Department of Computer Science
University of Cyprus

EPL646 – Advanced Topics in Databases

Lecture 11

Big Data Management I

(Introduction)

Demetris Zeinalipour
http://www.cs.ucy.ac.cy/~dzeina/courses/epl646

Lecture Outline

- ***Big Data Definitions and Background***
- *Big Data Definition by 3V Examples*
  - ***Velocity***
    - *Sensor Monitoring, Network Monitoring, Web2.0 Media, Smartphone Services*
  - ***Volume***
    - *Text<Multimedia<Sciences, Web Data, Filesystems*
  - ***Variety***
    - *The New Database Landscape*
    - *NoSQL (Document Stores, Replication, Consistency, Map-Reduce, Column Stores)*
    - *NewSQL Trends*
- *Big Data Research Prototypes @ UCY*

Big Data Definitions

- *"Refers to data sets whose **size** and **structure strains** (stretches) the **ability** of commonly used **relational DBMSs** to **capture**, **manage**, and **process** the data within a **tolerable elapsed time**."*
  - *Hoffer, Ramesh, Topi: Modern Database Management, 11E, 2013.*
- *Similar from Wikipedia, Feb. 2013*
  - *"big data is a collection of data sets **so large** and **complex** that it becomes **difficult** to process **using on-hand database management** tools or **traditional data processing** applications."*

Big Data Characteristics

- **Size:** from a few **dozen terabytes** to many **petabytes** in a single database.
- **Data model:** anything from *structured* (relational or tabular) to *semi-structured* (XML or JSON) or even *unstructured* (Web text and log files).
- **Architectures:** highly *parallel* and *distributed* in order to cope with the inherent I/O and CPU limitations.
- **Hardware:** mid-scale *private clouds* (datacenters), offering higher privacy, to *large-scale public clouds*.
- **Functionality:** *operational (OLTP)* and *analytic (OLAP)* functionality *stand-alone* or *as-a-Service.*

Big Data Characteristics

2013 IEEE International Conference on Big Data (IEEE BigData 2013), October 6-9, 2013, Silicon Valley, CA, USA

Wordle.net

Background: Public Clouds

Google's Datacenter in Oregon

Background: Public Clouds

Microsoft Azure in Chicago

Background: *-as-a-Service

To Amazon RDS (Relational Database Service)

Pay by the hour your DB Instance runs.

US – N. Virginia | US – N. California | EU – Ireland | APAC – Singapore

DB Instance Class	Price Per Hour
Small DB Instance	$0.11
Large DB Instance	$0.44
Extra Large DB Instance	$0.88
Double Extra Large DB Instance	$1.55
Quadruple Extra Large DB Instance	$3.10

963$ / year

↓

27,165 $ / year

DB Instance Classes

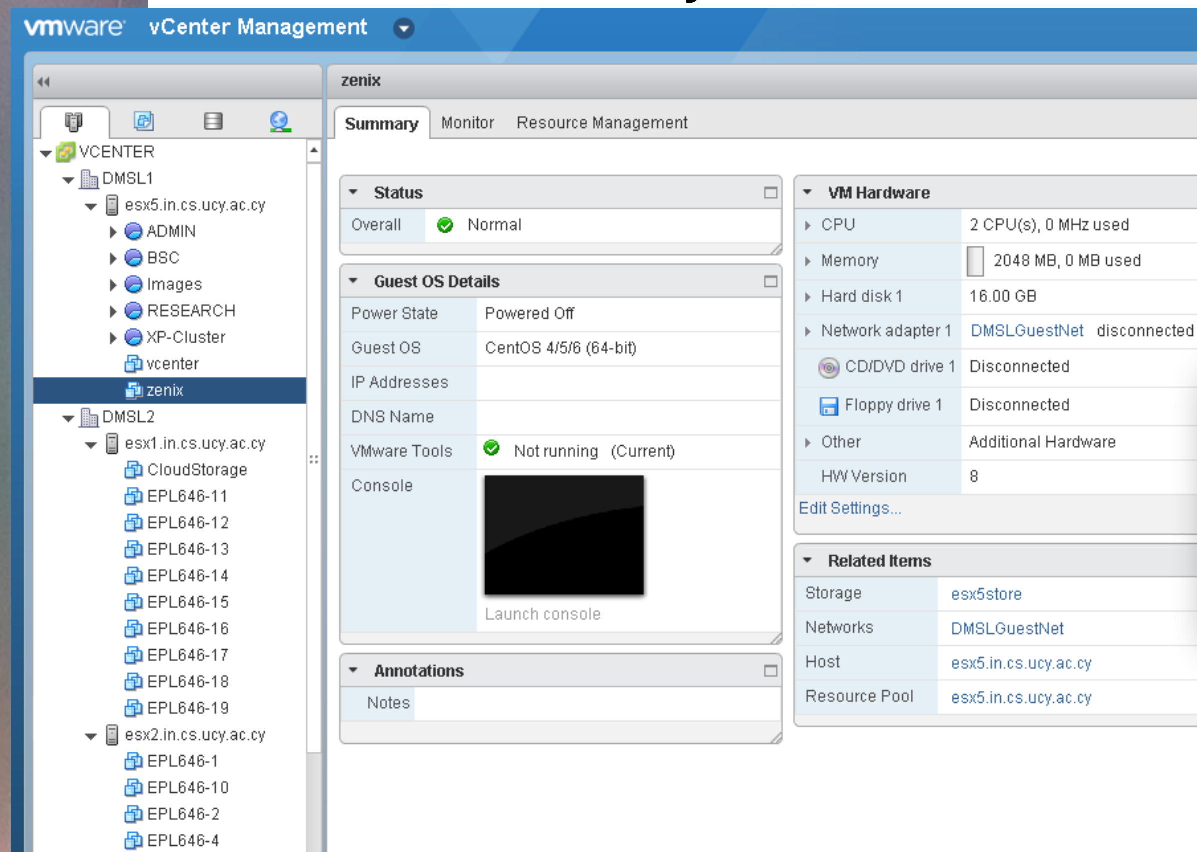
Amazon RDS currently supports five DB Instance Classes:

- Small DB Instance: 1.7 GB memory, 1 ECU (1 virtual core with 1 ECU), 64-bit platform, Moderate I/O Capacity
- Large DB Instance: 7.5 GB memory, 4 ECUs (2 virtual cores with 2 ECUs each), 64-bit platform, High I/O Capacity
- Extra Large DB Instance: 15 GB of memory, 8 ECUs (4 virtual cores with 2 ECUs each), 64-bit platform, High I/O Capacity
- Double Extra Large DB Instance: 34 GB of memory, 13 ECUs (4 virtual cores with 3,25 ECUs each), 64-bit platform, High I/O Capacity
- Quadruple Extra Large DB Instance: 68 GB of memory, 26 ECUs (8 virtual cores with 3.25 ECUs each), 64-bit platform, High I/O Capacity

For each DB Instance class, RDS provides you with the ability to select from 5GB to 1TB of associated storage capacity.
One ECU provides the equivalent CPU capacity of a 1.0-1.2 GHz 2007 Opteron or 2007 Xeon processor.

Background: Private Clouds

Our Laboratory Private IaaS



Big Data: Velocity-Volume-Variety

- **Velocity**
  - ***how fast data*** *is being* ***produced*** *and* ***how fast*** *the data must be* ***processed to meet demand****.*
    - *How to deal with torrents of data, in near-real time, streaming from* ***RFID tags*** *and* ***smart metering systems****?*
    - *How to* ***identify fraud*** *in 5 million trade events created* ***each day****?*
    - **Reacting** quickly **enough** to deal with velocity is a **challenge** to most organizations.

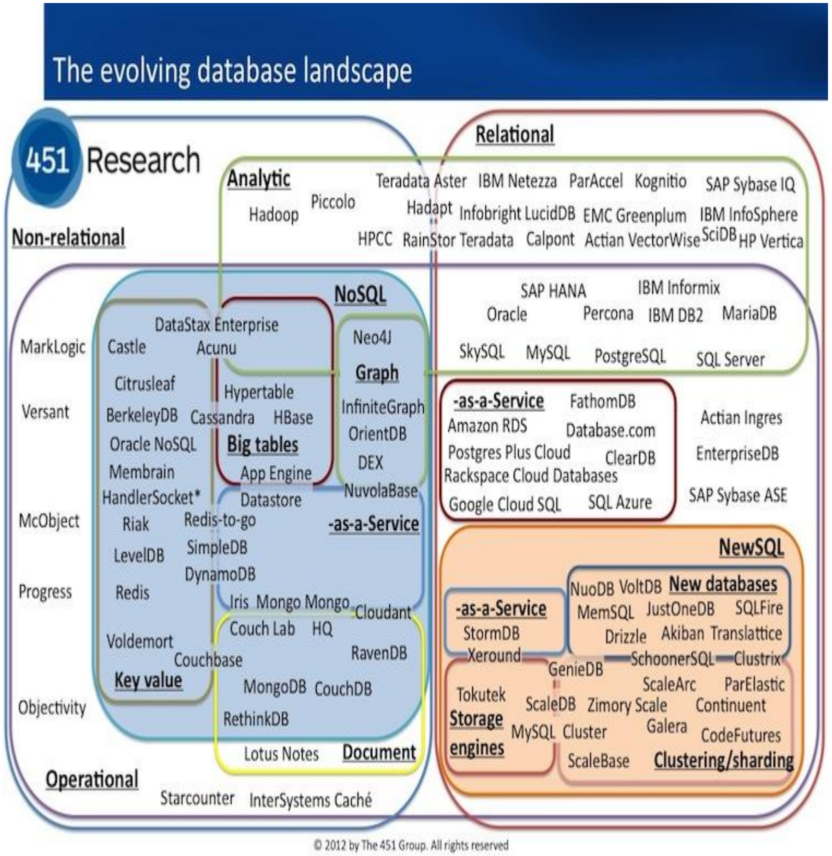
Source: IDC. "Big Data Analytics: Future Architectures, Skills and Roadmaps for the CIO," September 2011.

