyse and VMware Users and Channel Part](#)
Reuters - Aug 31 2009
- F** [Google looks to be 'cloud-computing' rainma](#)
Winnipeg Free Press - Mar 10 2010

[More news results »](#)

Rank by

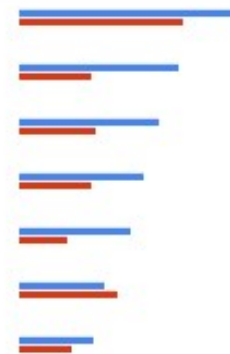
Regions

1. [India](#)
2. [Singapore](#)
3. [Hong Kong](#)
4. [South Korea](#)
5. [Taiwan](#)
6. [Malaysia](#)
7. [United States](#)



Cities

1. Bangalore, India
2. Mahape, India
3. Mumbai, India
4. Chennai, India
5. Delhi, India
6. San Jose, CA, USA
7. Singapore, Singapore



Language

1. Kore
2. Engli
3. Dutc
4. Germ
5. Chin
6. Portu
7. Thai

Are the trends telling the truth ?

你確定沒有被圖表晃點嗎？

Searches Websites

All regions

cloud computing does not have enough search volume for ranking

民眾比較想了解『端』

Scale is based on the average worldwide traffic of **iphone** in all years. [Learn more](#)

cloud computing 1.0 **iphone** 1.00 **ipad** 0.10 **android** 0.10 **chrome** 0.25



- A** [iPhone 3G Success!!!](#)
Dallas Morning News - Jul
- B** [iPod...iPhone...now, iPad](#)
Economic Times - Jan 27 2
- C** [Apple unveils the "iPad"](#)
AFP - Jan 28 2010
- D** [iPad Gag Apps Missing: N](#)
PC World - Apr 5 2010
- E** [Apple unveils iPhone 4 and OS4](#)
Myjoyonline.com - Jun 7 20
- F** ['iPhone 4 major leap on iPh](#)
Times of India - Jun 24 20

[More news results »](#)

找『雲』的只有服務提供者

雲端運算

Regions

Cities

Languages

善用雲端架構 打造企業人才庫

對於雲端的運用，多半仍停留在創造新商機的層次，然而善用雲端運算，可以替組織創造更多業務、行銷和人才培訓的機會。

作者：麥肯錫 出處：天下雜誌

過去五年，麥肯錫觀察重要科技發展，其中雲端、大量資訊 (big data)、智慧裝置 (smart assets) 三項，以超乎想像的速度發展。這三大技術，帶來五大趨勢，可被應用在企業營運及組織運作。先分別來看這三項技術：

第一、雲端運算。「雲端」在台灣已被一般民眾熟知。但我認為大家多半仍停留在雲端運算如何能創造新商機，卻很少好好思索，該怎麼運用雲端運算來替組織創造更多機會。特別是服務提供者，譬如電信業者、有線電視業者等，都應更有效應用雲端運算，為業務帶來更多機會。

第二、大量資訊。目前，絕大多數台灣企業，分析大量龐雜資料，仍使用類似微軟工具如 excel 等來整理。事實上，大量資訊經過快速運算分析，能更省時、省費用、有效的進行行銷活動。

第三、智慧裝置。如何善用監控器、智慧電表這類智慧裝置，來更優化公司營運。

參考來源：善用雲端架構 打造企業人才庫，作者：麥肯錫，出處：天下雜誌 455 期 (2010/09)
<http://www.cw.com.tw/article/print.jsp?id=41776>

雲端運算

大量資訊分析

智慧裝置

New Data Science : Social Network + Realtime Search

當「社交網路」遇上「即時搜尋」 = 即時市場行銷分析

創意行銷 / 臉書行銷 每天400萬顧客在線上

【經濟日報/潘俊琳】

2010.10.11 02:20 am

社交網站臉書Facebook的興起，重新定義了網路行銷的概念，大量的人潮讓業者彷彿看到滾滾錢潮，但臉書「開放平台」的模式，讓習慣有規則可循的行銷業者，必須開始學習全新的社群行銷，試著擁抱這項利器並串連消費者。

根據美國comScore的統計，美國網友8月分共花了1,140萬分鐘在臉書上，首次超越停留在Google旗下網站的時間，而臉書全球已經有超過5億的使用者，其中有35%的人每天登入。

快速分享 即時知道顧客反應

聖洋科技執行長邱繼弘表示，台灣臉書每個月約有700萬的累計使用人次，以60%每天上臉書的人口來算，就有420 萬人天天上線。

邱繼弘指出，臉書最大的行銷價值在於「開放平台」，只要符合它的基本規範，任何人、任何公司都可以在上面「免費」發揮自己的行銷創意。過去想要利用網路行銷，企業必須自己架站，林林總總的後台建設非常繁瑣，有多少人會來也是個問號？

但臉書幫企業解決了後台建設以及人潮，不論是企業或個人，只要成立自己的「粉絲專頁」，然後發揮行銷創意，回收可能比自己架站還更豐碩。因為臉書玩家只要在粉絲專頁按「讚」，就成為「粉絲團」的一員，往後企業發布在粉絲專頁的訊息，所有粉絲團成員都會收到，如果粉絲團的成員覺得某個行銷訊息不錯，只要按「分享」這個訊息就會出現在粉絲個人的臉書上，他所有的朋友就會看到這則行銷訊息，這是目前最高明的病毒式行銷。

社交網路

即時搜尋

評價排行榜

facebook

Facebook helps you connect and share with the people in your life.



