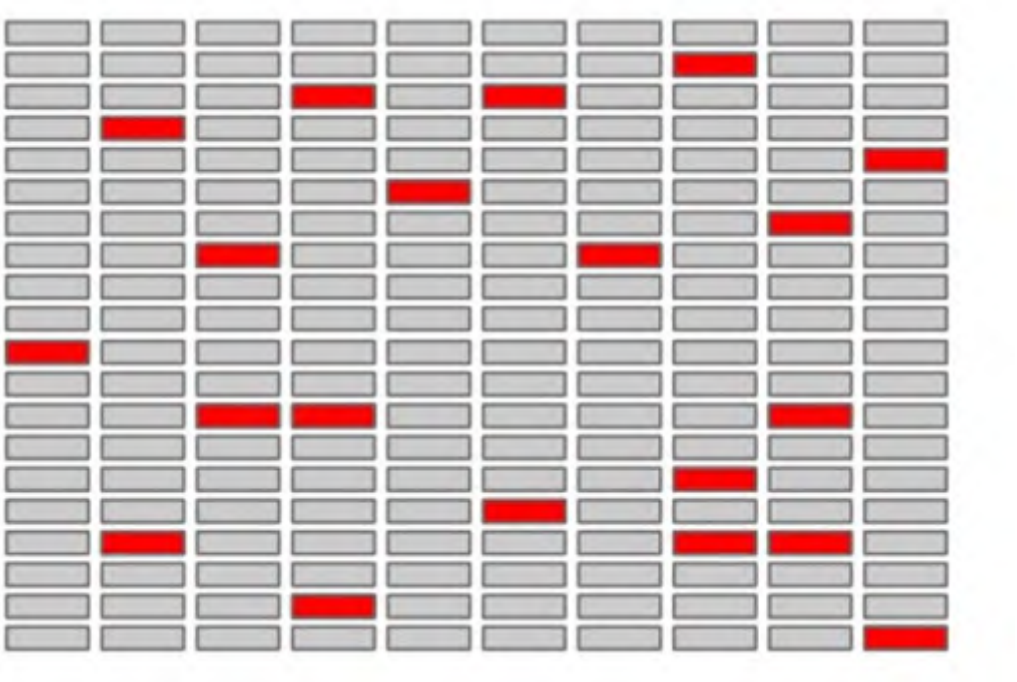
Stronger Consistency Simplified w/ **Apache DistributedLog**

