



ElasTraS: An Elastic Transactional Data Store in the Cloud

Sudipto Das, Divy Agrawal, Amr El Abbadi

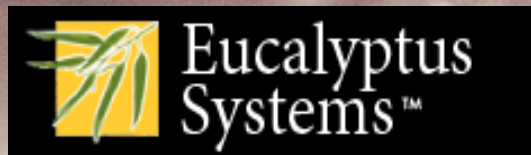
Department of Computer Science

University of California at Santa Barbara

ElasTraS: An Elastic Transactional Data Store in the Cloud

Paradigm Shift in Computing

Azure[™] Services Platform



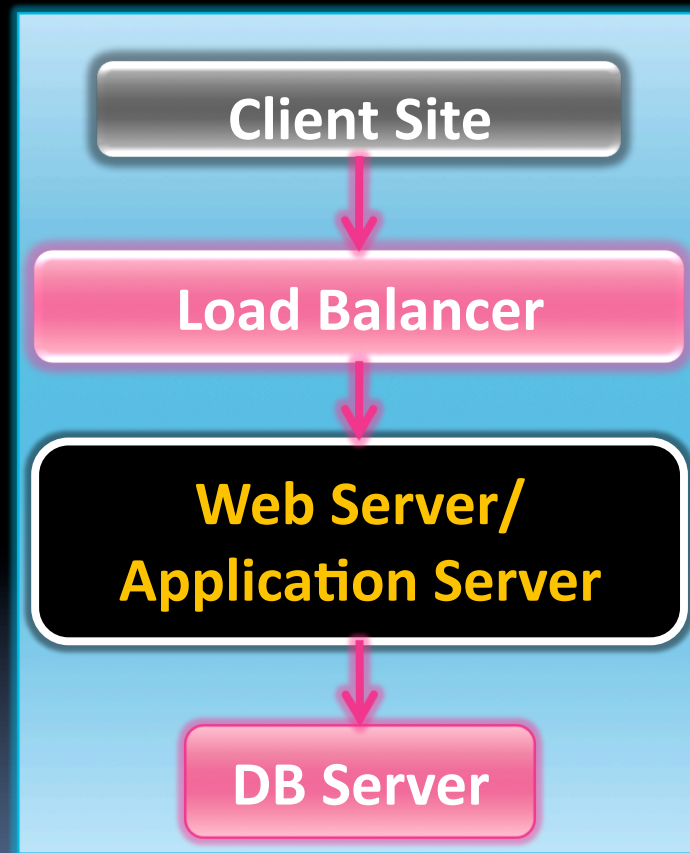
FreeStockPhotos.com



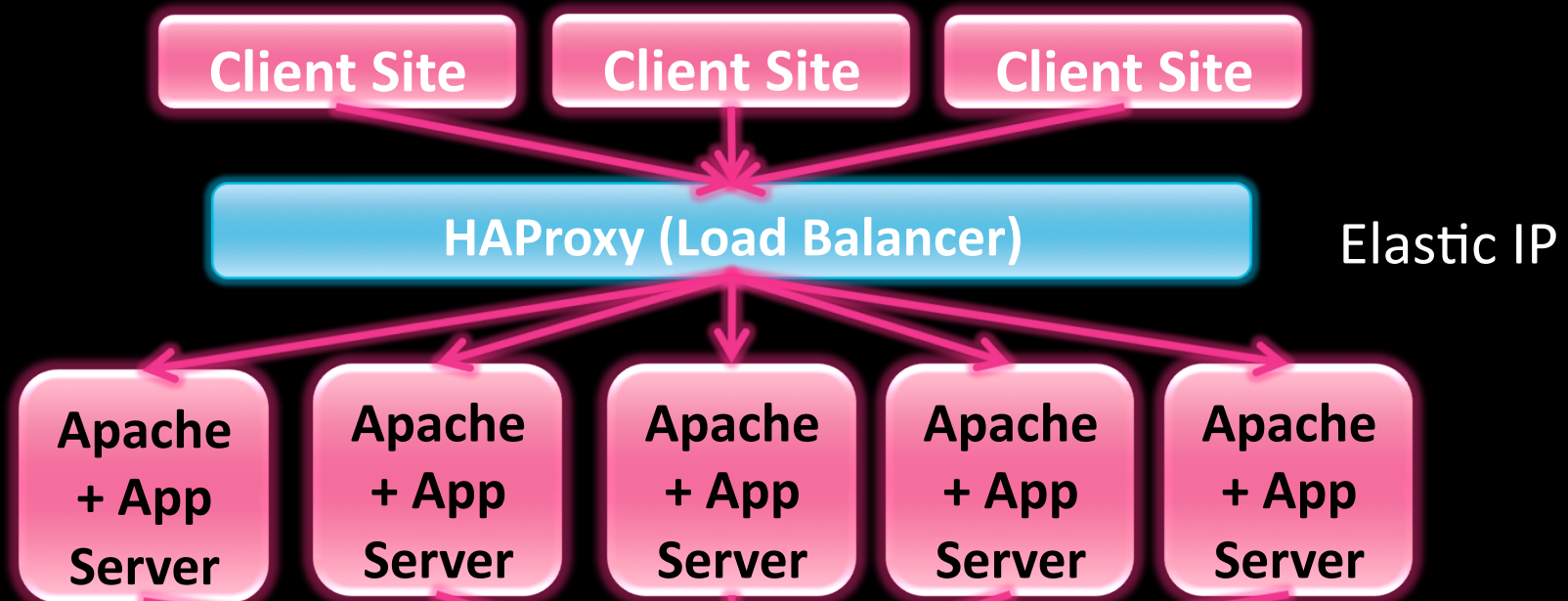
Cloud Computing

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)
- Utility Computing: pay-as-you-go computing
 - No up-front cost
 - **Elasticity** and illusion of infinite resources
 - Transfer of risks
- Scalable and Elastic infrastructure for hosting applications and services

