雲端運算簡介

王耀聰 陳威宇
Jazz@nchc.org.tw
waue@nchc.org.tw

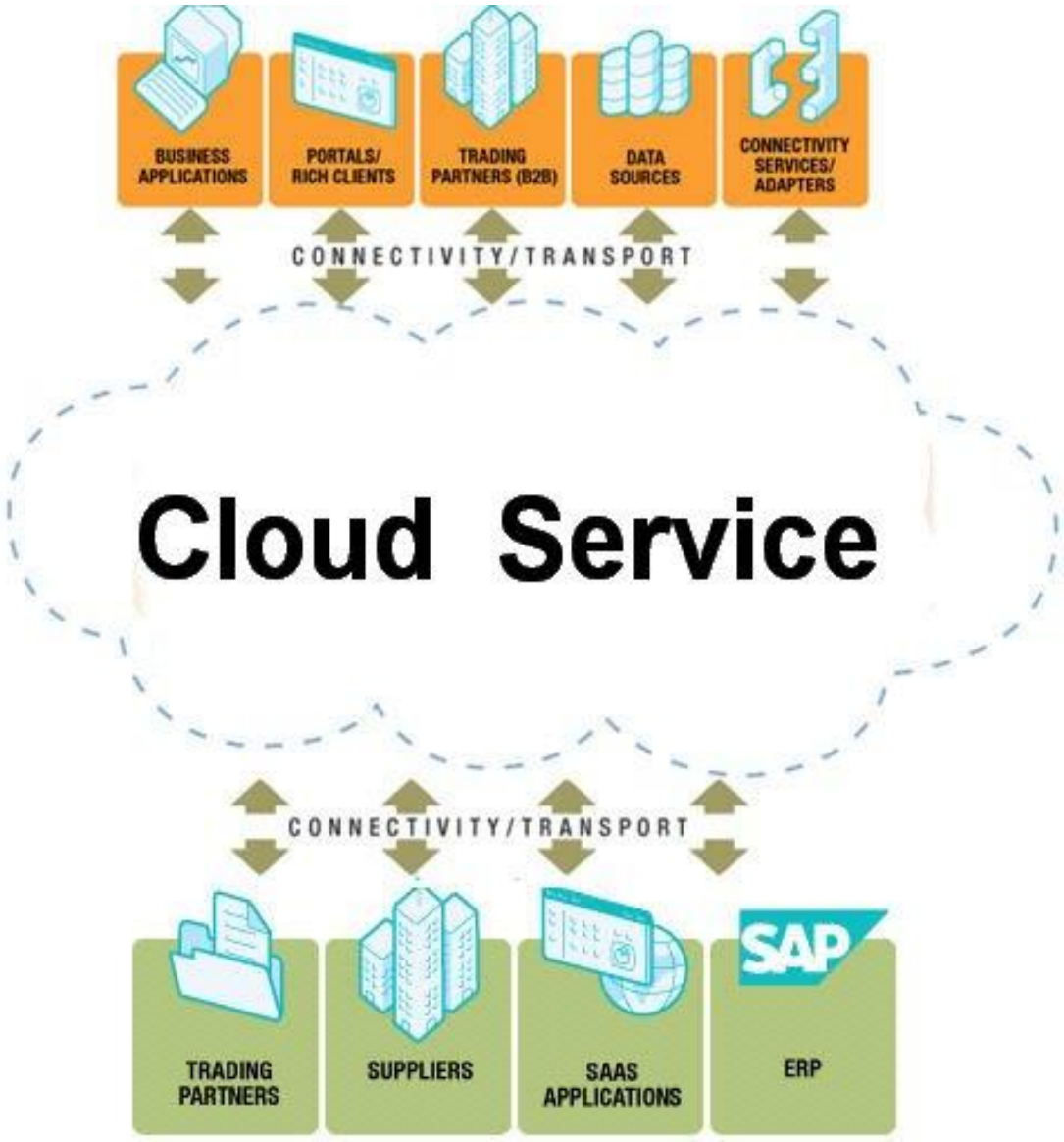
2008. 04 . 27-28

國家高速網路與計算中心(NCHC)


