ToroDB

A new, open-source,
document-oriented, JSON database,
built on Postgres

Álvaro Hernández <*aht@8kdata.com*>

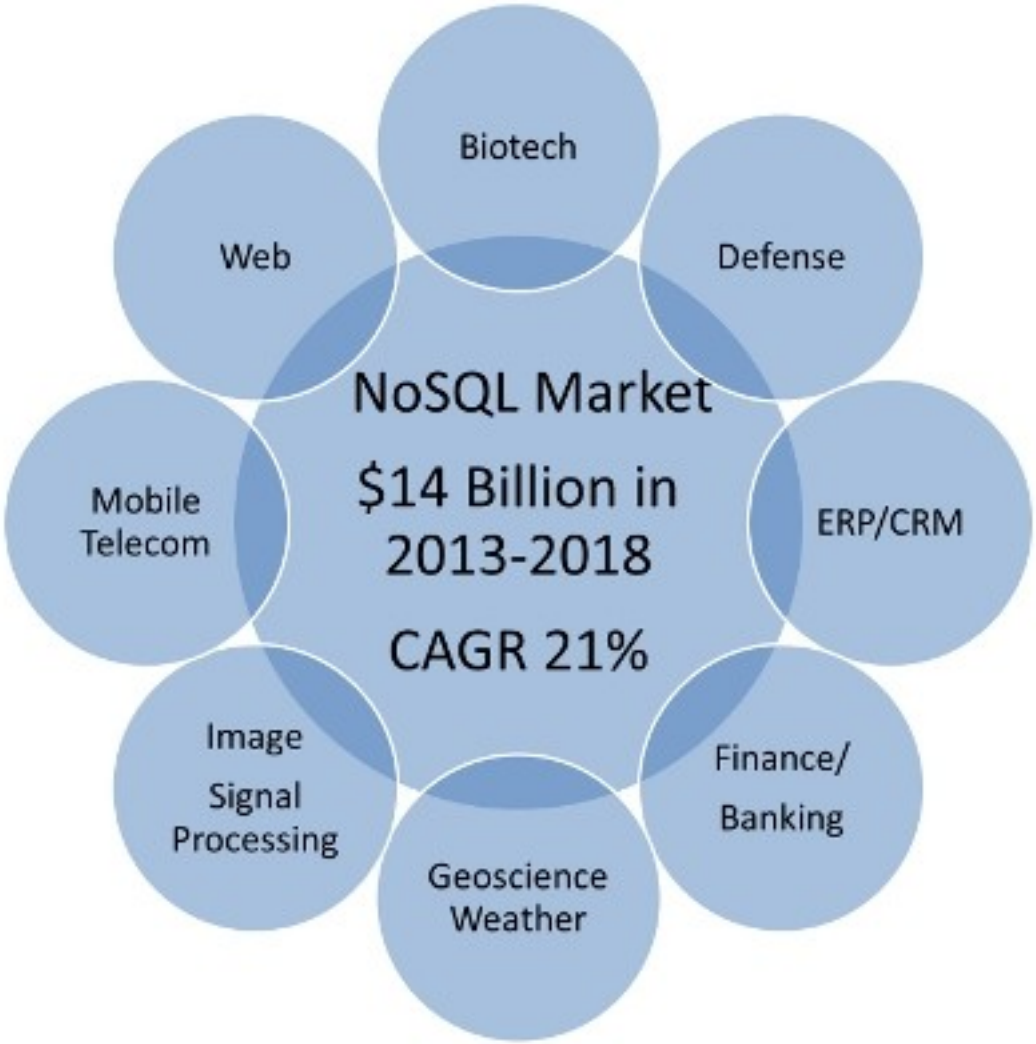
About ⟨8K⟩ data

www.8kdata.com

- Research & Development in databases
- Consulting, Training and Support in PostgreSQL
- Founders of PostgreSQL España, 3rd largest PUG in the world (>350 members as of today)
- About myself: CTO at 8Kdata:

@ahachete
http://linkd.in/1jhvzQ3

How big is “NoSQL”?



Source: 451 Research

Why people want “NoSQL”?

- Schema-less
- High availability
- It's cool

The schema-less fallacy

```
{
  “name”: “Álvaro”,
  “surname”: “Hernández”,
  “height”: 200,
  “hobbies”: [
    “PostgreSQL”, “triathlon”
  ]
}
```

The schema-less fallacy

```
{
  "name": "Álvaro",
  "surname": "Hernández",
  "height": 200,
  "hobbies": [
     "PostgreSQL", "triathlon"
  ]
}
```

metadata → Isn't that... **schema**?

