



Distributed, cross-datacenter SQL Database and Storage at ownCloud

Insight into integration with CockroachDB and IBM Spectrum Scale

Piotr Mrówczyński

*Student at Royal Institute of Technology
and Software Developer at ownCloud*

piotr@owncloud.com

Thomas Müller, Jörn Friedrich Dreyer



Research on benefits of integration with:

- 1. Objectstore for multi-region
deployments**
- 2. Distributed SQL Database**

Globally Distributed Organisation requiring relational SQL database and scaling storage

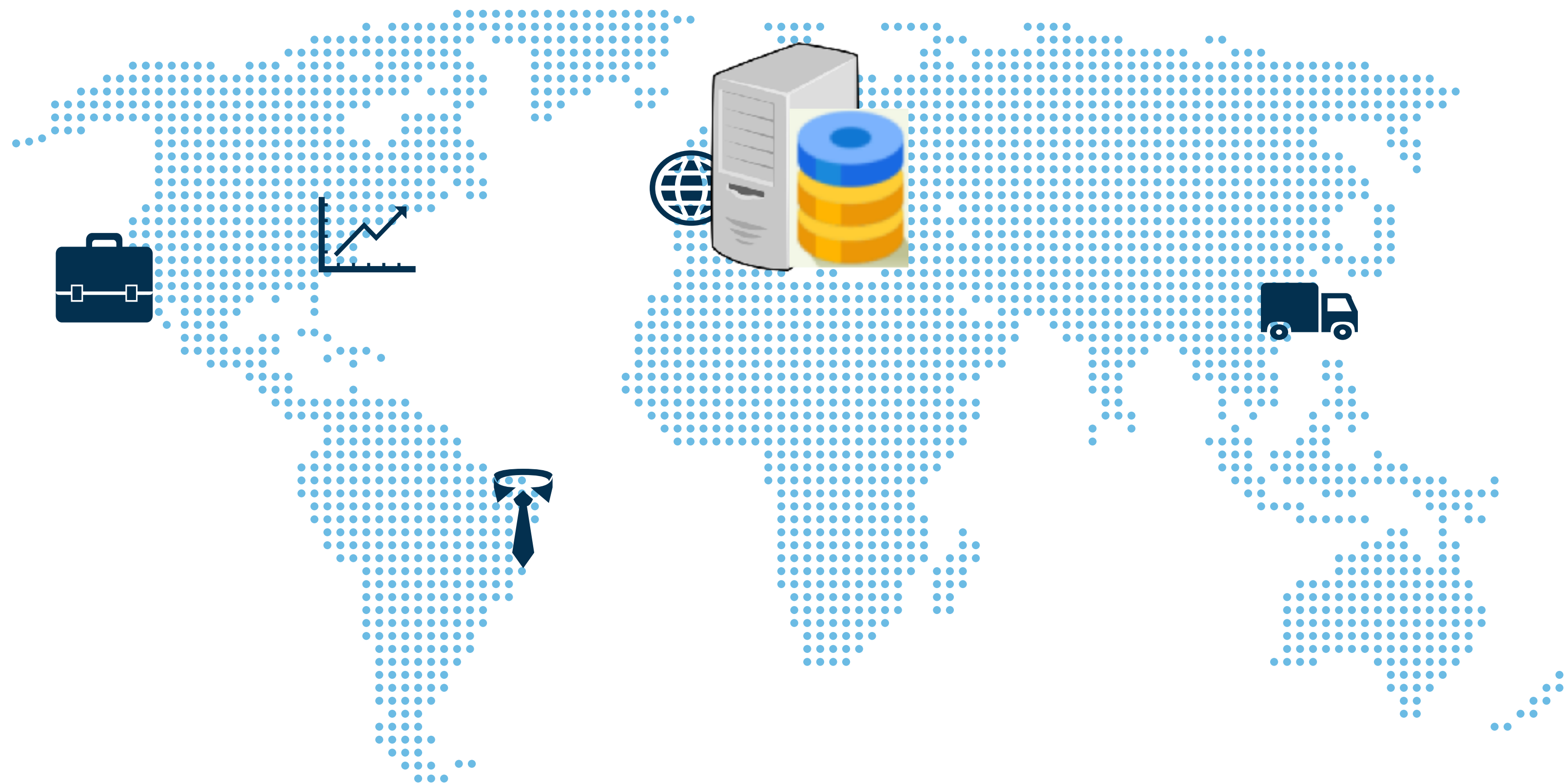



- ✓ Data needs to be **available**
- ✓ Data needs to be **protected**
- ✓ Data needs to be **close to users**
- ✓ Data management needs to be **cost efficient**
- ✓ Data needs to **scale**

3

Data Management Issues

Single datacenter



Let's bring servers and data storage in one location 

Slow, bad user experience!

