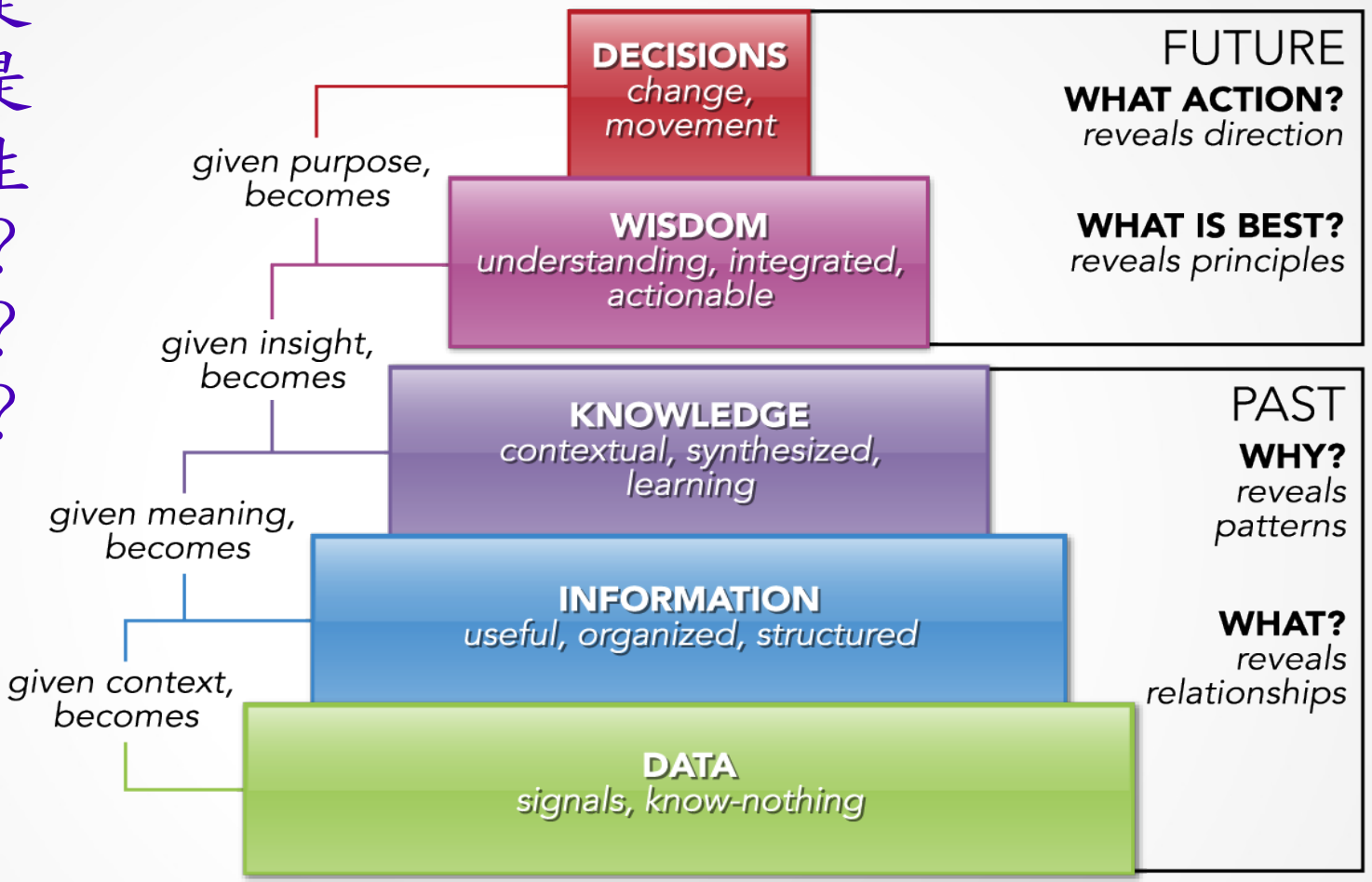
Three New Trends of Big Data
即時・安全・易用

王耀聰 / 國家高速網路與計算中心
Jazz Yao-Tsung Wang / NCHC
<jazz@nchc.narl.org.tw>

知識源自彙整過去，智慧在能預測未來

資料多寡不是重點，重點是我們想要產生什麼價值呢？時效合理嘛？成本合理嘛？

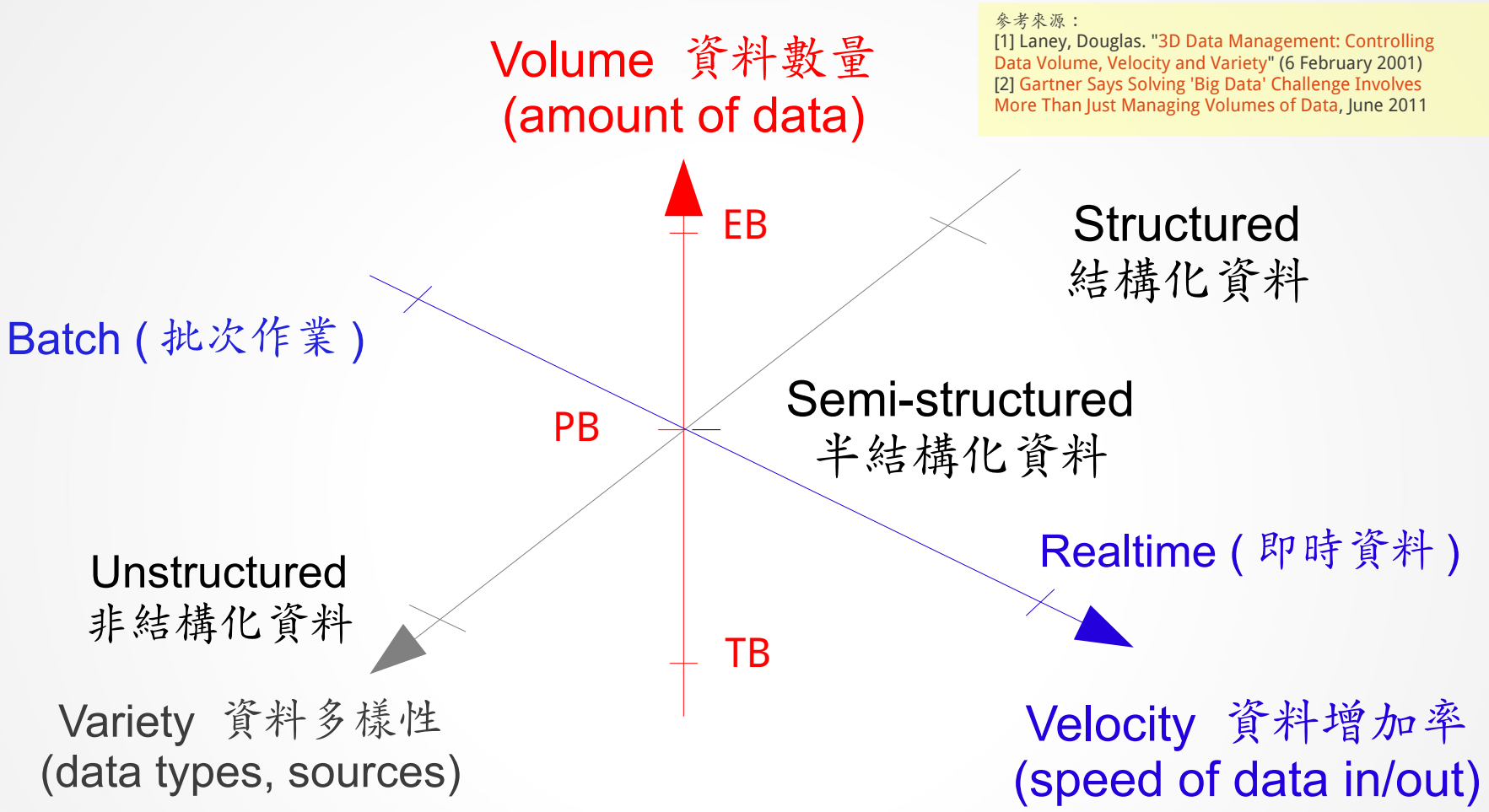


http://www.pursuantgroup.com/blog/tag/dikw-model/

大家都說「資料是金礦」，那就讓我們拿採礦當類比吧！

國際金價	提供給客戶的價值	產品通路
開採成本	總擁有成本	軟硬體投資
提煉廠	分析平台與工具軟體	**SMAQ**
含金度	資料鑑價？	商業模式
開採權	分析資料的合法性	個資法
金礦	資料集	**Open Data**

