

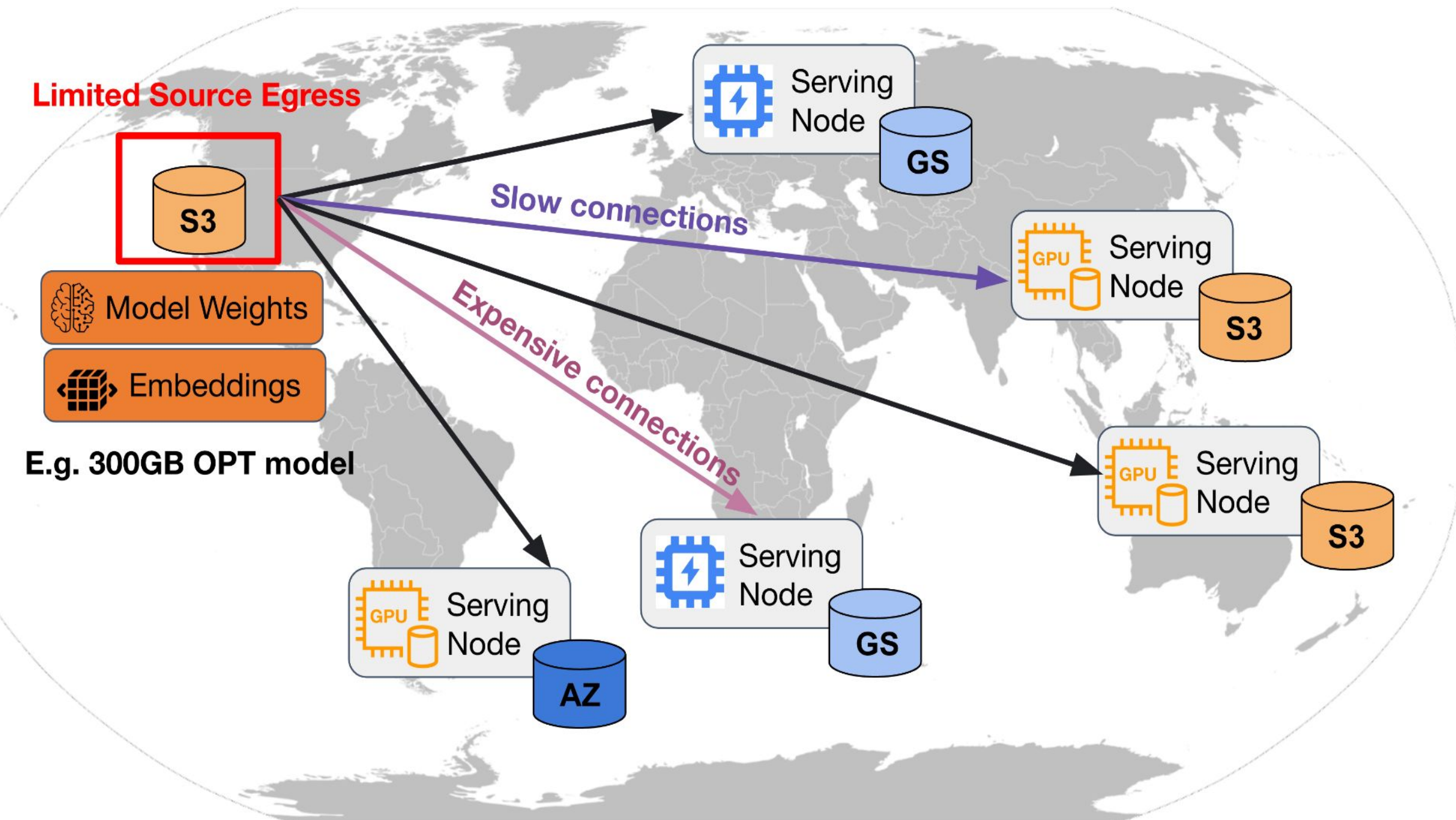
# Skystore: Unified Storage Across Clouds