- ~~Data needs to be available~~
- ~~Data needs to be protected~~
- ~~Data needs to be close to users~~
- Data management needs to be **cost efficient**
- Data needs to **scale**

Data Management Issues

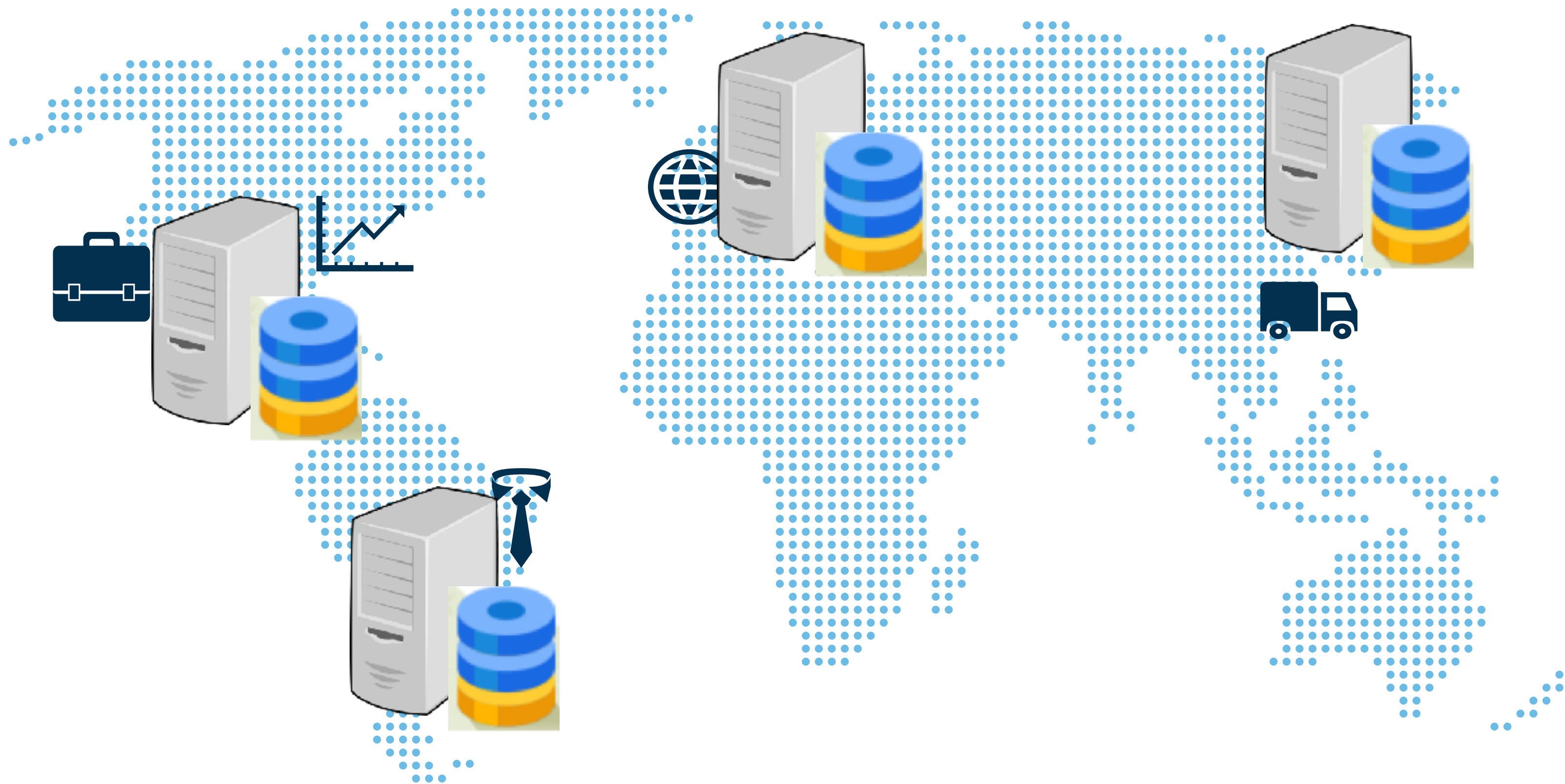
Multi-region SQL/Storage or Application Sharding

Let's divide the data (shard) 

Introduces very complex application logic and maintenance, problems with CAP Theorem, downtimes and more!

Google for AdWords dropped sharded MySQL in 2012..

- ~~Data needs to be available~~
- ~~Data needs to be protected~~
- Data needs to be **close to users**
- Data management needs to be **cost efficient**
- Data needs to **scale**



