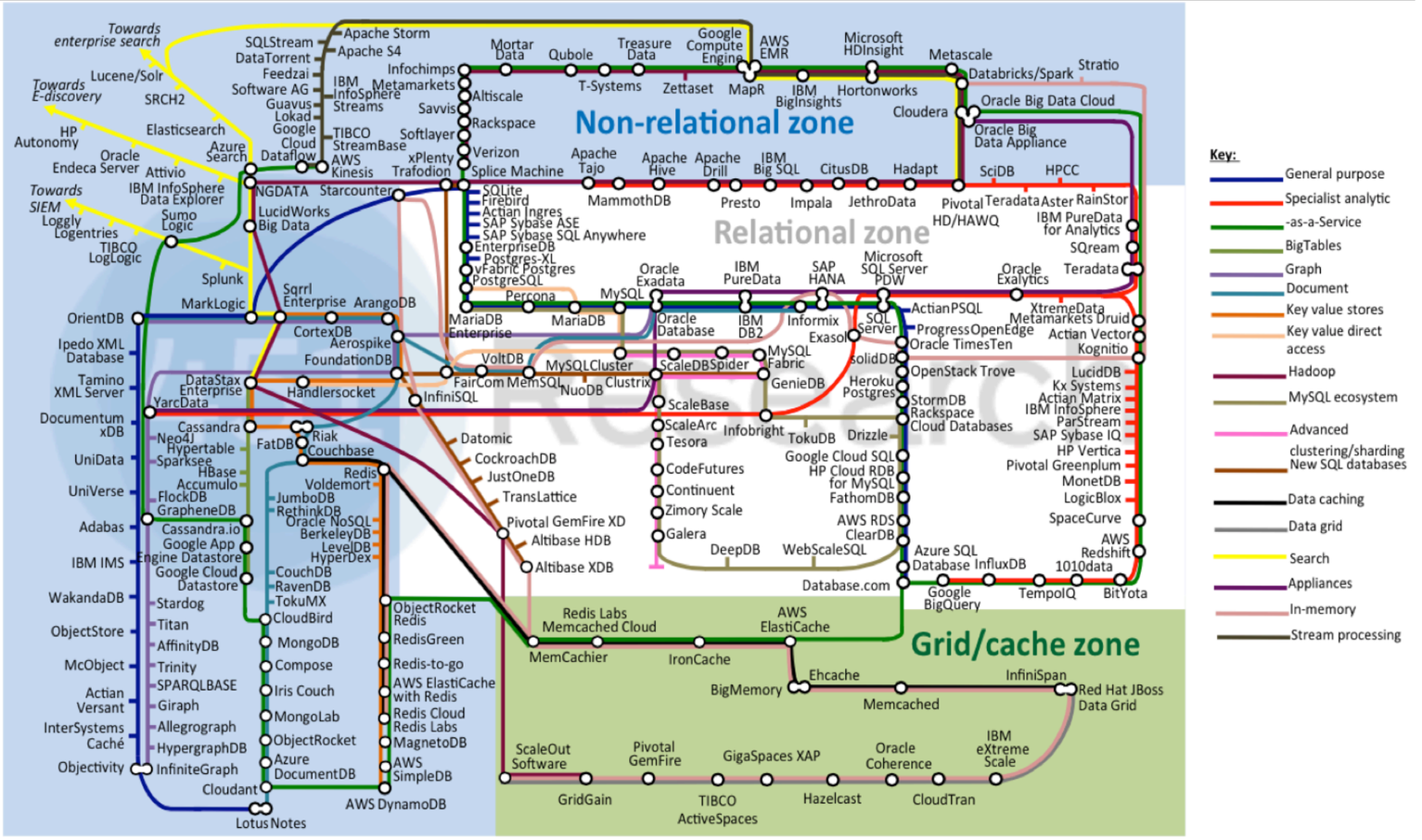
Databases - Part II

SNE Master: Essential Skills



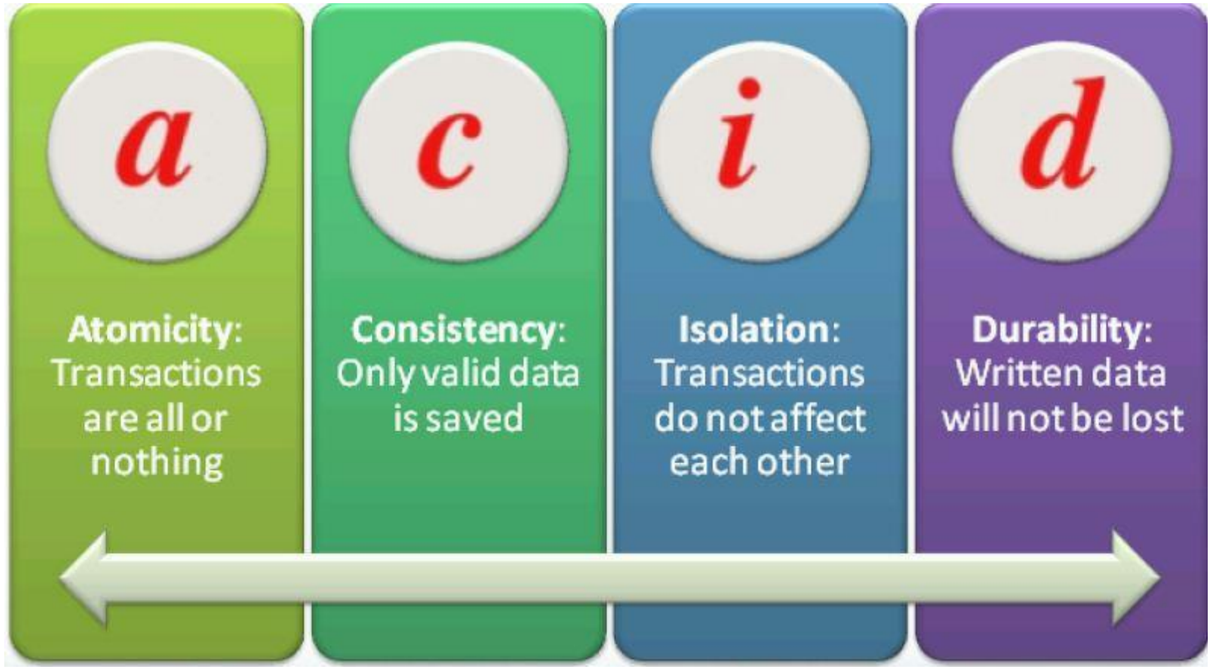
Part II: NoSQL and NewSQL

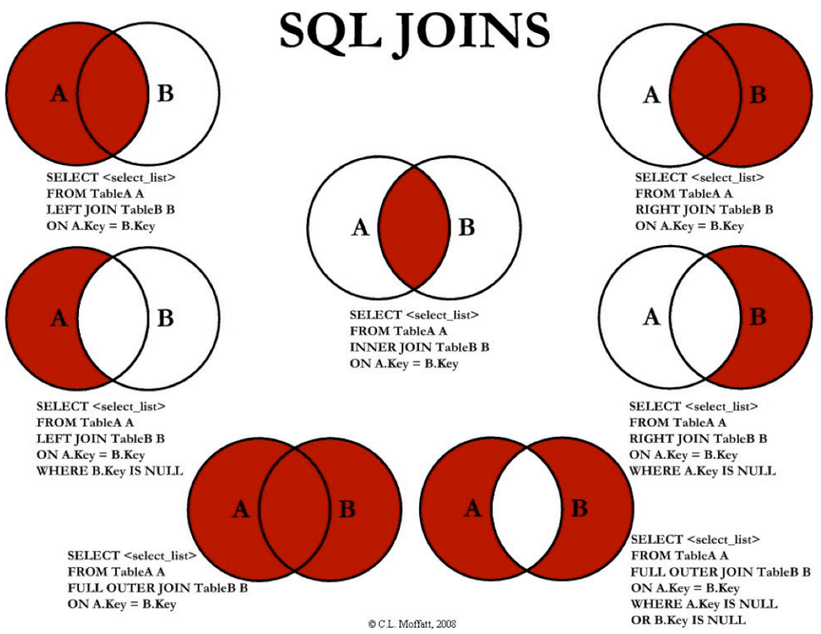
- NoSQL driving forces
- CAP Theorem
- Data models
- Hype?
- NewSQL

NoSQL - driving forces

1. *Scaling out*: partitioning
2. *Performance*: remove complexity, simplify assumptions
3. *Impedance mismatch*: from relational data structures to programming language constructs

Partitioning for RDBMS?





Performance: The cost of not being on time

"... because even the slightest outage has significant financial consequences and impacts customer trust."

"... an extra tenth of a second response time will cost 1% in sales."

amazon

"... half a second time increase in latency caused traffic to drop by a fifth."

Google

Performance: The cost of not being on time

Storing ING

Internetbankieren ING onbereikbaar door onderhoud

AMSTERDAM - Vanwege uitgelopen onderhoud aan internetbankieren bij ING was de dienst dinsdagochtend niet bereikbaar. Inmiddels kunnen klanten...

ING wil alternatief voor bankieren bij storing

AMSTERDAM – ING gaat vanwege de vele storingen in januari de technische splitsing van Mijn ING, mobiel bankieren en iDEAL versneld doorvoeren.

Problemen bij internetbankieren ING

AMSTERDAM - Internetbankieren kampt donderdagmiddag weer met een kleine storing. Een fractie van de klanten heeft problemen met inloggen of het...

ING-storingen gevolg van menselijke fouten

AMSTERDAM - Twee van de vele storingen die het internetbankieren van ING de afgelopen weken hebben getroffen zijn veroorzaakt door menselijke...

Geen standaard compensatie voor ING-storing

AMSTERDAM - ING is niet van plan klanten of webwinkels te compenseren voor diverse storingen in de afgelopen vier dagen.

Internetbankieren ING beperkt toegankelijk

AMSTERDAM - Klanten van ING kunnen woensdag soms geen gebruik maken van internetbankieren omdat de bank een maximum aan verkeer heeft ingesteld.

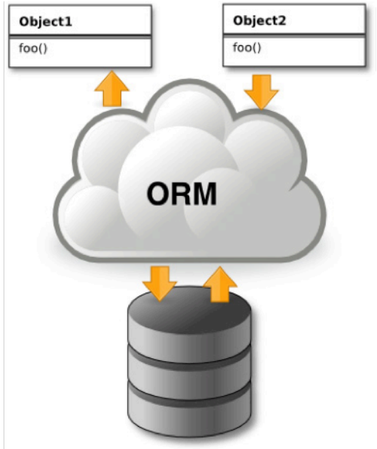
Internetbankieren ING voor derde dag offline

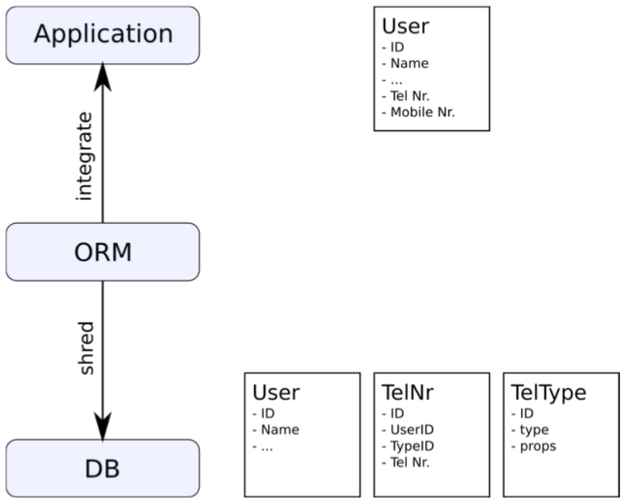
AMSTERDAM - Internetbankieren van ING kampt dinsdagmiddag weer met een storing. Zondag en maandag was de dienst ook al offline.

