



**hadoop**

in

2013

Taiwan

即時 · 安全 · 易用



# 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>

# 教師節快樂！謝謝各位蒞臨！

祝台下的老師們教師節快樂！

**Happy Teacher's Day !!**



值日生



主辦單位： **TREND MICRO**  
趨勢科技 [ 全程護航  
邁向雲端 ]

 **NCHC**  
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

 **TCloud**  
騰雲計算股份有限公司

協辦單位： 台灣雲端運算產業協會  
Cloud Competing Association in Taiwan

 財團法人資訊工業策進會  
INSTITUTE FOR INFORMATION INDUSTRY  
雲端系統軟體研究所  
CLOUD SYSTEM SOFTWARE INSTITUTE

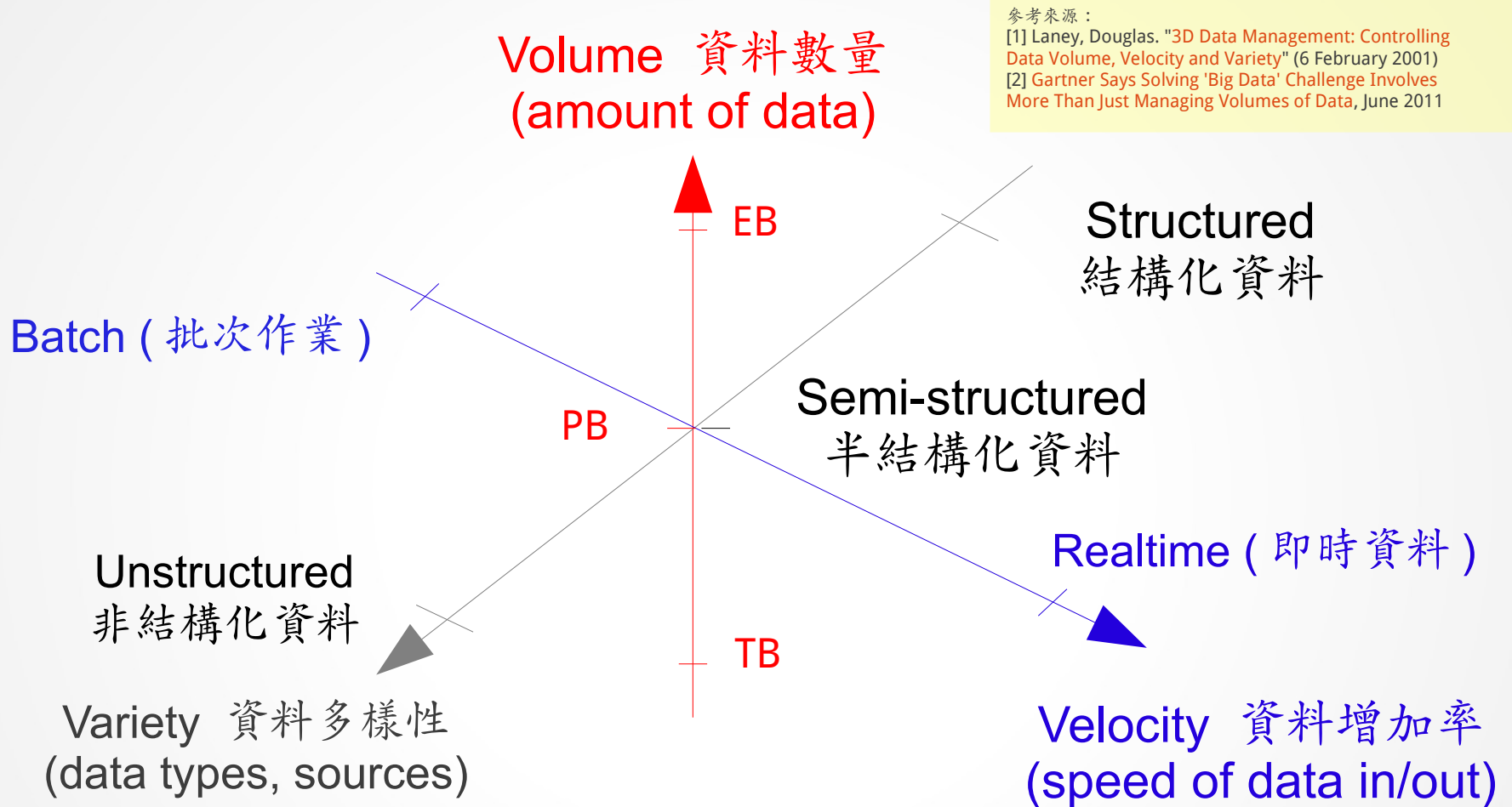
 **OpenFoundry**  
Your project hosting

贊助廠商白金級： **etu**  
nexus enlightenment

 **Microsoft**

感謝主辦單位  
與贊助廠商

# 3 Vs of Big Data

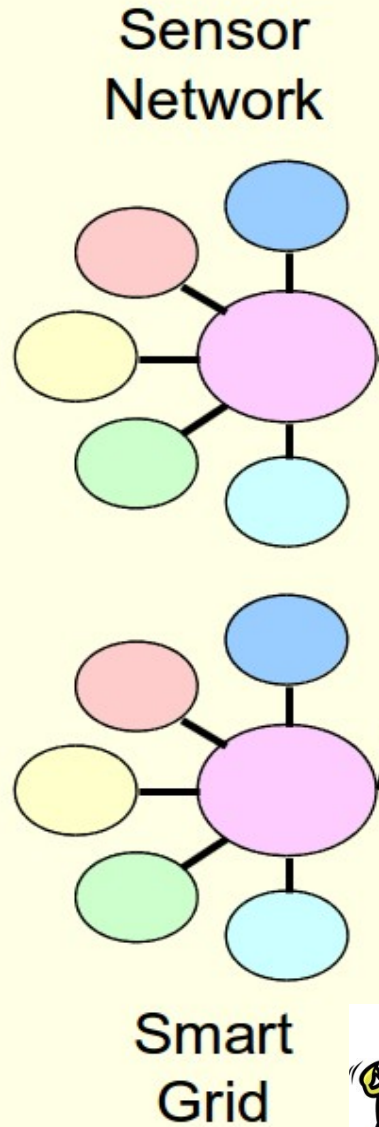


巨量資料的挑戰在於如何管理「數量」、「增加率」與「多樣性」

3

# Life of Big Data : 蒐、存、取、析、用

Internet of Things  
物聯網



雲 資料中心  
提供服務



Public Data Hub  
Data as a Service

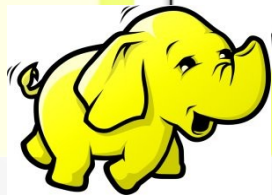
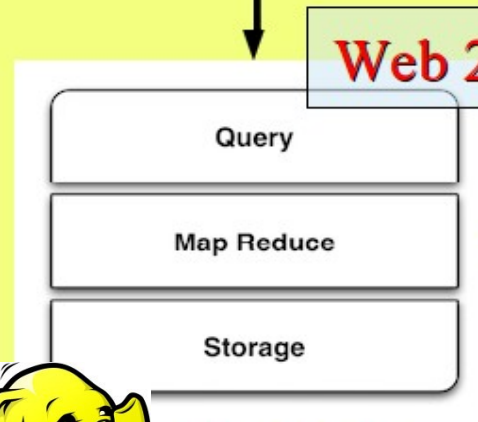
開放資料  
Open Data