Data Management Issues

Multi-region SQL/Storage Replication

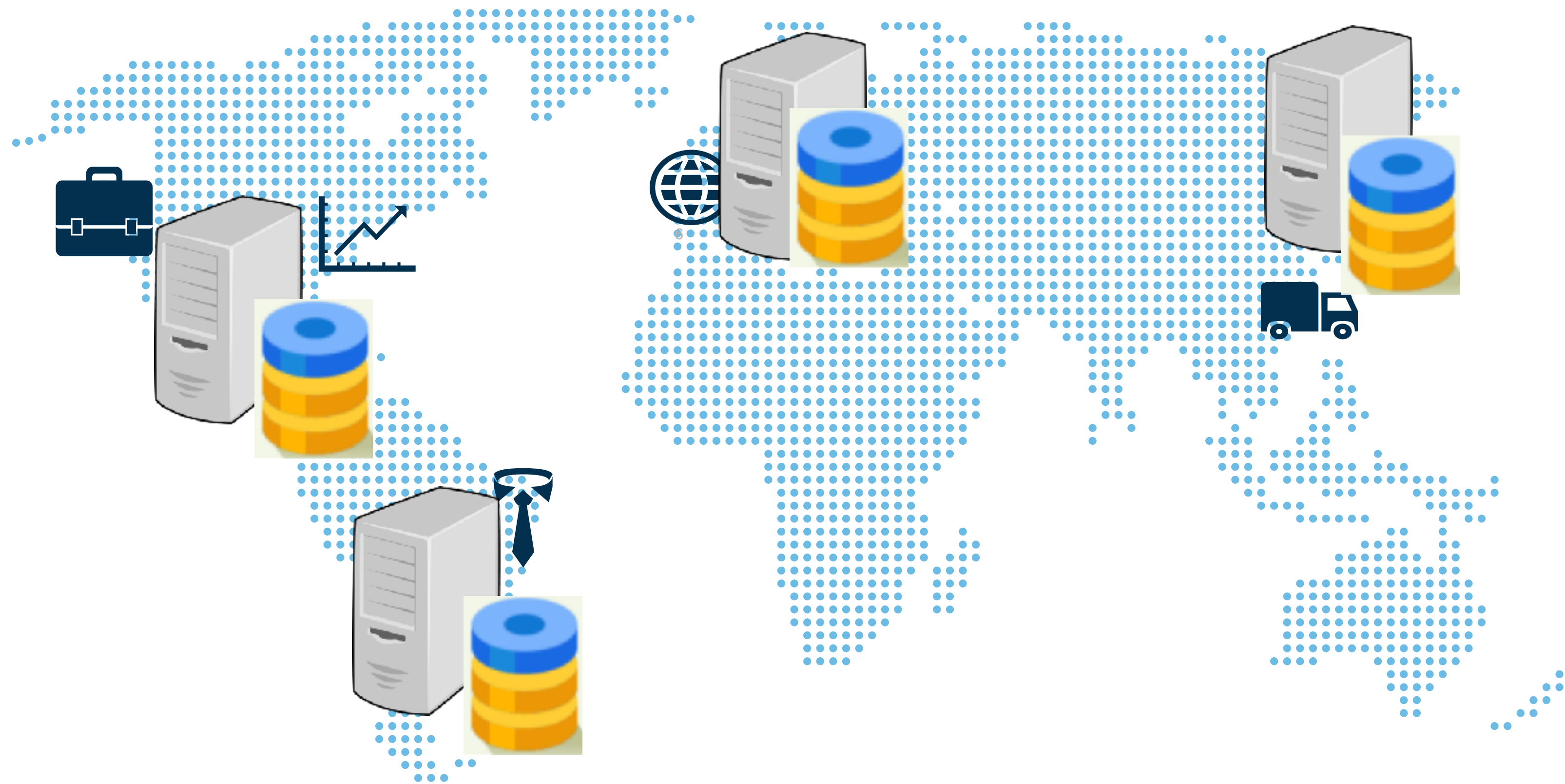
Let's replicate the data (make copies everywhere)

Needs to make application data location aware - it is OK



Costly, insert/mod/delete is slow!

- ✓ Data needs to be **available**
- ✓ Data needs to be **protected**
- ✓ Data needs to be **close to users**
- ✓ ~~Data management needs to be **cost efficient**~~
- ✓ ~~Data needs to **scale**~~



