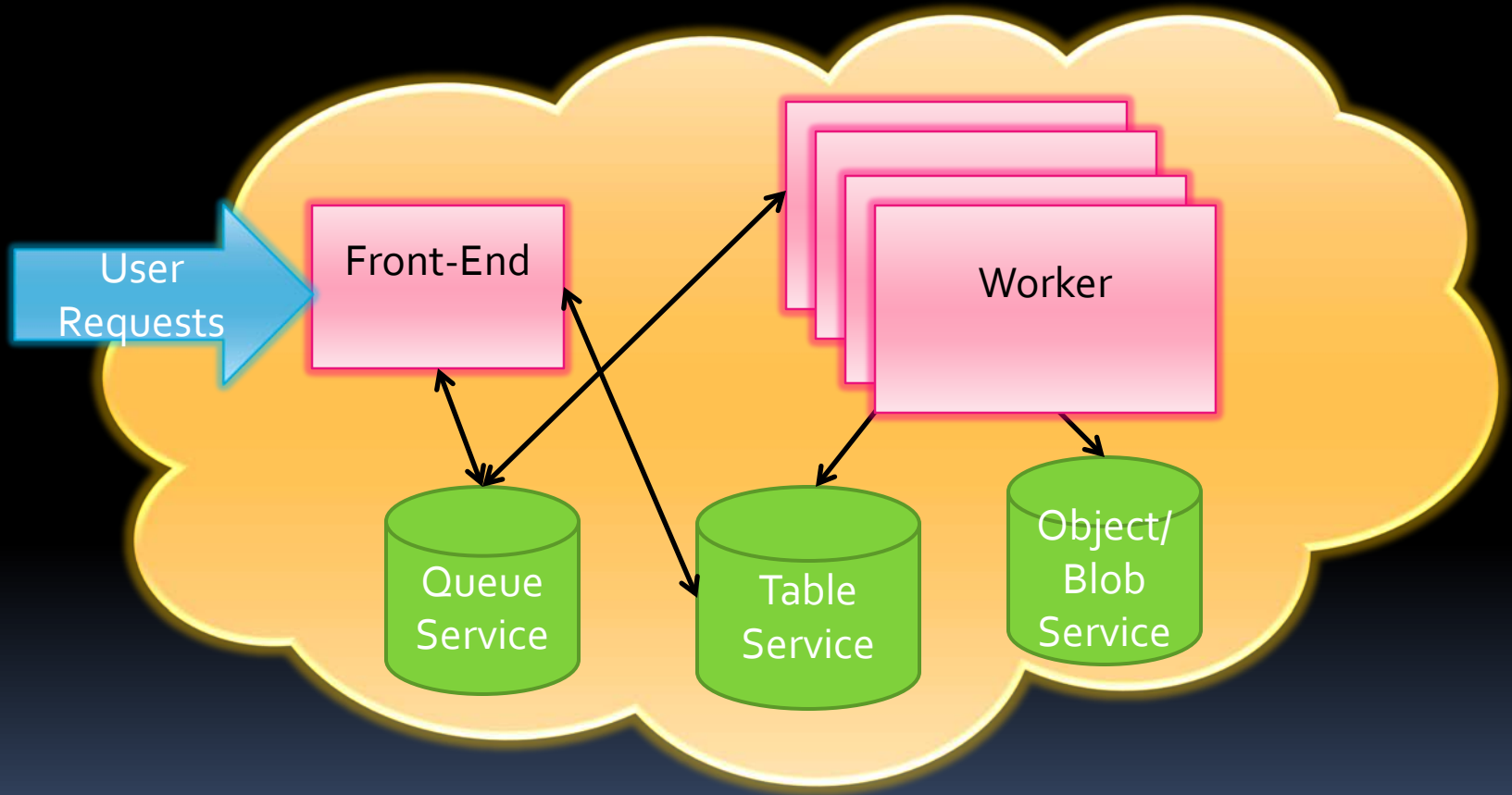


CSAL: A CLOUD STORAGE ABSTRACTION LAYER TO ENABLE PORTABLE CLOUD APPLICATIONS