雲端運算？？

雲端服務

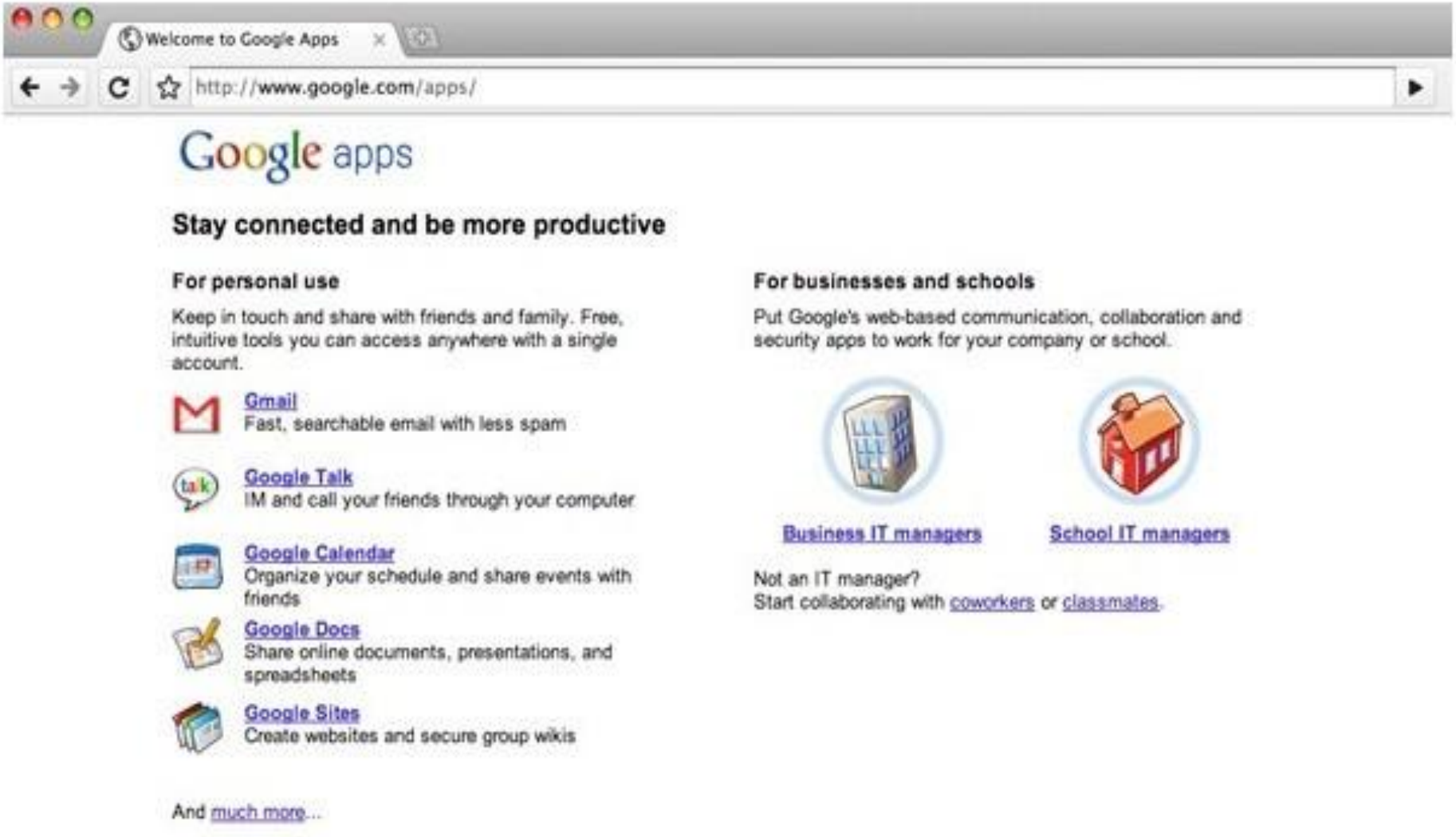
- 信件
- 影音
- 文書處理
- 相簿
- …

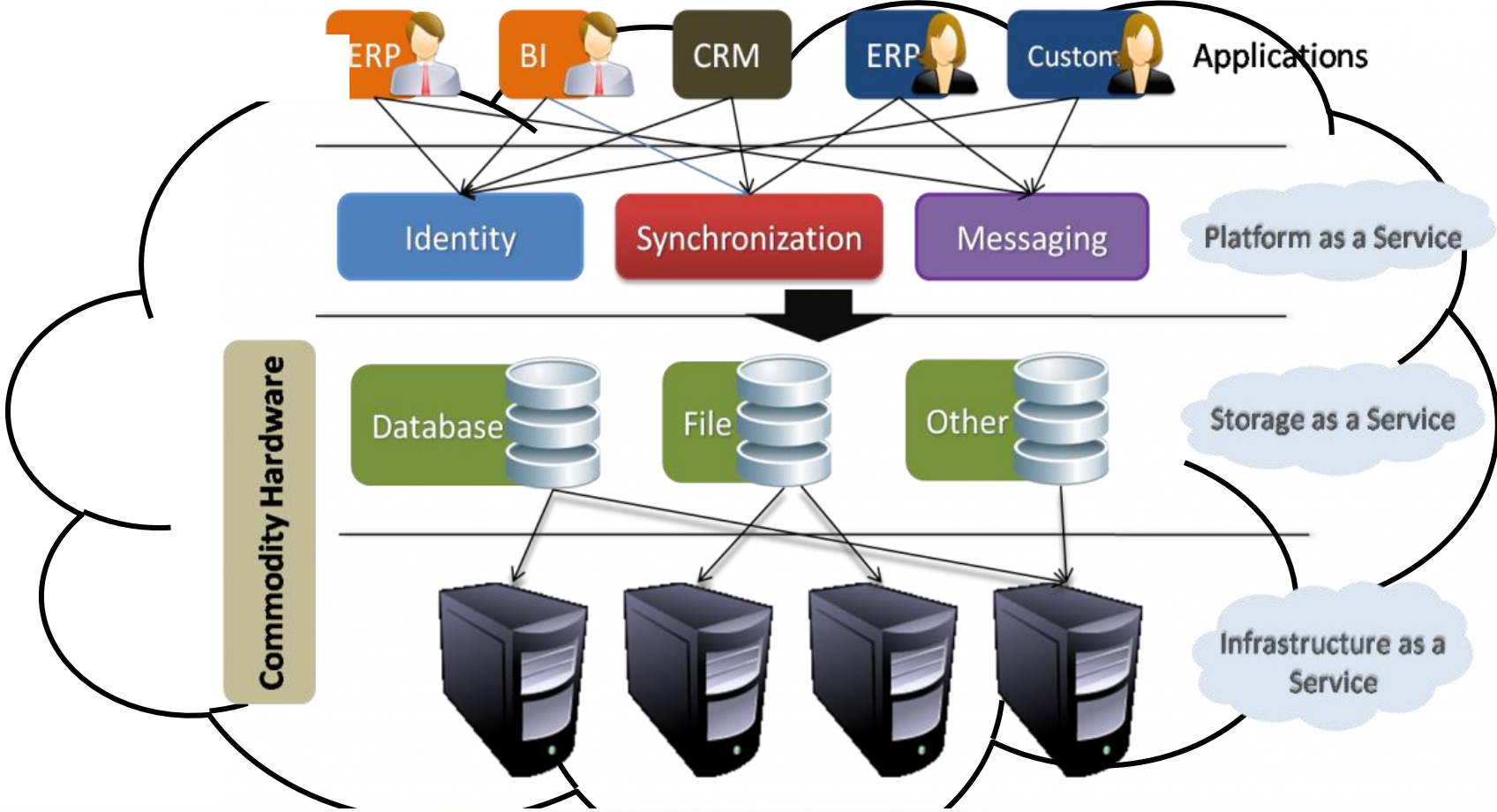


Google 發表 Chrome OS 作業系統：2010 年應用在輕省筆電（追加官方QA）

由 Atticus Wu 於 2 days 之前發表

文章分類: 網際網路, 軟體應用




ERP
BI
CRM
ERP
Custom
Applications
Identity
Synchronization
Messaging
Platform as a Service
Commodity Hardware
Database
File
Other
Storage as a Service
Infrastructure as a Service

雲端運算的架構

User Level

User-Level Middleware

Core Middleware

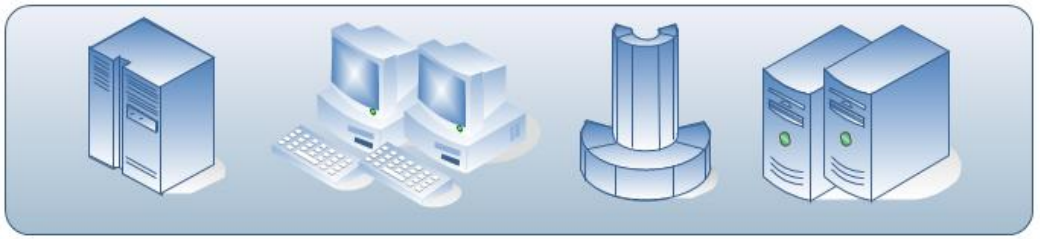
System Level

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



Amazon : Web Service

- AWS
- 虛擬化的技術：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

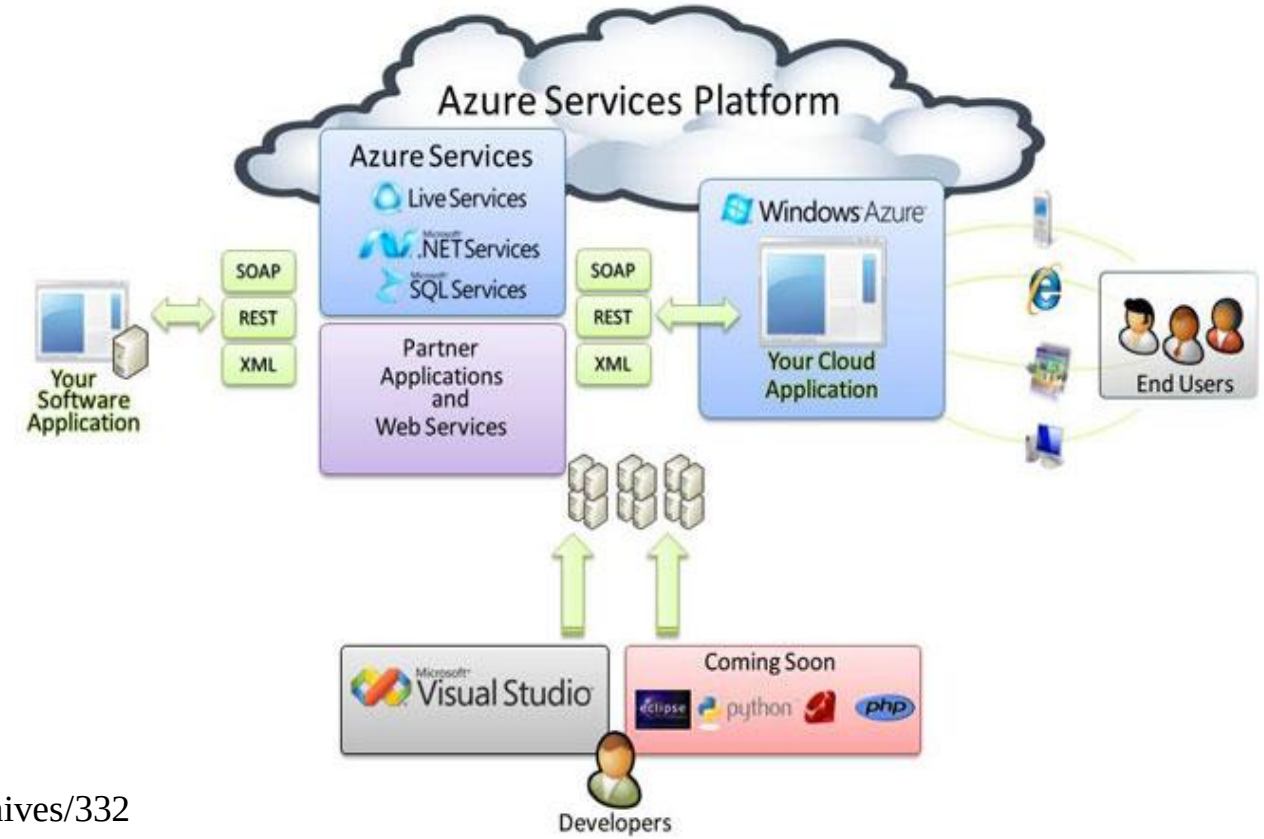
http://eblog.cisanet.org.tw/post/Cloud-Computing.aspx

Google : App Engine

- 網路平台，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
- Limit：
  - Language: Python、Java
  - web applications

