財團法人國家實驗研究院
國家高速網路與計算中心
NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING

Hadoop 簡介

王耀聰 陳威宇
jazz@nchc.org.tw
waue@nchc.org.tw
國家高速網路與計算中心(NCHC)

自由軟體實驗室

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??

The Other Open Source Projects:

Eucalyptus	University of California, Santa Barbara	http://open.eucalyptus.com
Sector	The National Center for Data Mining (NCDM)	http://sector.sf.net
Thrift	Facebook	http://developers.facebook.com/thrift

Hadoop ?

Hadoop is a software platform that lets one easily write and run applications that process vast amounts of data

什麼是
Hadoop

Hadoop

- 以 Java 開發
- 自由軟體
- 上千個節點
- Petabyte 等級的資料量
- 創始者 Doug Cutting
- 爲 Apache 軟體基金會的 top level project

有什麼
特色

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
  —藉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

歷史背景

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

歷史背景

起源：Google 論文

- Google File System
  - —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”
- 可擴充的分散式檔案系統
- 大量的用户提供總體性能較高的服務
- 對大量資訊進行存取的應用
- 運作在一般的普通主機上
- 提供錯誤容忍的能力

歷史背景

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 爲獨立項目
- Yahoo 於 2006 年僱用 Dong Cutting 組隊專職開發
  - Team member = 14 （engineers, clusters, users, etc.）
- 2009 年 8 月 Dong Cutting 跳槽到 Cloudera

有誰
在用

誰在用 Hadoop

- Yahoo 爲最大的贊助商
- 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

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

Sort benchmark, every nodes with terabytes data.

實用案例

Hadoop 於 Yahoo 的佈屬情形

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 的佈屬情形

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	MapReduce
Resource	open document	open source
File System	GFS	HDFS
Storage System (for structure data)	big-table	Hbase
Search Engine	Google	Nutch
OS	Linux	Linux / GPL

動手安裝囉！

財團法人國家實驗研究院
國家高速網路與計算中心
NATIONAL CENTER FOR HIGH-PERFORMANCE COMPUTING

Hadoop Overview

王耀聰 陳威宇
jazz@nchc.org.tw
waue@nchc.org.tw
國家高速網路與計算中心(NCHC)

自由軟體實驗室

作業系統的最核心！

儲存空間的資源管理

記憶體空間與
行程分配

名詞

- Job
  —任務
- Task
  —工作
- JobTracker
  —任務分派者
- TaskTracker
  —工作的執行者
- Client
  —發起任務的客户端
- Map
  —對應
- Reduce
  —彙整

- Namenode
  —名稱節點
- Datanode
  —資料節點
- Namespace
  —名稱空間
- Replication
  —副本
- Blocks
  —檔案區塊 (64M)
- Metadata
  —屬性資料