3 Vs of Big Data

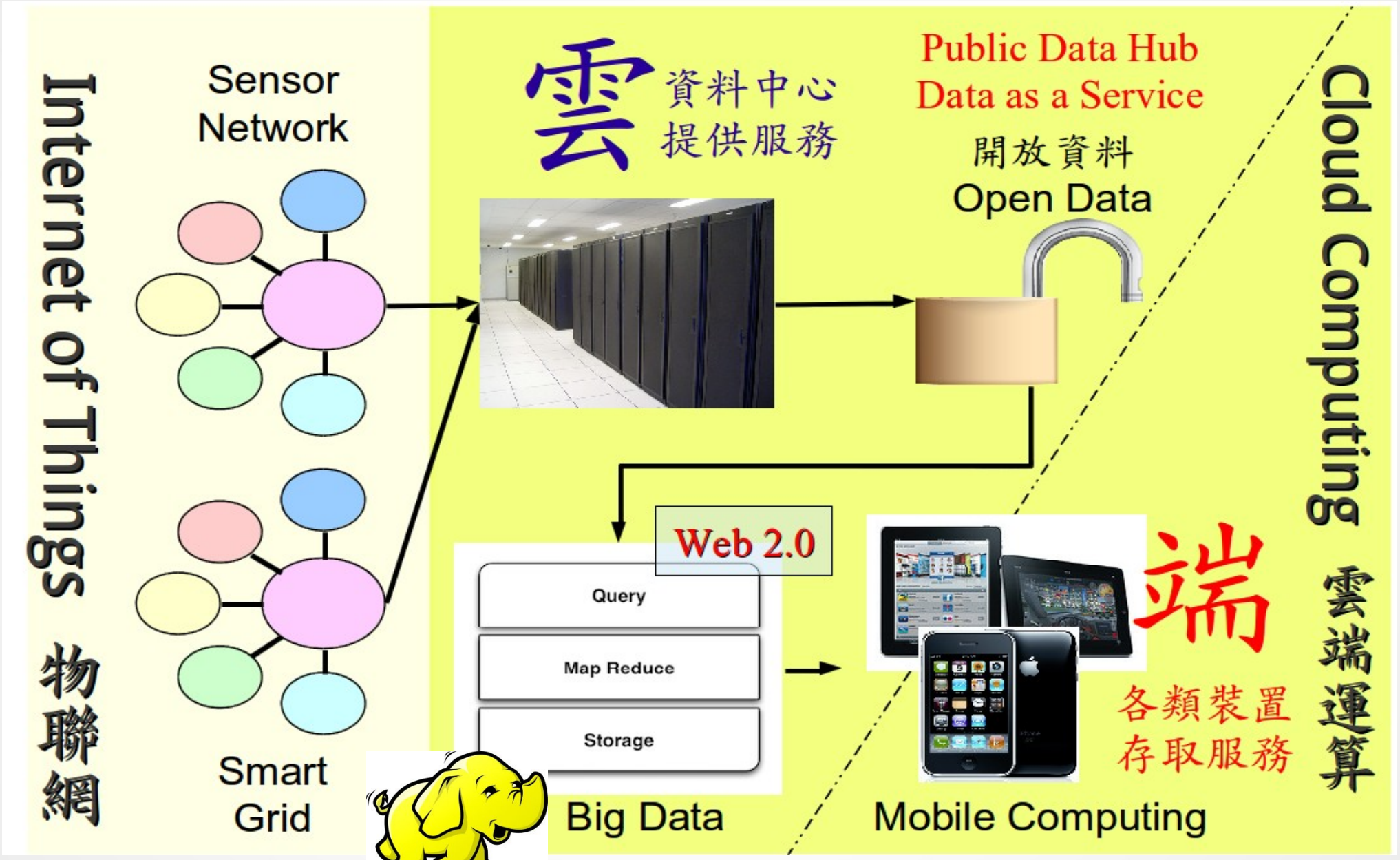


參考來源：
[1] Laney, Douglas. "3D Data Management: Controlling Data Volume, Velocity and Variety" (6 February 2001)
[2] Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data, June 2011

巨量資料的挑戰在於如何管理「數量」、「增加率」與「多樣性」

4

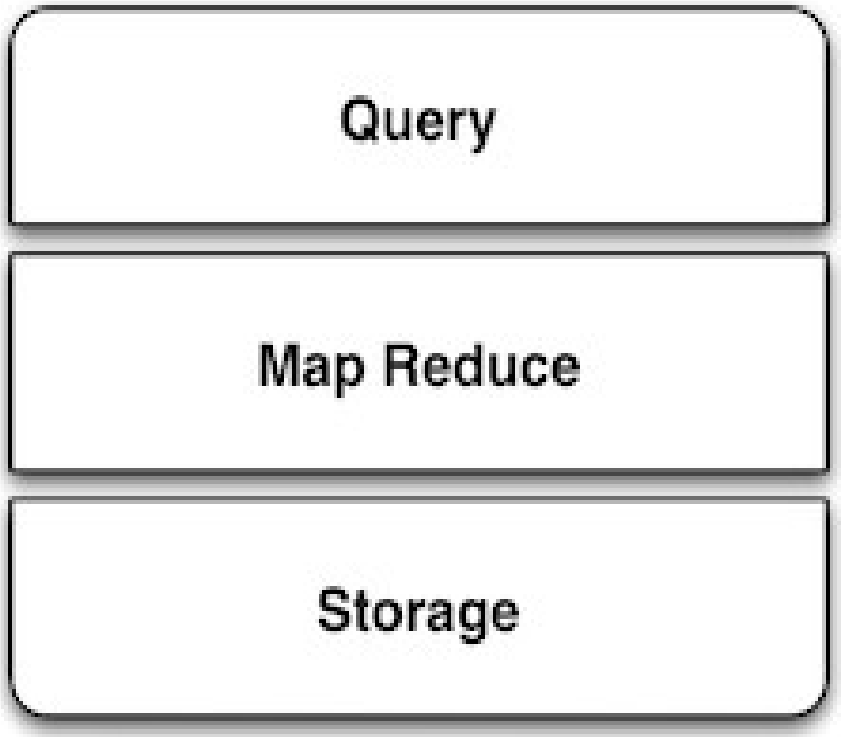
Life of Big Data ：蒐、存、取、析、用



巨量資料處理的資訊架構
The SMAQ stack for big data

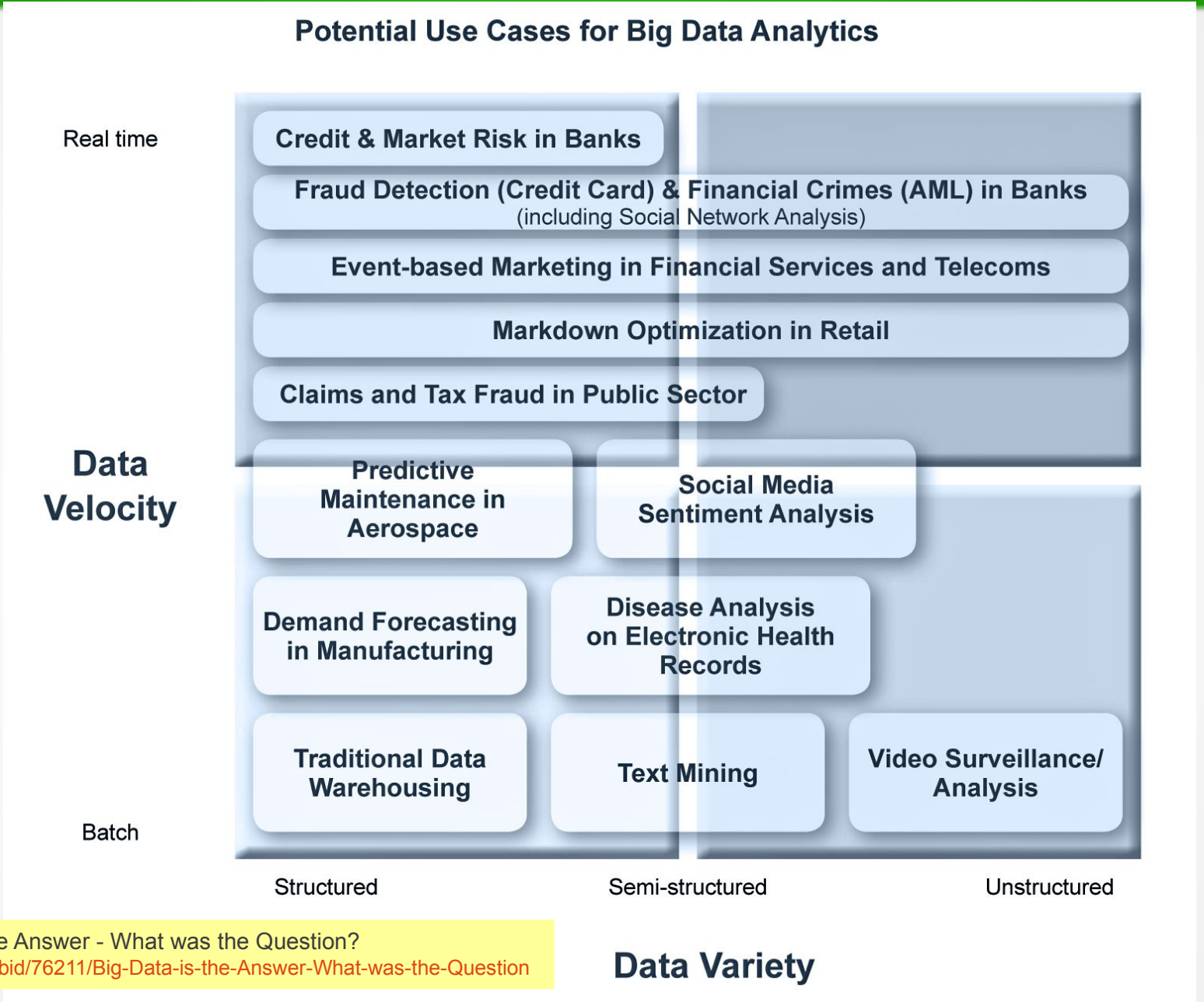
做網頁相關的人可能聽過 LAMP

未來處理海量資料的人必需知道
SMAQ（Storage, MapReduce and Query）



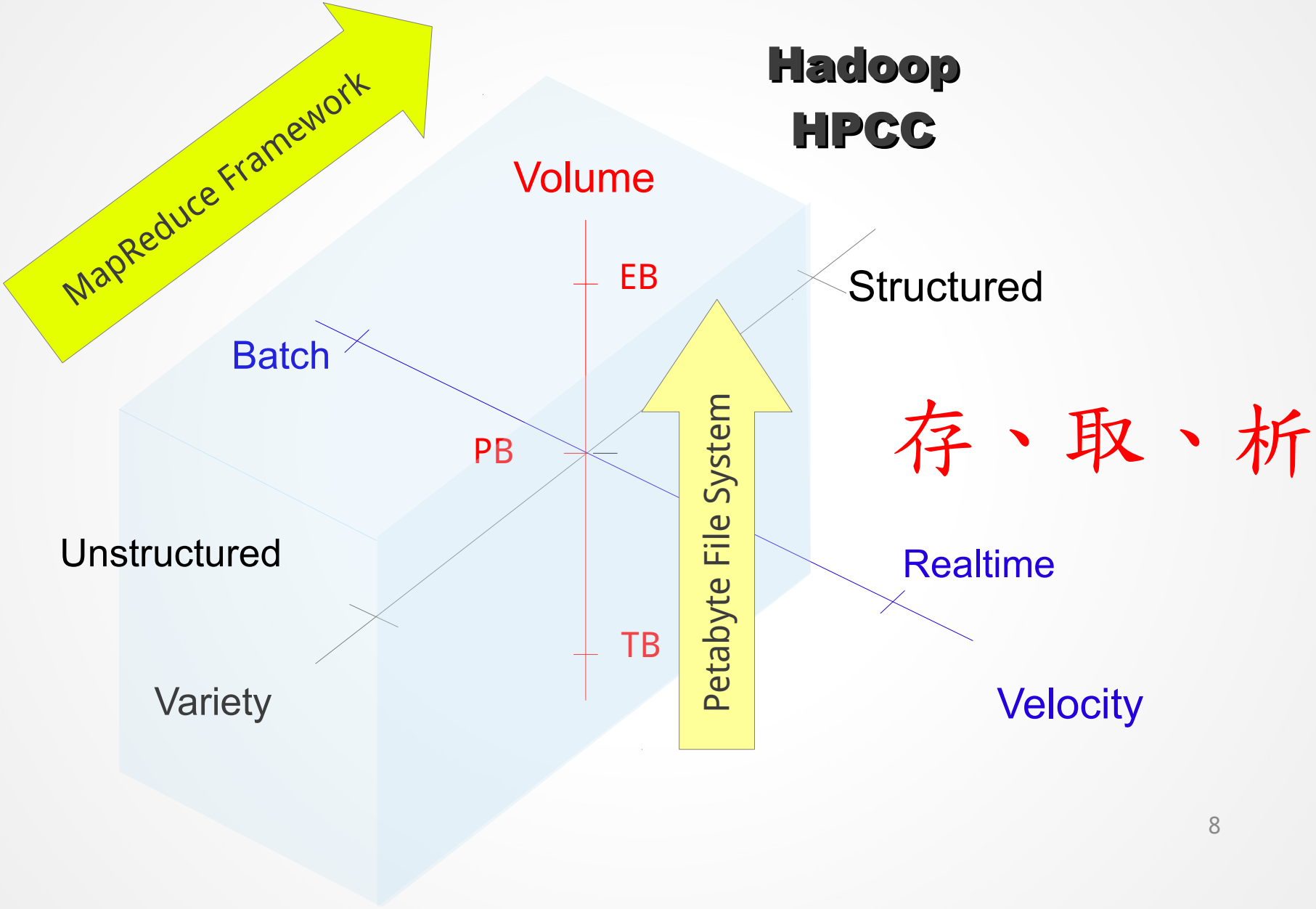
參考來源：The SMAQ stack for big data，Edd Dumbill，22 September 2010，
http://radar.oreilly.com/2010/09/the-smaq-stack-for-big-data.html
圖片來源：http://smashingweb.ge6.org/wp-content/uploads/2011/10/apache-php-mysql-ubuntu.png

