Cloud data store services and NoSQL databases

Ricardo Vilaça

Universidade do Minho

Portugal

Context

- Traditional RDBMS were not designed for massive scale.

- Storage of digital data has reached unprecedented levels.

- Centralized storage and processing making extensive and flexible data partitioning unavoidable.

- Emergence of Cloud Computing paradigm/ business model.

Cloud data store services

- Elasticity. Scale at any moment.

- Reduce the need to have IT specialized staff to implement reliable and scalable data management solutions.

- Reduce the power demands and consumption on data centers.

Cloud data store services

- Dependency on Internet connectivity.

- Dependency on cloud provider.

- Data security (privacy).

- Drawbacks and benefits have to be carefully balanced.

Amazon SimpleDB

- Amazon SimpleDB is one of the Amazon Cloud services, Amazon Web Services.

- Web service that allows to store and run simple queries on structured data.

- Requires no schema, automatically indexes data, and provides real time lookup.

- Highly scalable system optimized for real time and read only applications.

Semana da LEI 2011

Amazon SimpleDB

- Data is organized into domains.

- Domains are collections of items that are described by attribute-value pairs.

- A domain is like a spreadsheet.

- However multiple values can be associated with each attribute.

- Each item can have its own unique set of associated attributes.

Amazon SimpleDB

- CreateDomain

- DeleteDomain

- ListDomains

- PutAttibutes

- GetAttributes

- DeleteAttributes

- Select : =, !=, <, > <=, >=, starts-with, and, or, not, intersection and union.

Semana da LEI 2011

7

Google's App Engine Datastore API

- Google's App Engine is a toolkit that allows developers to build scalable apps. Python or Java.

- App Engine Datastore API provides a query engine and transactional storage.

- Includes a data modeling API and a SQL-like query language.

Google's App Engine Datastore API

- An entity has one or more properties and a unique key.

- The entity properties can be a reference to another entity.

- Every entity is of a particular kind, a group of entities that can be returned by a query.

- Datastore API allows to define data models, and create instances of those models.

- Entities are organized into entity groups, sets of one or more entities that can be manipulated in a single transaction.

Google's App Engine Datastore API

- A query object interface, and a query language called GQL (Python).

- Query language to filter, order, etc. entities in a procedural style: property conditions, **filter()**; ancestor conditions, **ancestor()**; and ordering, **order()**.

SELECT * FROM <kind>

[WHERE <condition> [AND <condition> ...]]

[ORDER BY <property> [ASC | DESC] [, <property> [ASC | DESC] ...]]

[LIMIT [<offset>,]<count>]

[OFFSET <offset>]

<condition> := <property> {< | <= | > | >= | = | != } <value>

<condition> := <property> IN <list>

<condition> := ANCESTOR IS <entity or key>

Semana da LEI 2011

Microsoft

- SQL Azure database

- Windows Azure storage

  - Blob

  - Queue storage

  - Table storage

Large Scale Tuple Stores

- Scalable, elastic and dependable systems.

- Amazon's Dynamo, Yahoo's PNUTS, Google's Bigtable, Facebook's Cassandra

- Solve internal data management problems and support their current or future Cloud services.

- A simple tuple store interface, that allows applications to insert, query, and remove individual elements.

Requirements

- Scale both in the sheer volume of data that can be held but also in how required resources can be provisioned dynamically and incrementally.

- Rely on distributed systems designed from the beginning to be elastic and highly available.

- The CAP theorem states that under network partitions it is impossible to achieve both strong consistency and availability.

Dynamo and PNUTS

- Amazon's Dynamo

  - Properties of both databases and distributed hash tables (DHTs)

  - Building block of some of the Amazon Web Services, such as S3.

  - Several distributed systems concepts in a production system.

- Yahoo's PNUTS

  - Hosted data management service that allows multiple applications to concurrently store and query data.

  - Is a component of the Yahoo!'s Sherpa, an integrated suite of data services.

