雲端運算平台-Hadoop介紹與應用

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

關於我-楊順發

- 學經歷
  - 2009~ 國家高速網路與計算中心 軟體技術組 助理研究員
  - 2009 國立台灣科技大學 電子工程所計算機組 畢業
- 認證資格及參賽經歷
  - Cloudera Certified Hadoop Developer
  - SCJP
  - 2010年開放原始碼創新應用開發大賽職業組冠軍
- 目前
  - Crawlzilla專案維護與開發
  - Big Data Research

Thinking in Big Data

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

TAIWAN www.nchc.org.tw National Applied Research Laboratories NCHC

Big Data Age

Inside Ancestry.com's Top-Secret Data Center
Posted by Diane

Inside the unassuming building that is the data center for Ancestry.com and other Generations Network properties, rows and rows of cabinets house the 5,328 servers that hold the Web site, all those indexes and digital images, and users' family trees.

In all, it's 2.5 petabytes of data (one petabyte is equivalent to 283,000 DVDs).

A lot of security protects that data. A guard watches cameras 24/7. Windows are bulletproof. Sensors monitor windows and doors. The Ancestry.com guy walking us around had to swipe his badge at several doors, then lay his palm in a *Mission: Impossible*-like handprint reader to enter the server rooms.

I can't disclose the location and photographs weren't permitted (darn it, I forgot my hidden-camera lapel pin), but the folks at Ancestry.com sent these approved images:

Some rows of server-filled cabinets:

- Ancestry.com, the genealogy site, stores around 2.5 petabytes of data.

Big Data Age

New York Stock Exchange Ticks on Data Warehouse Appliances

Netezza deployment replaces mega data warehouses while cutting query times from hours to seconds.

By **Doug Henschen** InformationWeek
May 16, 2008 04:59 PM

It's not a typical enterprise data warehouses story, but then NYSE Euronext (NYSE), the parent company of the New York Stock Exchange, is not a typical enterprise. For one thing, NYSE has not one but three warehouses, each approaching 100 terabytes. Then consider NYSE's queries, some of which interrogate more than 40 terabytes of data. The extreme data volumes and extreme query complexity led to an upgrade onto data warehouse appliances.

After a period of rampant growth and mergers with two smaller exchanges, NYSE knew its large and aging Oracle data warehouses needed replacement. After exploring alternatives in 2006, the company concluded a successful 45-day proof-of-concept project on a Netezza Performance Server (NPS) appliance in early 2007. The main warehouse for the New York Stock Exchange was migrated within two and a half months and went into production in May 2007. A second device, consolidating what had been two separate warehouses for the Chicago-based Arca Equities and Options markets, went into production in July. Yet another warehouse, one housing legacy data, will be migrated onto a third Netezza NPS.

Big Data Age

Facebook Trumps Most Photo Sharing Sites With 10 Billion Photos

October 15, 2008 by Stan Schroeder

74

Sometimes I forget how big the Internet is, and then something reminds me just how flabbergastingly, enormously huge the damn thing is. This time it's Facebook, whose software engineer Doug Beaver announced that it now hosts a total of 10 billion photos. And it's not even exclusively a photo sharing site!

Furthermore, since Facebook stores four sizes for each stored photo, this actually translates to 40 billion files. However, what I find fascinating about this is that Facebook is no Photobucket; it's not just some repository (at least from what I've seen) where people dump all kinds of images just because they can. Vast majority of photos I see on Facebook are actually real life photographs taken by users; I'm not sure if this means much in the grand scheme of life, universe and everything, but it seems like quite an accomplishment.

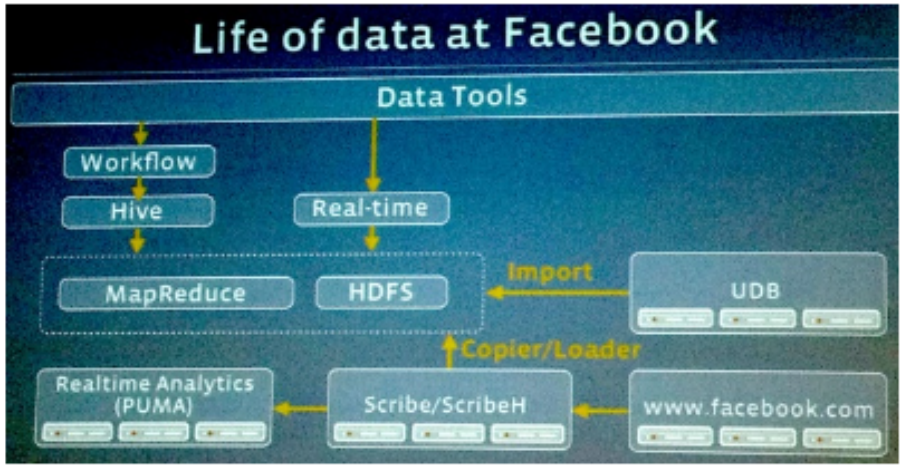
Big Data Age

FB大數字：每天處理3億張照片、25億則發文、27億個「讚」

數位時代網站｜撰文者：陳荻雅編譯　發表日期：2012-08-25

讚　23 人說這讚。成為你朋友中第一個說讚的人。　+1　5

Facebook對媒體揭露了身為社群網站龍頭的網站資料量實際數據－包括每天系統需處理超過25億則的發文、每天乘載超過500TB的流量、每天會有27億個「讚」、每日總上傳照片數約3億張、每半小時就要掃描約105TB的資料。



What Is Big Data?

From Databases to Big Data

Sam Madden • *Massachusetts Institute of Technology*

- Among all the definitions offered for “big data,” my favorite is that it means data that’s too big, too fast, or too hard for existing tools to process.

資料存儲

- 容量：1,370MB, 傳輸速度：4.4M/s
  - 1990年代，傳輸完一顆硬碟的速度約5分鐘
- 100M/s的傳輸速度
  - 2010年代
  - 每天在社群網站約產生500T的資料量
  - 複習一下：K -> M -> G -> T
  - 傳輸完一天所產生的資料需要約**1388888.89**小時，也就是**57870.37**天，也就是**158.55**年…
  - 這！還不包括資料分析的時間！

資料分析 - Traditional

- 以美國氣象資料為例：

0057
332130 # USAF weather station identifier
99999 # WBAN weather station identifier
19500101 # observation date
0300 # observation time
4
+51317 # latitude (degrees x 1000)
+028783 # longitude (degrees x 1000)
FM-12
+0171 # elevation (meters)
99999
V020
320 # wind direction (degrees)
1 # quality code
N
0072
1
00450 # sky ceiling height (meters)
1 # quality code
CN
010000 # visibility distance (meters)
1 # quality code
N9
-0128 # air temperature (degrees Celsius x 10)
1 # quality code
-0139 # dew point temperature (degrees Celsius x 10)
1 # quality code
10268 # atmospheric pressure (hectopascals x 10)
1 # quality code

資料分析 – Traditional

- 利用awk進行分析 1901~2001的所有資料

```
#!/usr/bin/env bash
for year in all/*
do
echo -ne `basename $year .gz`"\t"
gunzip -c $year | \
awk '{ temp = substr($0, 88, 5) + 0;
q = substr($0, 93, 1);
if (temp !=9999 && q ~ /[01459]/ && temp > max) max = temp }
END { print max }'
done
```

- 執行結果

```
1901 317
1902 244
1903 289
1904 256
1905 283
```

- ***費時：42min on EC2 High-CPU Extra Large Instance***

Hadoop 簡介

楊順發

shunfa@nchc.narl.org.tw

自由軟體實驗室

國家高速網路與計算中心

Outline

- Hadoop?
- Why?
- Hadoop History
- Hadoop Ecosystem
- Hadoop Releases

Hadoop

- Hadoop是Apache軟體基金會所研發的開放源碼並行運算編程工具和分散式檔案系統，與MapReduce和Google檔案系統的概念類似。

About

- 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
- 為Apache 軟體基金會的 top level project

Why Hadoop?

- **Need to process Multi Petabyte Datasets**
- **Data may not have strict schema**
- **Expensive to build reliability in each application.**
- **Nodes fail every day**
  - Failure is expected, rather than exceptional.
  - The number of nodes in a cluster is not constant.
- **Need common infrastructure**
  - Efficient, reliable, Open Source Apache License

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

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
  - 大量的用戶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

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.)
- 2009 年跳槽到Cloudera

起源

- **Nutch**
  - nutch是基於開放原始碼所開發的web search
  - 利用Lucene函式庫開發
- **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

History

- Dec 2004 – Google GFS paper published
- July 2005 – Nutch uses MapReduce
- Feb 2006 – Starts as a Lucene subproject
- Apr 2007 – Yahoo! on 1000-node cluster
- Jan 2008 – An Apache Top Level Project
- May 2009 – Hadoop sorts Petabyte in 17 hours
- Aug 2010 – World's Largest Hadoop cluster at Facebook
  - 2900 nodes, 30+ PetaByte

Who Use Hadoop?

- Amazon/A9
- Facebook
- Google
- IBM
- Joost
- Last.fm
- New York Times
- PowerSet
- Veoh
- Yahoo!

- **http://wiki.apache.org/hadoop/PoweredBy**

用途

- Search
  - *Yahoo, Amazon, Zvents*
  - *關於Yahoo!與Hadoop*
    - http://www.slideshare.net/ydn/hadoop-yahoo-internet-scale-data-processing
- Log processing
  - *Facebook, Yahoo, ContextWeb. Joost, Last.fm*
- Recommendation Systems
  - *Facebook*
- Data Warehouse
  - *Facebook, AOL*
- Video and Image Analysis
  - *New York Times, Eyealike*

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

Features supported by Hadoop release series

Feature	1.x	0.22	2.X
Secure authentication	Y	N	Y
Old configuration names	Y	Deprecated	Deprecated
New configuration names	N	Y	Y
Old MapReduce API	Y	Y	Y
New MapReduce API	Y	Y	Y
MapReduce 1 runtime (Classic)	Y	Y	N
MapReduce 2 runtime (YARN)	N	N	Y
HDFS federation	N	N	Y
HDFS high-availability	N	N	Y

Hadoop Overview

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

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**
  - 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)
- **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