Big Data is the Answer - What was the Question?



參考來源： Big Data is the Answer - What was the Question?
http://www.saama.com/blog/bid/76211/Big-Data-is-the-Answer-What-was-the-Question

Big Data at Rest – MapReduce Framework



8

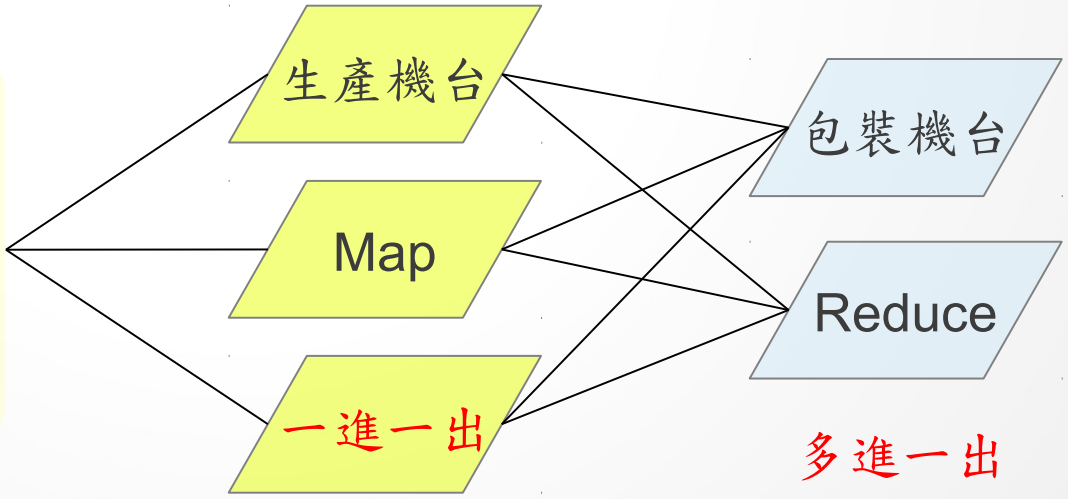
高資料通量處理平台 Hadoop
Key Concept : Data Locality

Hadoop 是一個讓使用者簡易撰寫並執行處理海量資料應用程式的軟體平台。

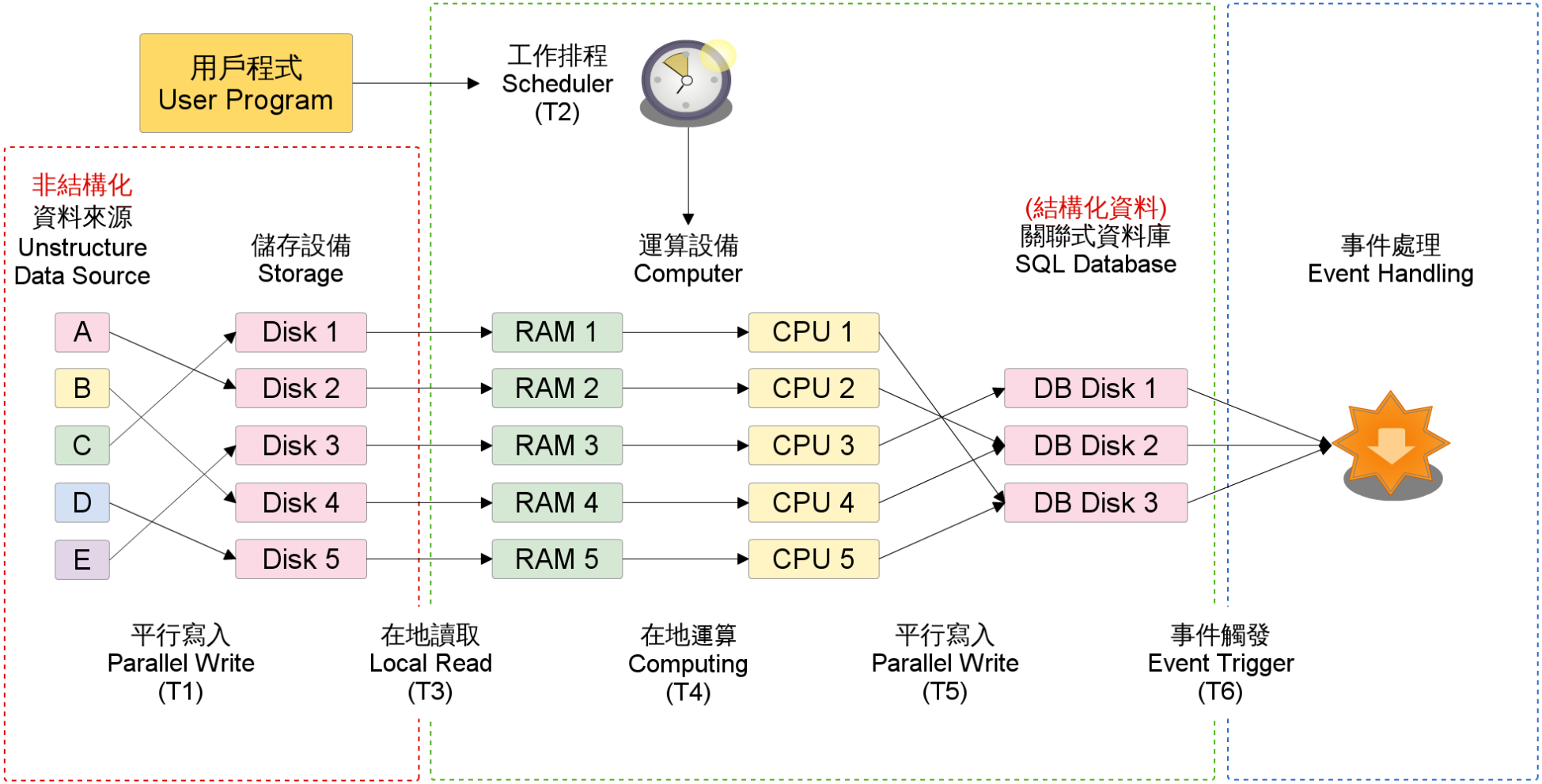
亦可以想像成一個處理海量資料的生產線，只須學會定義 **map** 跟 **reduce** 工作站該做哪些事情。