Cloud Computing

雲端運算

Web 2.0



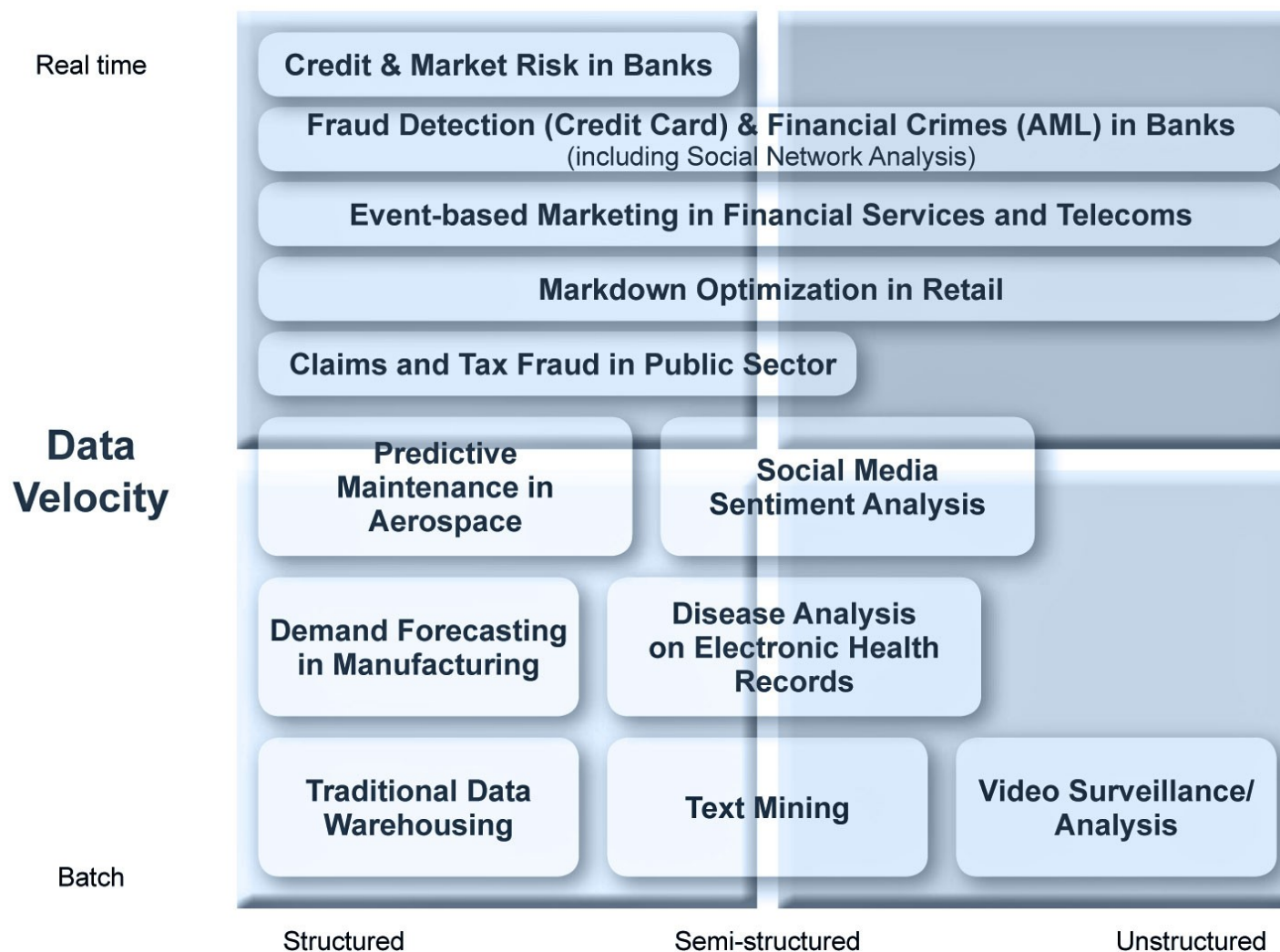
端

各類裝置  
存取服務

Mobile Computing