Windows : Azure

- Windows 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。
- .Net services
- SQL services
- Live services



http://tech.cipper.com/index.php/archives/332

雲端運算產業類型

SaaS
Software as a Service

PaaS
Platform as a Service

IaaS
Infrastructure as a Service

www.spoutingshite.com/wp-content/uploads/2008/12/saas_and_cloud_computing.ppt

雲端運算產業

IaaS

架構即服務

- 提供了核心計算資源和網絡架構的服務
- infrastructure stack:
  - Full OS access
  - Firewalls
  - Routers
  - Load balancing

雲端運算產業

PaaS

IaaS

平台即服務

- 提供平台給系統管理員和開發人員，以為它構建、測試及部署定製應用程序
- 管理系統的成本昂貴
- Popular services
  - Storage
  - Database
  - Scalability

雲端運算產業

SaaS

PaaS

IaaS

軟體即服務

- 不用管理硬體與軟體
- 操作簡單 (瀏覽器)
- Pay per use
- Instant Scalability
- Security
- Reliability

比較表

服務 / 屬性	Amazon EC2	Google App Engine	Microsoft Azure	Yahoo Hadoop
架構	Iaas/Paas	Paas	Paas	Software
服務型態	Compute/ Storage	Web application	Web and non-web	Software
管理技術	OS on Xen hypervisor	Application container	OS through Fabric controller	Map / Reduce Architecture
使用者介面	EC2 Command-line tools	Web-based Administration console	Windows Azure portal	Command line and web
APIs	yes	yes	yes	yes
收費	yes	maybe	yes	no
程式語言	AMI (Amazon Machine Image)	Python	.NET framework	Java,

看了這麼多雲端服務

但.....

是否有一套能夠

開放給大家使用

的雲端平台呢??

就是你了
Hadoop !

Hadoop

- 以Google平台為仿效對象
- 創始者 Doug Cutting
- 以Java開發
- 自由軟體
- 上千個節點與Petabyte等級的資料量
- 為Apache 軟體基金會的 top level project

起源:2002-2004

- Lucene
  - 用Java設計的高效能文件索引引擎API
  - 索引文件中的每一字，讓搜尋的效率比傳統逐字比較還要高的多
- Nutch
  - nutch是基於開放原始碼所開發的web search
  - 利用Lucene函式庫開發

起源：Google論文

- Google File System
  - 可擴充的分散式檔案系統
  - 設計目的在於可以給大量的用戶提供總體性能較高的服務
  - 適用於分散式、對大量資訊進行存取的應用
  - 可運作在一般的普通主機上，且提供錯誤容忍的能力
- “The Google File System “ 發表於SOSP' 03 October，並將設計的概念公開

怎麼來的

起源：Google論文

- Google's GFS & MapReduce papers published:
  - SOSP 2003 : "The Google File System"
    - OSDI 2004 : "MapReduce : Simplifed Data Processing on Large Cluster"
    - OSDI 2006 : "Bigtable: A Distributed Storage System for Structured Data"
- directly address Nutch's scaling issues

http://research.google.com/pubs/papers.html

怎麼來的

起源:2004~

- Dong Cutting 開始參考論文來實做
- Added DFS & MapReduce implement to Nutch
- Nutch 0.8版之後，Hadoop為獨立項目
- Yahoo 於2006年僱用Dong Cutting 組隊專職開發
  - Team member = 14 （engineers, clusters, users, etc.)

系統特色

- 巨量
  - 擁有儲存與處理大量資料的能力
- 經濟
  - 可以用在由一般PC所架設的叢集環境內
- 效率
  - 籍由平行分散檔案的處理以致得到快速的回應
- 可靠
  - 當某節點發生錯誤，系統能即時自動的取得備份資料以及佈署運算資源

有誰在用

誰在用Hadoop

- Yahoo 為最大的贊助商
- IBM 與 Google 在大學開授雲端課程的主要內容
- Hadoop on Amazon Ec2/S3
- More…:

- A9.com
- ADSDAQ by Contextweb
- EHarmony
- Facebook
- Fox Interactive Media
- IBM
- ImageShack
- ISI
- Joost
- Last.fm
- Powerset
- The New York Times
- Rackspace
- Veoh
- Metaweb

Yahoo : Hadoop

- Apache 項目，Yahoo 資助、開發與運用
  - 2006年開始參與開源的雲端運算框架Hadoop，並將其使用在內部服務中。
  - 2008年2：目前最大的Hadoop應用
  - 2千臺伺服器
  - 執行超過1萬個Hadoop虛擬機器
  - 5 Petabytes的網頁內容
  - 分析1兆個網路連結

實用案例

Hadoop於yahoo的運作資訊

年份	日期	節點數	耗時（小時）
2006	四月	188	47.9
2006	五月	500	42
2006	十一月	20	1.8
2006	十一月	100	3.3
2006	十一月	500	5.2
2006	十一月	900	7.8
2007	七月	20	1.2
2007	七月	100	1.3
2007	七月	500	2
2007	七月	900	2.5

Sort benchmark, every nodes with terabytes data.

實用案例

Hadoop於yahoo的部屬情形

資料標題：Yahoo! Launches World's Largest Hadoop Production Application

資料日期：February 19, 2008

Number of links between pages in the index	roughly 1 trillion links
Size of output	over 300 TB, compressed!
Number of cores used to run single Map-Reduce job	over 10,000
Raw disk used in the production cluster	over 5 Petabytes

實用案例

Hadoop於yahoo的部屬情形

資料標題：Scaling Hadoop to 4000 nodes at Yahoo!

資料日期：September 30, 2008

Total Nodes	**4000**
Total cores	**30000**
Data	**16PB**

	500-node cluster		4000-node cluster	
	write	**read**	**write**	**read**
number of files	990	990	14,000	14,000
file size (MB)	320	320	360	360
total MB processes	316,800	316,800	5,040,000	5,040,000
tasks per node	2	2	4	4
avg. throughput (MB/s)	**5.8**	**18**	**40**	**66**

瞭解更多

Hadoop 與google的對應

Develop Group	Google	Apache
Sponsor	Google	Yahoo, Amazon
Algorithm Method	MapReduce	Hadoop
Resource	open document	open source
File System (MapReduce)	GFS	HDFS
Storage System (for structure data)	big-table	Hbase
Search Engine	Google	nutch
OS	Linux	Linux / GPL

