



UAB

Universitat Autònoma
de Barcelona

Andrés Cencerrado

Tecnologías de BBDD NoSQL (I)

Contenido

- **Tecnologías BBDD NoSQL (I)**
 - Origen, necesidad
 - Situación actual
 - Tipos de aproximación: key-value, wide column store, document store
 - Introducción a tecnologías actuales: HBASE, Cassandra, MongoDB
- **Tecnologías BBDD NoSQL (II):**
 - Exploración tecnologías: modelo de datos
 - Exploración tecnologías: arquitectura
 - Consideraciones: ¿cuál escoger?



Origen y necesidad

- Recordando las “4 V”



- **Volumen:** Enorme cantidad de datos.
- **Velocidad:** Generación y movimiento a gran velocidad.
- **Variedad:** Distintos tipos.
- **Veracidad:** El grado de confianza de los datos.

Origen y necesidad

- Necesidad de implementación sistemas NoSQL a gran (enorme) escala



Google
BigTable

"A Bigtable is a sparse, distributed, persistent multidimensional sorted map. The map is indexed by a row key, column key, and a timestamp; each value in the map is an uninterpreted array of bytes."

Fay Chang, Jeffrey Dean, et al. "Bigtable: A Distributed Storage System for Structured Data" – OSDI 2006