Big Data: Velocity-Volume-Variety

- **Volume**
  - **Past Challenge: Store data.**
    - transaction-based data stored through the years.
    - sensor data being collected
    - Integration with web applications & social media
  - **New Challenge: Create value from data**
    - Turn 12 TB of Tweets each day into a *sentiment analysis* (opinion mining) product.
      - e.g., People feel positive/negative/neutral about brand X.
    - Turn 350 billion annual smart meter readings to knowledge that helps predicting power consumption.

Big Data: Velocity-Volume-Variety

- **Variety:**
  - By some estimates, **80 percent** of an organization's data is **not numeric**!
  - Different data format: *unstructured, structured, semi-structured*
    - text, sensor data, audio, video, click streams, log files, etc.



Velocity #1: Smart Meters

- **Smart meter:** records consumption of **electric energy** in **intervals** and communicates that information to the utility for **monitoring** and **billing purposes.**

Every 15m

Velocity #1: Smart Meters

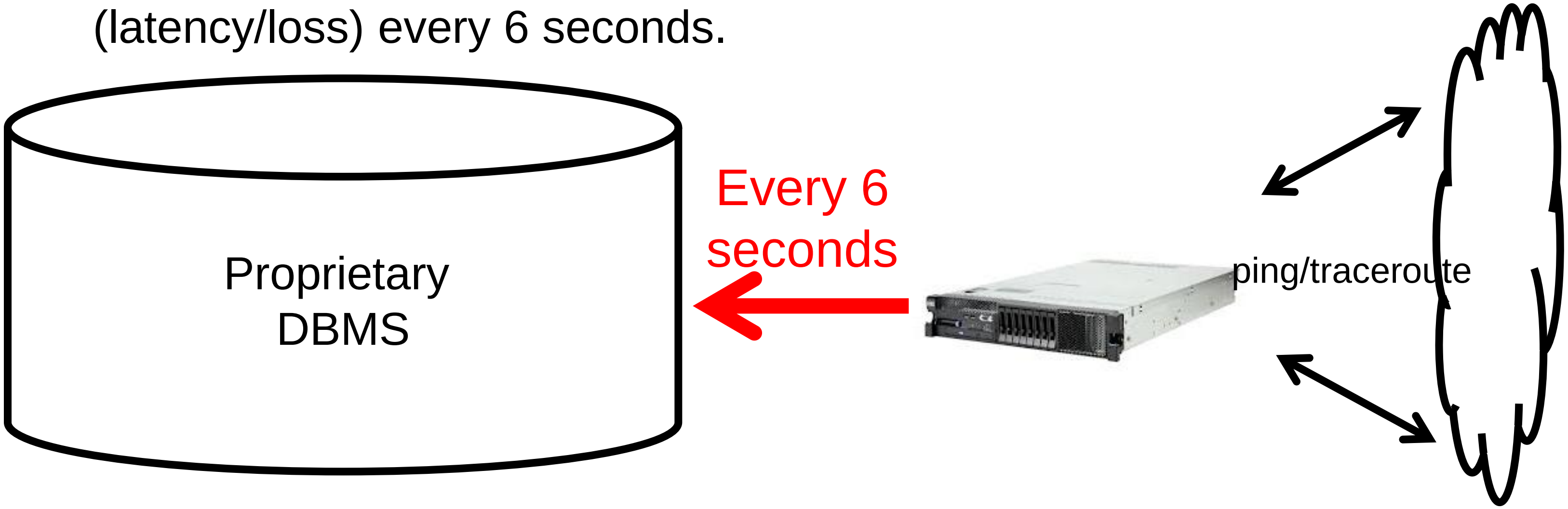
- **Ontario's Meter Data Management and Repository (MDM/R):** storing, processing and managing all smart meter data in Ontario, Canada
- **Characteristics:**
  - Provides hourly billing quantity and extensive reports.
  - 4.6 million smart meters.
    - Storage/Bandwidth: 4.6M meters x 0.5K message (typical HTTP) = 2.3 GB / round
  - 110 million meter reads per day
    - on an annual basis, exceeds the number of debit card transactions processed in the **country** (Canada!)

Source: Smart Metering Entity: http://www.smi-ieso.ca/mdmr

Velocity #2: Network Monitoring

- **Akamai:**
  - CDN serving 15-30% of all Web traffic (10TB/sec)
    - One out of every three Global 500® companies
    - All of the top Internet portals
  - Has a picture of the global traffic every **6 seconds**

- **How?**
  - 119,000 servers in 80 countries within over 1,100 networks.
  - Servers report to a proprietary database network health information (latency/loss) every 6 seconds.



Velocity #2: Network Monitoring

Big Data Akamai Jobs – Bosto | Careers Center – More infor

https://jobs-akamai.icims.com/jobs/6493/senior-system-software-engineer/job

Location:	US-MA-Cambridge	**Posted Date:**	8/23/2012
Cost Center:	215	**Category:**	Engineering
ID:	6493		

Companies started seeking Big data engineers.

Apply for this job:

Your application choices are:

- Apply for this job online

If you would like to include a cover letter, please be sure combine your cover letter and resume into one document.

More information about this job:

Overview:

About the Job
Be a Big Data software engineer. Extract statistics to provide insight into usage logs collected by Akamai's edge services. You will use our extensible distributed processing cluster to parse, aggregate and generate reports from logfiles with volume of more than Peta-Byte per day. You will work with the QA team to ensure that the resulting data products are highly accurate and available to all consumers.

Velocity #3: Web2.0 Media

- Analyze online conversations in Social Nets.
- Accelerated responses to marketplace shifts.

Continously

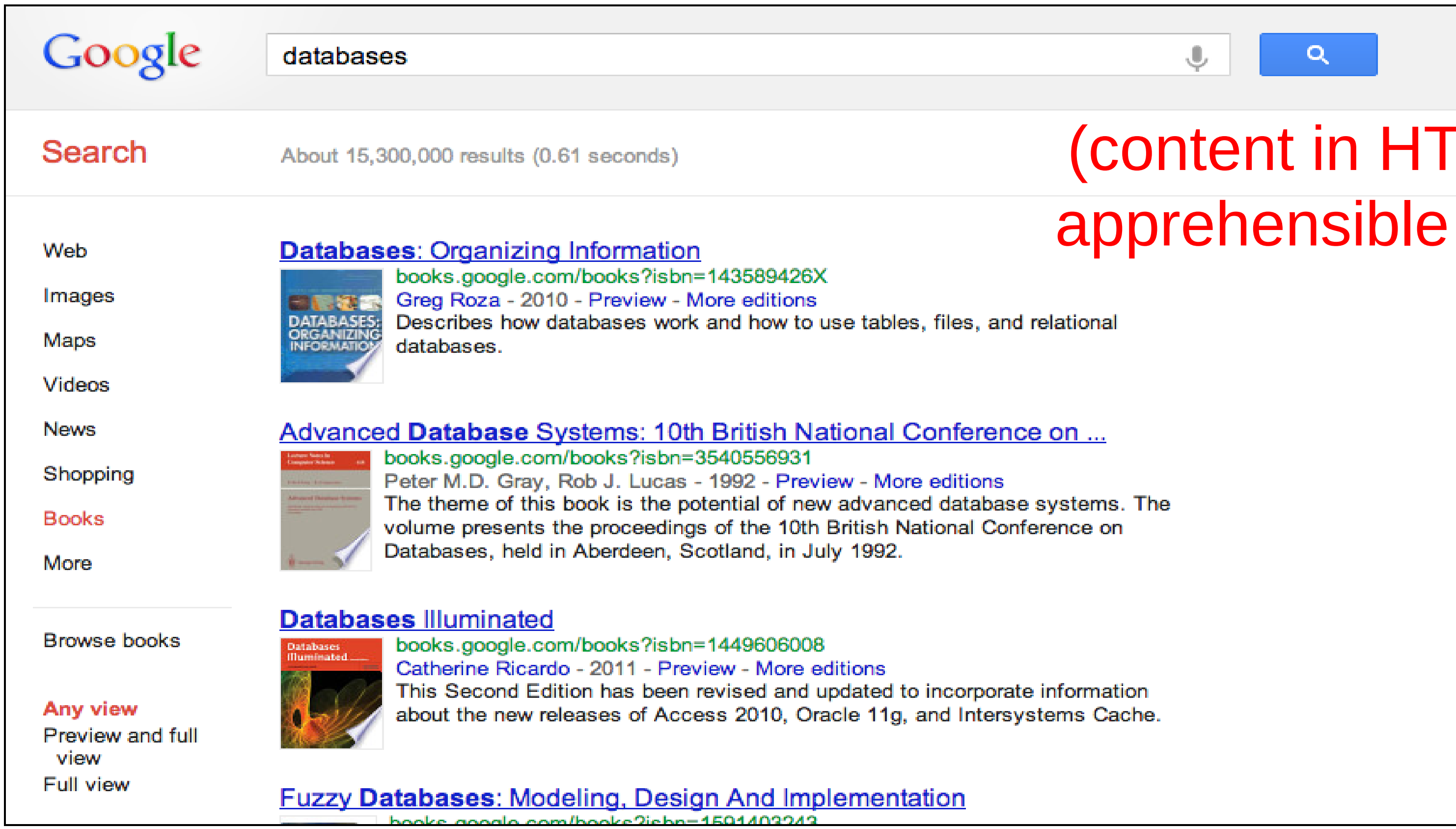
←

Over Web2.0 protocols

Velocity #3: Web2.0 Media

Web1.0: The Unstructured Web

http://books.google.com/



(content in HTML only apprehensible to User)