The schema-less fallacy: BSON

```
{
  “name”: (string) “Álvaro”,
  “surname”: (string) “Hernández”,
  “height”: (number) 200,
  “hobbies”: {
        “0”: (string) “PostgreSQL” ,
        “1”: (string) “triathlon”
  }
}
```

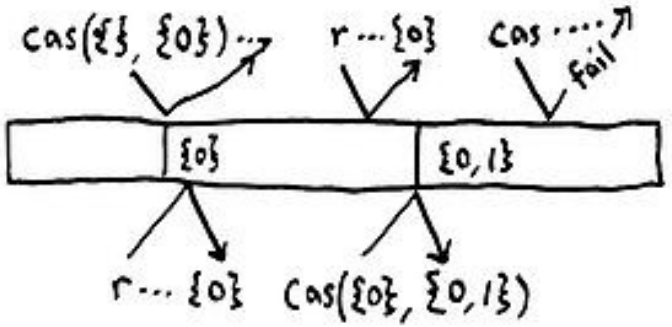
metadata → Isn't that... **schema**?

The schema-less fallacy

- It's not schema-less
- It is *"attached-schema"*
- It carries an overhead which is not 0

High availability: at what cost?

Jepsen!!! :)



cas, cas, cas, cas, -.. read final
+0 +1 +2 +3 ↑ {0,1,2,3,...}
?

MongoDB:

- Unacknowledged: 42% data loss
- Safe: 37% data loss
- Only majority is safe

http://aphyr.com/posts/284-call-me-maybe-mongodb

More NoSQL struggle

- Durability is sometimes not guaranteed on a single node
- Programming for AP systems may be a big burden
- Most (all?) NoSQL databases wrote their storage from scratch. **Journaling, concurrency are really hard**

Can we do a better "NoSQL"?

- Document model is very appealing to many. Let's offer it
- DRY: why not use relational databases? They are proven, durable, concurrent and flexible
- Why not base it on relational databases, like PostgreSQL?

Schema-attached repetition

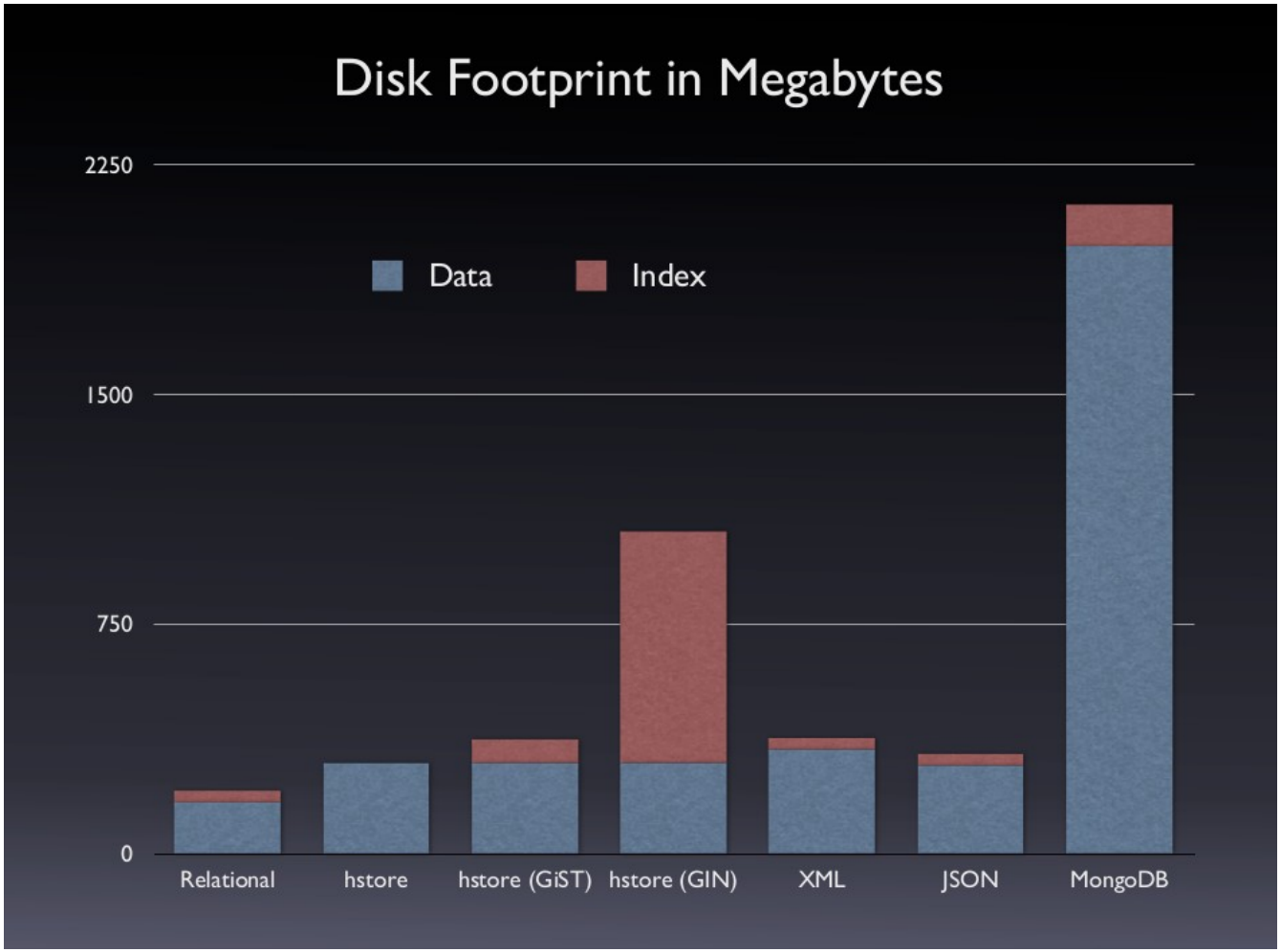
```
{ "a": 1, "b": 2 }
{ "a": 3 }
{ "a": 4, "c": 5 }
{ "a": 6, "b": 7 }
{ "b": 8 }
{ "a": 9, "b": 10 }
{ "a": 11, "b": 12, "¡": 13 }
{ "a": 14, "c": 15 }
```