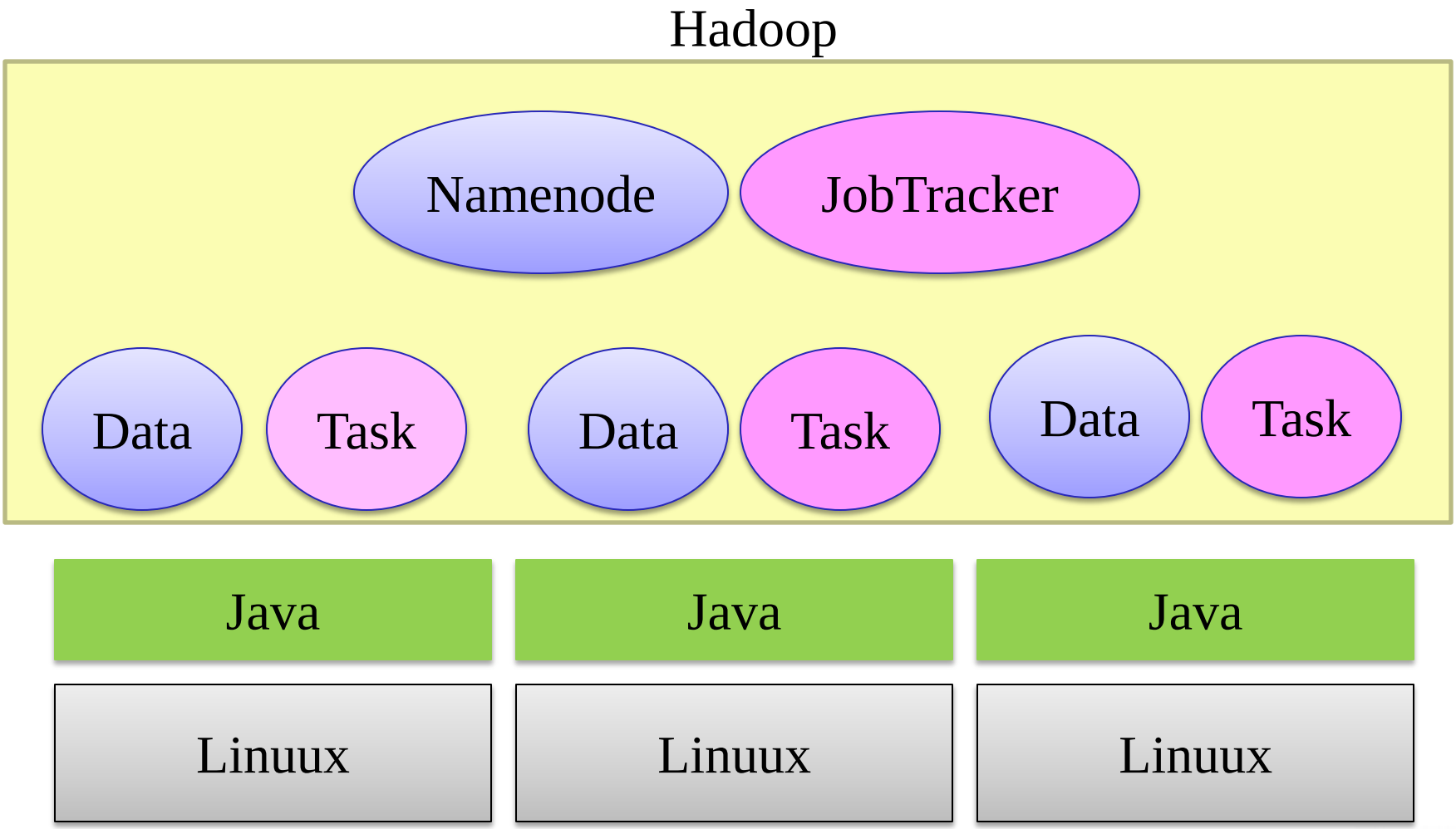
怎麼安裝？

連安裝都免！來hadoop.nchc.org.tw 申請個帳號就能用囉！

（完整安裝於此 http://trac.nchc.org.tw/cloud/wiki/0428Hadoop_Lab1）

作業系統的最核心！

Hadoop 平台架構圖



Hadoop 的資料儲存篇

Hadoop Distributed File System

HDFS？

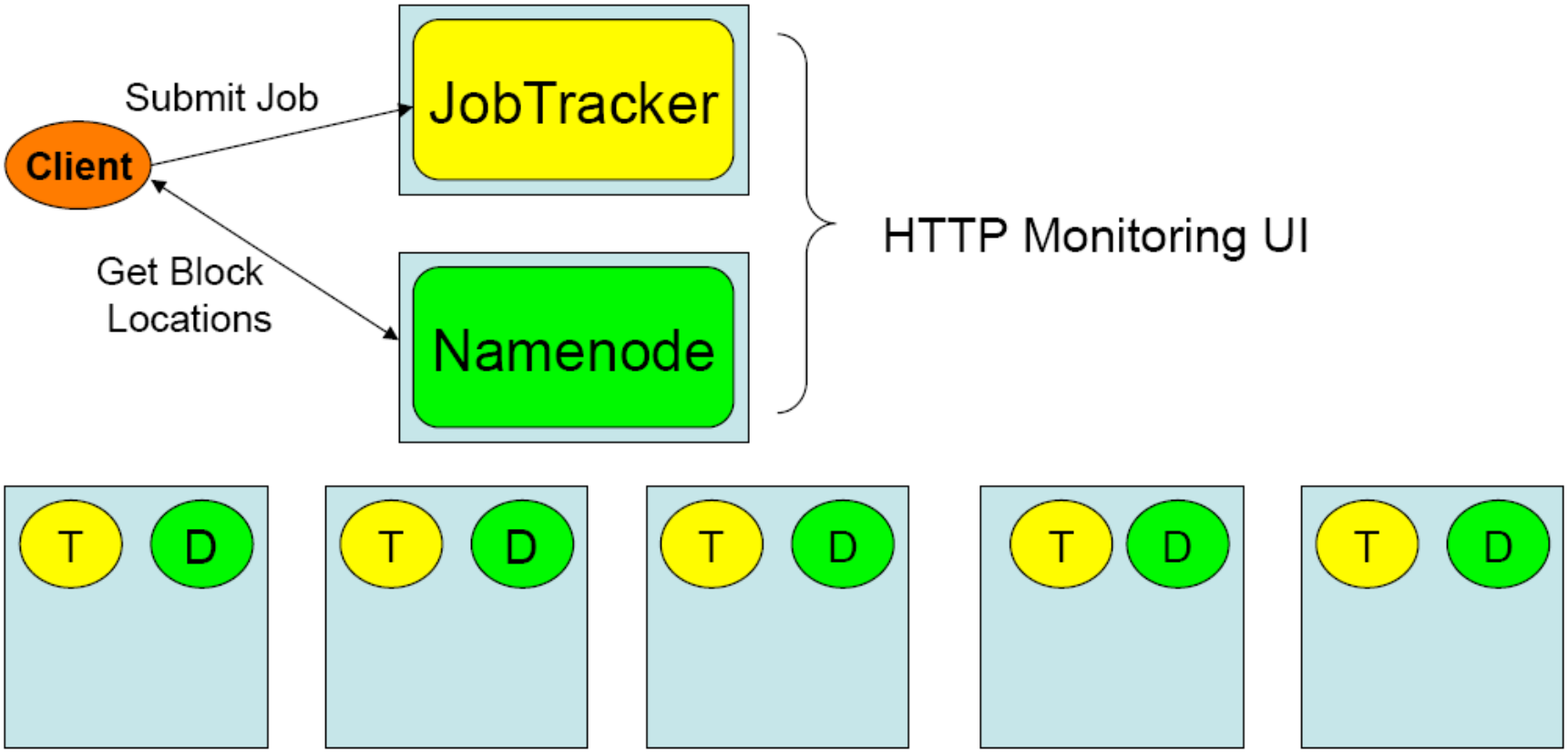
- Hadoop Distributed File System
  - Hadoop：自由軟體專案，為實現Google的MapReduce架構
  - HDFS: Hadoop專案中的檔案系統
- 實現類似Google File System
  - GFS是一個易於擴充的分散式檔案系統，目的為對大量資料進行分析
  - 運作於廉價的普通硬體上，又可以提供容錯功能
  - 給大量的用戶提供總體性能較高的服務

名詞

- Job
  - 任務
- Task
  - 小工作
- JobTracker
  - 任務分派者
- TaskTracker
  - 小工作的執行者
- Client
  - 發起任務的客戶端
- Map
  - 應對
- Reduce
  - 總和
- Namenode
  - 名稱節點
- Datanode
  - 資料節點
- Namespace
  - 名稱空間
- Replication
  - 副本
- Blocks
  - 檔案區塊 (64M)
- Rack awareness
  - 用來告知網路拓樸狀況
- Metadata
  - 屬性資料

HDFS的架構？

架構


Submit Job
Client
JobTracker
Get Block Locations
Namenode
HTTP Monitoring UI
T
D
T
D
T
D
T
D
T
D