Origen y necesidad

- Necesidad de implementación sistemas NoSQL a gran (enorme) escala

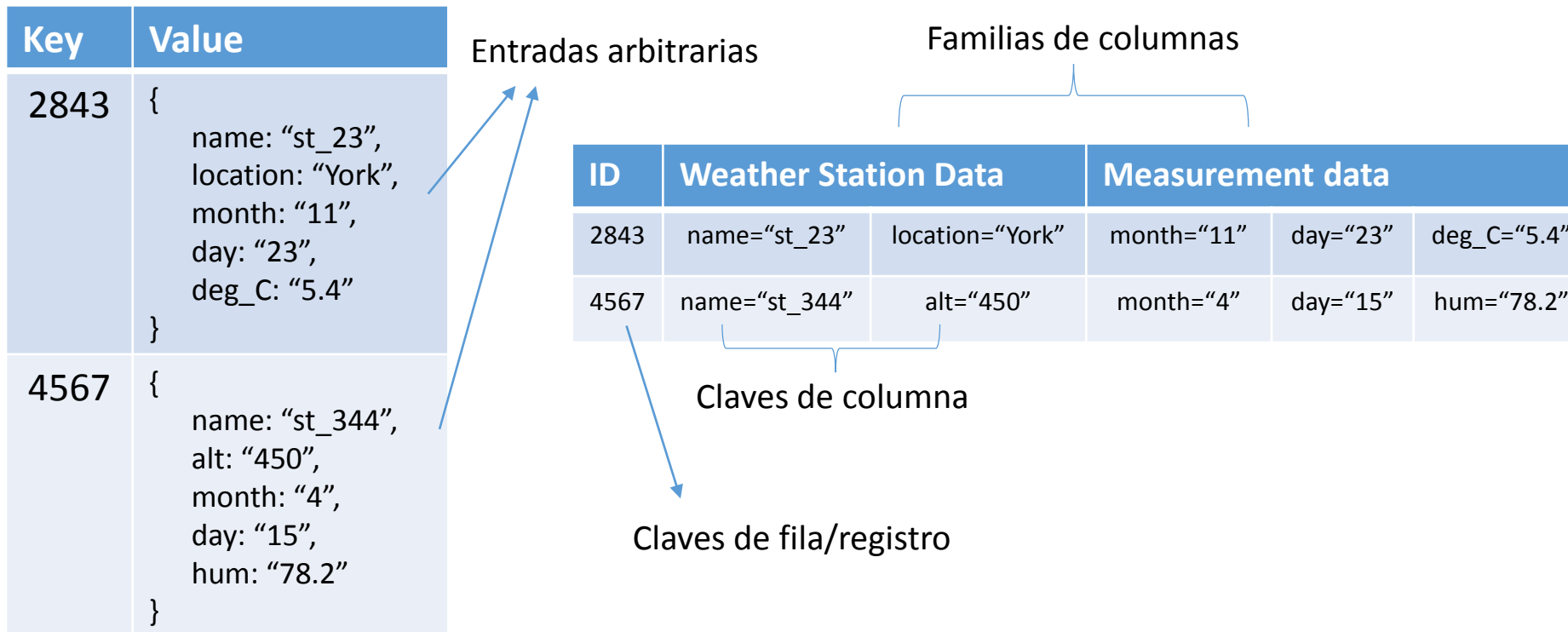


"Dynamo is used to manage the state of services that have very high reliability requirements and need tight control over the tradeoffs between availability, consistency, cost-effectiveness and performance."

Giuseppe DeCandia, Deniz Hastorun, et al. "Dynamo: Amazon's Highly Available Key-value Store" – SOSP 2007



Origen y necesidad



Situación actual

Key-value store:

Key1	V1
Key2	V2
Key3	V3
Key4	V4
Key5	V5

Document store:

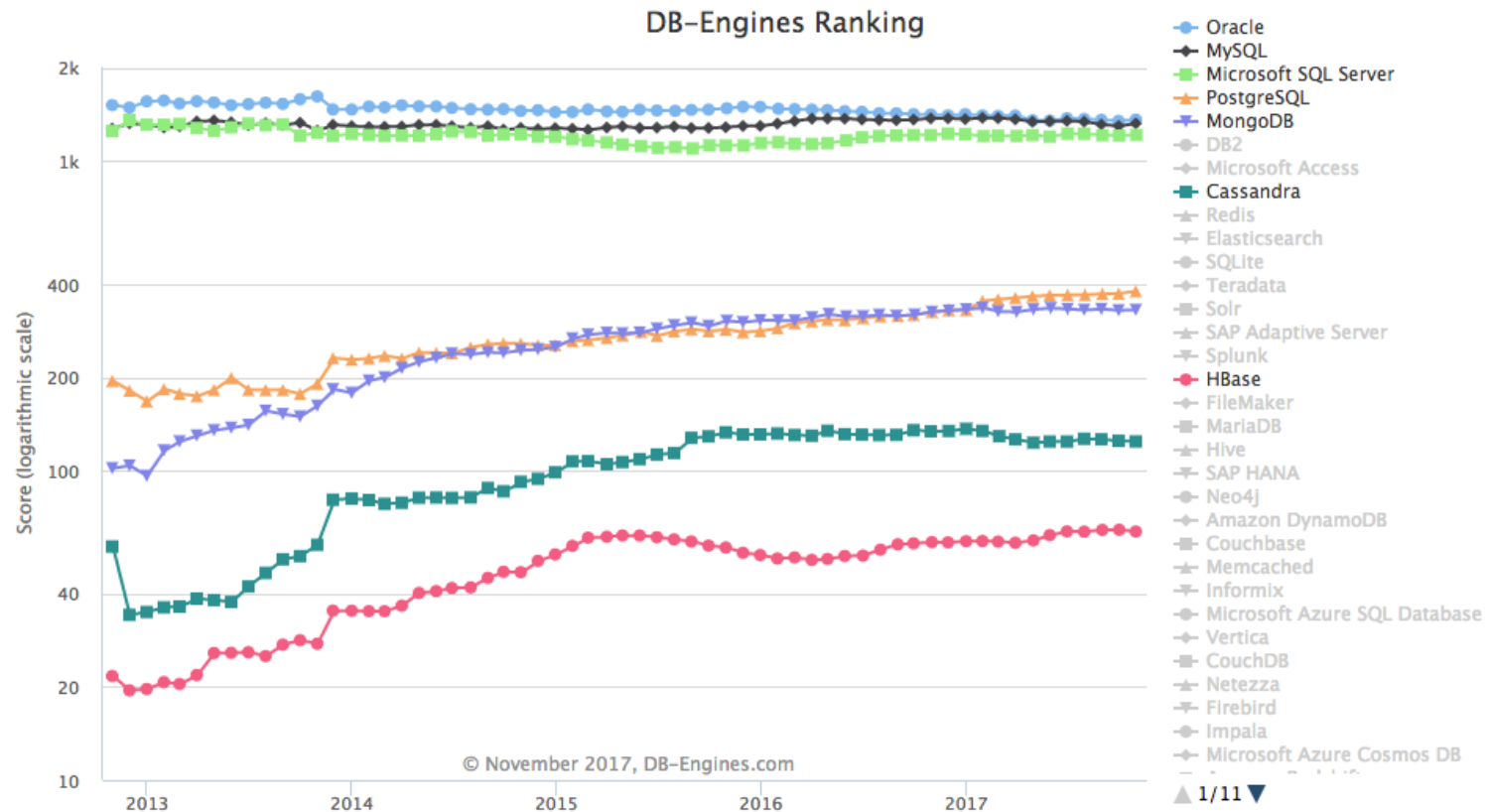
```
ID {
  "created_at": "Thu Jun 22 21:00:00 +0000 2017",
  "id": 877994604561387500,
  "id_str": "877994604561387520", "text": "Creating a
  Grocery List Manager Using Angular, Part 1: Add &
  Display Items https://t.co/xFox78juL1 #Angular",
  "truncated": false,
  "entities": {
    "hashtags": {
      "text": "Angular",
      "indices": [103, 111] },
      "symbols": [], "user_mentions": [],
      "urls": {
        "url": "https://t.co/xFox78juL1",
        "expanded_url": "http://buff.ly/2sr60pf",
        "display_url": "buff.ly/2sr60pf"
      },
    },
    "user": {
      "id": 772682964,
      "id_str": "772682964",
      "location": "Melbourne, Australia"
    },
  },
  "description": { "urls": [] } },
  "protected": false,
  "followers_count": 2145,
  "created_at": "Wed Aug 22 02:06:33 +0000 2012",
  "favourites_count": 57, "utc_offset": 43200,
  "time_zone": "Wellington"
}
```