Data Management Issues

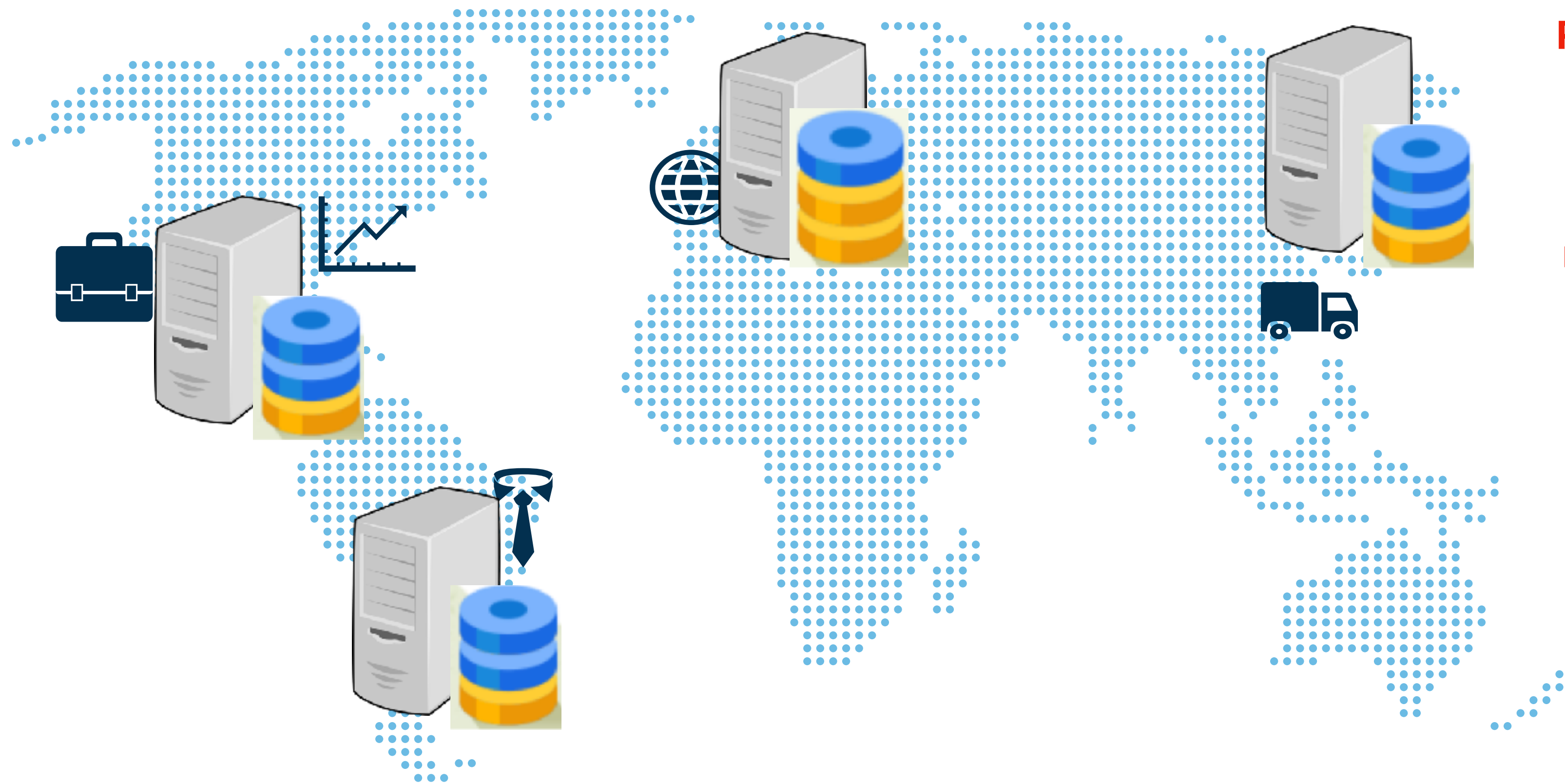
Multi-region NoSQL workaround

Let's use benefits of NoSQL

Problem with data integrity, problems with ACID

Difficult for applications requiring relations - e.g. Google's AdWords, or ownCloud