@sijieg

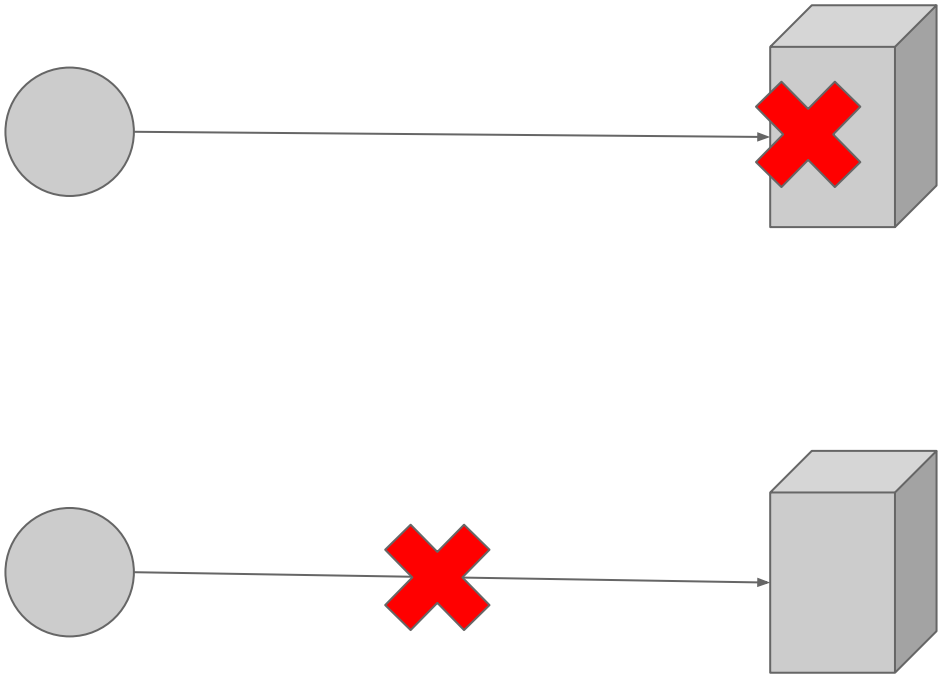
Challenges in Distributed Systems

Expect Failures

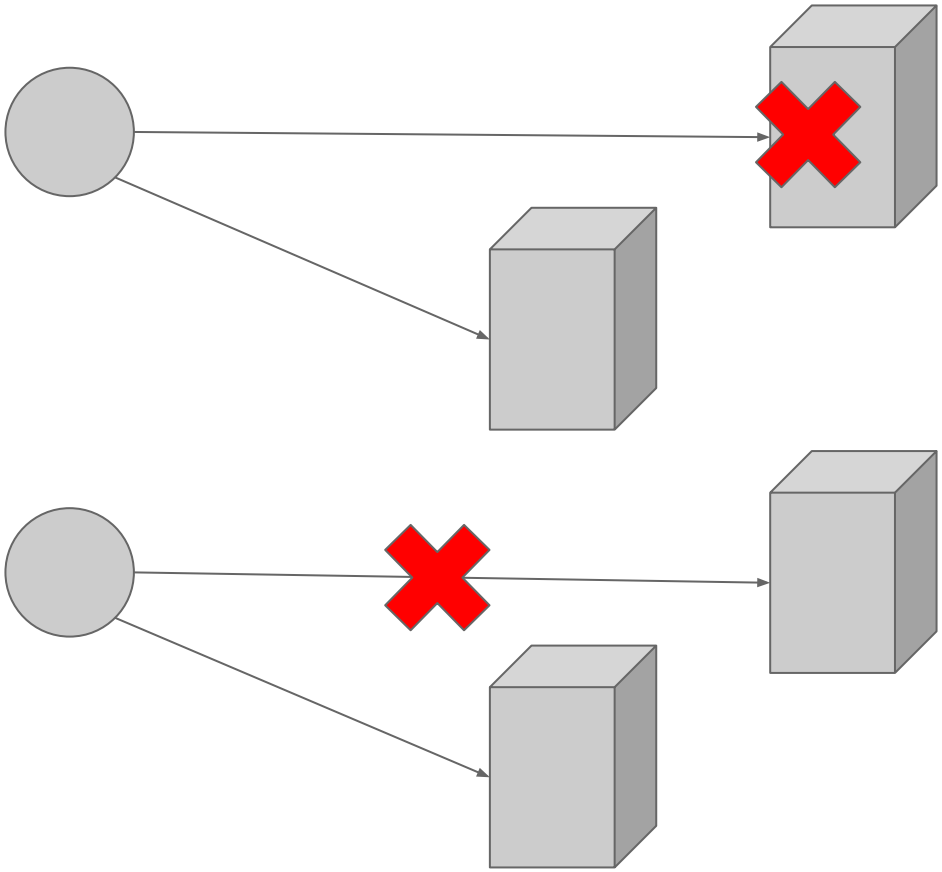


up to 10% annual failure rates for disks/servers

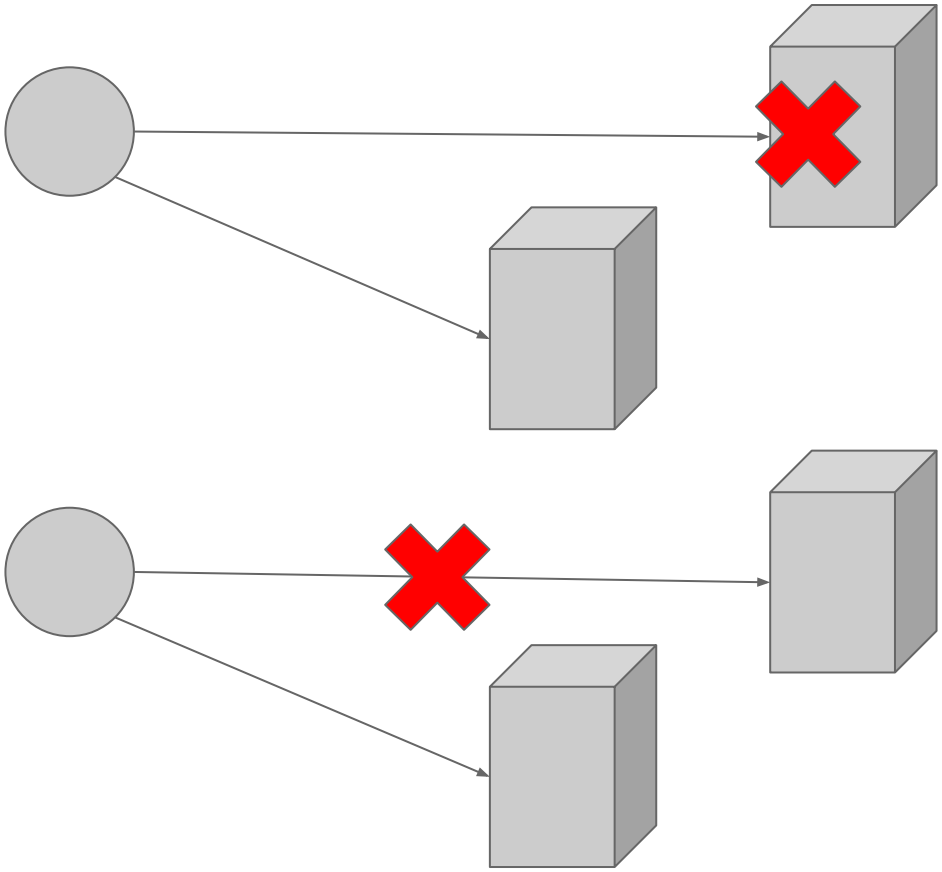
Problem 1: Not Available



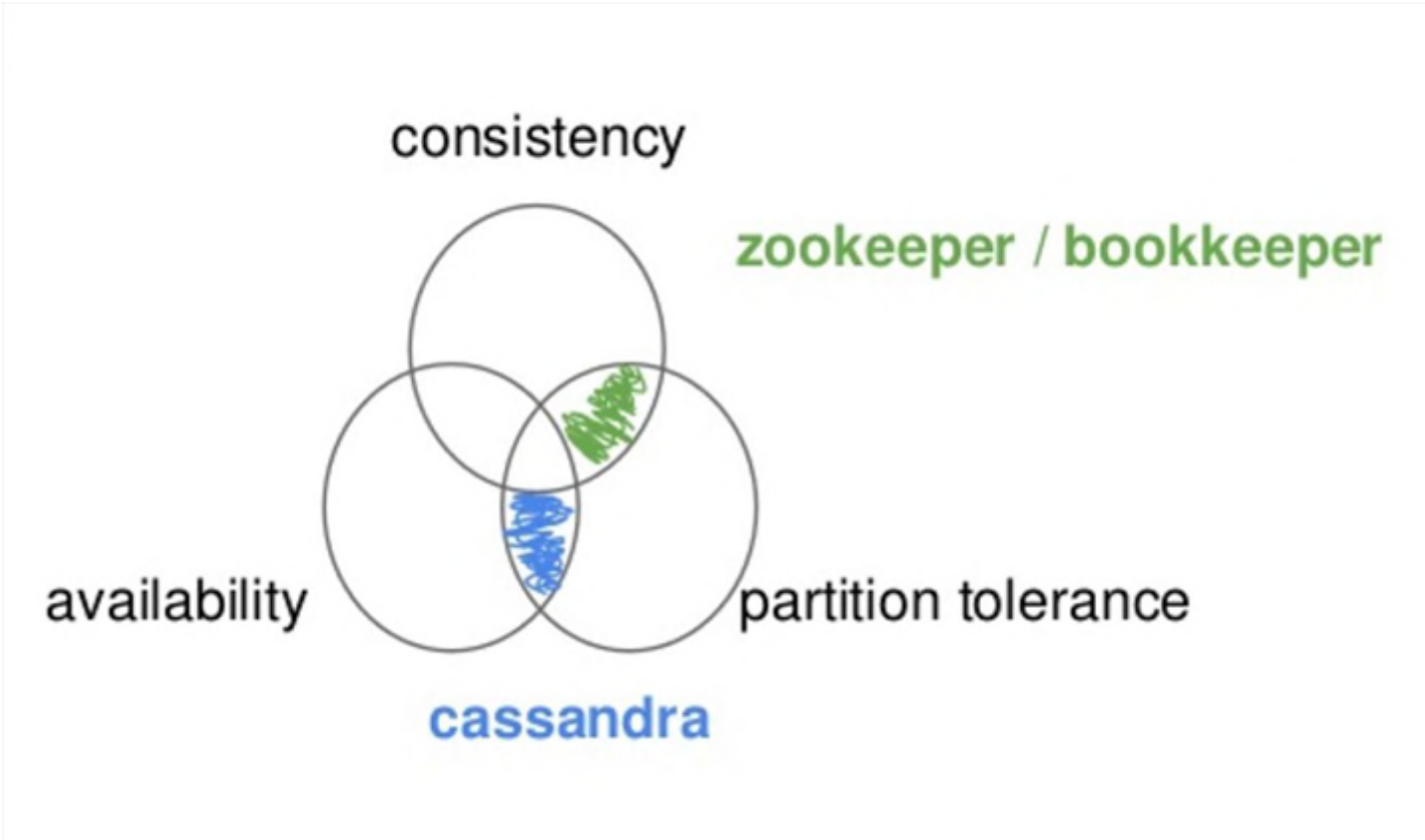
Problem 1: Not Available



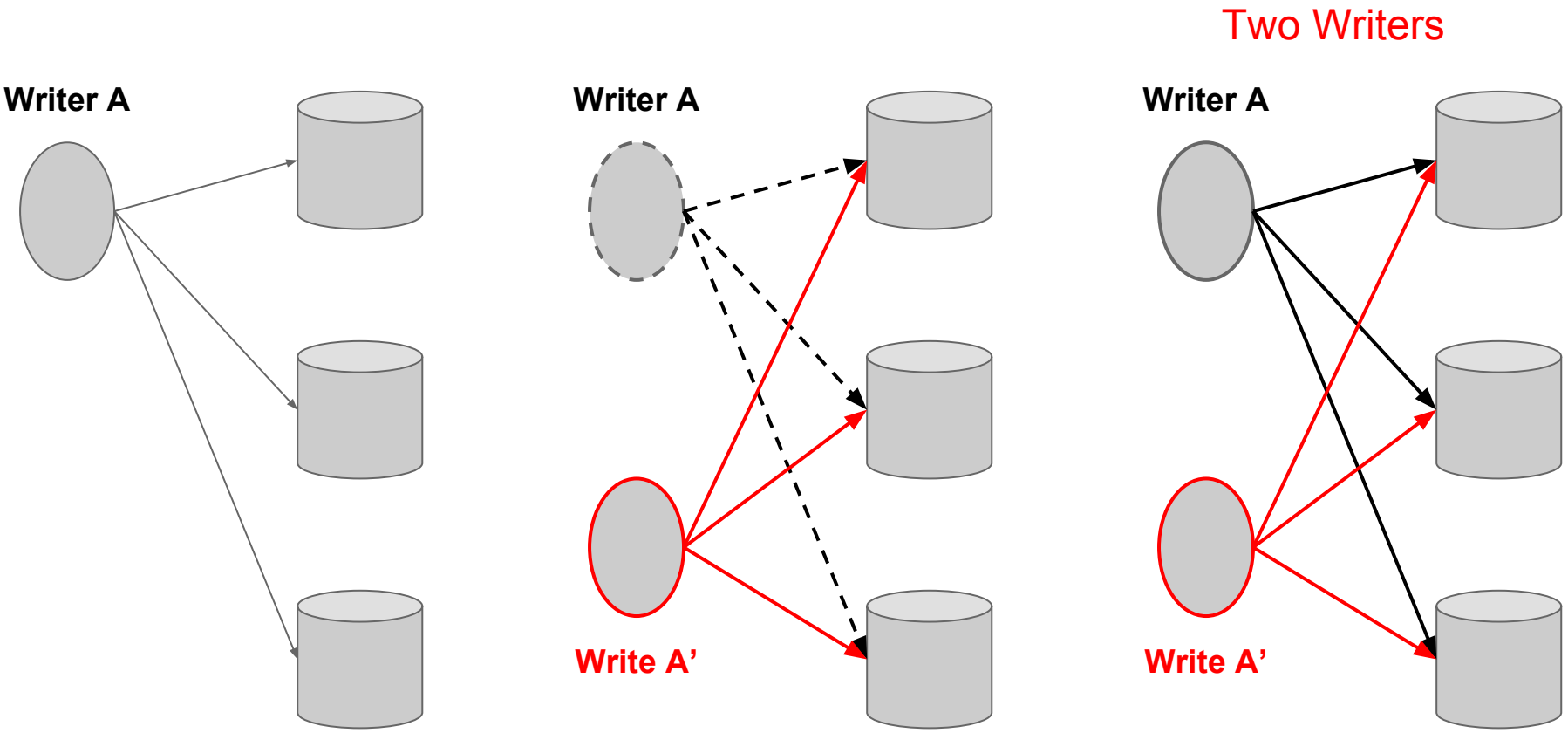
Problem 2: Inconsistencies



CAP


consistency
zookeeper / bookkeeper
availability
partition tolerance
cassandra

Problem 3: Split Brains



Solutions?

Consensus Algorithms

- Paxos
- Zab
- Raft
- ...

It is hard ...

Behind Consensus ...

- Order: which change comes first?
- Deterministic: Order won’t change even read multiple times
- How to keep a consistent replicated log?

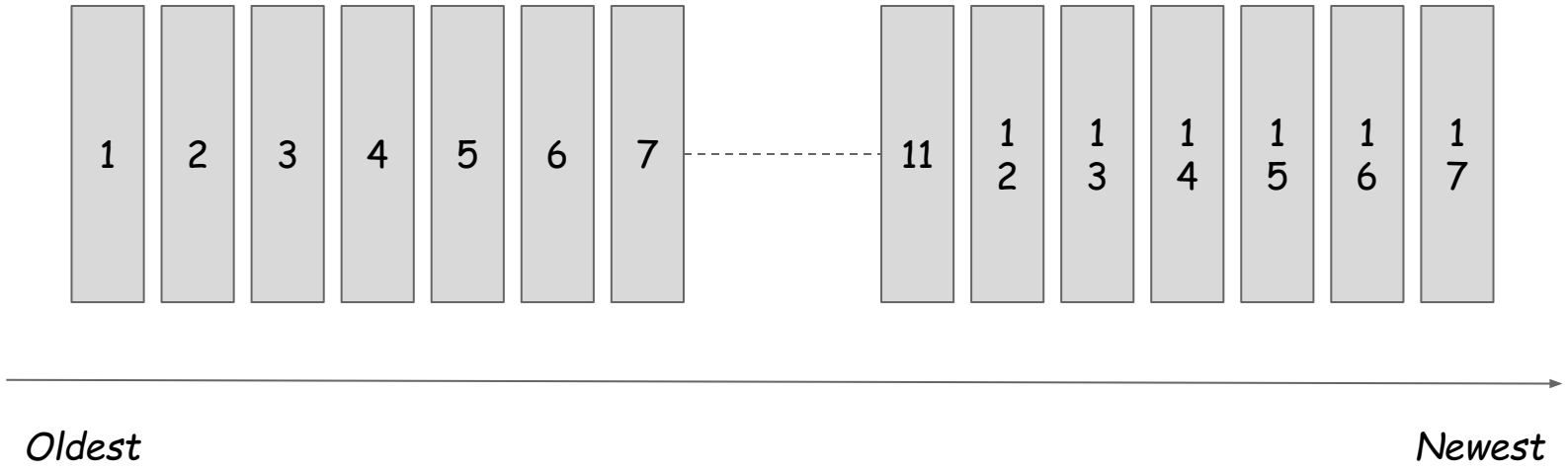
Solutions!!

Apache DistributedLog

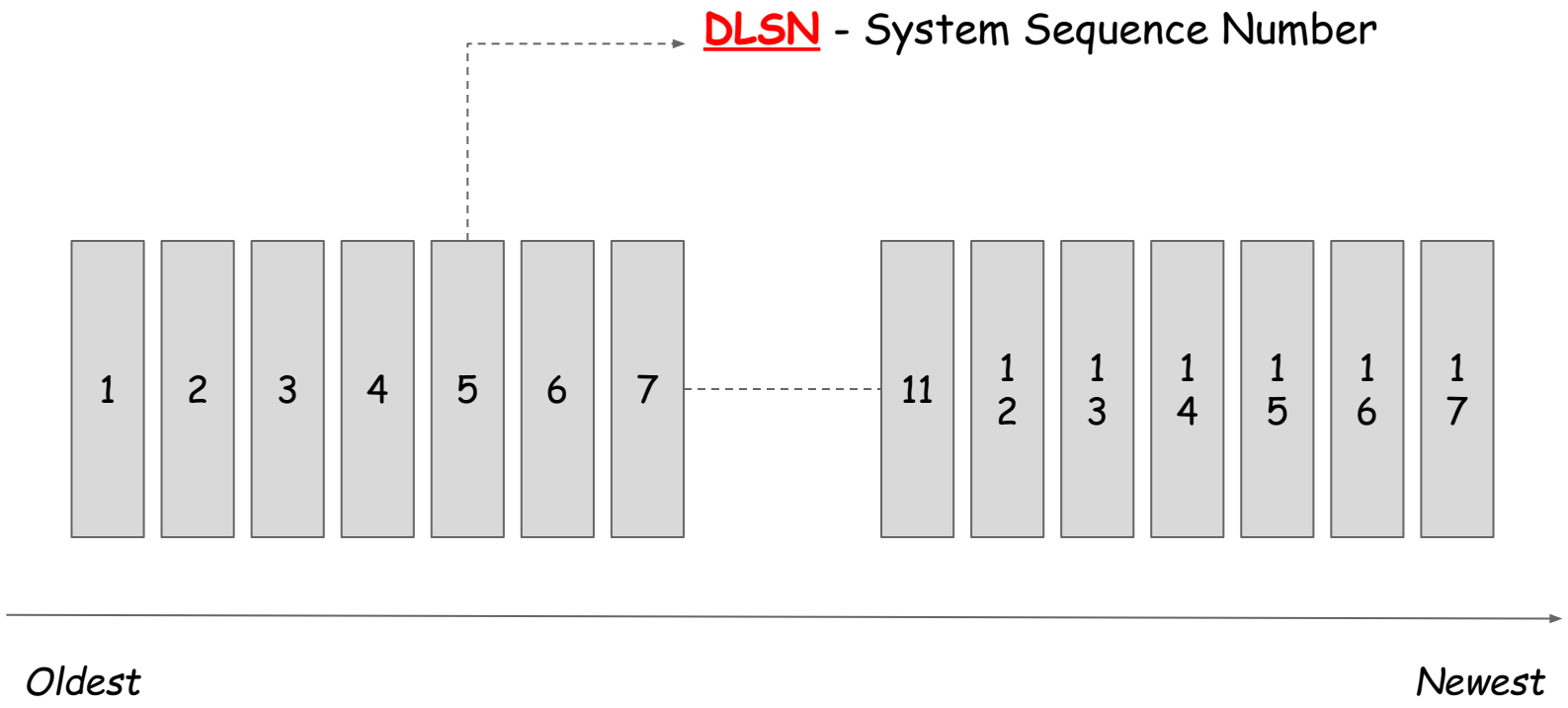
Apache DistributedLog

- Low Latency, High Performance Replicated Log Store
- Durable, Consistent
- Efficient Fan-in and Fan-out
- From journal/wal to general pub/sub messaging
- Multi Tenant
- Layered Architecture

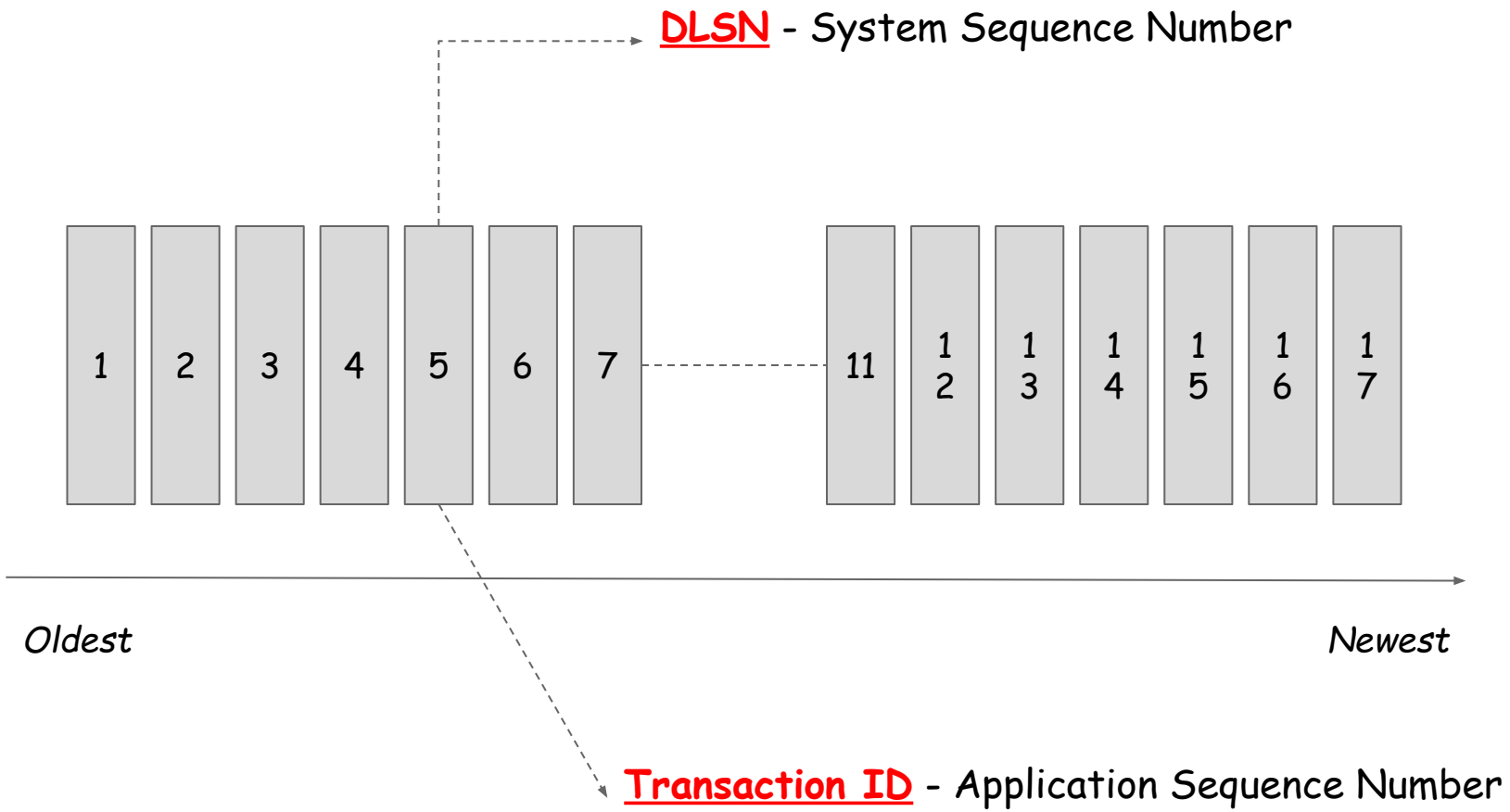
Log Stream



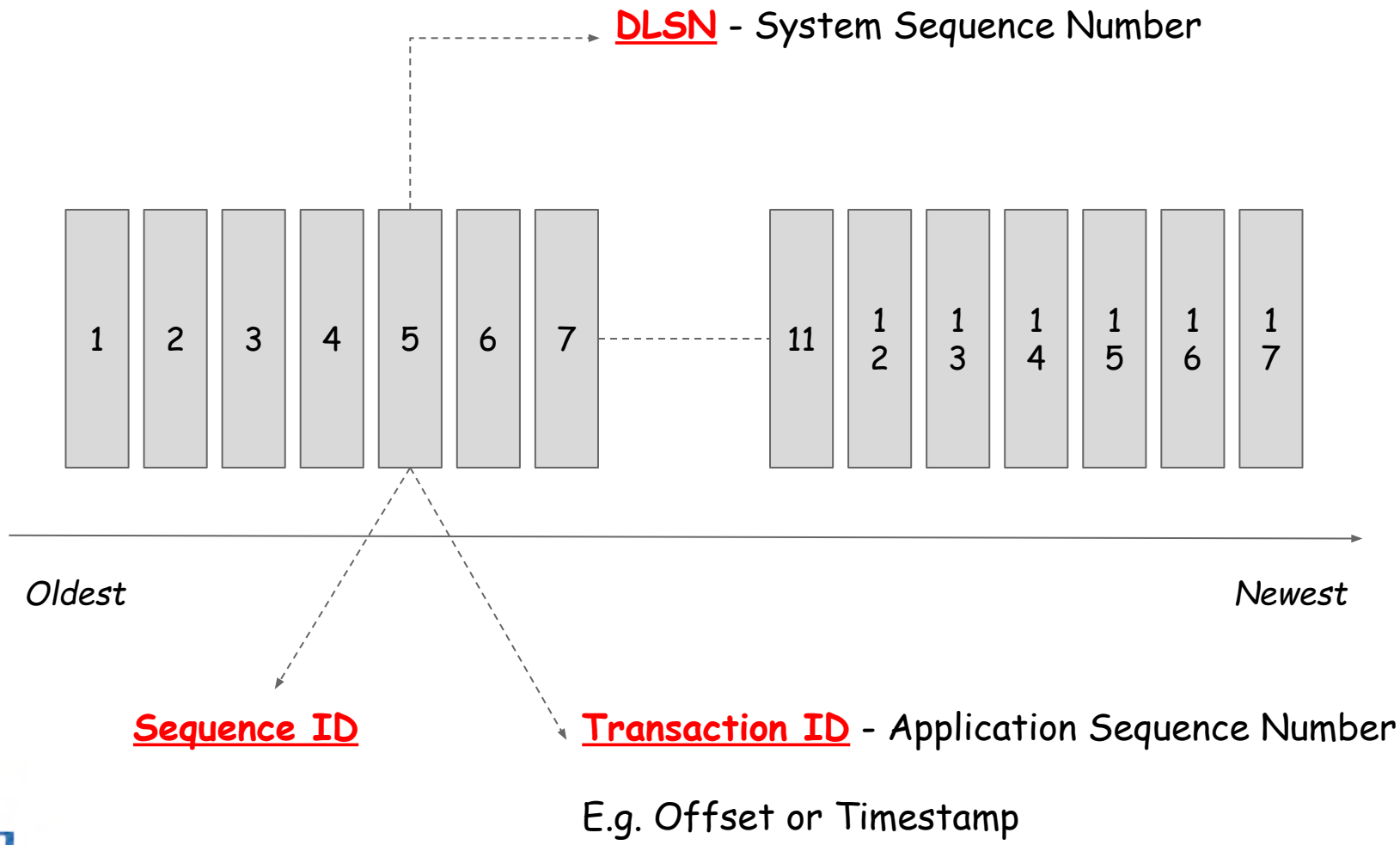
Sequence Numbers - DLSN



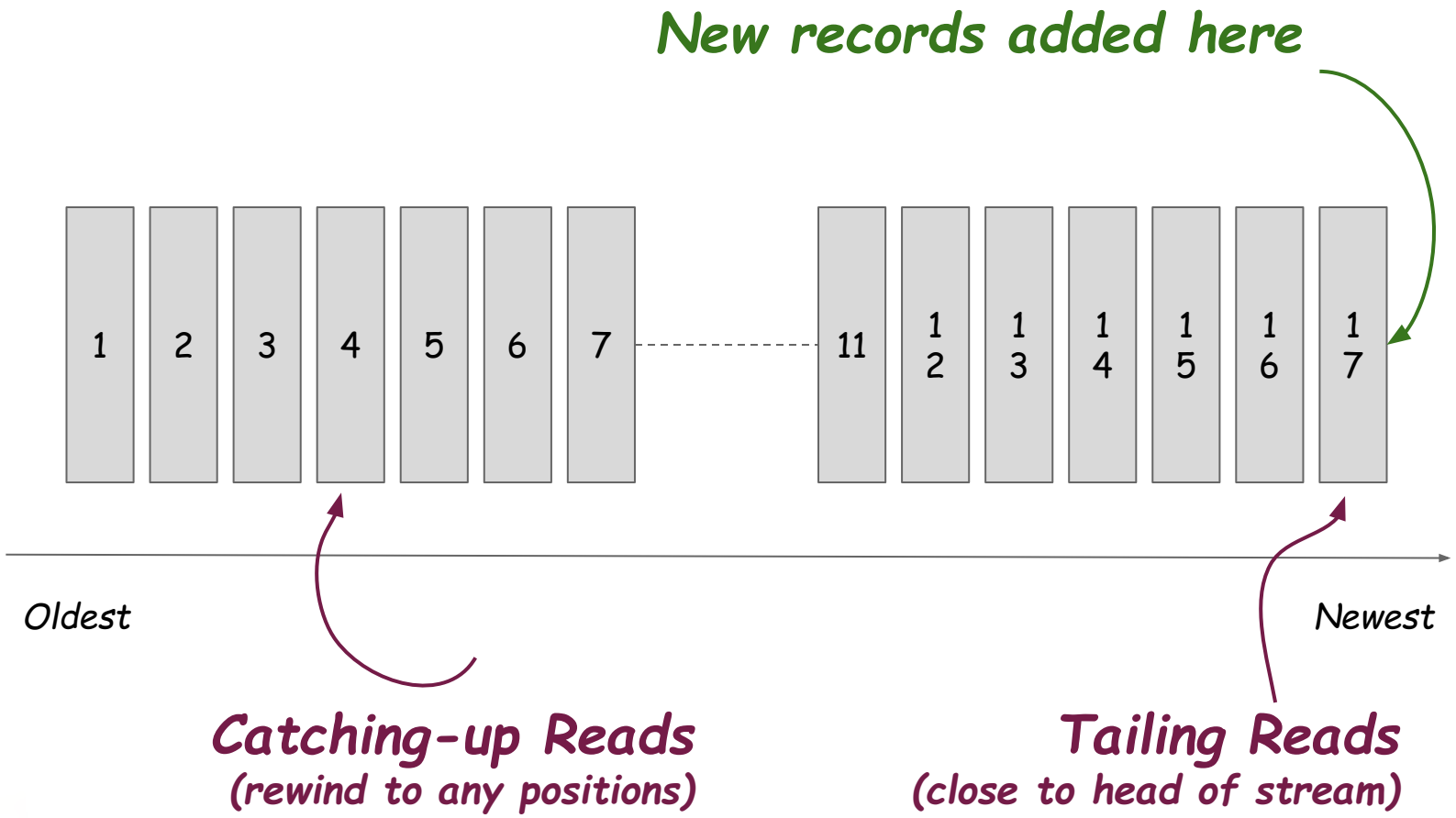
Sequence Numbers - Transaction ID



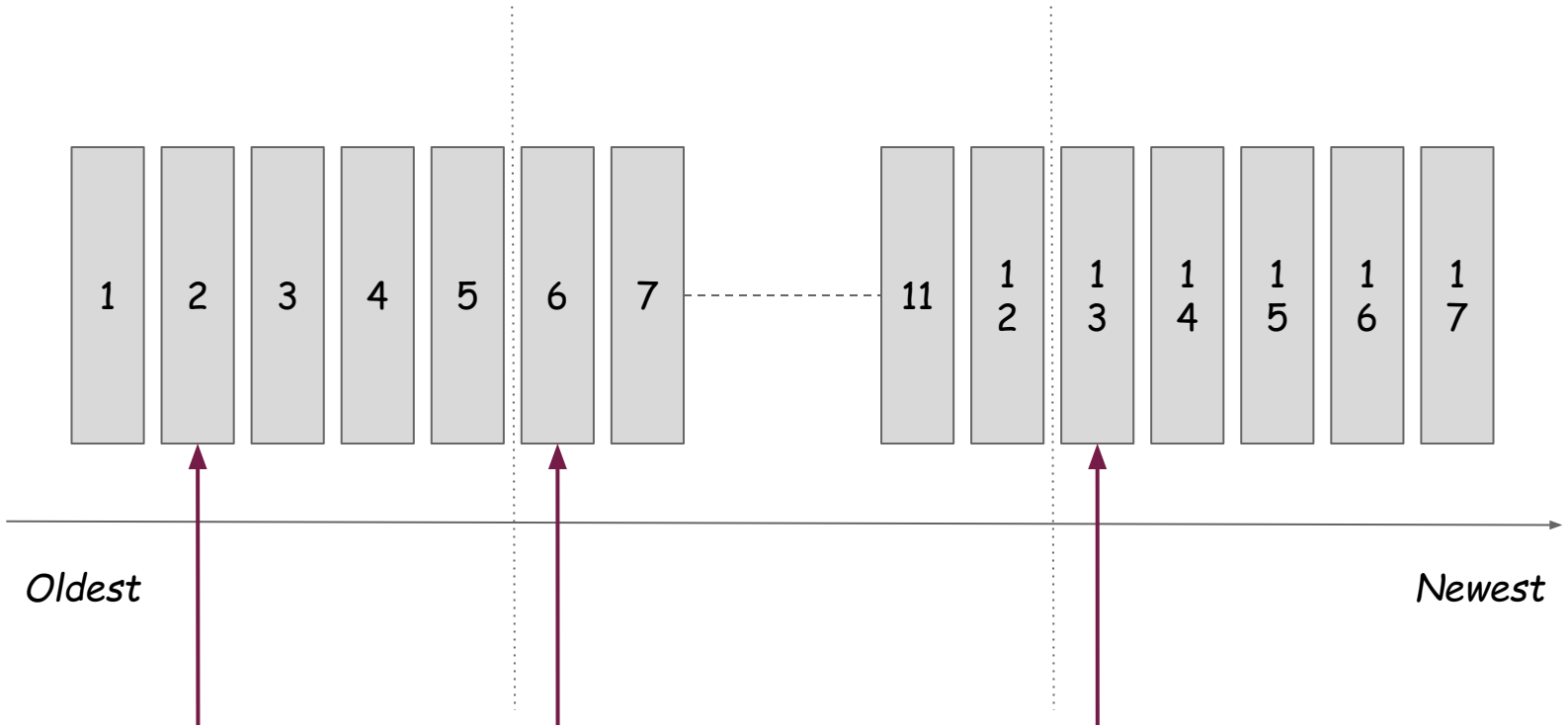
Sequence Numbers - Sequence ID



Writer & Readers

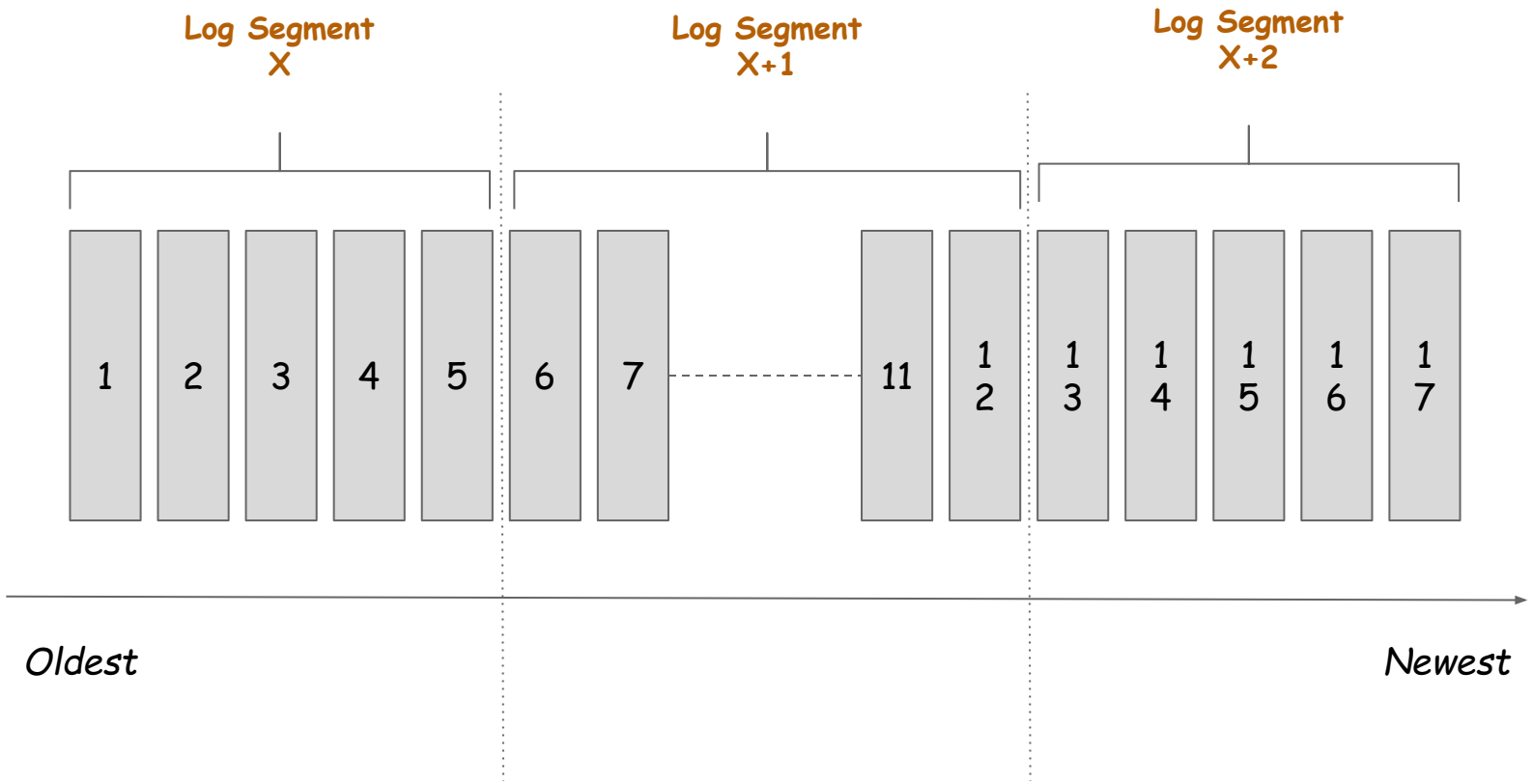


Read Parallelism

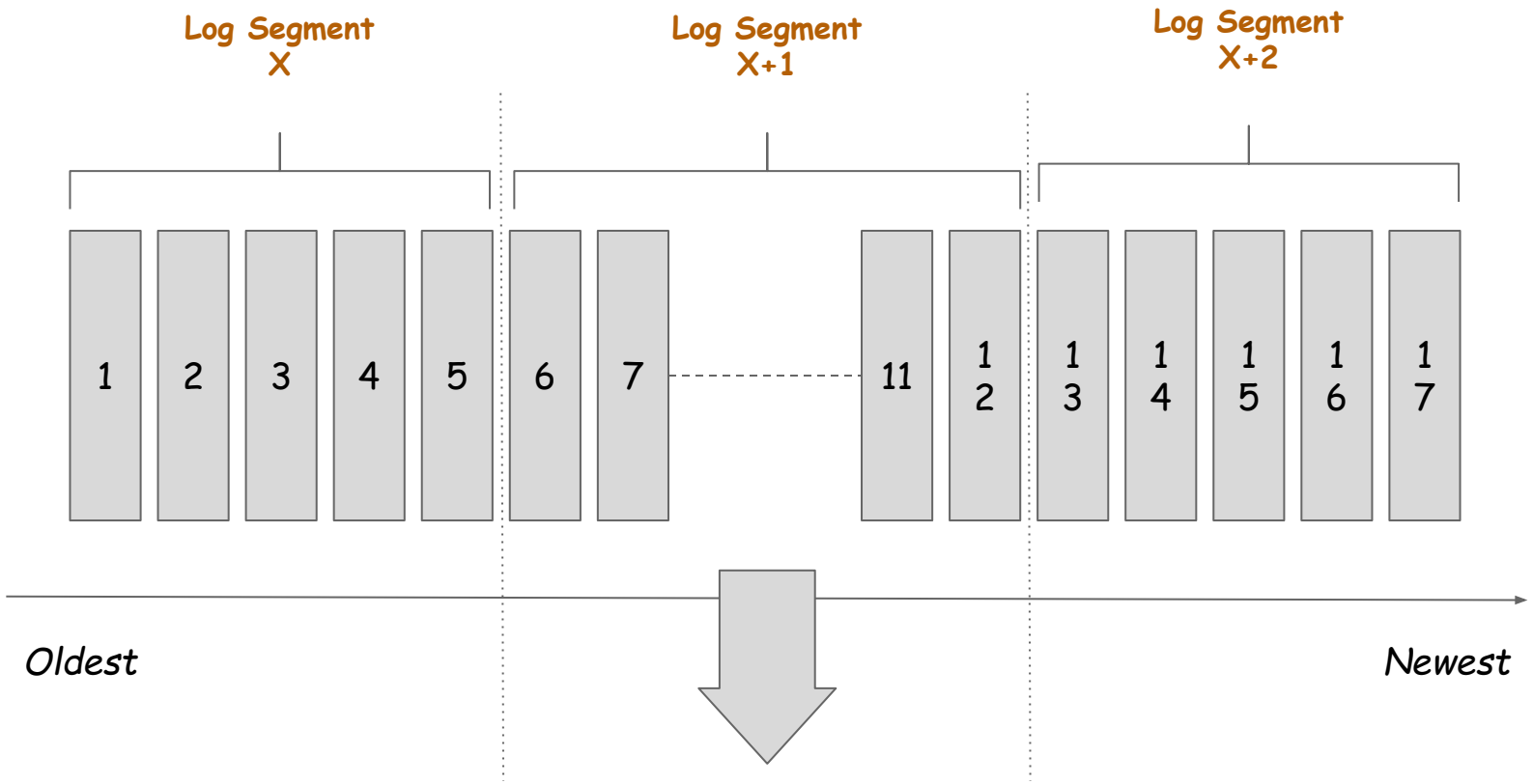


Read from multiple positions in parallel

Log Segments

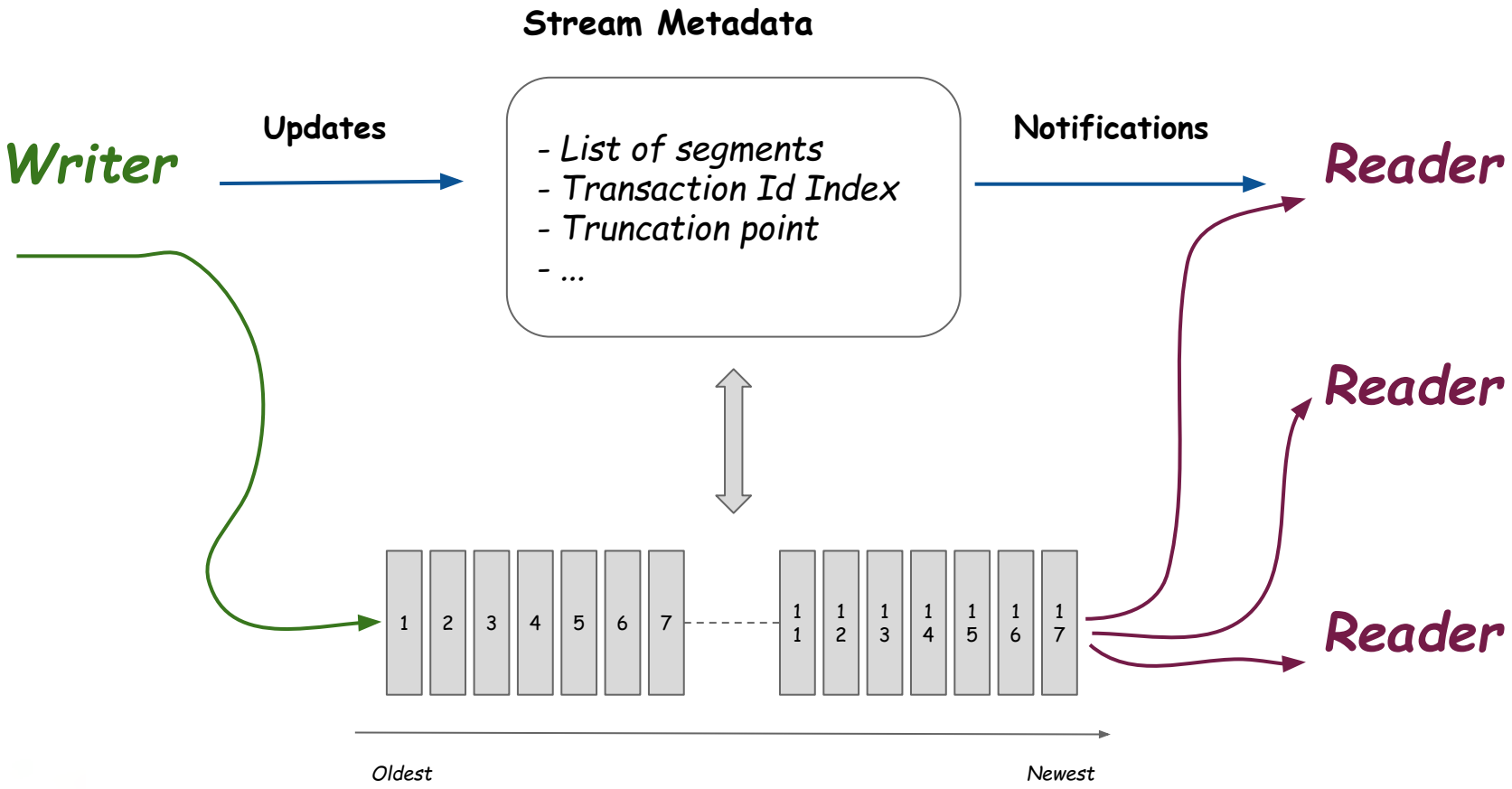


Log Segment Store

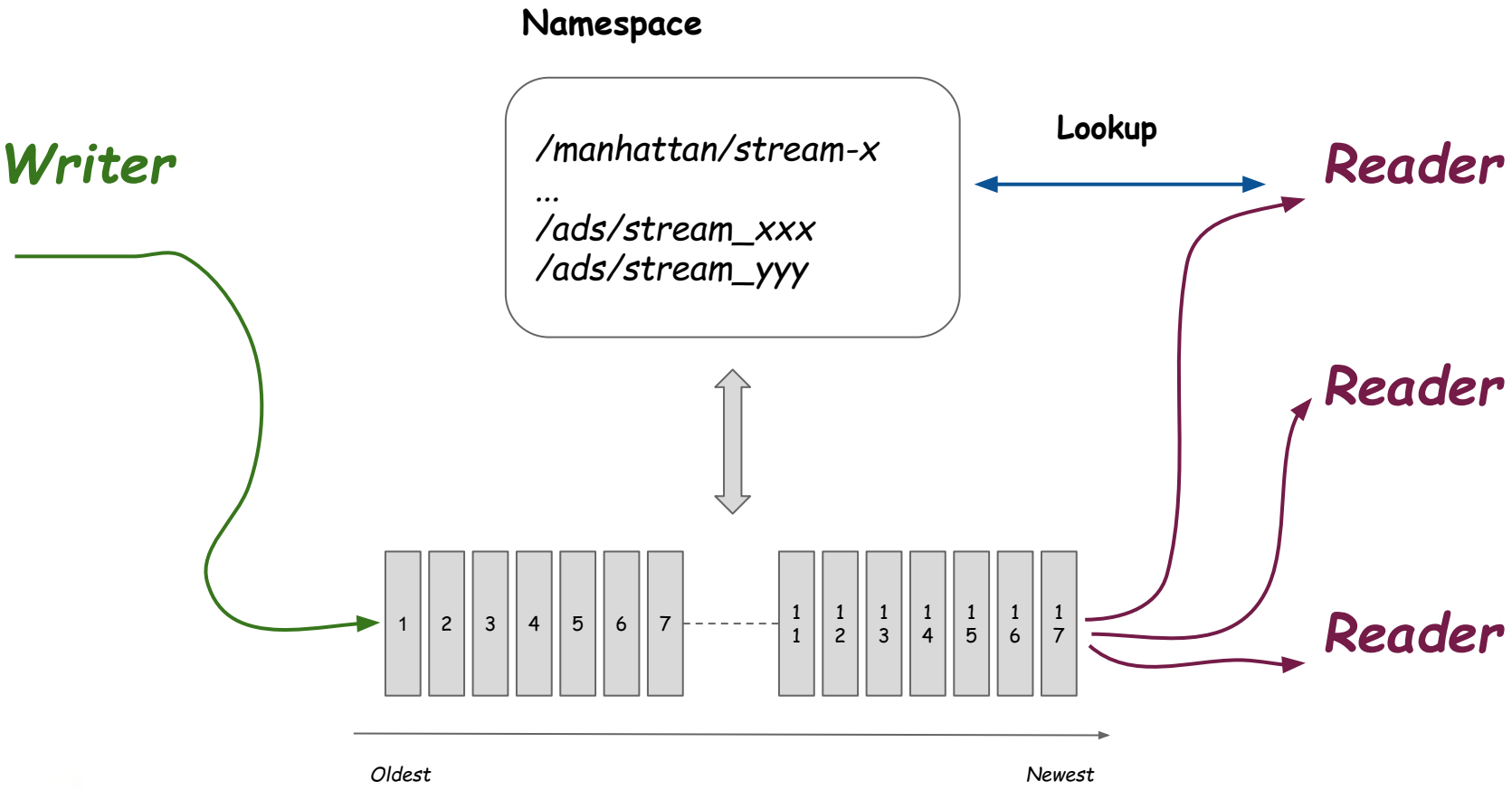


Apache BookKeeper

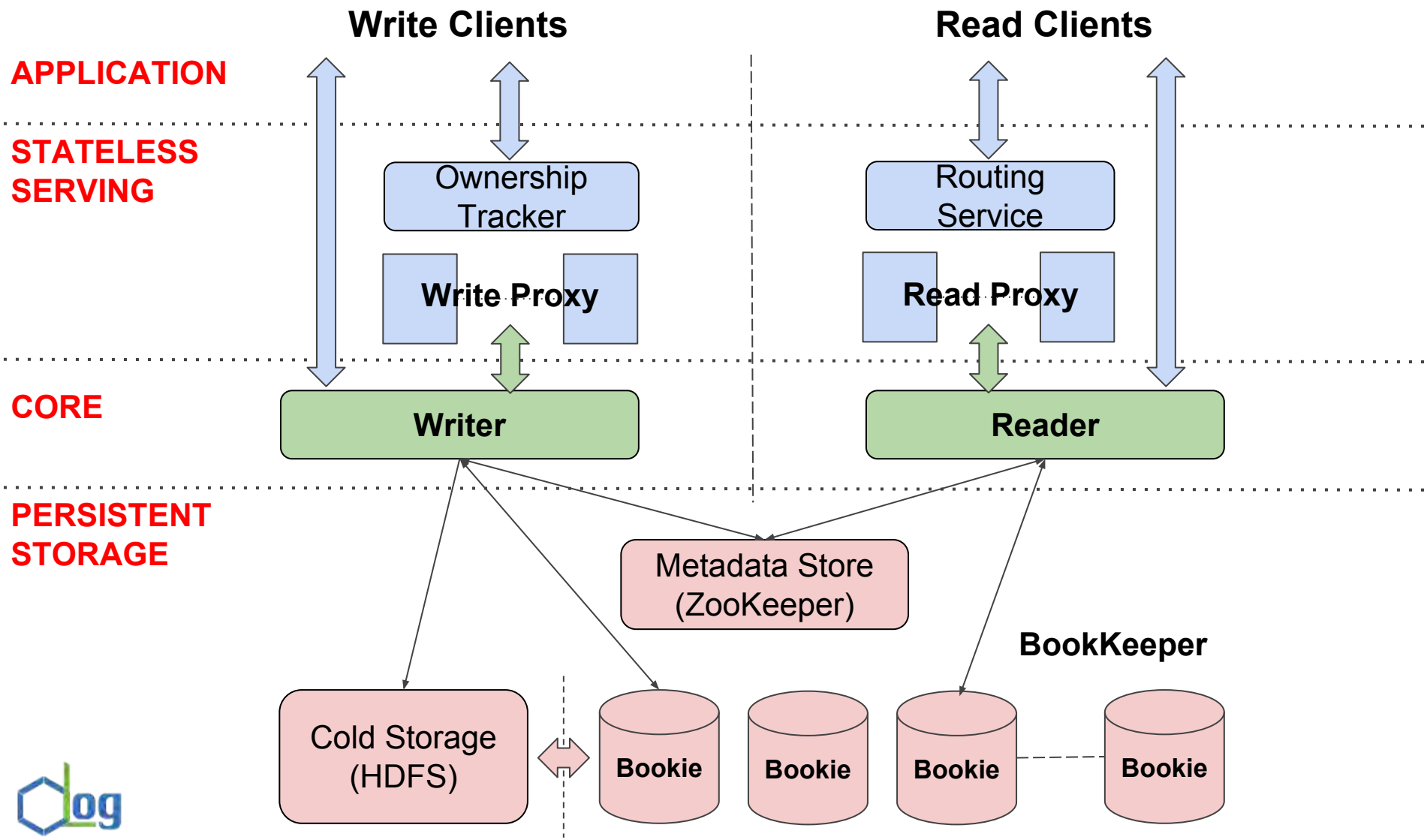
Log Stream Metadata



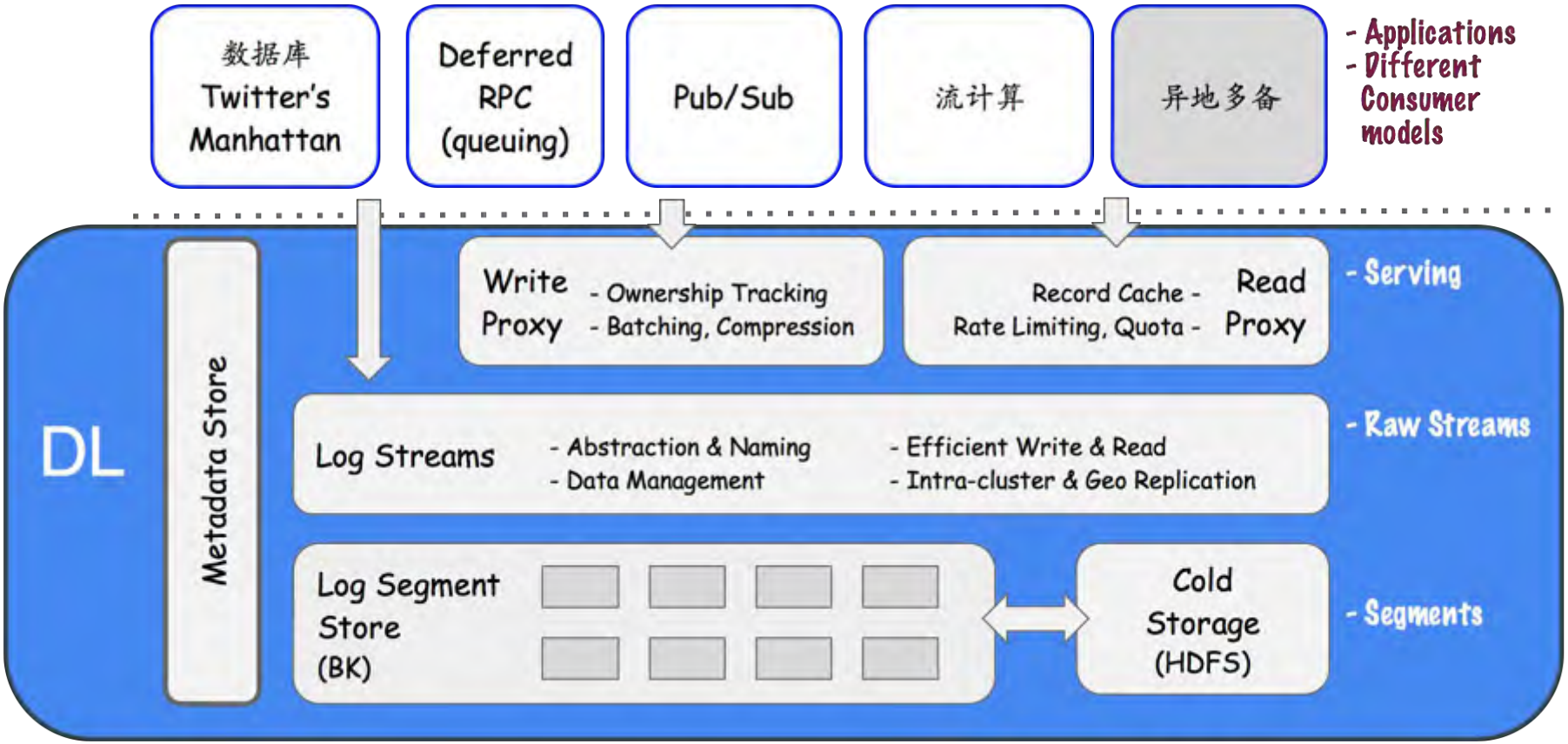
Namespace


Namespace
/manhattan/stream-x
...
/ads/stream_xxx
/ads/stream_yyy
Lookup
Writer
Reader
Reader
Reader
1
2
3
4
5
6
7
11
12
13
14
15
16
17
Oldest
Newest