14

Bigtable and Cassandra

- Google's Bigtable

  - Being used internally at Google for web indexing, Google Earth and Google Finance.

  - Used to store Google's App Engine Datastore entities.

- Facebook's Cassandra

  - Initially developed by Facebook and now an Apache open source project.

  - Uses most of the ideas of the Dynamo architecture to offer a data model based on Bigtable.

DataDroplets

- Offers a simple application interface providing the atomic manipulation of tuples and the flexible establishment of arbitrary relations among tuples.

- Multi-tuple operations leverage disclosed data relations to manipulate sets of comparable or arbitrarily related elements.

- Provides additional consistency guarantees.

Semana da LEI 2011

Dynamo

Key	Value
1	a
3	b
4	c

$$K \rightharpoonup (V \times C)$$

Semana da LEI 2011

PNUTS

Key	Attribute 1	Attribute 2	Attribute 3
1	a		b
3		b	c
4	a	d	

$$K \rightharpoonup \mathcal{P}(String \times V)$$

BigTable

Key	Column1		Column Family1				
			Column2		Column3		
1	a	v1	g	v1	e	v1	
	b	v2	h	v2	f	v2	
	c	v3			g	v3	
3	d	v1			u	v1	
4			e	v1	e	v1	
			f	v2	f	v2	

$$K \rightarrow (String \rightarrow (String \times Long \rightarrow V))$$

Cassandra

Super Column Family1				
	Column Family1		Column Family2	
Key	Column1	Column2	Column3	Column4
1	g	e	x	v
3		u	u	
4	e	e	v	k

$$K \rightarrow (String \rightarrow (String \rightarrow (String \rightarrow V)))$$

DataDroplets

Key	Value	Tags		
1	a	Blue	Oval	
3	b	Red	Rectangular	
4	c	Oval	Red	Rectangular

$$String \to (K \to (V \times \mathcal{P}String))$$

Semana da LEI 2011

Dynamo

- **get**(key)

- **put**(key, context, value)

PNUTS

- **get-any**(tableName,key), **get-critica**l (tableName,key,version), **get-lates**t(tableName,key)

- **put**(tableName,key,value)

- **delete**(tableName,key)

- **test-and-set-put**(tableName,key,value,version)

- **scan**(tableName,selections,projections)

- **rangeScan**(tableName, rangeSelection, projections)

- **multiget**(keys)

Semana da LEI 2011

BigTable

- **put**(key,rowMutation)

- **get**(key,columns)

- **delete**(key,columns)

- **scan**(startKey,stopKey,columns)

- **put**(key,rowMutation)

- **get**(key,columns)

- **range**(startKey,endKey,columns)

- **delete**(key,columns)