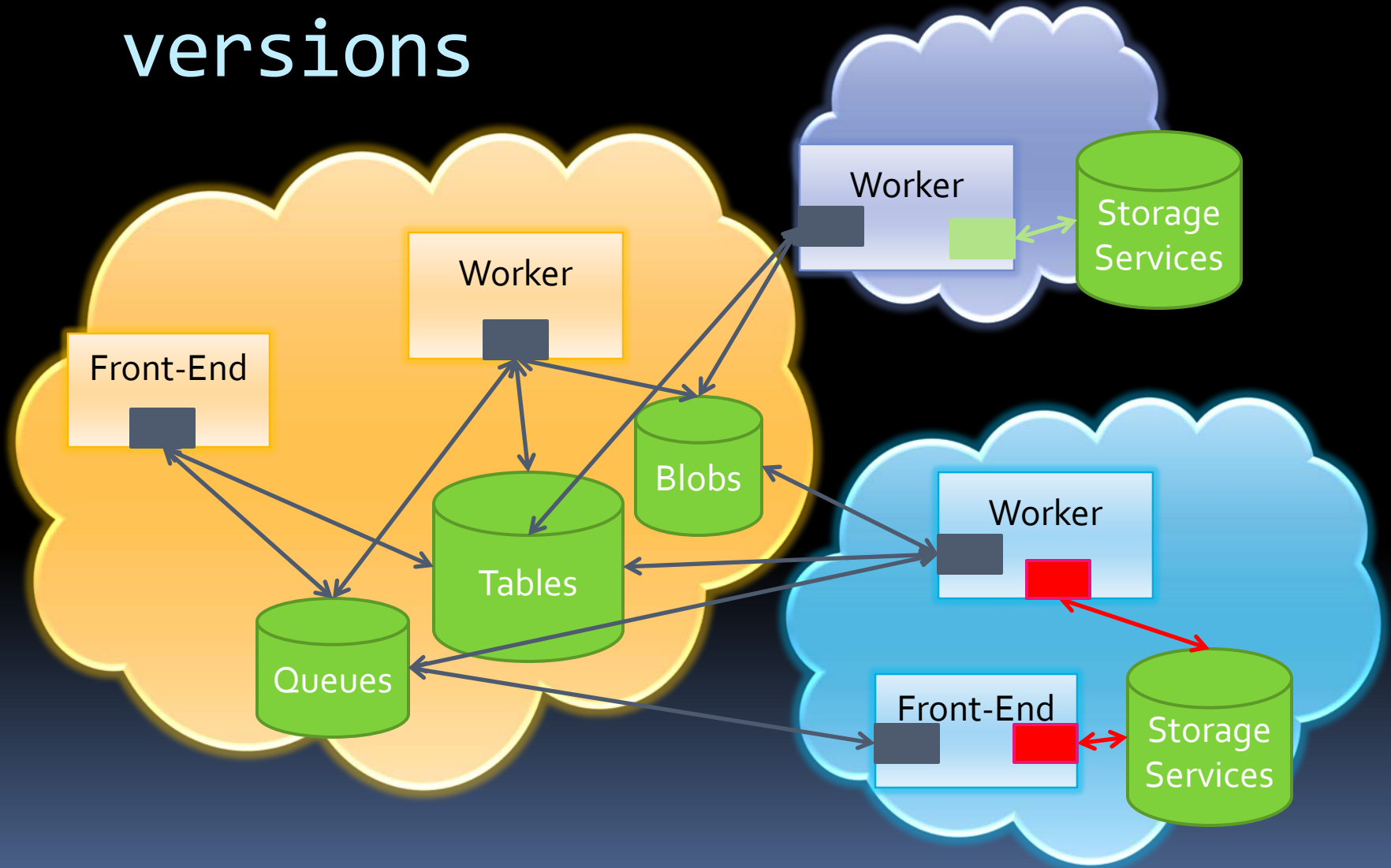
Zach Hill & Marty Humphrey
Dept. of Computer Science, University of Virginia