# Big Data is the Answer - What was the Question?

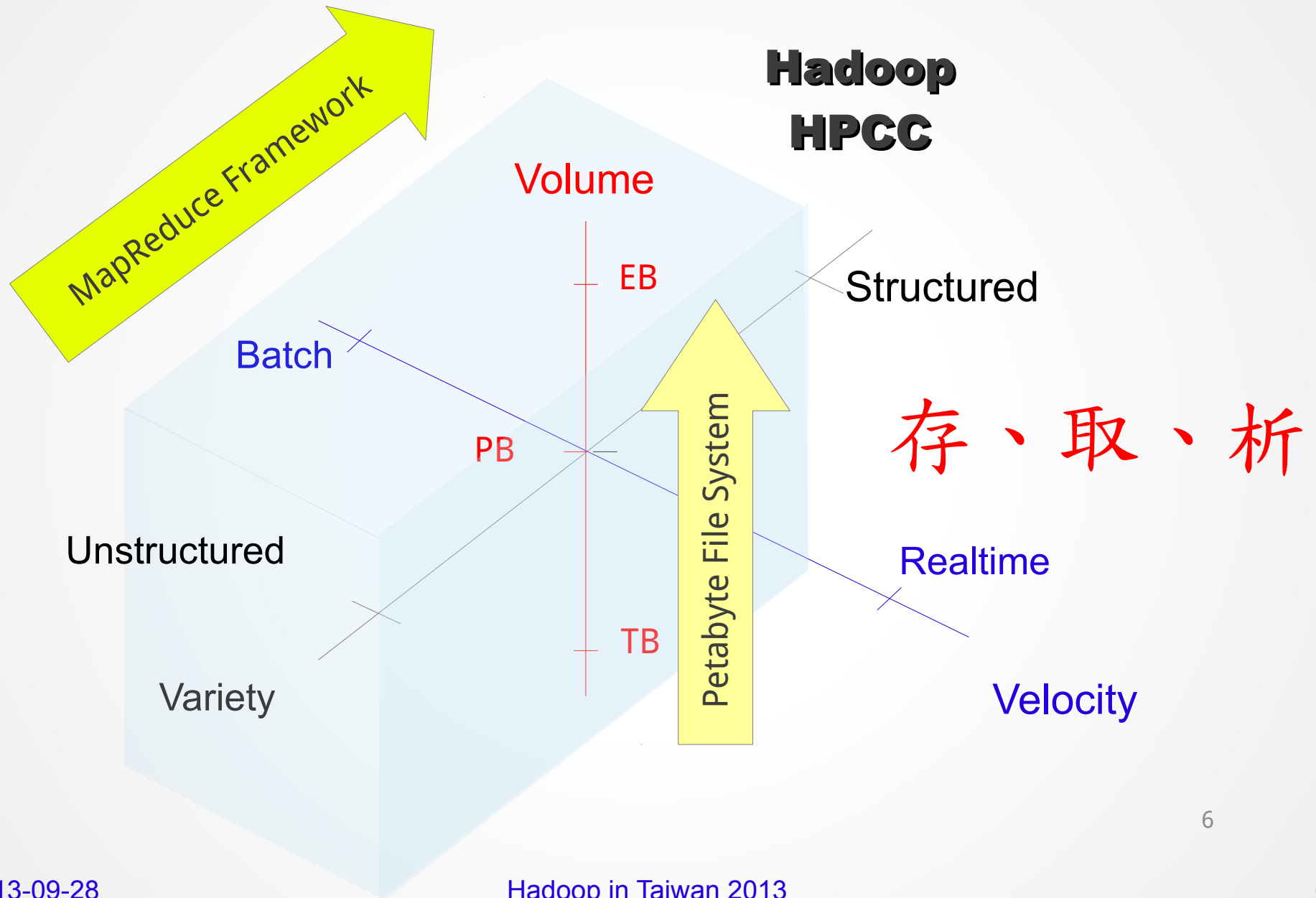
## Potential Use Cases for Big Data Analytics