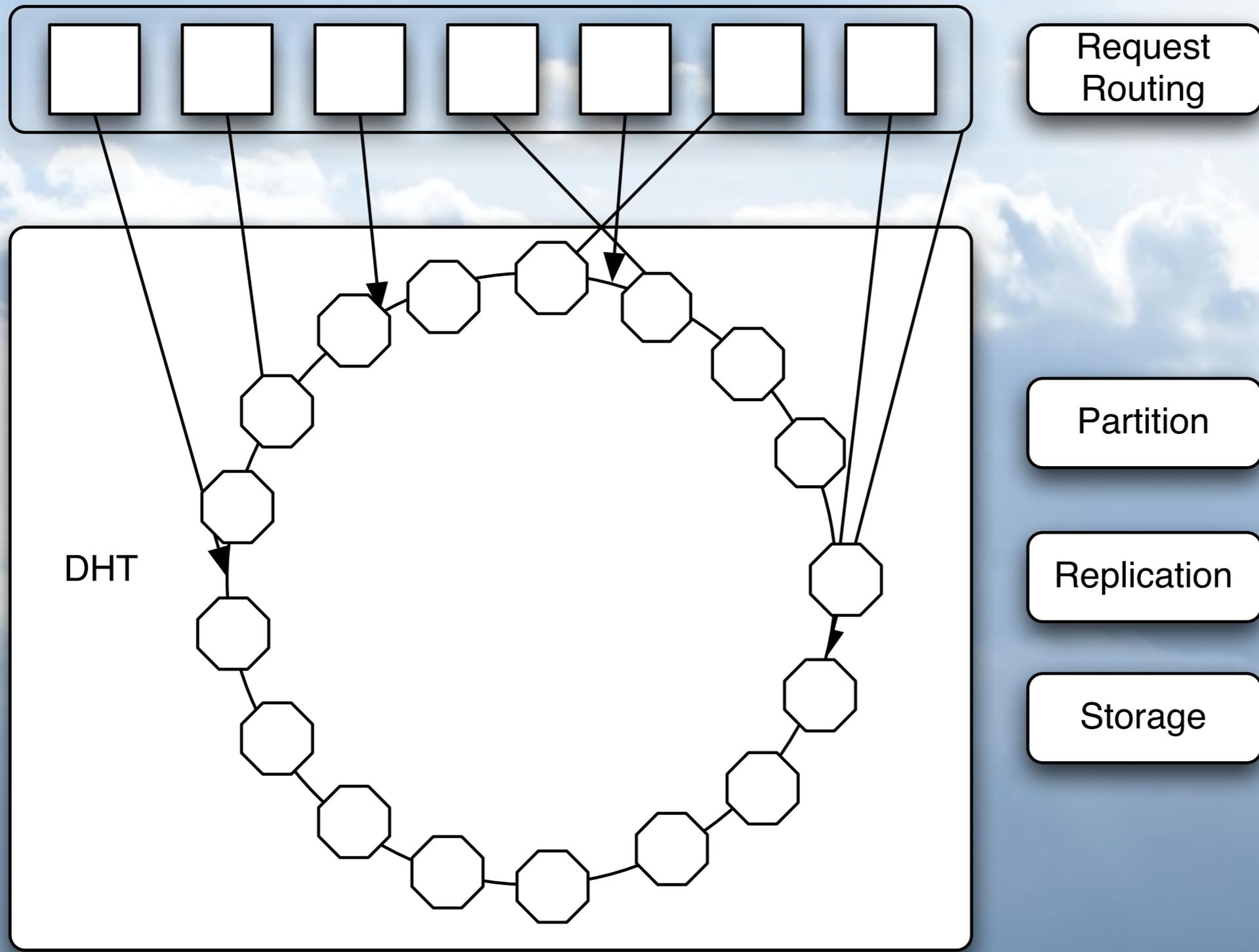# DataDroplets

- **put**(collection,key, value, tags)

- **get**(collection,key)

- **delete**(collection,key)

- **multiPut**(mapKeys)

- **multiGet**(keys)
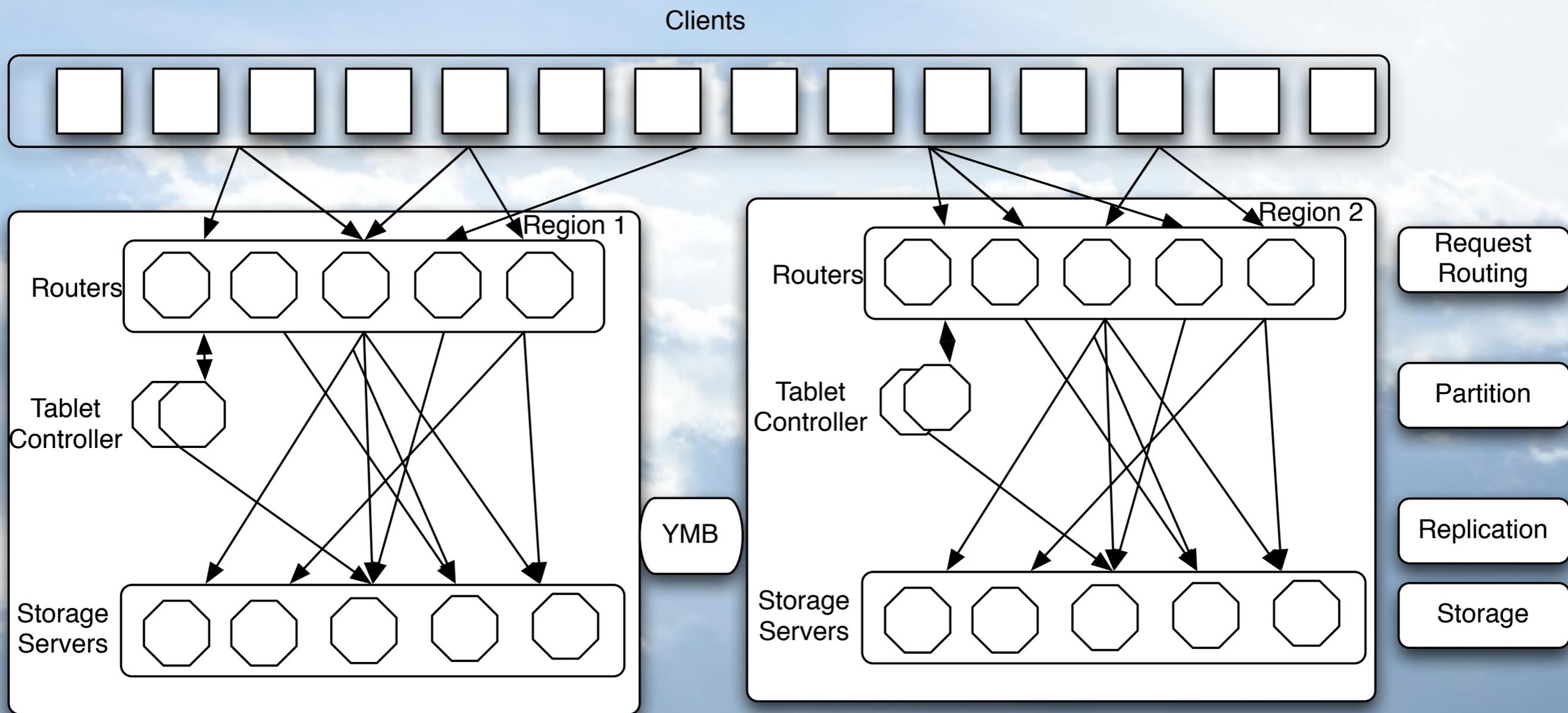