Counting "document types" in collections of millions: at most, 1000s of different types

Schema-attached repetition

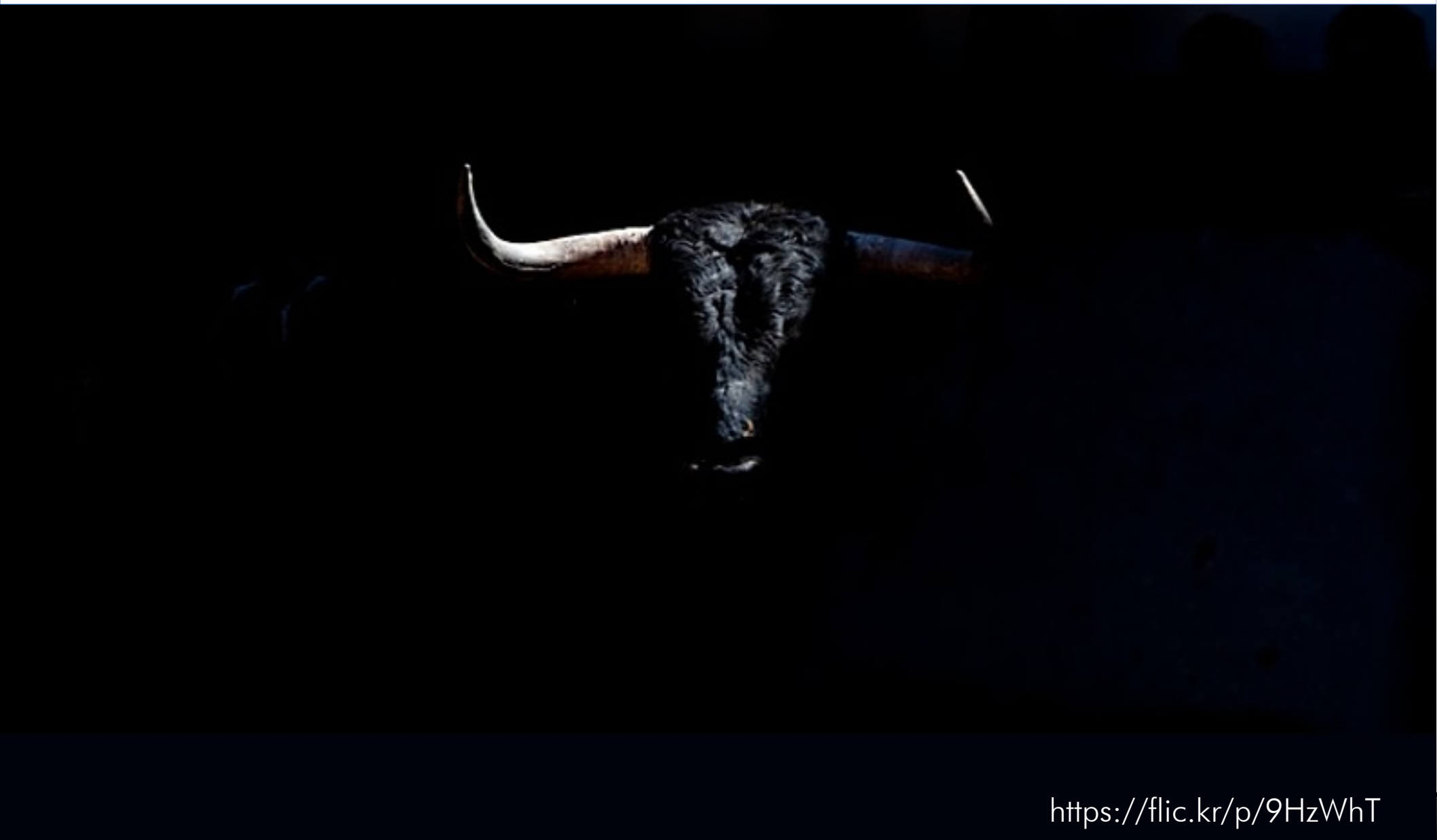
How data is stored in schema-less

Pettus and BTP inspired us



https://wiki.postgresql.org/images/b/b4/Pg-as-nosql-pgday-fosdem-2013.pdf
http://www.slideshare.net/nosys/billion-tables-project-nycpug-2013

