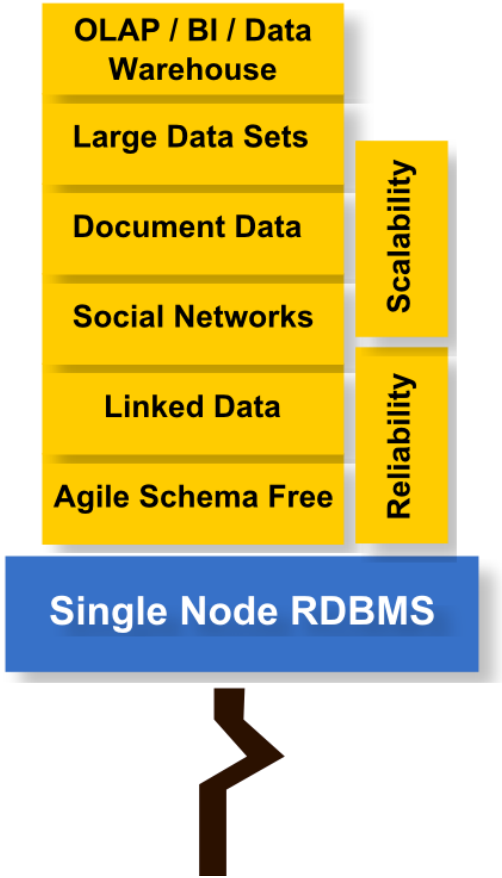
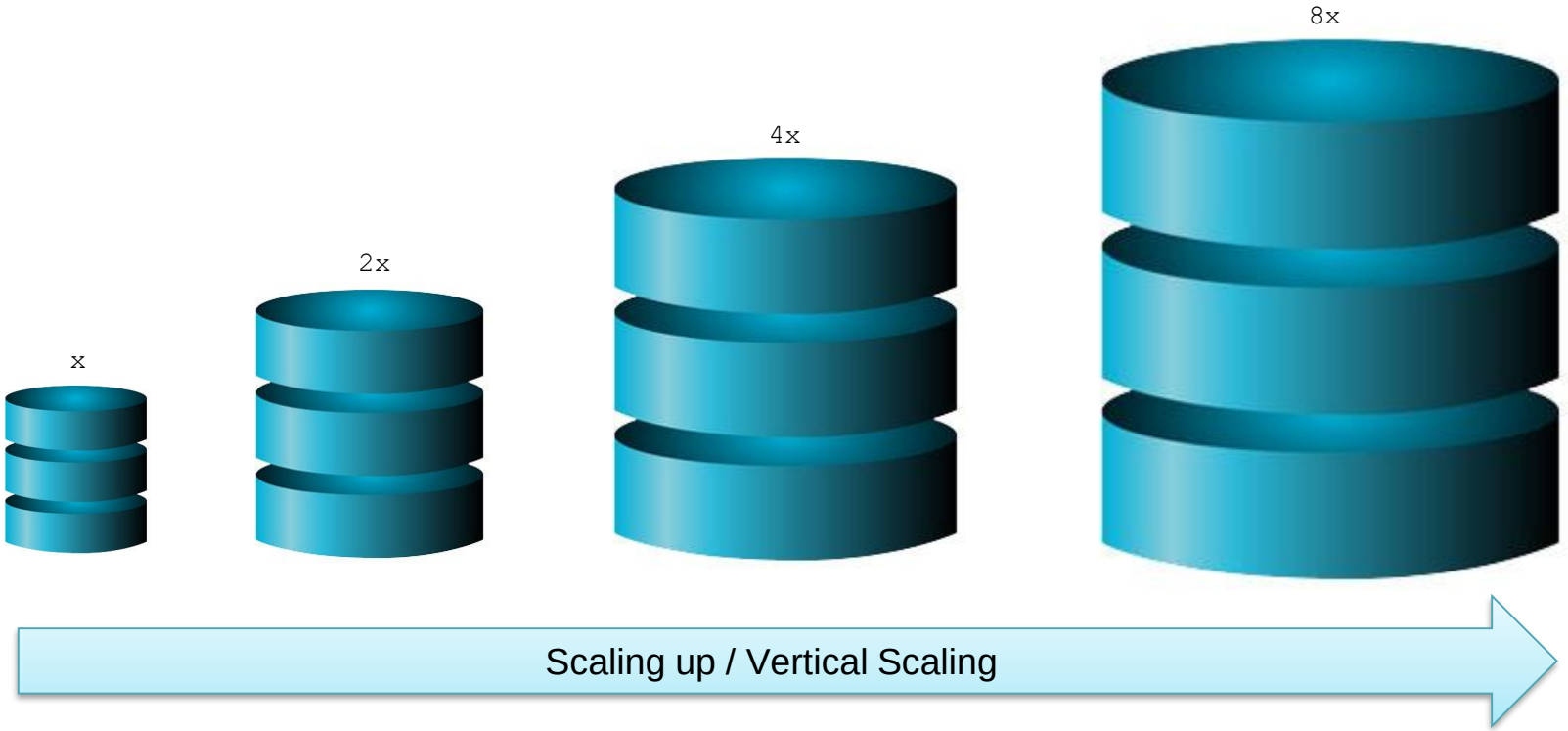
Database Classification