Impedance mismatch

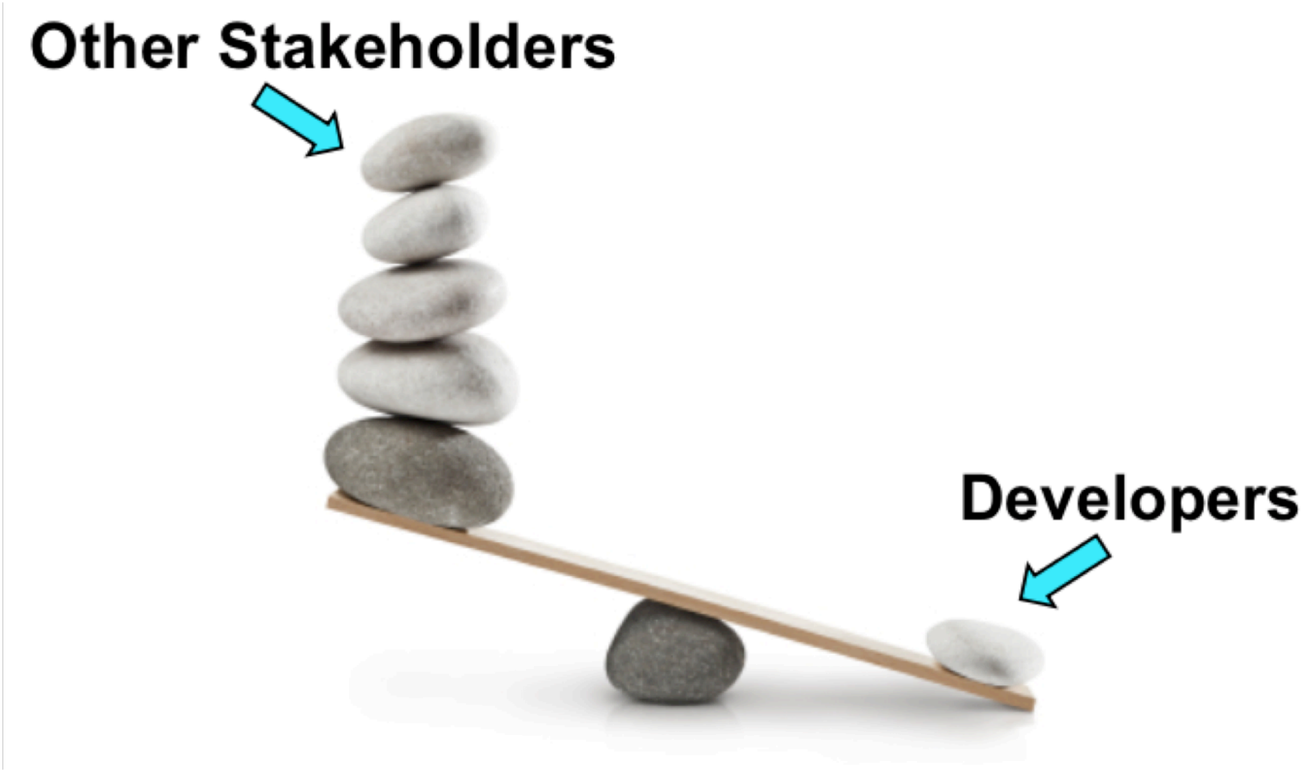
Object oriented

vs. **Relational** data structures

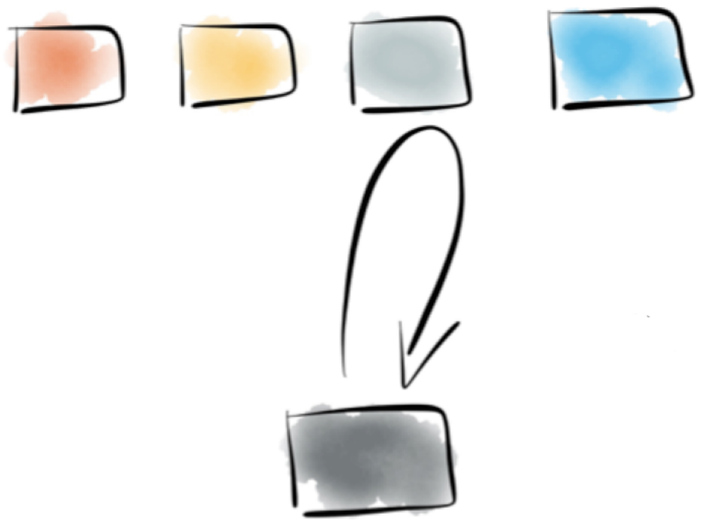




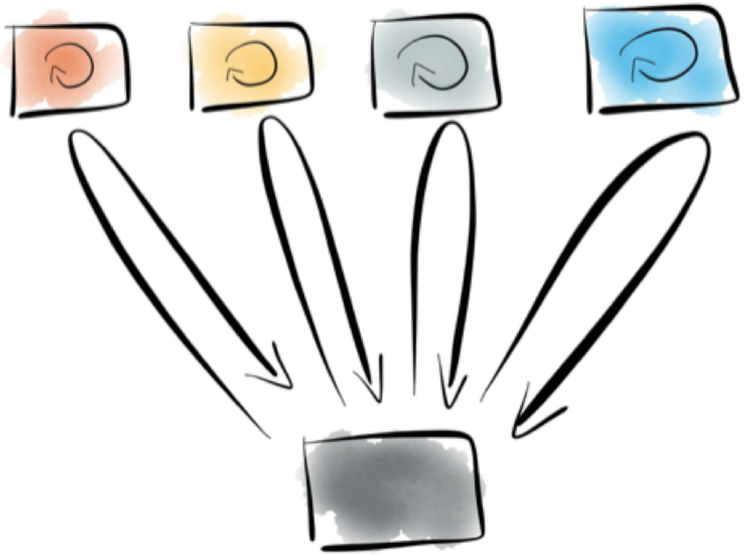
Impedance mismatch



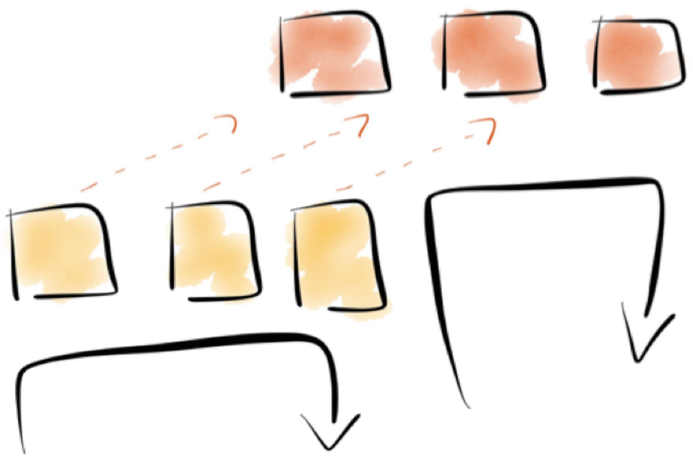
Partitioning



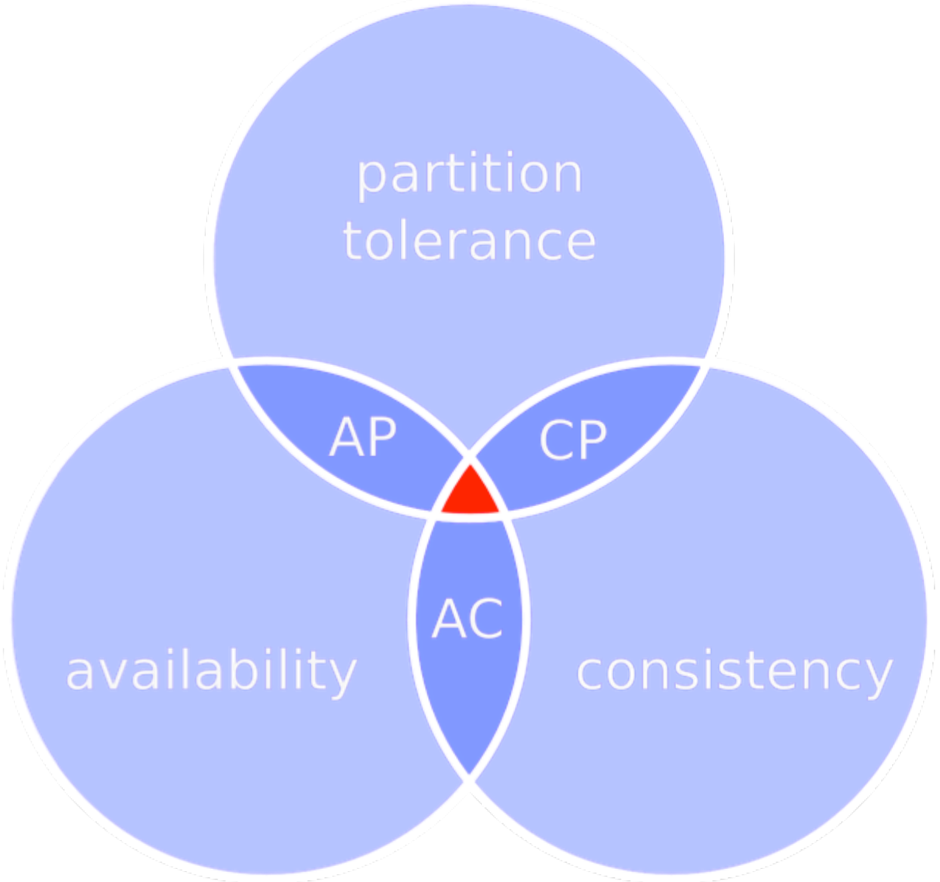
Partitioning



Partitioning



Partitioning and CAP



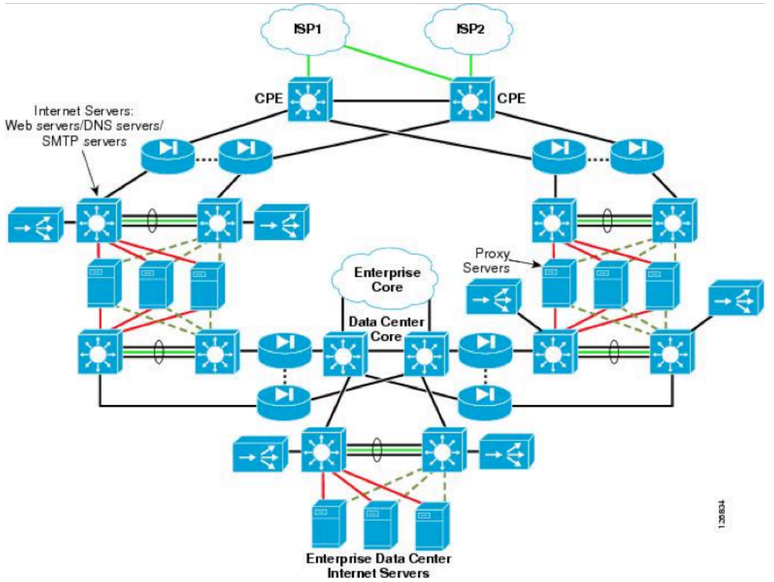
Consistency: All nodes see the same data at the same time

Availability: A guarantee that every request receives a response about whether it succeeded or failed