Typical Software Stack

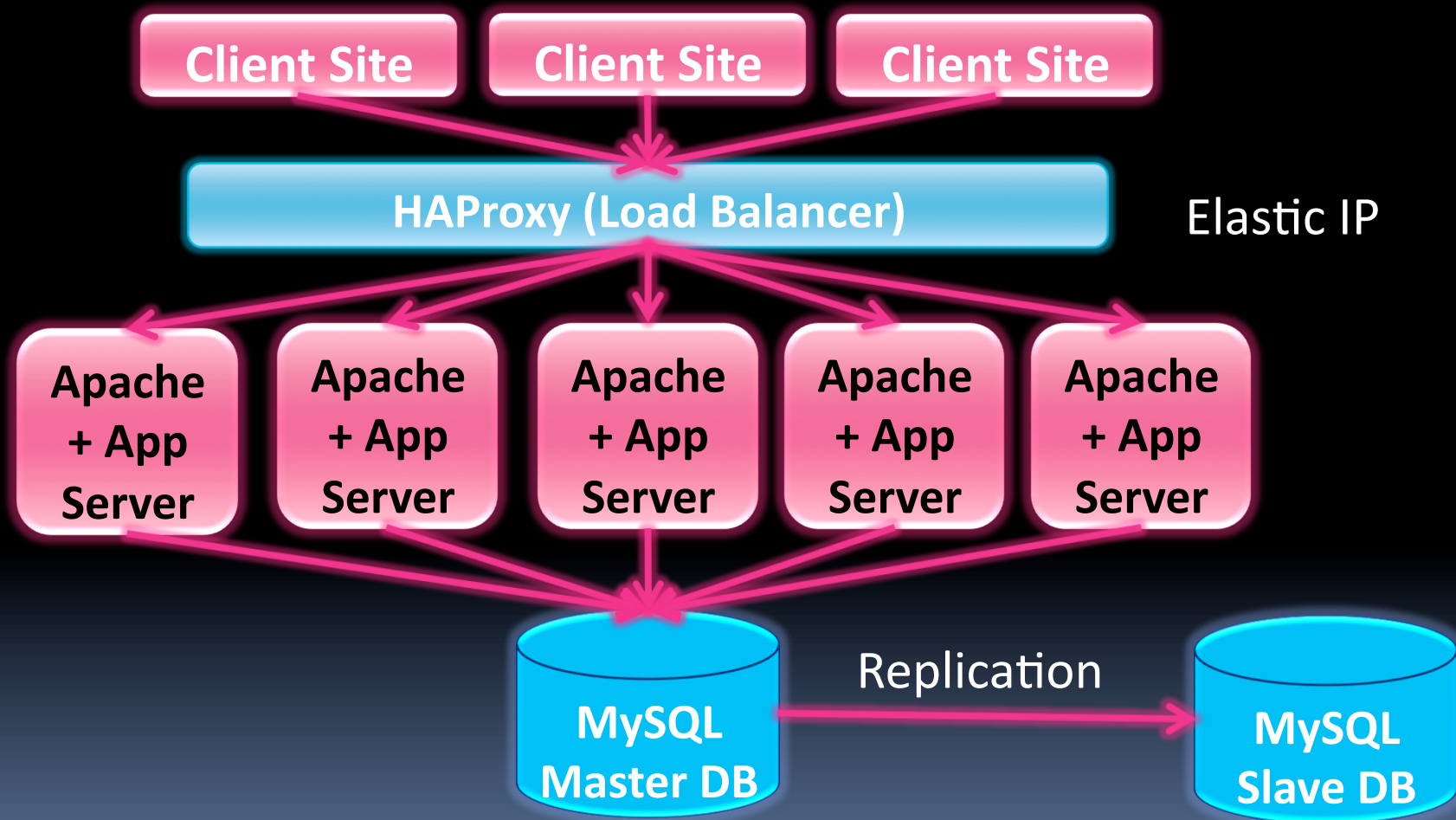


Scaling in the Cloud

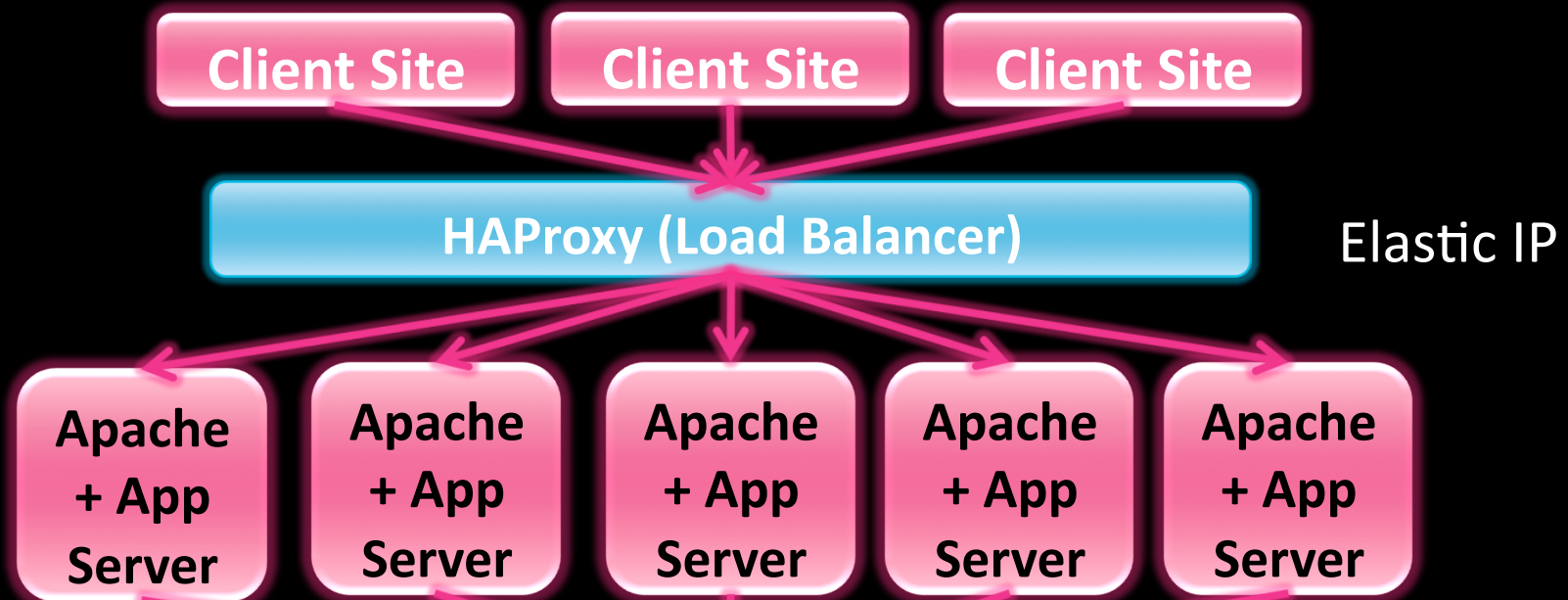


**Database becomes the
Scalability Bottleneck
Cannot leverage elasticity**

Scaling in the Cloud

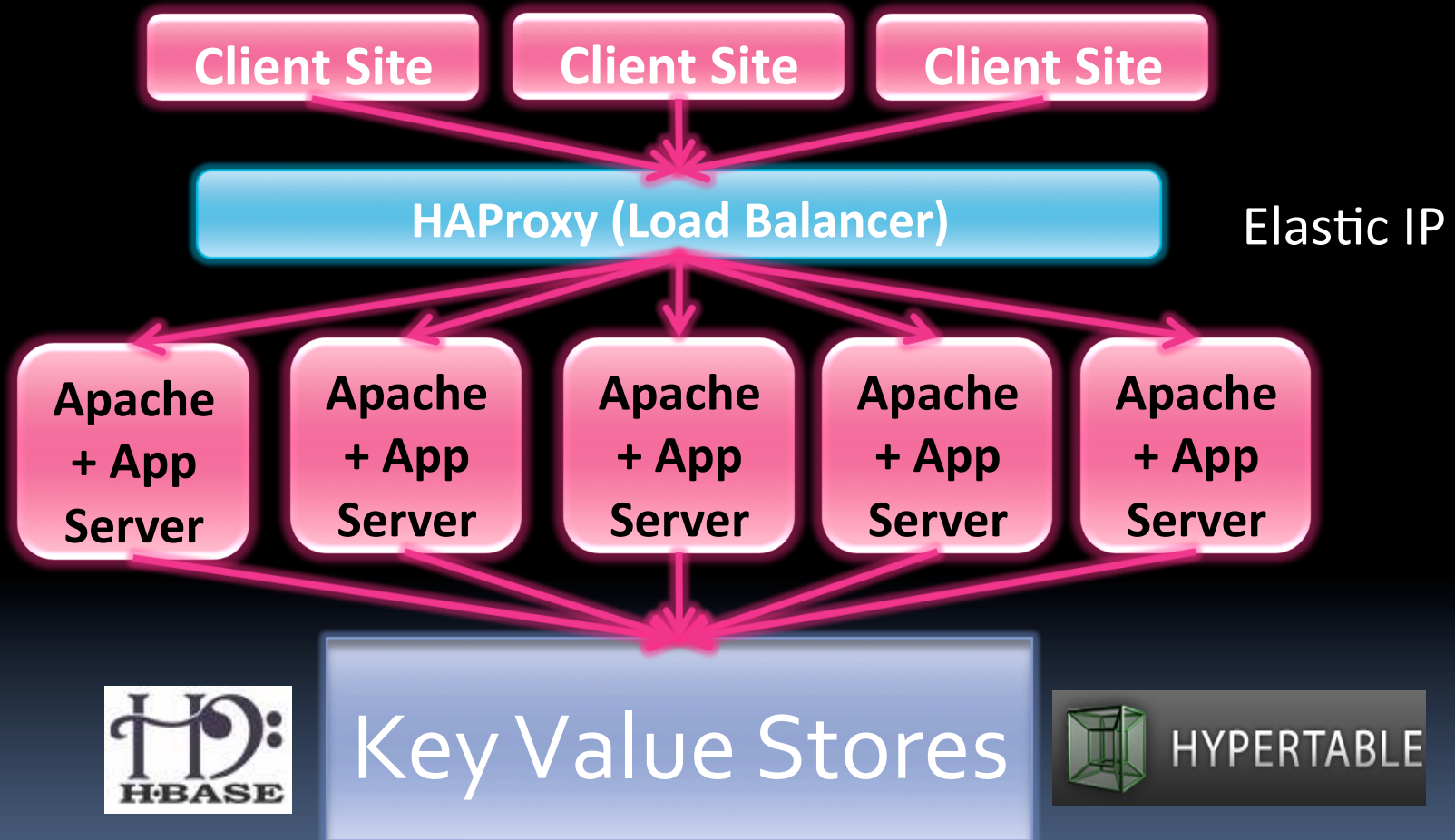


Scaling in the Cloud

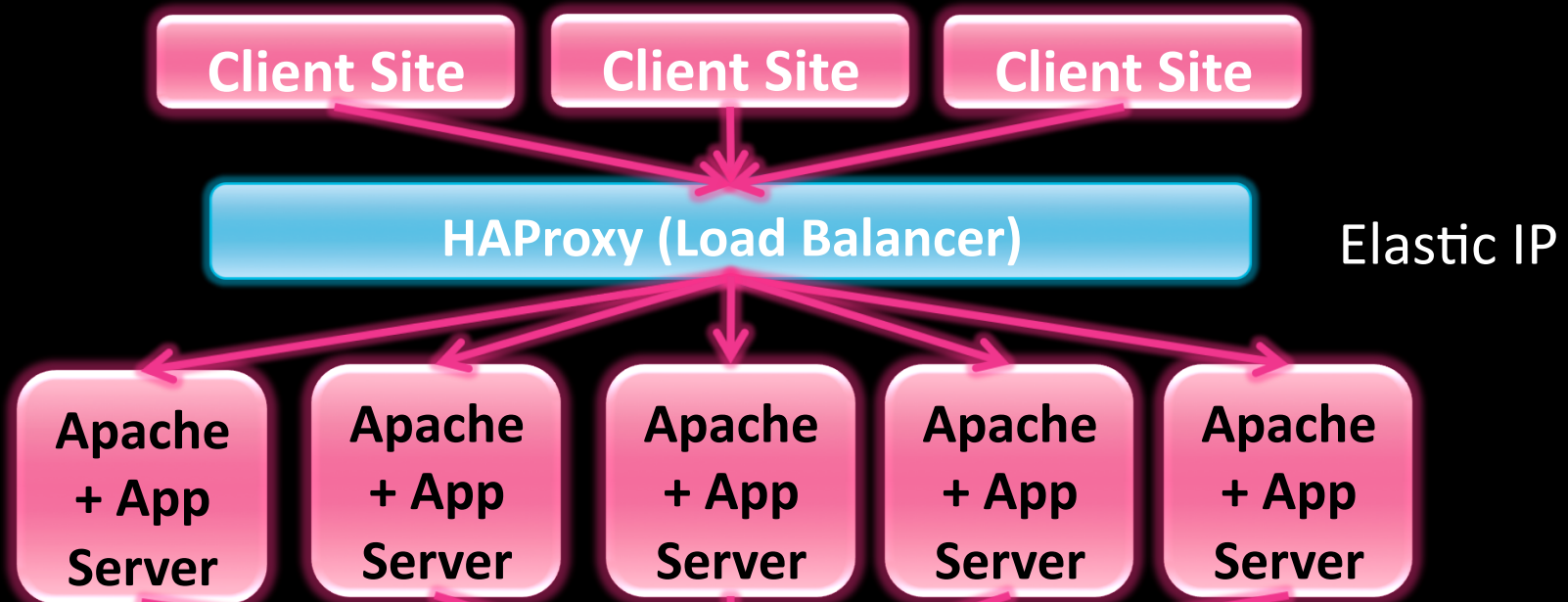


**Scalable and Elastic
But minimal consistency and
operational flexibility**

Scaling in the Cloud



Scaling in the Cloud



ElasTraS

- Is it yet another Distributed Database?
-
- It's a Lightweight Distributed Data Store
- Scalable design principles borrowed from Bigtable
- Transaction Management & Failure Recovery borrowed from Databases
- **An Elastic and Scalable Transactional Store**

- Exclusive access rights to database partitions
- Executes client transactions for the partitions it owns
- Aggressive caching of partitions locally
- No interaction across OTMs
- Interacts with Metadata Manager for obtaining leases

0

OTM



Distributed Storage (S3)

Raw
Storage

- Heart and Brain of the system
- Distributed lease management for the partitions
- Maintaining persistent mapping of partitions to owning transaction managers
- Monitoring the health of the system
- Synchronous replication of metadata for fault-tolerance

On



...



Distributed Storage (S3)

Raw
Storage

6/30/09

- Absorb read only query workload
 - Cache partial database contents
 - Answer queries from the database cache
- Read-only access to the entire database
- No state associated with an HTM
- Acts as coordinator for execution of *minitransactions*