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
- 只能有一個

Datanode

- Workers
- 執行讀 / 寫動作
- 執行 Namenode 的副本策略
- 可多個

分派程序

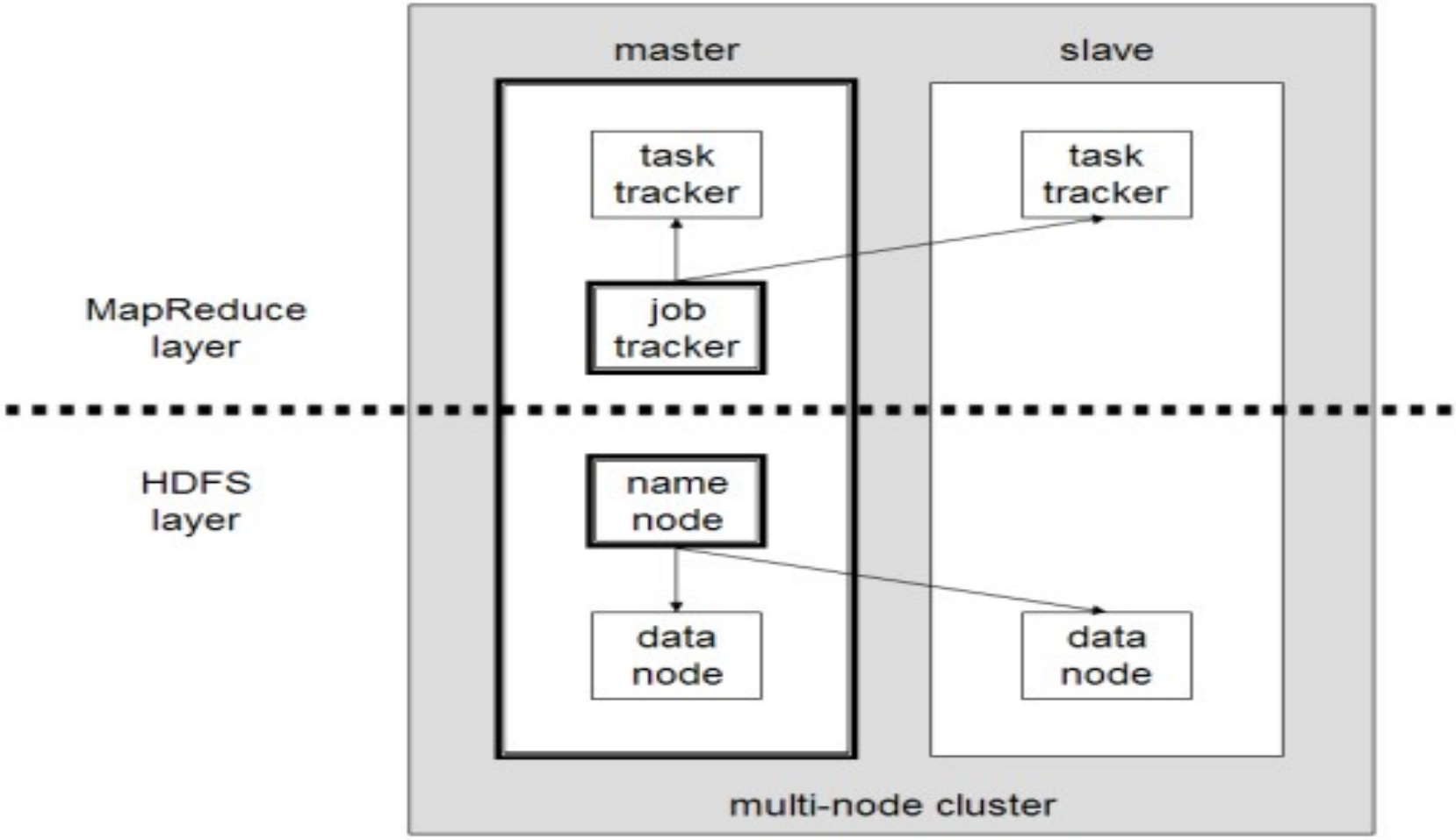
Jobtracker

- 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