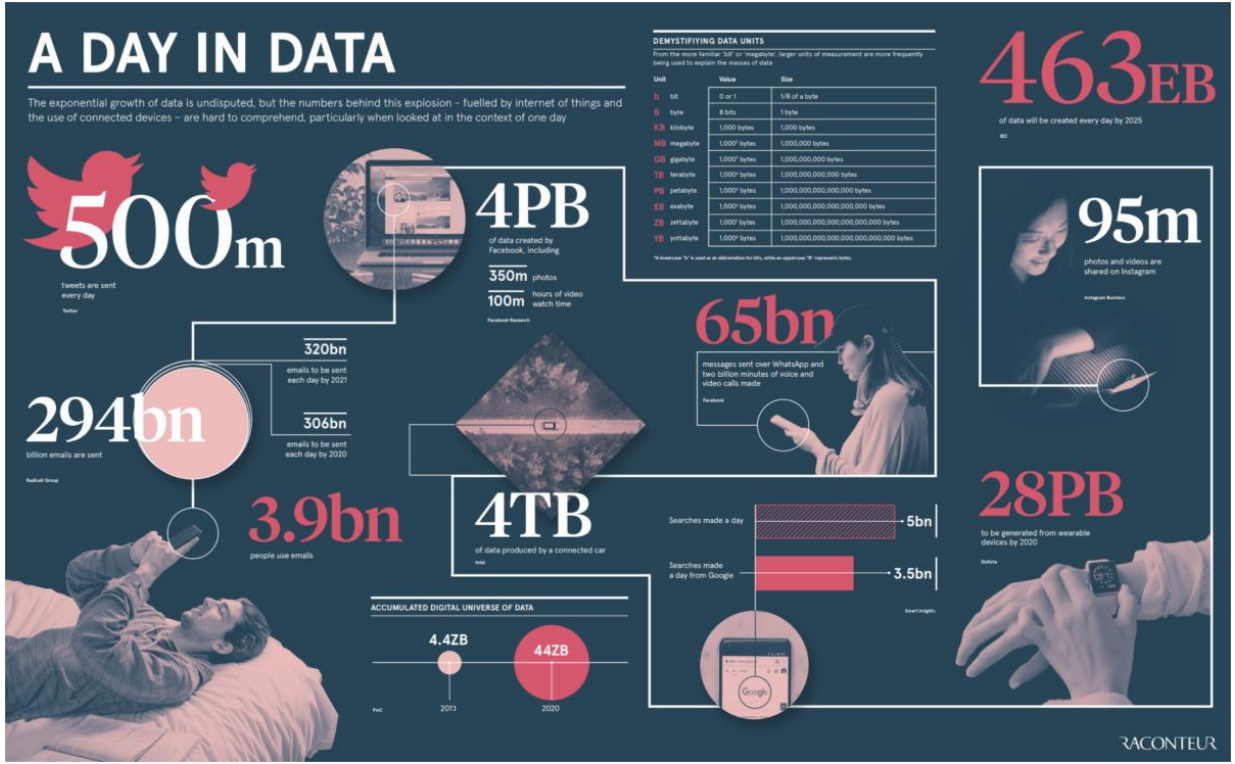
Pressure on Single Node RDBMS Architectures



Scaling the RDBMS Architectures



Data all around us !!!!