管理資料

Namenode

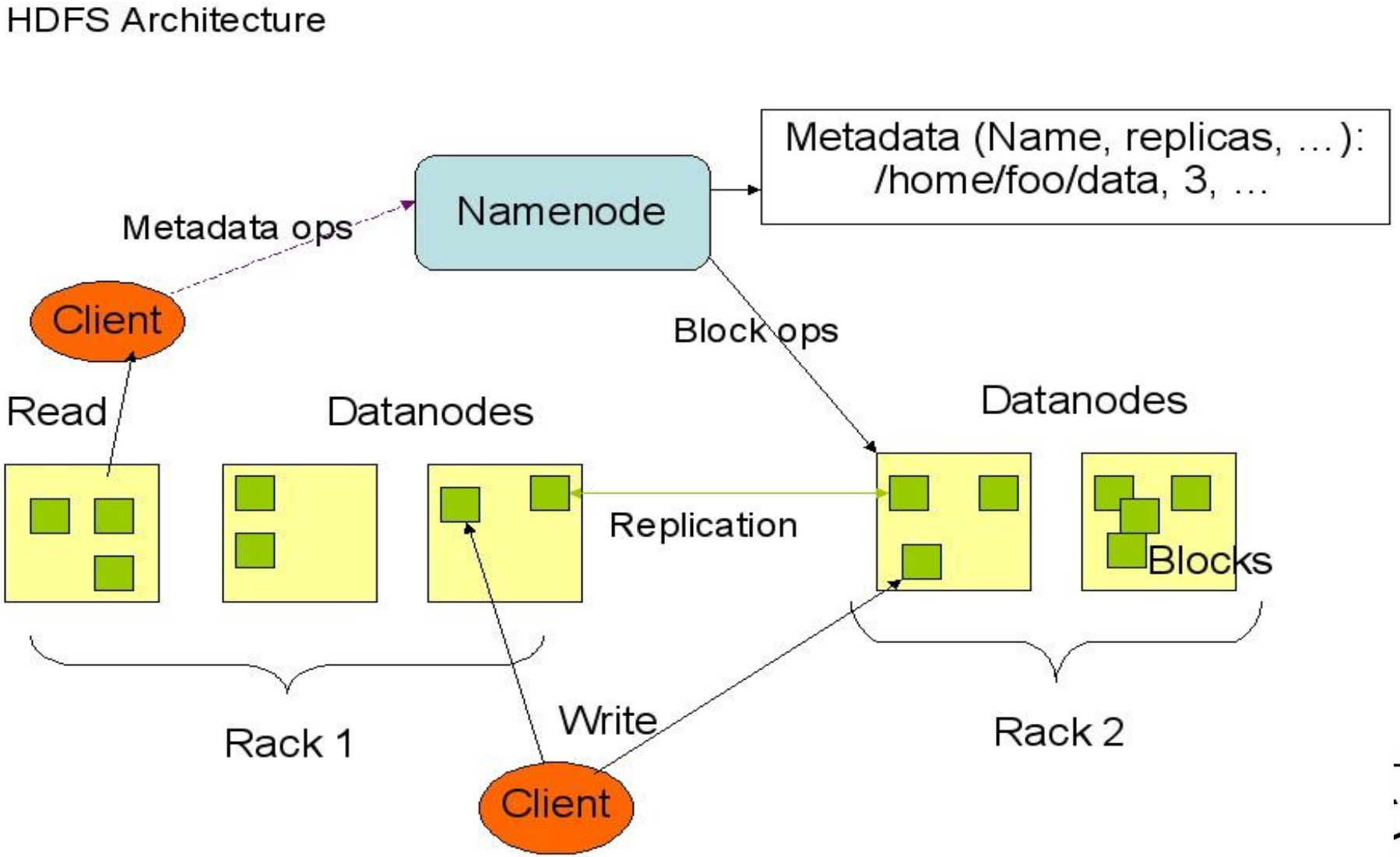
- Master
- 管理HDFS的名稱空間
- 控制對檔案的讀/寫
- 配置副本策略
- 對名稱空間作檢查及紀錄

Datanode

- Workers
- 執行讀/寫動作
- 執行Namenode的副本策略

HDFS的架構？

管理資料

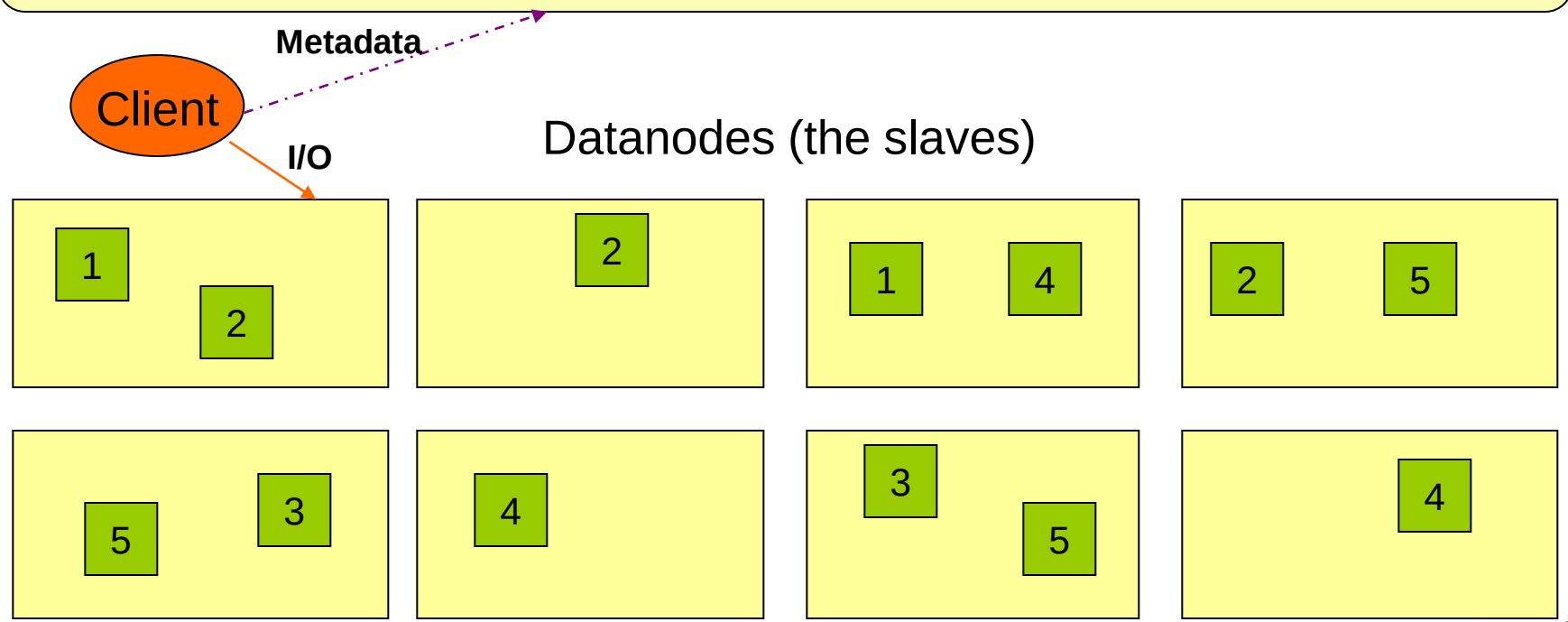


HDFS 運作

Namenode (the master)

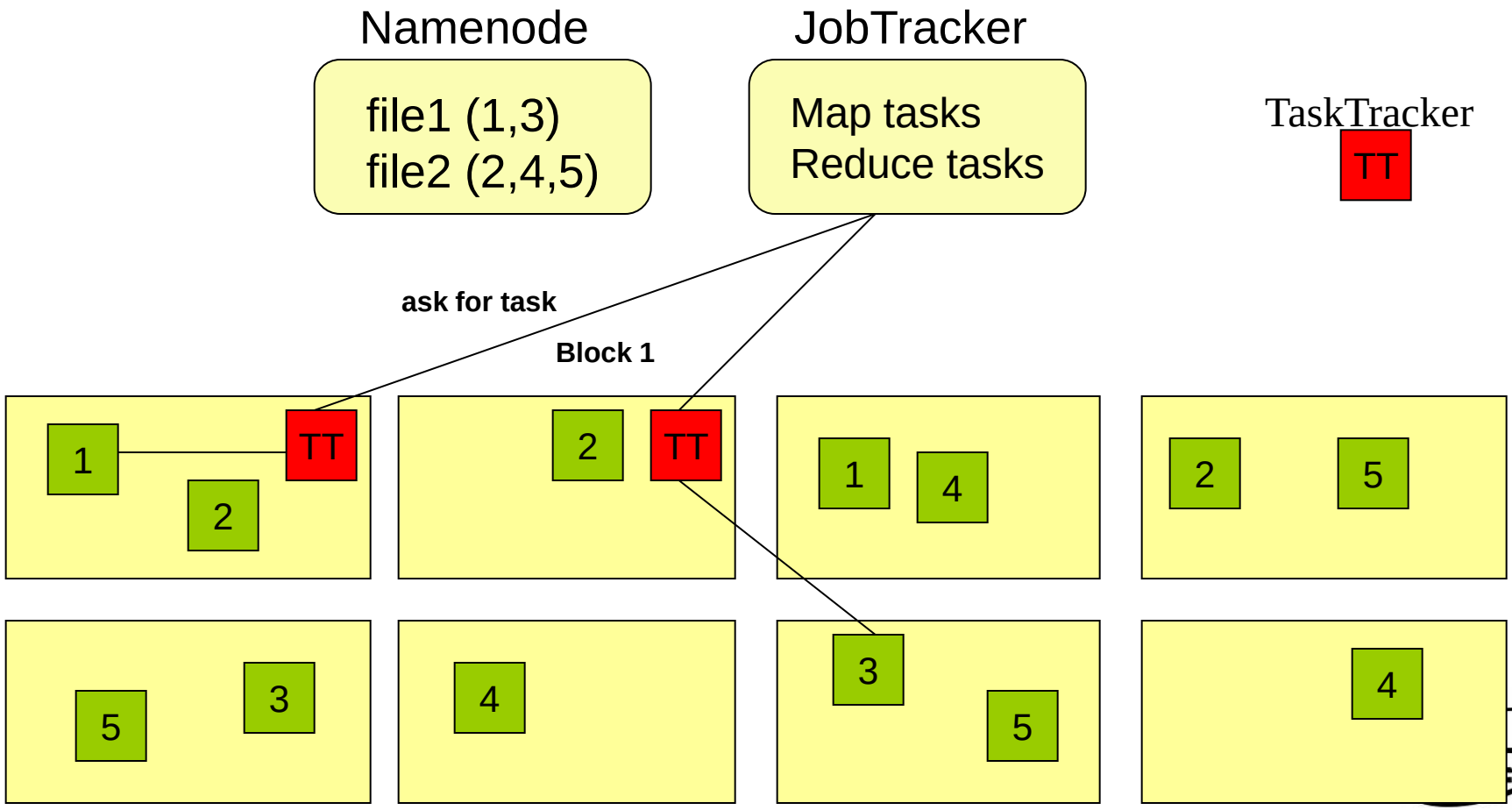
檔案路徑– 副本數 , 由哪幾個block組成

name:/users/joeYahoo/myFile - copies:2, blocks:{1,3}
name:/users/bobYahoo/someData.gzip, copies:3, blocks:{2,4,5}