Velocity #3: Web2.0 Media

Web2.0: The Semi-structured Web!

https://www.googleapis.com/books/v1/volumes?q=databases

https://www.googleapis.com/books/v1/volumes?q=databases

content in XML/JSON apprehensible to Computer


```
{
 "kind": "books#volumes",
 "totalItems": 899,
 "items": [
  {
   "kind": "books#volume",
   "id": "4Z6tfpuBmmgC",
   "etag": "urgGiT9Q1G4",
   "selfLink": "https://www.googleapis.com/books/v1/volumes/4Z6tfpuBmmgC",
   "volumeInfo": {
    "title": "Databases",
    "subtitle": "Organizing Information",
    "authors": [
     "Greg Roza"
    ],
    "publisher": "Rosen Central",
    "publishedDate": "2010-08-15",
    "description": "Describes how databases work and how to use tables, files, and relational databases.",
    "industryIdentifiers": [
     {
      "type": "ISBN_10",
      "identifier": "143589426X"
     },
     {
      "type": "ISBN_13",
      "identifier": "9781435894266"
     }
    ],
    "pageCount": 48,
    "printType": "BOOK",
    "contentVersion": "preview-1.0.0",
    "imageLinks": {
     "smallThumbnail": "http://bks3.books.google.com/books?id=4Z6tfpuBmmgC&printsec=frontcover&img=1&zoom=5&edge=cu
     "thumbnail": "http://bks3.books.google.com/books?id=4Z6tfpuBmmgC&printsec=frontcover&img=1&zoom=1&edge=curl&so
    },
    "language": "en",
    "previewLink": "http://books.google.com/books?id=4Z6tfpuBmmgC&printsec=frontcover&dq=databases&hl=&cd=1&source=
    "infoLink": "http://books.google.com/books?id=4Z6tfpuBmmgC&dq=databases&hl=&source=gbs_api",
```

Velocity #3: Web2.0 Media

Twitter API

https://twitter.com/users/dmslucy.json

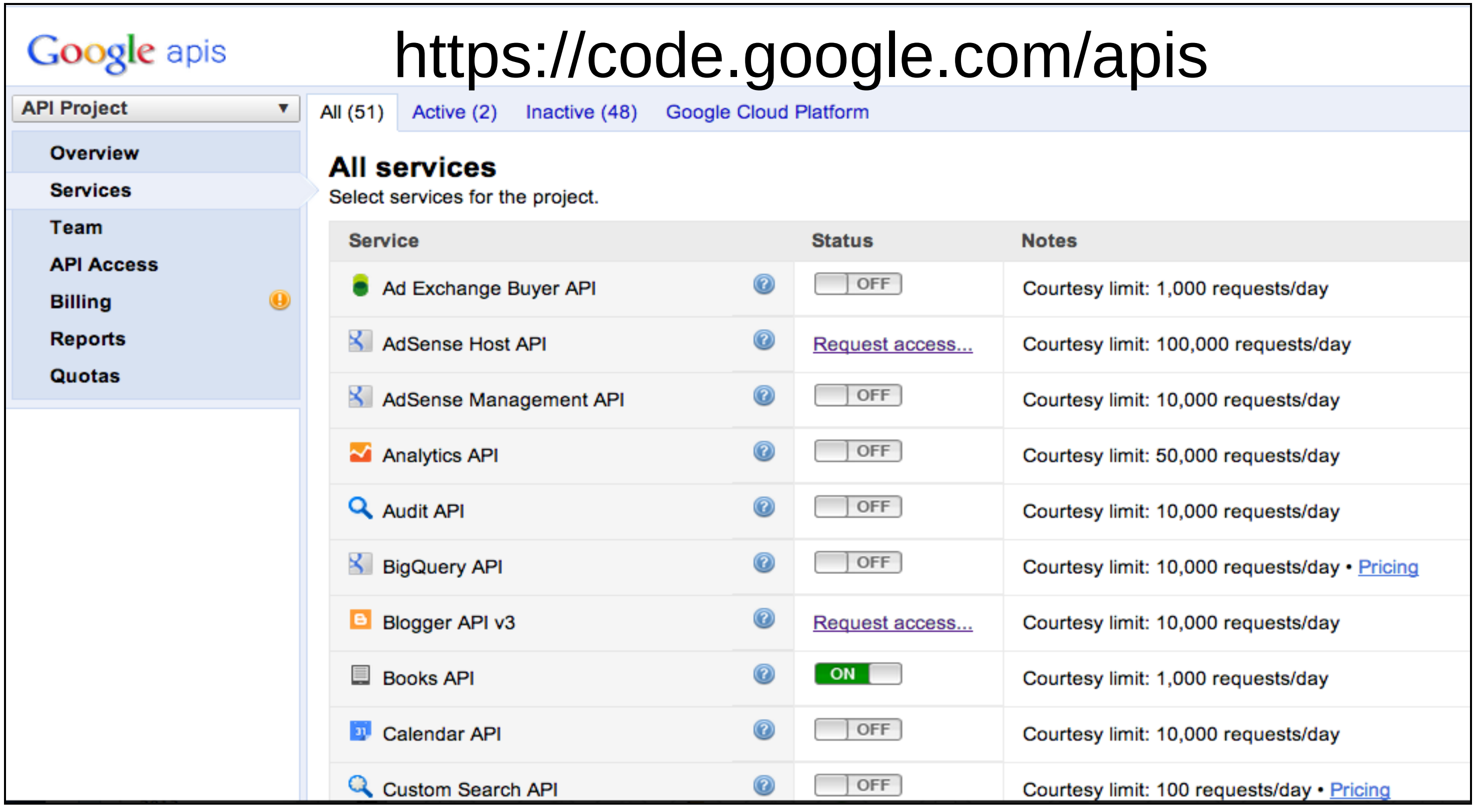

```
- {
    "id":742558014,
    "follow_request_sent":null,
    "following":null,
    "screen_name":"DMSLUCY",
    "url":"http:\/\/dmsl.cs.ucy.ac.cy\/",
    "profile_use_background_image":true,
    "created_at":"Tue Aug 07 09:36:30 +0000 2012",
    "profile_text_color":"333333",
    "utc_offset":7200,
    "statuses_count":10,
    "default_profile_image":false,
    "verified":false,
    "name":"DMS Laboratory, UCY",
    "favourites_count":10,
    "profile_sidebar_border_color":"C0DEED",
    "friends_count":0,
    "profile_image_url_https":"https:\/\/si0.twimg.com\/profil
    e_images\/2728729106\/130bc7921970a06228d1ad0d352260de_nor
    mal.png",
    "description":"DMSL belongs to the Computer Science
    Department at the University of Cyprus. We focus on Data
    Engineering Systems and Knowledge Discovery Solutions. ",
    "profile_image_url":"http:\/\/a0.twimg.com\/profile_images
    \/2728729106\/130bc7921970a06228d1ad0d352260de_normal.png"
    ,
```

Velocity #3: Web2.0 Media

In fact, Web2.0 Services are omnipresent!

(Google, Twitter, Facebook, Youtube, Linkedin, …)

http://www.programmableweb.com/ - 7800 APIs!!! + 6800 Mashups!



Velocity #4: Smartphone Services

Request Format (request.json)

```
{
 "homeMobileCountryCode": 310,
 "homeMobileNetworkCode": 260,
 "radioType": "gsm",
 "carrier": "T-Mobile",
 "cellTowers": [
  {
   "cellId": 39627456,
   "locationAreaCode": 40495,
   "mobileCountryCode": 310,
   "mobileNetworkCode": 260,
   "age": 0,
   "signalStrength": -95
  }
 ],

}
```

```
"wifiAccessPoints": [
 {
  "macAddress": "01:23:45:67:89:AB",
  "signalStrength": 8,
  "age": 0,
  "signalToNoiseRatio": -65,
  "channel": 8
 },
 {
  "macAddress": "01:23:45:67:89:AC",
  "signalStrength": 4,
  "age": 0
 }
]
```

Response Format

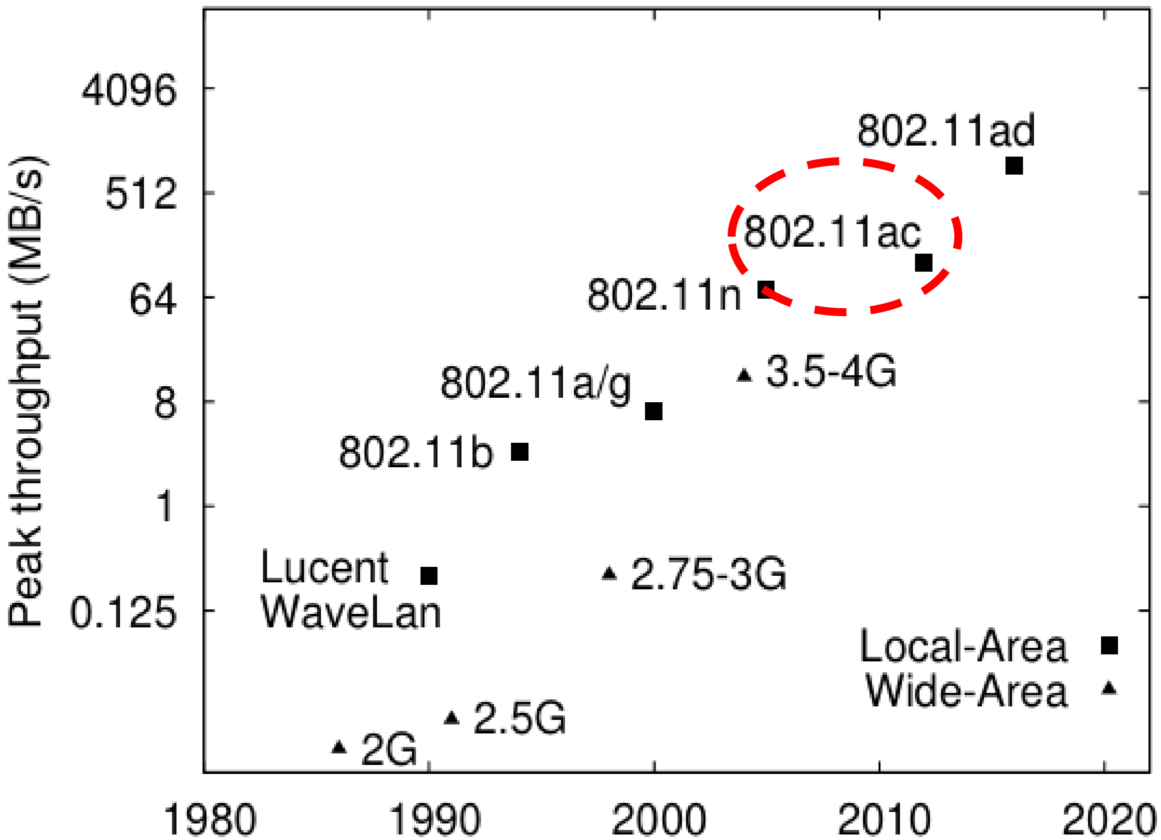
The response format is also JSON.

```
{
 "location": {
 "latitude": 51.0,
 "longitude": -0.1,
 },
 "accuracy": 1200.4,
}
```