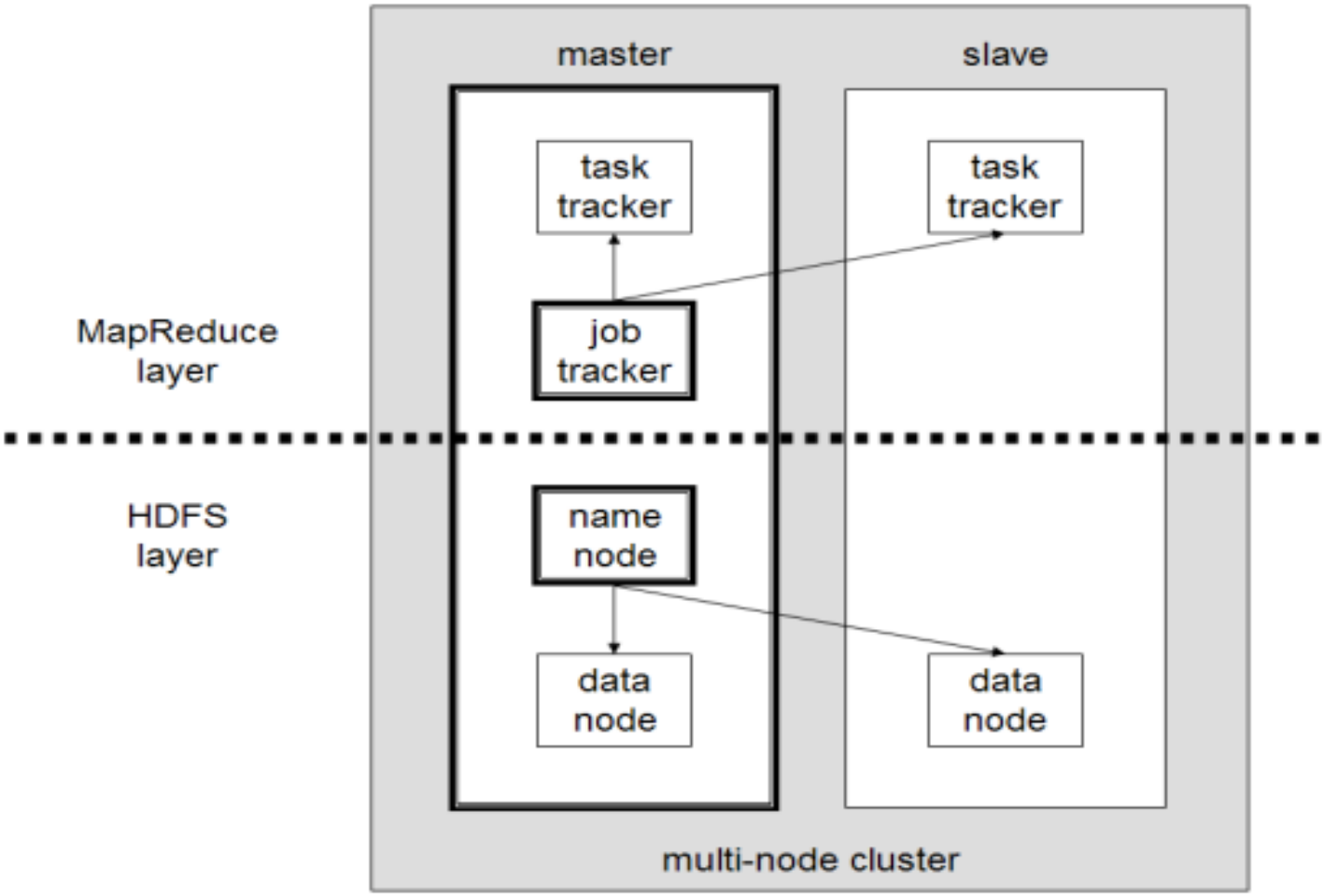
Jobtracker

- 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

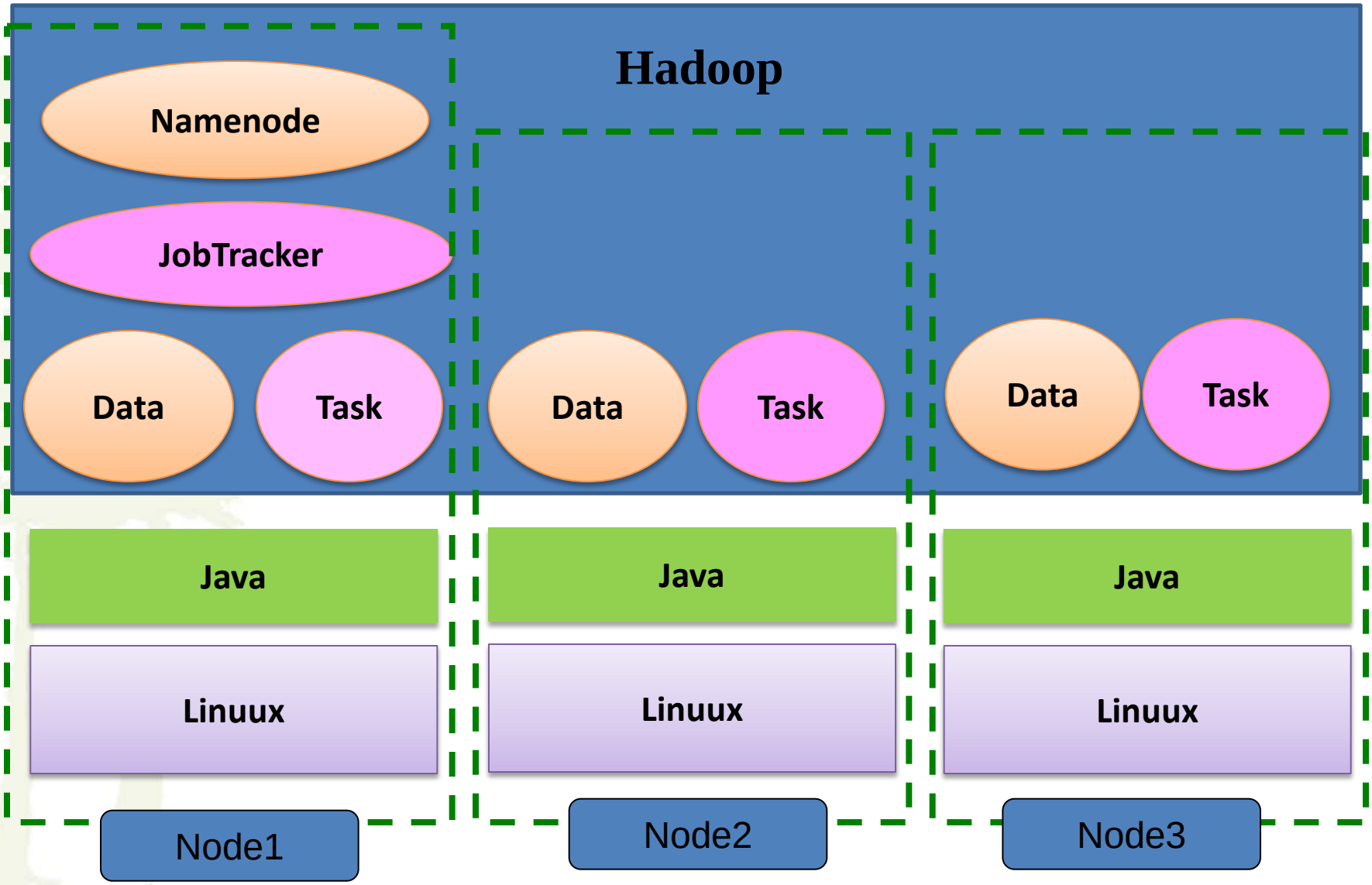
Tasktrackers

- 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



HDFS介紹

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

Outline

- HDFS?
- The Design of HDFS
- HDFS Concepts
- HDFS High-Availability
- HDFS Federation
- POSIX Like
- Data Flow

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

Goals of HDFS

- **Very Large Distributed File System**
  - 10K nodes, 1 billion files, 100 PB
- **Assumes Commodity Hardware**
  - Files are replicated to handle hardware failure
  - Detect failures and recovers from them

Goals of HDFS

- **Optimized for Batch Processing**
  - Data locations exposed so that computations can move to where data resides
  - Provides very high aggregate bandwidth
- **User Space, runs on heterogeneous OS**
- Streaming data access
  - Write Once, read many times.

不適合HDFS的情況

- Low-latency data access
- Lots of small files
- Multiple writers, arbitrary file modifications

Blocks

- 檔案存儲的最小單位
- **In HDFS, Block Size is 64MB by default**
- **Why?**
  - to minimize the cost of seeks
- 查看檔案**block**分配情形
  - % hadoop fsck / -files -blocks

Block Placement

- **Policy**
  - 在本端機架的本端節點上放置一份複本
  - 在本端機架的不同節點上放置第二份複本
  - 在遠端機架上放置第三份複本
  - 他複本則隨機放置
- **用戶端會讀取位置最近的複本**

Namenodes and Datanodes

- **Namenode**
  - 只能有一個(Master)
  - 負責存儲檔案系統的metadata
  - In Local Disk
    - namespace image
    - edit log
- **Datanode**
  - 可多個
  - 檔案系統的workhorses

Heartbeats

- DataNode 傳送Heartbeats給 NameNode
  - 每 3 秒傳送一次
- NameNode 使用Heartbeats來偵測DataNode問題。

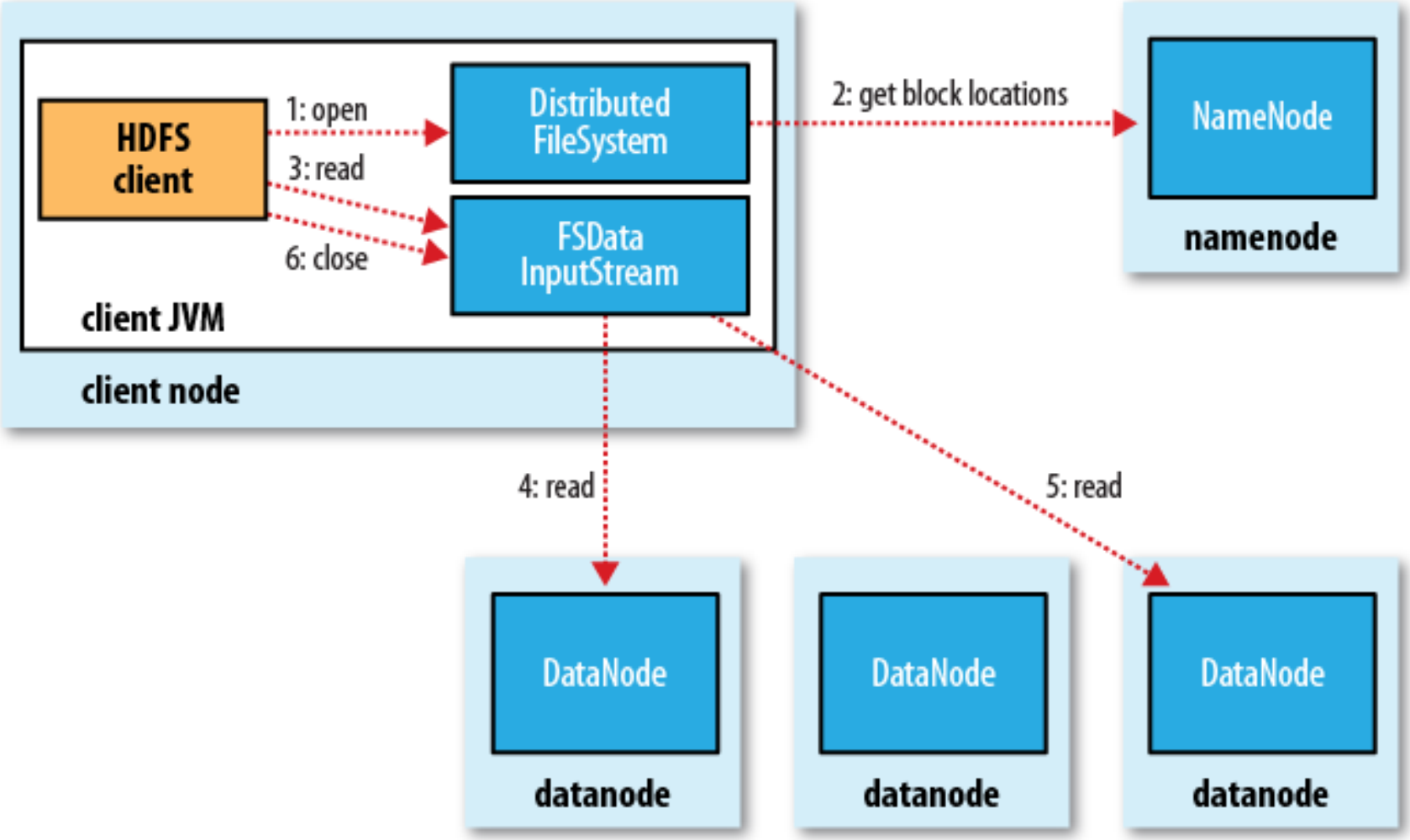
Others

- **Secondary Namenode**
  - Not act as a namenode
  - periodically merge the namespace image with the edit log to prevent the edit log from becoming too large
  - 須大量的運算資源，通常運作於另一台主機上

Data Correctness

- 使用Checksum來驗證資料
  - Cyclic Redundancy Check (CRC32)
- 檔案建立
  - 用戶端每隔 512 個位元組就會計算Checksum
  - DataNode 儲存Checksum的資訊
- 檔案存取時
  - 用戶端會同時擷取資料與Checksum
  - 如果驗證失敗，用戶端會嘗試使用其他複本

Data Flow - Read

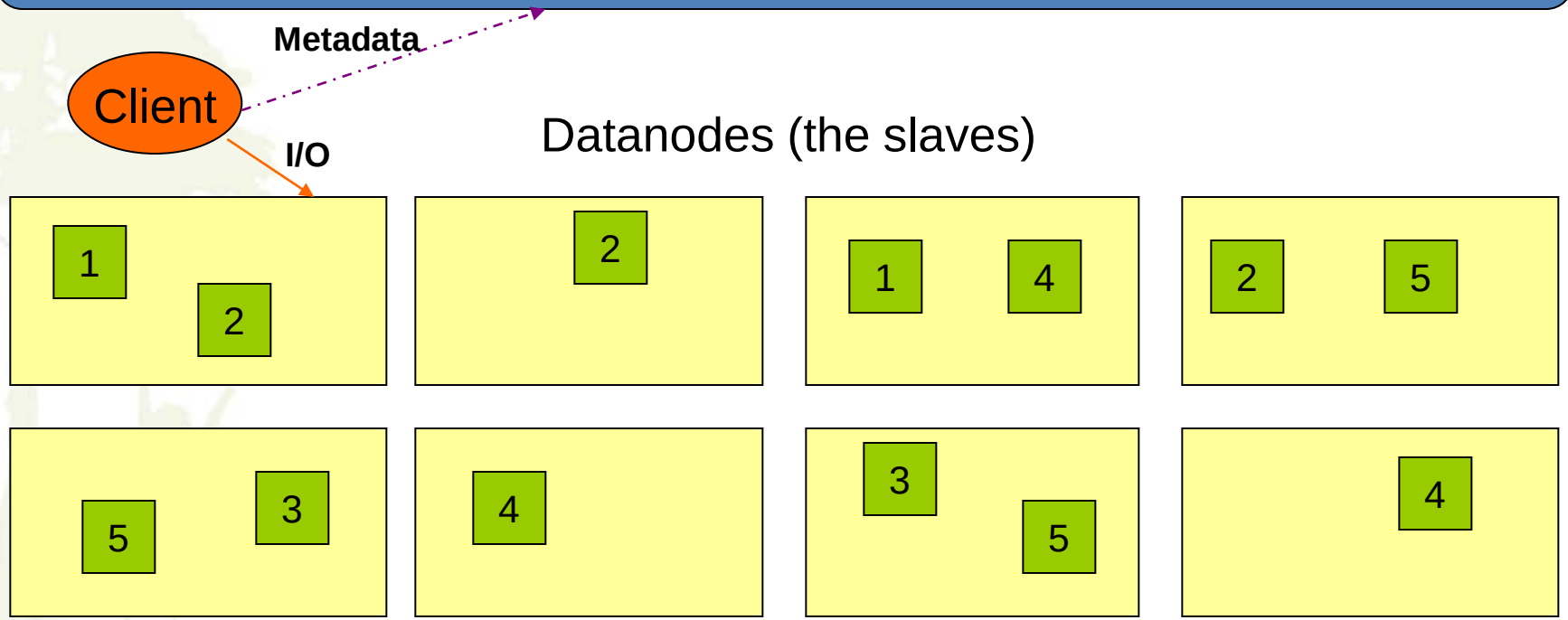


HDFS 運作

Namenode (the master)