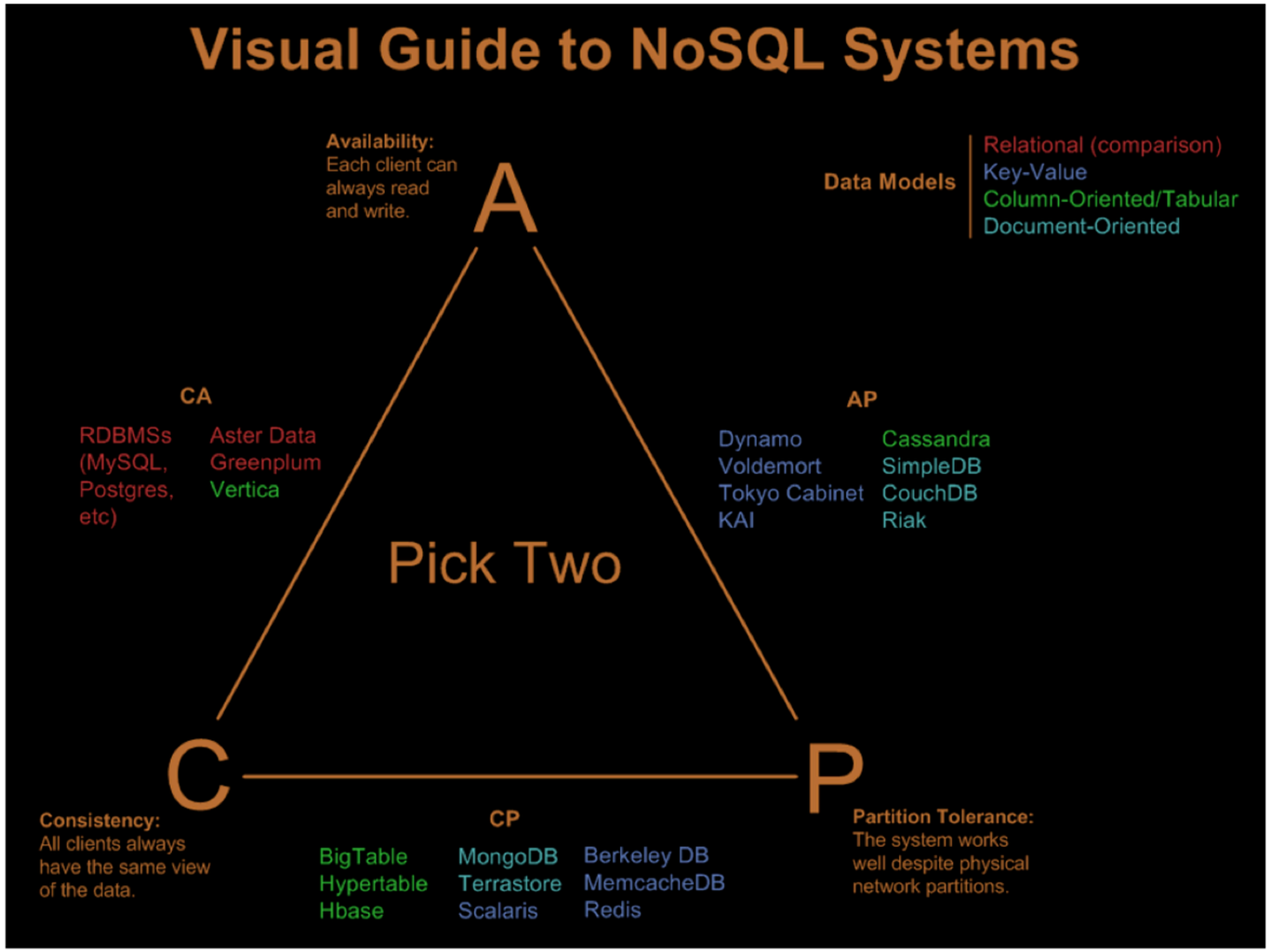
Partition tolerance: Ability to cope with a partitioned network of system nodes

Partition tolerance

1. The network is reliable
2. Latency is zero
3. Bandwidth is infinite
4. The network is secure
5. Topology doesn't change
6. There is one administrator
7. Transport cost is zero
8. The network is homogeneous



Visual Guide to NoSQL systems



CAP Misconceptions

1. CA?
2. No CA when P?
3. C is all or nothing
4. CAP is all about eventual consistency

PACELC

If Partioned:

Tradeoff Availability and Consistency

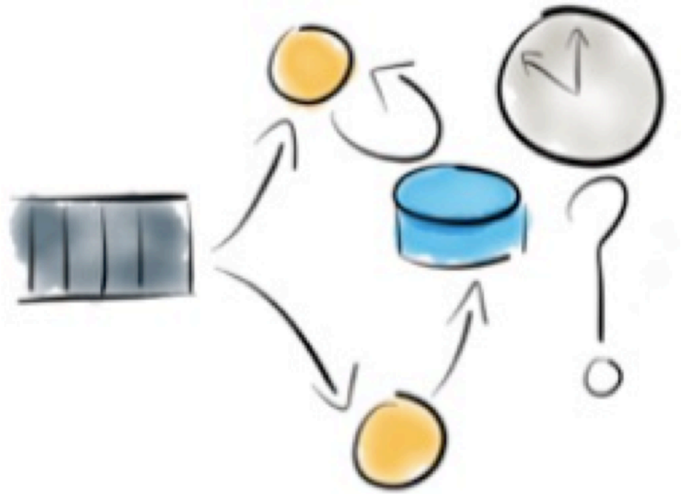
Else:

Tradeoff Latency and Consistency

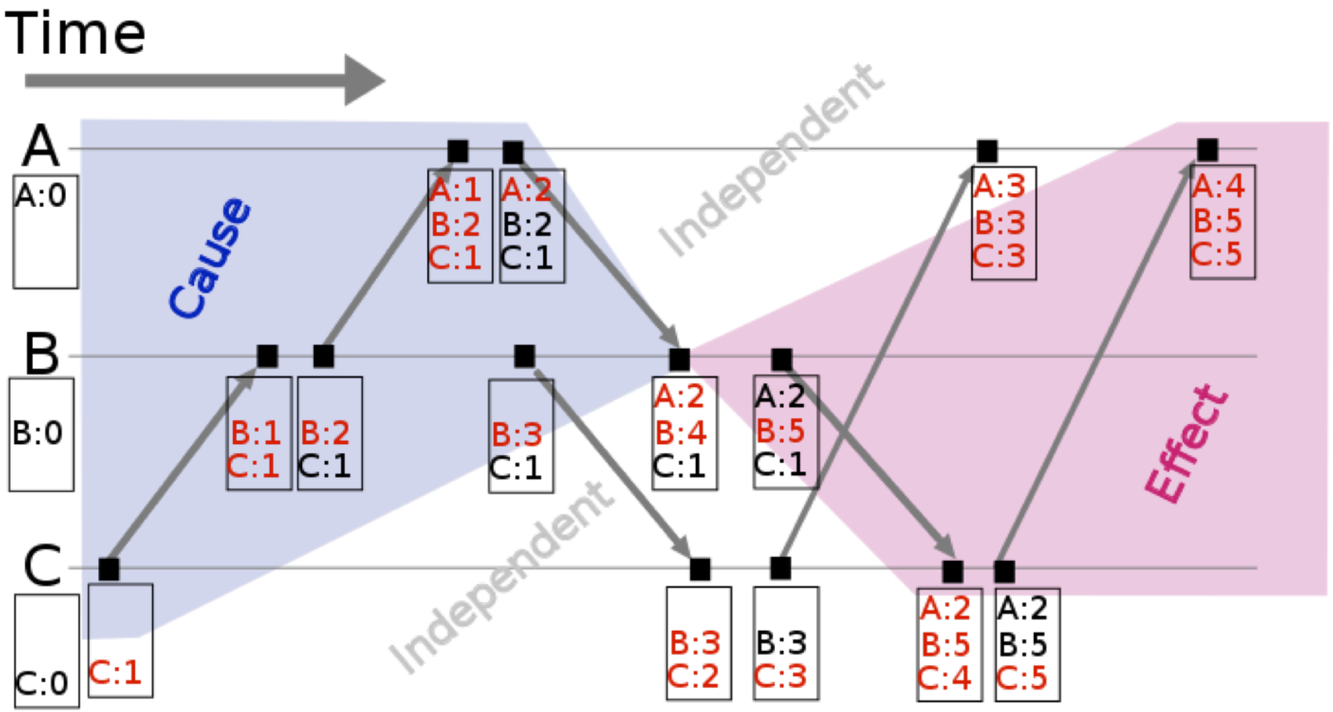
Types of systems:

- PC/EC -> Most consistent
- PA/EL -> No consistency but low latency
- PA/EC -> Give up consistency when partitioned
- PC/EL -> Madness? See PNUTS

Consistency: atomicity and ordering



Consistency: vector clocks



Consistency: tunable CAP

Riak:

N: number of total copies

R: minimal number of responding clients when reading

W: minimal number of responding clients when writing

Cassandra:

Tunable write consistency

Tunable read consistency

MongoDB:

Write result setting

Read result setting

Tunable CAP: Amazon shopping Basket

Amazon uses DynamoDB for shopping baskets. DynamoDB, like Riak is a distributed key-value store where N-R-W can be set for operation.

If you were Amazon, how would you choose N-R-W for a shopping basket?

Eventual consistency

Key property of non-ACID

- If no further changes
- All nodes will end up consistent

Weak guarantee

- When is eventually? It doesn’t say..
- In practice: expect inconsistency; always!

In practice:

- Stronger guarantees: predicting/measuring behavior
- Systems often appear strongly consistent

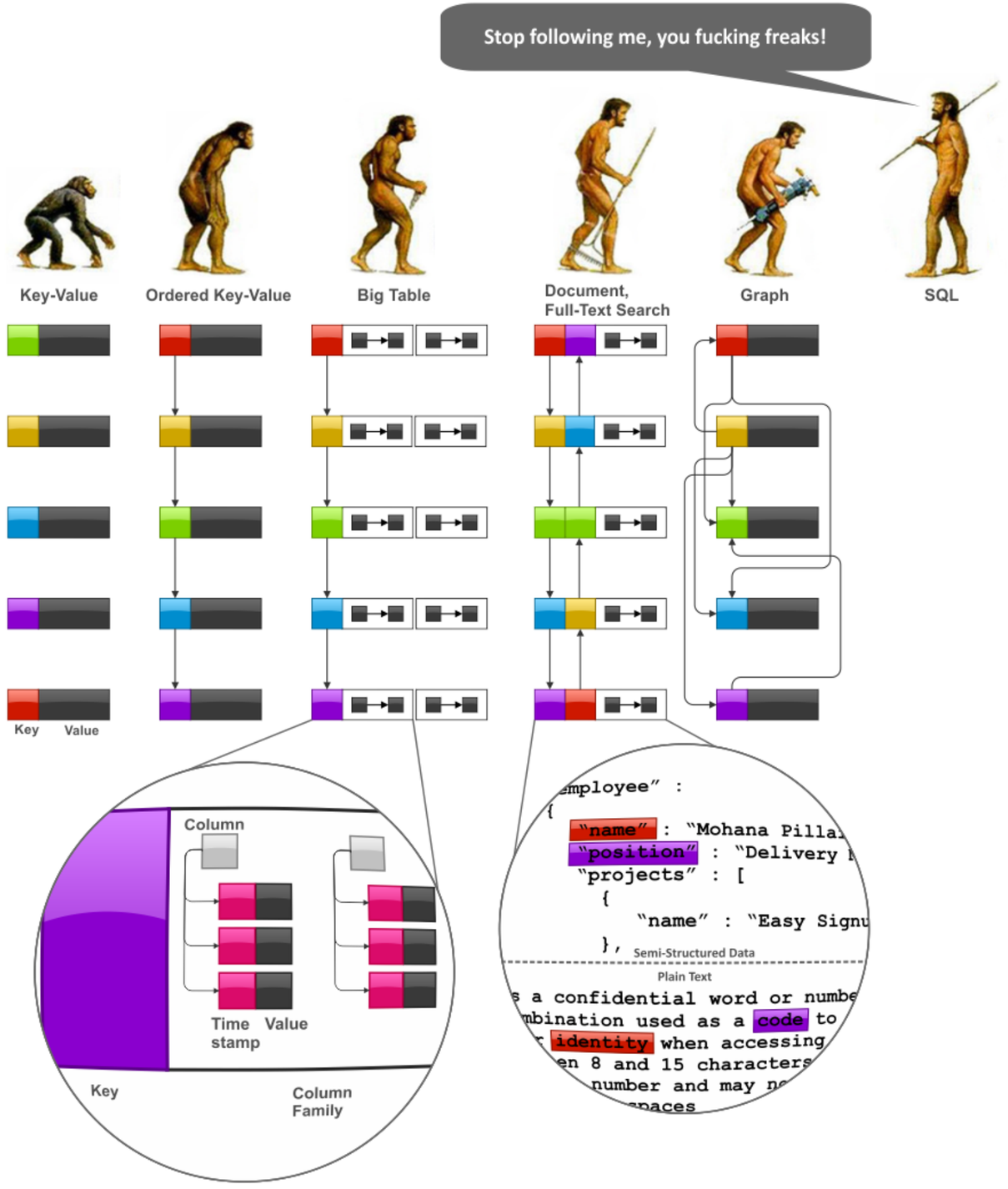
Nodes must exchange information on writes:

- Inform replicas
- Inform client

Deal with conflicts:

- Last write win?
- Vector clocks
- Multiversion storage
- Hardware clocks

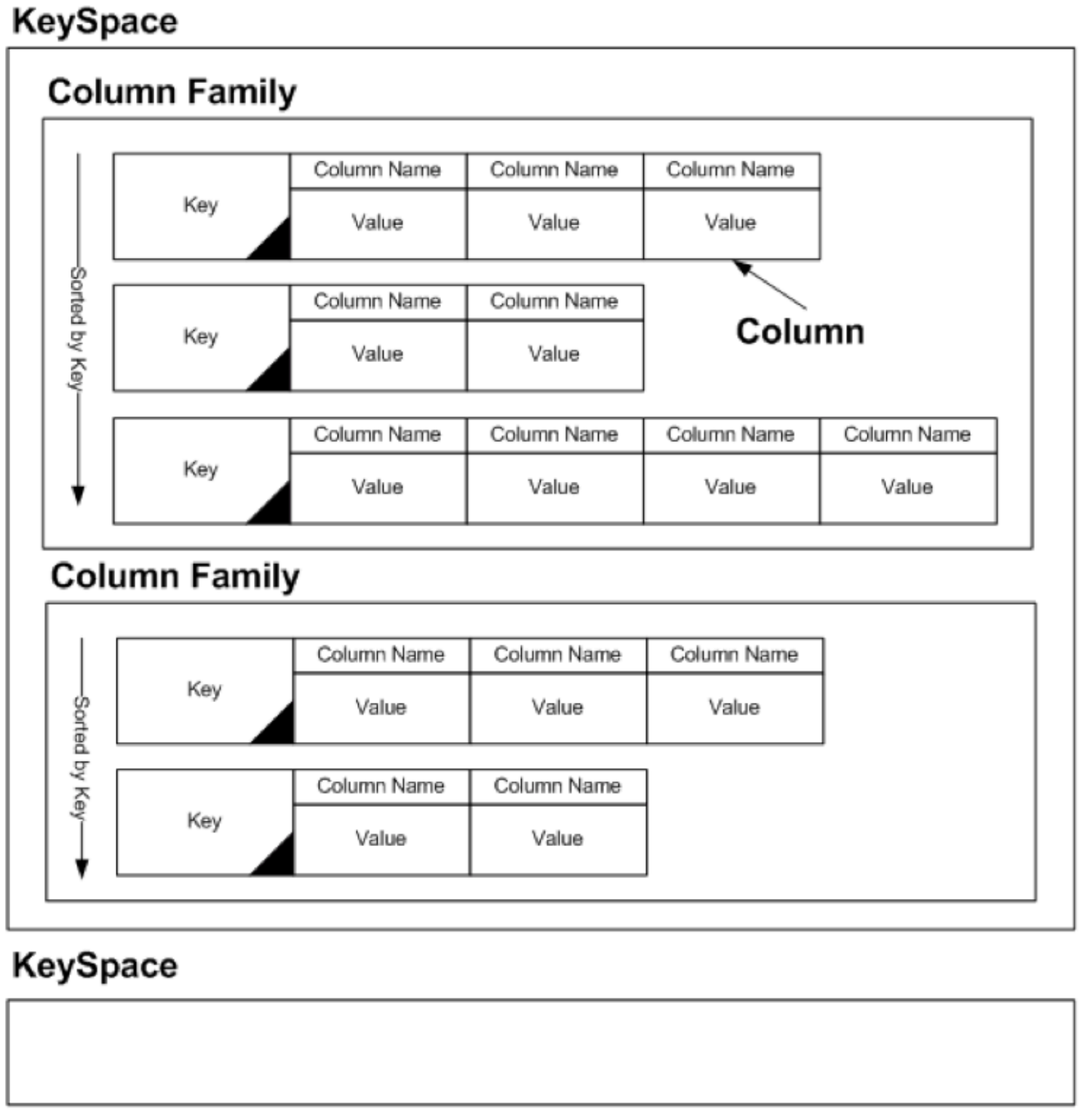
Data models


Stop following me, you fucking freaks!
Key-Value
Ordered Key-Value
Big Table
Document,
Full-Text Search
Graph
SQL
Key
Value
Column
Time stamp
Value
Key
Column Family
"name" : "Mohana Pilla
"position" : "Delivery
"projects" : [
"name" : "Easy Signu
Semi-Structured Data
Plain Text
a confidential word or numbe
bination used as a code to
identity when accessing
en 8 and 15 characters
number and may n
paces

Data model: Column based

Column stores:

- HBase
- Cassandra
- MonetDB
- Google: BigTable

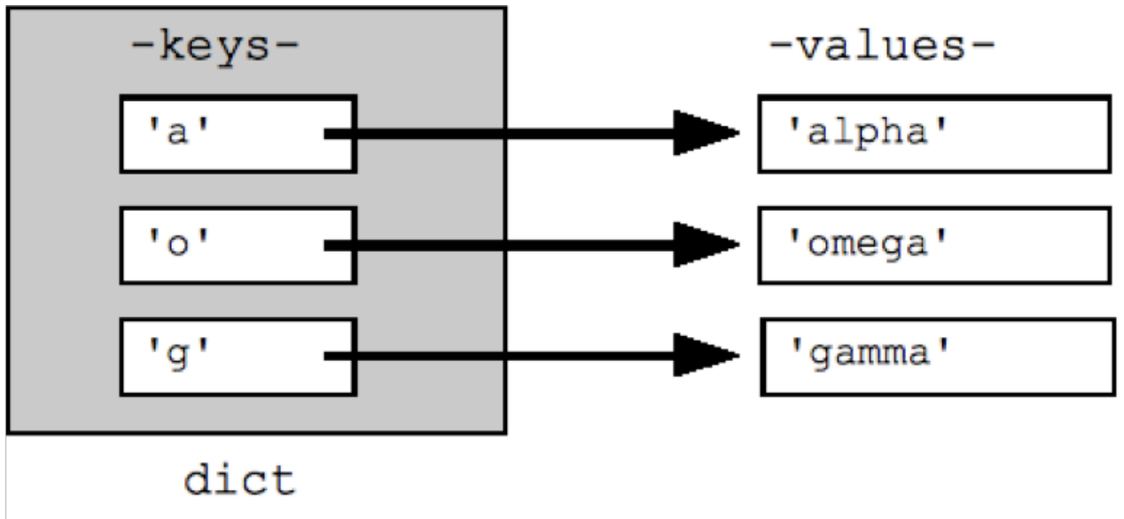


Data model: Key-value

Similar to dictionaries and (hash)maps

Key-value stores:

- Riak
- DynamoDB
- Redis
- memcached



Data model: Key-value

JSON Documents as data.

Document stores:

- MongoDB
- CouchDB
- CouchBase

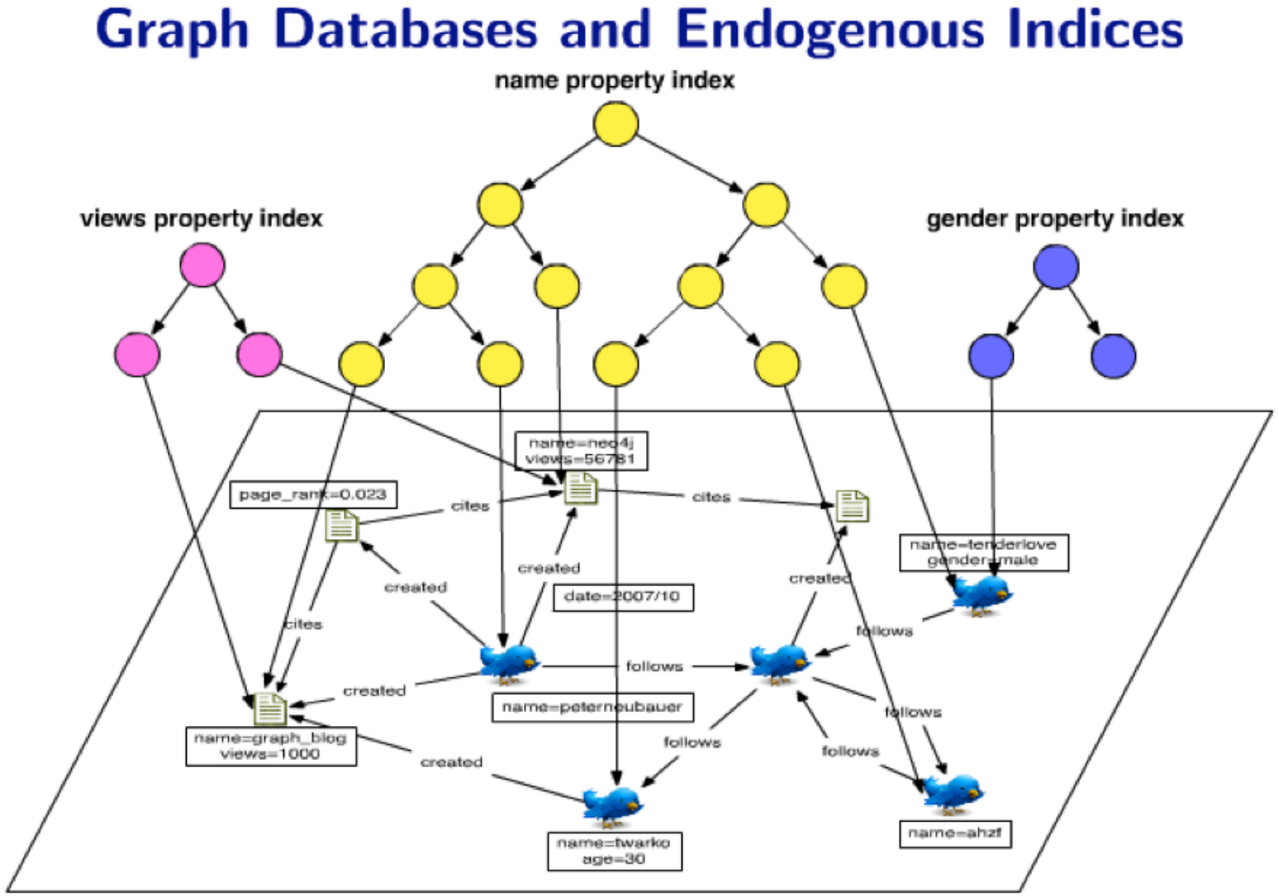
```
{
  "arguments" : { "number" : 10 },
  "url" : "http://localhost:8080/restty-tester/collection",
  "method" : "POST",
  "header" : {
    "Content-Type" : "application/json"
  },
  "body" : [
    {
      "id" : 0,
      "name" : "name 0",
      "description" : "description 0"
    },
    {
      "id" : 1,
      "name" : "name 1",
      "description" : "description 1"
    }
  ],
  "output" : "json"
}
```

Data model: Graphs

Graphs and graph queries as first class citizens.

Graph stores:

- Neo4J
- OpenLink
- Titan



Data model: Specialized

Distributed

Lucene

Indexes

elasticsearch.

