CLOUD COMPUTING APPLICATIONS

SWIFT – An Object Store

Roy Campbell & Reza Farivar

Definition

A **Binary Large OBject** (**BLOB**) is a collection of binary data stored as a single entity in a database management system.

(The blob was a science fiction movie featuring Steve McQueen.)

https://en.wikipedia.org/wiki/Binary_large_object

Use case

- Store unstructured object data like text or binary data
- Images
- Movies
- Audio, Signal data
- Large queue of messages
- Example is LinkedIn data in a user page (Uses Ambry)
- Usually accessible over the web

Examples

- Windows Azure Blob Storage
- Ambry – LinkedIn
- Facebook’s Warm BLOB Storage System
- Amazon Simple Storage Service (S3)*
- Apache Open Stack Blob Service (SWIFT)

Goals

- Data growth ~ 50% a year
- 50%-70% data is unstructured or archival
- RESTful API (HTTP)
- High availability (no single point of failure)
- Agile data centers
- Open Source
- Multi-region, geographic distribution of data
- Storage policies
- Erasure Coding

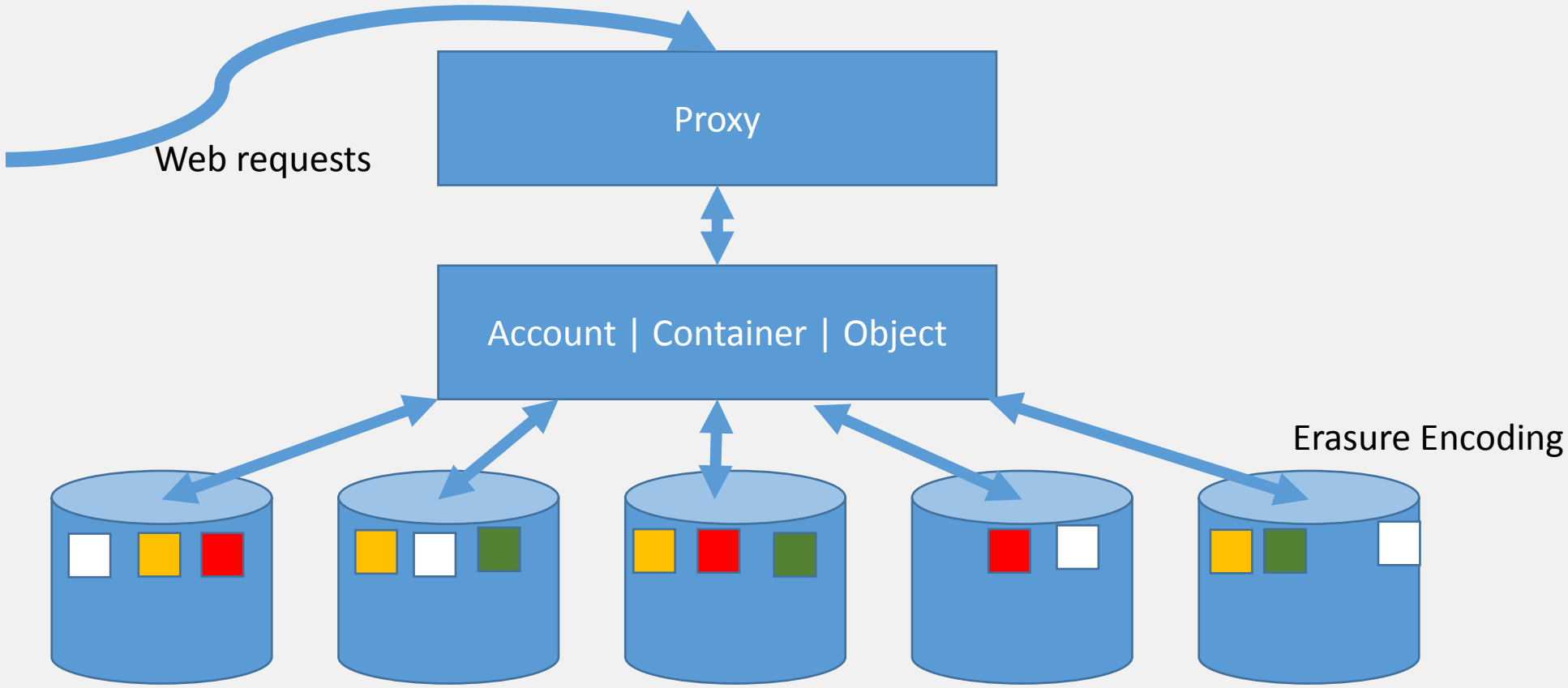
Swift API

https://swift.illinois.edu/version2/auth_account/container/object