Will be discussing some furtherin-house applications in a while

Velocity #4: Smartphone Services

Wireless Data Transfer Rates



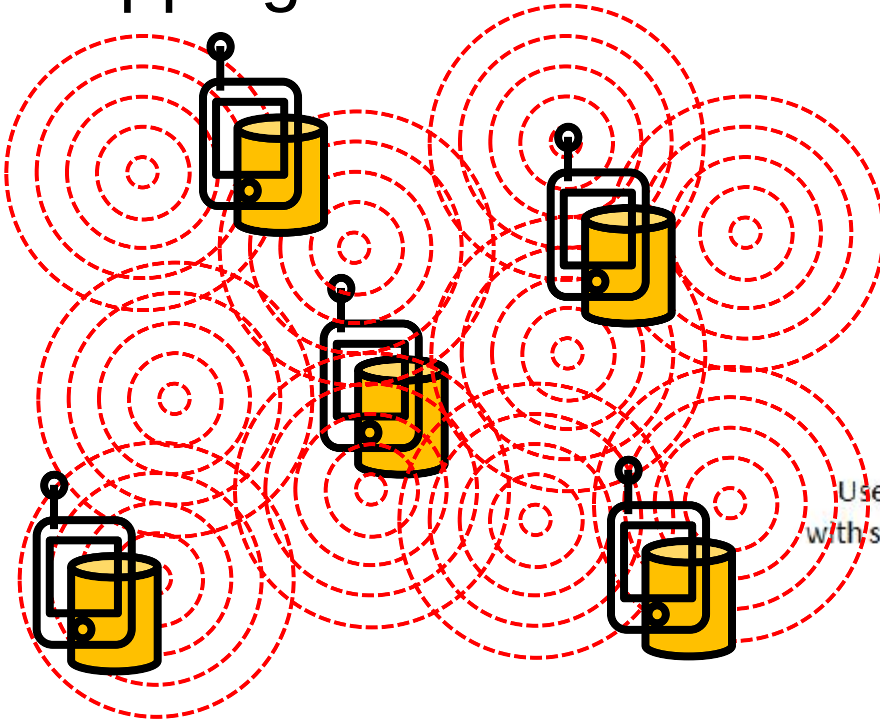
4G ITU peak rates:

- 100 Mbps (high mobility, such as trains and cars)
- 1Gbps (low mobility, such as pedestrians and stationary users)

*Plot Courtesy of H. Kim, N. Agrawal, and C. Ungureanu, "Revisiting Storage for Smartphones", The 10th USENIX Conference on File and Storage Technologies (FAST'12), San Jose, CA, February 2012. *** Best Paper Award ****

Velocity #4: Smartphone Services

Mapping the Road traffic by collecting WiFi signals.



Received Signal Strength (RSS):
power present in WiFi radio signal

Graphics courtesy of: A .Thiagarajan et. al. "Vtrack: Accurate, Energy-Aware Road Traffic Delay Estimation using Mobile Phones, In Sensys'09, pages 85-98. ACM, (Best Paper) MIT's CarTel Group

Volume #1: Text<Multimedia<Sciences

- **From the TB-era to the PB-era.**

Human Generated

  - The U.S. Library of Congress (April 2011): **235 TB**
  - Anchestry.com: Genealogical data **600 TB**

Multimedia/ Streaming

  - **Games:** World of Warcraft uses **1.3 PB** of storage to maintain its game.
  - **Internet Video:** will account for 61% of total Internet Data by 2015 (**966 Exabytes** or nearly 1 **Zettabyte!)**

Sciences/ Sensors

  - **Climate science:** The German Climate Computing Centre (DKRZ) has a storage capacity of **60 PB** of climate data.
  - **Physics:** The experiments in the Large Hadron Collider produce about **15 PB** of data per year, which is distributed over the LHC Computing Grid (Our department is part of the EGEE – Enabling Grids for E-sciencE, now EGI - European Grid Infrastructure).

Source: Petabyte, from Wikipedia: http://en.wikipedia.org/wiki/Petabyte

Volume #2: Web Data

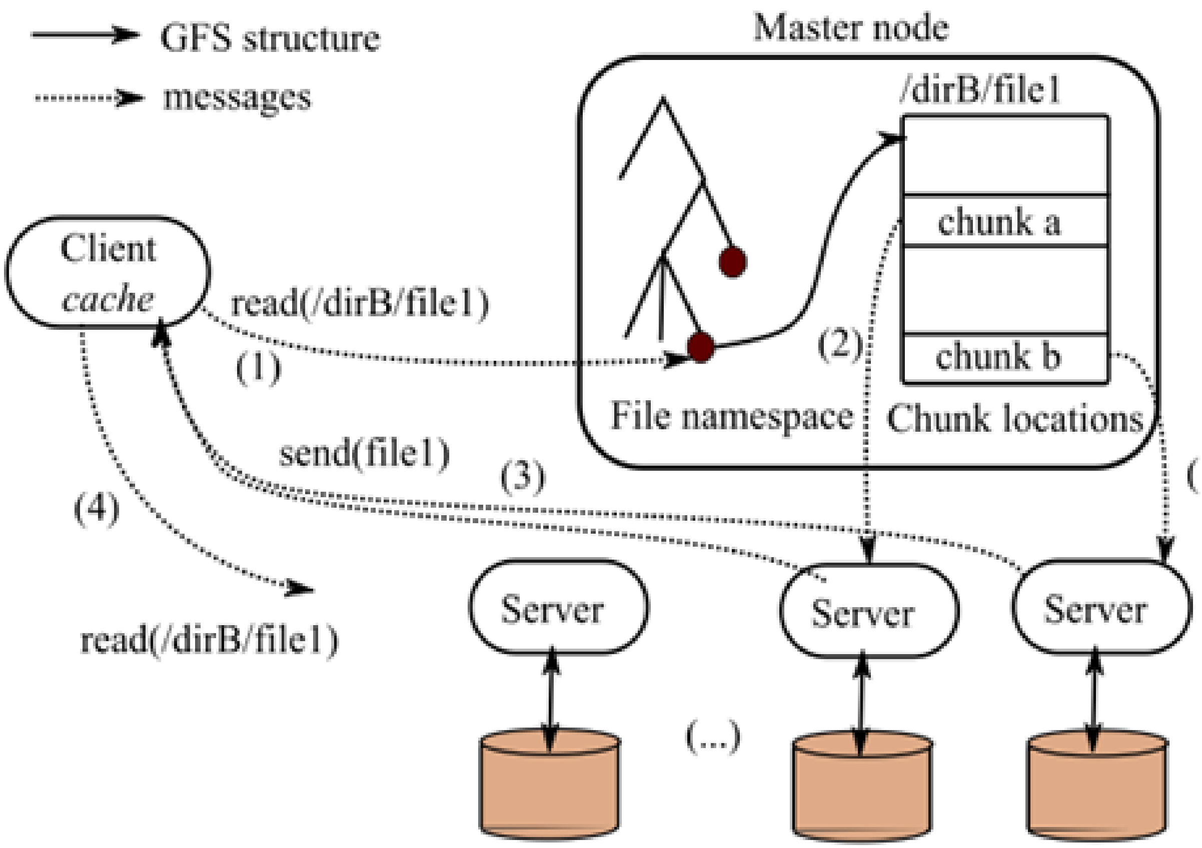
Google Volume (in 2006)
IDC: The total amount of global data is expected to grow to 2.7 zettabytes during 2012. This is 48% up from 2011. http://en.wikipedia.org/wiki/Zettabyte

Project name	Table size (TB)	Compression ratio	# Cells (billions)	# Column Families	# Locality Groups	% in memory	Latency-sensitive?
Crawl	800	11%	1000	16	8	0%	No
Crawl	50	33%	200	2	2	0%	No
Google Analytics	20	29%	10	1	1	0%	Yes
Google Analytics	200	14%	80	1	1	0%	Yes
Google Base	2	31%	10	29	3	15%	Yes
Google Earth	0.5	64%	8	7	2	33%	Yes
Google Earth	70	–	9	8	3	0%	No
Orkut	9	–	0.9	8	5	1%	Yes
Personalized Search	4	47%	6	93	11	5%	Yes

Bigtable: A Distributed Storage System for Structured Data,
OSDI'06: Seventh Symposium on Operating System Design and Implementation, Seattle, WA, November, 2006.