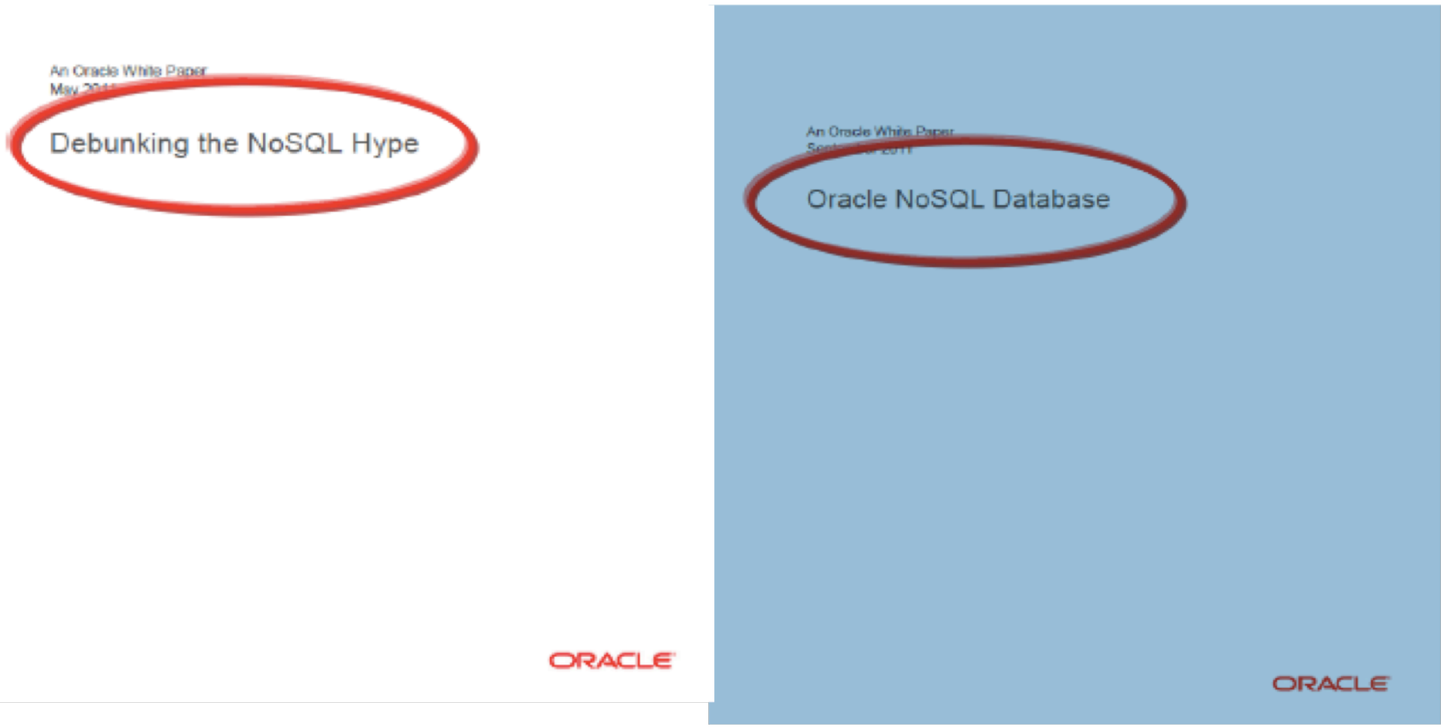
kibana

logstash

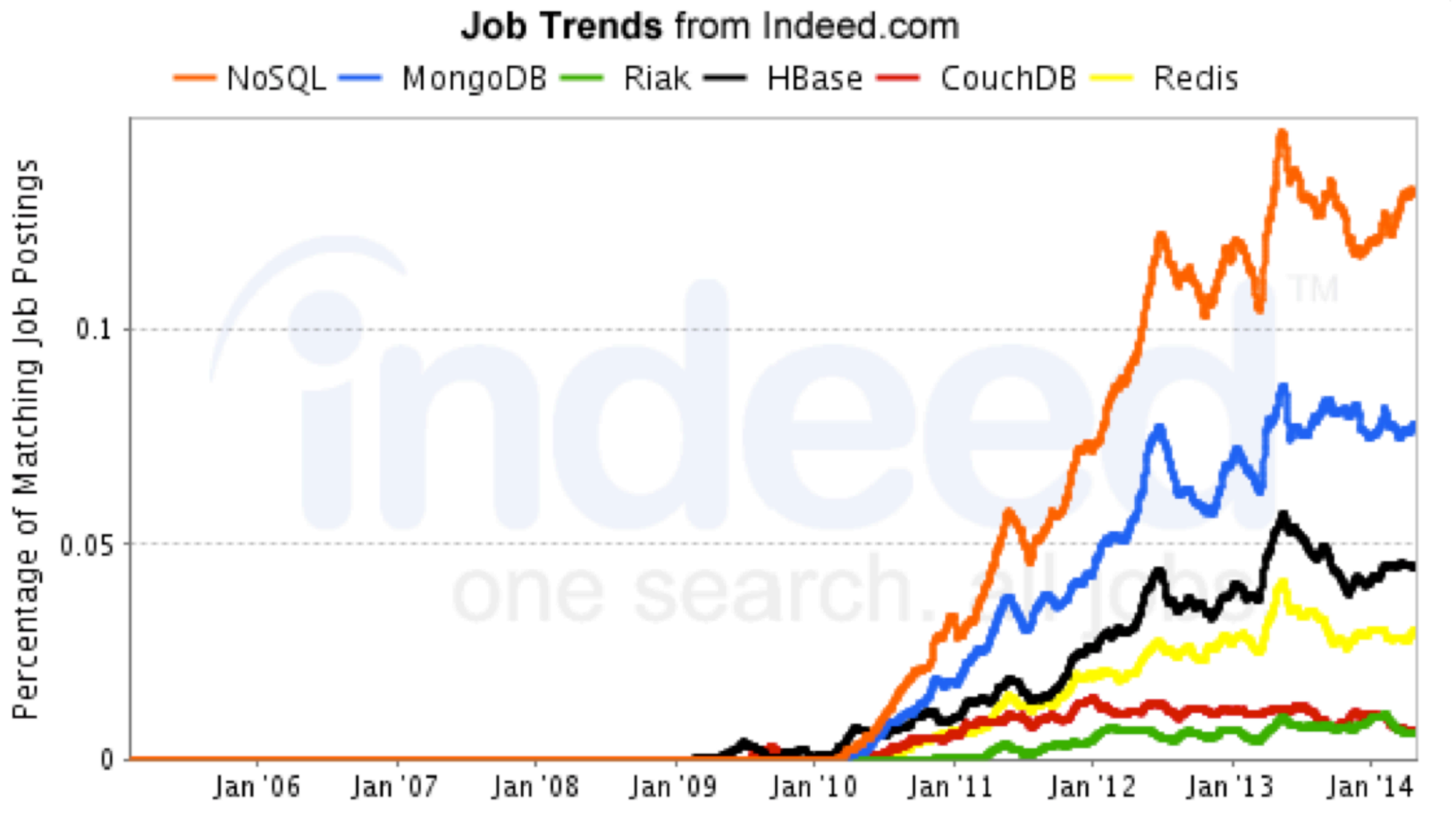
<p class="cleartext"> Products: Laser Printers The fundamental everyday requirement for mono and colour laser printing throughtout today's offices is perfectly met with the extensive Epson laser printer range. The latest AcuLaser printer range offers users exceptionally Epson AcuLaser C1900 Networked compact colour laser printer for professional enterprises Businesses have been denied simple and affordable colour laser printing for far too long. The traditionally high costs and poor speeds of colour lasers has left many offices looking a bit, well, grey. But not any more: with the Epson AcuLaser C1900, Epson brings both colour and monochrome laser printing together at a black and white price. more Where to Buy Support Epson AcuLaser C3000 The fastest colour laser printer in its class The perfect printer for small businesses and work groups, the Epson AcuLaser C3000 prints high volumes in black and white and vibrant colour, at high speed and with low running costs... more Where to Buy High quality resolution: 2400dpi RIT* Large paper capacity: 600 sheets, expandable up to 1,600 sheets Compatible Windows and Mac High speed USB and EpsonNet 10/100 Base Tx Ethernet interfaces as standard** * Epson AcuLaser Resolution Improvement Technology **EpsonNet 10/100 Base Tx Ethernet standard with Epson AcuLaser C3000N model only. AcuLaser C3000: 64MB Memory, 100 sheet MP Tray, 500 sheet cassette, Duplex printing as standard AcuLaser C3000N: 64MB Memory, 100 sheet MP Tray, 500 sheet cassette, Duplex printing, 10/100BaseTX Ethernet Interface Networked compact colour laser printer for professional enterprises Businesses have been denied simple and affordable colour laser printing for far too long. The traditionally high costs and poor speeds of colour lasers has left many offices looking a bit, well, grey. But not any more: with the Epson AcuLaser C1900, Epson brings both colour and monochrome laser printing together at a black and white price. Key Features cost effective mono printing for day to day business needs and vivid versatile colour when required. search Search Epson UK Epson AcuLaser C900 Outstanding professional colour printing for business Add colour to your business with the Epson AcuLaser C900 from Epson. Its perfect for the smaller workgroup, being a compact and cost effective laser printing workhorse that offers amazing colour output as well as high performance black and white production. more Where to Buy Support As cost efficient to run as a mono-only laser printer Paper capacity of 700 sheets from two media sources Easy to operate with advanced printer driver Memory expandable from 32Mb to 1024Mb Pre-configured models available with Wireless 802.11b, Adobe ® PostScript ® Level 3™ and two-sided printing The AcuLaser C1900 is available in 5 configurations: - AcuLaser C1900S: with 32MB, 200 Sheet MP Tray, 10/100BaseTX Networking - AcuLaser C1900: with 32MB, 200 Sheet MP Tray, 500 Sheet Cassette, 10/100BaseTX Networking Support Epson AcuLaser C4100 High performance colour lasers for all your business printing needs The Epson AcuLaser C4100 provides businesses with a high performance colour and monochrome printing solution. It adds crucial colour to your business, while producing high quality monochrome output at lower costs than many monochrome-only printers, and it's just as easy to operate. So now there's no reason to buy two printers, because perfect monochrome and colour solutions are available in one. more Where to Buy Support Epson AcuLaser C8600 Professional high performance A3W colour laser printer Epson AcuLaser C8600 is the perfect professional printing solution for users who require exceptional quality colour and mono output on a range of media formats from C5 up to A3W in size. The Epson AcuLaser C8600 is able to achieve superb print quality by utilising a combination of Epson's exclusive AcuLaser Color Laser Technologies. more Where to Buy Support - AcuLaser C1900PS: with Adobe® PostScript® 3™, 96MB, 200 Sheet MP Tray, 500 Sheet Cassette, 10/100BaseTX Networking - AcuLaser C1900D: with Duplex unit (two sided printing) 96MB, 200 Sheet MP Tray, 500 Sheet Cassette, 10/100BaseTX Networking - AcuLaser C1900 WiFi: with 32MB, 200 Sheet MP Tray, 500 Sheet Cassette, Wireless Networking facility Add colour to your business with the Epson AcuLaser C900 from Epson. Its perfect for the smaller workgroup, being a compact and cost effective laser printing workhorse that offers amazing colour output as well as high Support Epson AcuLaser C4000 High performance colour laser The Epson AcuLaser C4000 provides businesses with high performance colour and monochrome printing solutions more Where to Buy Epson AcuLaser C9100 High speed A3 colour laser printer Why have separate black and white and colour printers when you can have the Epson AcuLaser C9100? Epson has taken the lead in laser technology to deliver a complete high-performance solution for all your colour and mono printing needs. Support EPL-6200L High performance A4 mono laser professional printers The Epson EPL-6200 and EPL-6200L are the ideal printing solutions for small to medium workgroups and personal users. They deliver professional performance quickly, easily, reliably and cost-effectively, and are perfect for users who need high levels of laser quality and productivity at a low investment. more Where to Buy Support EPL-6200 High performance A4 mono laser professional printers The Epson EPL-6200 and EPL-6200L are the ideal printing solutions for small to medium workgroups and personal users. They deliver professional performance quickly, easily, reliably and cost-effectively, and are perfect for users who need high levels of laser quality and productivity at a low investment. more performance black and white production. For the first time, you can now bring the power of high quality colour to your documents without suffering the high costs or low speeds traditionally associated with colour

Hype?