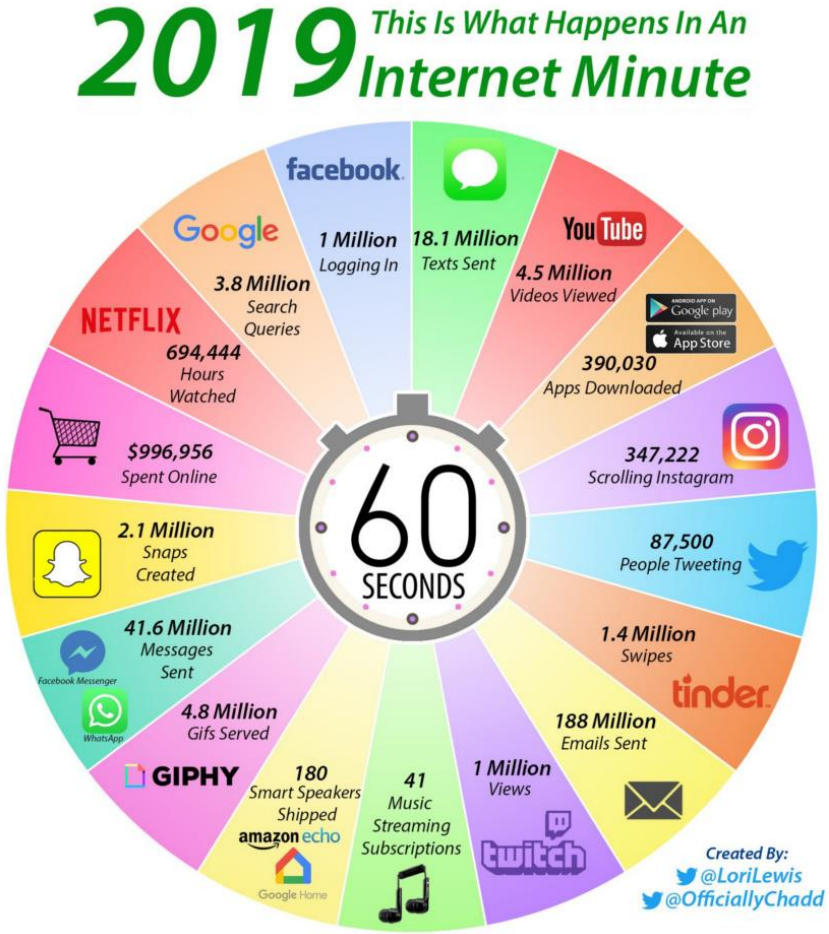


Source: https://www.visualcapitalist.com

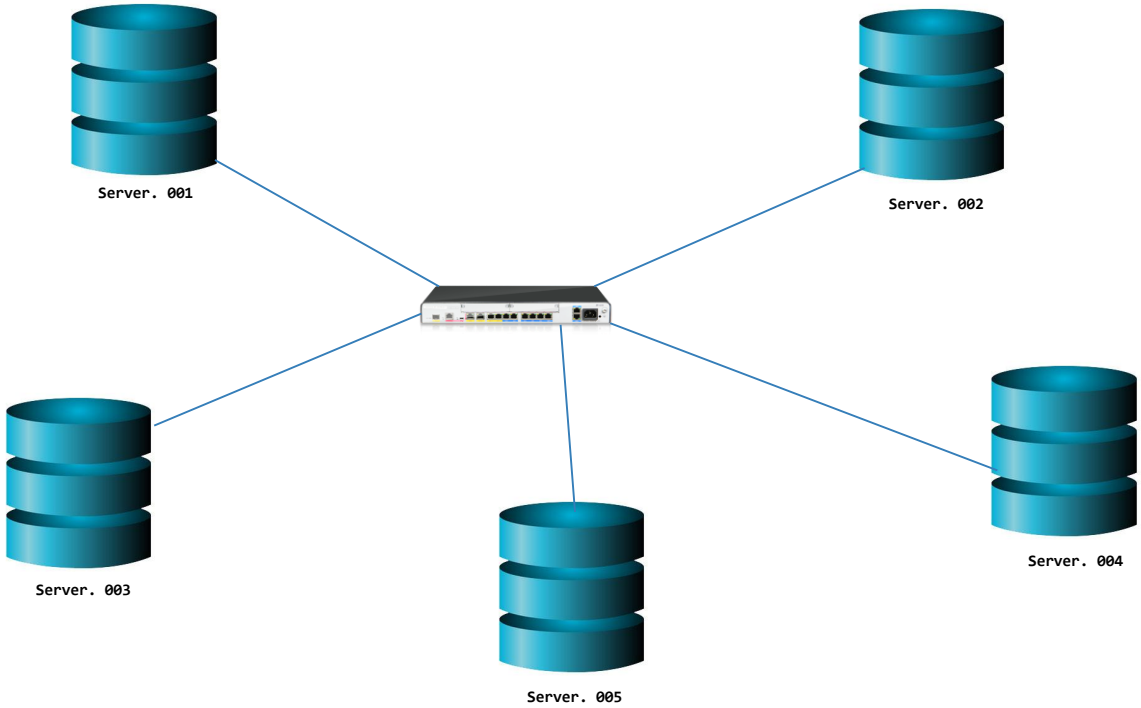
Per Day

- 500 million tweets are sent
- 294 billion emails are sent
- 4 petabytes of data are created on Facebook
- 4 terabytes of data are created from each connected car
- 65 billion messages are sent on WhatsApp
- 5 billion searches are made

Per minute!!!!