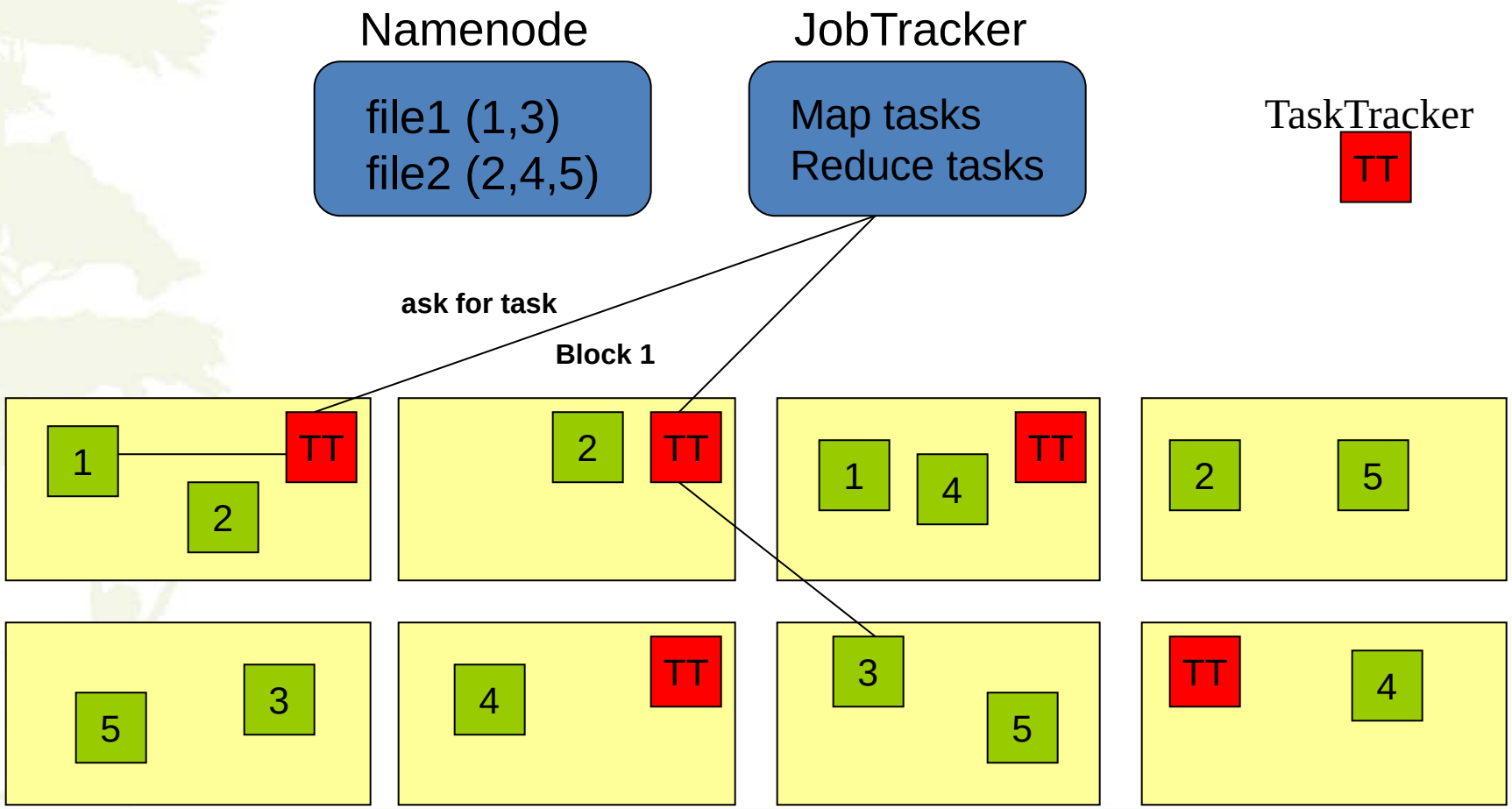
檔案路徑－ 副本數 , 由哪幾個block組成

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

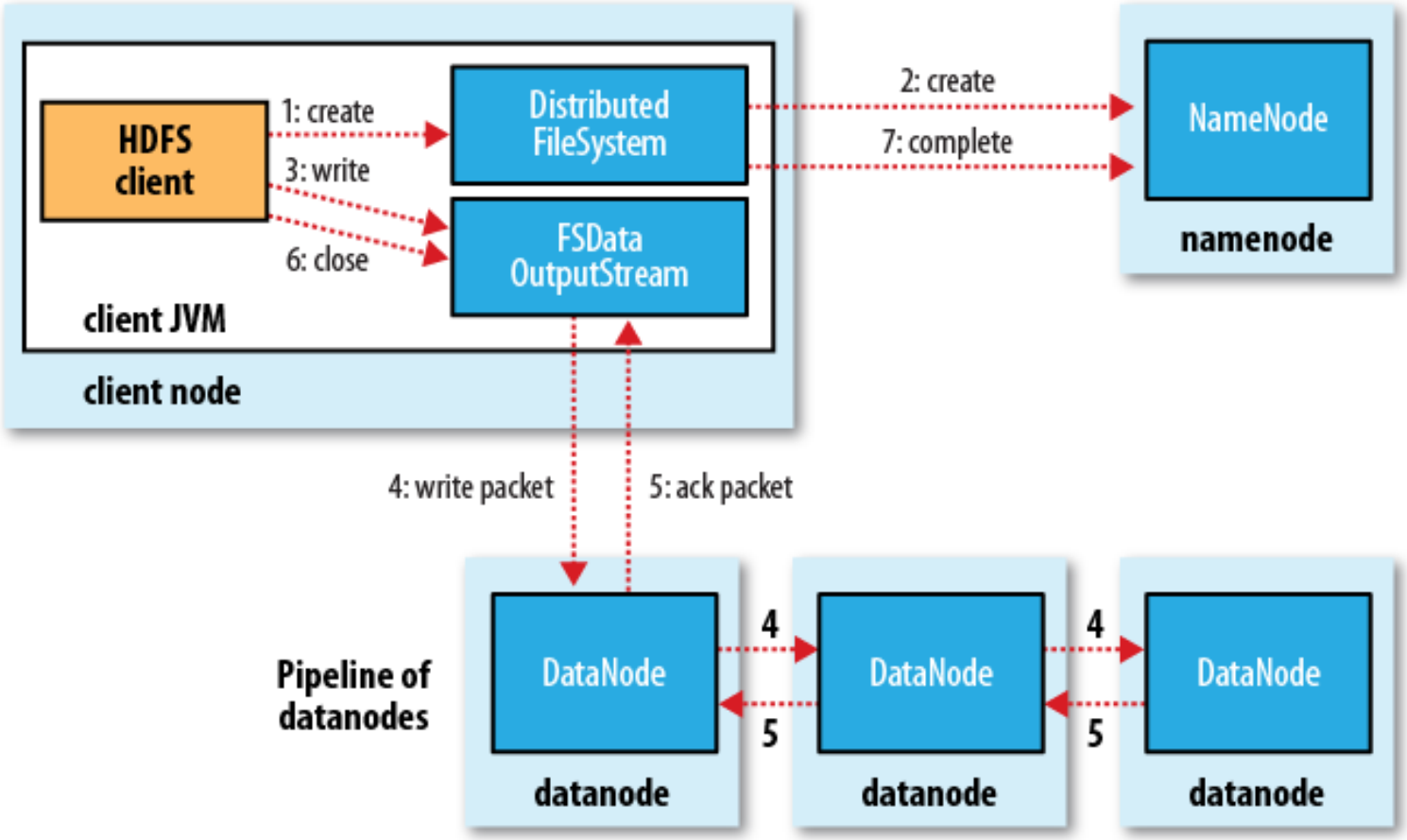


HDFS 運作

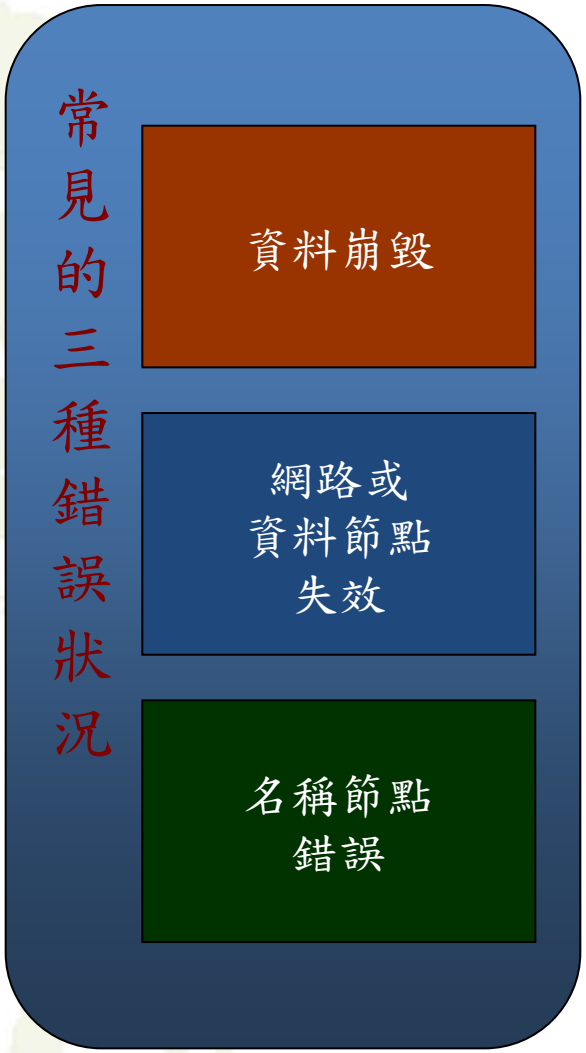
- 目的：提高系統的可靠性與讀取的效率
  - 可靠性：節點失效時讀取副本已維持正常運作
  - 讀取效率：分散讀取流量（但增加寫入時效能瓶頸）



Data Flow - Write



可靠性機制



- 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

一致性與效能機制

- **檔案一致性機制**
  - 刪除檔案＼新增寫入檔案＼讀取檔案皆由Namenode負責
- **巨量空間及效能機制**
  - 以Block為單位： 64M為單位
  - 在HDFS上得檔案有可能大過一顆磁碟
  - 大區塊可提高存取效率
  - 區塊均勻散佈各節點以分散讀取流量

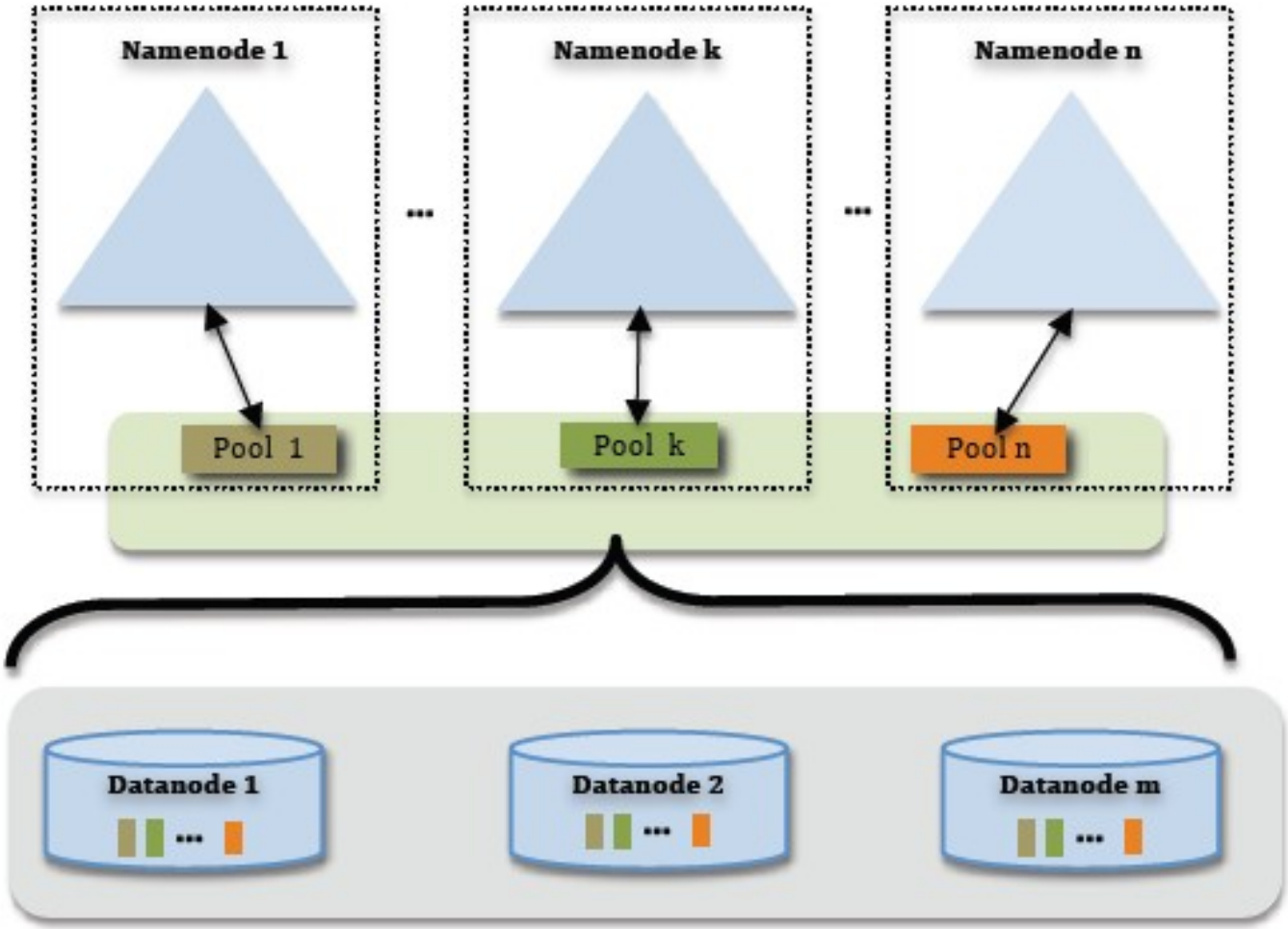
Rebalancer

- **Goal: % disk full on DataNodes should be similar**
  - Usually run when new DataNodes are added
  - Cluster is online when Rebalancer is active
  - Rebalancer is throttled to avoid network congestion
- **Disadvantages**
  - Does not rebalance based on access patterns or load
  - No support for automatic handling of hotspots of data

HDFS Federation(v2.x)

- 解決單一Namenode的問題
  - 拓展性
  - 性能
    - 一個namenode約可support 60k tasks
    - 未來Hadoop將可支援超過100k tasks
  - 隔離性
    - 不同的Group使用相同權限，共享同一個檔案系統

HDFS Federation(v2.x)



User Interface

- **API**
  - Java API
  - C language wrapper for the Java API is also avaiable
- **POSIX like command**
  - hadoop dfs -mkdir /foodir
  - hadoop dfs -cat /foodir/myfile.txt
  - hadoop dfs -rm /foodir myfile.txthadoop dfs -rm /foodir myfile.txt
- **DFSAdmin**
  - bin/hadoop dfsadmin –safemode
  - bin/hadoop dfsadmin –report
  - bin/hadoop dfsadmin -refreshNodes
- **Web管理介面**
  - http://host:port/dfshealth.jsp

POSIX Like

```
hadoop fs [-fs <local | file system URI>] [-conf <configuration file>]
        [-D <property=value>] [-ls <path>] [-lsr <path>] [-du <path>]
        [-dus <path>] [-mv <src> <dst>] [-cp <src> <dst>] [-rm <src>]
        [-rmr <src>] [-put <localsrc> <dst>] [-copyFromLocal <localsrc> <dst>]
        [-moveFromLocal <localsrc> <dst>] [-get <src> <localdst>]
        [-getmerge <src> <localdst> [addnl]] [-cat <src>]
        [-copyToLocal <src><localdst>] [-moveToLocal <src> <localdst>]
        [-mkdir <path>] [-report] [-setrep [-R] [-w] <rep> <path/file>]
        [-touchz <path>] [-test -[ezd] <path>] [-stat [format] <path>]
        [-tail [-f] <path>] [-text <path>]
        [-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
        [-chown [-R] [OWNER][:[GROUP]] PATH...]
        [-chgrp [-R] GROUP PATH...]
        [-help [cmd]]
```

Web管理介面

NameNode 'localhost:9000'

Started:	Thu Sep 13 09:40:41 CST 2012
Version:	0.20.2, r911707
Compiled:	Fri Feb 19 08:07:34 UTC 2010 by chrisdo
Upgrades:	There are no upgrades in progress.

Browse the filesystem
Namenode Logs

Cluster Summary

8 files and directories, 1 blocks = 9 total. Heap Size is 15.31 MB / 966.69 MB (1%)

Configured Capacity	:	18.82 GB
DFS Used	:	36 KB
Non DFS Used	:	6.8 GB
DFS Remaining	:	12.02 GB
DFS Used%	:	0 %
DFS Remaining%	:	63.86 %
Live Nodes	:	1
Dead Nodes	:	0

MapReduce 介紹

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

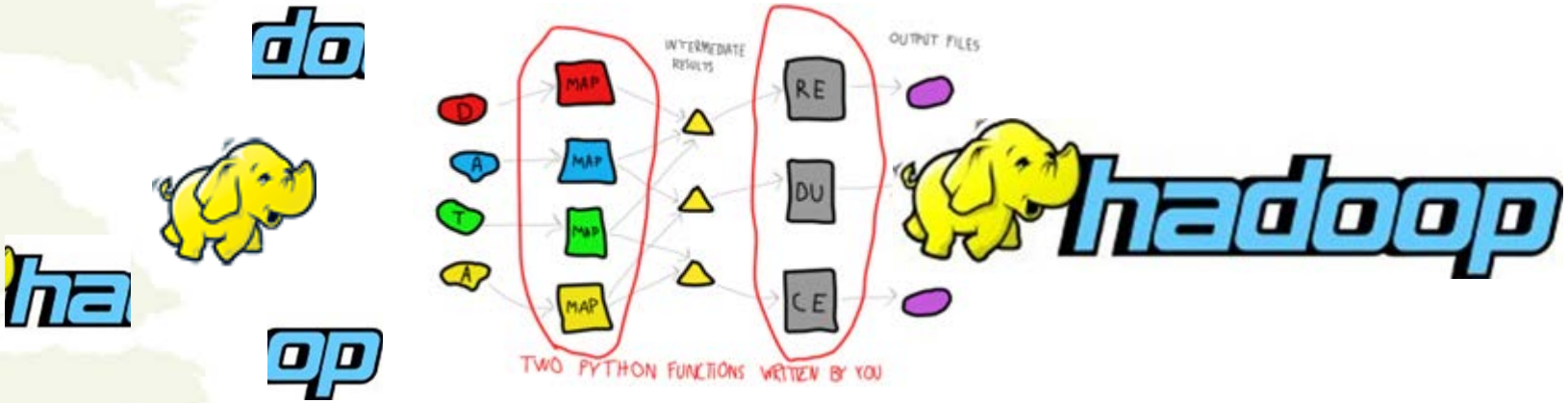
Outline

- 概念
  - Divide and Conquer
- **MapReduce起源**
- **MapReduce 運作流程**
- 資料分析範例 **– Using MapReduce**
- **MapReduce Works Detial**

Map Reduce 起源

- 演算法(Algorithms)
  - Divide and Conquer
  - 分而治之
- 在程式設計的軟體架構內，適合使用在大規模數據的運算中

Hadoop MapReduce 定義

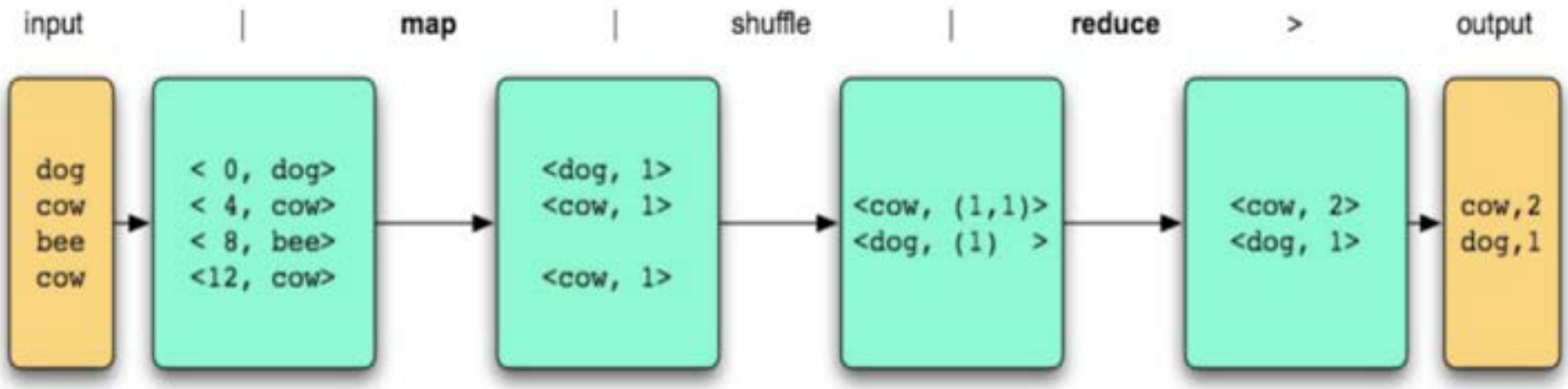


Hadoop Map/Reduce是一個易於使用的軟體平台，以MapReduce為基礎的應用程序，能夠運作在由上千台PC所組成的大型叢集上，並以一種可靠容錯的方式平行處理上Peta-Bytes數量級的資料集。

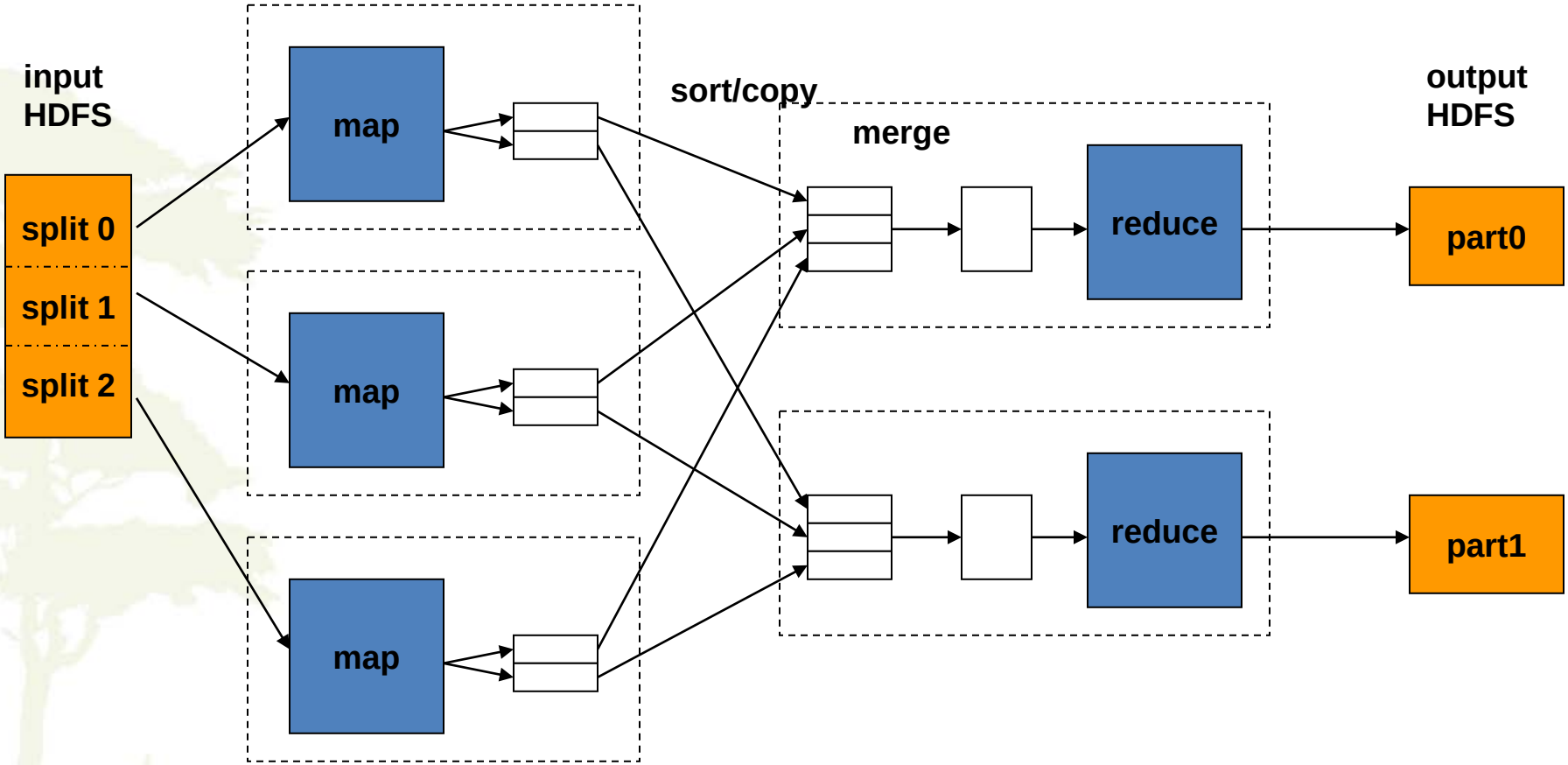
適用範圍

- 大規模資料集
- 可拆解
  - Text tokenization
  - Indexing and Search
  - Data mining
  - machine learning
  - …

MapReduce Overview



MapReduce Overview



JobTracker跟NameNode取得需要運算的blocks

JobTracker選數個TaskTracker來作Map運算，產生些中間檔案

JobTracker將中間檔案整合排序後，複製到需要的TaskTracker去

JobTracker派遣TaskTracker作reduce

reduce完後通知JobTracker與Namenode以產生output

Map

- 輸入是”一個” key/value 序對 ，輸出則為另”一組”intermediate key/value 序對組。

Example: Upper-case Mapper

```
let map(k, v) =
  emit(k.toUpper(), v.toUpper())

(“foo”, “bar”) ➔ (“FOO”, “BAR”)
(“Foo”, “other”) ➔ (“FOO”, “OTHER”)
(“key2”, “data”) ➔ (“KEY2”, “DATA”)
```

Example: Explode Mapper

```
let map(k, v) =
  foreach char c in v:
    emit(k, c)
```

(“A”, “cats”) ➔ (“A”, “c”), (“A”, “a”),
(“A”, “t”), (“A”, “s”)

(“B”, “hi”) ➔ (“B”, “h”), (“B”, “i”)

Example: Filter Mapper

```
let map(k, v) =
  if (isPrime(v)) then emit(k, v)
```

(“foo”, 7) ➔ (“foo”, 7)

(“test”, 10) ➔ (nothing)

Example: Changing Keyspaces

```
let map(k, v) = emit(v.length(), v)
```

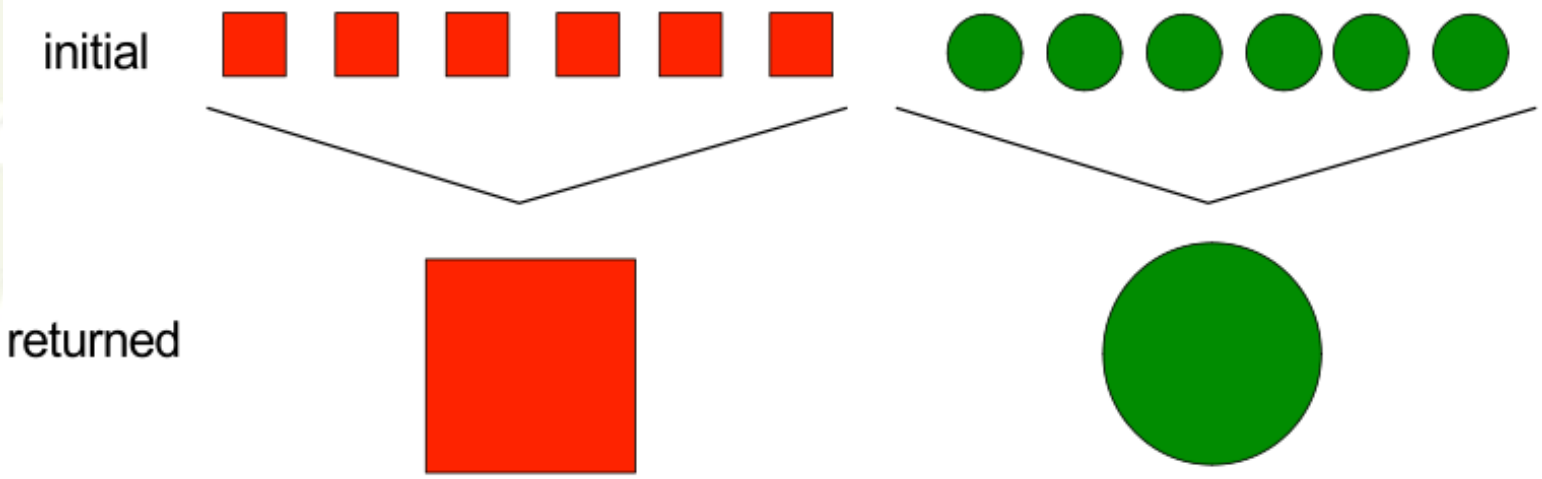
(“hi”, “test”) ➔ (4, “test”)
(“x”, “quux”) ➔ (4, “quux”)
(“y”, “abracadabra”) ➔ (10, “abracadabra”)

Reduce

- 數負責針對相同的 intermediate key 合併其所有相關聯的intermediate values。並產生輸出結果的key/value序對

Reduce

```
reduce (inter_key, inter_value list) ->
      (out_key, out_value) list
```



Reduce

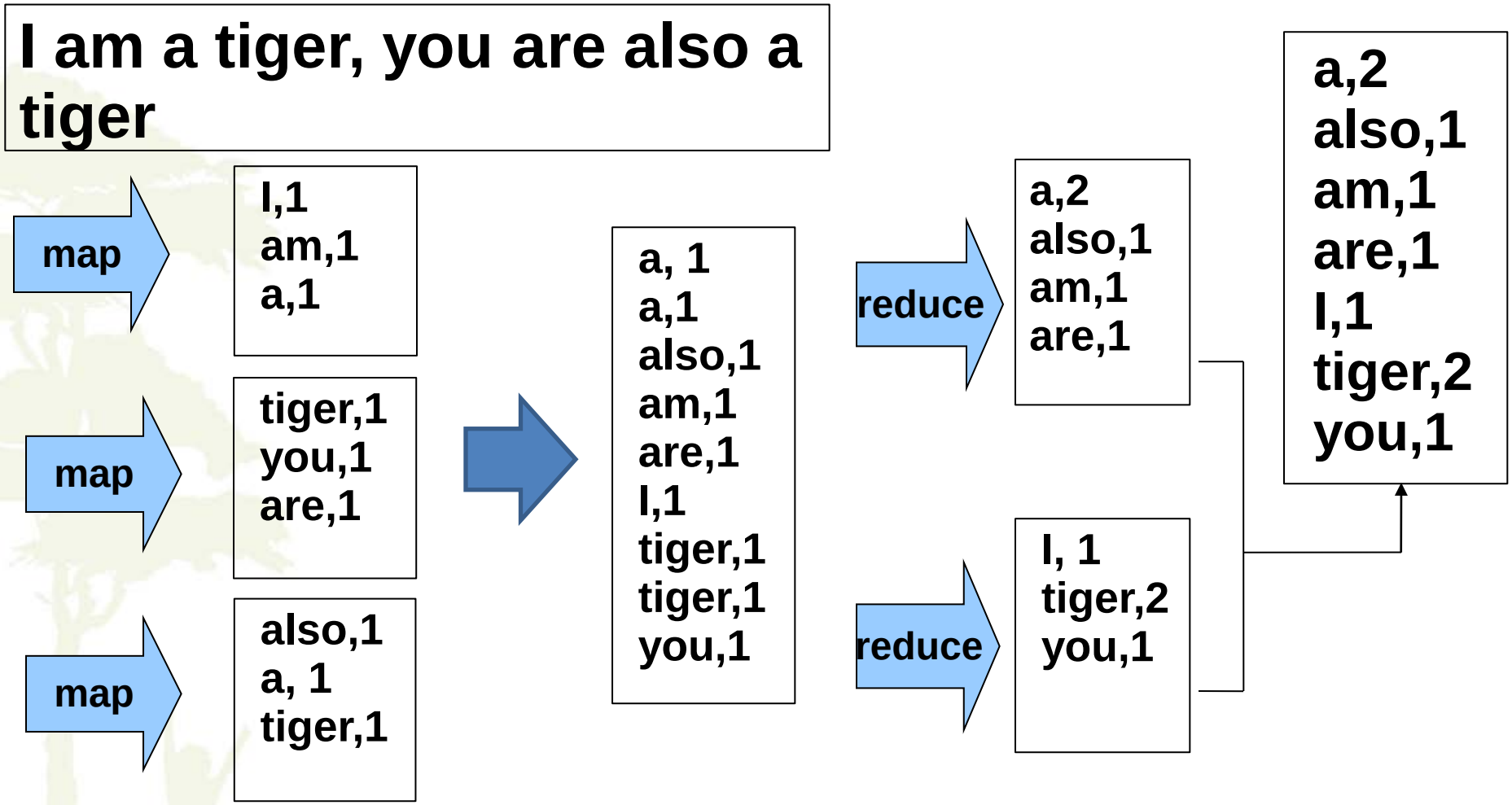
- **Example: sum reducer**

```
let reduce(k, vals)=
sum = 0
foreach int v in vals:
sum +=v
emit(k,sum)
```

("A", [42, 100, 312]) → ("A", 454)

("B", [12, 6, -2]) → ("B", 16)

MapReduce範例



JobTracker先選了三個Tracker做map

Map結束後，hadoop進行中間資料的整理與排序

JobTracker再選兩個TaskTracker作reduce

回到以美國氣象資料為例：

0057
332130 # USAF weather station identifier
99999 # WBAN weather station identifier
19500101 # observation date
0300 # observation time
4
+51317 # latitude (degrees x 1000)
+028783 # longitude (degrees x 1000)
FM-12
+0171 # elevation (meters)
99999
V020
320 # wind direction (degrees)
1 # quality code
N

0072
1
00450 # sky ceiling height (meters)
1 # quality code
CN
010000 # visibility distance (meters)
1 # quality code
N9
-0128 # air temperature (degrees Celsius x 10)
1 # quality code
-0139 # dew point temperature (degrees Celsius x 10)
1 # quality code
10268 # atmospheric pressure (hectopascals x 10)
1 # quality code

資料分析 – MapReduce

- 原始資料

0067011990999991950051507004...9999999N9+00001+99999999999...
0043011990999991950051512004...9999999N9+00221+99999999999...
0043011990999991950051518004...9999999N9-00111+99999999999...
0043012650999991949032412004...0500001N9+01111+99999999999...
0043012650999991949032418004...0500001N9+00781+99999999999...

- **Map – 擷取所需的資訊**

(0, 0067011990999991**950**051507004...9999999N9+**0000**1+99999999999...)
(106, 0043011990999991**950**051512004...9999999N9+**0022**1+99999999999...)
(212, 0043011990999991**950**051518004...9999999N9-**0011**1+99999999999...)
(318, 0043012650999991**949**032412004...0500001N9+**0111**1+99999999999...)
(424, 0043012650999991**949**032418004...0500001N9+**0078**1+99999999999...)

資料分析 – MapReduce

- **Map - shuffle**

(1950, 0)
(1950, 22)
(1950, −11)
(1949, 111)
(1949, 78)

- **Map - Output:<Key, Value>**

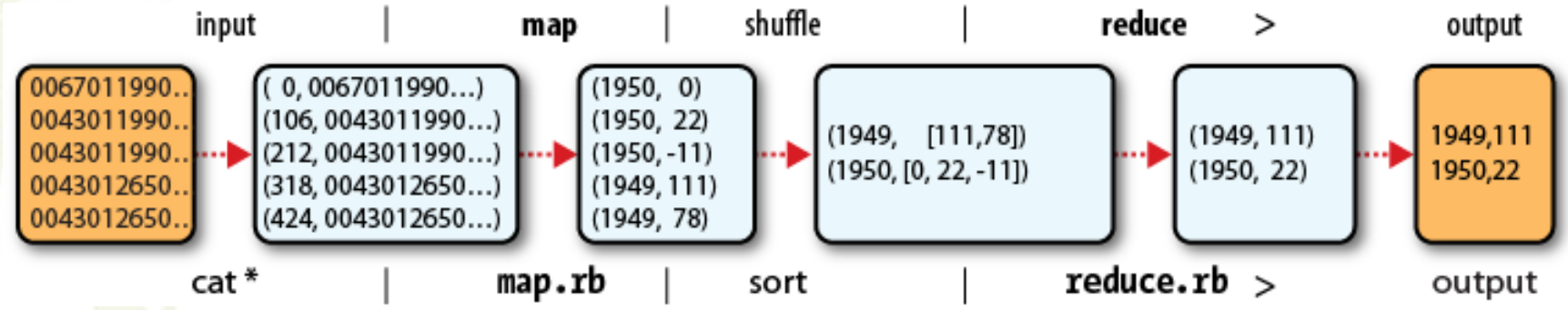
(1949, [111, 78])
(1950, [0, 22, −11])

資料分析 – MapReduce

- **Reduce - Output: <Year, Max>**

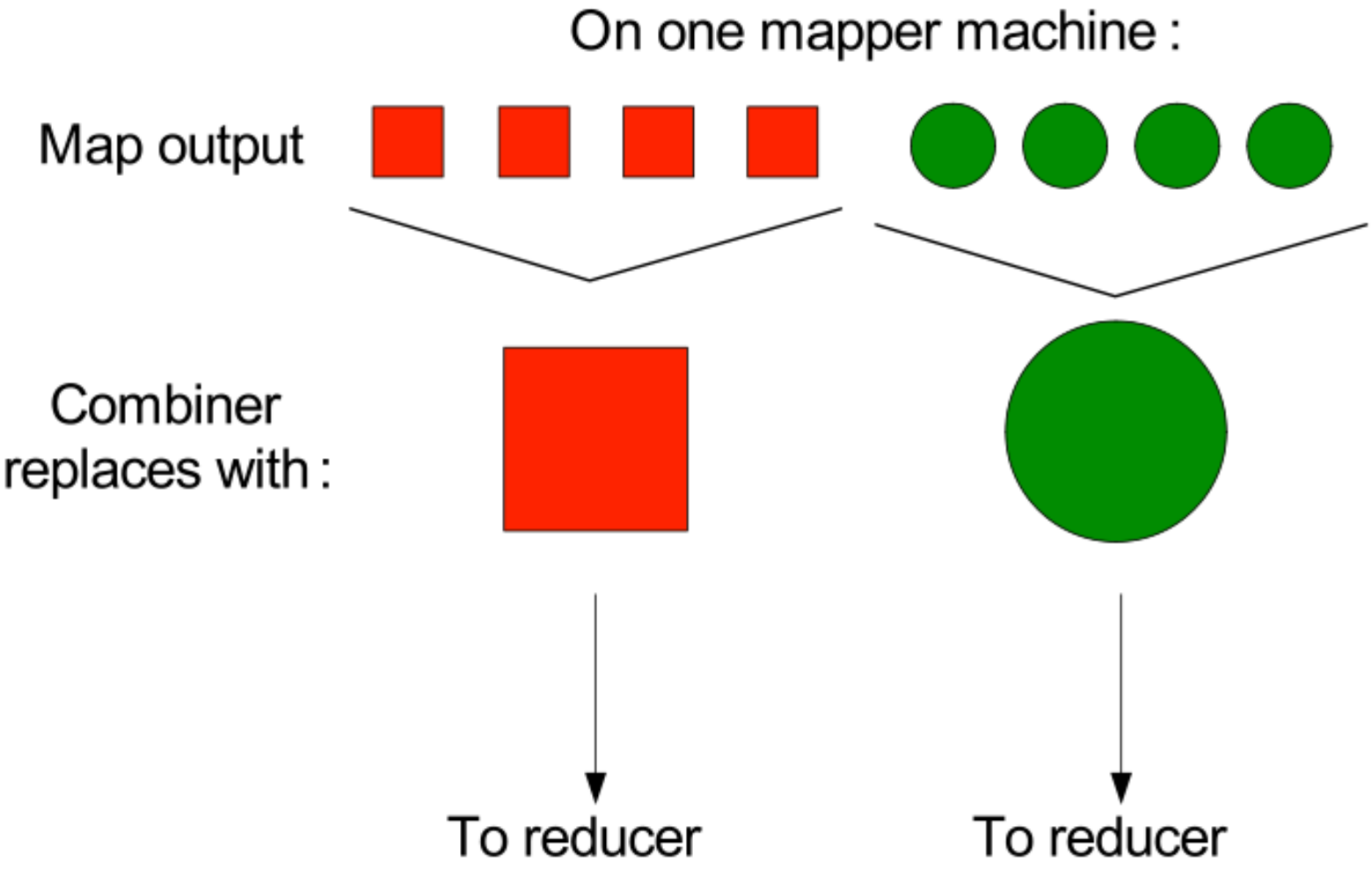
  (1949, 111)
  (1950, 22)

- **Logical Flow**



- ***費時：6 min on EC2 High-CPU Extra Large Instance***

Combiner


On one mapper machine :
Map output
Combiner
replaces with :
To reducer
To reducer

以 ICAS 提供雲內網路入侵偵測日誌分析

楊順發

國家高速網路與計算中心

- **Goal**
  - IDS Log Cloud Analysis System (called ICAS)
  - Provide fast and high reliability of the system.

RELATED WORKS

- **Alert Correlation**
- **Existing IDS Types**
- **Hadoop**

ALERT CORRELATION

- **Alert correlation is an analysis process that takes the alerts generated by IDS and creates reports under its surveillance network.**

EXISTING IDS TYPES

- **NIDS: To monitor network spigot and to detection exceptionally transmission behavior by connecting to network hubs or network switch**
- **HIDS: This is inseparable from operating system and to detect and monitor malicious actives such as system calls, file system changes, application logs.**

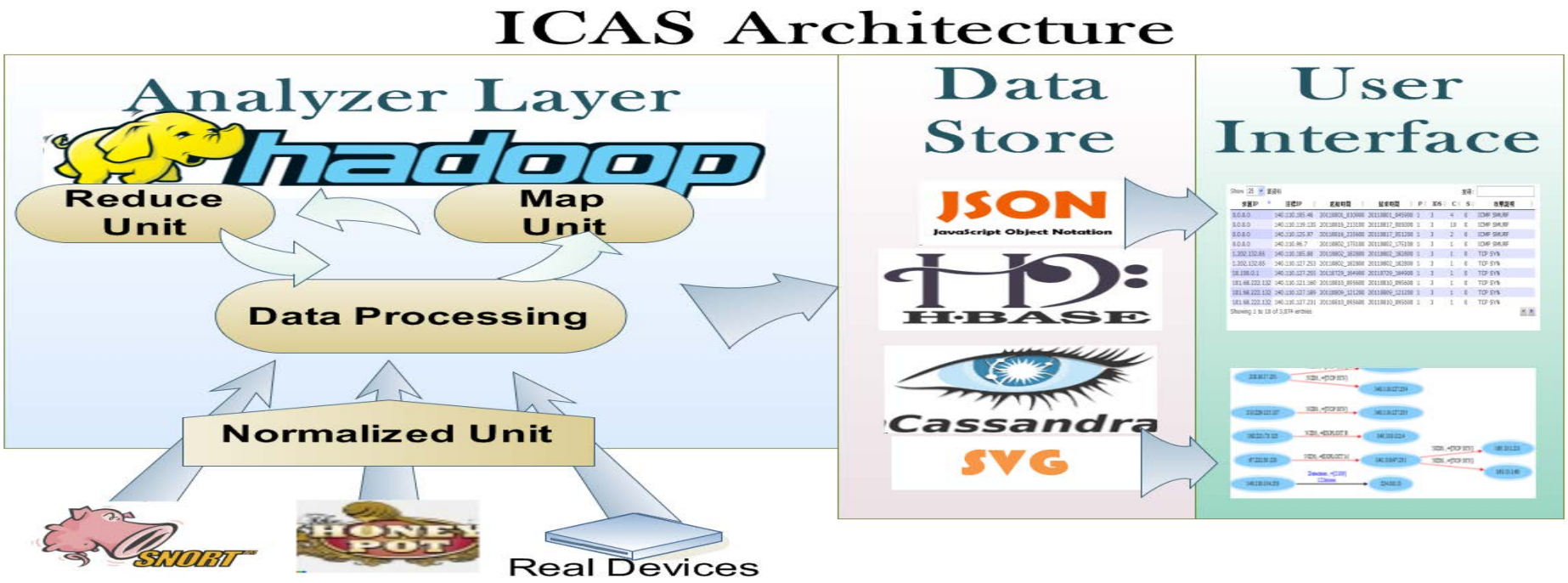
HADOOP

- **Hadoop was inspired by Google's MapReduce and Google File System (GFS) papers.**
- **Hadoop is a powerful computing platforms, which is used to process large-scale computation and includes distributed file system.**
- **Enables applications to work with thousands of nodes and petabytes of data.**

ALERT INTEGRATION PROCEDURE

- **System Architecture**
- **Analysis Format**
  - Input Format and Output Format
- **Integrating IDS into Cloud Computing Platforms**

SYSTEM ARCHITECTURE



ANALYSIS FORMAT

- **Snort**

  [**] [1:2189:3] BAD-TRAFFIC IP Proto 103 PIM [**]
  [Classification: Detection of a non-standard protocol or event] [Priority: 2]
  05/17-08:30:14.750704 140.110.138.253 -> 224.0.0.13
  PIM TTL:1 TOS:0xC0 ID:4076 IpLen:20 DgmLen:58
  [Xref => http://cve.mitre.org/cgi-bin/cvename.cgi?name=2003-0567][Xref => http://www.securityfocus.com/bid/8211]

- **IDP8200**

  Time Received ## Src Addr ## Dst Addr ## Action ## Protocol ## Dst Port ## interface ## Description ## Severity
  2003/8/11 13:05,140.113.130.221,0.0.0.0,Accepted,TCP,65432,'interface=eth2',FTP: Format String in Command,Major

- **NK7Admin**

  NO. ## name ## from(address) ## to(address) ## start time ## total ## from(port) ## to(port)
  1,TCP SYN,60.173.26.116,140.110.127.253,2011/3/1 14:41,1,6000,9415

INPUT FORMAT AND OUTPUT FORMAT

- **Output as <Key, Value> Pair**

Key:

- <src_ip – dst_ip>

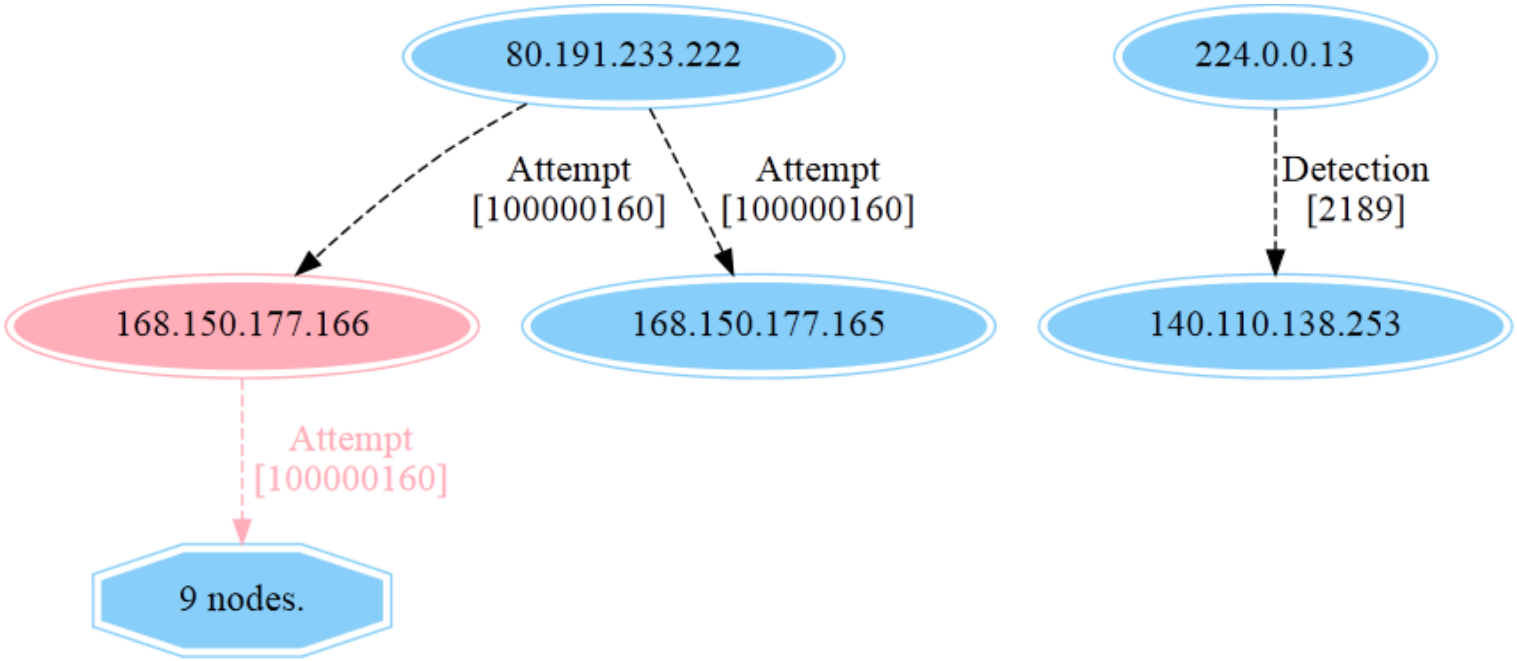
Value:

- <date @@ time @@ class_id @@ ids @@ s-id @@ priority @@ port @@ description>

- **ICAS will count these attack record as < *IDS Source; Alert Identification; Attack; Class; Severity(1~3); Date; Time; from(IP Addr.); to(IP Addr.); Port NO.>***

SYSTEM VISUALIZATION DEMO

- **The Graph of ICAS Output**



DEMO WEBSITE

HTTP://CRAWLWEB2.NCHC.ORG.TW/ICAS-EN/INDEX.PHP

Explanation

Warning List

IDS means Intrusion Detection System, support as follows:

* snort = 1
* IDP8200 = 2
* NK7Admin = 3

P means priority, 1~3 , 1 is the most serious.

C means merge count

S means Snort Signature ID, for more detial please see http://www.snortid.com/ps : only Snort support this function

List

Show 25 筆資料 搜尋:

來源IP	目標IP	起始時間	結束時間	P	IDS	C	S	攻擊說明
0.0.0.0	140.110.105.46	20110801_010900	20110801_045900	1	3	4	0	ICMP SMURF
0.0.0.0	140.110.96.7	20110802_175100	20110802_175100	1	3	1	0	ICMP SMURF
1.202.132.85	140.110.105.80	20110802_182800	20110802_182800	1	3	1	0	TCP SYN
1.202.132.85	140.110.127.253	20110802_182800	20110802_182800	1	3	1	0	TCP SYN
10.100.0.1	140.110.127.255	20110729_164900	20110729_164900	1	3	1	0	TCP SYN
101.68.222.132	140.110.127.189	20110809_121200	20110809_121200	1	3	1	0	TCP SYN
108.59.3.28	140.110.117.141	20110729_131500	20110802_120600	1	3	2	0	VULN MS Windows SChannel Security Remote Code Execution
108.61.17.196	140.110.102.27	20110809_081700	20110809_081700	1	3	1	0	VULN MS Windows SChannel Security Remote Code Execution
109.0.0.0	140.110.111.125	20110730_132800	20110730_132800	1	3	3	0	ICMP SMURF

Hadoop應用
- Nutch 簡介

王耀聰 楊順發
jazz@nchc.org.tw
shunfa@nchc.org.tw
國家高速網路與計算中心(NCHC)

Outline

- What is Nutch
- Why Nutch
- Nutch’s Details
- Let’s go

What's Nutch

- Nutch是一個open source，以Java來實做的搜索引擎，它提供了架設自己的搜索引擎所需的全部工具。
- 利用Lucene為函式庫
- **架構於Hadoop之上**

Nutch’s goals

- 每個月抓取幾十億網頁
- 為這些網頁維護索引
- 對索引文件進行每秒上千次的搜索
- 提供高質量的搜索結果
- 以最小的成本運作

Why Nutch ?

- 透明
  - Opensource，資訊不隱藏
- 擴充
  - 有各種函式庫應用於分析不同檔案
- 隱私
  - 可應用於搜尋專屬資料
- 客製化
  - 可以之為基礎設計自己的data mining 工具

Who use Nutch

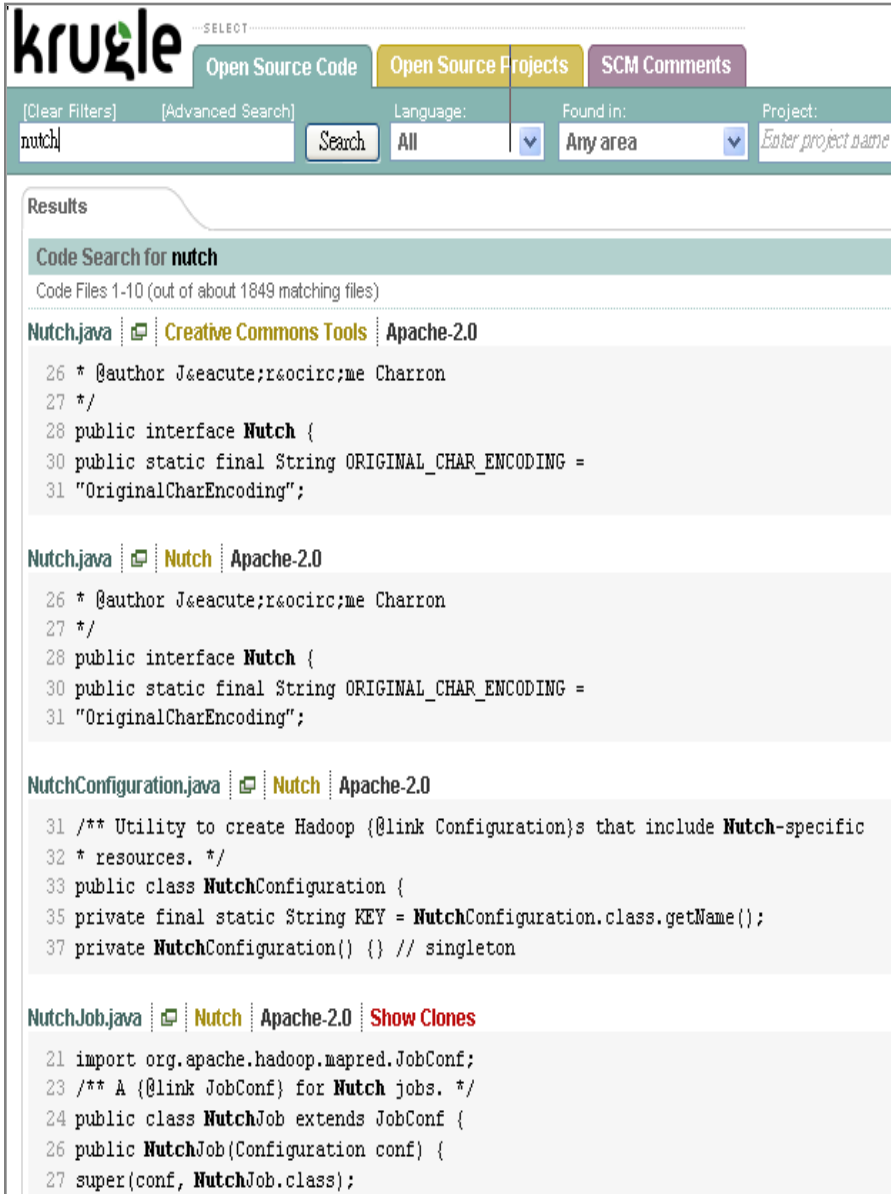
Public search engines using Nutch

Please sort by name alphabetically

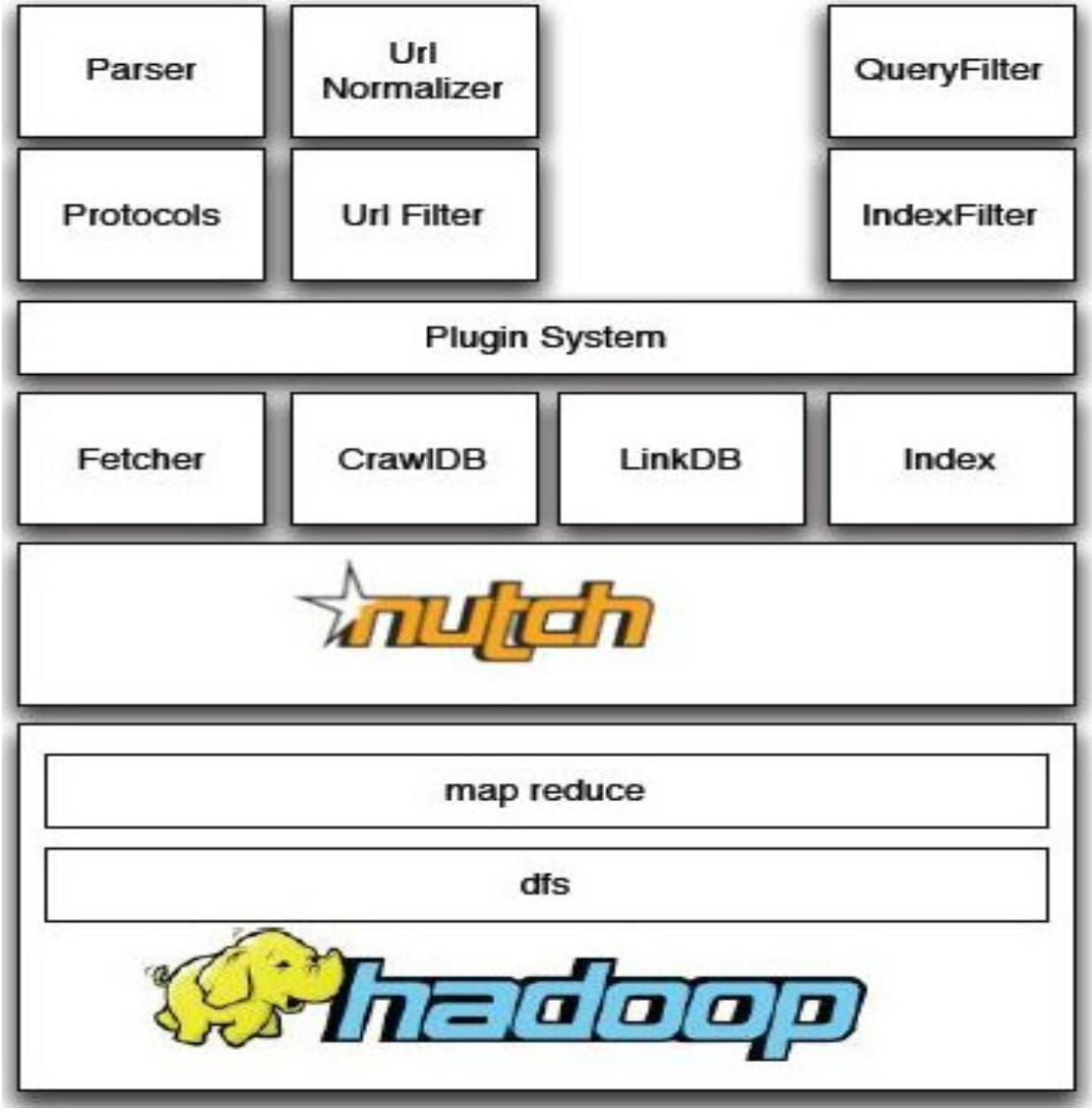
- AskAboutOil is a vertical search portal for the petroleum indust
- Baynote provides free hosted Nutch search for businesses.
- BeThere BeSquare is an Event Search Engine for the San Fra category and get details about events in 4 different views.
- Bigsearch.ca uses nutch open source software to deliver its se
- BusyTonight: Search for any event in the United States, by key from original source Web sites.
- Central Budapest Search is a search engine for English langu events.
- Circuit Scout is a search engine for electrical circuits.
- Comtec Search is a search engine for UK Tour Operator Pack
- Coder-Suche.de searchs for coding stuff like apis, documenta english.
- Cornell University Library is collaborating with the research gro pages based on Nutch. The nutch-based search engine is near t
- Creative Commons is a search engine for creative commons l
- Dadi360 Usee nutch search engine for providing search of Ch
- Ecolhub Web Search an E. coli specific search engine based thereby reducing the number of spurious hits. Searches can be o More resources getting added.
- Epivista is a search engine of epilepsy related web sites.
- eroscanner is a search engine for german adult stuff. Watching **NSFW**)

…..more

(http://wiki.apache.org/nutch/PublicServers)

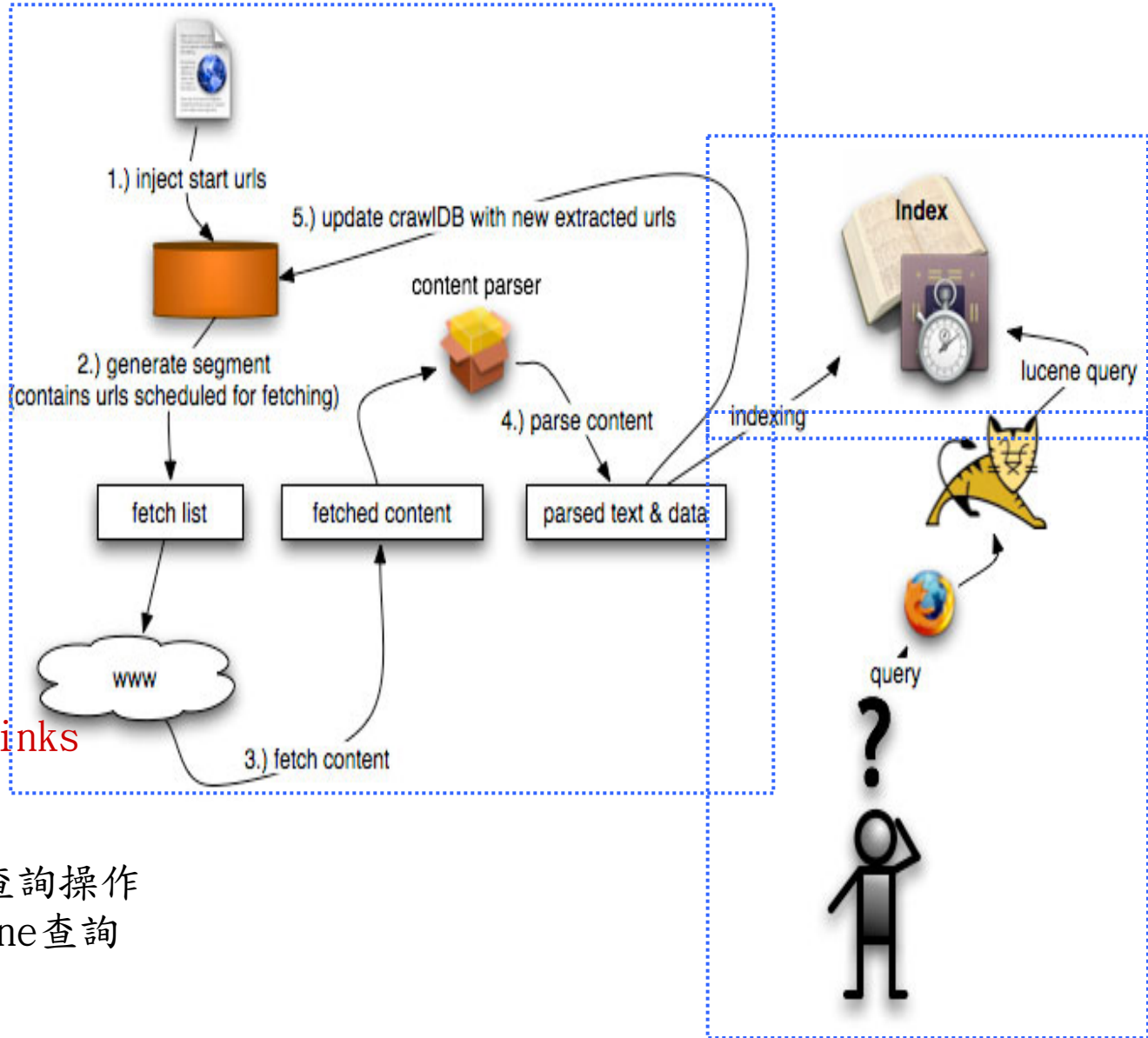


架構


Parser
Url Normalizer
QueryFilter
Protocols
Url Filter
IndexFilter
Plugin System
Fetcher
CrawlDB
LinkDB
Index
nutch
map reduce
dfs
hadoop

運作流程

1）建立初始URL集
2）將URL集注入crawldb---inject
3）根據crawldb建立抓取清單---generate
4）執行抓取，獲取網頁內容---fetch
5)用獲取到的頁面資訊更新crawlDB---updatedb
6）重複進行3～5的步驟，直到預先設定的抓取深度



7）更新linkdb ---invertlinks
8）建立索引---index

9）用戶通過用戶接口進行查詢操作
10）將用戶查詢轉化為lucene查詢
11）返回結果

Plugin

- 修改 conf/nutch-site.xml的plugin.includes屬性
- 在nutch基本功能之上擴充其功能
  - “parse-xx”：加入解析xx檔案類型的能力
  - “protocol -xx”：加入在此協定內的檔案也處理

parse-text
parse-ext
parse-html
parse-js
parse-mp3
parse-zip
parse-rtf

parse-msword
parse-msexcel
parse-pdf
parse-rss
parse-oo
parse-swf
parse-mspowerpoint

protocol-file
protocol-ftp
protocol-http
protocol-httpclient

http://wiki.apache.org/nutch/PluginCentral

International

- 已有多國語言版可選，但若還要客製化…
- the page header
  - src/web/include/*language*/header.xml
- the "about" page
  - src/web/pages/*lang*/about.xml
- the "search" page
  - src/web/pages/*lang*/search.xml
- the "help" page
  - src/web/pages/*lang*/help.xml
- text for search results
  - src/web/locale/org/nutch/jsp/search_*lang*.properties

No！ Nutch

- 告訴網頁機器人是否允許進入爬網
- 將robots.txt放在web上
- robots.txt

```
User-agent: Nutch
Disallow: /
```

Home Page


nutch
About
FAQ
Search
help
nutch
powered by
ca | de | en | es | fi | fr | hu | it | jp | ms | nl | pl | pt | sh | sr | sv | th | zh

References..

- Nutch Website
  - http://lucene.apache.org/nutch/
- Nutch wiki
  - http://wiki.apache.org/nutch/
- Nutch API
  - http://lucene.apache.org/nutch/apidocs-1.0/index.html

Start

- **23 March 2009 - Apache Nutch 1.0 Released**

Let’s Go

Stepssssssssssssss!

前言
環境
step 1 安裝好Hadoop叢集
step 2 下載與安裝
  2.1 下載 nutch 並解壓縮
  2.2 部屬hadoop,nutch目錄結構
step 3 編輯設定檔
  3.1 hadoop-env.sh
  3.2 hadoop-site.xml
  3.3 nutch-site.xml
  3.4 slaves
  3.5 crawl-urlfilter.txt
  3.6 regex-urlfilter.txt
  3.7 整個移植到另一台node
step 4 執行nutch
  4.1 編輯url清單
  4.2 上傳清單到HDFS
  4.3 執行nutch crawl
step 5 瀏覽搜尋結果
  5.1 安裝tomcat
  5.1 tomcat server設定
  5.3 下載crawl結果
  5.4 設定nutch的搜尋引擎頁面到tomcat
  5.5 設定搜尋引擎內容的來源路徑
  5.6 啓動tomcat
step 6 享受結果

The Other Choose is…

Crawlzilla!!!

Hadoop應用 - Crawlzilla 搜尋引擎安裝與實作

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

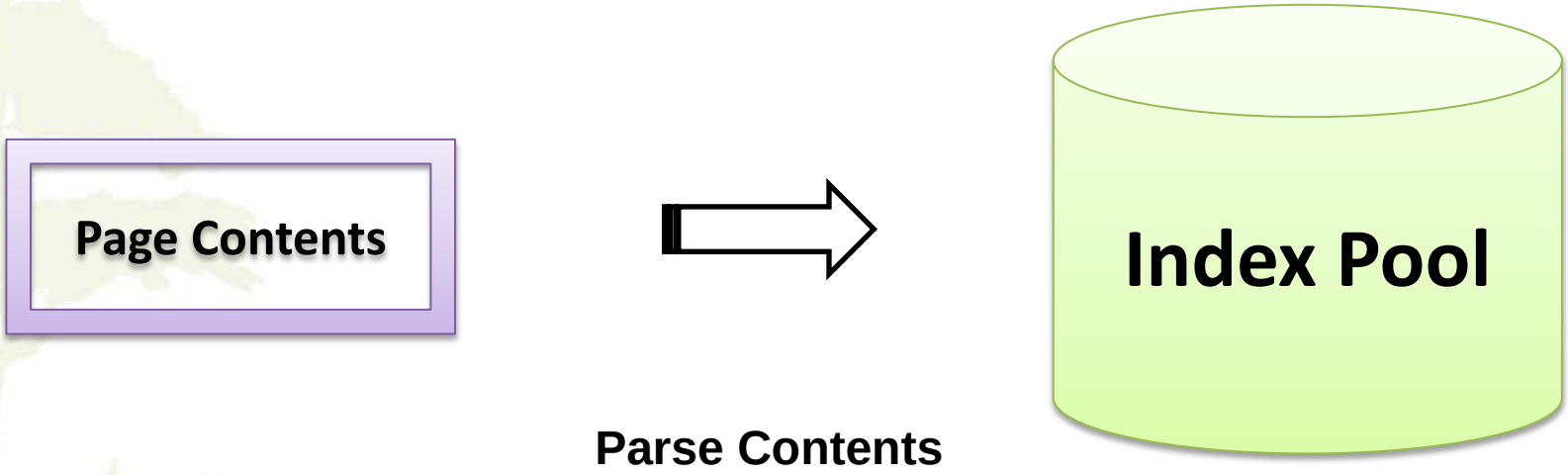
搜尋引擎運作原理 – Phase1

- **Crawling the Web**

Crawler visits the web pages of the links

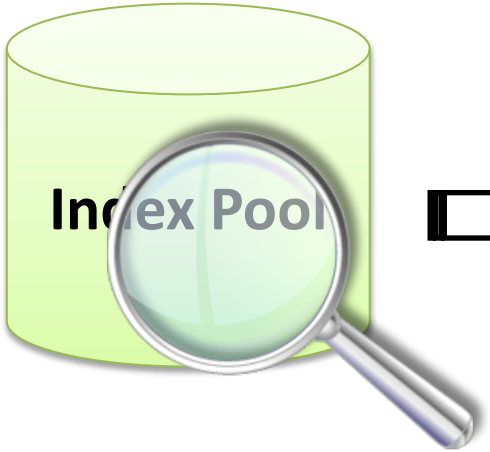
搜尋引擎運作原理 – Phase2

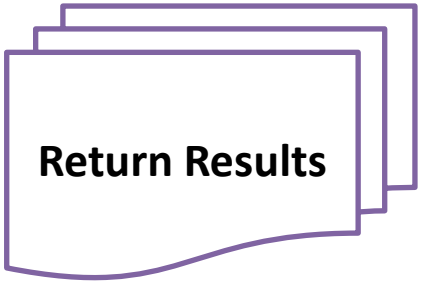
- **Building the Index Pool**



搜尋引擎運作原理 – Phase3

- **Serving Queries**





User Sent a Query **Search from Index Pool**

What is Crawlzilla?

- Crawlzilla 簡介
  - 於2009推出實驗版
  - Crawlzilla 於2010更名並延續實驗版開發更多新功能
  - 提供簡單安裝及操作管理介面，輕鬆建立搜尋引擎的套件工具
  - 提供索引資料庫瀏覽功能，搜尋引擎資料庫資訊一目了然

Why Crawlzilla?

- 開放式搜尋引擎不適用於企業內部網站
- 使用Opensource建立搜尋引擎的技術門檻太高
- 叢集環境架設不易
- 使用Crawlzilla優點
  - Opensource專案，使用者可依自己的需求修改源始碼
  - 使用簡單，可輕鬆建立叢集環境
  - 友善的操作環境，節省適應系統時間
  - 支援中文分詞，提高搜尋精準度

Crawlzilla 操作介面特色

```
check_sunJava
Crawlzilla need Sun Java JDK 1.6.x or above version
System has Sun Java 1.6 above version.
System has ssh.
System has ssh Server (sshd).
System has dialog.
Welcome to use Crawlzilla, this install program will create a new accunt and to
assist you to setup the password of crawler.
Set password for crawler:
password:

keyin the password again:
password:

Master IP address is:  140.110.138.186
Master MAC address is:   08:00:27:99:4d:09
Please confirm the install infomation of above : 1.Yes 2.No
```

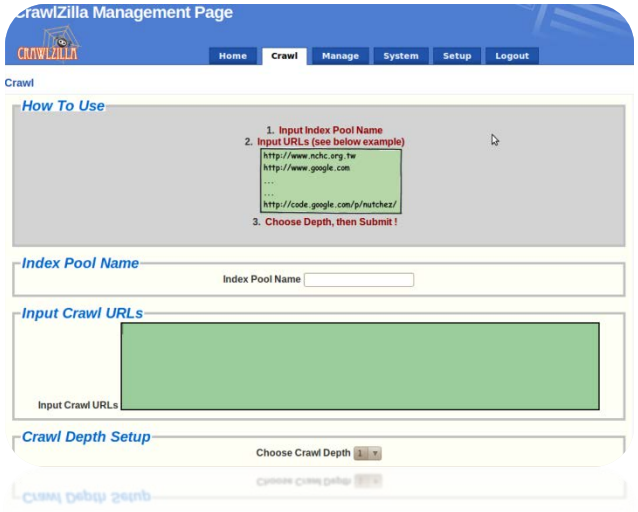
(1) Easy to Deploy Crawling Cluster Environment

```
[Management Options]
Please choose what you want to do:

cluster_status   Check cluster state
fast_manage      Fast Startup/Stop all Sevices
cluster_setup    Set datanode & tasktracker
server_setup     Set namenode & jobtracker
tomcat_switch    Startup/Shutdown/Restart Tomcat
tomcat_port      Change Tomcat port
lang_switch      Change language
client_install   Client Install Steps
exit             Exit

< 確定 >        < 取消 >
```

(2) Easy to Manage

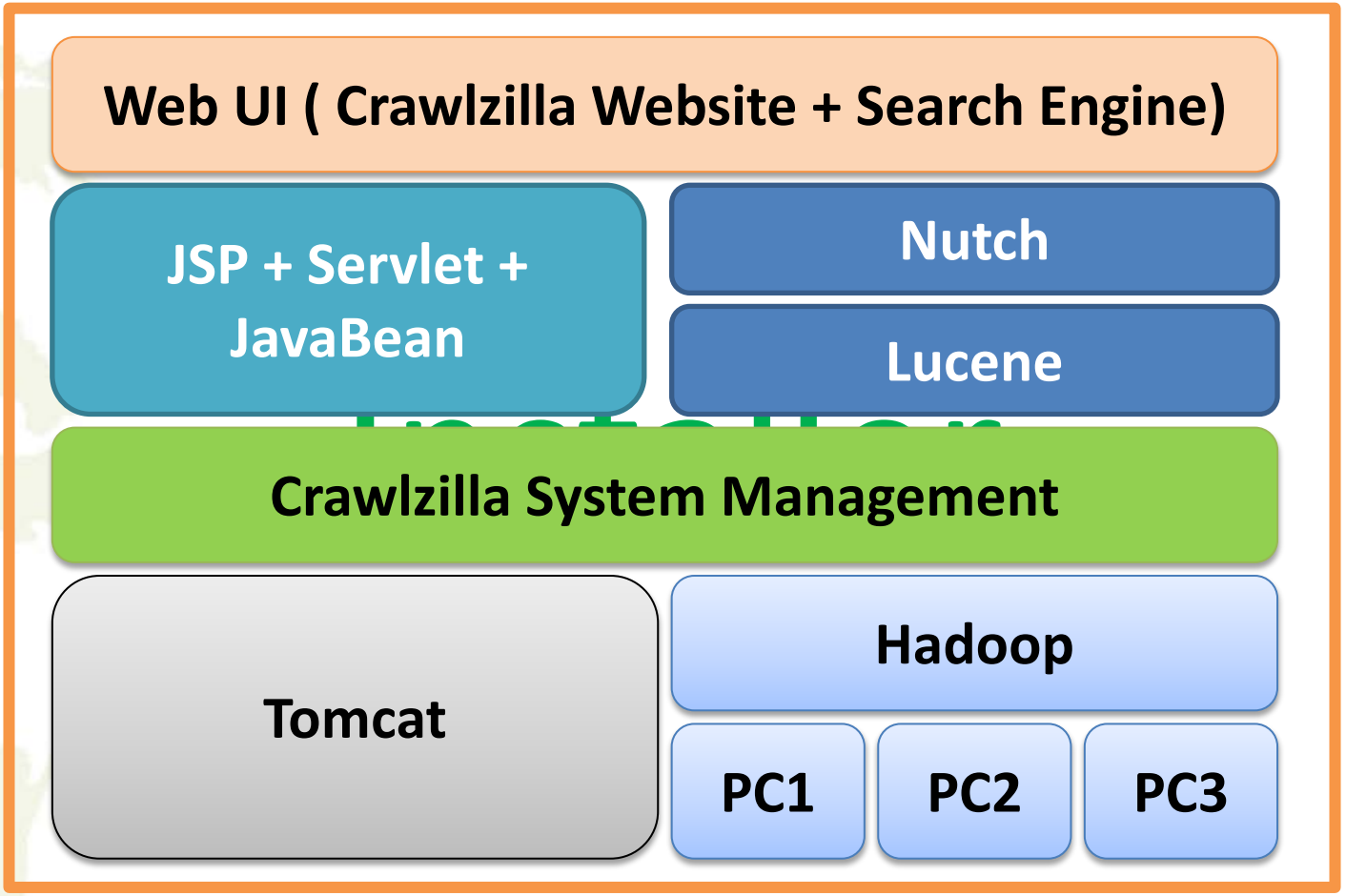


(3) Easy to Use

Crawlzilla 系統功能

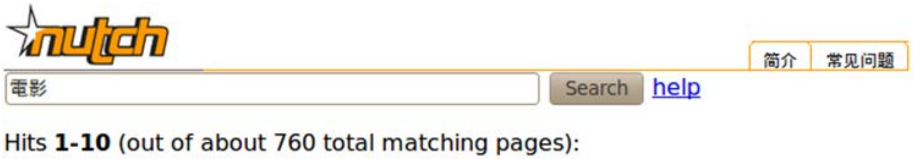
- 支援叢集運算及顧全安全性
- 支援中文分詞功能
- 支援多工網頁爬取
- 支援多重搜尋引擎
- 即時瀏覽資料庫資訊
- 解決中文亂碼及中文支援
- 支援多國語言
- 網頁管理

系統架構

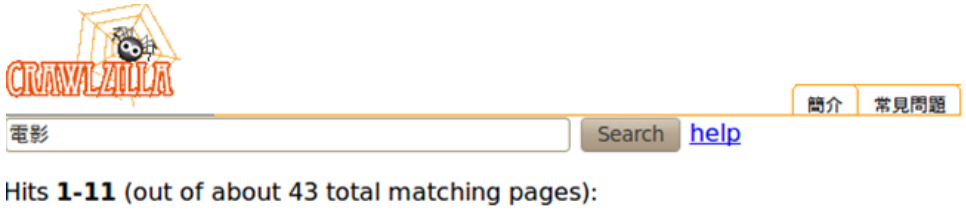

Web UI (Crawlzilla Website + Search Engine)
JSP + Servlet + JavaBean
Nutch
Lucene
Crawlzilla System Management
Tomcat
Hadoop
PC1
PC2
PC3

搜尋引擎加入中文分詞功能

- 索引資料庫會以中文字詞為基本單位建立索引
- 加入中文分詞針對同一網站爬取進行搜尋
  - 搜尋引擎無中文分詞功能時，搜尋關鍵字 - 電影
    - *760* 筆搜尋結果



  - 搜尋引擎加入中文分詞功能時，搜尋關鍵字 - 電影
    - *43* 筆搜尋結果
    - 可提高搜尋的精準度



Crawlzilla - 叢集環境需求

- 如果你覺得…
  - 一台電腦無法滿足你的運算需求
  - 閒置電腦太多
  - 解：讓多台電腦分工運算
- 但是…
  - 架設叢集環境很麻煩!?
  - 解：Crawlzilla 提供叢集安裝模式，只要三分鐘即可建立叢集式搜尋引擎!!!

Resources

- **Crawlzilla @ Google Code Project Hosting (中文說明頁)**
  - http://code.google.com/p/crawlzilla/
- **Crawlzilla @ SourceForge(英文說明頁)**
  - http://sourceforge.net/p/crawlzilla/home/
- **Crawlzilla User Group @ Google**
  - http://groups.google.com/group/crawlzilla-user
- **NCHC Cloud Computing Research Group**
  - http://trac.nchc.org.tw/cloud

總結

楊順發
shunfa@nchc.narl.org.tw
自由軟體實驗室
國家高速網路與計算中心

Big Data Programming using Map/Reduce-HDFS in Hadoop

- **Big Data**
- **Hadoop**
  - MapReduce
  - HDFS
  - Namenode, DataNode
  - Jobtracker, Tasktracker
  - Programming

Resources

- **NCHC Cloud Trac**
  - http://trac.nchc.org.tw/cloud
- **Hadoop 實驗叢集**
  - http://hadoop.nchc.org.tw/
- **台灣 Hadoop 技術討論區**
  - http://forum.hadoop.tw/
- **臉書粉絲團**
  - https://www.facebook.com/groups/hadoop.tw/

Ref.

- **Apache™ Hadoop®!**
  - http://hadoop.apache.org/
- **hdfs-federation-hadoop-summit2011**
  - http://www.slideshare.net/huguk/hdfs-federation-hadoop-summit2011
- **Hadoop: The Definitive Guide**
  - http://www.amazon.com/Hadoop-Definitive-Guide-Tom-White/dp/0596521979
- **Hadoop in Action**
  - http://manning.com/lam/