ToroDB



https://flic.kr/p/9HzWhT

What is ToroDB

- Open source, document-oriented, JSON database that runs on top of PostgreSQL

- JSON documents are stored relationally, not as a blob: significant storage and I/O savings

- Wire-protocol compatibility with Mongo

ToroDB benefits

- 100% durable database
- High concurrency and performance
- Compatible with existing mongo API programs, clients
- Full set of JSON operations (MongoDB's "SELECT" API)

ToroDB storage

- Data is stored in tables
- JSON documents are split by hierarchy levels, and each (plain) level goes to a different table
- Subdocuments are classified by “type”, which maps to tables

ToroDB storage (II)

- A "structure" table keeps the subdocument "schema"
- Keys in JSON are mapped to attributes, which retain the original name
- Tables are created dinamically and transparently to match the exact types of the documents

ToroDB storage (III)

How data is stored in ToroDB

ToroDB storage internals

```
{
    "name": "ToroDB",
    "data": {
        "a": 42, "b": "hello world!"
    },
    "nested": {
        "i": 42,
        "deeper": {
            "a": 21, "b": "hello"
        }
    }
}
```

ToroDB storage internals

The document is split into the following subdocuments:

{ "name": "ToroDB", "data": {}, "nested": {} }

{ "a": 42, "b": "hello world!"}

{ "j": 42, "deeper": {}}

{ "a": 21, "b": "hello"}

ToroDB storage internals

```
select * from demo.t_3
```

did	index	_id	name
0	¤	\x5451a07de7032d23a908576d	ToroDB

```
select * from demo.t_1
```