zjh5f@cs.virginia.edu

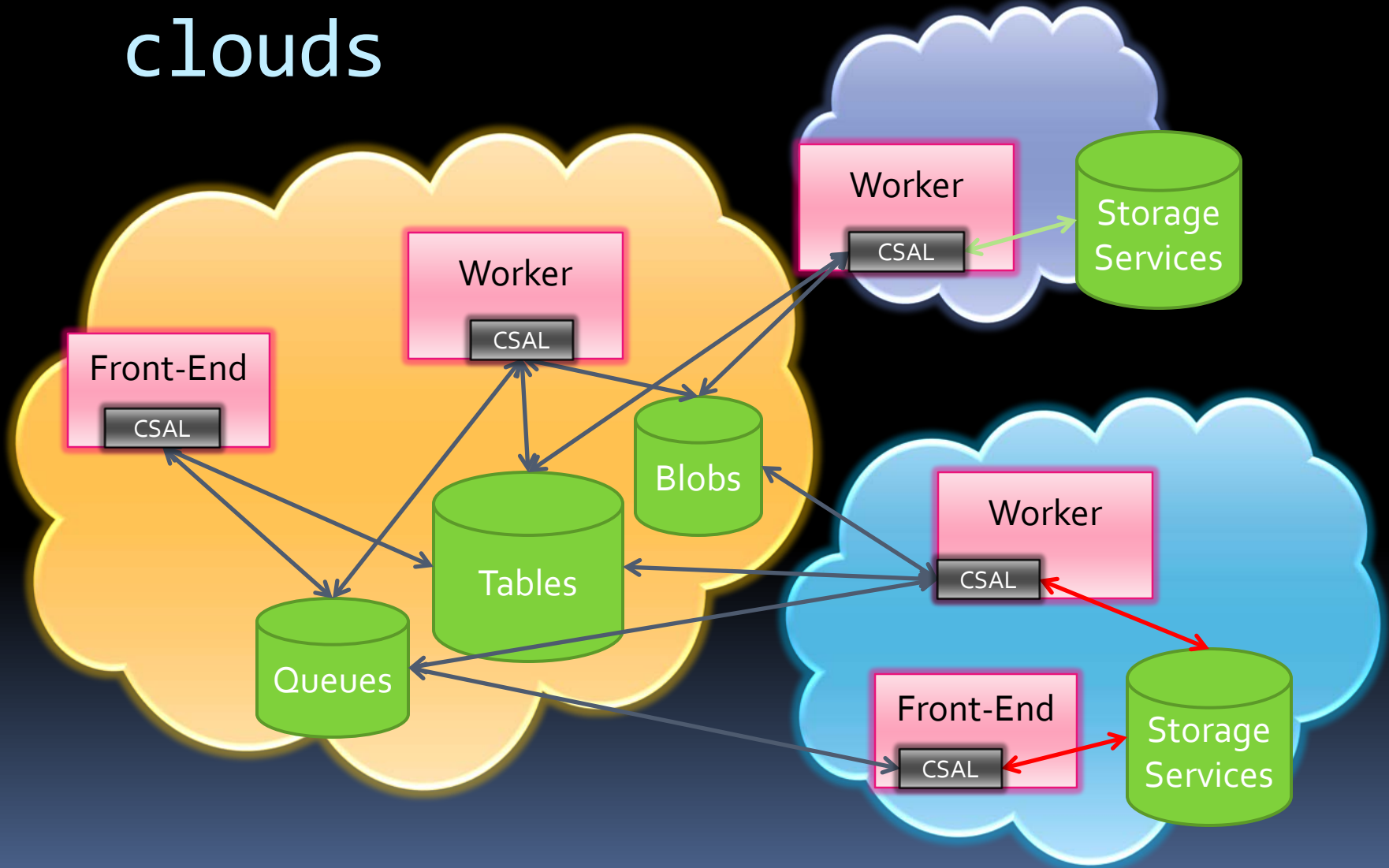
A Cloud Application



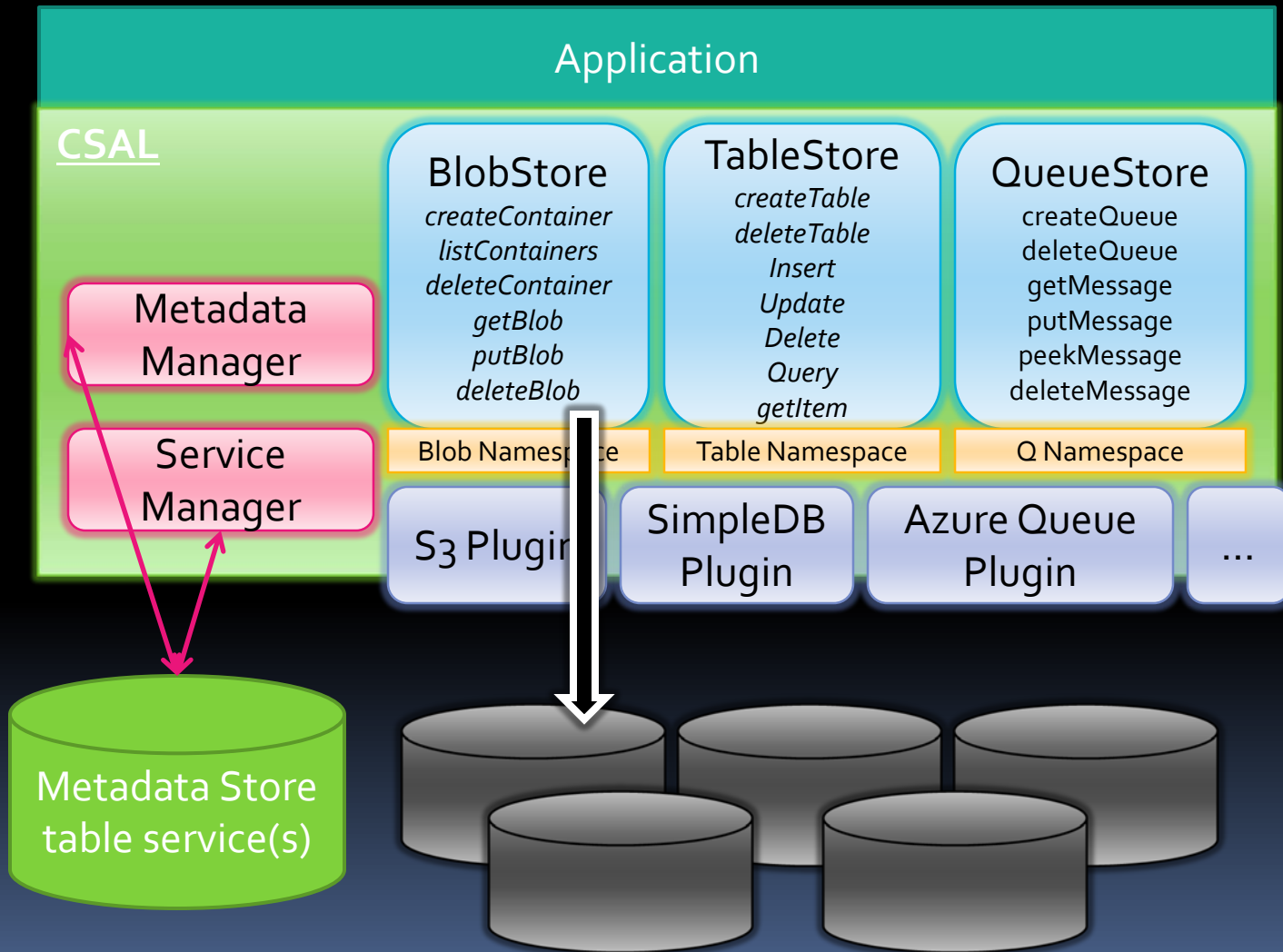
Many clouds, many code versions



Single code version, many clouds