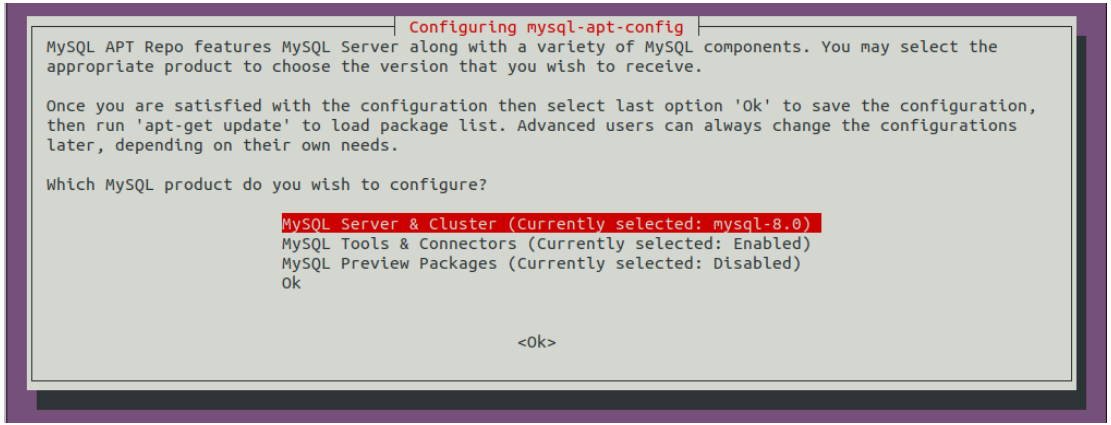

2019 This Is What Happens In An Internet Minute
facebook
1 Million
Logging In
18.1 Million
Texts Sent
YouTube
4.5 Million
Videos Viewed
390,030
Apps Downloaded
347,222
Scrolling Instagram
87,500
People Tweeting
1.4 Million
Swipes
tinder
188 Million
Emails Sent
1 Million
Views
twitch
41
Music Streaming Subscriptions
180
Smart Speakers Shipped
amazon echo
Google Home
4.8 Million
Gifs Served
GIPHY
41.6 Million
Messages Sent
Facebook Messenger
WhatsApp
2.1 Million
Snaps Created
$996,956
Spent Online
NETFLIX
694,444
Hours Watched
Google
3.8 Million
Search Queries
60
SECONDS
Created By:
@LoriLewis
@OfficiallyChadd

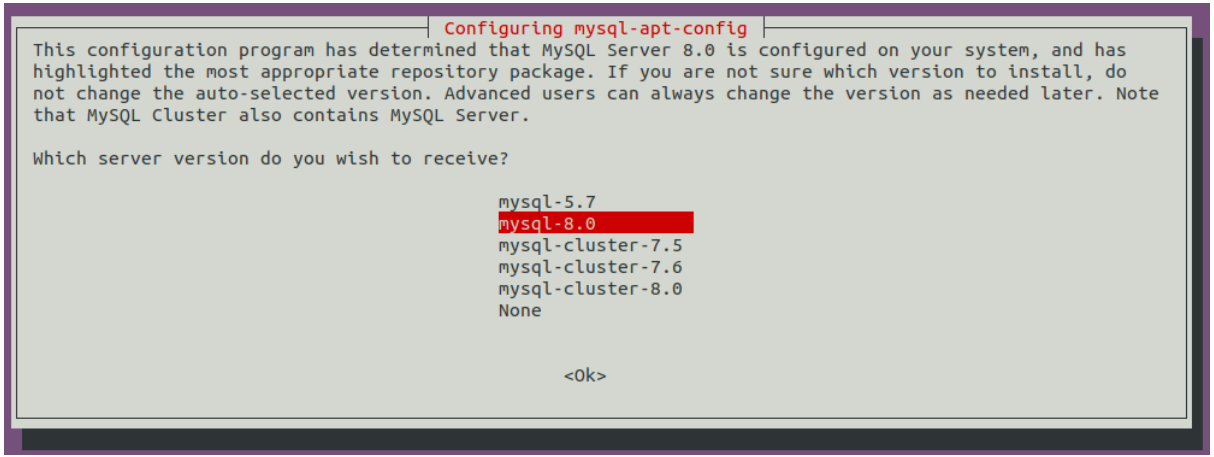
Distributed Databases ???