did	index	a	b
0	¤	42	hello world!
0	1	21	hello

```
select * from demo.t_2
```

did	index	j
0	¤	42

ToroDB storage internals

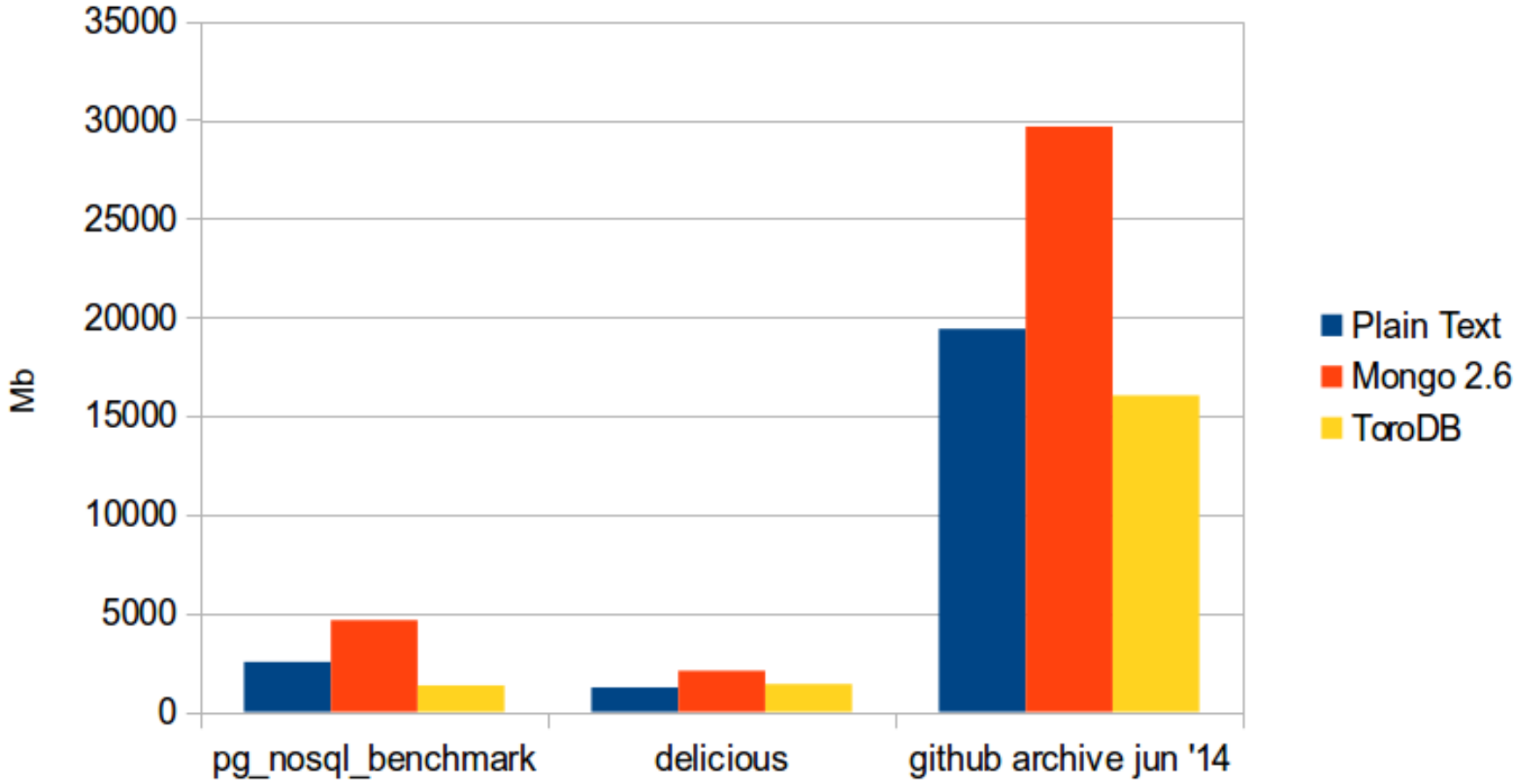
```
select * from demo.structures
```

sid	_structure
0	{"t": 2, "data": {"t": 1}, "nested": {"t": 3, "deeper": {"i": 1, "t": 1}}}

```
select * from demo.root;
```