Junhao Hu, Shaopu Song  
University of California - Berkeley

## Problem Statement

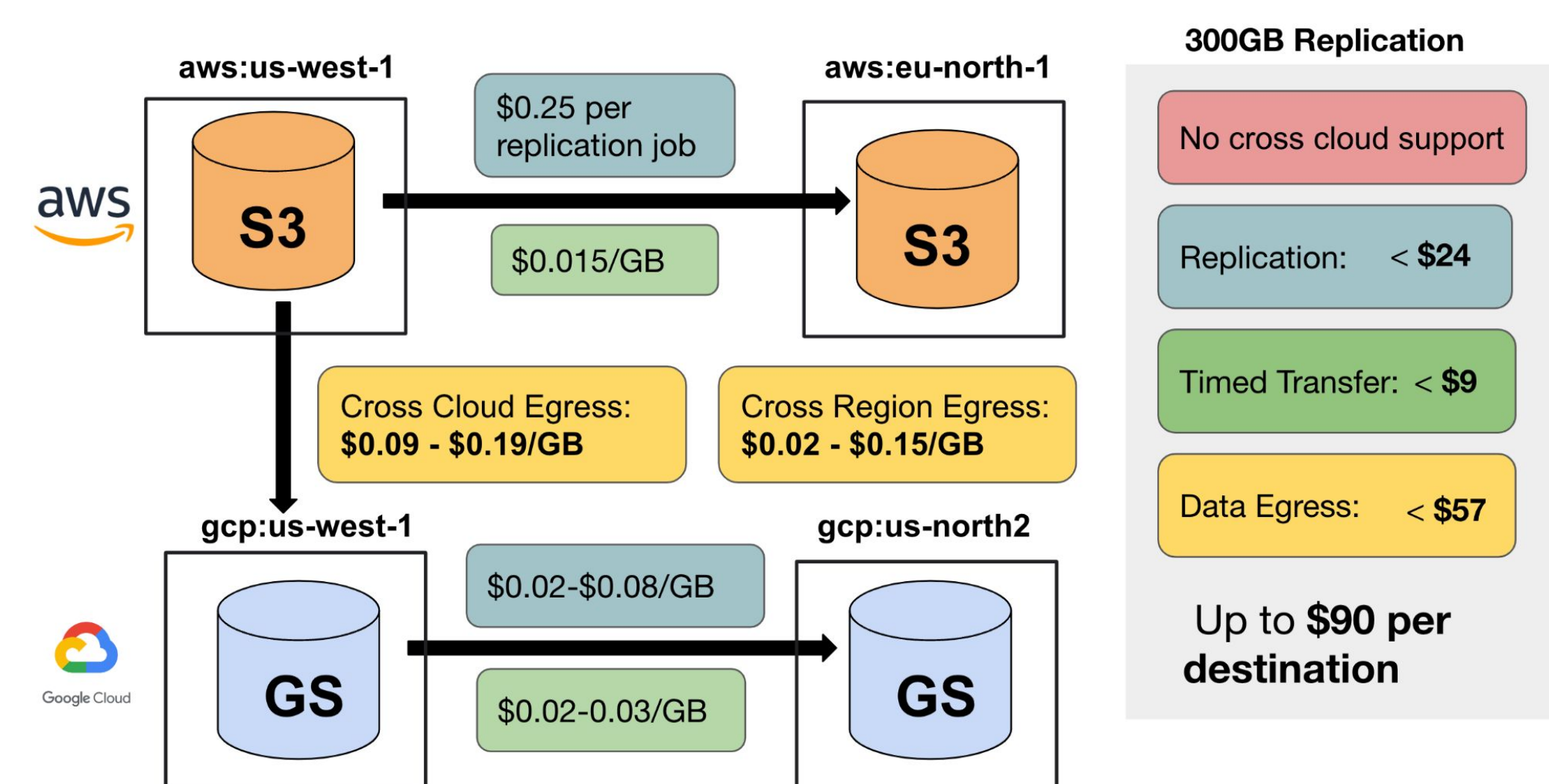
- Cloud applications (e.g. model propagation, feature replication) are increasingly replicating data across multiple regions and providers
- Unfortunately, wide-area cloud data dissemination are often **complex**, **slow** and **expensive**, bottlenecking the applications
- Goal: achieve **simpler**, **faster**, and **cheaper** data store across regions & clouds