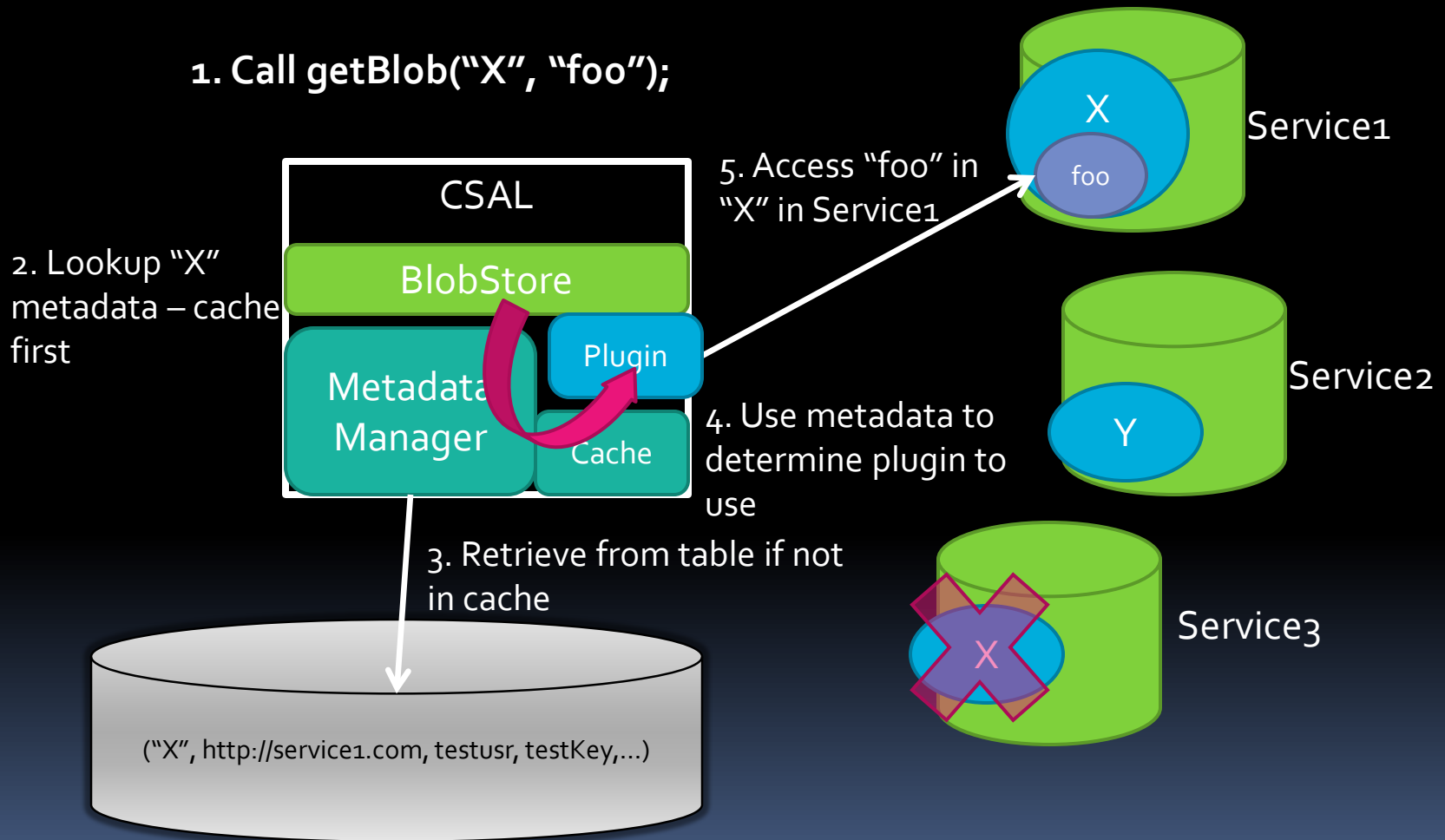
CSAL Overview



CSAL Namespaces

- One namespace for each abstraction type
- Metadata only for containers
 - Service endpoint, identifier, user credentials
- Each abstraction has an independent metadata store
- Metadata caching
 - Container ops are not very common
 - If data is stale, simply re-fetch and retry