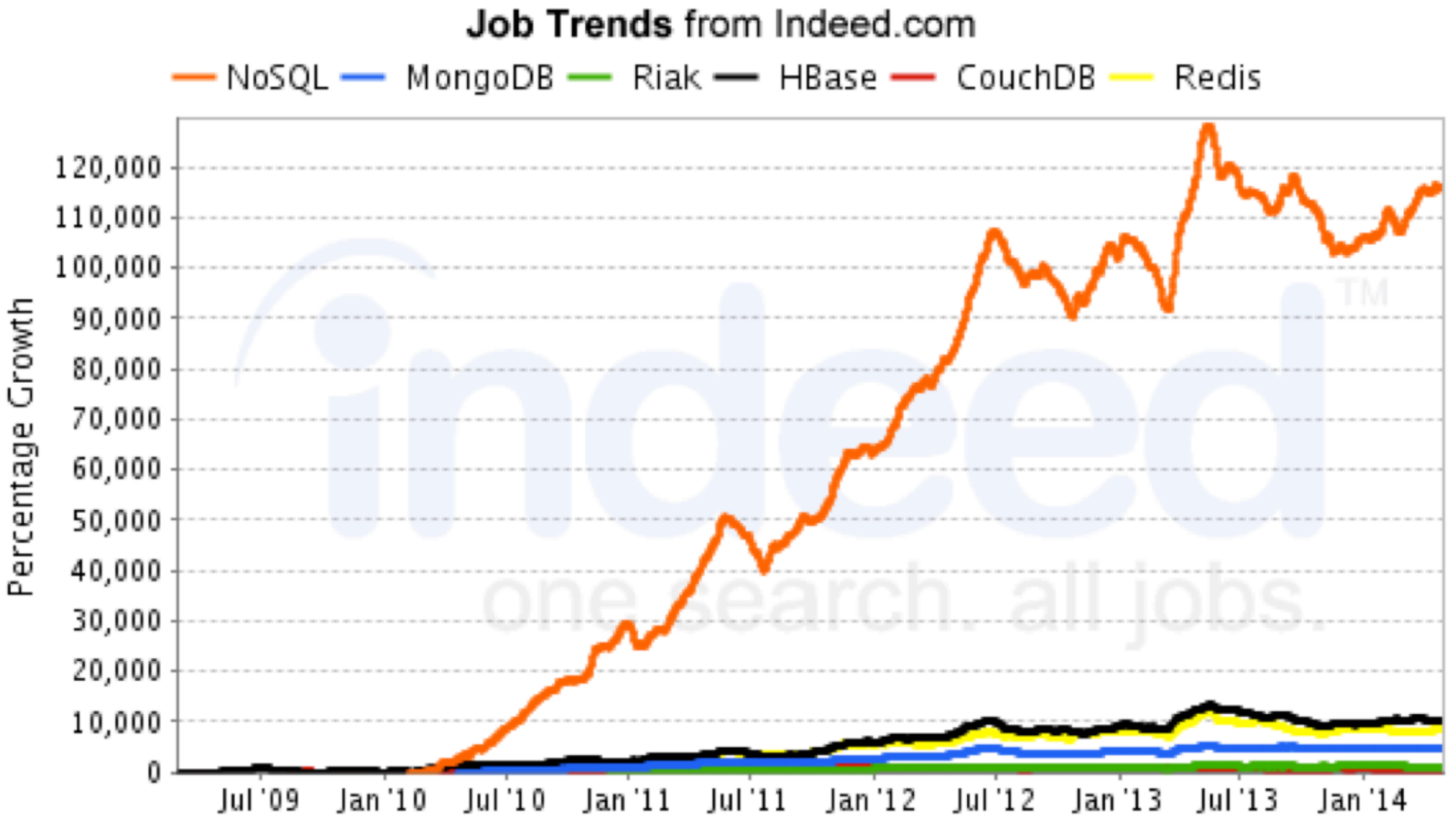
Hype?


An Oracle White Paper
Debunking the NoSQL Hype
ORACLE
An Oracle White Paper
Oracle NoSQL Database
ORACLE

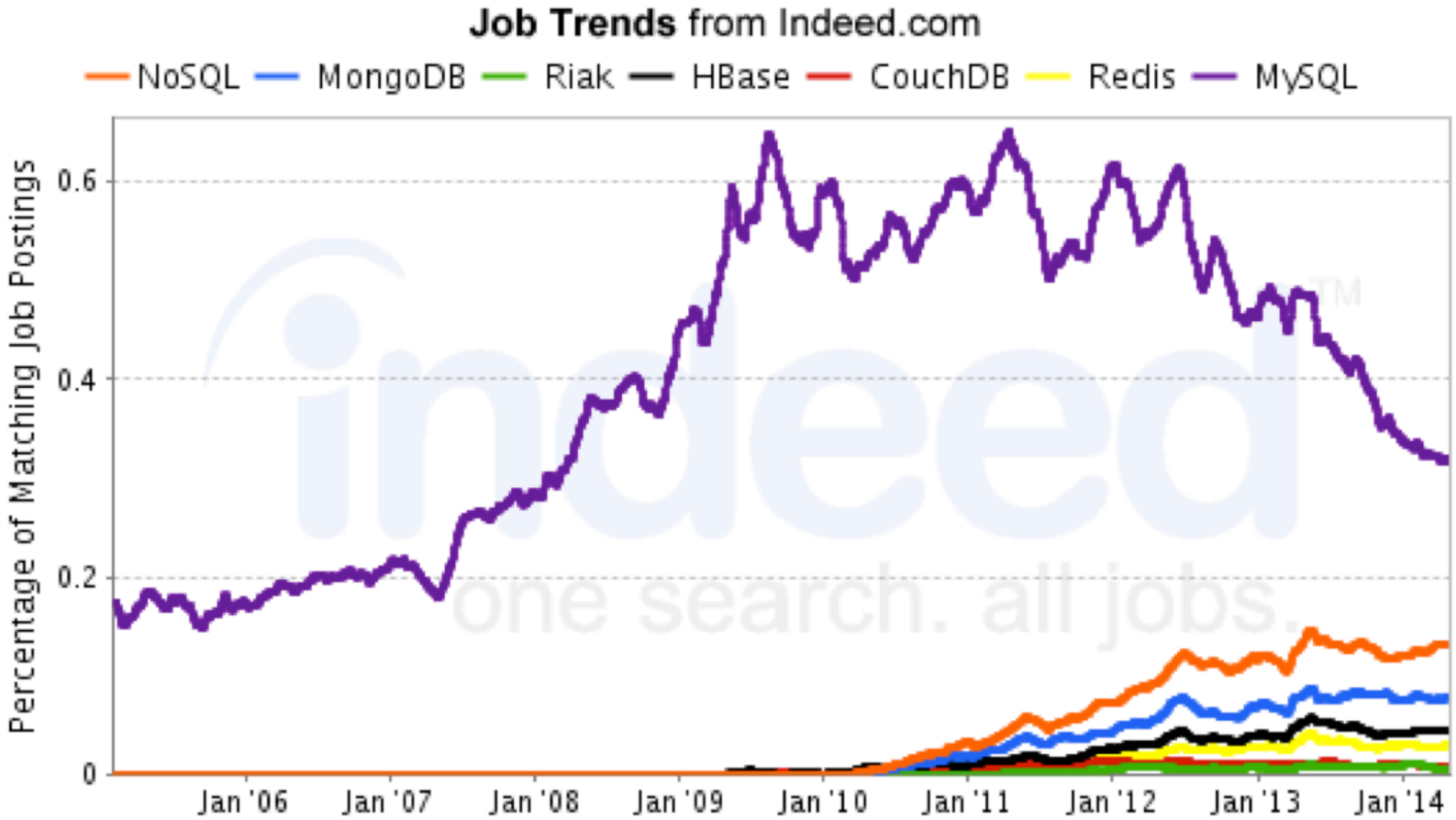
Hype?


Job Trends from Indeed.com
NoSQL
MongoDB
Riak
HBase
CouchDB
Redis
Percentage of Matching Job Postings
0.1
0.05
0
Jan '06
Jan '07
Jan '08
Jan '09
Jan '10
Jan '11
Jan '12
Jan '13
Jan '14

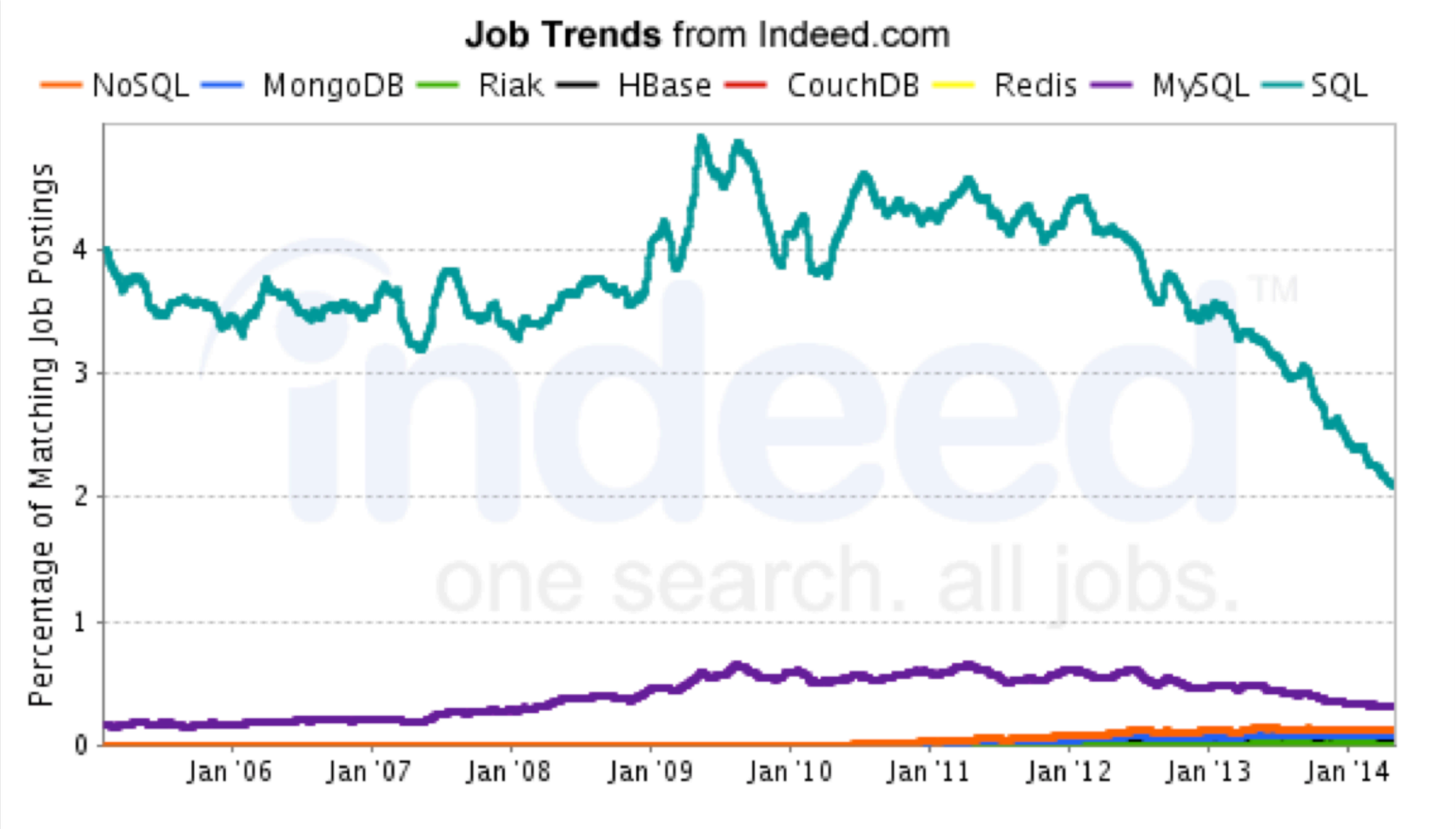
Hype?