- 只能有一個

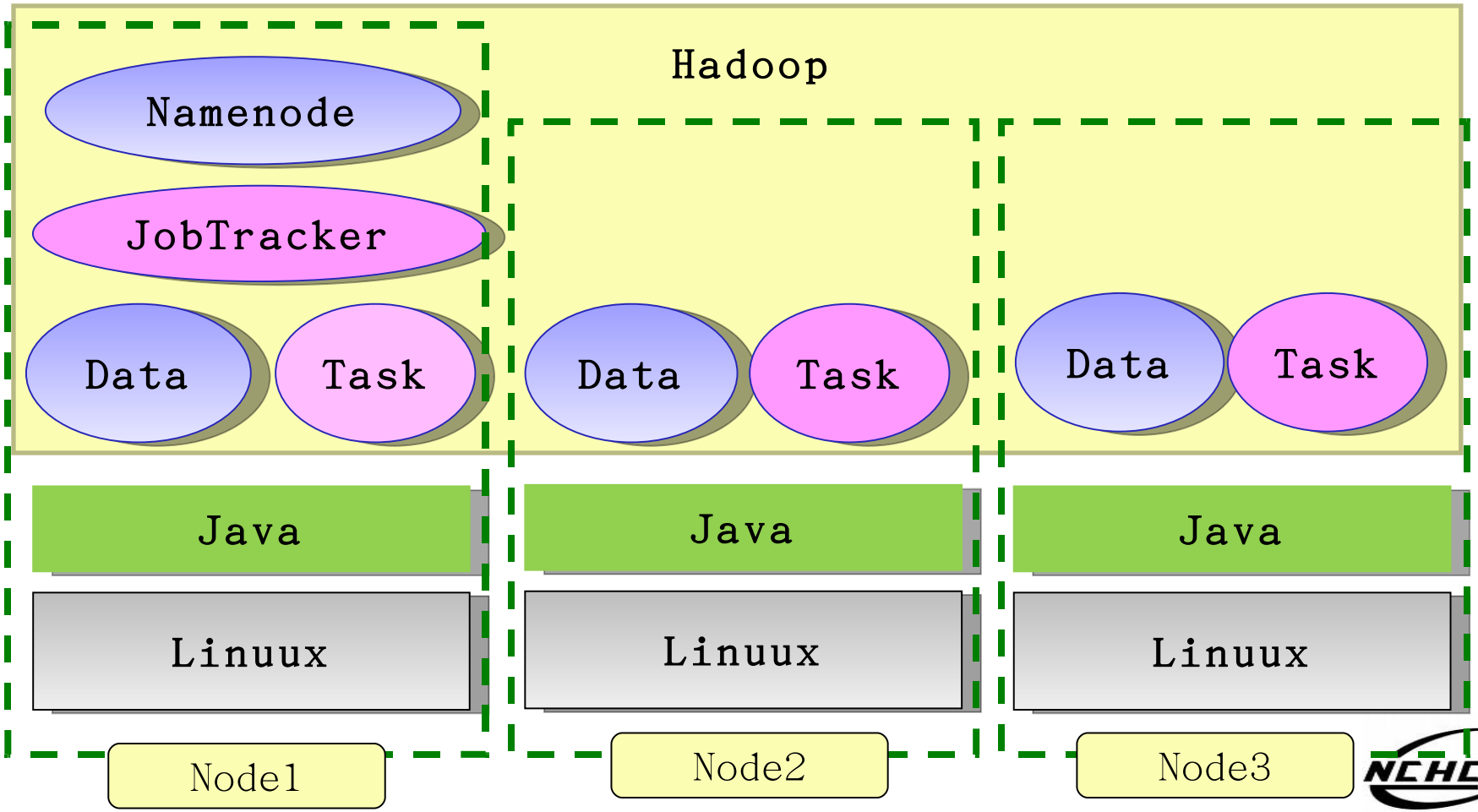
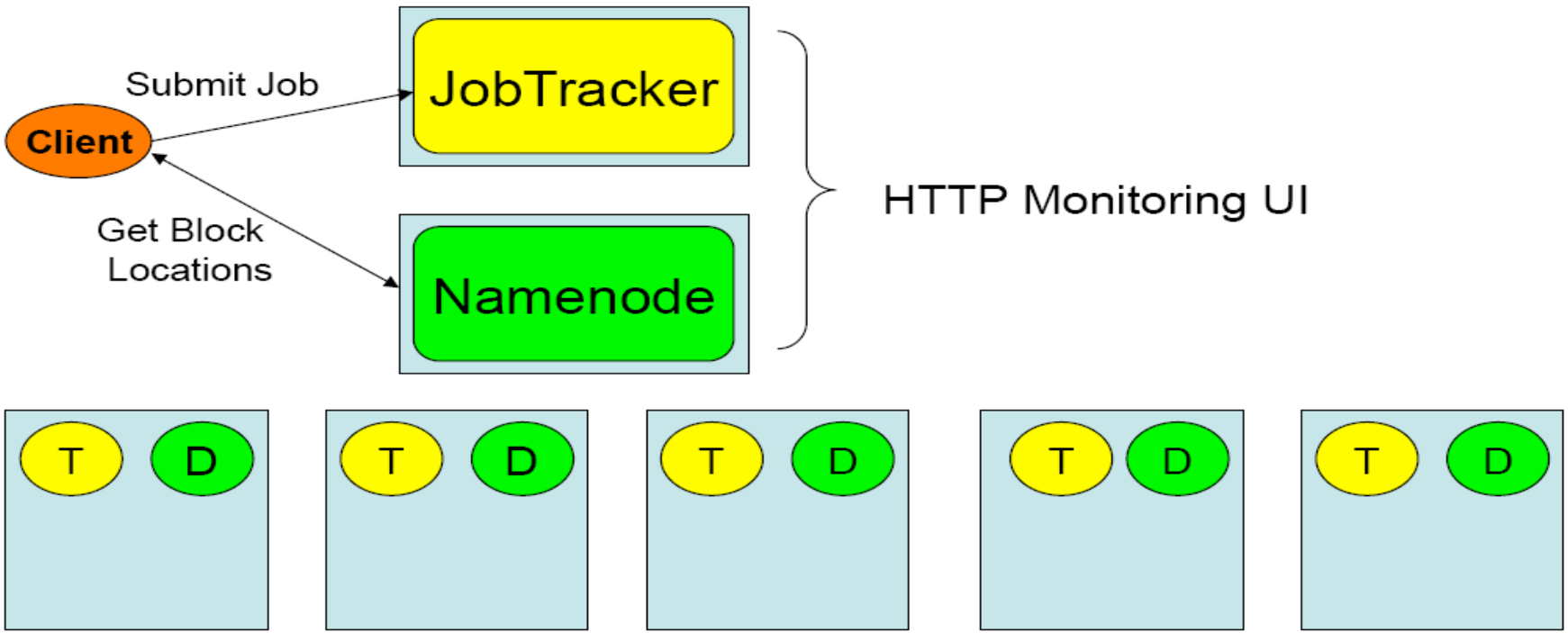
Tasktrackers

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

Hadoop 的各種身份



Building Hadoop



不在雲裡的 Client