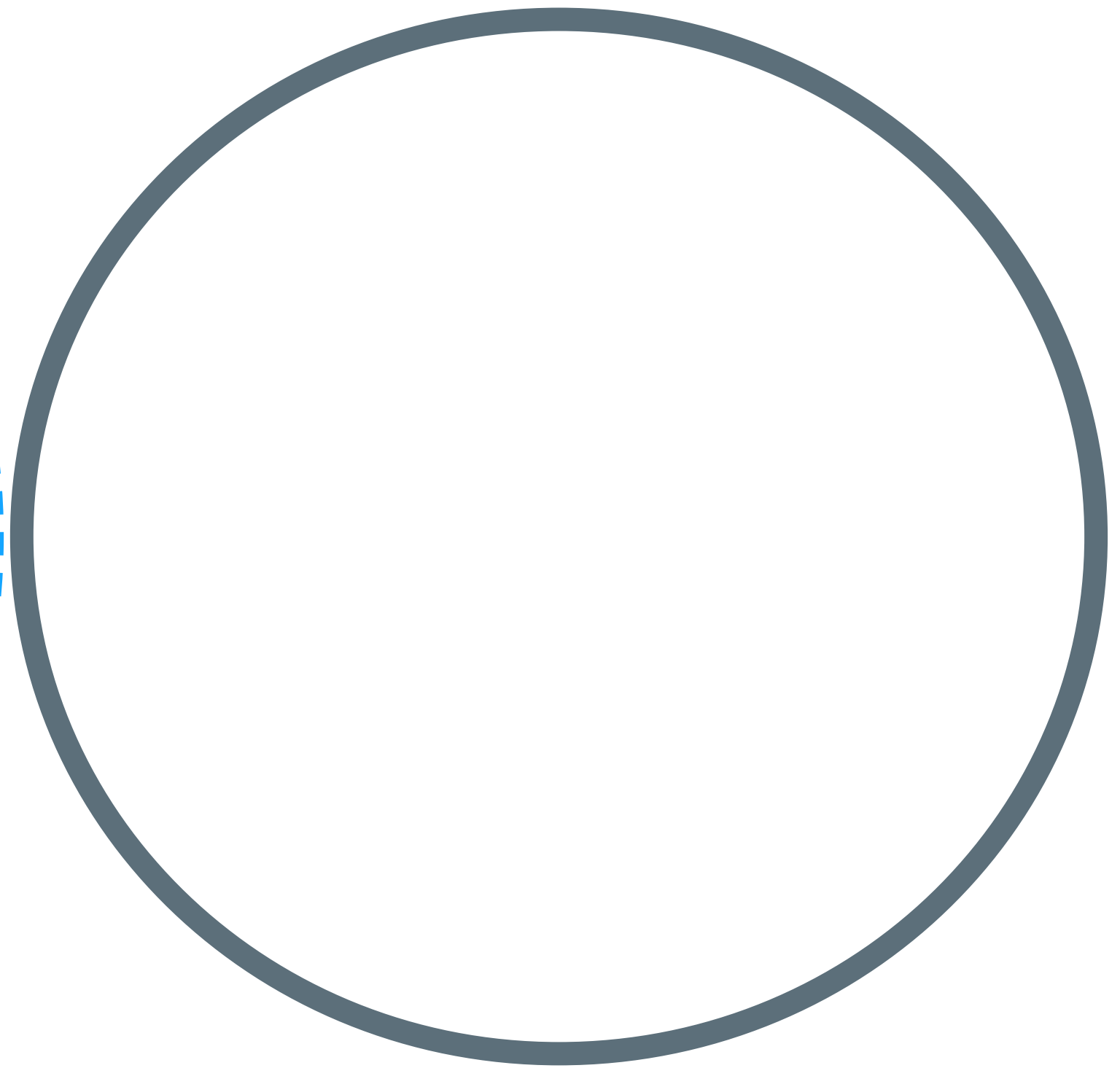
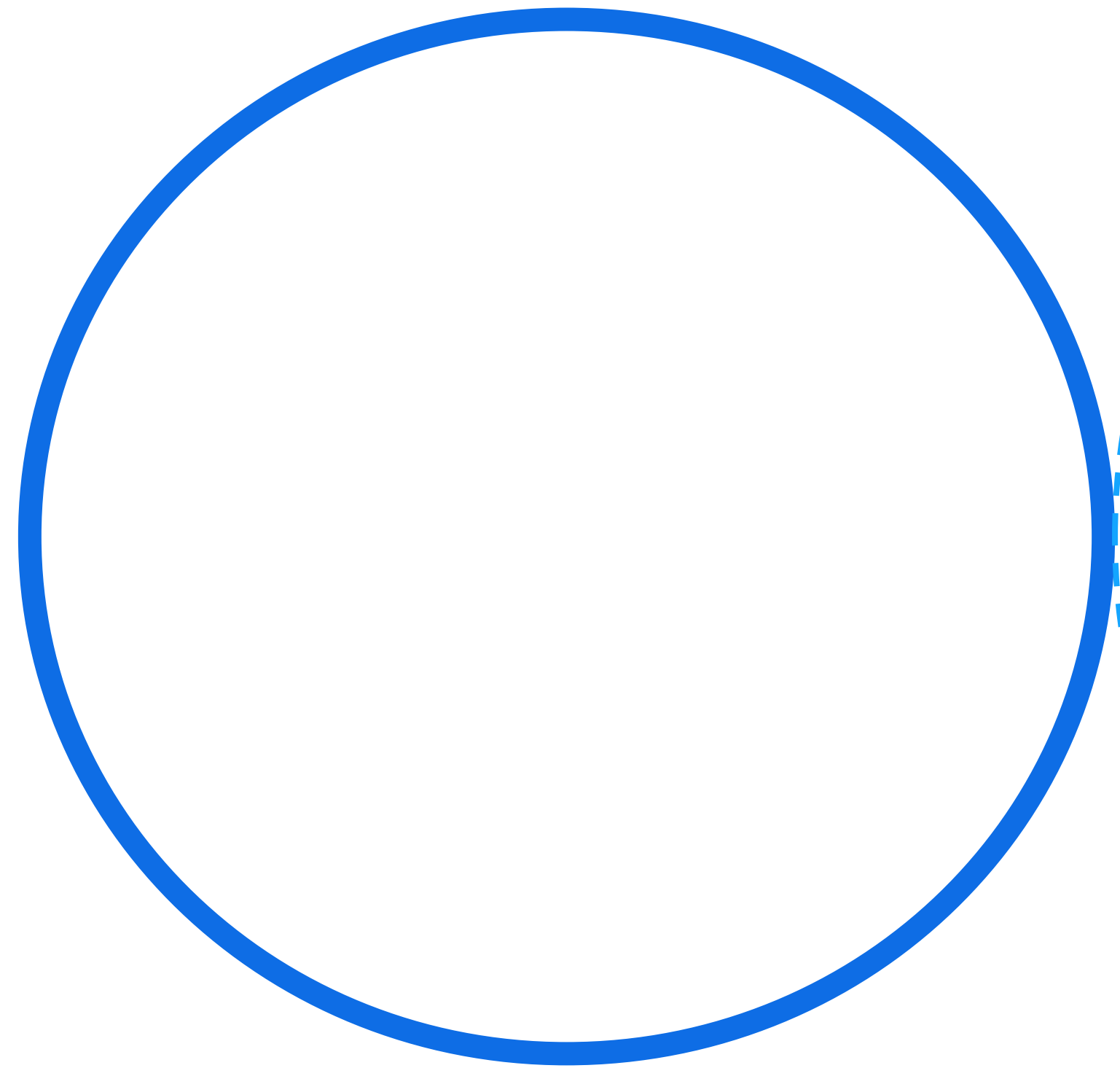
- ✓ Data needs to be **available**
- ✓ Data needs to be **protected**
- ✓ Data needs to be **close to users**
- ✓ Data management needs to be **cost efficient**
- ✓ Data needs to **scale**



“

“How can I scale my ownCloud’s storage system for my global business?”