Horizontal Scaling or Scaling out

MySQL Cluster

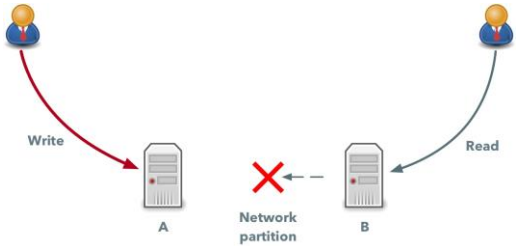




Concerns of Distributed Databases

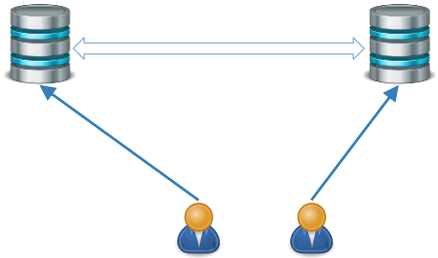
Network Partition Tolerance

System continues to operate despite network partition



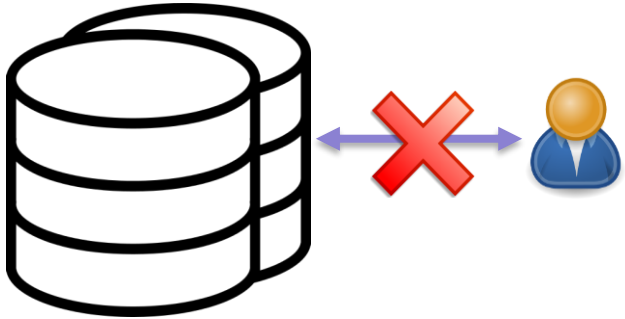
Consistency

Client always has the same view of the data.



Availability

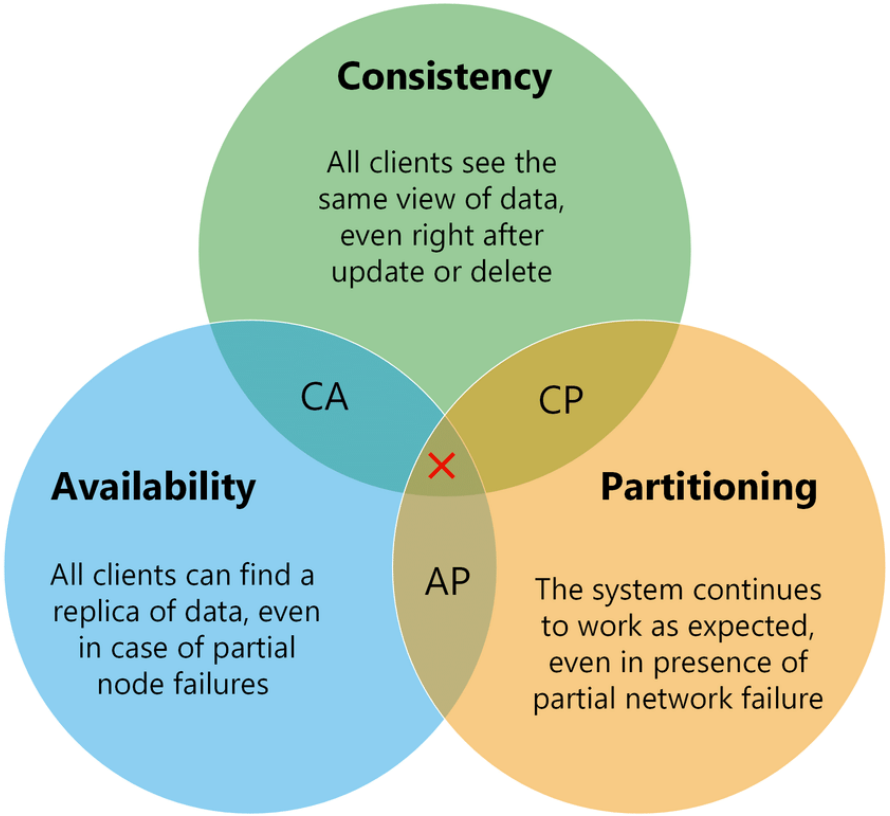
Non-technical error response for read/write request even in the event of not having most latest data.