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>

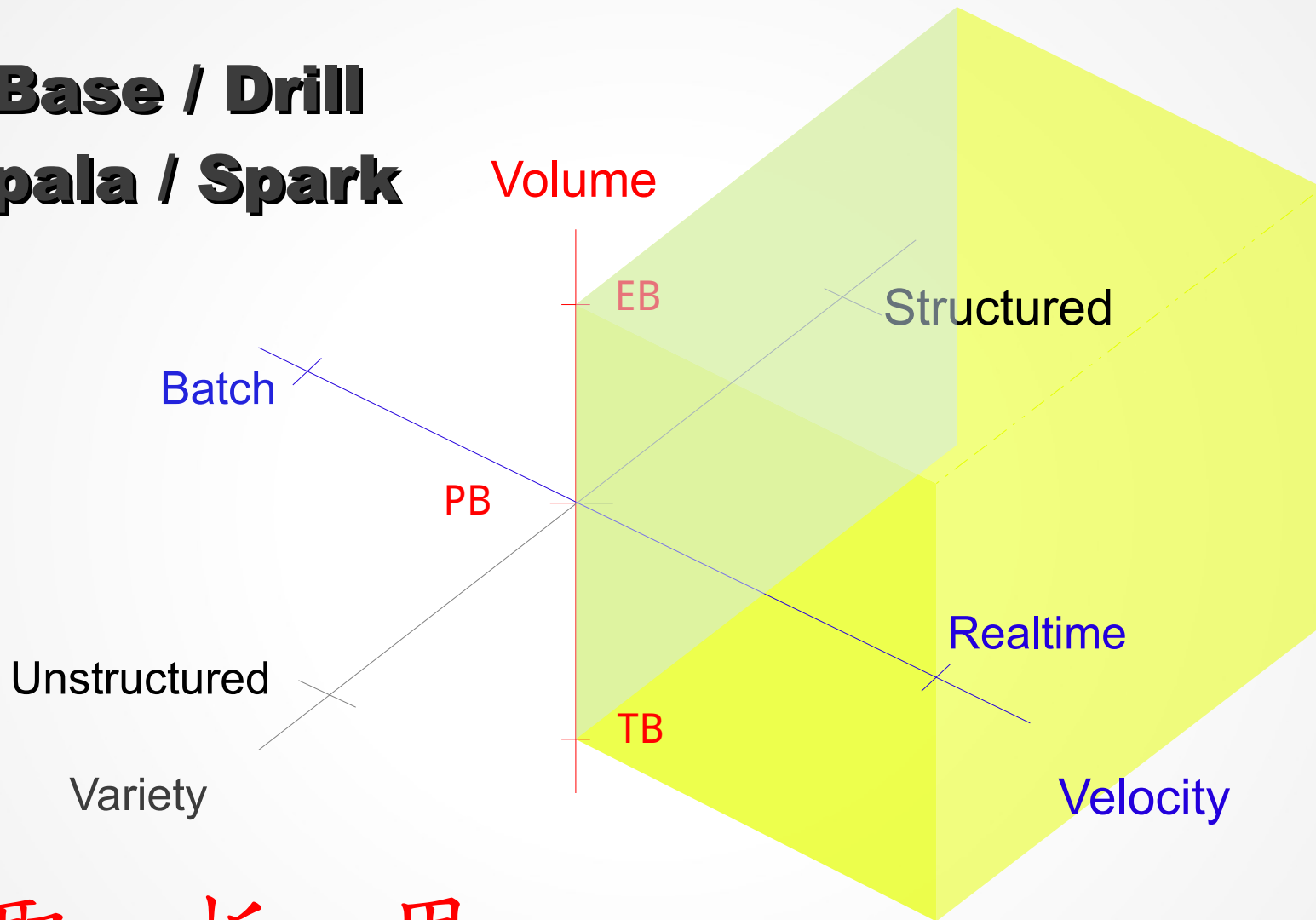
**Data Variety**

# Big Data at Rest – MapReduce Framework