IBM Spectrum Scale example

✓ Data needs to be **available**

✓ Data needs to be **protected**

✓ Data needs to be **close to users**

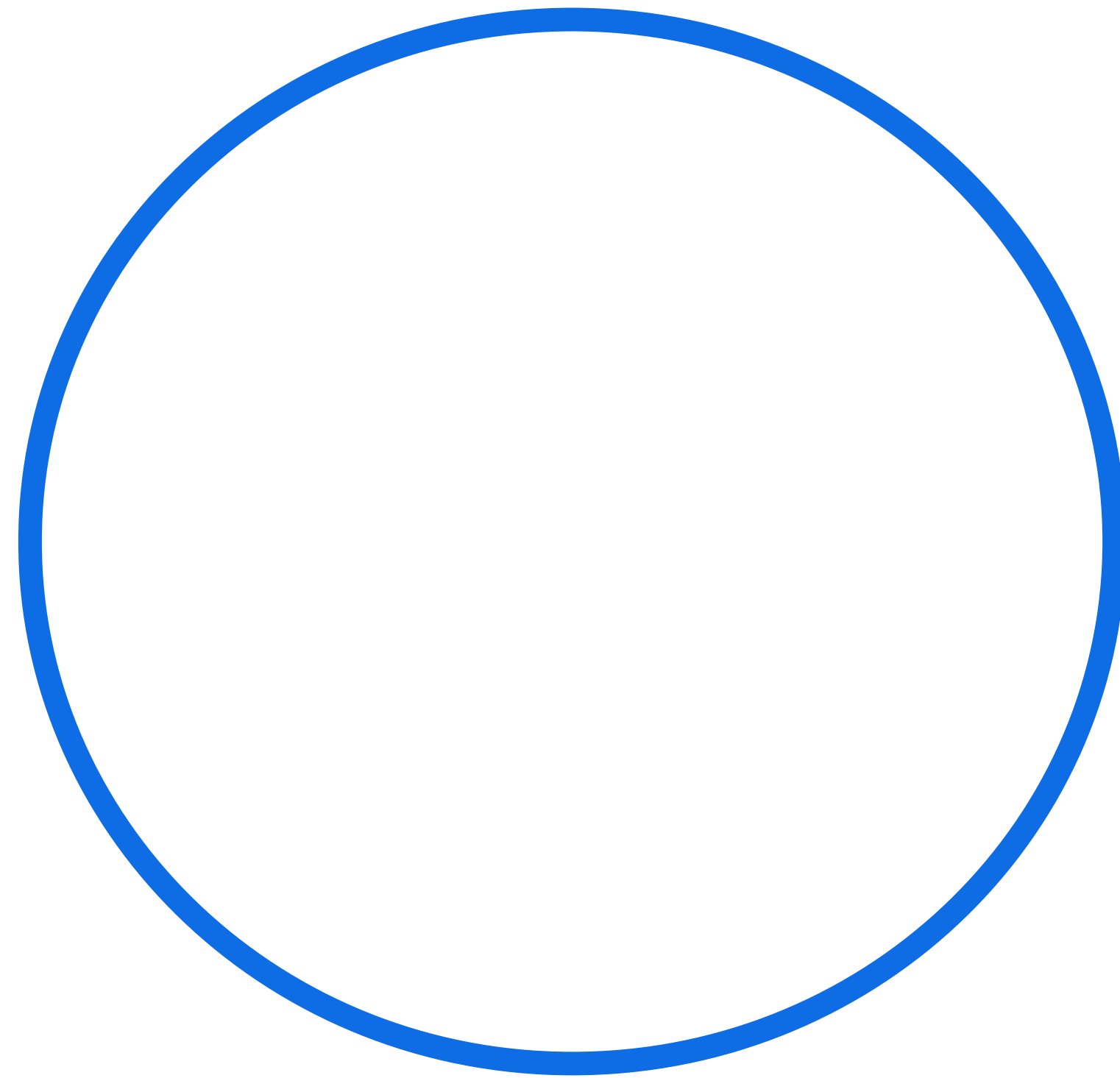
✓ Data management needs to be **cost efficient**

✓ Data needs to **scale**

Multi-site replication and flexible restore capacity. Location aware file-tree in ownCloud

High-performance deduplication, compression and incremental capabilities. Storage options including flash, disk, tape, object stores.

Expands to manage billions of objects



Scalable, multi-site, cost-efficient object store
e.g. IBM Spectrum Scale



Requires new hierarchical file-tree representation in ownCloud database.

File-tree aware of location of files

Better integration
for object storages
in ownCloud

Watch Thomas talk about
object storage at
owncloud

Scalable, multi-site, cost-
efficient object store
e.g. IBM Spectrum Scale

Requires new
hierarchical file-tree
representation in
ownCloud database.

File-tree aware of
location of files



“

“How can I scale my ownCloud’s relational database for my global business? Without sharding or NoSQL workarounds?”



It all started with AdWords at Google...

- [1] F1: A Distributed SQL Database That Scales, Jeff Shute et al., 2013, VLDB'13
- [2] Spanner: Becoming a SQL System, David F. Bacon et al., 2017, SIGMOD'17



... where Google's F1 and Spanner were developed in 2013-2017. Open source CockroachDB released in 2017



**... Distributed SQL
Queries!**

The SQL database for global cloud services.



Data integrity

SQL layer converts SQL statements into a plan of KeyValue operations, which pass to Transaction Layer. Full support for ACID transaction semantics in the Transaction Layer.



SQL API

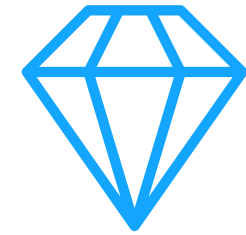
PostgreSQL client drivers

The SQL database for global cloud services.



Data integrity

SQL layer converts SQL statements into a plan of KeyValue operations, which pass to Transaction Layer. Full support for ACID transaction semantics in the Transaction Layer.



Availability and Protection

Always-on services with multi-active availability. Groups of symmetric nodes intelligently agree on write success, once consensus is reached, writes are instantly readable from any node in the cluster.



SQL API

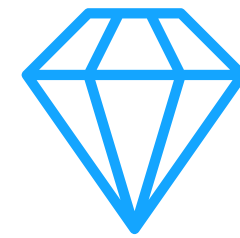
PostgreSQL client drivers

The SQL database for global cloud services.