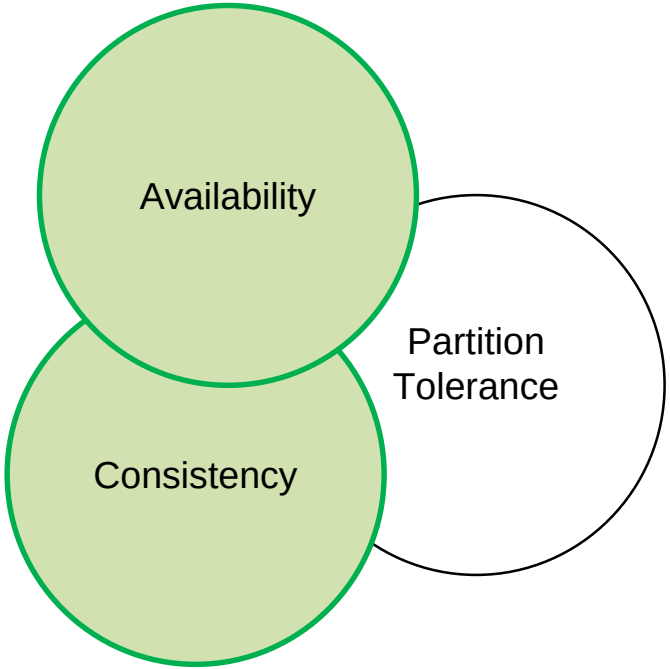
CAP Theory

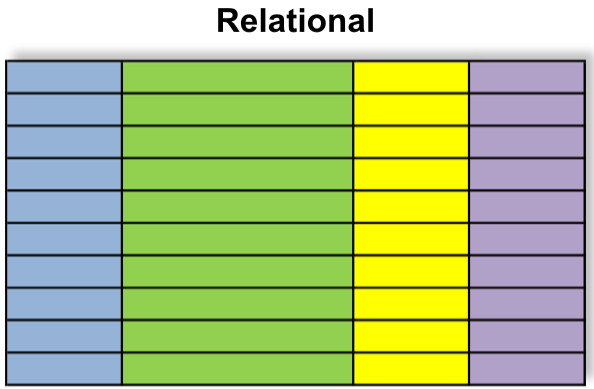
A distributed system can satisfy any two of CAP guarantees at the same time but not all three:

- **Consistency + Availability**
- **Consistency + Partition Tolerance**
- **Availability + Partition Tolerance**



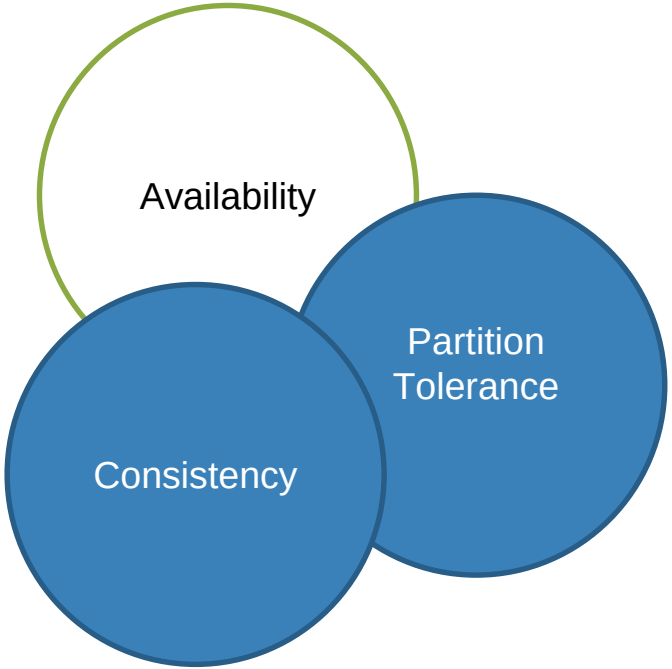
Relational Databases: Consistency + Availability





Examples:
Oracle, MySQL, PostgreSQL, Microsoft SQL Server, IBM DB/2

CAP Theorem - Consistency + Partition Tolerance



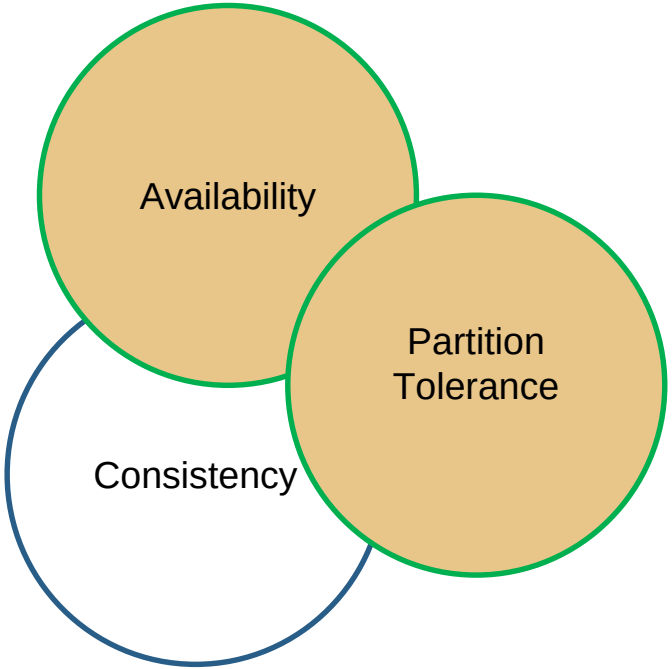
Examples:

- **HBase, MongoDB, Redis, BigTable**
- **Distributed Locking (Apache Zookeeper)**
- **Majority protocols**

Traits::

- **Pessimistic locking**
- **Make minority partitions unavailable**

CAP Theorem - Availability + Partition Tolerance



Examples:

- **Cassandra, Riak, CouchDB**
- **DNS**

Traits::

- **NSPF (No Single Point of Failure)**
- **Conflict resolution**
- **Optimistic**

BASE

Other classification

- Data model
- Operational Capability
- Data variability (Unstructured and semi-structured data)

Eric Evans

“The whole point of seeking alternatives [to RDBMS systems] is that you need to solve a problem that relational databases are a bad fit for.”