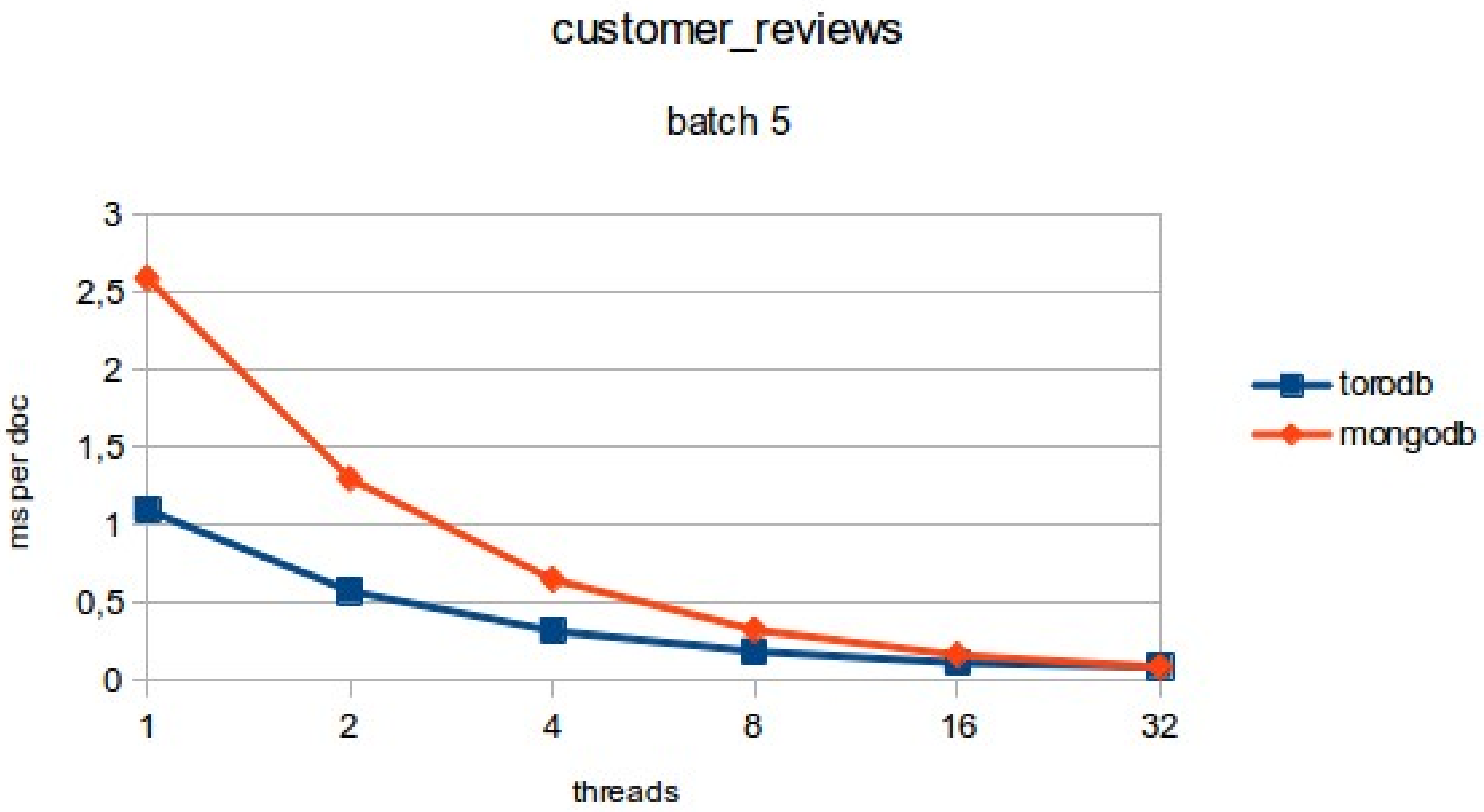
did	sid
0	0

ToroDB storage and I/O savings

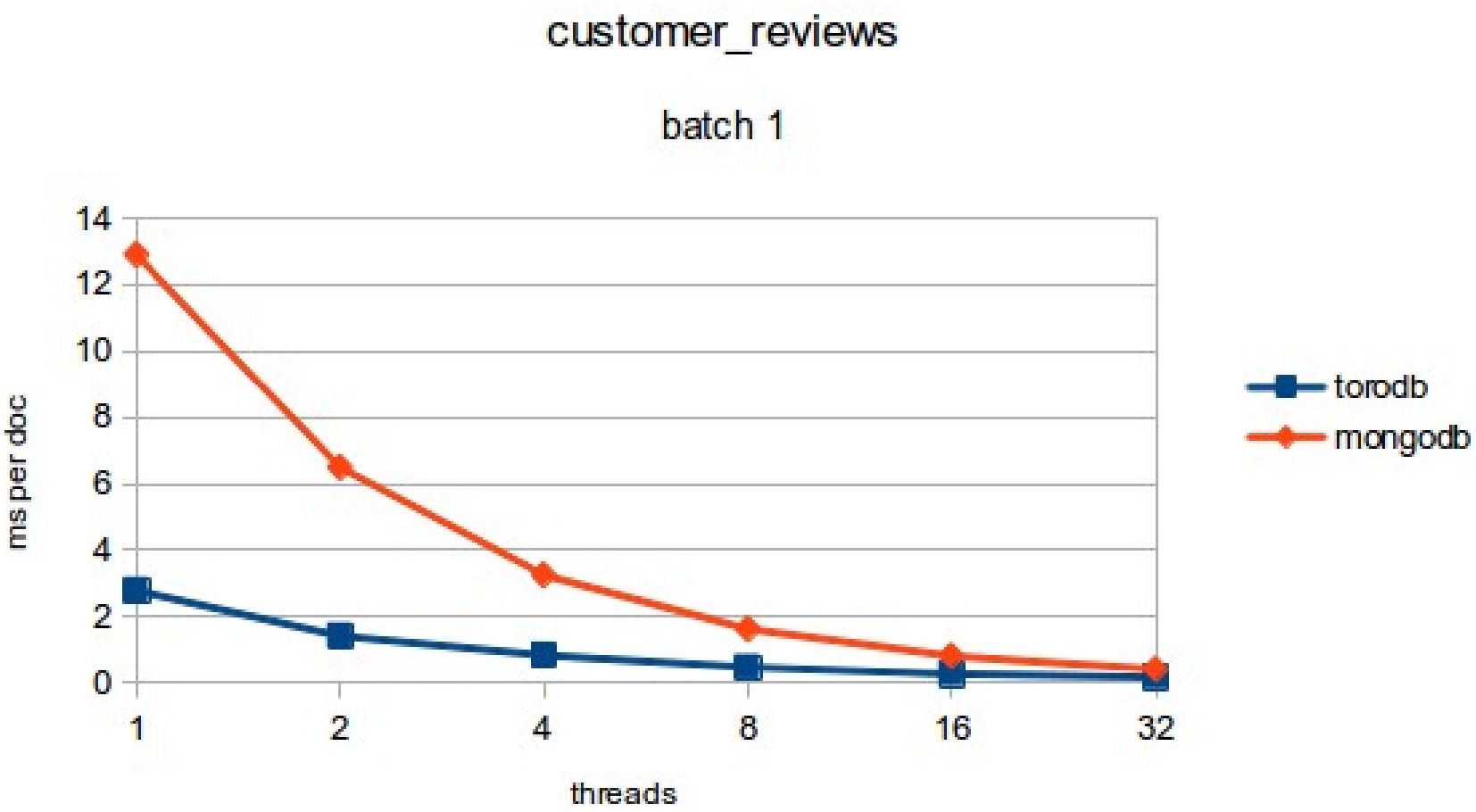


29% - 68% storage required,
compared to Mongo 2.6

ToroDB performance



ToroDB performance (II)



ToroDB: query “by structure”

- ToroDB is effectively **partitioning by type**
- Structures (schemas, partitioning types) are cached in ToroDB memory
- Queries only scan a subset of the data.
- Negative queries are served directly from memory.