Volume #3: Big Data File Systems

- ***Big Data Filesystems: HDFS***



Namespace lookup are fast (1 Master enough!) [1GB Metadata = 1PB Data]

In NFS Metadata + Transfers going through same server => Not Scalable

HDFS designed for unreliable hardware (2-3 failures / 1000 nodes / day)

Volume #3: Big Data File Systems

- ***Big Data Filesystems: How Big?***
- Results from 2010:

YAHOO!

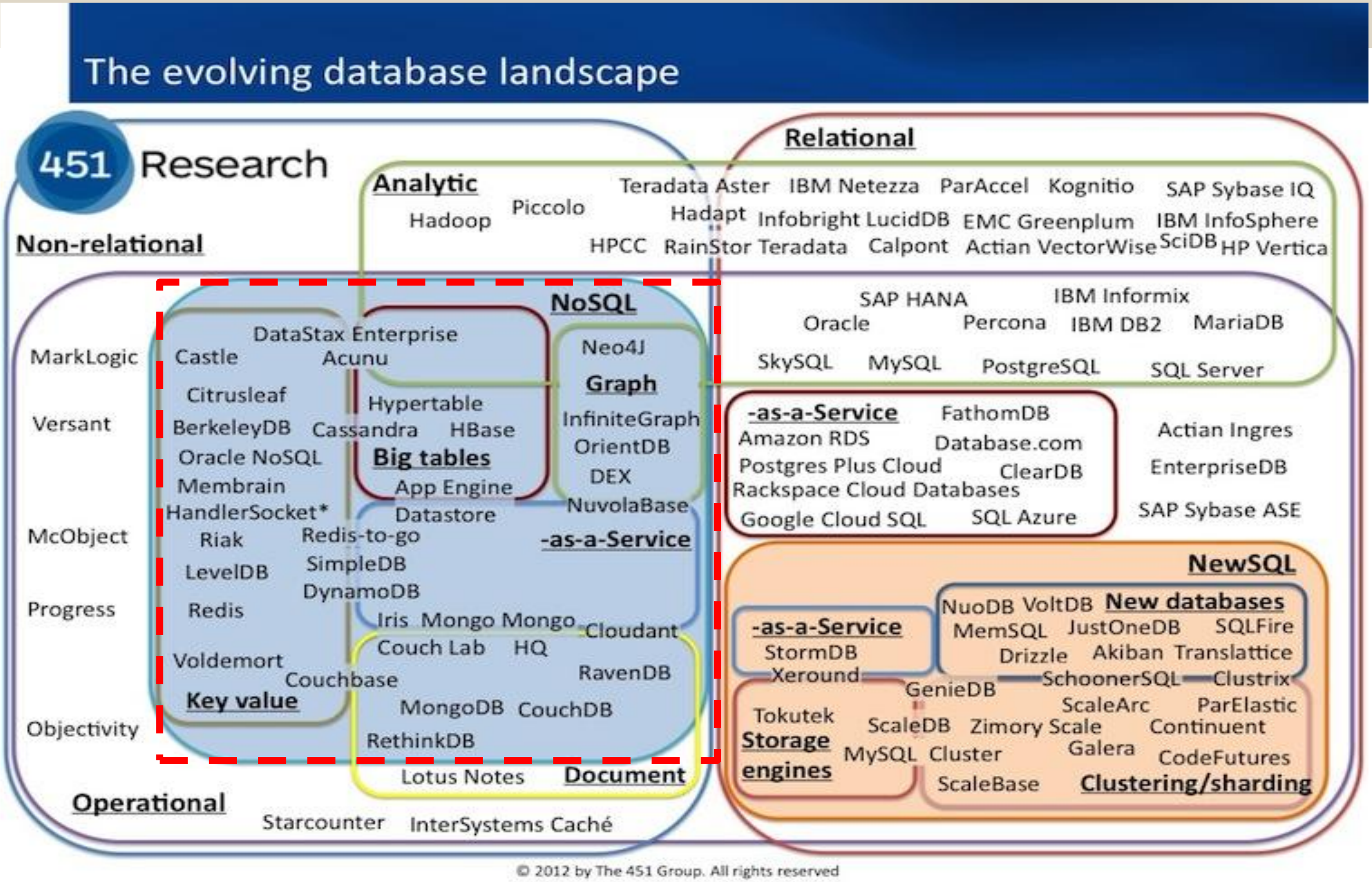
	Target	Deployed
Capacity	10PB	14PB
Nodes	10,000	4000
Clients	100,000	15,000
Files	100,000,000	60,000,000

facebook

- 21 PB of storage in a single HDFS cluster
- 2000 machines
- 12 TB per machine (a few machines have 24 TB each)
- 1200 machines with 8 cores each + 800 machines with 16 cores each
- 32 GB of RAM per machine
- 15 map-reduce tasks per machine

HDFS scalability: the limits to growth
http://static.usenix.org/publications/login/2010-04/openpdfs/shvachko.pdf

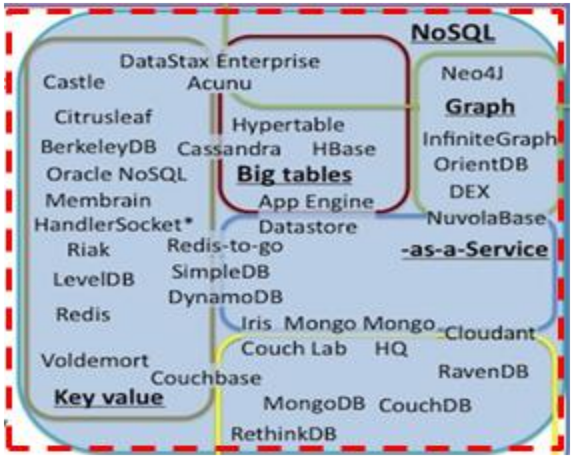
Variety Overview



451 Research, Matthew Aslett, http://goo.gl/GYcEx

Variety #1: NoSQL

- ***NoSQL ("not only SQL")** is a broad class of database management systems identified by n**on-adherence** to the widely used **relational database management system model.***
- **NoSQL** databases are **NOT built** primarily on **tables**, and generally **DO NOT** use SQL for data.
- NoSQL => Not Relational!
  - Key Value (e.g., BerkeleyDB – emb, Oracle NoSQL - Distributed)
  - Document Stores (e.g., JSON stores)
  - BigTables (i.e., Column-stores)
  - Graph Databases (e.g., FlockDB)
  - … potentially much longer list but I will only focus on a few trends



Variety #1: NoSQL / Document Stores

Document in CouchDB

```
{
   "_id": "book10.json",
   "_rev": "1-d0cc2ae0ab3211314a65a5c5244df221",
   "type": "Book",
   "title": "The AWK Programming Language",
   "year": "1988",
   "publisher": "Addison-Wesley",
   "authors": [
       "Alfred V. Aho",
       "Brian W. Kernighan",
       "Peter J. Weinberger"
   ],
   "source": "DBLP"
}
```

Map Function

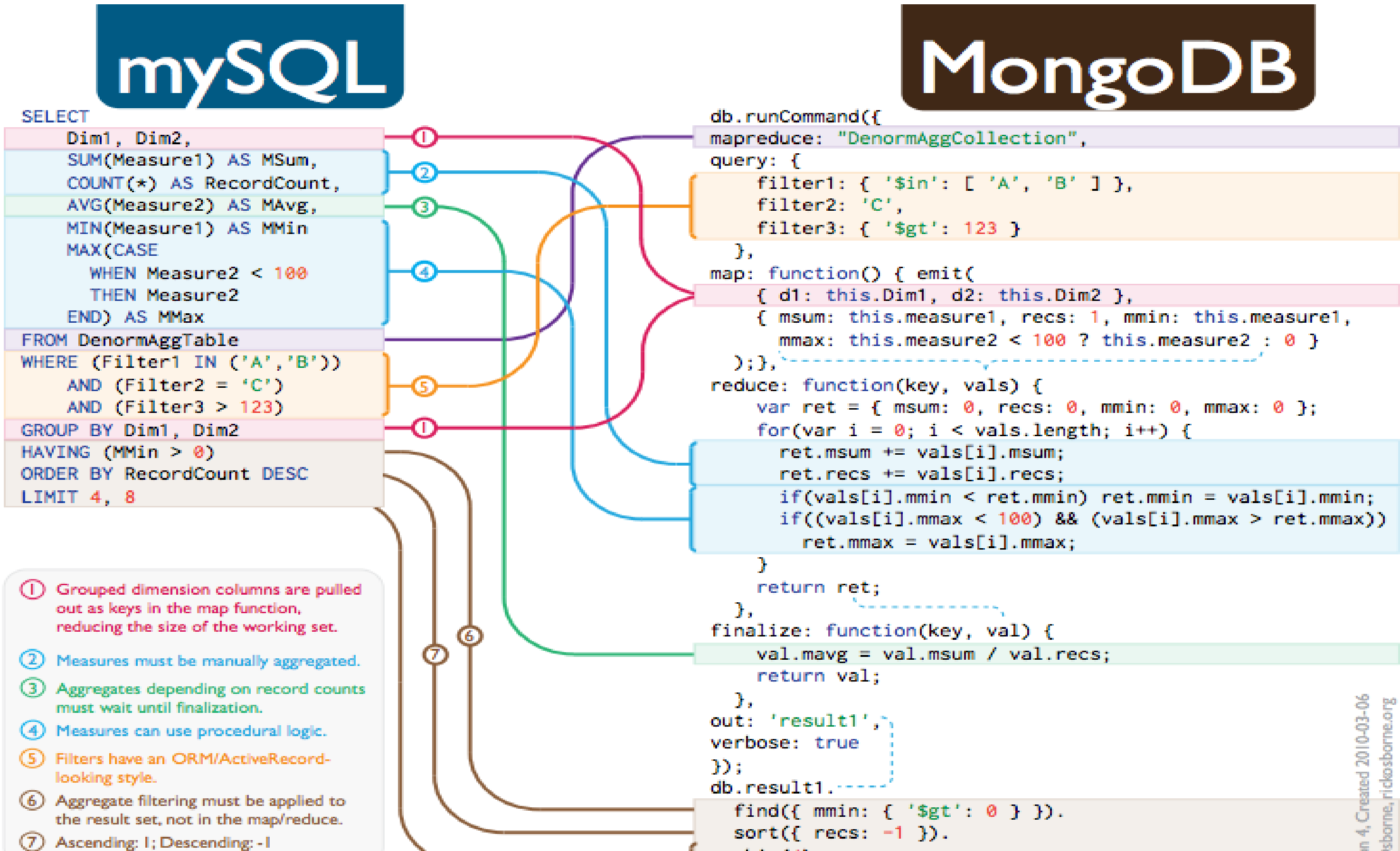
```
function(doc) {
    for (i in doc.authors) {
      author = doc.authors[i];
      emit(doc._id, author);
    }
}
```

Results (through REST/HTTP or Futon)

Key ▲ Grouping: exact	Value
"book10.json" ID: book10.json	"Alfred V. Aho"
"book10.json" ID: book10.json	"Brian W. Kernighan"
"book10.json" ID: book10.json	"Peter J. Weinberger"

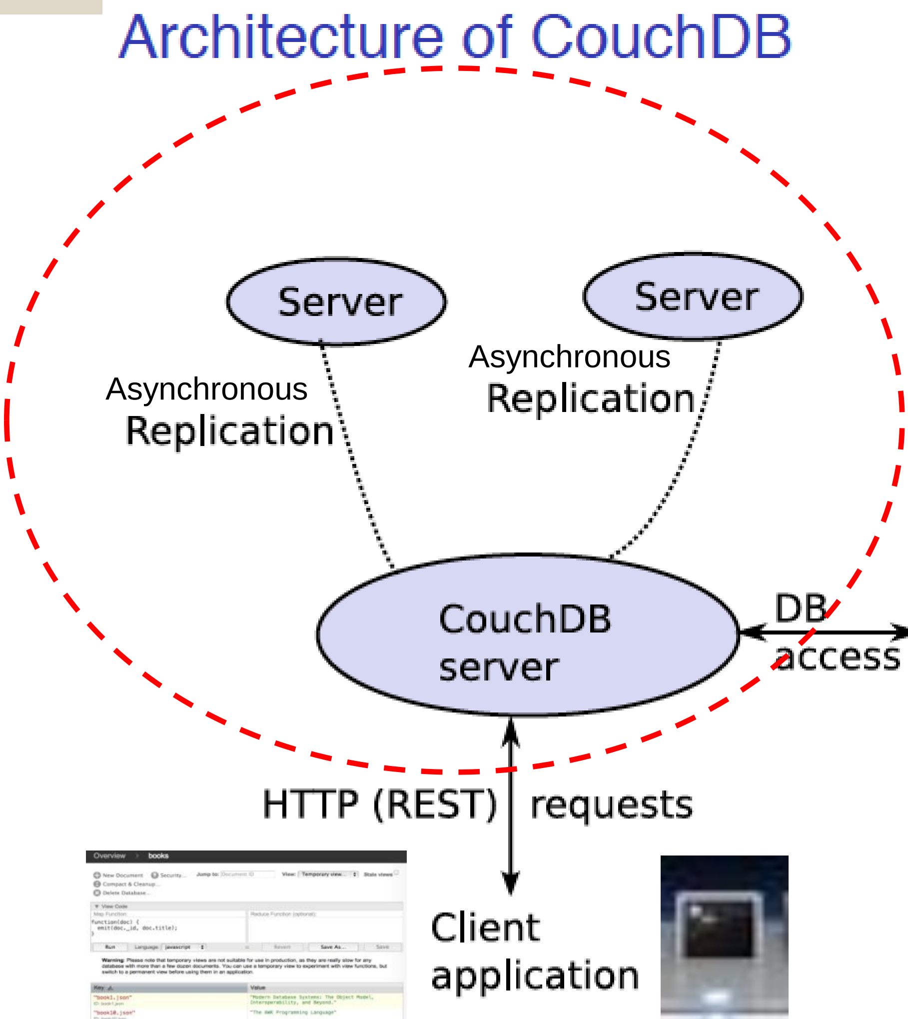
Variety #1: NoSQL / Document Stores

For a real app we could envision much more complex queries.

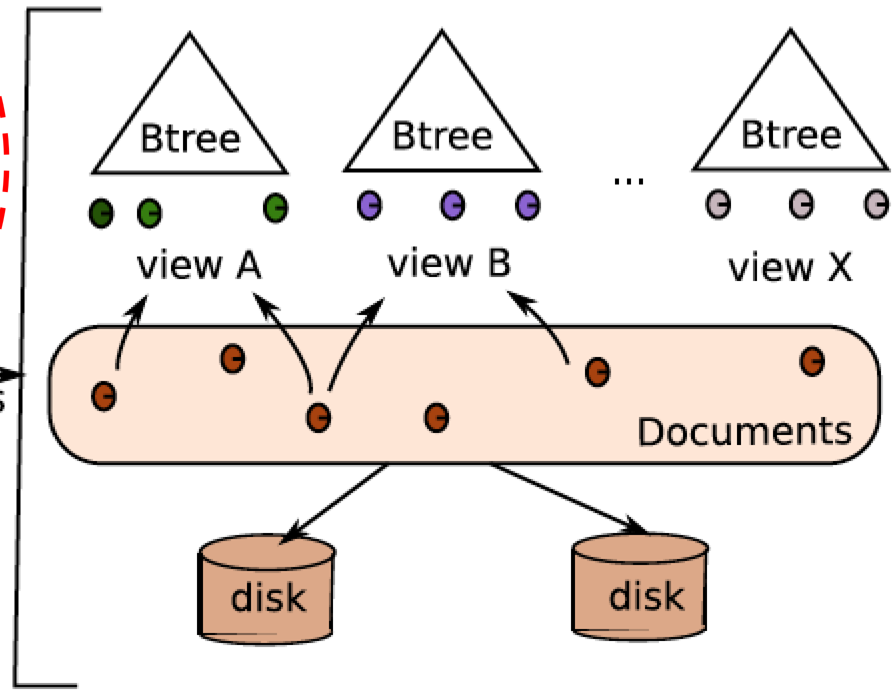


http://rickosborne.org/download/SQL-to-MongoDB.pdf

Variety #1: NoSQL / Replication



Asynchronous Replication means Eventually Consistent



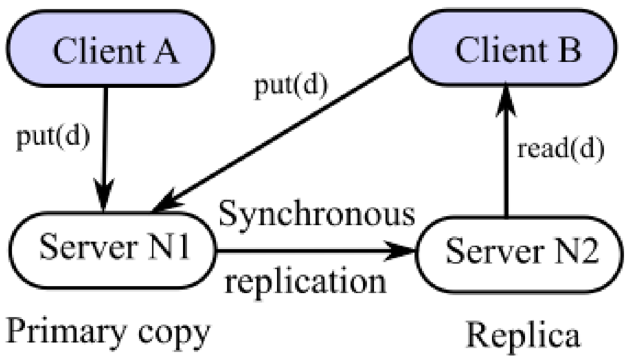
Variety #1: NoSQL / Consistency ☹

SQL RDBMSs

(Most) NoSQL DBMSs

Some illustrative scenarios

Strongly consistent!



a) Eager replication with primary copy

ntually
tent!

Variety #2: NoSQL / Map Reduce Analytics

- ***Map-Reduce:*** *a* ***programming model for processing*** *large data sets (Not online like Warehouses ☹).*
  - *Invented by Google! "MapReduce: Simplified Data Processing on Large Clusters, Jeffrey Dean and Sanjay Ghemawat, OSDI'04: Sixth Symposium on Operating System Design and Implementation,San Francisco, CA, December, 2004."*
  - *Can be implemented in any language (Java, example nex)*
- ***Hadoop:*** *Apache's open-source* ***software framework*** *that supports* ***data-intensive distributed applications***
  - *Derived from Google's MapReduce + Google File System (GFS) papers.*
  - *Enables applications to work with thousands of computation-independent computers and petabytes of data.*
  - Download: http://hadoop.apache.org/

Variety #2: NoSQL / Map Reduce Analytics

Example: term count in MapReduce (input)

Count the distinct words in all documents

cat *.txt | sort | uniq -c

term	count
jaguar	5
mammal	1
family	3
available	1
...	

URL	Document
u_1	the jaguar is a new world mammal of the felidae family.
u_2	for jaguar, atari was keen to use a 68k family device.
u_3	mac os x jaguar is available at a price of us $199 for apple's new "family pack".
u_4	one such ruling family to incorporate the jaguar into their name is jaguar paw.
u_5	it is a big cat.

1 TB on 1 PC = 2 hours!!!

1TB on 100 PCs = 1min!!!

Variety #2: NoSQL / Map Reduce Analytics

Example: term count in MapReduce

Example uses 1 mapper / 1 reduce only!

Map

list(K', V')

term	count
jaguar	1
mammal	1
family	1
jaguar	1
available	1
jaguar	1
family	1
family	1
jaguar	2
...	

Shuffle

$(K', \text{list}(V'))$

term	count
jaguar	1,1,1,2
mammal	1
family	1,1,1
available	1
...	