Data integrity

SQL layer converts SQL statements into a plan of KeyValue operations, which pass to Transaction Layer. Full support for ACID transaction semantics in the Transaction Layer.



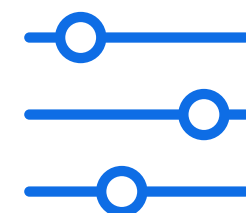
Availability and Protection

Always-on services with multi-active availability. Groups of symmetric nodes intelligently agree on write success, once consensus is reached, writes are instantly readable from any node in the cluster.



SQL API

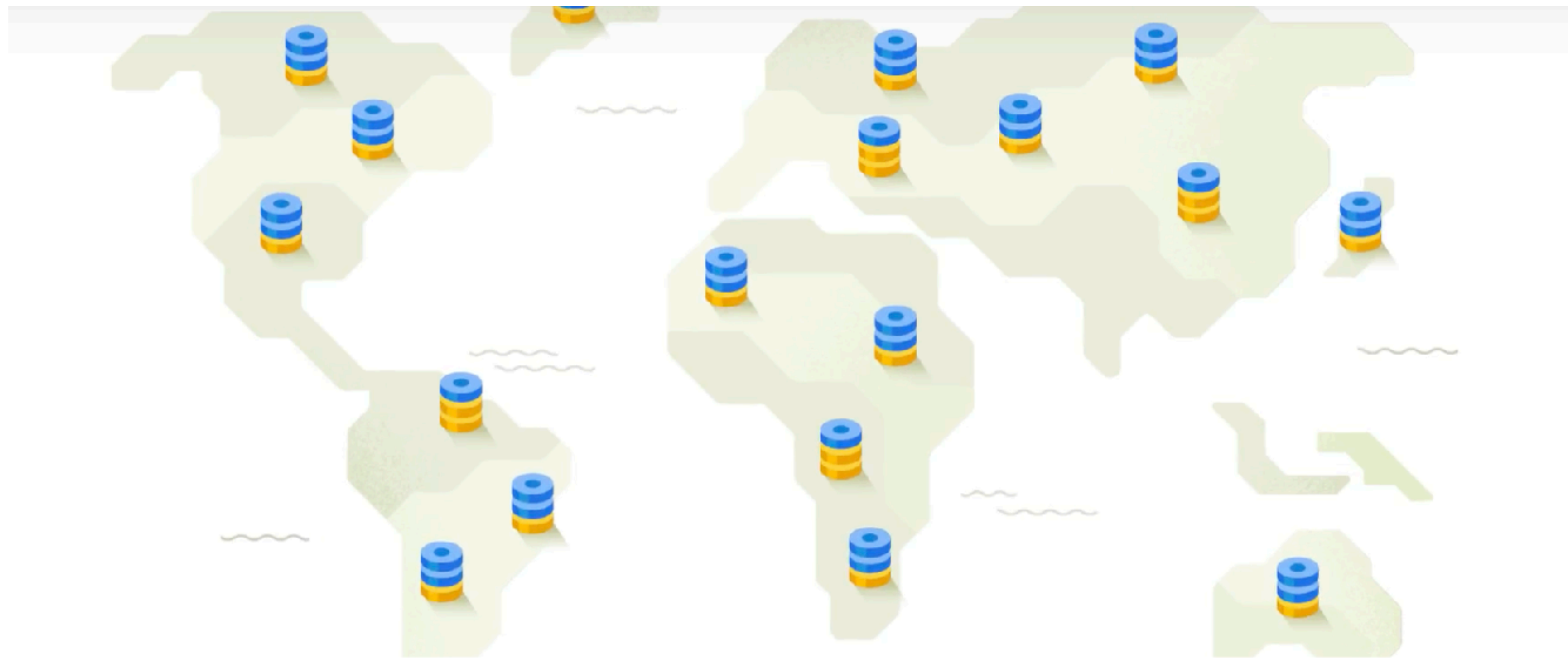
PostgreSQL client drivers



Flexible operations

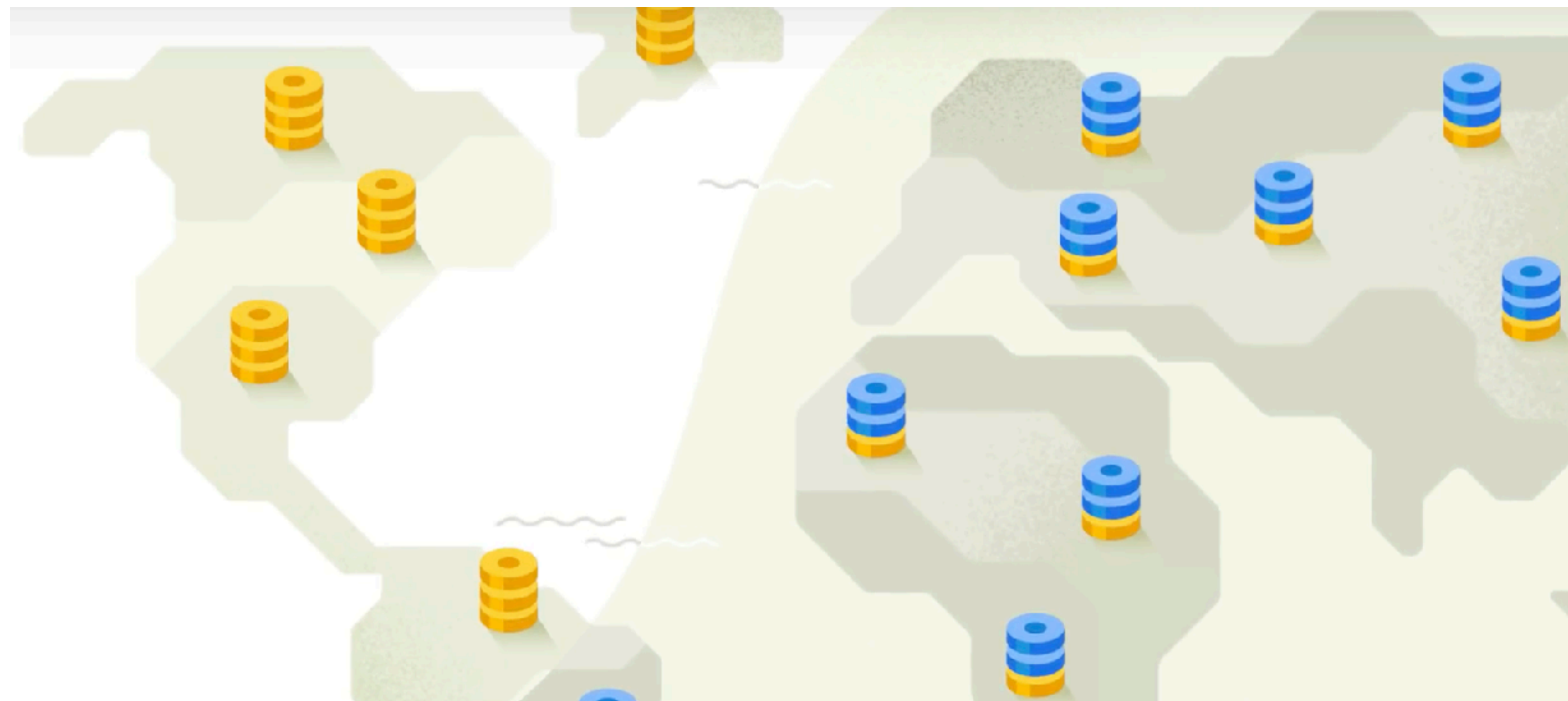
Reduces operational overhead with self-organizing nodes that support built-in scaling, failover, replication, and repair. Control data placement with zone configurations.

The SQL database for global cloud services.



Self-organising nodes

The SQL database for global cloud services.



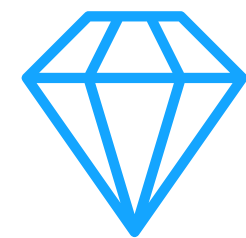
Self-organising nodes

The SQL database for global cloud services.



Data integrity

SQL layer converts SQL statements into a plan of KeyValue operations, which pass to Transaction Layer. Full support for ACID transaction semantics in the Transaction Layer.



Availability and Protection

Always-on services with multi-active availability. Groups of symmetric nodes intelligently agree on write success, once consensus is reached, writes are instantly readable from any node in the cluster.



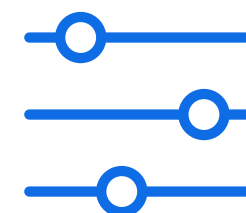
Scalability

Horizontal scalability with symmetric nodes that run on commodity hardware. Orchestrated with Kubernetes.



SQL API

PostgreSQL client drivers



Flexible operations

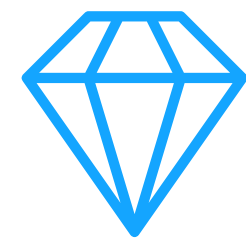
Reduces operational overhead with self-organizing nodes that support built-in scaling, failover, replication, and repair. Control data placement with zone configurations.