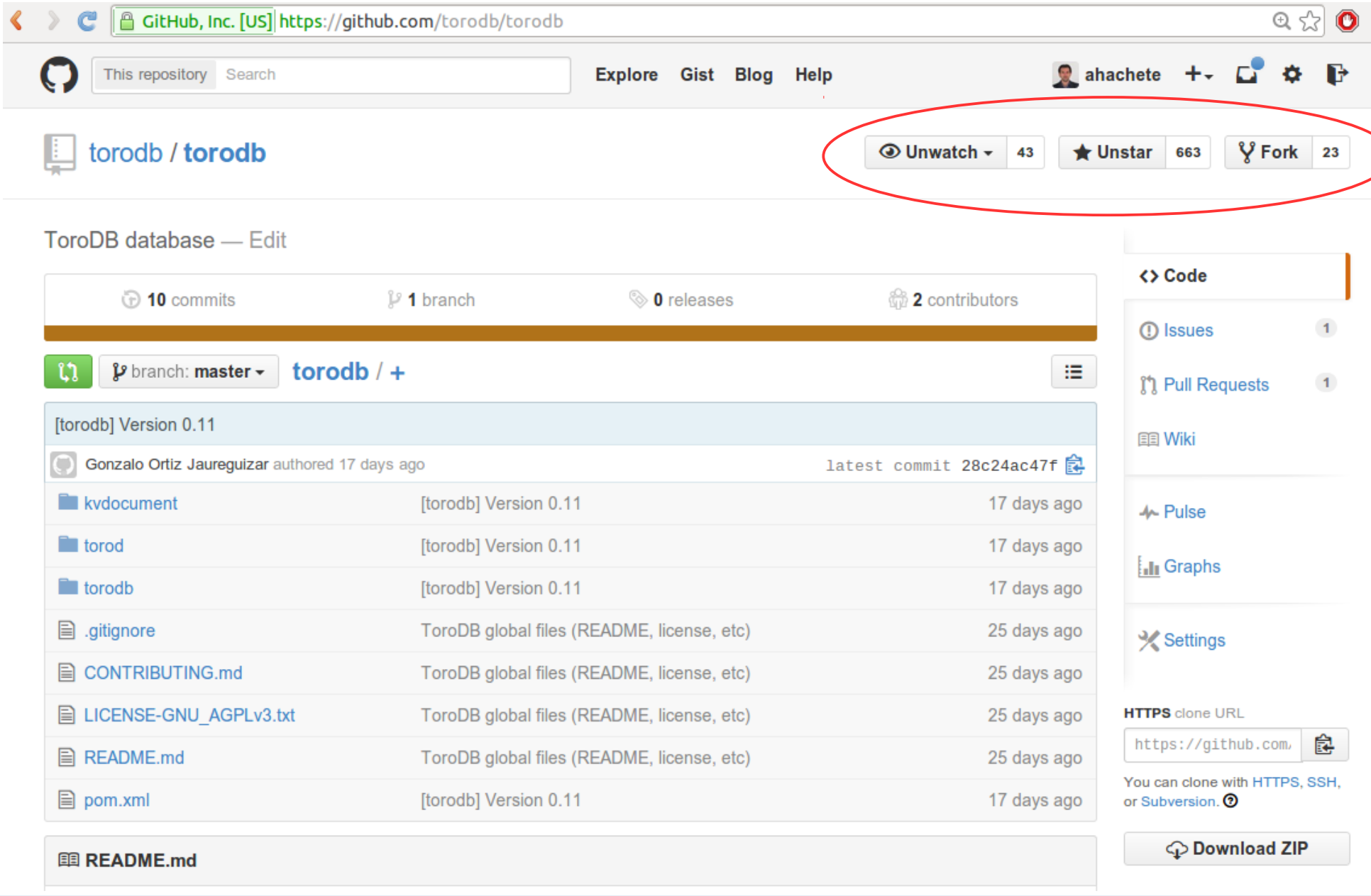
ToroDB: Developer Preview

- ToroDB launched on October 2014, as a Developer Preview. Support for CRUD and most of the SELECT API
- **github.com/torodb**
- RERO policy. Comments, feedback, patches... greatly appreciated
- AGPLv3

ToroDB: Developer Preview

- Clone the repo, build with Maven
- Or download the JAR:
  http://maven.torodb.com/jar/com/torodb/torodb/0.15/torodb.jar
- Usage:
  java -jar torodb-*version*.jar –help
  java -jar torodb/target/torodb-*version*.jar -d dbname -u dbuser -P 27017
  Connect with normal mongo console!

ToroDB: Community Response



ToroDB: Community Response

Jerónimo López
@jerolba

Following

WOW! ToroDb es trending en GitHub! github.com/trending Une lo mejor de MongoDb con lo mejor de las BD de verdad! github.com/torodb/torodb

@nelhage
@nelhage

Follow

As nice as SQL storage engines are these days, projects like ToroDB speak to a real problem: SQL is a fucking terrible API.

Lukas Eder
@lukaseder

Follow

Wow, kick-ass start for ToroDB after Hackernews link! github.com/torodb/torodb

Charity Majors
@mipsytipsy

Follow

more proof that everyone wants mongodb's data interface layer with better way of laying bits on disk. github.com/torodb/torodb via @nelhage

Gunboat Laundromat
@peschkaj

Follow

@cheeseplus Why not ToroDB? github.com/torodb/torodb It's like Mongo, but it's not built on a platform of fail!

Alvaro Hernandez T
@ahachete

ToroDB (github.com/torodb/torodb) is #2 in trending repos today in github!!! github.com/trending Star it!