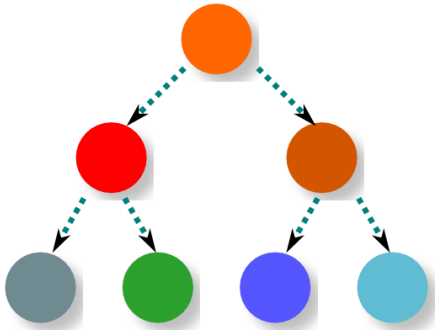
Eric Evans

NoSQL

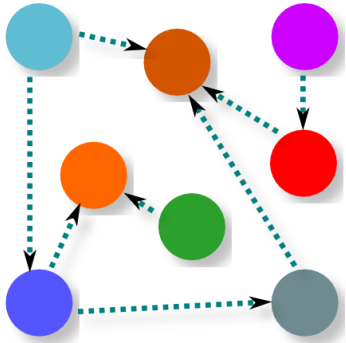
Key - Value

Key	Value
Key	Value
Key	Value
Key	Value

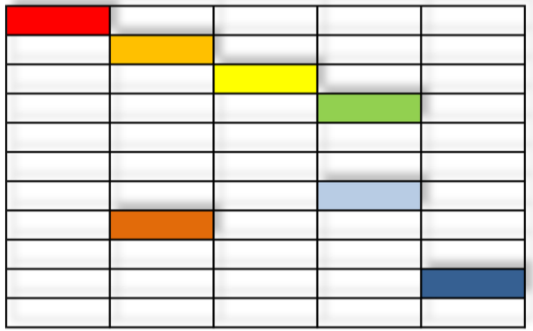
Document



Graph



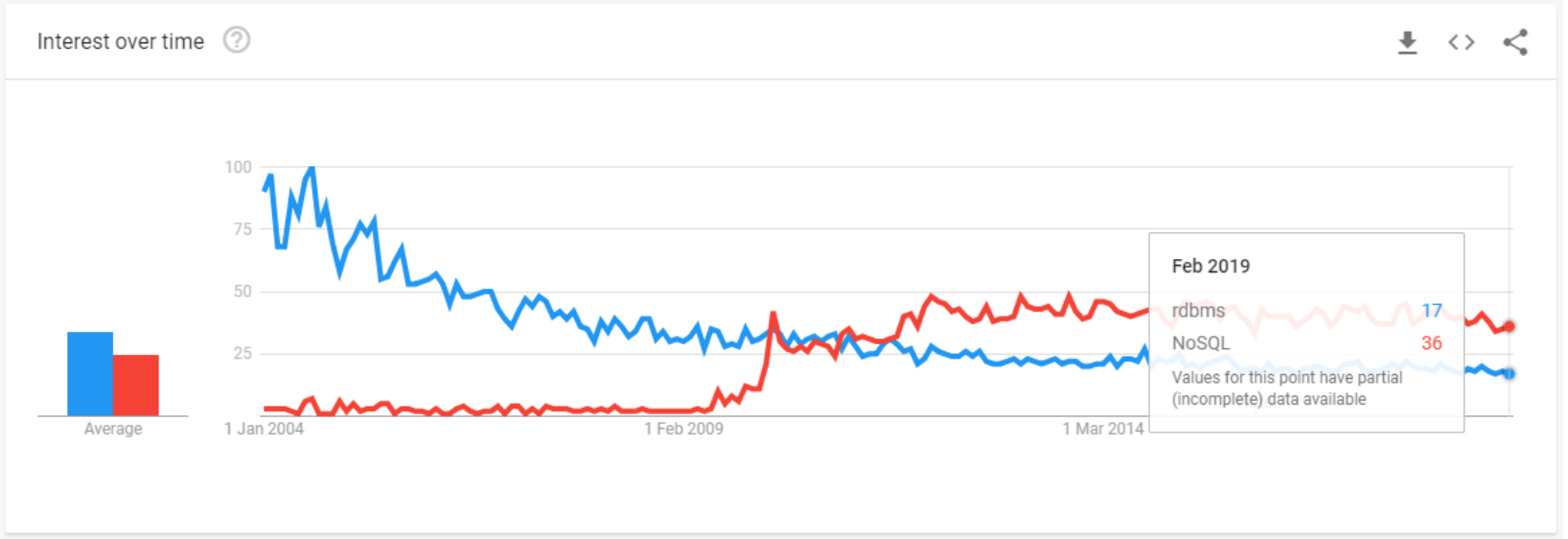
Column Family



The NO-SQL Universe

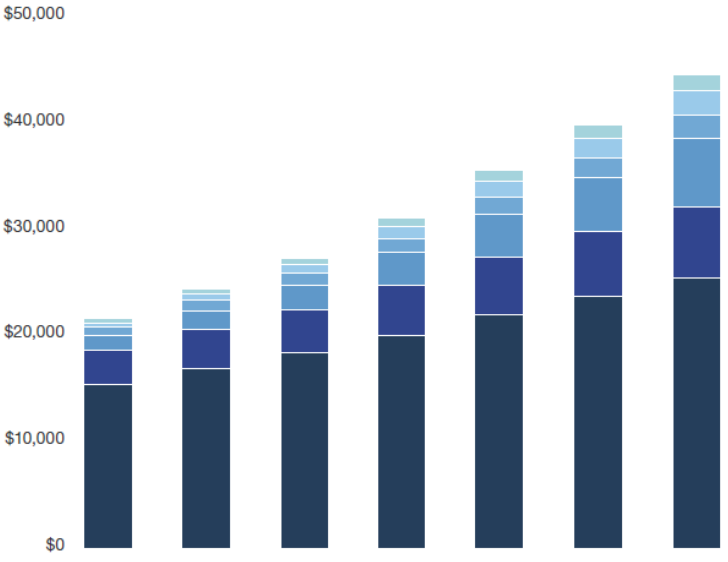
RDBMS vs. NoSQL

https://trends.google.com/trends/explore?date=all&q=rdbms,%2Fm%2F076tfwq



Estimated Big Data and NoSQL Sales

Big data management solutions (US$ millions):



	2015	2016	2017	2018	2019	2020	2021
NoSQL	$415	$518	$642	$819	$1,037	$1,264	$1,579
Hadoop	$435	$554	$768	$1,071	$1,431	$1,800	$2,297
Big data integration	$830	$994	$1,140	$1,353	$1,619	$1,897	$2,192
Data virtualization	$3,257	$3,686	$4,094	$4,663	$5,402	$6,056	$6,716
In-memory data fabric	$1,273	$1,745	$2,286	$3,070	$4,019	$5,069	$6,368
Enterprise data warehouse	$15,472	$16,959	$18,428	$20,113	$22,062	$23,797	$25,478

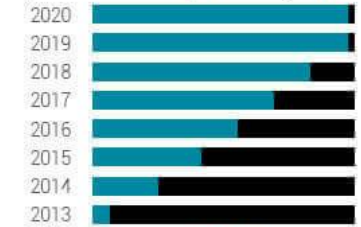
Source: Forrester Data: Big Data Management Solutions Forecast, 2016 To 2021 (Global)

Global NoSQL Market

Size and Forecast (2013 - 2020)