Job Trends from Indeed.com
NoSQL
MongoDB
Riak
HBase
CouchDB
Redis
Percentage Growth
120,000
110,000
100,000
90,000
80,000
70,000
60,000
50,000
40,000
30,000
20,000
10,000
0
Jul '09
Jan '10
Jul '10
Jan '11
Jul '11
Jan '12
Jul '12
Jan '13
Jul '13
Jan '14
indeed
one search. all jobs.

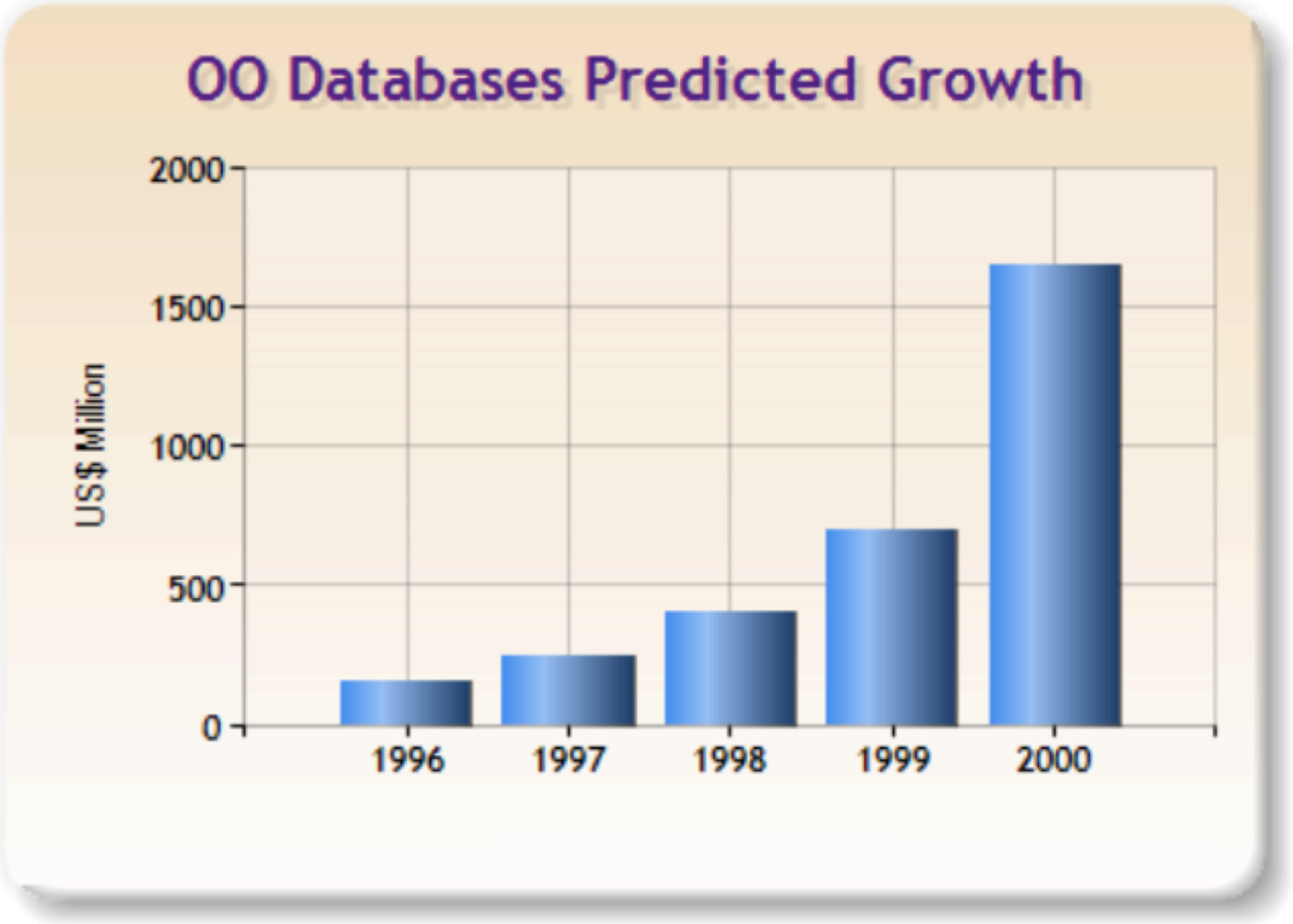
Hype?


Job Trends from Indeed.com
NoSQL
MongoDB
Riak
HBase
CouchDB
Redis
MySQL
Percentage of Matching Job Postings
0.6
0.4
0.2
0
Jan '06
Jan '07
Jan '08
Jan '09
Jan '10
Jan '11
Jan '12
Jan '13
Jan '14
indeed
one search. all jobs.

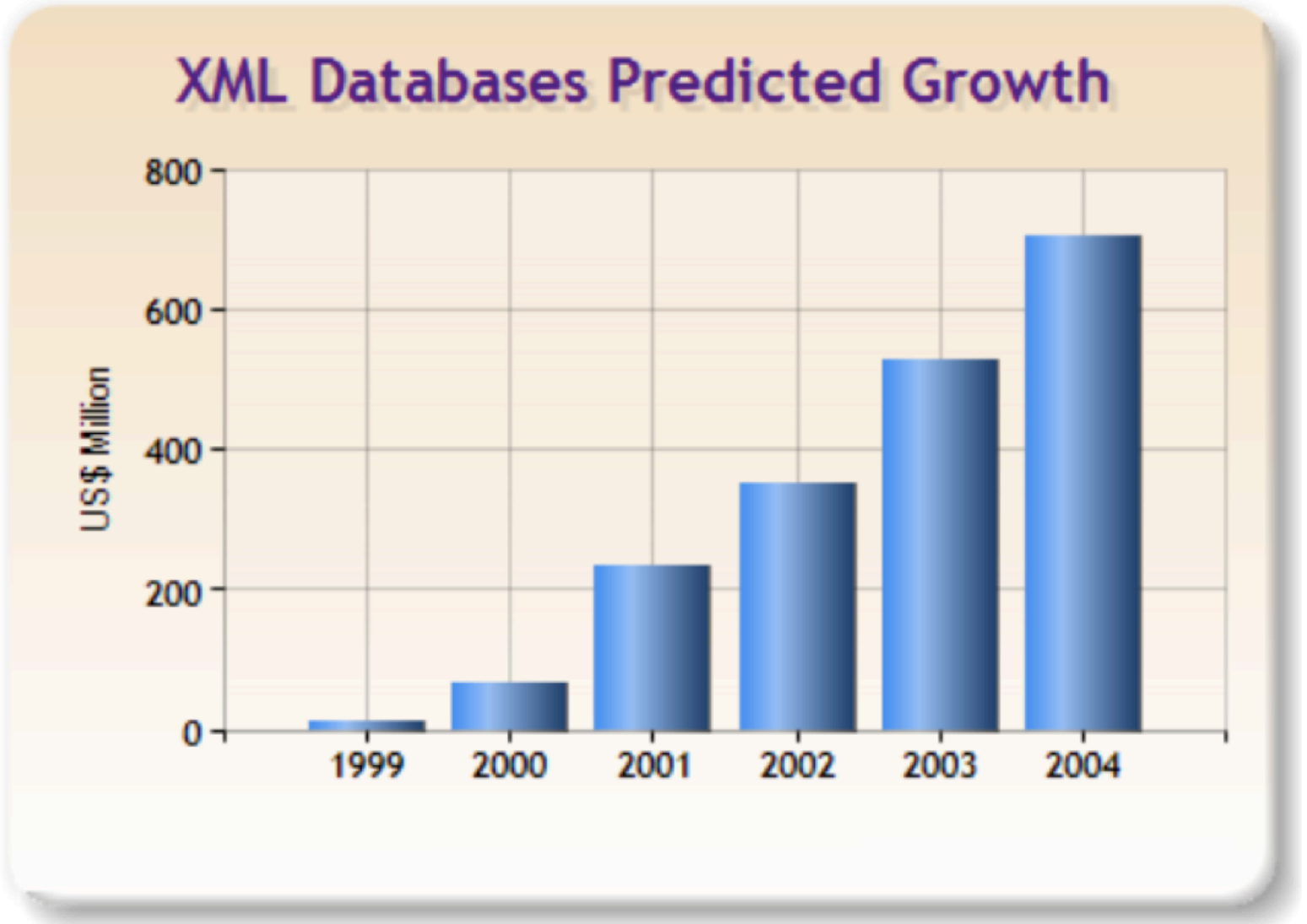
Hype?


Job Trends from Indeed.com
NoSQL
MongoDB
Riak
HBase
CouchDB
Redis
MySQL
SQL
Percentage of Matching Job Postings
0
1
2
3
4
Jan '06
Jan '07
Jan '08
Jan '09
Jan '10
Jan '11
Jan '12
Jan '13
Jan '14
indeed
one search. all jobs.

Hype?


OO Databases Predicted Growth
US$ Million
2000
1500
1000
500
0
1996
1997
1998
1999
2000

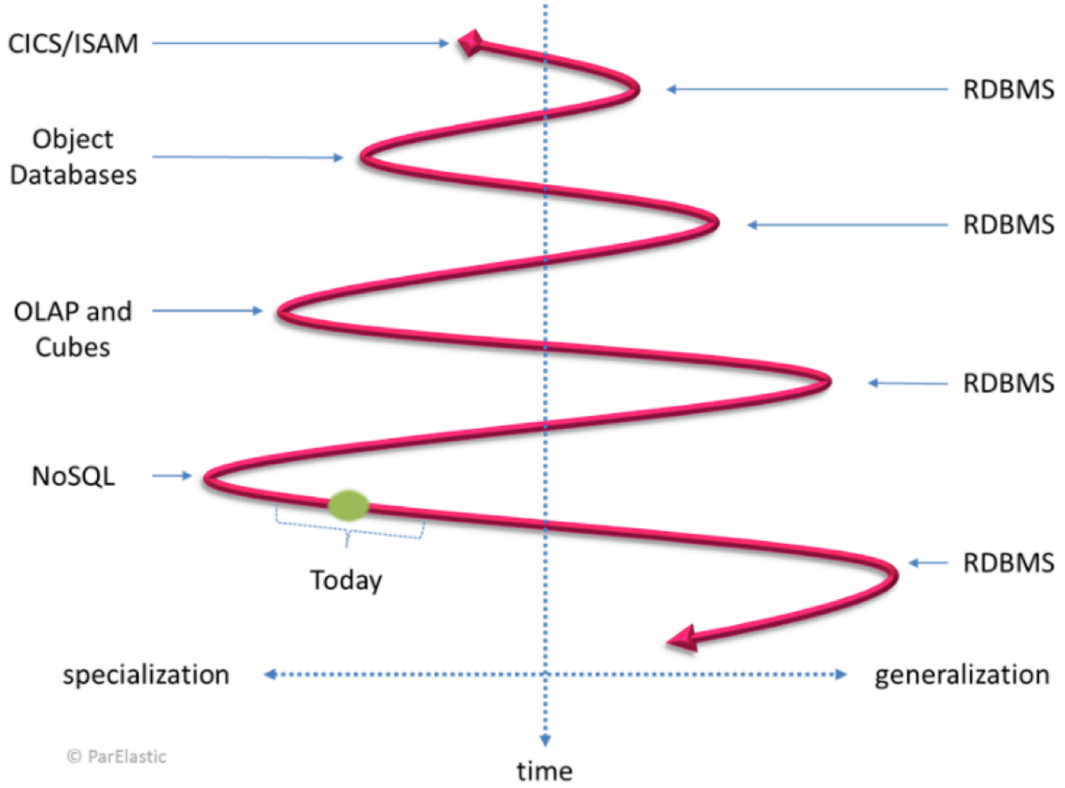
Hype?