CSAL Namespaces



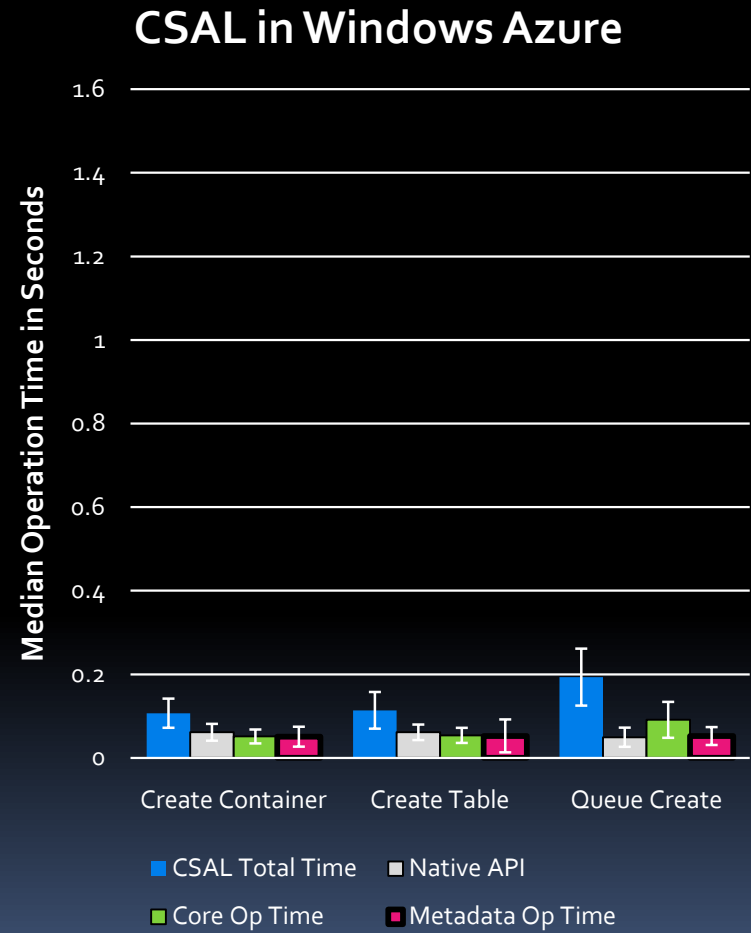
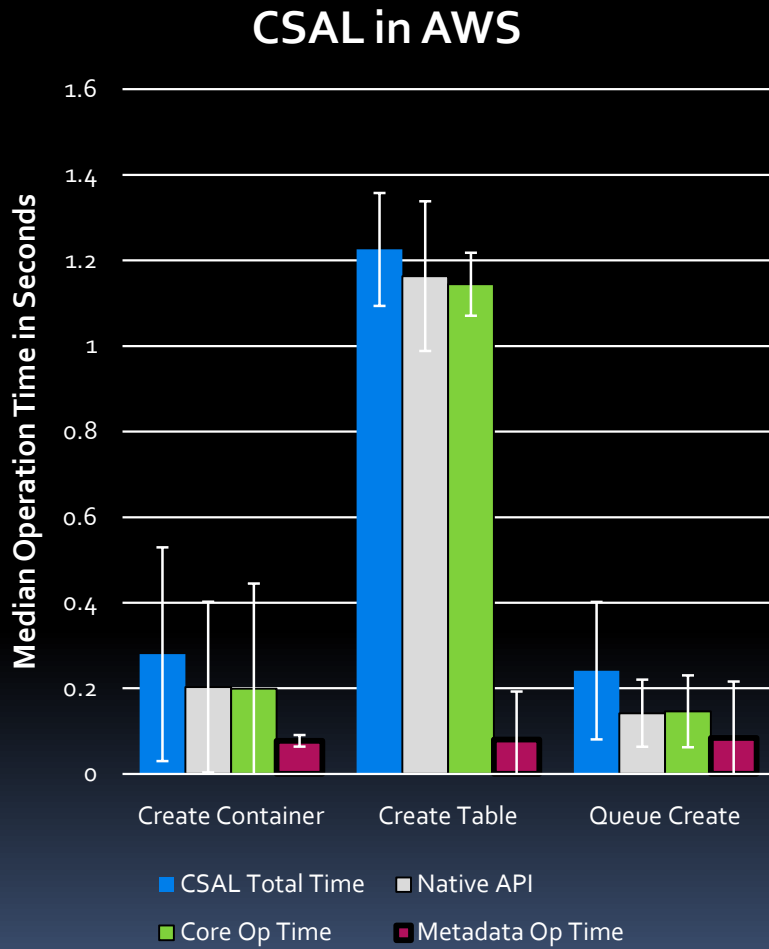
CSAL Implementation

- Client-side java library
 - i.e. *BlobStore.getBlob("Container","foo");*
- Metadata backing store in the cloud
- Currently supports Azure & AWS storage
 - Both SOAP and REST

Performance of CSAL

- Adding software layers isn't free
- Compare CSAL to Azure's and Amazon's SDK APIs
- Set of micro-benchmarks to test operation latency
- Container Ops and Data Ops
 - Expect a slowdown for container ops due to metadata