就像工廠的倉庫
存放生產原料跟待售貨物

HDFS 存放
待處理的非結構化資料
與處理後的結構化資料



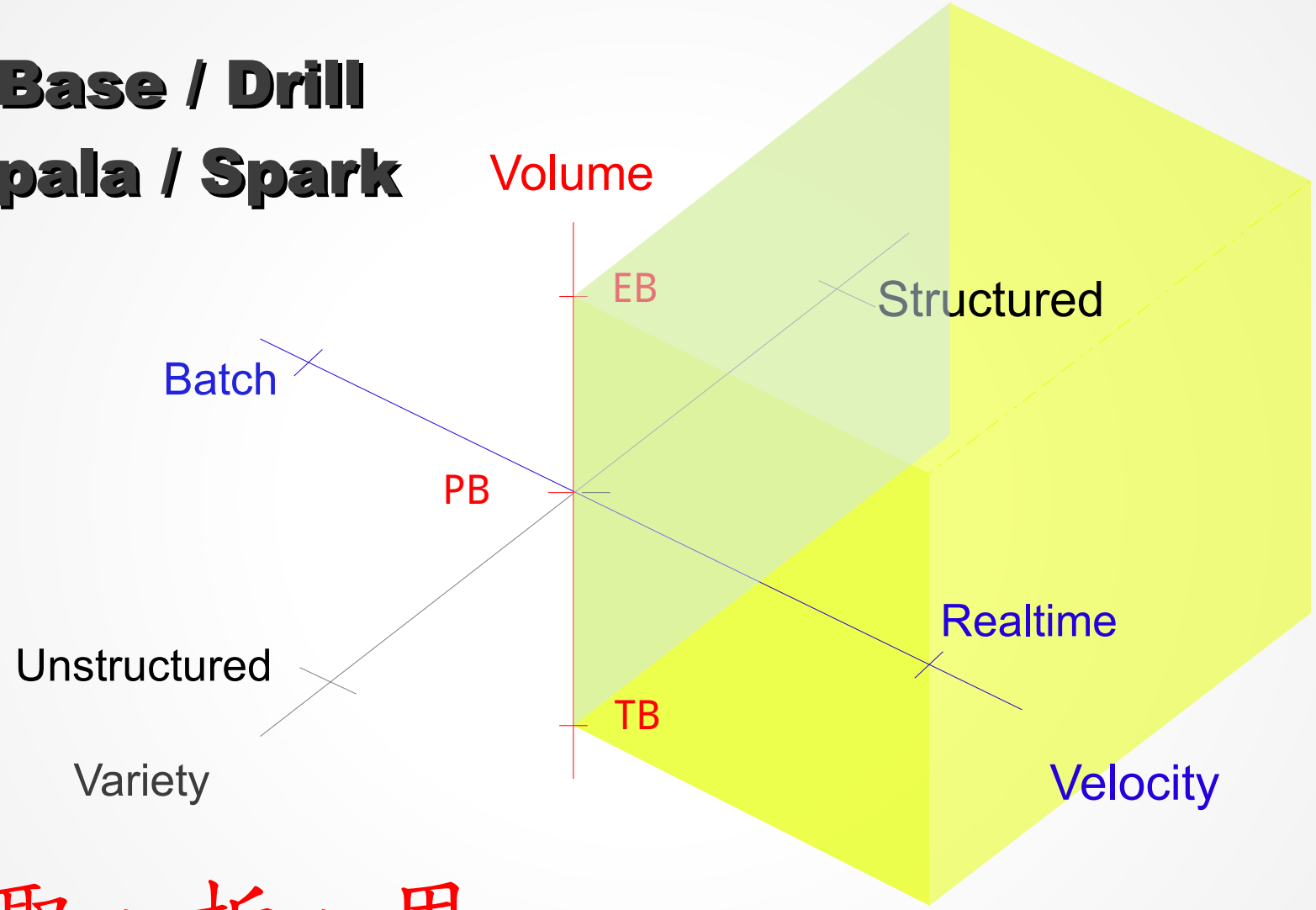
批次作業的運算時間
Processing Time of Batch Jobs



Big Data in Motion – In-Memory Processing 、 Predictive Analytics

HBase / Drill

Impala / Spark

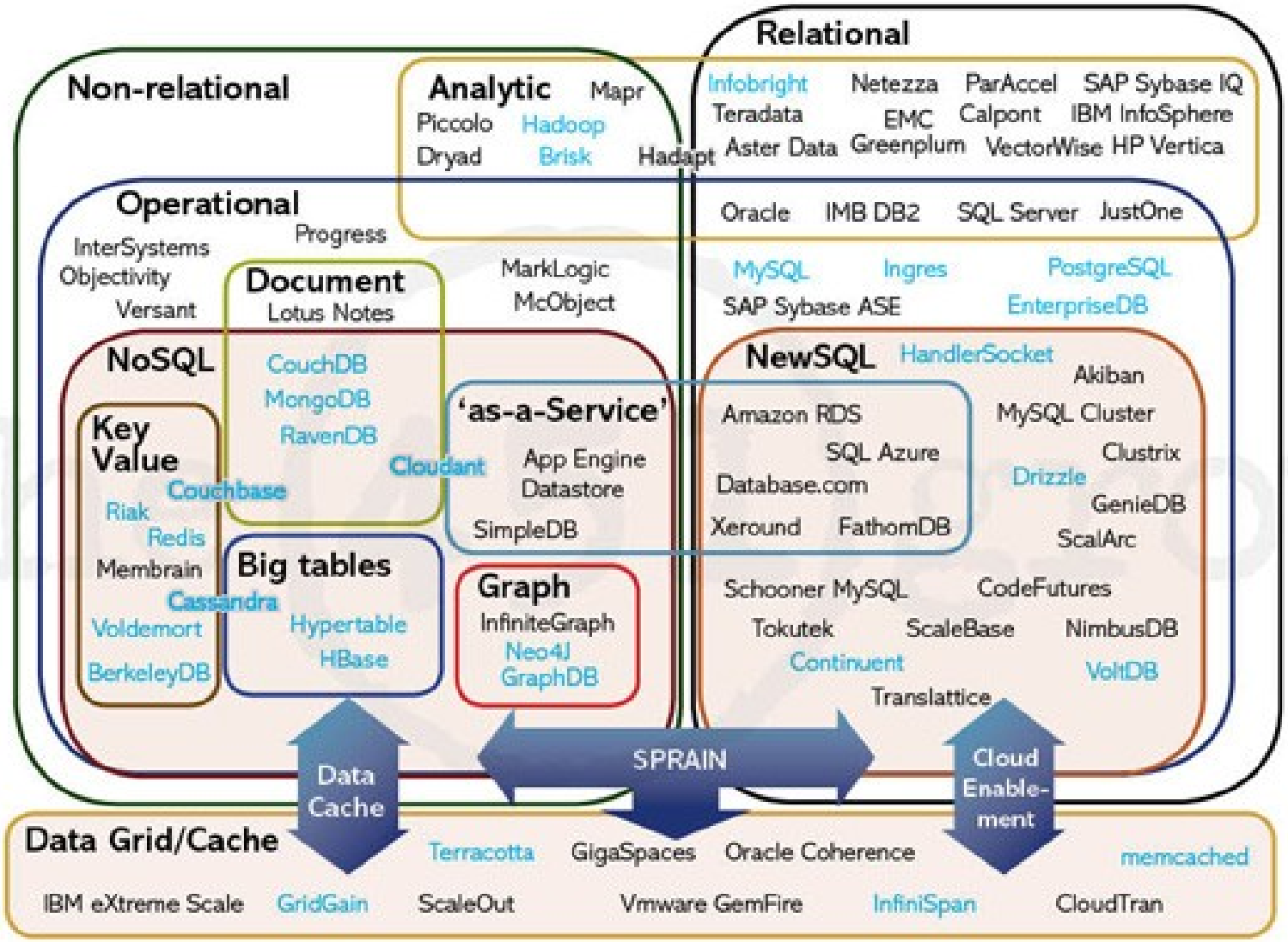


取、析、用

Google 的技術演進 vs Apache 專案

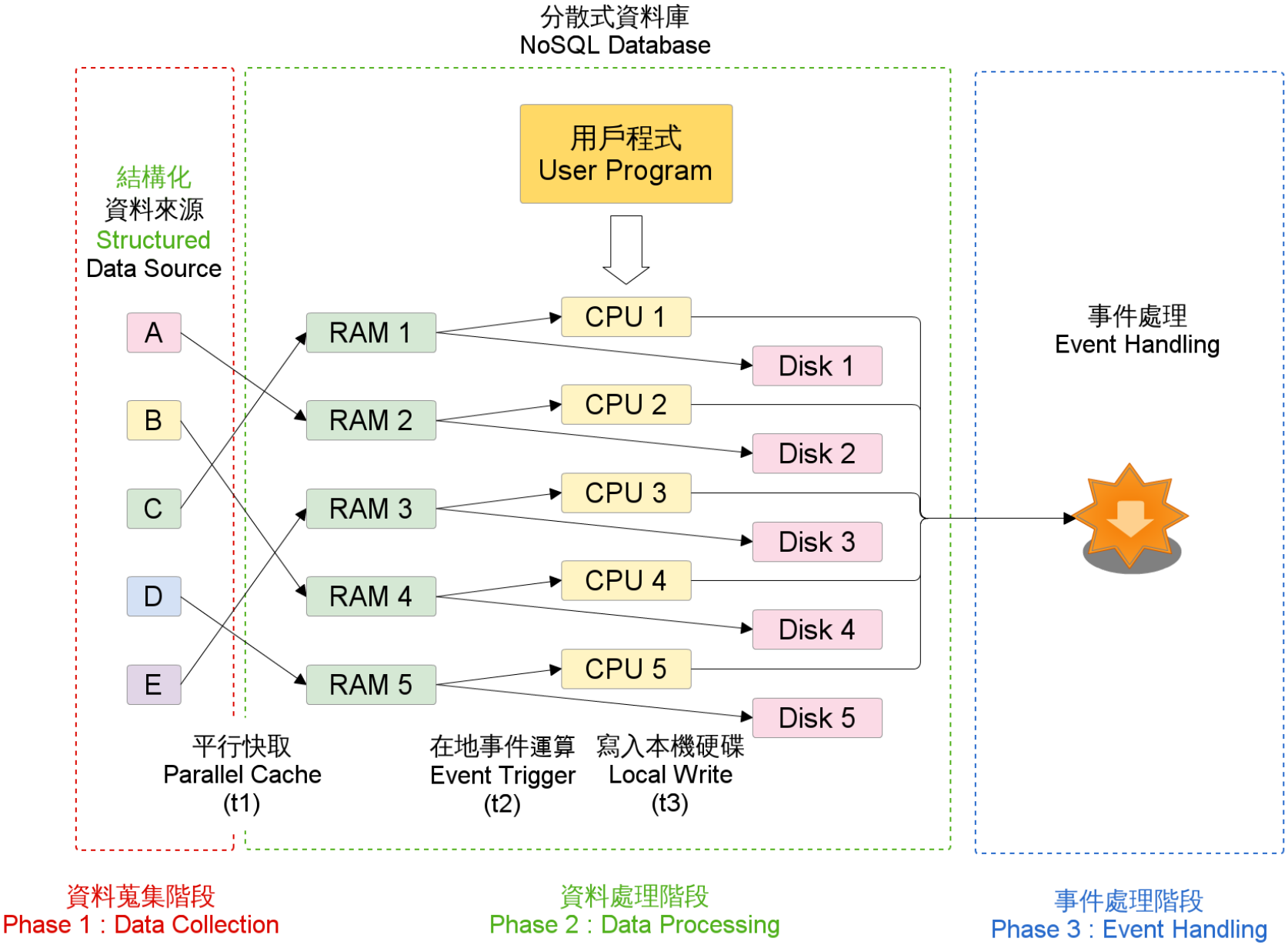
Big Query (JSON, SQL-like)	Dremel (2010)	Apache Drill (2012)
Incremental Index Update (Caffeine)	Percolator (2010)	
Graph Database	Pregel (2009)	Apache Giraph (2011)
Query	BigTable (2006)	Apache HBase (2007)
Map Reduce	MapReduce (2004)	Hadoop MapReduce (2006)
Storage	Google File System (2003)	HDFS (2006)