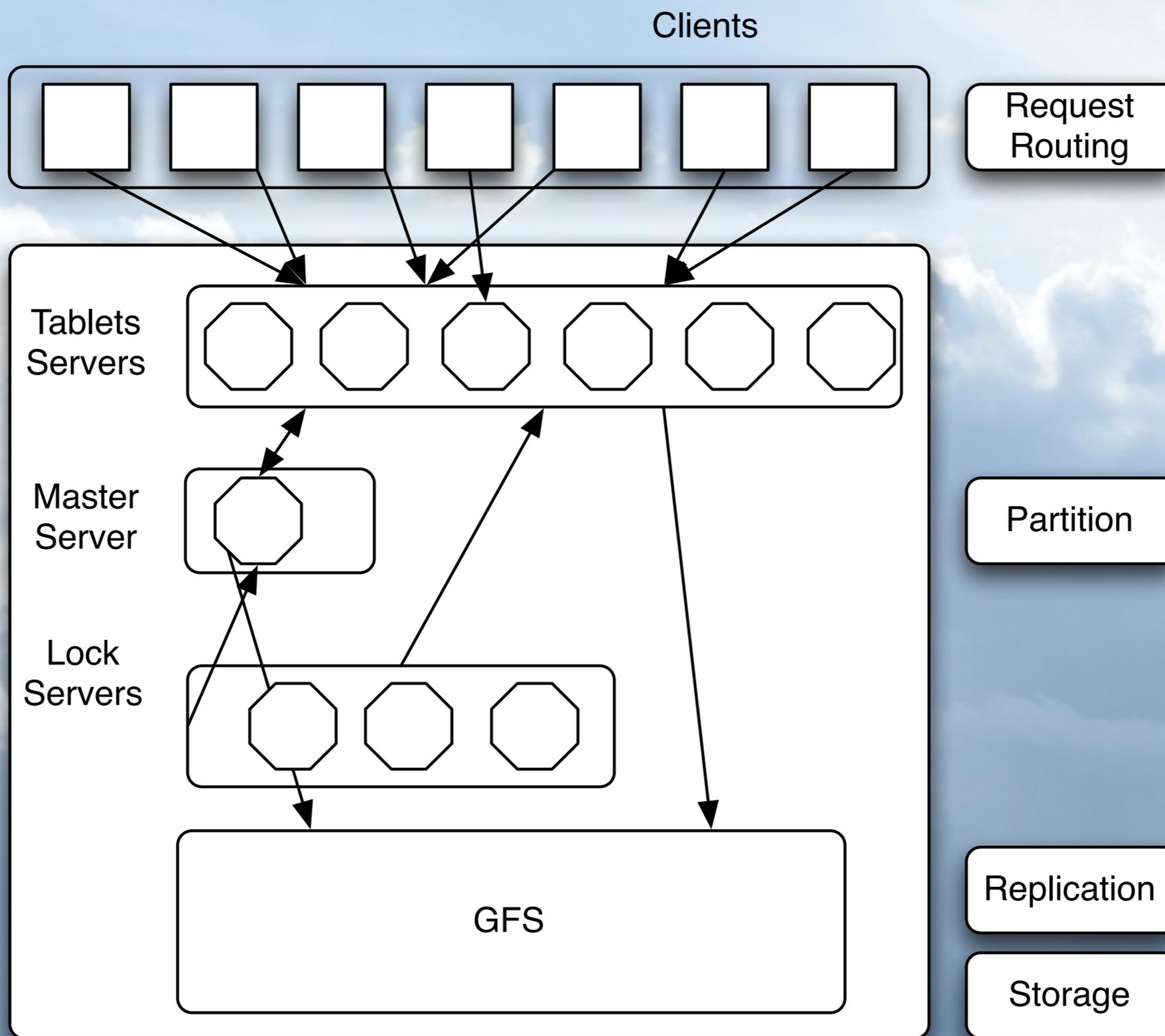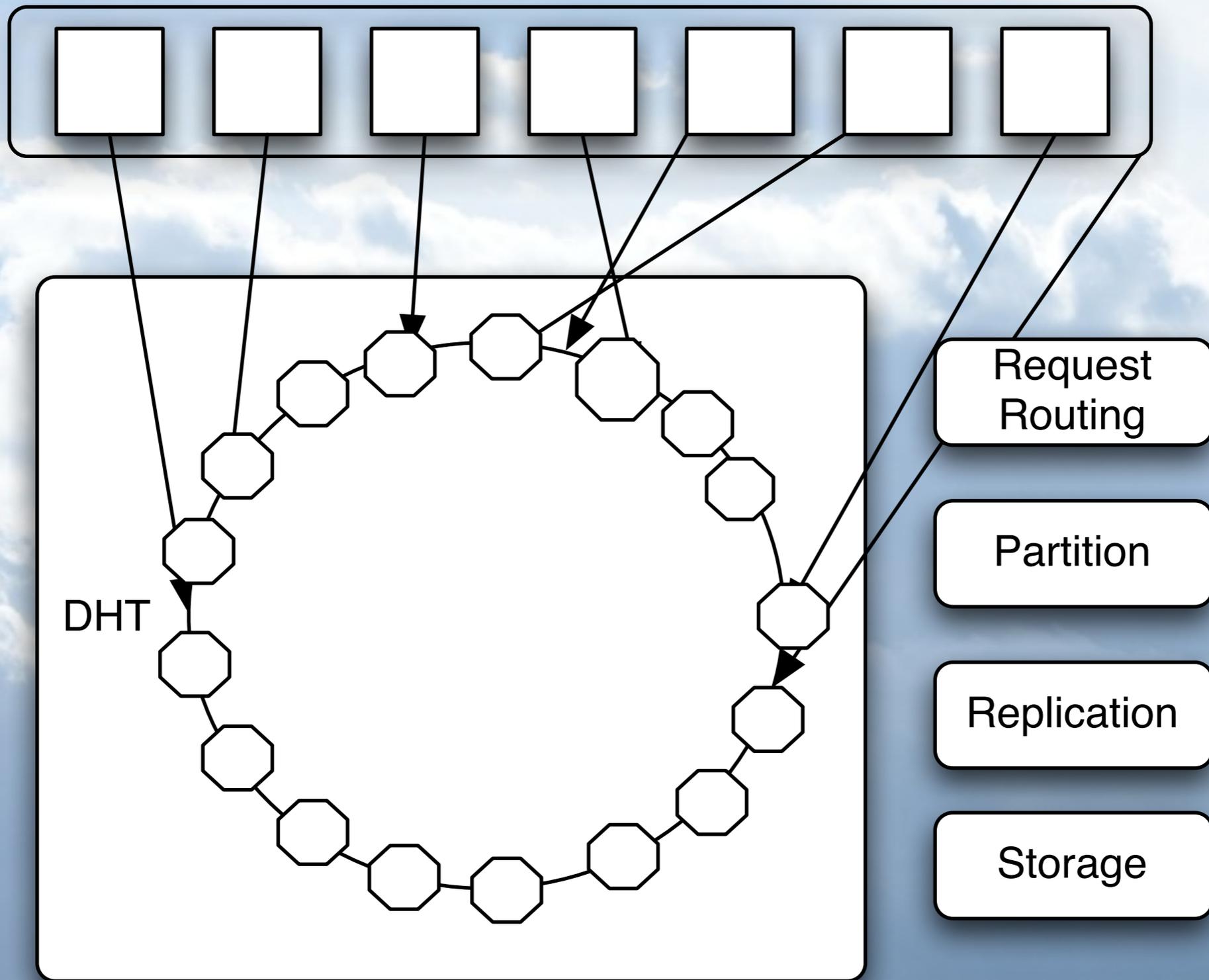- **getByRange**(min, max)

- **getByTags**(tags)

Dynamo

PNUTS

BigTable
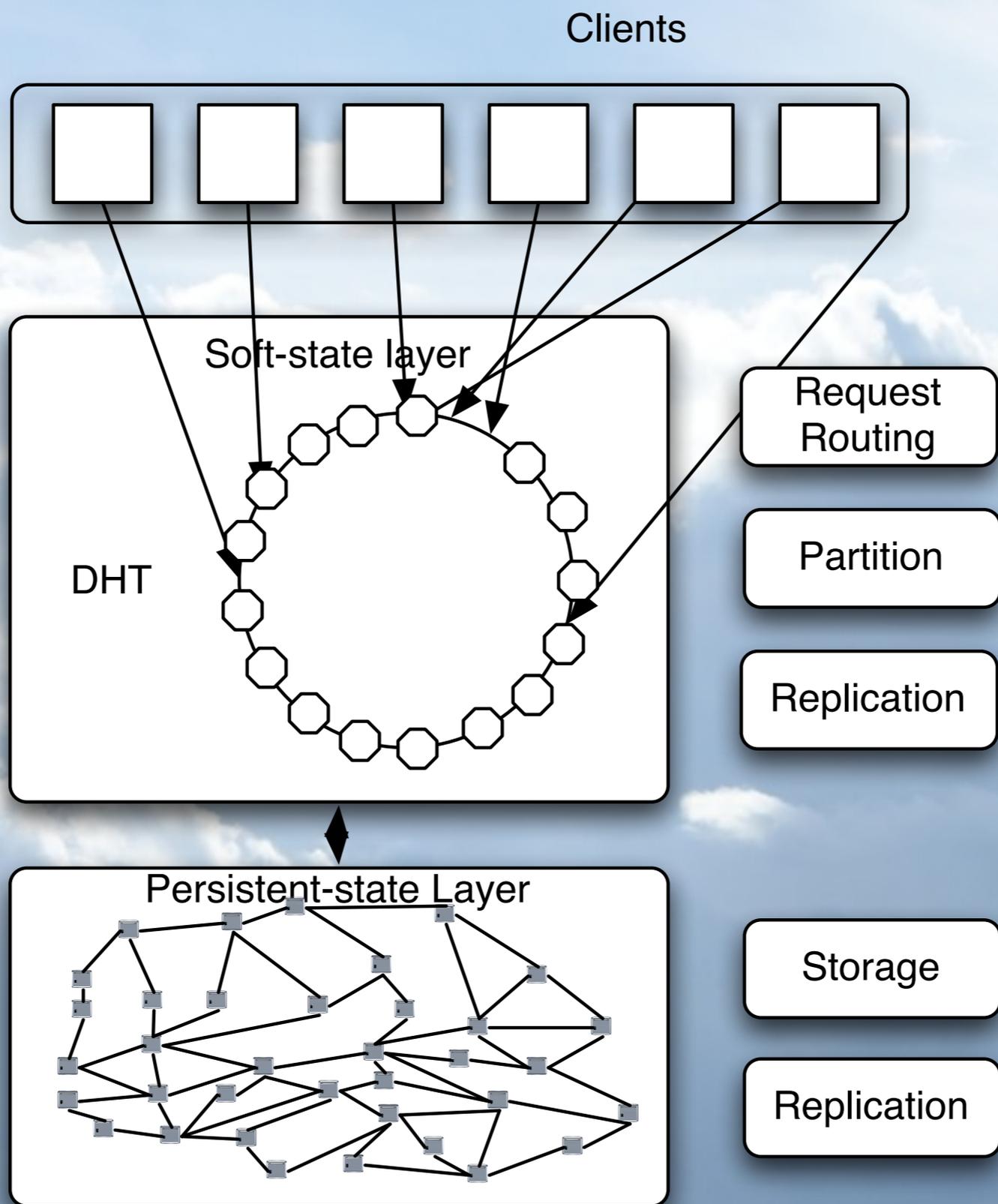

Clients

Request Routing

Tablets Servers

Master Server

Lock Servers

Partition

GFS

Replication

Storage

Cassandra

DataDroplets

Clients

Soft-state layer

DHT

Request Routing

Partition

Replication

Persistent-state Layer

Storage

Replication

Tradeoffs

	Dynamo	PNUTS	BigTable	Cassandra	DataDroplets
Data Partition	random	random and ordered	ordered	random and ordered	random, ordered, and correlation
Consistency	eventual	atomic or stale reads	atomic	eventual	atomic or stale reads
Multiple versions	version	version	timestamp	timestamp	none
Replication	quorum	async message broker	file system	quorum	sync or async

Large scale tuple stores

- Each system offers a different data modeling and a distinct expressiveness on operations.

  - Flexible data models with dynamic attributes.

- Replace the traditional transactional serializability by eventual consistency.

  - Specific trade-offs regarding consistency, availability and migration cost.

Semana da LEI 2011

Large scale tuple stores

- Focus on applications that have minor consistency requirements and can favor availability.

  - Increase complexity at the application logic.

- In most enterprises, it is hard to add this complex layer to the application.

DataDroplets

- Provides additional consistency guarantees and higher level data processing primitives.

- Smooths the migration path for existing applications.

- Fits the access patterns required by most current applications.

Open Issues

- Automatic elasticity.

- Multi-tenancy.

- More complex operations.

- Storage of massive data in commodity hardware.

Ongoing Cloud Research Projects

- Stratus: A Layered Approach to Data Management in the Cloud.

- CumuloNimbo: A highly scalable transactional multi-tier platform as a service

Semana da LEI 2011