Neil Brennan, Esq.
@nellophonic

Follow

ToroDB. Speaks natively the MongoDB protocol AND saves your data. github.com/torodb/torodb

Andrew Martin
@sublimino

Follow

Loving the look of github.com/torodb/torodb - speaks Mongo protocol with a Postgres relational backend. Best of both? #nosql

ToroDB: Roadmap

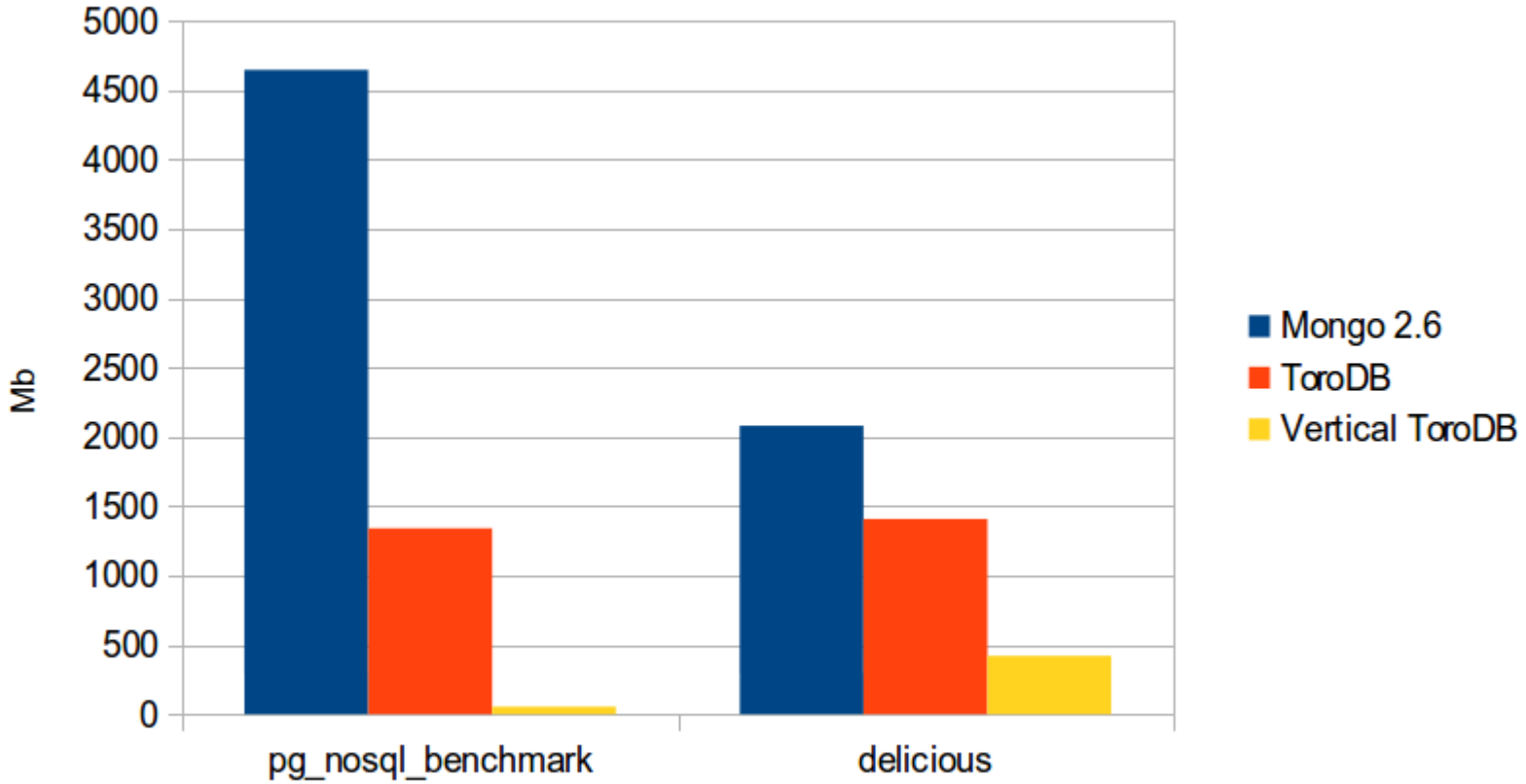
- Current Developer Preview is single-node

- Version 1.0:
  - ➔ Expected Q1 2015
  - ➔ Production-ready
  - ➔ MongoDB Replication support (Paxos-based replication protocol?)
  - ➔ Very high compatibility with Mongo API

Big Data speaking mongo: Vertical ToroDB

What if we use CitusData's cstore to store the JSON documents?

Big Data speaking mongo: Vertical ToroDB



1.17% - 20.26% storage required, compared to Mongo 2.6

"Software acknowledgements"

- PostgreSQL!
- The Netty framework
- jOOQ
- Guava, guice, findbugs
- Hikari CP


8K
data