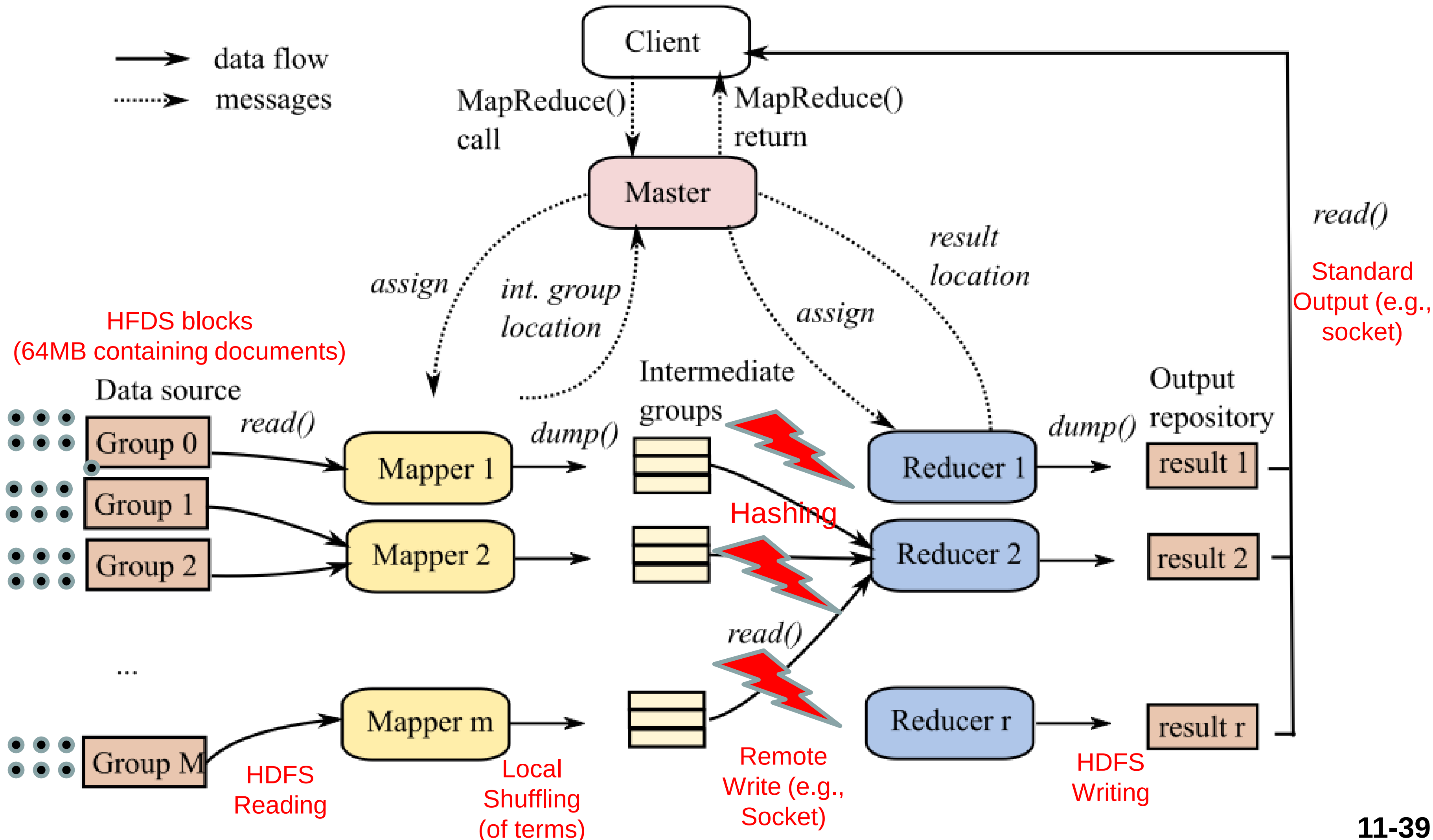
Reduce

list(K'', V'')

term	count
jaguar	5
mammal	1
family	3
available	1
...	

final

Variety #2: NoSQL / Map Reduce Analytics



Variety #3: NoSQL / Column Stores

- A **column-oriented DBMS** is a database management system (DBMS) that **stores data tables** as sections of **columns rather than as rows**, like most relational DBMSs

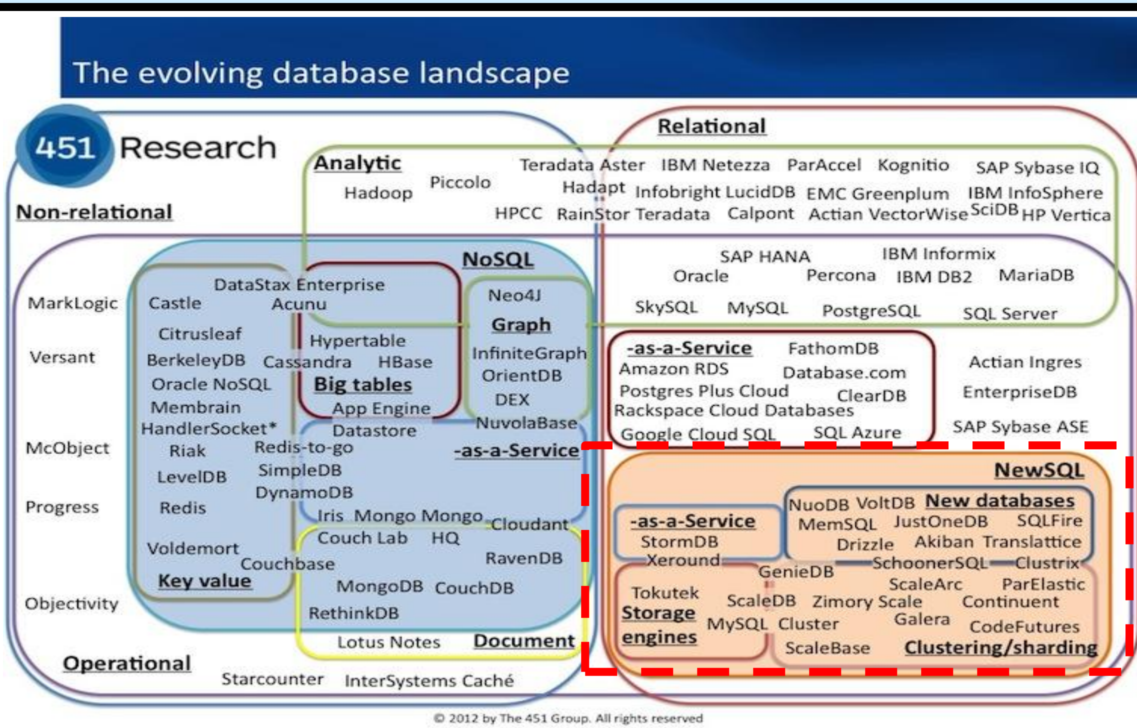
Row-Store OLTP-workloads!
1,Smith,Joe,40000;
2,Jones,Mary,50000;
3,Johnson,Cathy,44000;

Column-Store OLAP-workloads!
1,2,3;
Smith,Jones,Johnson;
Joe,Mary,Cathy;
40000,50000,44000;

- Suggested for data warehouses, customer relationship management (CRM) systems and other ad-hoc inquiry systems **where aggregates or scans are carried out over large numbers** of **similar data items**

Variety #4: NewSQL

- ***"NewSQL"** is a class of modern relational database management systems that seek to provide the **same scalable performance** of **NoSQL** systems for **OLTP** workloads while still maintaining the **ACID guarantees (i.e., offering transactions)** of a traditional DBMS.*