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

**HBase / Drill**  
**Impala / Spark**



取、析、用

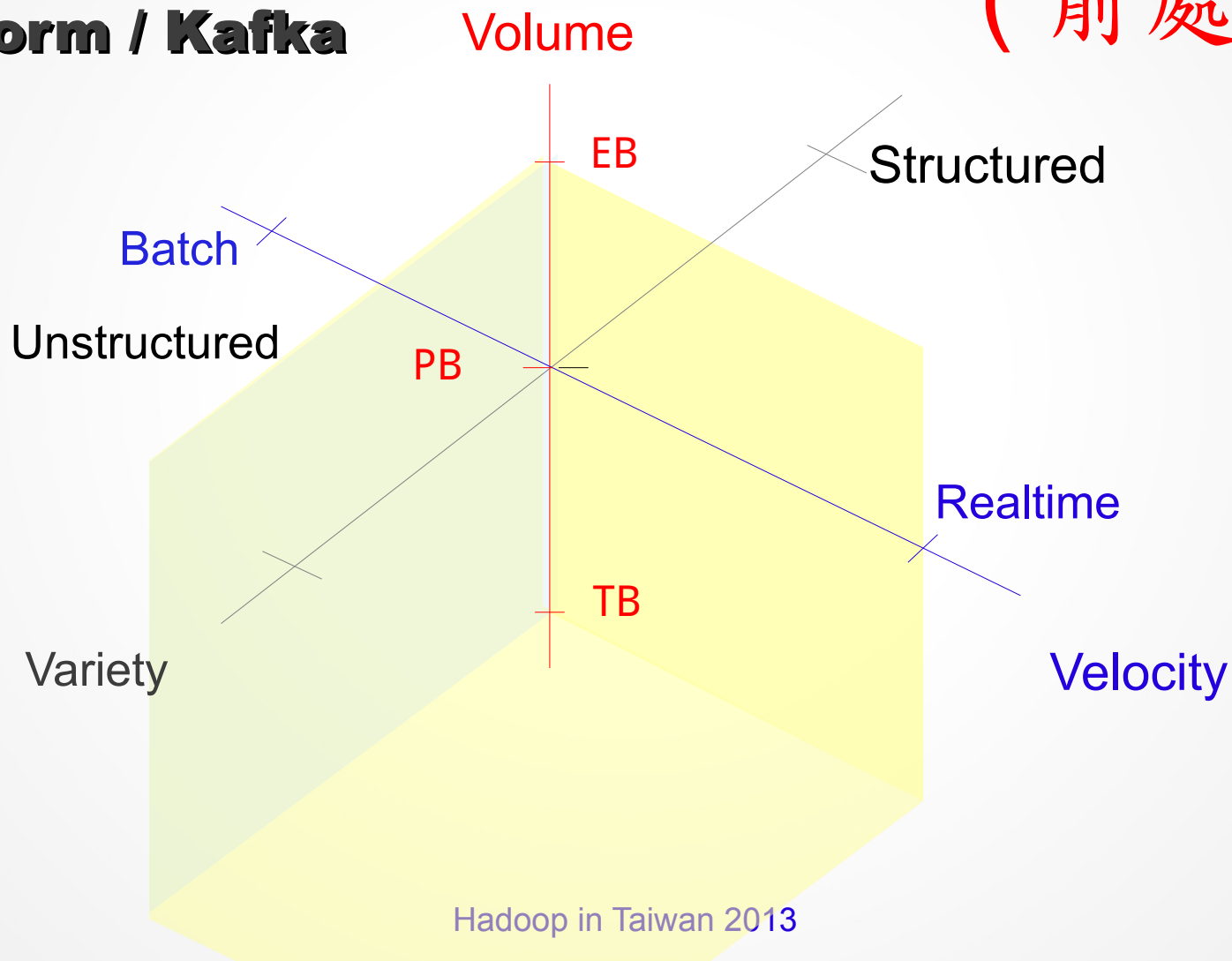
# Big Data in Motion – Streaming Data Collection / Data Cleaning

## Message Queue

( AMQP , RabbitMQ )

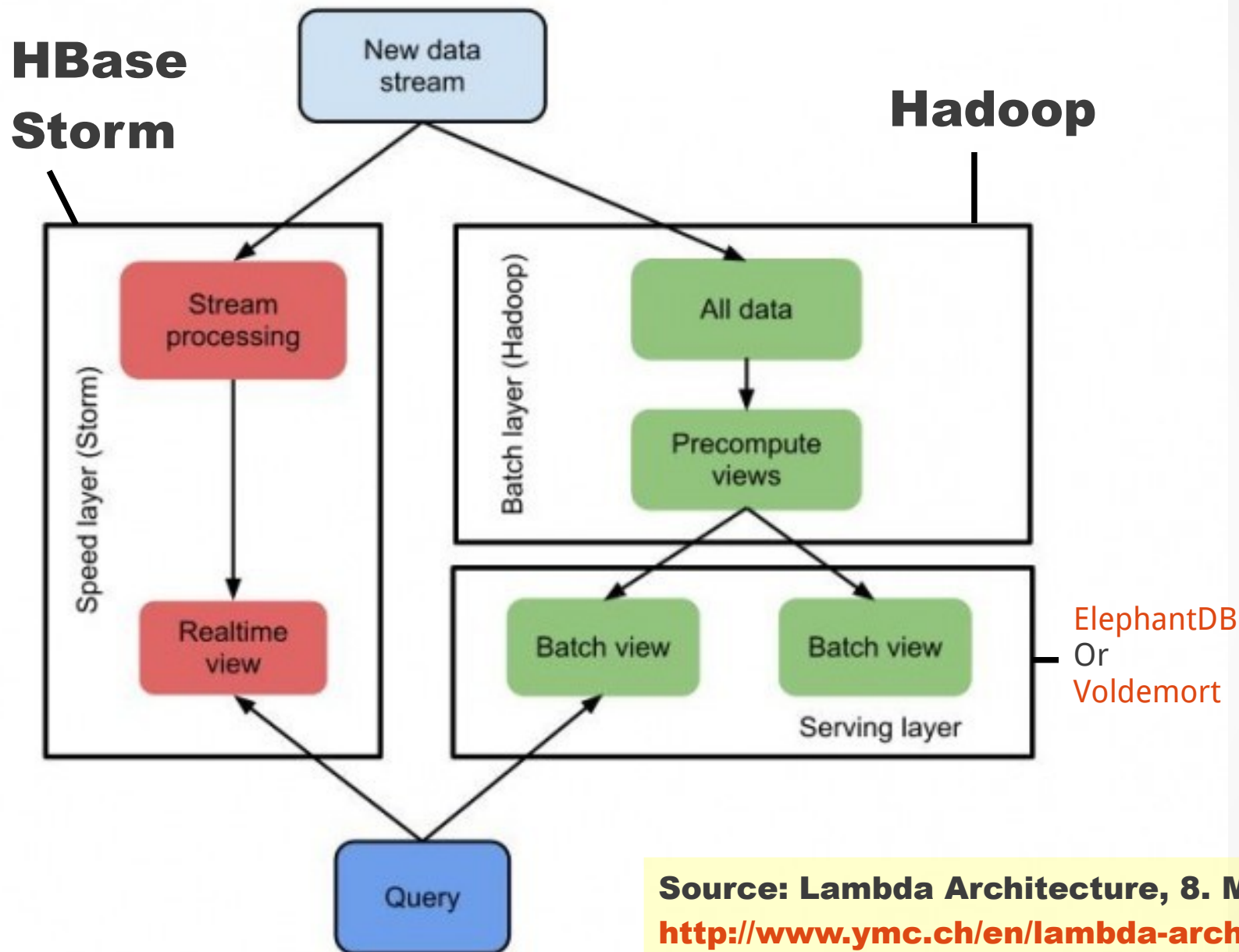
**Storm / Kafka**

蒐、存  
( 前處理 )