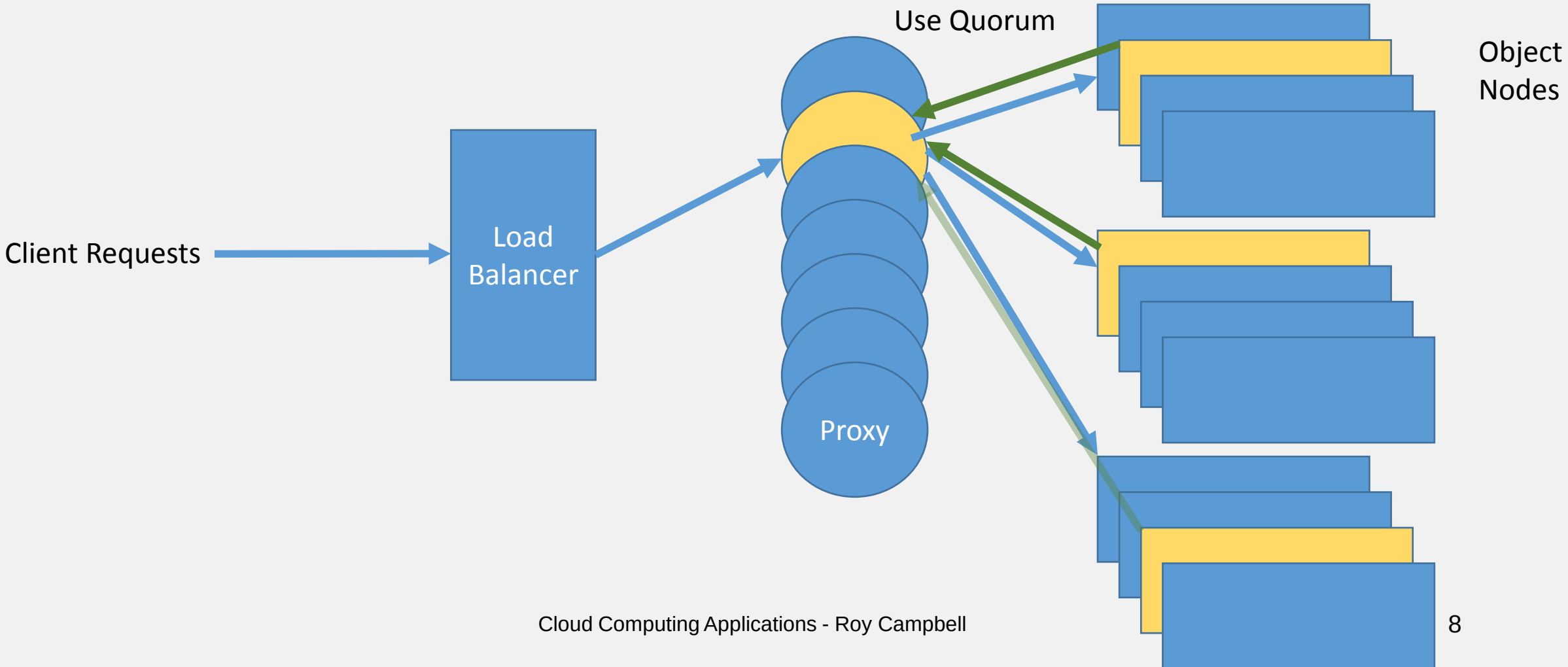
PUT /version2/roy/myblobs/classvideo1

GET /version2/reza/hisblobs/yesterdaysdataforhadoop

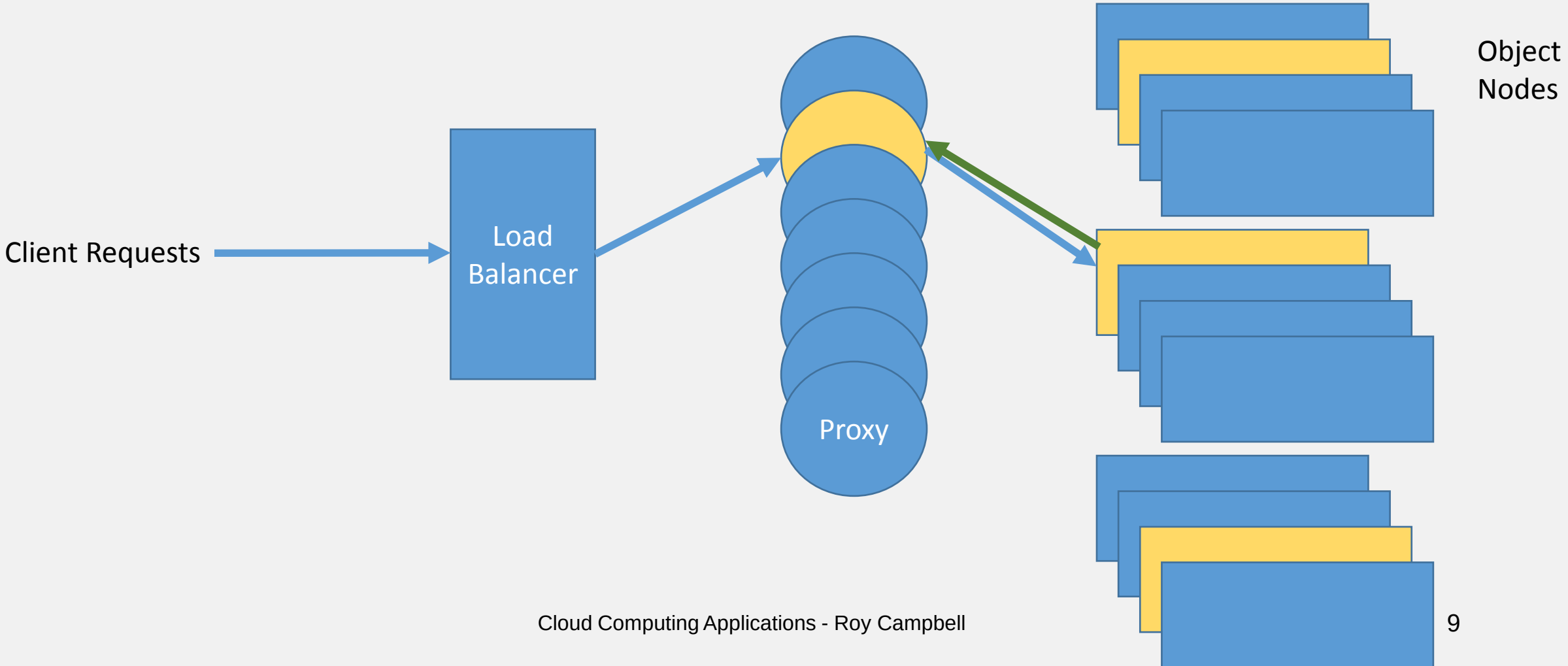
Swift Components



Write Requests - Load Balancer and Proxy



Read Requests - Load Balancer and Proxy



Details

- MD5 Checksums with each object
- Auditing and active replication
- Any sized disks

Swift Partitions

- 1 Node, 8 Disks, 16 Partitions per disk, 8*16 = 128 partitions
- 2 Nodes, 8 Disks each, 8 Partitions per disk, 8*16 = 128 partitions
- Use Hash into Ring to map objects into storage partitions