NewSQL= NoSQL+Transactions

Variety #4: NewSQL

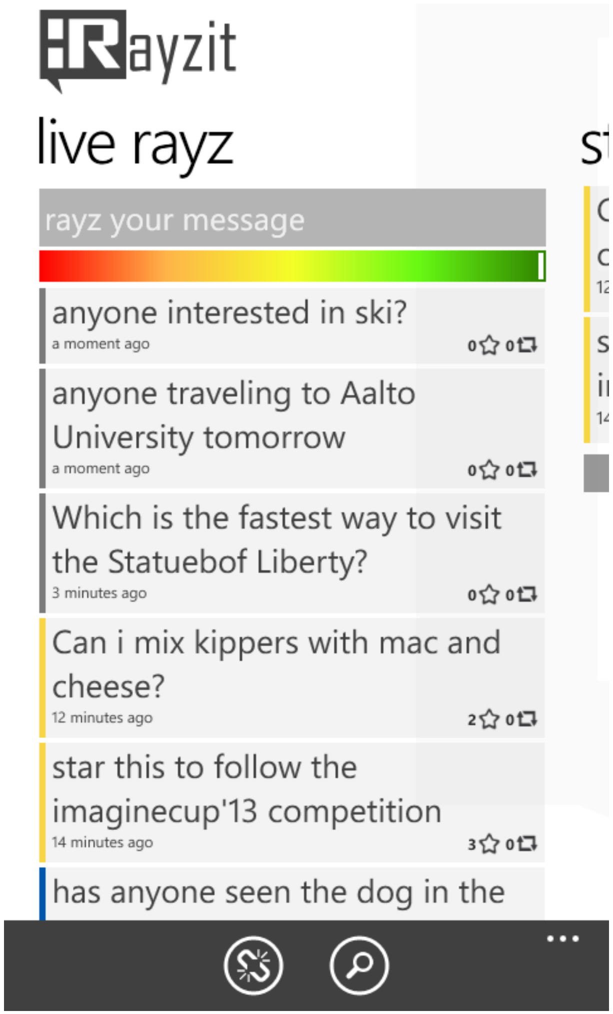
Google's Trajectory

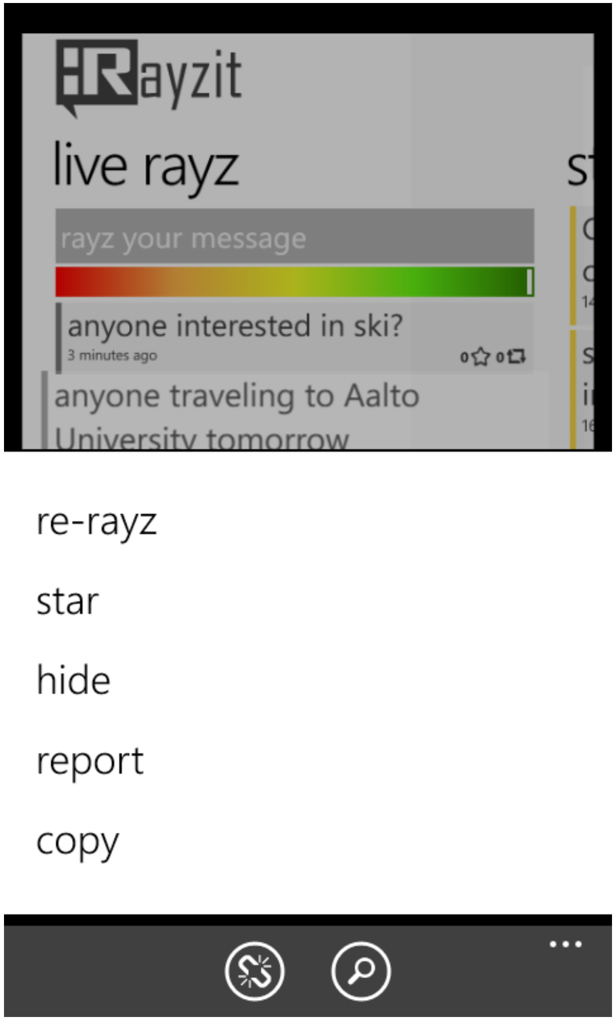
- **(2003) Google GFS Paper (SOSP'03)**
  - **Objective:** Create a Google-scale Filesystem
  - Apache HDFS is GFS open-source implementation.
- **(2004) Google's Map-Reduce Paper (OSDI'04)**
  - **Objective:** Enable big-data analytics over non-tabular data (e.g., XML or text) … with the assistance of GFS.
  - Apache's MapReduce: An open source implementation of the paper
- **(2006) Google BigTable Paper (OSDI'06)**
  - **Objective:** Enable big-data analytics over tabular data (i.e., tables)
  - (2008) Apache's Hbase: An open-source implementation of the paper
  - (2010): Facebook Messaging moves from Cassandra to HBase
- (2012) Google's F1 RDBMS (SIGMOD'12) & Spanner Storage Papers **(OSDI'12)**

Big Data Research @ UCY

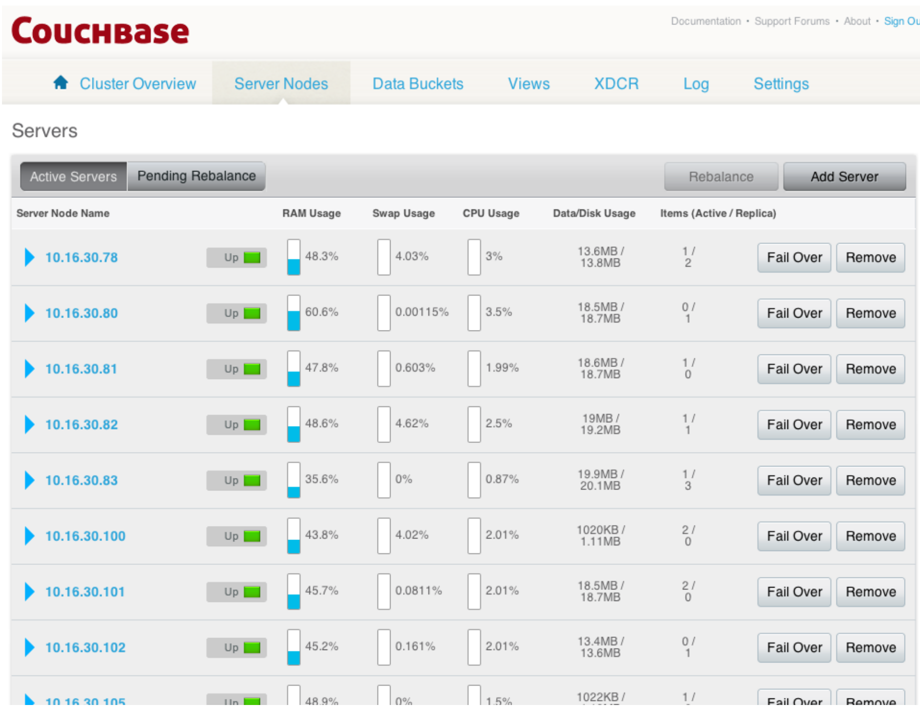
- **Rayzit.com:** Build an innovative Windows Phone messaging platform for a Finnish alliance, backed by Microsoft & Nokia.
- **Problem:** Millions of users querying their K closest smartphones continuously.
  - Query executed every few seconds.
  - Currently state-less service
- **Setup:** A 14-node Couchbase cluster (i.e., distributed - shared-nothing architecture - NoSQL document-oriented database that is optimized for interactive applications

Big Data Research @ UCY
http://rayzit.com/





Big Data Research @ UCY

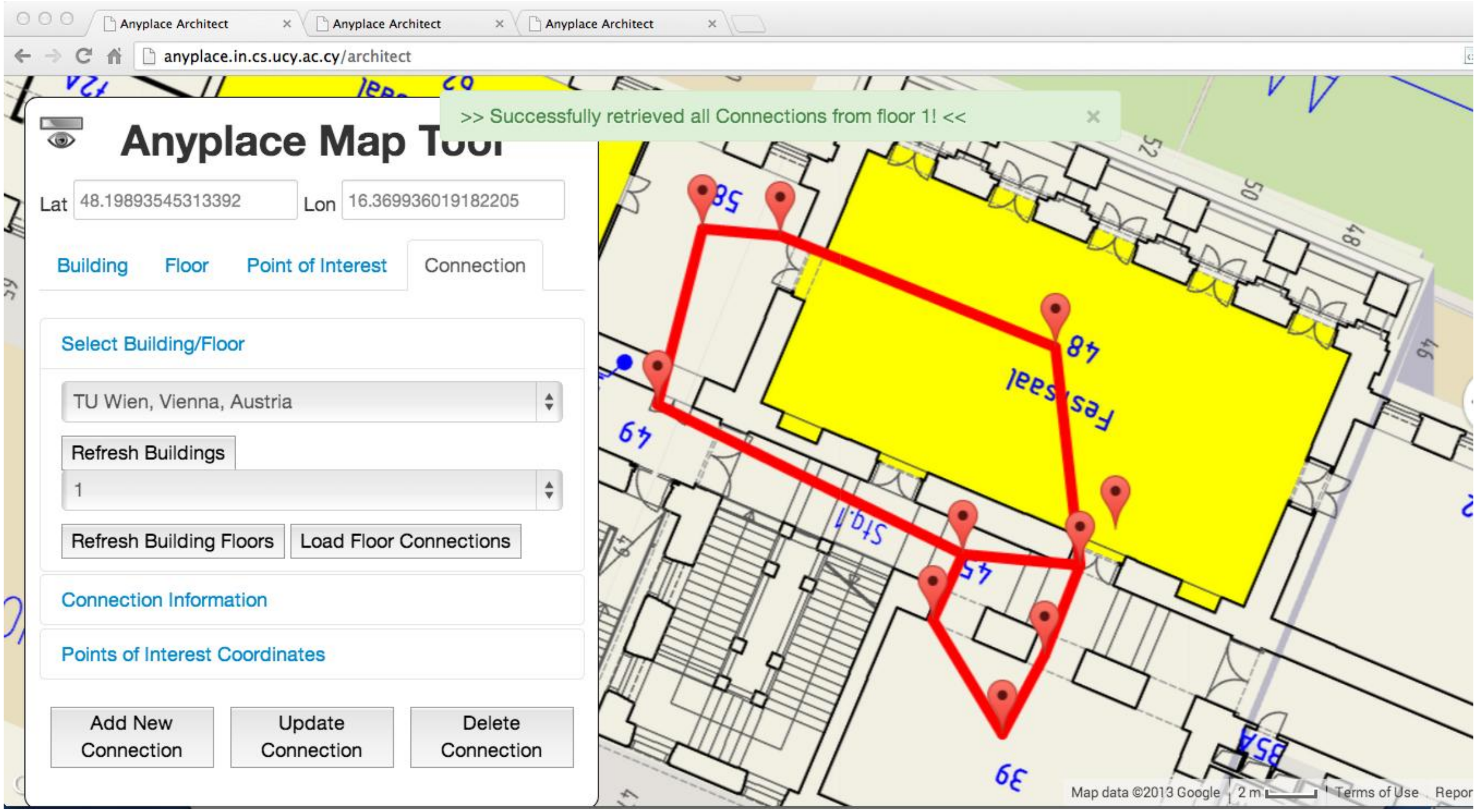


Native JSON Store + JSON RESTful API

Big Data Research @ UCY (Anyplace)

- **Anyplace:** Build an innovative indoor localization & navigation platform
- **Problem:** Radiomaps of indoor environments are fairly large structures considering that those become massively available.
- **Setup:** A 4-node Apache Hbase cluster (i.e., distributed, non-relational, shared-nothing architecture modeled after Google's BigTable and is written in Java.
- Best Demo Award at IEEE MDM'12, covered on Euronews and local media.

Big Data Research @ UCY (Anyplace)



http://anyplace.cs.ucy.ac.cy/

Big Data Research @ UCY (SmartLab)

SmartLab: Massive smartphone simulations with our first global open smartphone IaaS cloud – http://smartlab.cs.ucy.ac.cy/

Big Data Research @ UCY (SmartLab)

http://smartlab.cs.ucy.ac.cy/