HDFS 運作

- 目的：提高系統的可靠性與讀取的效率
  - 可靠性：節點失效時讀取副本已維持正常運作
  - 讀取效率：分散讀取流量（但增加寫入時效能瓶頸）



如何達成
其好處？

可靠性機制

常見的三種錯誤狀況

- 資料崩毀
- 網路或資料節點失效
- 名稱節點錯誤

- 資料完整性
  - checked with CRC32
  - 用副本取代出錯資料
- Heartbeat
  - Datanode 定期向Namenode送heartbeat
- Metadata
  - FSImage、Editlog為核心印象檔及日誌檔
  - 多份儲存，當NameNode壞掉可以手動復原

HDFS的功能

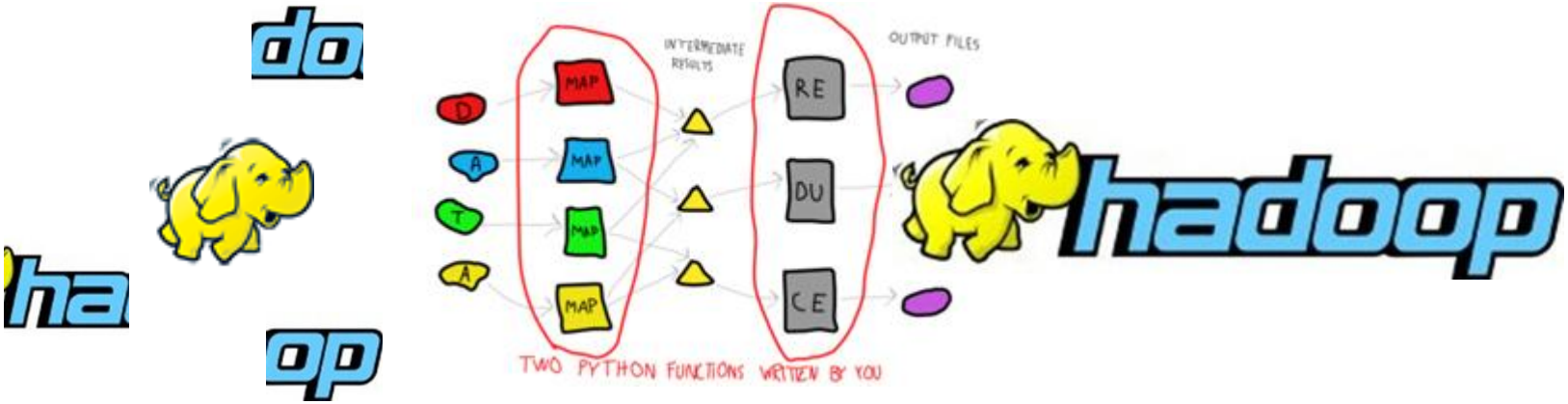
- 類POXIS指令
- 權限控管
- 超級用戶模式
- Web 瀏覽
- 用戶配額管理
- 分散式複製檔案

動手囉

Hadoop 的程序分配篇

Map / Reduce

Hadoop MapReduce定義



Hadoop Map/Reduce是一個易於使用的軟體平台，以MapReduce為基礎的應用程序，能夠運作在由上千台PC所組成的大型叢集上，並以一種可靠容錯的方式平行處理上P級別的資料集。

分派程序

Jobtracker

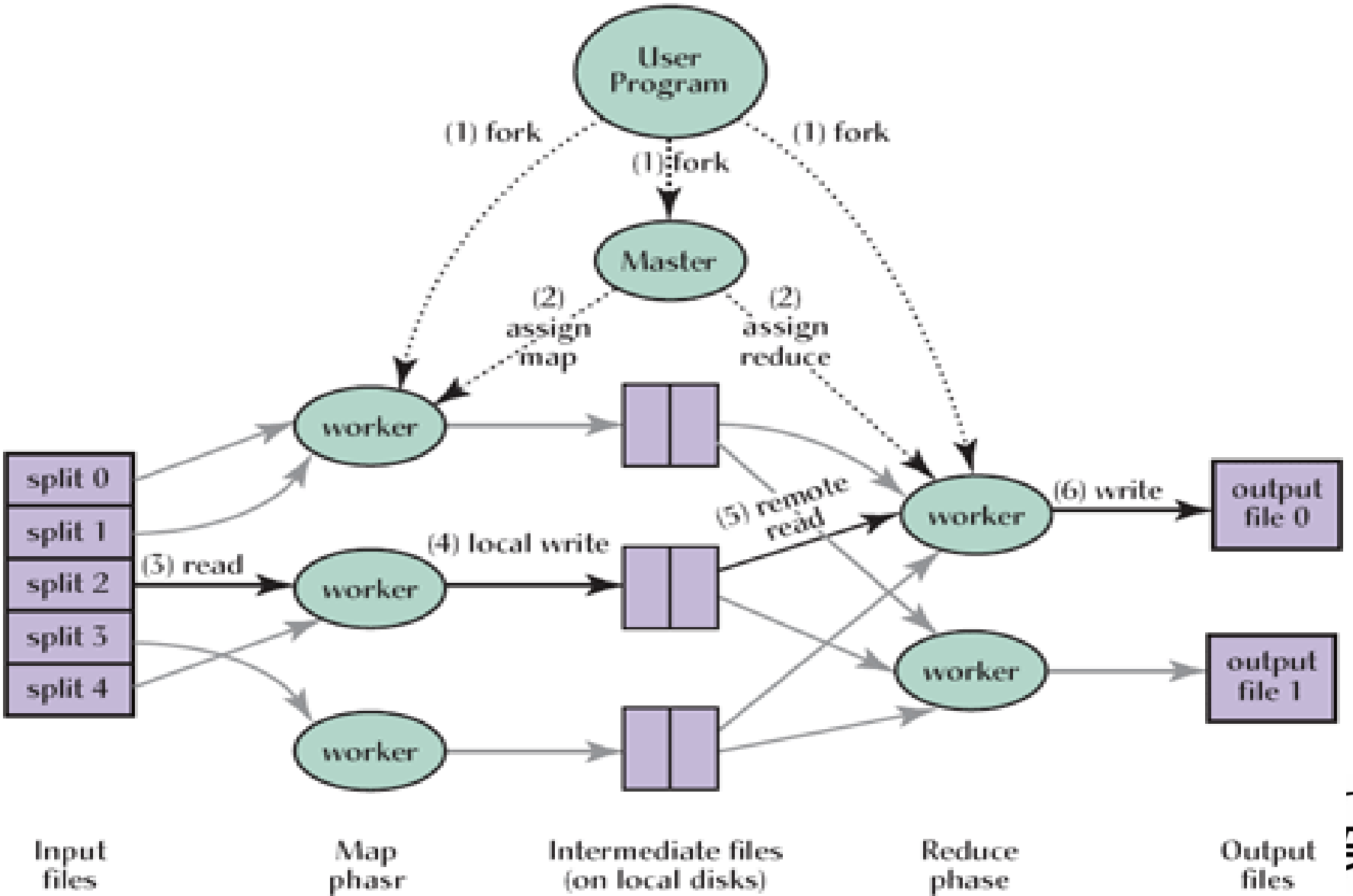
- Master
- 使用者發起工作
- 指派工作給 Tasktrackers
- 排程決策、工作分配、錯誤處理

Tasktrackers

- Workers
- 運作Map 與 Reduce 的工作
- 管理儲存、回覆運算結果

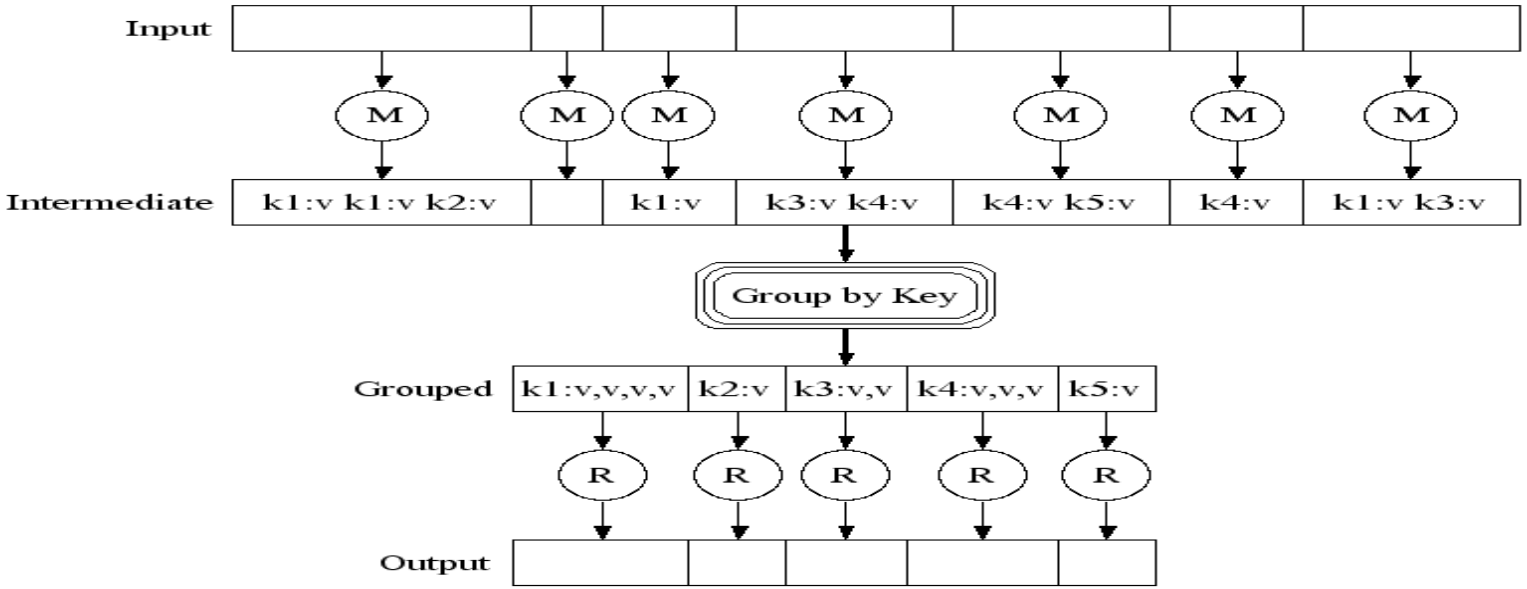
HDFS的架構？

分派程序



MapReduce

- 什麼是MapReduce
  - Map 將每個資料視為一個key，並作<key,value>的配對，Reduce再統合所有的Map結果做出<key, list(value)>
- 運作方法



