

# Who am I?

Vitaly Kushner

- Founder of **Astrails**
- 19 years of industry experience
- Building Web Apps since 2005

# NoSQL

when, why and how?



Vitaly Kushner  
astrails.com

Google

Yahoo

Google

Yahoo

Google

Amazon

Yahoo

Facebook

Google

Amazon

Yahoo

Facebook

Twitter

Google

Amazon

Yahoo

Digg

Facebook

Twitter

Google

Amazon

Yahoo

Digg

Facebook

Twitter

Google

Amazon

Rackspace

Yahoo

Digg

Facebook

Twitter

Google

LinkedIn

Amazon

Rackspace

Yahoo

Digg

Facebook

Twitter

Google

Everybody

LinkedIn

Amazon

Rackspace



# WTF is NoSQL?

and why should you care?

Non relational

Document based

Non relational

Document based

Key-Value store

Non relational

Document based

Key-Value store

Non relational

column-based

Document based

Key-Value store

# Non relational

column-based

Graph DB

Document based

Key-Value store

# Non relational

column-based

Distributed

Graph DB

Schema-less

Document based

Key-Value store

Non relational

column-based

Distributed

Graph DB

Schema-less

Document based

Key-Value store

# Non relational

column-based

Distributed

**BASE** is not ACID

Graph DB

# Why & When

# Massive Data Volume

100K servers in a cluster

# Massive Data Volume

100K servers in a cluster

Twitter: 7+T/day

# High query workload

MongoDB: 8M operations/sec

# Flexible Schema

on the fly schema changes

# Massive Scale

# Availability

Everyone want  
Availability

RDBMS  
can deliver

high price

Not ACID anymore

# CAP theorem

Pick two

- Consistent
- Available
- Partition tolerant

# Scale

How?

Throw ~~hardware~~  
money at it!

# Par-ti-tion

MySQL + Memcached

=

“square wheel”

cassandra

How?

# Which one?

- document based
- column or key-value store
- advanced storage schemas

# Cassandra

- built by Facebook
- very high write throughput
- OLTP
- automatic horizontal scaling
- no single point of failure

# HBase

- Apache project
- Consistent
- Optimized for analytics (OLAP)
- Has single point of failure

# MongoDB

- probably **easiest** to move to from SQL
- document based
- on-demand queries
- automatic sharding
- **no single-node durability**

# CoachDB

- document based
- map-reduce javascript querying/filtering.
- has some replication and scaling problems

# REDIS

- key-value store
- advanced data types: list, set
- atomic operations

# Schema

```
Users: {
```

```
  vitaly: {
```

```
    email: vitaly@astrails.com,
```

```
    company: astrails,
```

```
    password: secret
```

```
  },
```

```
  michael: {
```

```
    email: michael@astrails.com,
```

```
    company: astrails,
```

```
    password: superduper
```

```
  },
```

```
  ...
```

```
}
```

```
UsersByEmail: {  
  "vitaly@astrails.com": "vitaly",  
  "michael@astrails.com": "michael",  
  ...  
}
```

# Migrations

Start Slow

NoSQL can help  
You

# WTF is NoSQL?

Q & A



*astrails*

@astrails

Vitaly Kushner

astrails.com

@vkushner