# NoHadoop ?! Not Only Hadoop !!



# 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

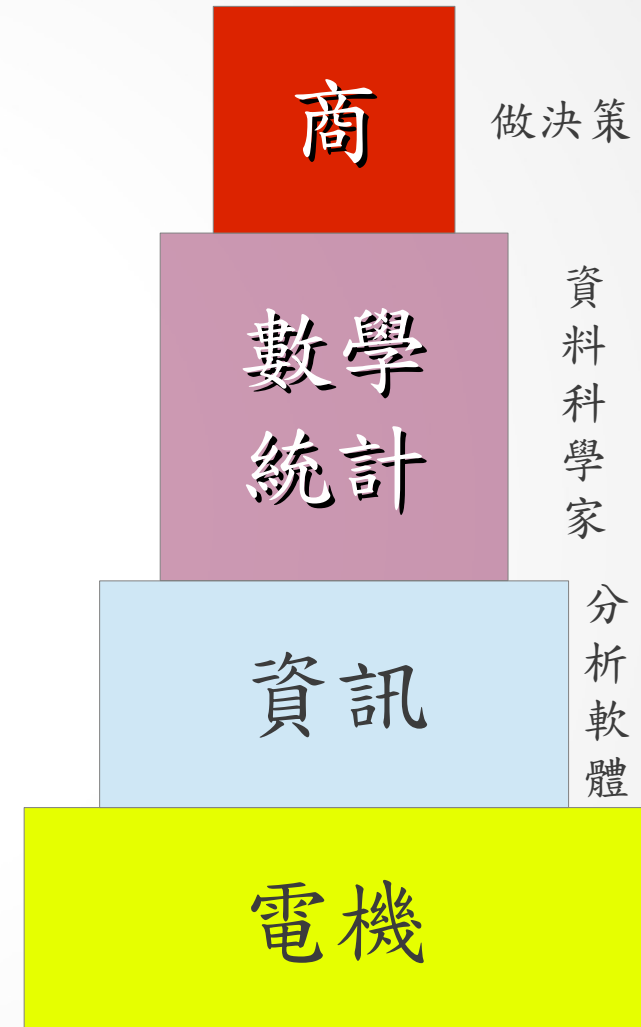
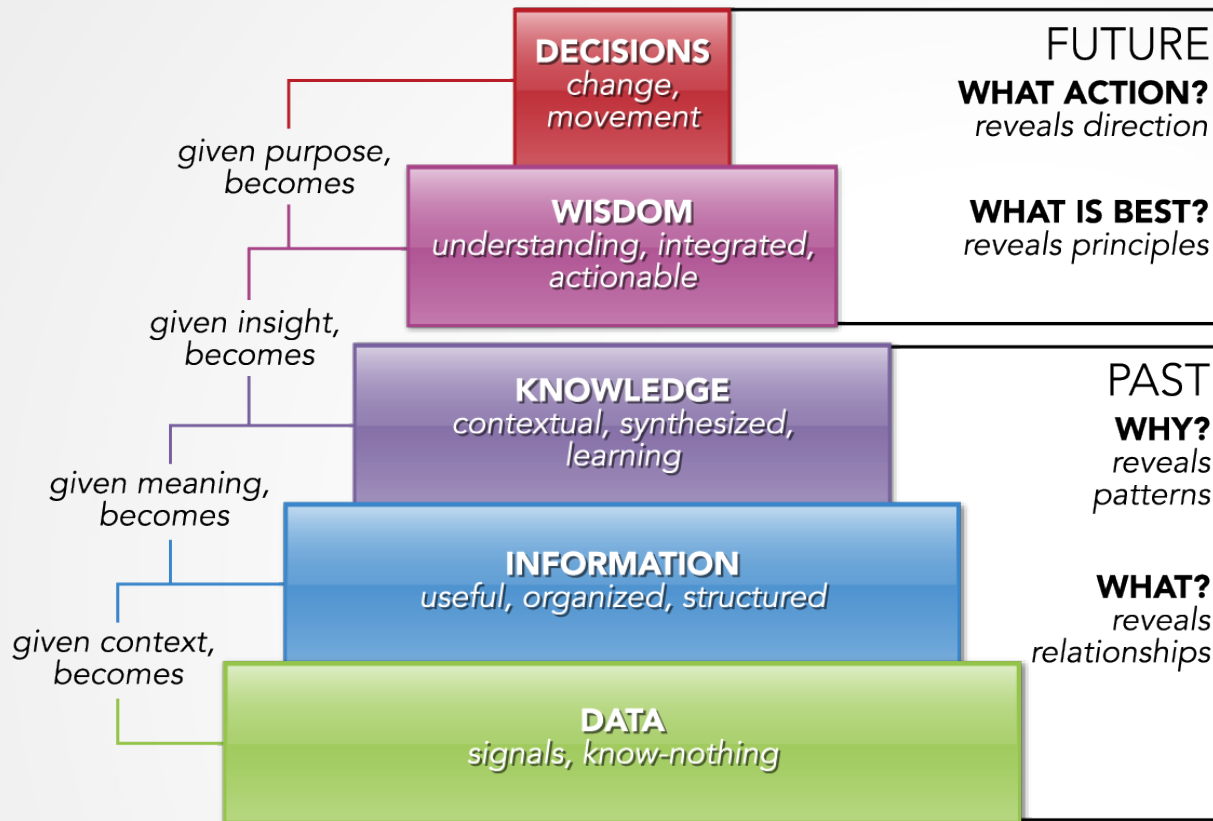
國家安全：美國 **PRISM** 計劃  
( 網軍！終極警探 4.0 )

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**

# 議程安排 (上午場次)

即時 · 安全 · 易用

時間	Keynote 議程-4F國際會議廳
08:30~09:00	來賓報到
09:00~09:10	<b>Opening</b>
09:10~09:50	<b>Architectural patterns for building real time applications with Apache HBase</b>   Andrew Purtell ( Apache HBase計畫提交人)
09:50~10:30	<b>Big Data Security</b>   Jason Yueh (Distinguished Technologist at Trend Micro)
10:30~10:50	休息 / 攤位參觀
10:50~11:30	<b>The Spark Stack: Lightning-Fast Analytics over Hadoop Data</b>   Jason Dai (Engineering Director & Principal Engineer, Intel Software and Services Group )
11:30~12:10	<b>基於 R 加Hadoop的高性能預測分析</b>   Liu Xiaohui (Revolution Analytics; Senior Developer)
12:10~13:30	午餐 / 攤位參觀

# 議程安排 (下午場次)

時間	Track A-Developer 「開發者」 4F國際會議廳(316人)	Track B-Administrator 「營運者」 3F第一會議室(113人)	Track C-Application 「應用案例」 3F第二會議室(113人)	廠商焦點對談 3F遠距會議室(36人) 每時段30分鐘
13:30~14:10	<b>Approaching real-time: things you can do before going Impala / 即時運算: 採用 Impala 之前你能做的事</b> Chris Huang / 趨勢科技	<b>Savanna - Hadoop跟 OpenStack的第一次接觸</b> macJack	跨時域巨量資安事件分析 毛敬豪 / 資策會	<b>時段A</b>
14:10~14:50	<b>High Speed HBase Query: Hare</b> 范姜冠宇 / 亦思科技	<b>HBase Operation</b> 劉一正 (Tony Liu) / 趨勢科技	<b>Apache Mahout 於電子商務的應用</b> 陳昭宇 / Etu	資策會 雲端系統軟體研究所
14:50~15:20	休息/攤位參觀			
15:20~16:00	用雲端實現巨量資料的商業智慧-HDInsight on <b>Azure</b> 張書源 / Microsoft全球策略夥伴推廣團隊	<b>Real-time Big Data Analytics Engine using Impala</b> Jason Shih / Etu	<b>Bowerbird: Real-time Pig Template for Big Data Query</b> 涂凱晏 (Kevin Twu) / 趨勢科技	ETU
16:00~16:40	<b>Real-time Big Data Processing with Storm: Using Twitter Streaming as Example</b> 謝良奇(Liang-Chi Hsieh) / 台大資工	<b>Hadoop Security: Now and future</b> 徐啓超 / TCloud Computing	無人空中載具(UAV)影像於雲端儲存與處理之實作及後續應用 林峰正 / 逢甲地理資訊中心	<b>時段D</b>