令人眼花撩亂的多樣化資料庫選擇
NoSQL vs NewSQL



http://www.infoq.com/news/2011/04/newsql

In-Memory Processing 的運算時間
以 HBase 為例



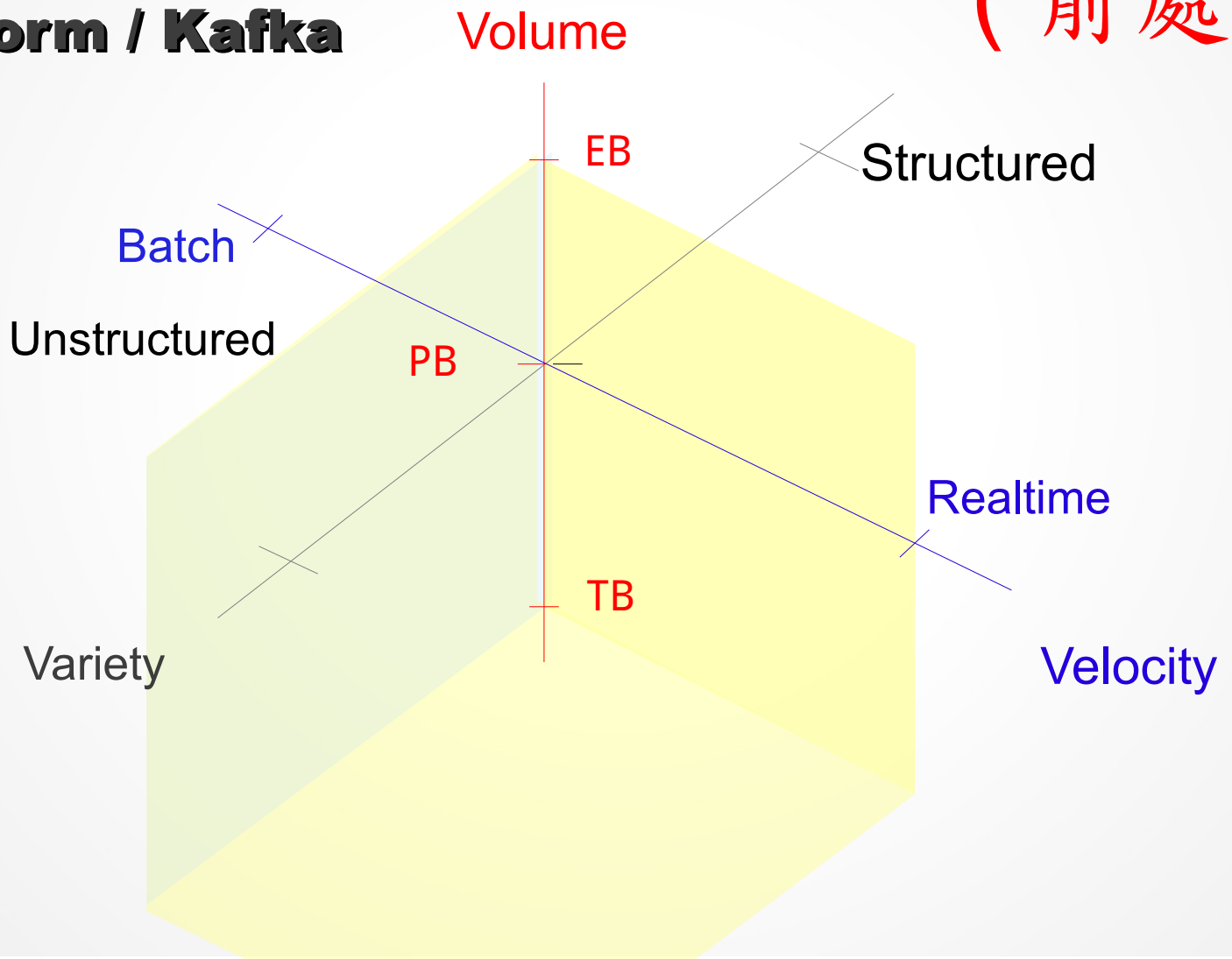
Big Data in Motion – Streaming Data Collection / Data Cleaning

Message Queue

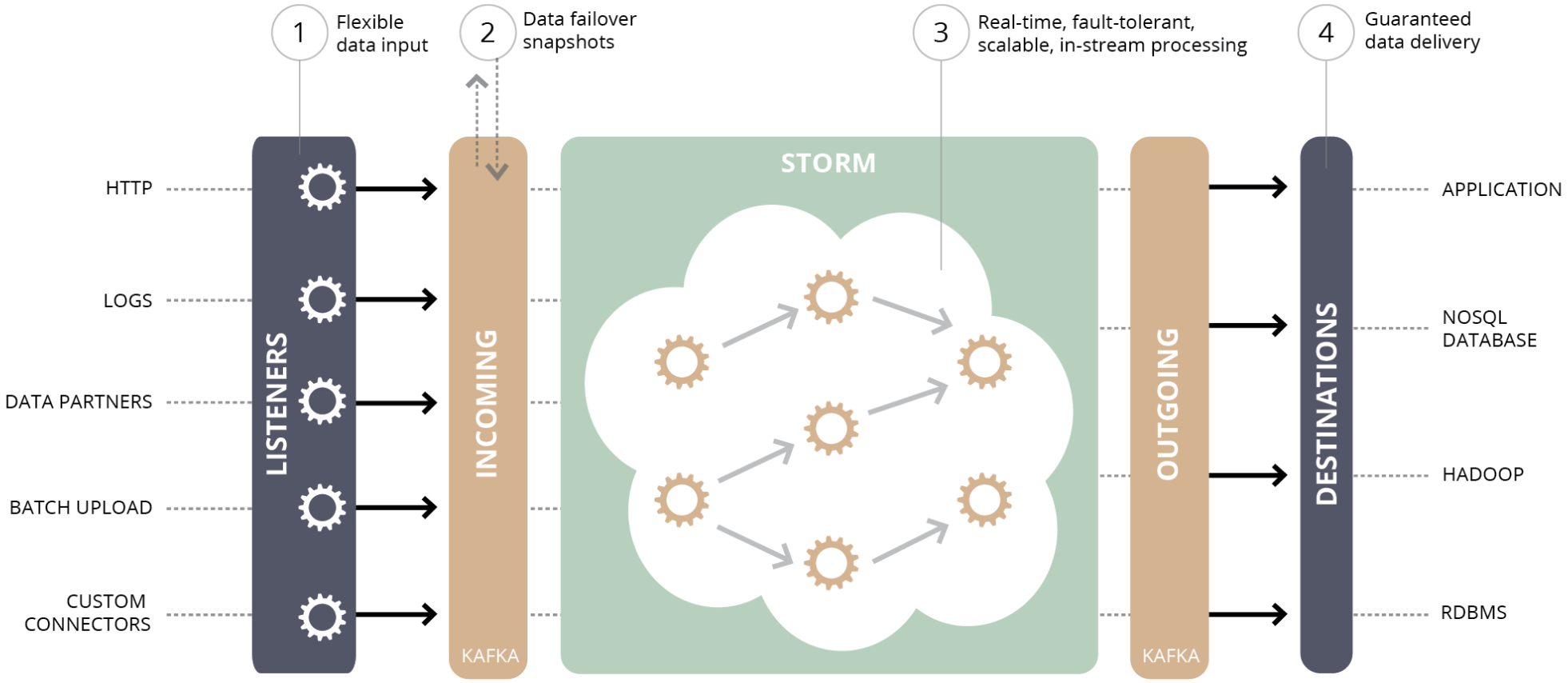
(AMQP , RabbitMQ)

Storm / Kafka

蒐、存
（前處理）

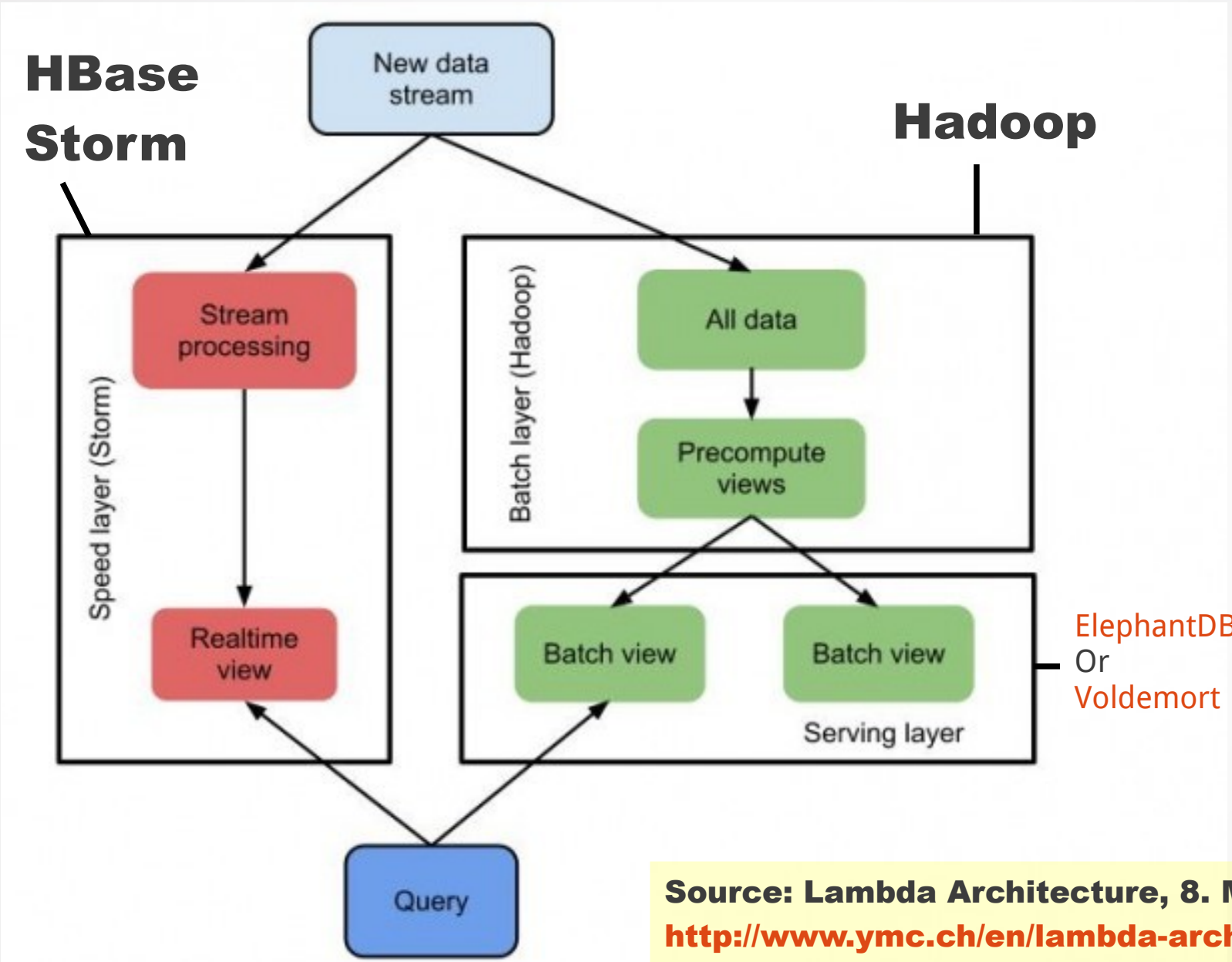


Twitter Storm + Apache Kafka



http://blog.infochimps.com/2012/10/30/next-gen-real-time-streaming-storm-kafka-integration/

混合模式的巨量資料處理架構
Lambda Architecture



Source: Lambda Architecture, 8. March 2013
http://www.ymc.ch/en/lambda-architecture-part-1

Next Step : Big Data Security

當我們緊密相連 .....