The SQL database for global cloud services.



Data integrity

SQL layer converts SQL statements into a plan of Key/Value operations, which pass to Transaction Layer. Full support for ACID transaction semantics in the Transaction Layer.



Availability and Protection

Always-on services with multi-active availability. Groups of symmetric nodes intelligently agree on write success, once consensus is reached, writes are instantly readable from any node in the cluster.



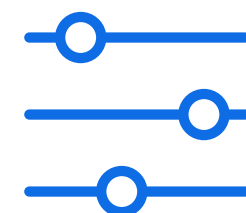
Scalability

Horizontal scalability with symmetric nodes that run on commodity hardware. Orchestrated with Kubernetes.



SQL API

PostgreSQL client drivers



Flexible operations

Reduces operational overhead with self-organizing nodes that support built-in scaling, failover, replication, and repair. Control data placement with zone configurations.



ownCloud

Requires changes in ownCloud core queries, might not work with all SQL constructs that apps use. Research in progress

- ✓ You have global business which requires data to be **available, protected, low-latency, cost-efficient and scale?**
- ✓ You don't want **application and maintenance complexity**, and to **sacrifice availability** and **data protection** using **sharded** architecture?
- ✓ Your first-class citizen is **data integrity** and **SQL/relational data?**

- ✓ You have global business which requires data to be **available, protected, low-latency, cost-efficient and scale**?
- ✓ You don't want **application and maintenance complexity**, and to **sacrifice availability** and **data protection** using **sharded** architecture?
- ✓ Your first-class citizen is **data integrity** and **SQL/relational data**?

Thus **object storage** and **distributed SQL** is something you might be looking at for your **ownCloud** service.



Thank you!

Any questions?