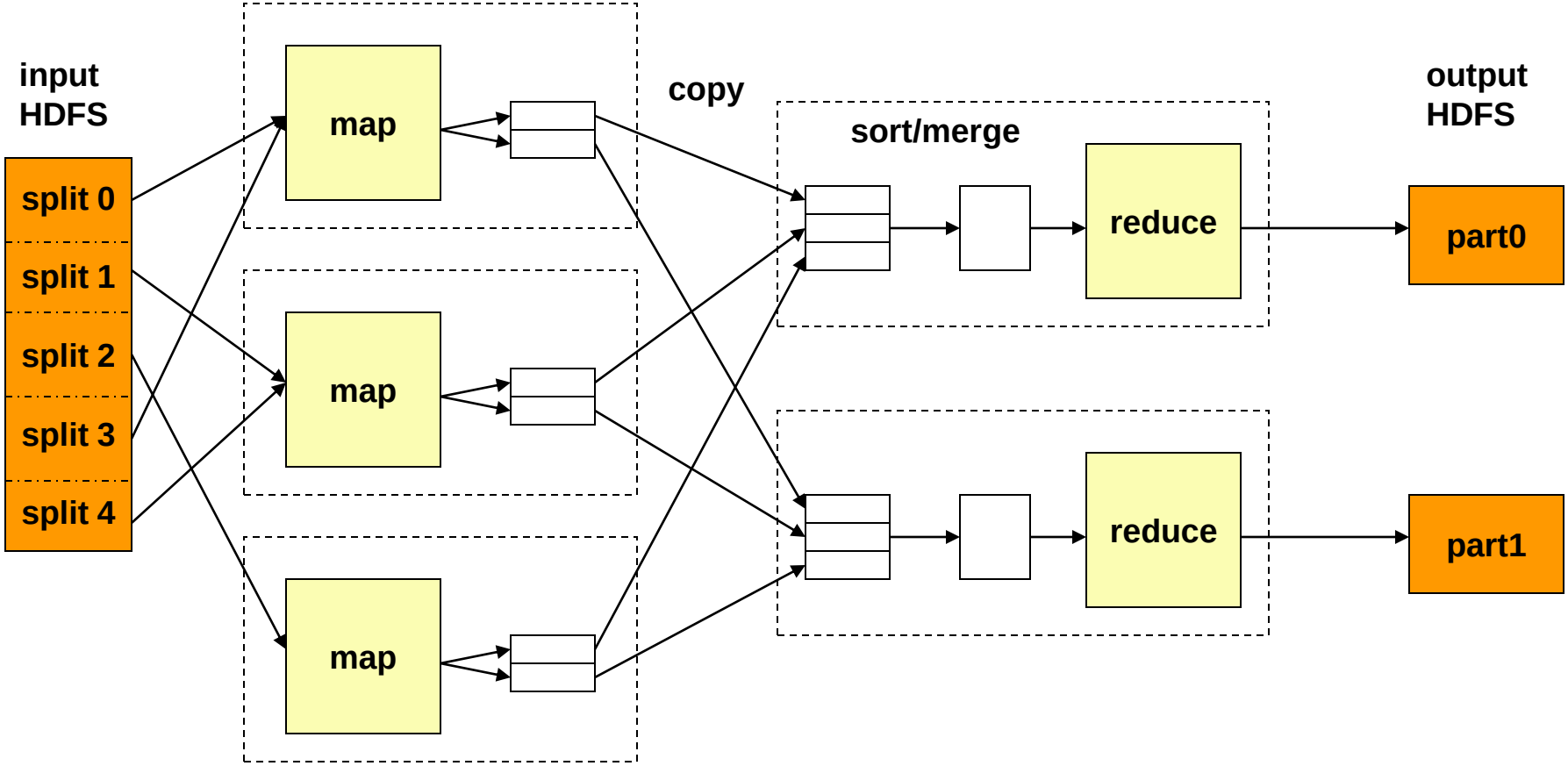
Key-Value 遙不可及？

- Text tokenization
- Indexing and Search
- Data mining
- machine learning
- …

來個範例吧！

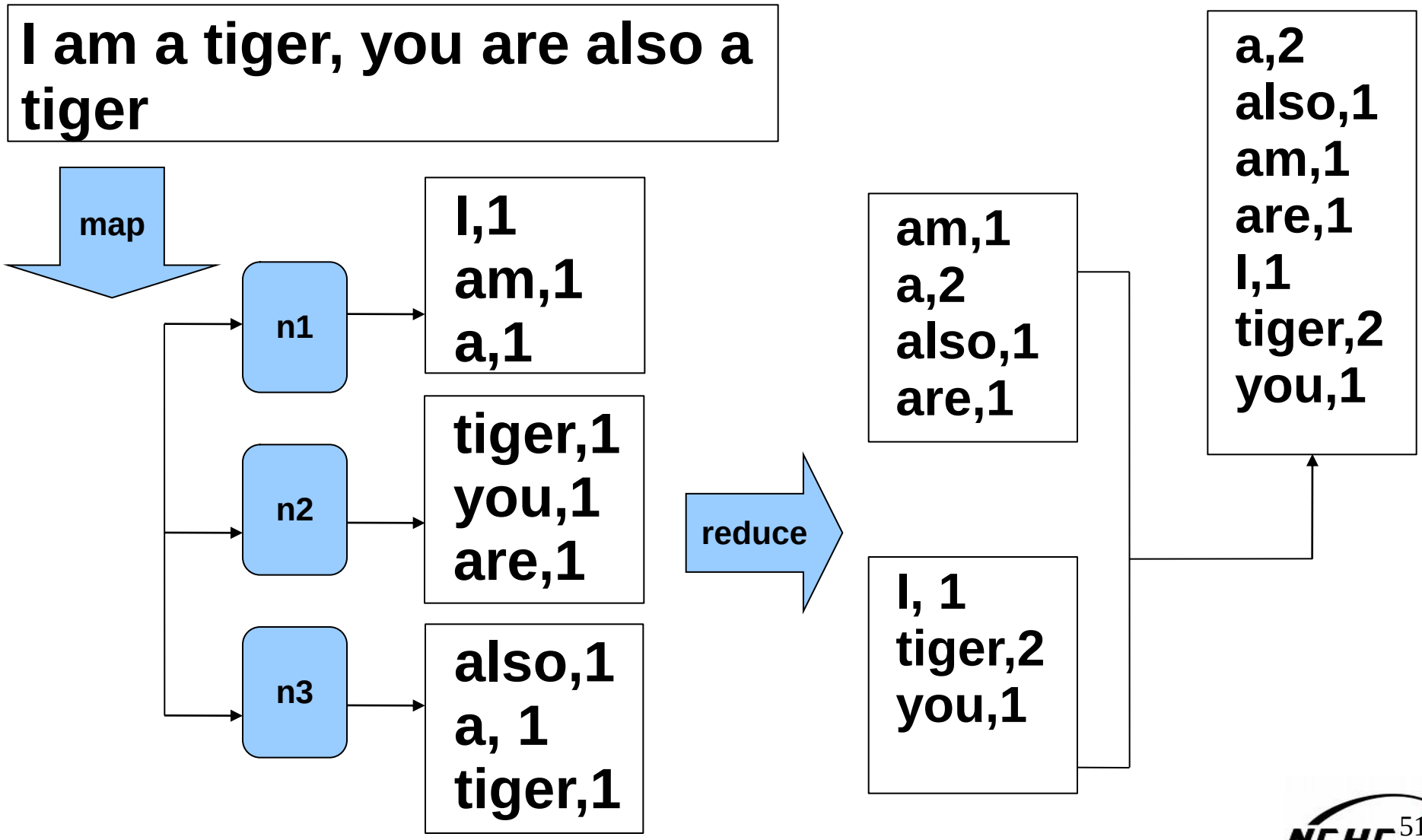
http://picasaweb.google.com.tw/cloudexam/

MapReduce 運作流程



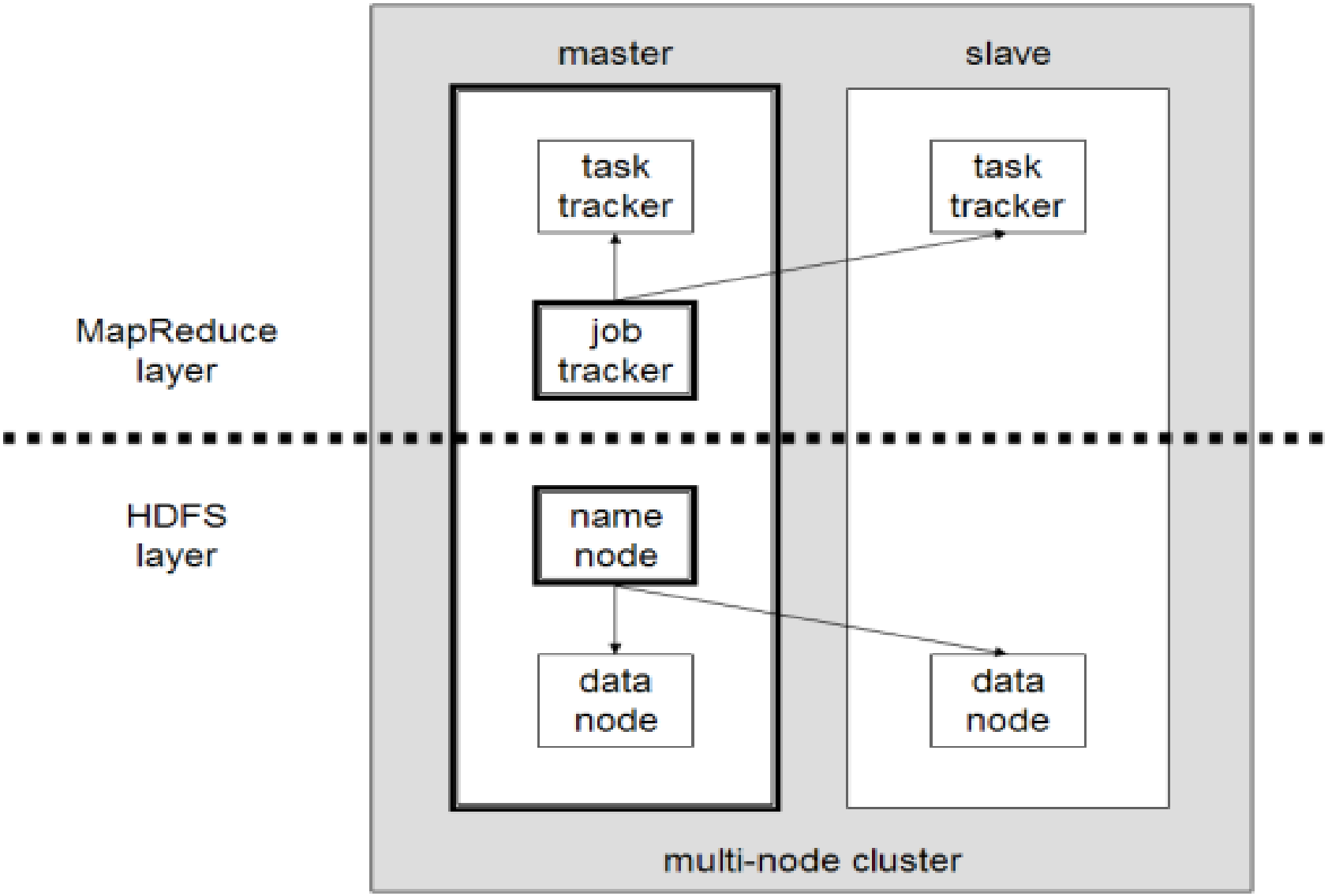
在hadoop平台上進行運算

範例


I am a tiger, you are also a tiger
map
n1
n2
n3
I,1
am,1
a,1
tiger,1
you,1
are,1
also,1
a, 1
tiger,1
reduce
am,1
a,2
also,1
are,1
I, 1
tiger,2
you,1
a,2
also,1
am,1
are,1
I,1
tiger,2
you,1

HDFS的架構？

HDFS + MapReduce



其他相關專案

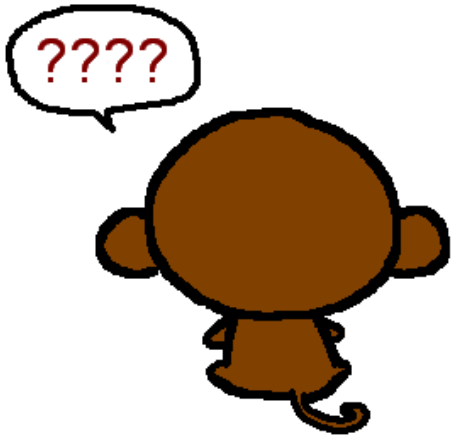
- HBase （http://hadoop.apache.org/hbase/）
  - 用Hadoop為基礎的雲端資料庫
- Nutch （http://lucene.apache.org/nutch/）
  - 以Hadoop 為基礎的搜尋引擎
- Pig (http://hadoop.apache.org/pig/)
  - 一個可用在Hadoop上的平台，提供一個全新語言(Pig Latin)以簡化撰寫分析的程式

- Disco （http://discoproject.org/）
  - Nokia所研發的MapReduce架構，用erlang 實做，使用者可以python驅動，類似Hadoop的自由軟體專案。

文獻參考

- Hadoop 官方網站
  - http://hadoop.apache.org/core/
- Hadoop API
  - http://hadoop.apache.org/core/docs/r0.18.3/api/index.html
- Hadoop Taiwan User Group
  - http://www.hadoop.tw/
- 中文Hadoop手冊
  - http://cn.hadoop.org/doc/index.html
- 維基百科
  - http://en.wikipedia.org/wiki/Hadoop
  - http://zh.wikipedia.org/wiki/Hadoop

Question ?



http://chinese.storylands.org/1329magic_al/story1c15.php

Thank You !

http://miumiu516.pixnet.net/album/photo/94262410