Architecture

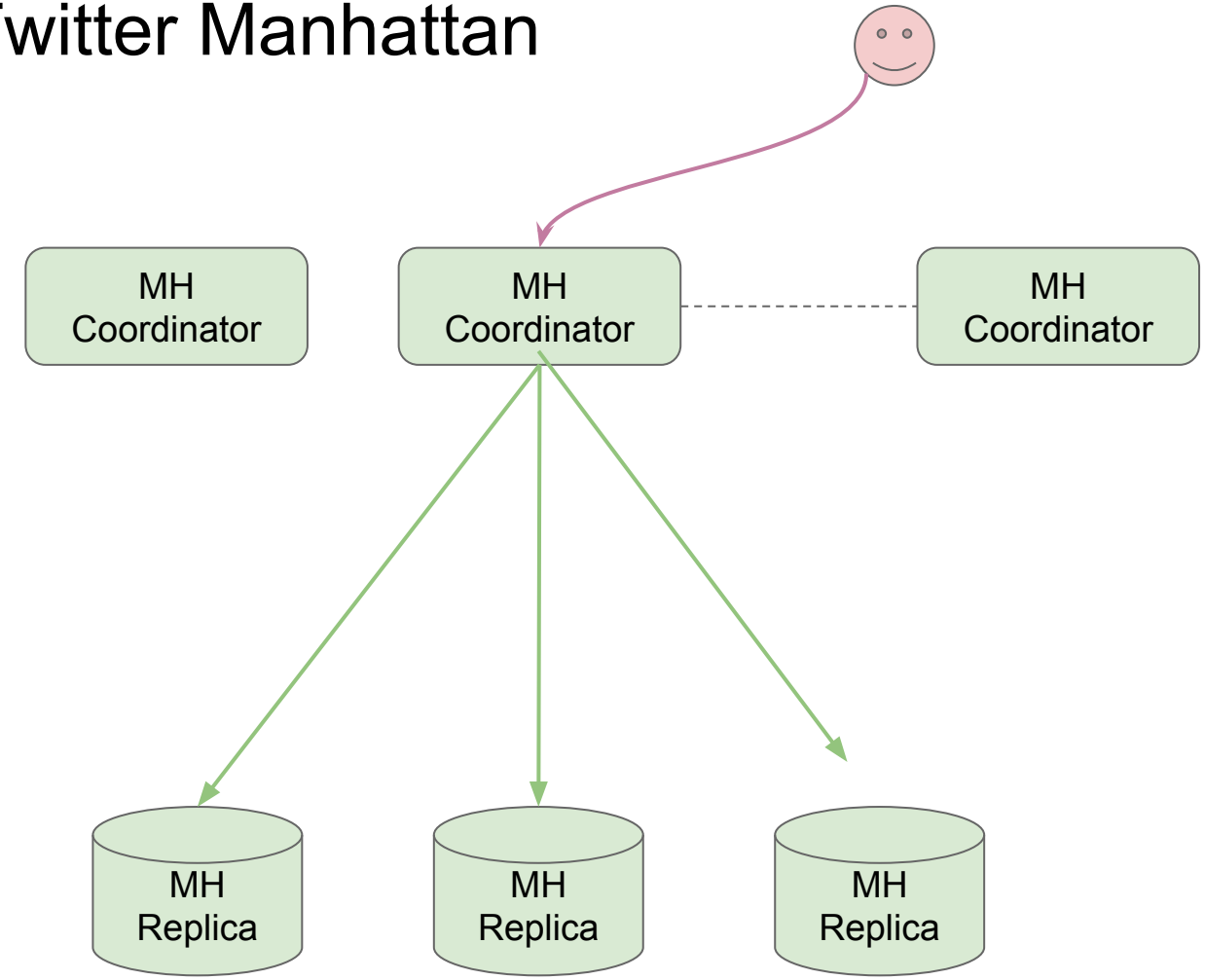


Software Stack

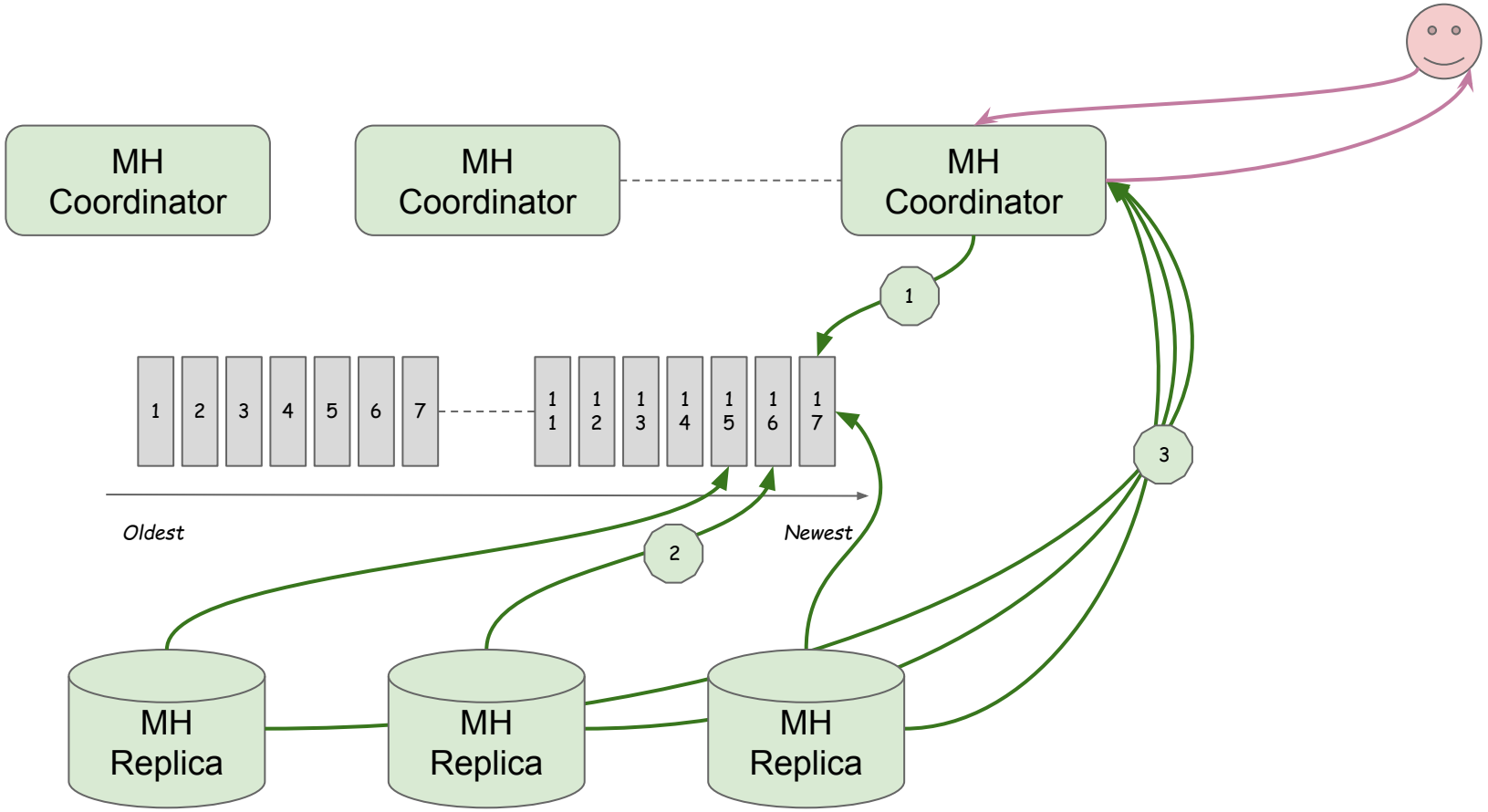


Use Case
-
Database

Twitter Manhattan



Stronger Consistency in Manhattan

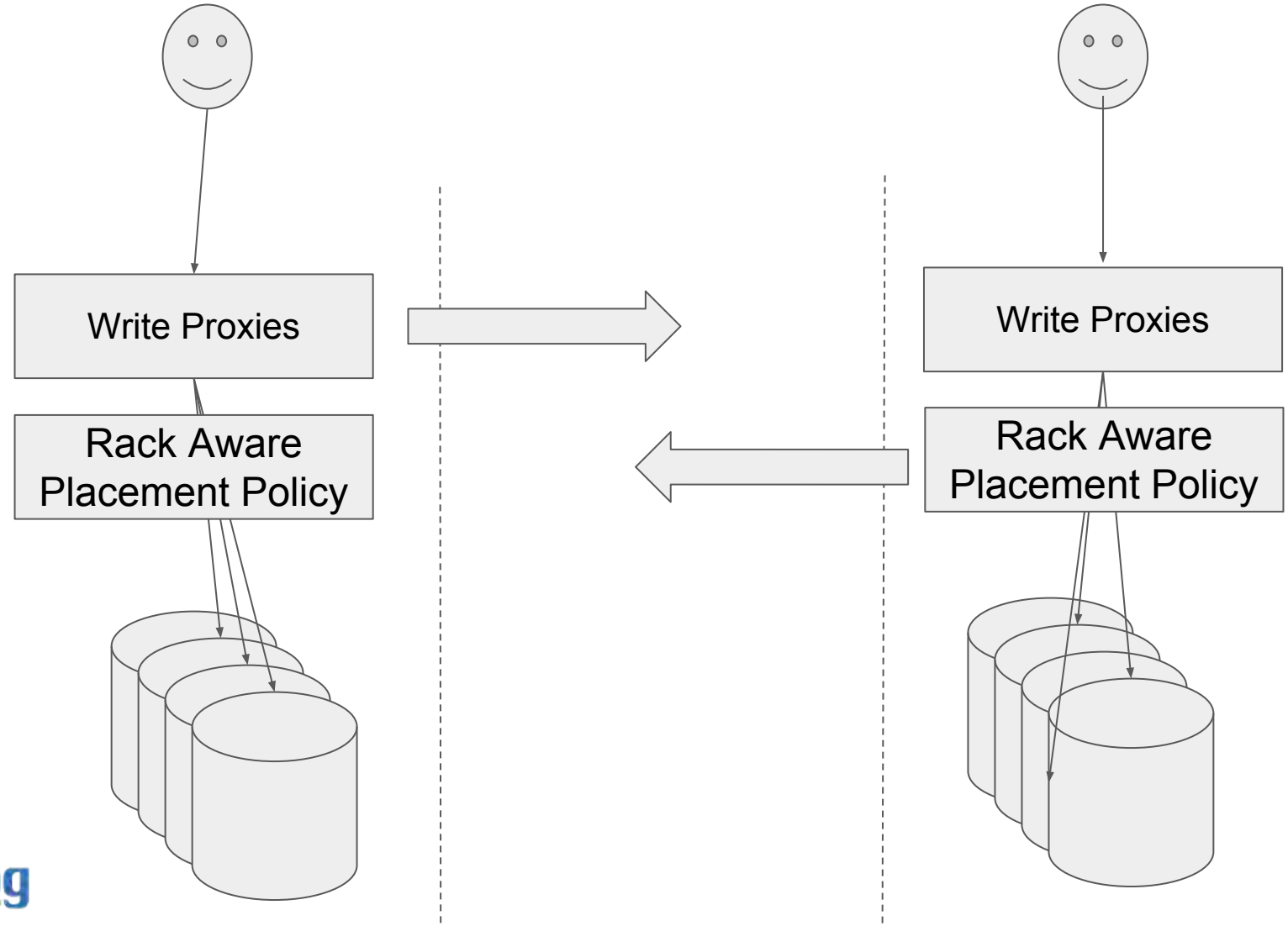


Use Case
-
Cross Datacenter Replication

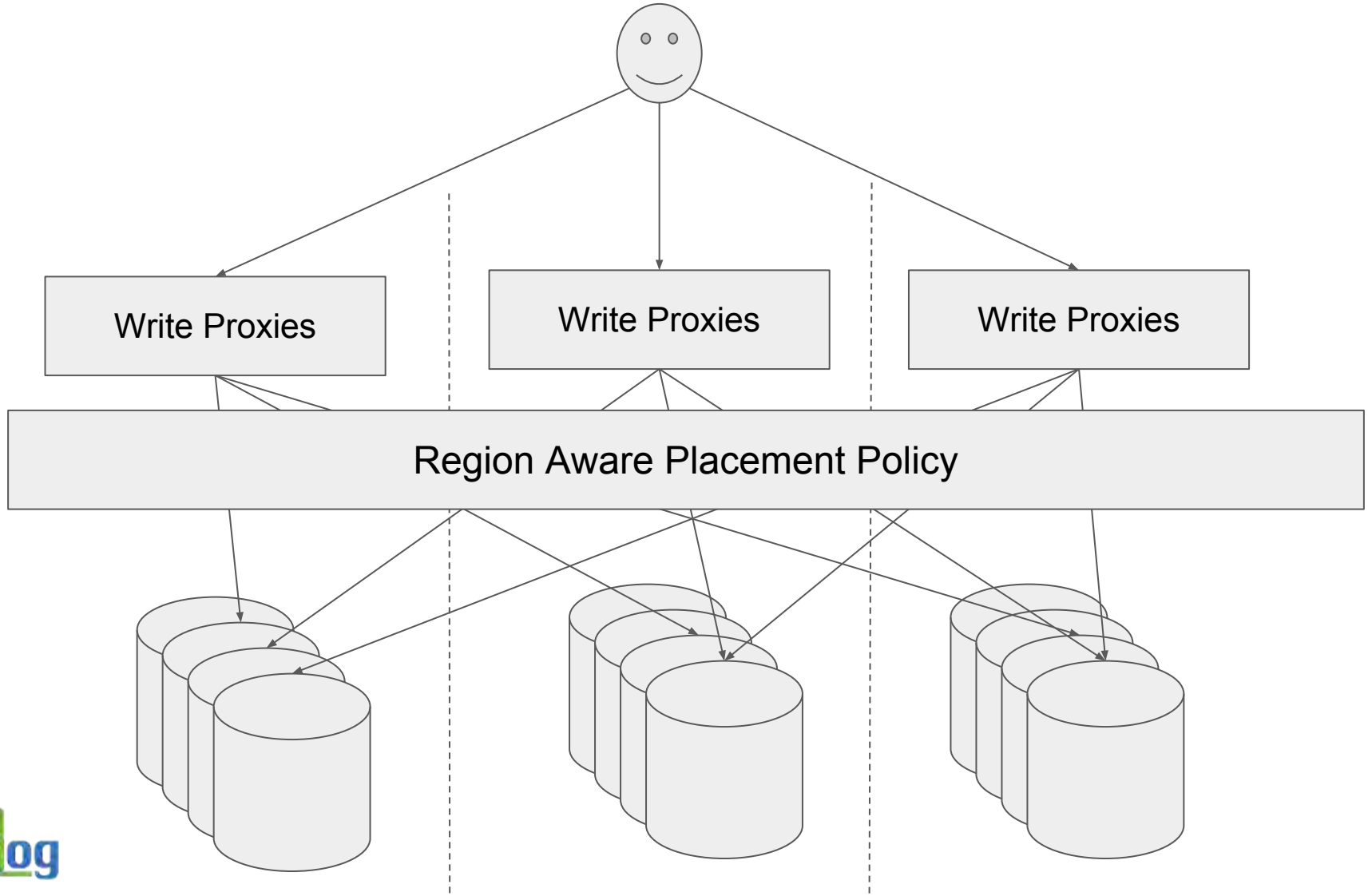
Cross Datacenter Replication

- Synchronous Replication
- Asynchronous Replication

Asynchronous Replication



Synchronous Replication



Summary

- Low latency and High performance
- Durable and Consistent
- Intra-cluster and geo replication
- Flexible replication use cases

Resources

- distributedlog.io
- https://github.com/apache/incubator-distributedlog
- ICDE 2017 Paper - “DistributedLog: A high performance replicated log service”
- Follow us @distributedlog

Thank you

- @sijieg
- Wechat: guosijie_