XML Databases Predicted Growth
US$ Million
800
600
400
200
0
1999
2000
2001
2002
2003
2004

Hype?


NoSQL Databases Predicted Growth
US$ Million
1200
1000
800
600
400
200
0
2012
2013
2014
2015
2016

Hype?


CICS/ISAM
Object Databases
OLAP and Cubes
NoSQL
Today
RDBMS
RDBMS
RDBMS
RDBMS
specialization
generalization
time
© ParElastic

Future directions

- Internal polyglot support
- Multi-model systems
- NewSQL: Can you have a scalable databases without going NoSQL? (“beating” CAP)
- Further support for NoSQL in RDBMs
- DBaaS and Baas

Meanwhile at

NewSQL

This backend was originally based on **a MySQL database that was manually sharded** many ways. The uncompressed dataset is tens of terabytes, which is small compared to many NoSQL instances, but was large enough to cause difficulties with sharded MySQL. The MySQL sharding scheme assigned each customer and all related data to a fixed shard. This layout enabled the use of indexes and complex query processing on a per-customer basis, but **required some knowledge of the sharding in application business logic**. Resharding this revenue-critical database as it grew in the number of customers and their data was extremely costly. **The last resharding took over two years of intense effort**, and involved coordination and testing across dozens of teams to minimize risk.

NewSQL

We store financial data and have **hard requirements on data integrity and consistency**. We also have a lot of experience with eventual consistency systems at Google. In all such systems, we find developers spend a significant fraction of their time **building extremely complex and error-prone mechanisms to cope with eventual consistency** and handle data that may be out of date. We think this is an unacceptable burden to place on developers and that consistency problems should be solved at the database level.

NewSQL

Some Requirements:

Scalable
- By adding hardware
- No manual sharding

Available
- No downtime, ever

Consistent
- Full ACID

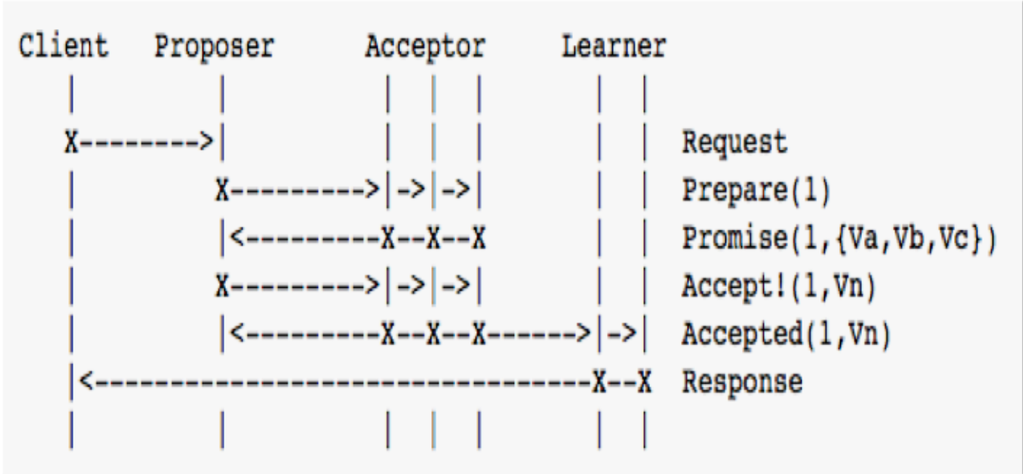
Usable
- Full SQL with indexes

At least 300 applications within Google use Megastore (despite its relatively low performance) because its **data model is simpler to manage** than Bigtable's, and because of its **support for synchronous replication** across datacenters. (Bigtable only supports eventually-consistent replication across datacenters.) Examples of well-known Google applications that use Megastore are **Gmail, Picasa, Calendar, Android Market, and AppEngine**.

NewSQL

Spanner:

- Semi-relational distributed DB
  - SQL queries
  - Versioned data
- Consistent reads/writes
- Atomic schema updates
- High availability
- Removing nodes has no effect except on throughput (PC/EC)

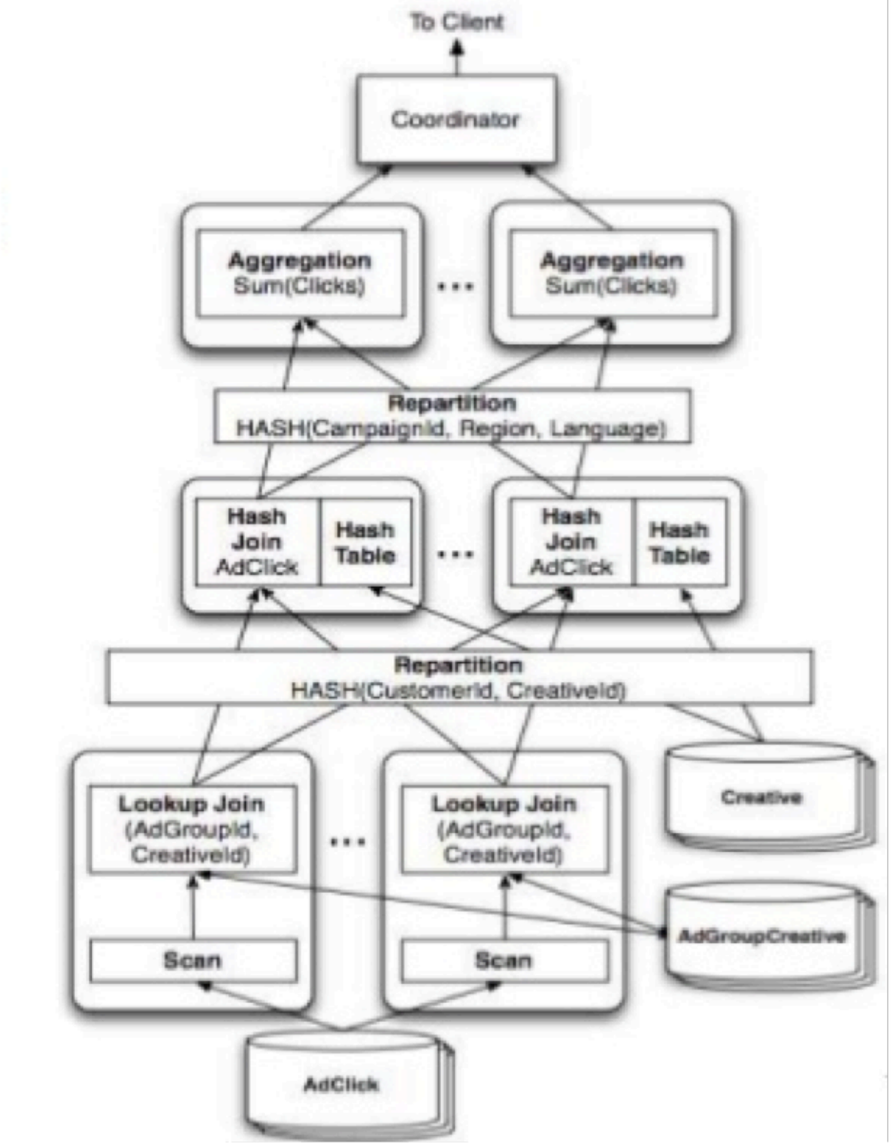


TrueTime:

- Timestamp: consistent ordering on transactions
- Uses: GPS and atomic clocks

F1:

- Distributed SQL queries
- Consistent indexes
- Automatic change history (triggers)



NewSQL

F1:

- 100 TB of uncompressed data
  - Over 5 data centers
  - Five nines uptime

— Hundreds of thousands of request/second

— SQL queries scan trillions of rows/day

— No observable increase of latency compared to MySQL

— NoSQL (key -> row) and full SQL

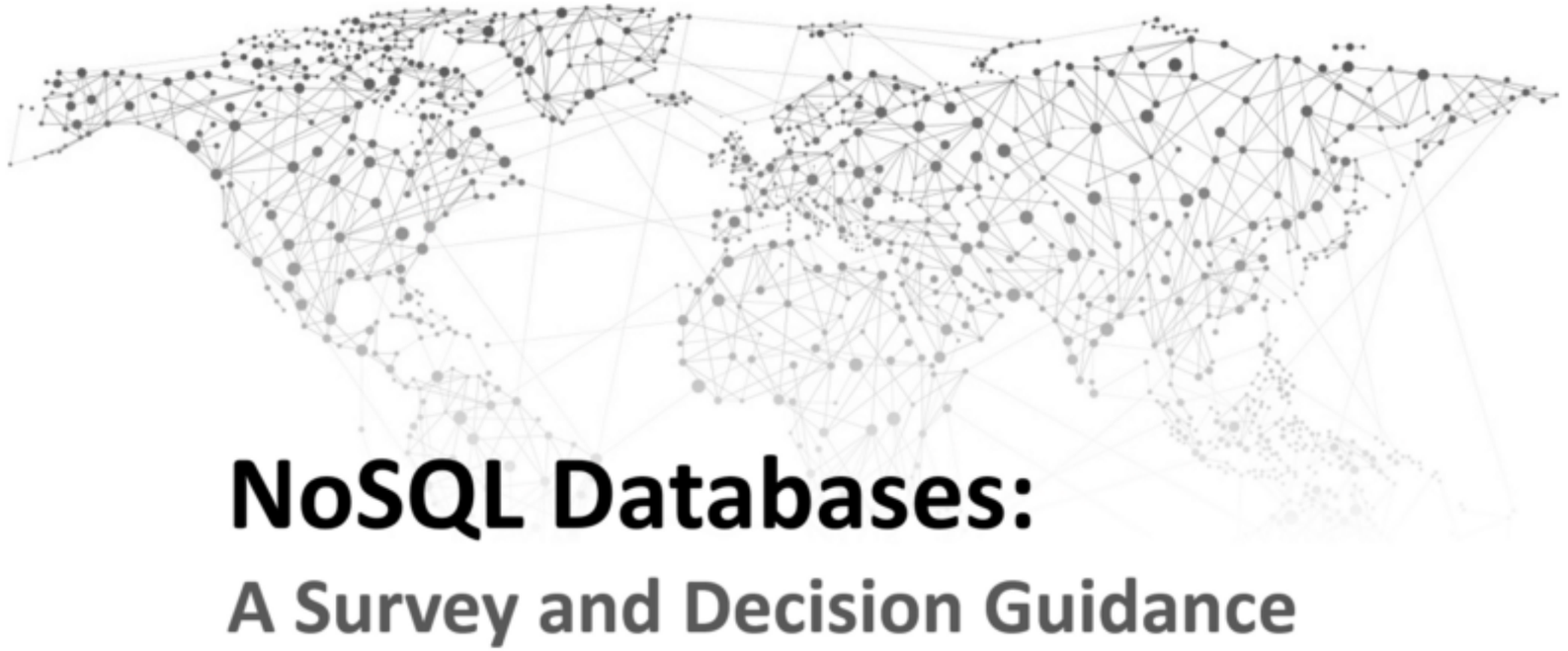
operation	latency (ms)		count
	mean	std dev	
all reads	8.7	376.4	21.5B
single-site commit	72.3	112.8	31.2M
multi-site commit	103.0	52.2	32.1M

Beware of vendor speak

What vendor says	What vendor means
The biggest in the world	The biggest one we've got
The biggest in the universe	The biggest one we've got
There is no limit to ...	It's untested, but we don't mind if you try it
A new and unique feature	Something the competition has had for ages
Currently available feature	We are about to start Beta testing
Planned feature	Something the competition has, that we wish we had too, that we might have one day
Highly distributed	International offices
Engineered for robustness	Comes in a tough box

Further reading

Recent NoSQL survey and decision guide



Questions?