Wide column store:

ID	Col Fam 1	Col Fam 2	Col Fam 3
	ColKey1: x1	ColKey1: z1	ColKey1: t1
	ColKey2: x2	ColKey2: z2	ColKey2: t2
	ColKey3: x3	ColKey3: z3	
	ColKey4: x4		
	ColKey5: x5		

Situación actual

Actualmente hay 300+ tecnologías DBMS



BIG DATA

Tecnologías BBDD NoSQL (I)

Tecnologías representativas



- 2007, Apache
- Integración Hadoop, HDFS
- Wide column store
- Escrita en Java, desarrollo a partir de BigTable
- Sistema CP (consistencia y tolerancia a particiones)



Tecnologías representativas



- 2007 Facebook – 2009 Apache
- Alta escalabilidad y disponibilidad
- Wide column store
- Escrita en Java, muy buena integración con Hadoop y otras plataformas Big Data
- Sistema AP (disponibilidad y tolerancia a particiones)

NETFLIX



Tecnologías representativas



- 2007 10gen Inc. (MongoDB Inc.)
- Simplicidad en el desarrollo
- Document store
- Escrita en C++, muy buena integración en distintos casos de uso
- Sistema CP (consistencia y tolerancia a particiones)



BOSCH

En la segunda parte...

- Tecnologías BBDD NoSQL (I)
 - Orígen, necesidad
 - Situación actual
 - Tipos de aproximación: key-value, wide-column, document-store
 - Introducción a tecnologías actuales: HBASE, Cassandra, MongoDB
- **Tecnologías BBDD NoSQL (II):**
 - **Exploración tecnologías: modelo de datos**
 - **Exploración tecnologías: arquitectura**
 - **Consideraciones: ¿cuál escoger?**



A green-tinted image of a tablet displaying the word "MOOC" in a bold, sans-serif font.

MOOC

Tecnologías de BBDD NoSQL (I)

A green-tinted image of a laptop keyboard with the UAB MOOC logo overlaid. The logo includes the UAB acronym, the full name of the university, and the name of the postgraduate school.

UAB
Universitat Autònoma de Barcelona
MOOC
Escola de Postgrau

UAB
Universitat Autònoma
de Barcelona