## Existing solutions

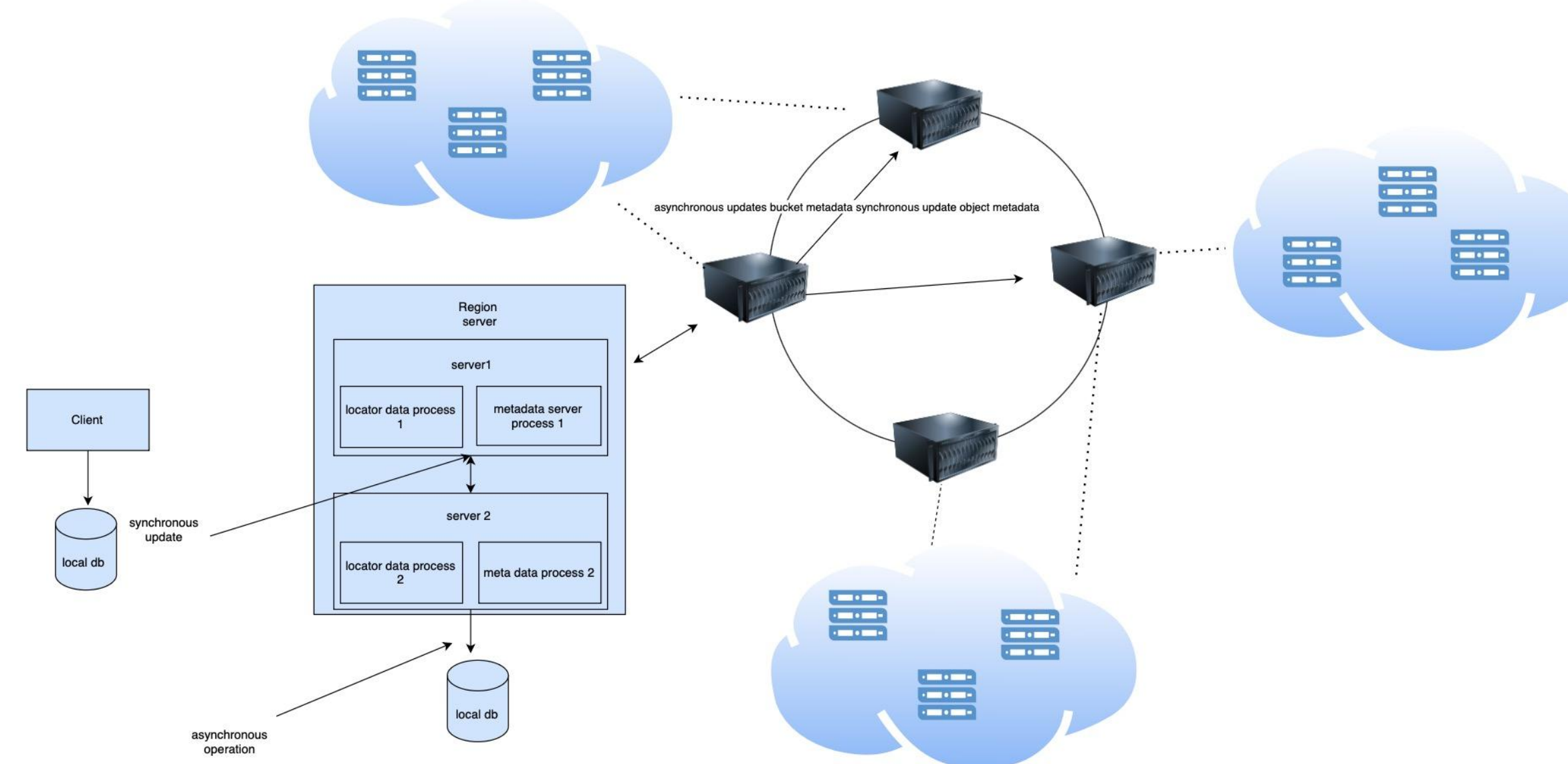
Store everywhere | Store in one place | Multi-region buckets

- Cons: high-cost (egress + service fees), high latency, lack of configurability and multi-cloud support
- None of these solutions are both cost- and throughput-aware