or

HTM



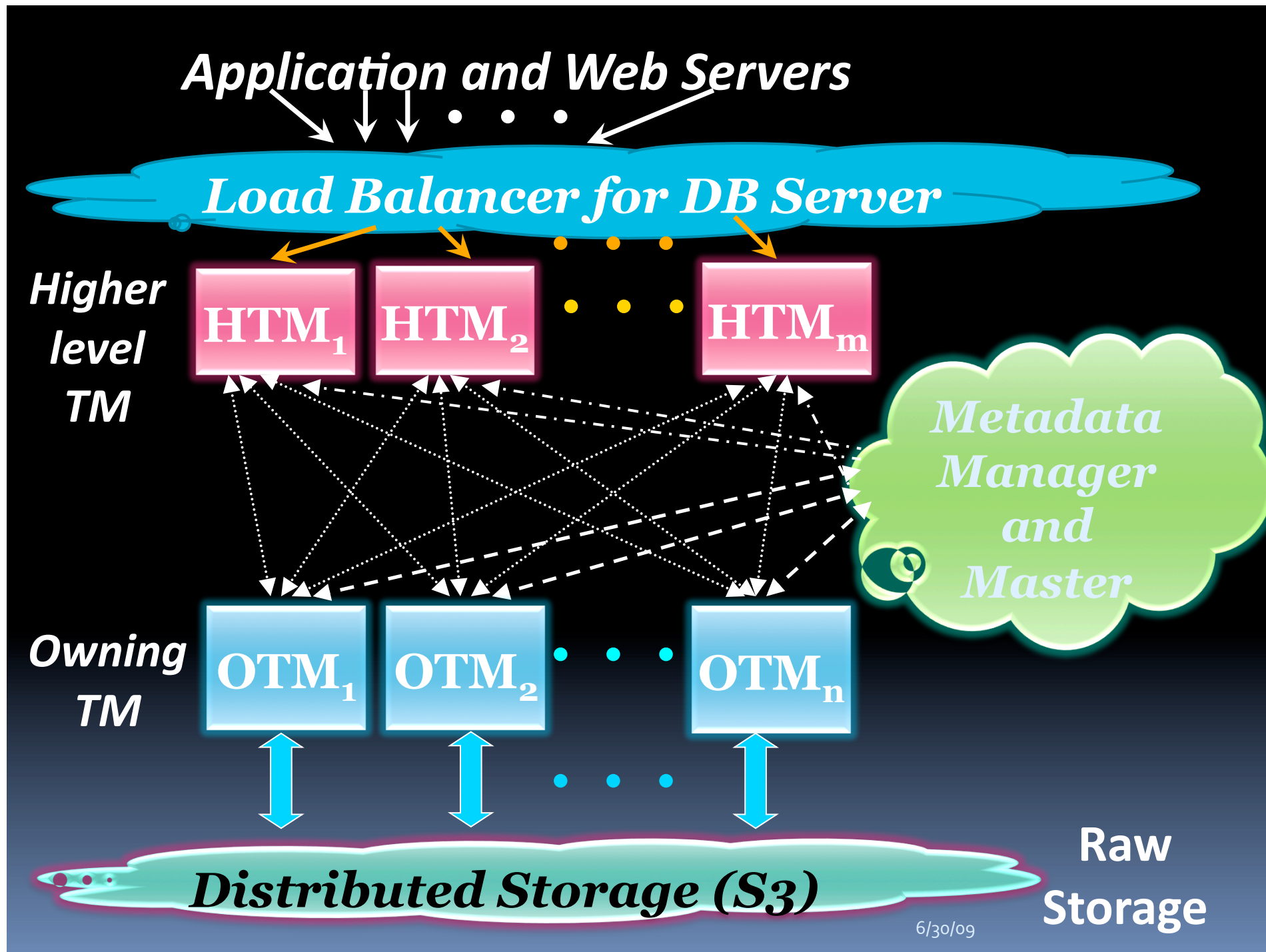
...



Distributed Storage (S3)

Raw
Storage

6/30/09



Digging into ElasTras

- Partitioned database
- Can be configured for static as well as dynamic partitioning
- Limited transactional semantics
- Support for minitransactions [Sinfonia]

Design Principles

- Segregate System metadata from Application Specific data
 - Different semantics – different requirements
- Limit Application Interaction to Single physical machine
 - Restricted transactional semantics
- Limited Distributed synchronization is practical
 - Synchronous replication and Distributed consensus for consistent and fault tolerant storage of meta data

Concluding Remarks

- Easy transition of partitioned Enterprise database systems into the cloud
- Flexible schema for supporting a wide variety of applications
- Static Partitioning: Can support transactions limited to partitions
- Dynamic Partitioning: Can support only *minitransactions*
- Elasticity through partitioning

ThanQ

Questions