Global NoSQL Market

Global NoSQL Market is expected to reach at **$4.2** billion by **2020**

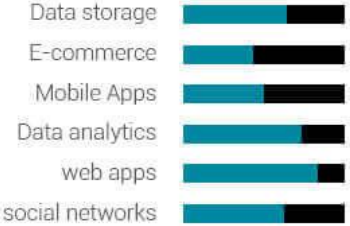


Growing at a CAGR of 35.1% (2014-2020)

Global NoSQL Market By Type

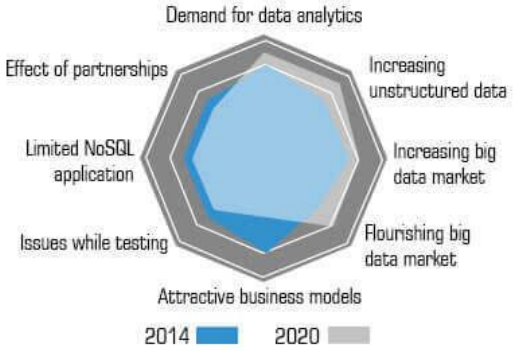
Key-value store to be highest revenue generating segment throughout 2013 - 2020

Global NoSQL Market By Application



Web Apps to be highest revenue generating segment throughout 2013 - 2020

Global NoSQL Market Top Impacting Factors



For More Details See Table of Contents

Global NoSQL Market By Geography

North America to be the highest revenue generating region throughout 2013 – 2020

Global NoSQL Market Dynamics

Drivers:

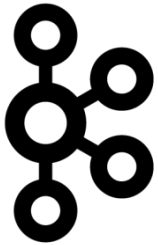
- Rise in unstructured data
- Attractive business models
- Demand for data analytics
- Growing App development business

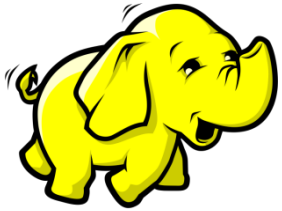
Restraints:

- Difficulties while testing of NoSQL applications
- Problems while adoption in structured data applications

Source : https://www.alliedmarketresearch.com/NoSQL-market

How do we go from here?





Source code