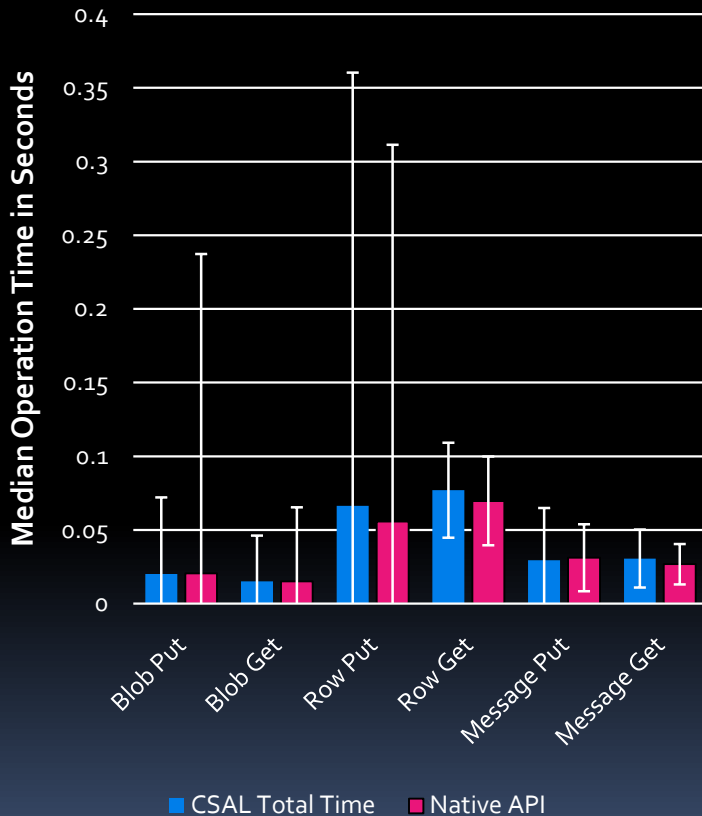
Performance – Container Ops



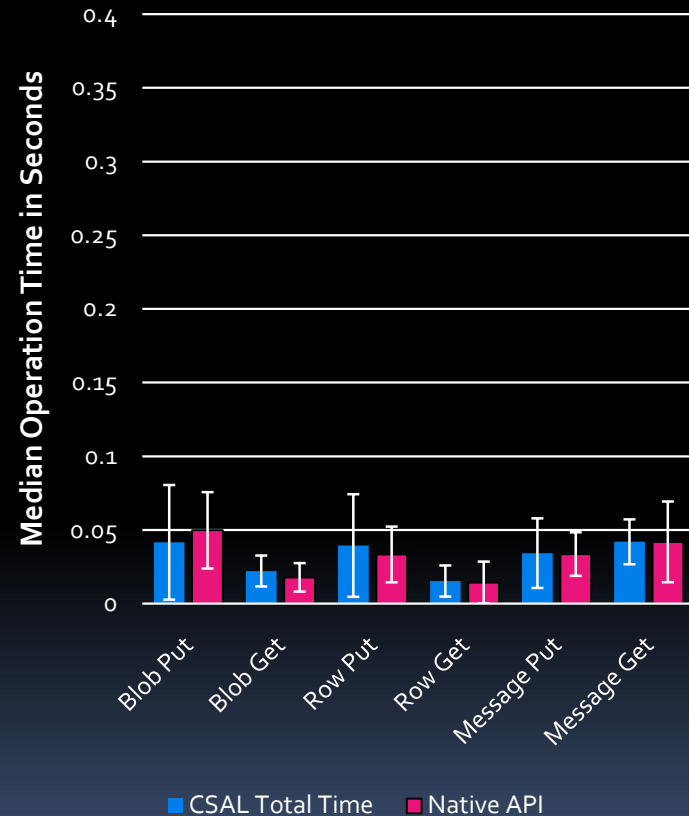
Note: Error Bars indicate 1 Standard Deviation

Performance – Data Ops

CSAL in AWS



CSAL in Windows Azure



Note: Error Bars indicate 1 Standard Deviation

What about Standards?

- Standards Efforts
 - OCCI
 - OVF
- Standards take time to develop and are resisted by vendors
- Multi-cloud APIs
 - SimpleCloud, jClouds, DeltaCloud, LibCloud
 - SAGA

Future work

- What if Cloud X doesn't have tables/blobs/queues?
 - Map one abstraction to other (i.e. filesystem)
 - 3rd party services: Hbase, HyperTable, Cassandra...
- Placing, replicating, and migrating data in real-time for performance and/or cost
- Real-world applications such as multi-cloud MR

Summary

- Application lock-in and portability are problems in clouds
- Standards are great, but don't hold your breath just yet
- CSAL provides storage abstractions to make the application code itself portable with little performance impact for common data operations

Questions?

zjh5f@cs.virginia.edu