參考來源：創意行銷 / 臉書行銷 每天 400 萬顧客在線上

【經濟日報 / 潘俊琳】

<http://udn.com/NEWS/FINANCE/FIN11/5901891.shtml>

2011 年 10 大策略科技

科技	影響
雲端運算	大型企業將會在 2012 年成立動態採購小組，專門負責雲端運算相關的決定以及管理。
媒體平板以及行動應用	2010 年將會有 12 億人使用具備上網能力的手機。隨著行動上網裝置以及應用程式日趨普及，與地點(location)、動作(motion)相關的應用軟體，可望進一步推動裝置的銷售。
社交溝通以及協作 (collaboration)	多數的公司在 2016 年已經把社交科技整合至多數的企業應用中，整合的範圍包含內部社交 CRM、溝通及協作以及外部社交網站。
影片	2013 年每位工作者看到的內容中，將有 25% 都是照片、影音。
次世代分析	隨著電腦、行動裝置運算能力、連結能力更強，影響企業如何決策，SAS 是長期領導廠商，IBM 以及甲骨文(Oracle) 事後起之秀。
社交分析	衡量人、主題以及想法的關係，範圍不限於社交網路，IBM 預計在 2011 年成為該領域的主要廠商之一。
情境感知運算 (context-aware computing)	較人工智慧更為寬廣，預計在 2013 年時 Fortune 500 大企業中超過半數會有相關採用方案。
儲存等級記憶體 (storage class memory)	快閃記憶體在消費性裝置、娛樂設備中的使用更多。
無所不在的運算 (ubiquitous computing)	儘管 Gartner 已經提及這個概念許多年，但隨著手機、射頻晶片更為普及，越多的物件可以連上網路。
架構化(fabric-based) 的基礎建設以及電腦	運算能力模組化，系統可以透過不同的模組來建構，可望提升效能。

資料來源：DIGITIMES 整理，2010/10

製表：雷佳宜、李盈瑩

雲端運算

平板行動應用

社交溝通協作

多媒體內容

次世代分析

社交分析

情境感知運算

儲存等級記憶體

無所不在的運算

模組化基礎建設

端

平板行動應用

社交溝通協作

多媒體內容

次世代分析

社交分析

情境感知運算

儲存等級記憶體

無所不在的運算

模組化基礎建設

雲端運算

SaaS :
Web 2.0

PaaS :
Big Data

IaaS :
Virtualization

社交網路

評價排行榜

即時搜尋

智慧裝置

大量資訊分析

雲端運算

雲

Two Type of Cloud Architecture ?

雲端架構的兩大陣營？

SaaS :
Web 2.0

IaaS :
Virtualization



想盡辦法誘你用計算跟網路
Computing Intensive

SaaS :
Web 2.0

PaaS :
Big Data



想盡辦法誘你提供資料作分析
Data Intensive

How can we build Cloud Services ??

觀察雲端關鍵提供者，找尋打造雲端服務的模式

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



可敬的對手
Reference Model

Hybrid
Cloud

以大型企業
為主要客戶
Enterprise is
key market



私有雲端
Private Cloud



- 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

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 介面, 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

User-Level

User-Level
Middleware

Core
Middleware

System Level

IaaS
PaaS
SaaS

Open Source to build Cloud Service

建構雲端服務的 自由軟體

應用軟體 Application

Social Computing, Enterprise, ISV, ...

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

程式語言 Programming

Web 2.0 介面, 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

- 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



Three Core Technologies of Google

Google 的三大關鍵技術

- Google 在一些會議分享他們的三大關鍵技術
- Google shared their design of web-search engine
 - SOSP 2003 :
 - “The Google File System”
 - <http://labs.google.com/papers/gfs.html>
 - OSDI 2004 :
 - “MapReduce : Simplified Data Processing on Large Cluster”
 - <http://labs.google.com/papers/mapreduce.html>
 - OSDI 2006 :
 - “Bigtable: A Distributed Storage System for Structured Data”
 - <http://labs.google.com/papers/bigtable-osdi06.pdf>



Open Source Mapping of Google Core Technologies

Google 三大關鍵技術對應的自由軟體

BigTable

A huge key-value datastore

HBase, Hypertable
Cassandra,

MapReduce

To parallel process data

Hadoop MapReduce API
Sphere MapReduce API, ...

Google File System

To store petabytes of data

Hadoop Distributed File System (HDFS)
Sector Distributed File System

更多不同語言的 MapReduce API 實作：

<http://trac.nchc.org.tw/grid/intertrac/wiki%3Ajazz/09-04-14%23MapReduce>

其他值得觀察的分散式檔案系統：

- IBM GPFS - <http://www-03.ibm.com/systems/software/gpfs/>
- Lustre - <http://www.lustre.org/>
- Ceph - <http://ceph.newdream.net/>