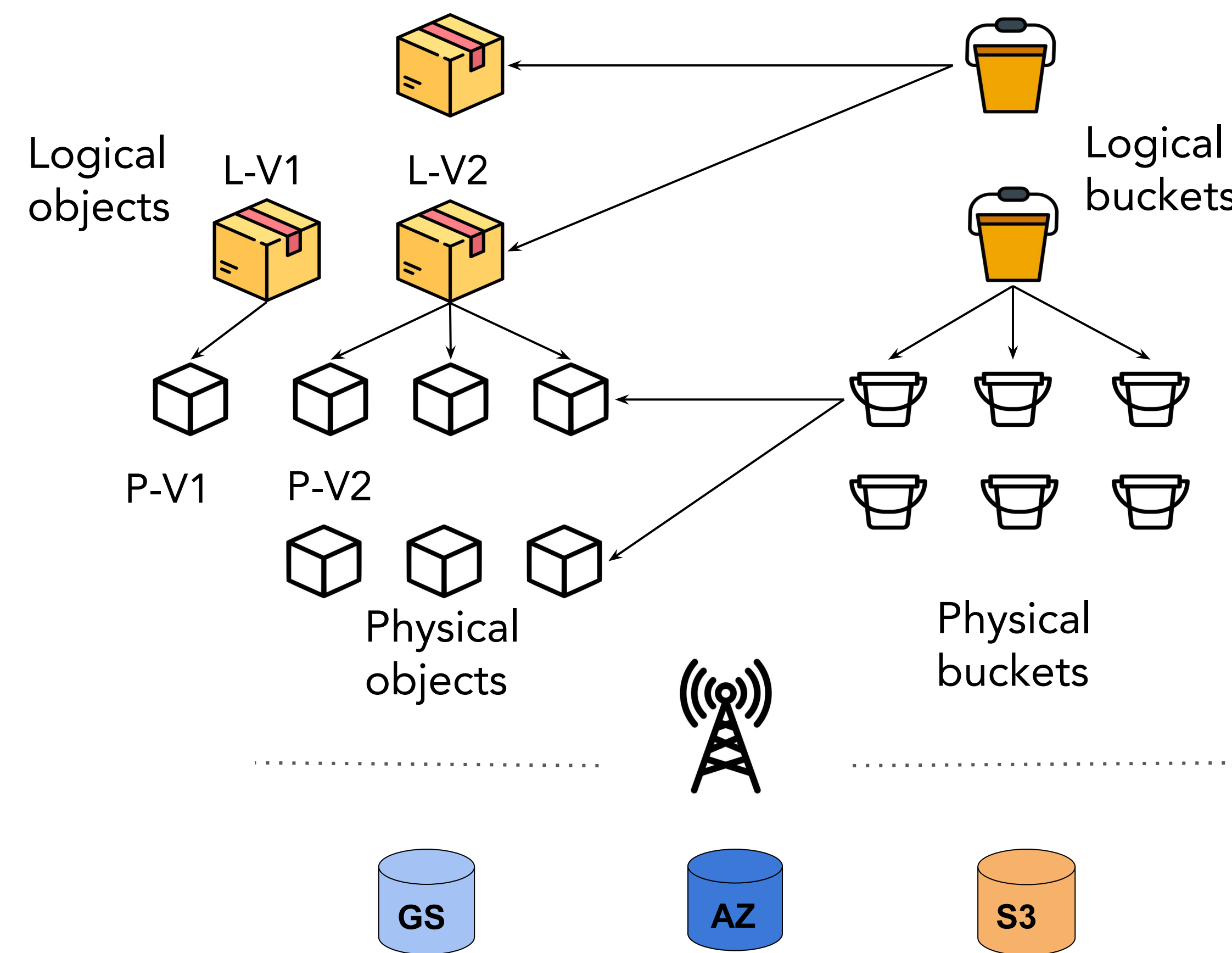


## Architecture

Our system: flexibly navigates the trade-off between price and performance for data store and data transfer. Specifically, there are two components of the system