世界政經：歐盟想分 **Tweeter**
找出經濟、政治的脈動

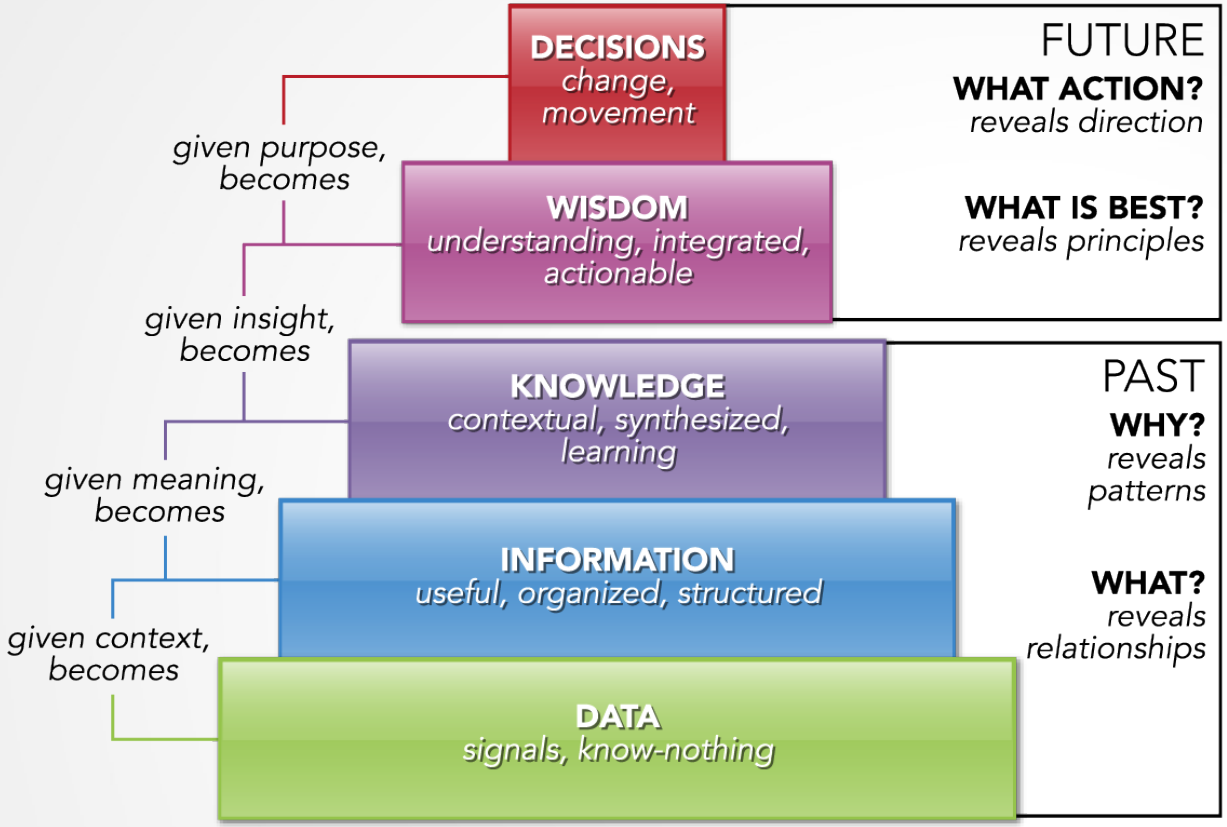
國家安全：美國 **PRISM** 計劃
(網軍！終極警探 4.0)

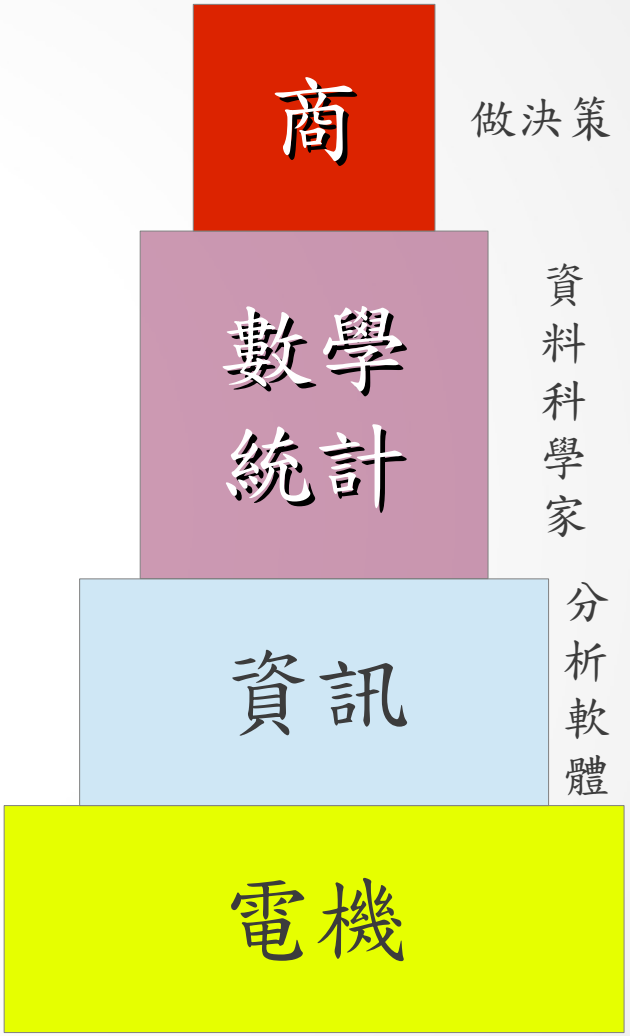
組織如何因應 **APT ?**
Big Data 平台本身的安全性 **?**

有太多安全的問題等待解決！

Source: Gartner (March 2011), 'Big Data' Is Only the Beginning of Extreme Information Management, 7 April 2011, http://www.gartner.com/id=1622715

To Find the Value of Big Data
We need Data Scientist Team !





重點在找到價值

Value

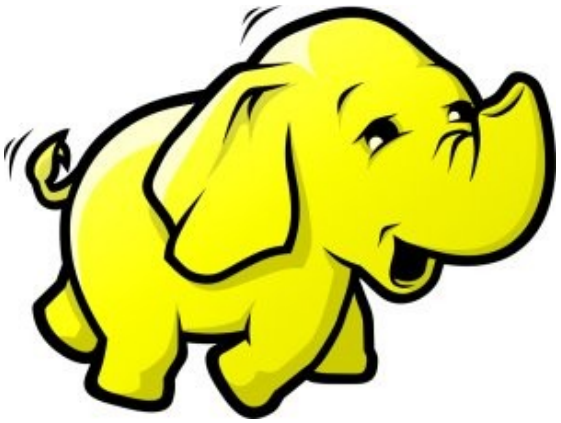
Hadoop 專業術語

Introduction to Hadoop Terminology

Jazz Wang

Yao-Tsung Wang

jazz@nchc.org.tw



Two Key Elements of Operating System
作業系統兩大關鍵組成元素

Scheduler
程序排程

File System
檔案系統

Terminologies of Hadoop
Hadoop 文件中的專業術語

- 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

Two Key Roles of HDFS
HDFS 軟體架構的兩種關鍵角色

名稱節點 NameNode

- Master Node
- Manage NameSpace of HDFS
- Control Permission of Read and Write
- Define the policy of Replication
- Audit and Record the NameSpace
- Single Point of Failure

資料節點 DataNode

- Worker Nodes
- Perform operation of Read and Write
- Execute the request of Replication
- Multiple Nodes

Two Key Roles of Job Scheduler
程序排程的兩種關鍵角色

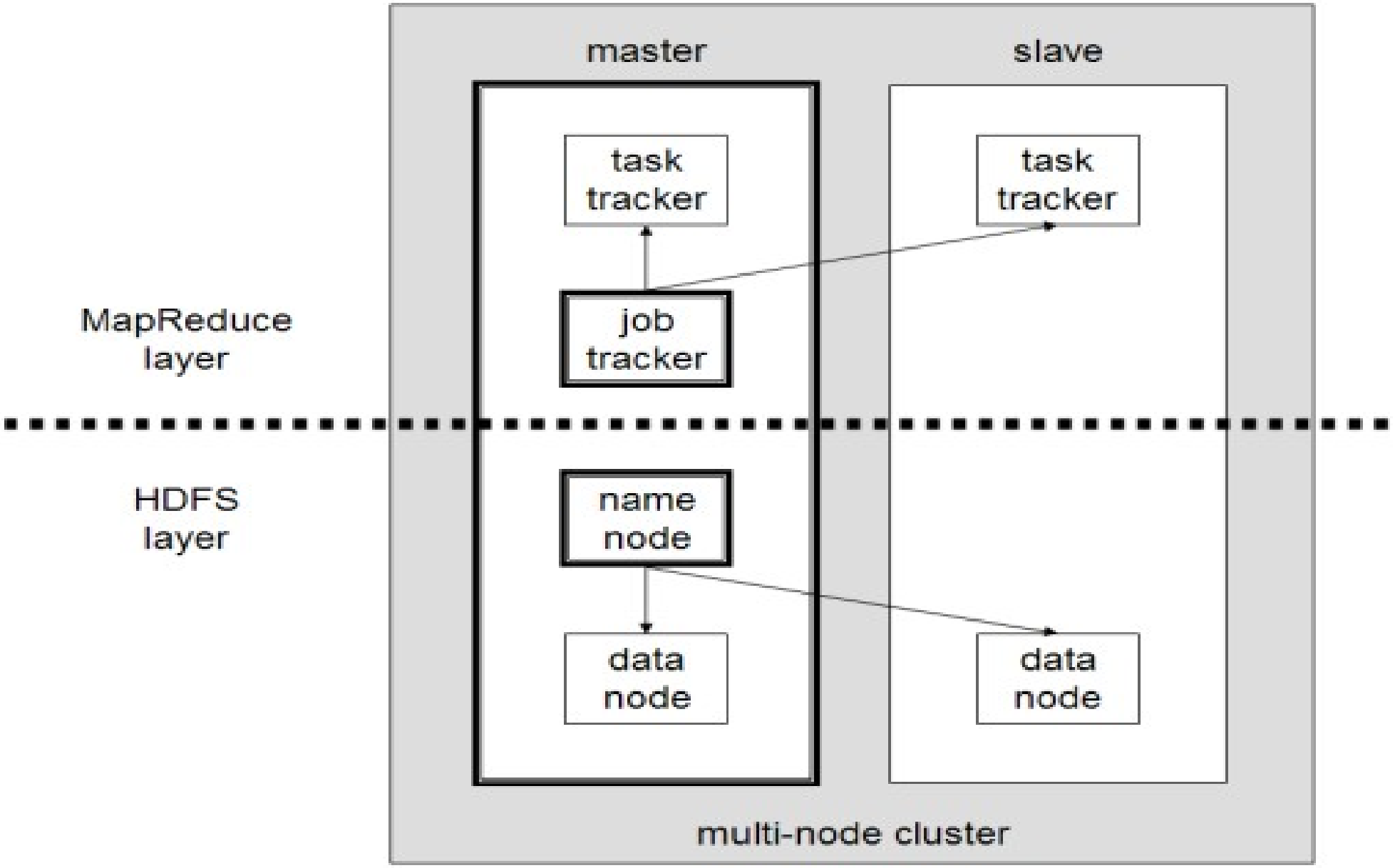
JobTracker

- Master Node
- Receive Jobs from Hadoop Clients
- Assigned Tasks to TaskTrackers
- Define Job Queuing Policy, Priority and Error Handling
- Single Point of Failure

TaskTracker

- Worker Nodes
- Excute Mapper and Reducer Tasks
- Save Results and report task status
- Multiple Nodes

Different Roles of Hadoop Architecture
Hadoop 軟體架構中的不同角色



Hadoop 除錯、效能監控與調校指標

How to debug, measure the performance and key index of performance

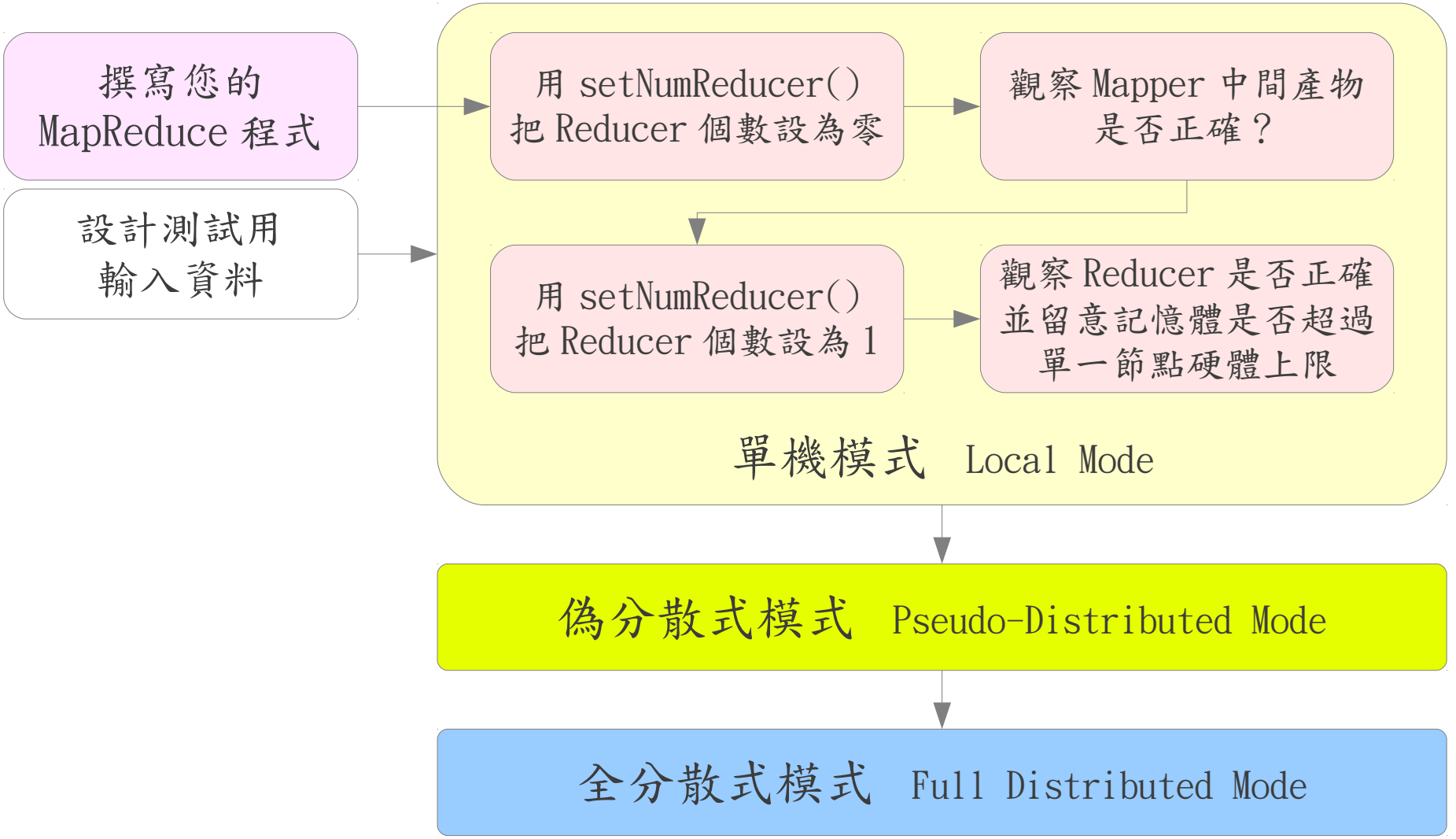
Jazz Wang

Yao-Tsung Wang

jazz@nchc.org.tw

Powered by DRBL

Hadoop Debug Process 標準除錯程序



http://wiki.apache.org/hadoop/HowToDebugMapReducePrograms

Java Remote Debug

- 有時靠 **System.out.println()** 是不夠的，有人想要 **Step Trace** 怎麼辦
- 先在 **Local Mode** 執行，啓動 **Java Remote Debug** 的參數，然後用 **Eclipse** 的 **Step Trace** 功能來觀察程式的行爲。

```
export HADOOP_OPTS=
"-agentlib:jdwp=transport=dt_socket,server=y,suspend=y,address=5000"
```

http://javarevisited.blogspot.com/2011/02/how-to-setup-remote-debugging-in.html
http://code.google.com/p/hadoop-clusternet/wiki/DebuggingJobsUsingEclipse

How to setup java remote debugging in eclipse

by JAVIN PAUL • FEB. 26, 2011

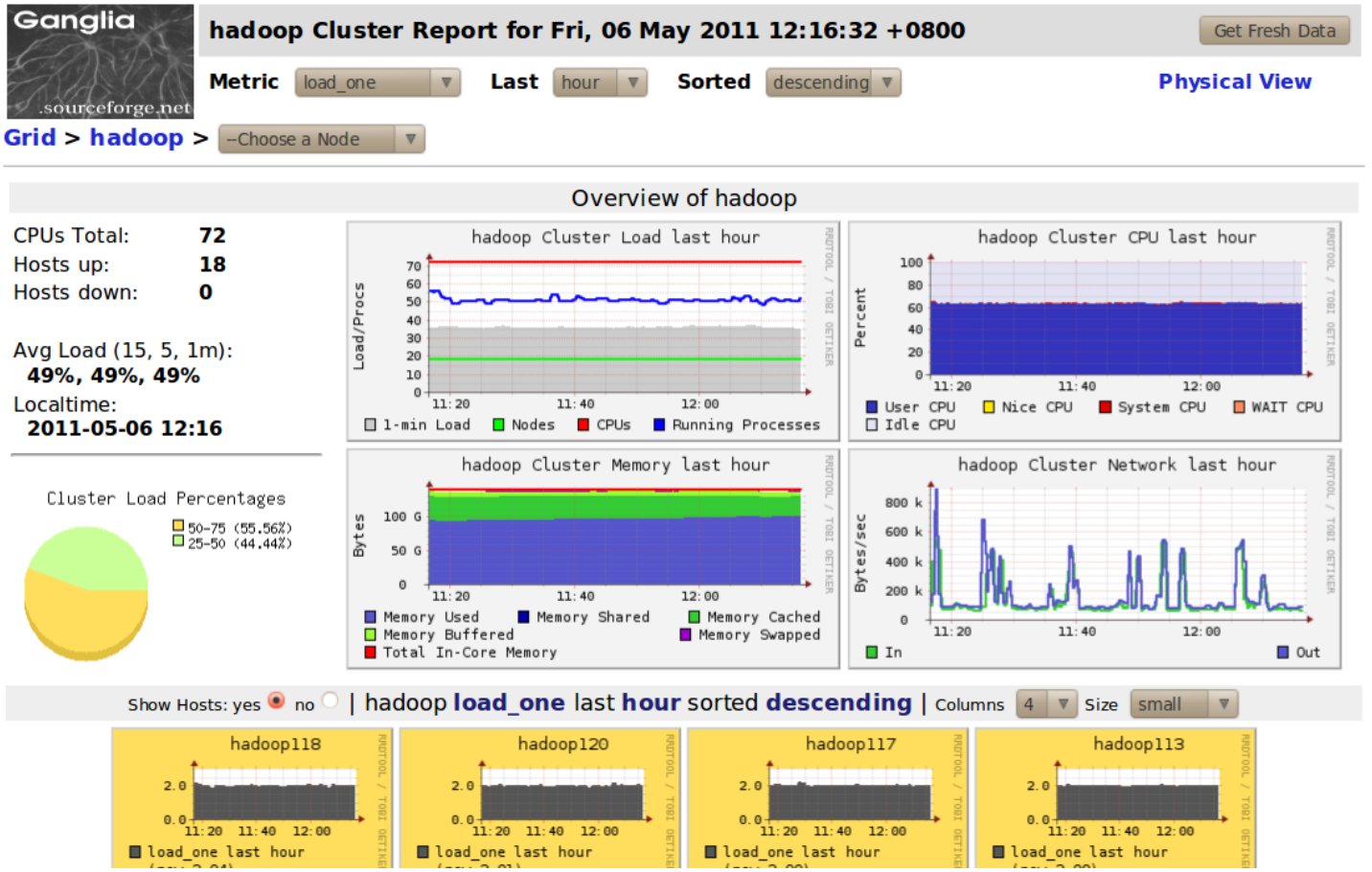
READ LATER

Remote debugging is not a new concept and many of you are aware of this just for who don't know what is remote debugging? It's a way of debugging any process could be **Java** or C++ running on some other location from your development machine. Since debugging is essential part of development and ability to debug your application not only

系統狀態監控 Ganglia

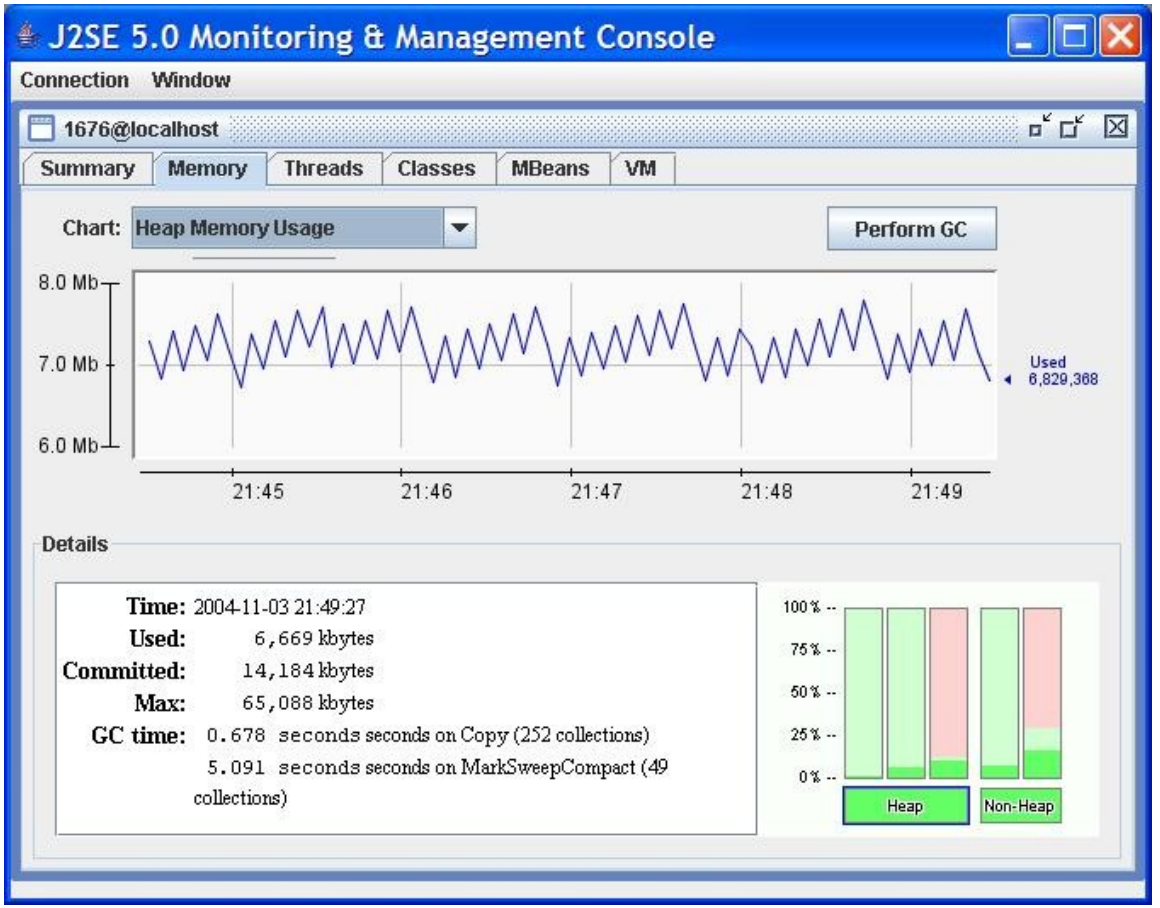
- **Hadoop** 預設可以產生效能數據（**Metrics**）給 **Ganglia**
- 請根據您的 **Ganglia** 安裝情形設定 **conf/hadoop-metrics.properties**

http://ganglia.sourceforge.net/



Hadoop 系統狀態監控 JMX

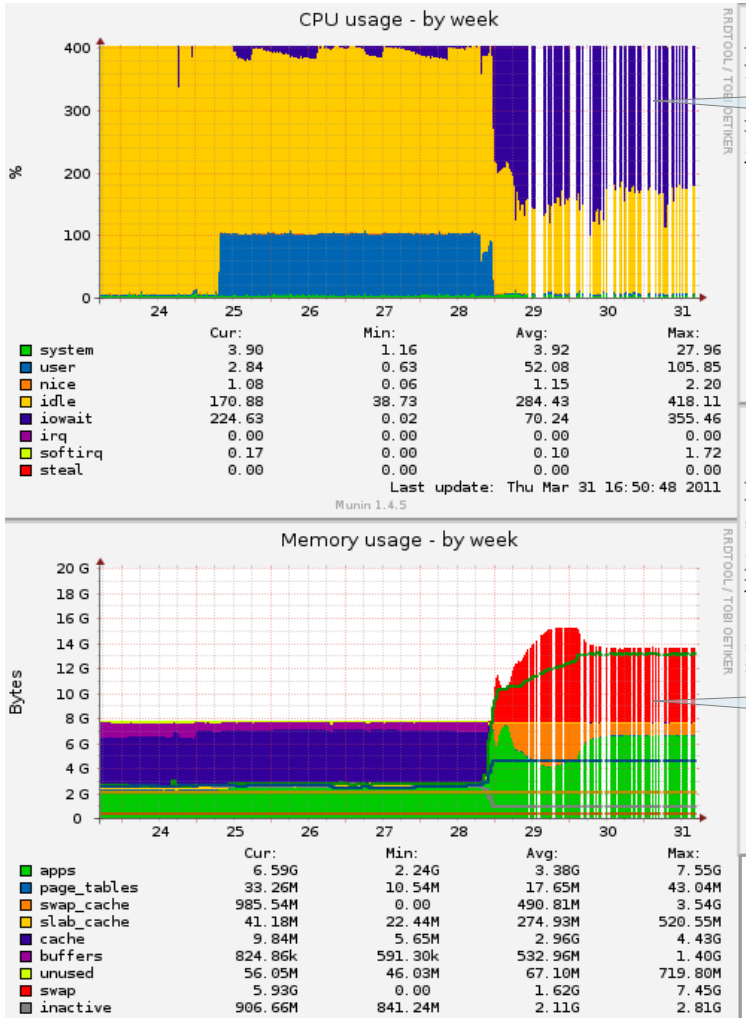
- 由於 **Ganglia** 的取樣頻率一般是 **10** 秒一次到一分鐘一次，若是需要更即時的狀態資料，可以使用 **JMX Client** 來讀取 **Hadoop** 送出的 **Metrics**
- 像是 **jconsole**、 **Hyperic**或 **Nagios**等。

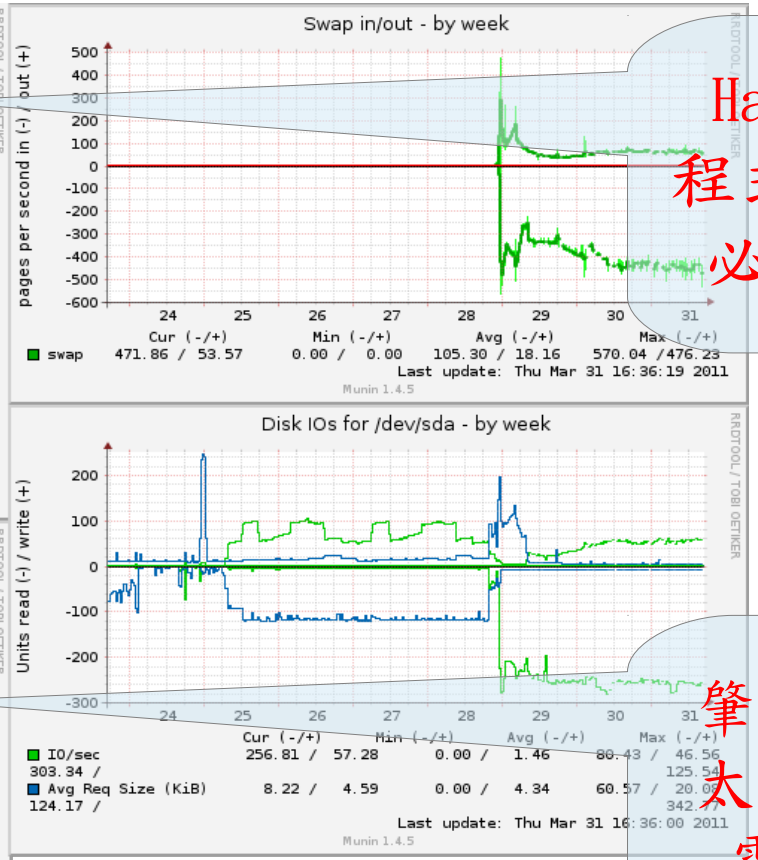


http://www.cloudera.com/blog/2009/03/hadoop-metrics/

系統 I/O 狀態監控 Munin

- 由於 **Ganglia** 所蒐集的資料並沒有每顆硬碟的 **I/O** 數據，有時會使用 **Munin** 這套軟體來了解每顆硬碟的 **I/O** 情形，進而分析讀寫效能。
- 當讀寫 **I/O** 遠慢於 **CPU** 運算時，會發生 **IOWAIT**





Hadoop MapReduce 程式若出現 IOWait 必然效能不彰！

肇事原因常是用了太多記憶體，結果需要改用 SWAP～

Optimization = I/O Impedance ? 資料通量達成匹配？

電路講究阻抗匹配，資料探勘講究計算與讀寫通量的匹配。