(1) **Control Plane:**

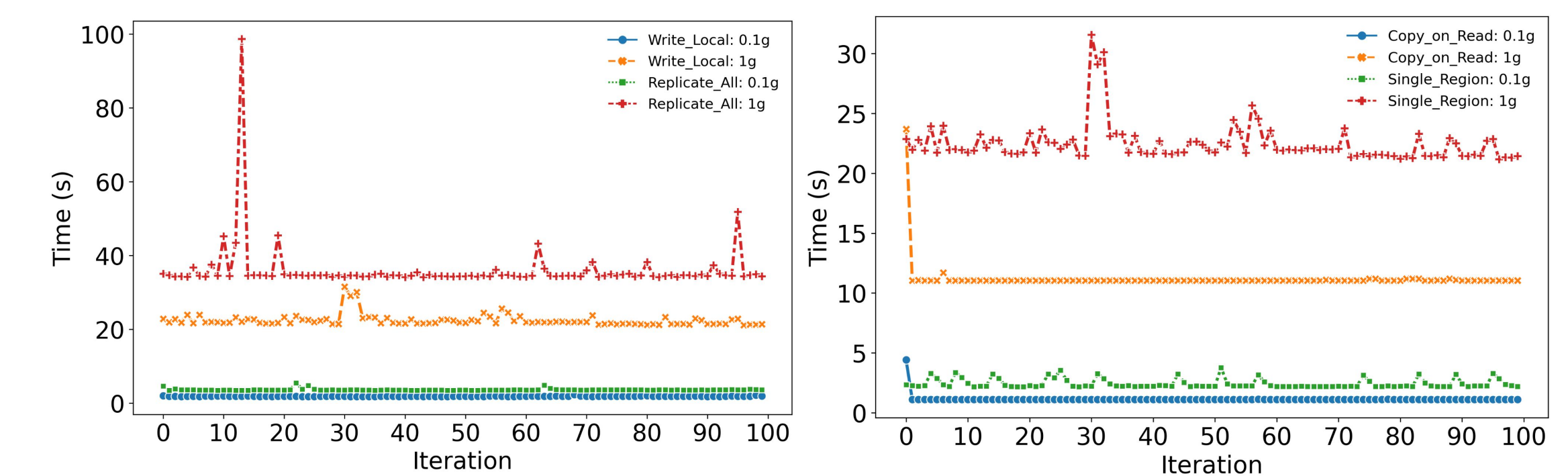


(2) **Data Plane:** execute data transfer/placement plan

- Unified interface for S3, GCS, Azure with full S3-compatible semantics API
- Stateless Design: not retain any internal state between different requests
- Interact with object stores and transfer/place data according to the control plane decision

## Skystore Policies

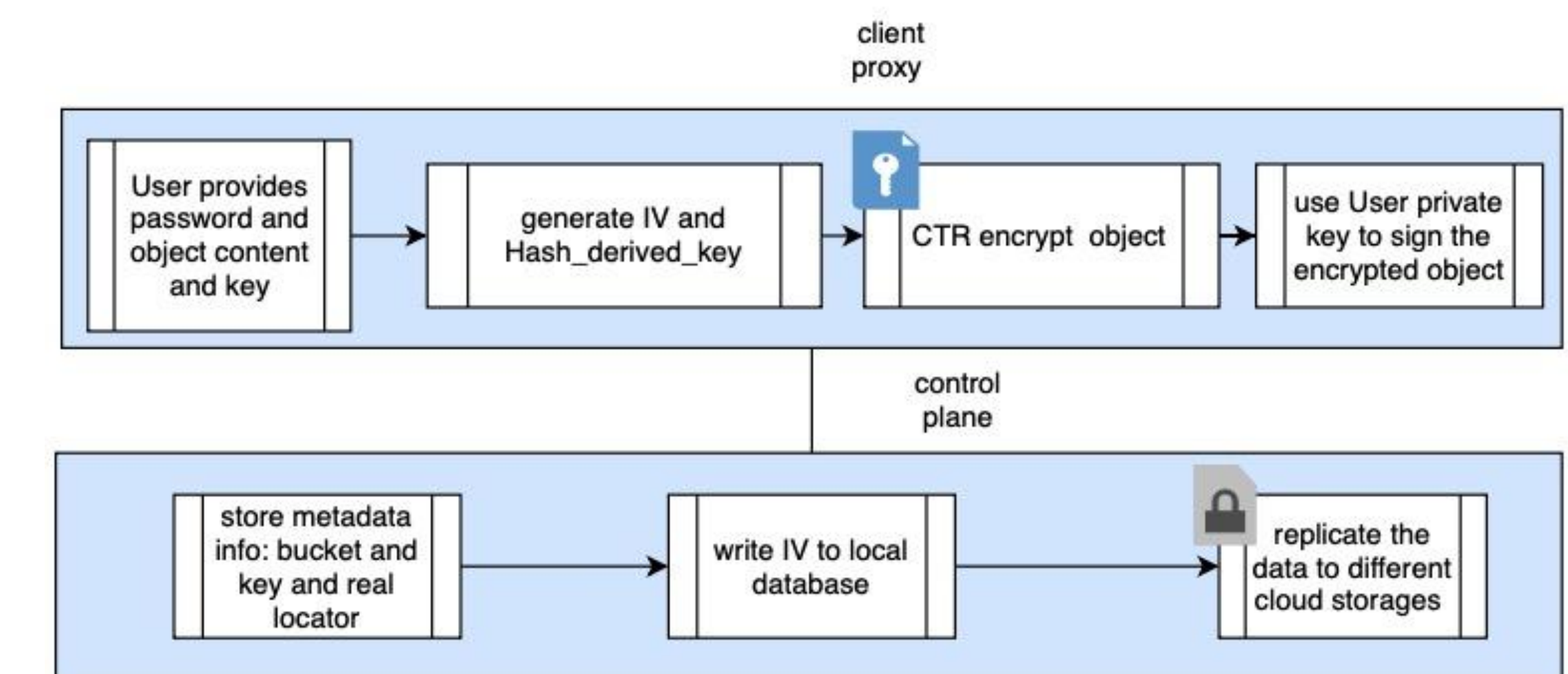
- Placement Policy (Write):**
  - Fast Write:** *Write-local* reduces latency and egress costs associated with the write operation; ensuring immediate availability of data
  - Cache:** *Copy-on-Read* lowers data transfer cost when repeated reads are often.



b. **Transfer Policy (Read):**

- Closest | Cheapest

## Security



## Evaluation

End to end evaluation:

