PolarDB Serverless: A Cloud Native Database for Disaggregated Data Centers

Wei Cao, Yingqiang Zhang, Xinjun Yang, Feifei Li, Sheng Wang, Qingda Hu, Xuntao Cheng, Zongzhi Chen, Zhenjun Liu, Jing Fang, Bo Wang, Yuhui Wang, Haiqing Sun, Ze Yang, Zhushi Cheng, Sen Chen, Jian Wu, Wei Hu, Jianwei Zhao, Yusong Gao, Songlu Cai, Yunyang Zhang, Jiawang Tong

WWW.ALIYUN.COM

Outline

The current challenges of cloud database

The design of PolarDB Serverless and optimization

Benchmark Result

The challenge of current cloud database

➢ Different resources coupled

  - resource limited by physical machine
  - independent resource provisioning
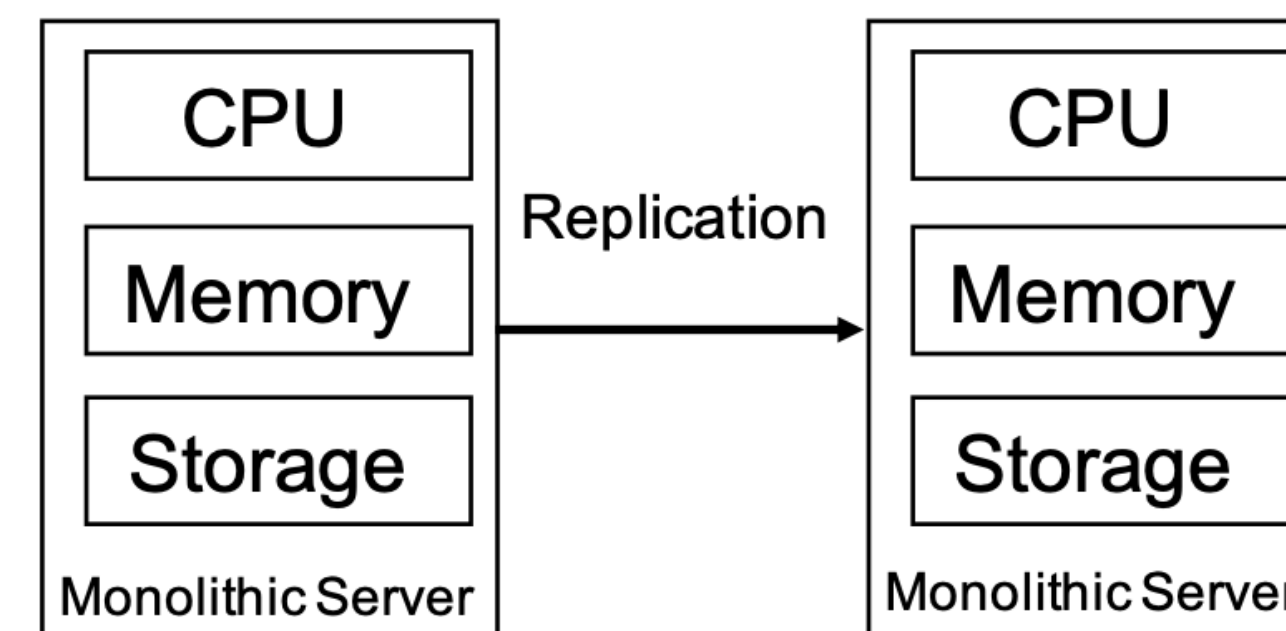 - fate sharing

➢ Redundant in-memory data copy

Figure 1: *monolithic machine*

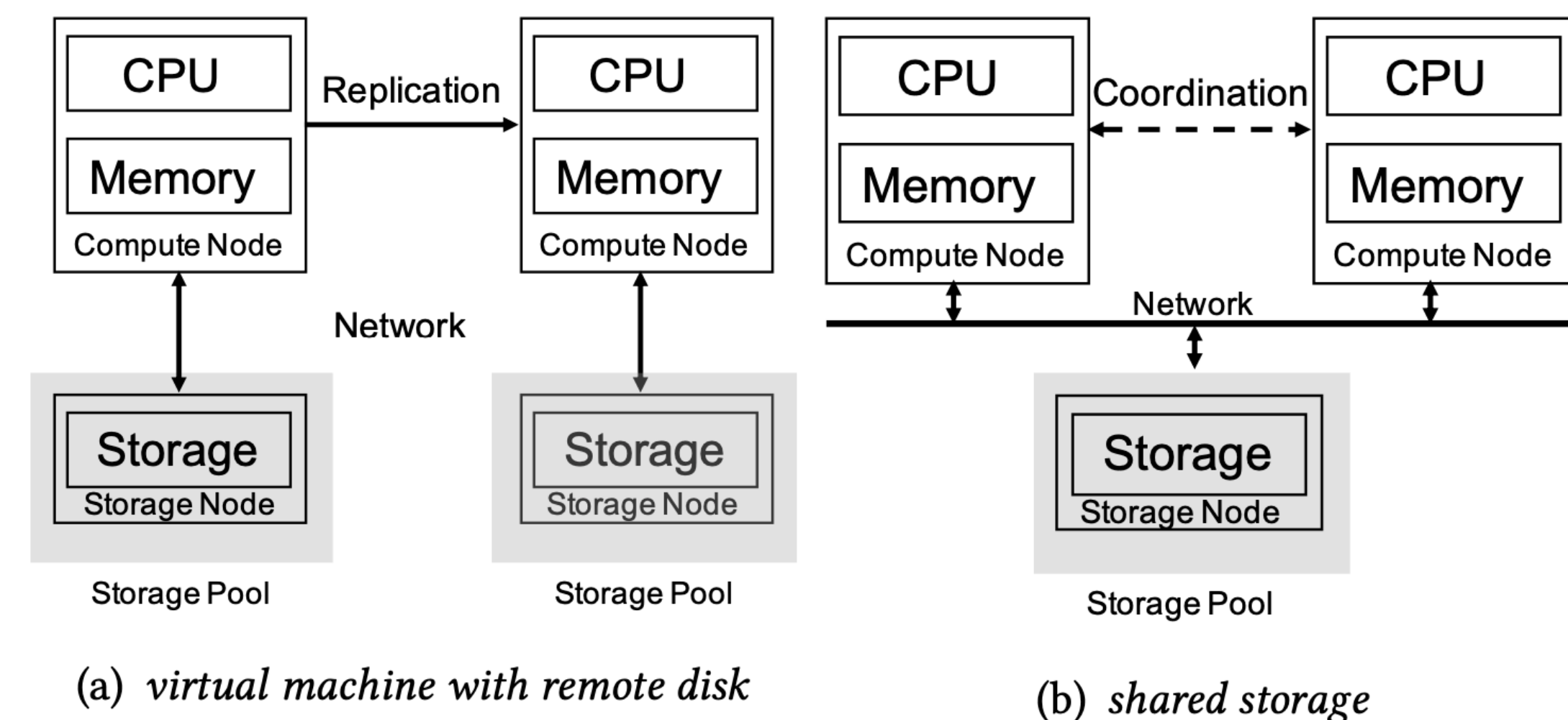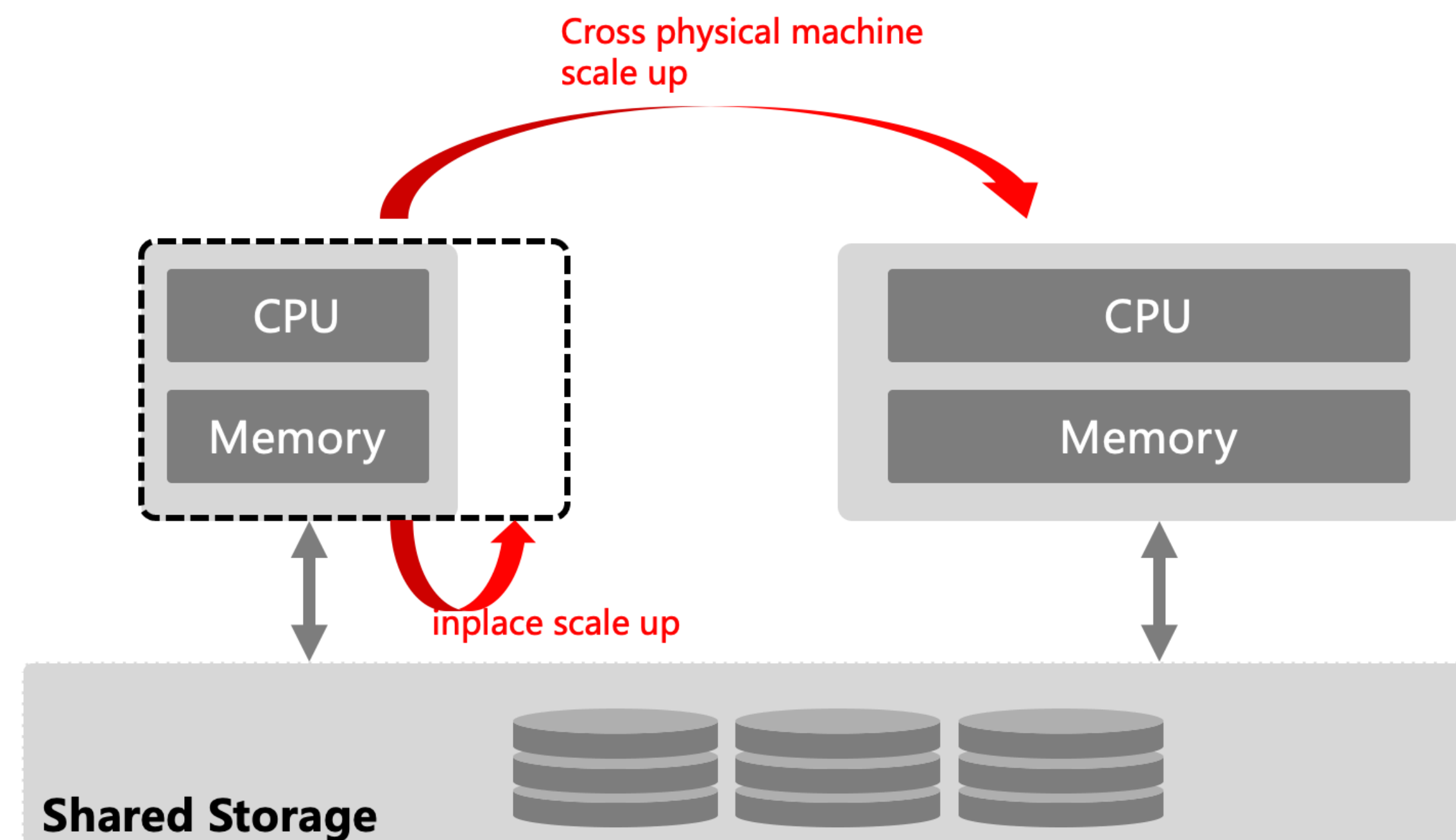(a) *virtual machine with remote disk*

(b) *shared storage*

Figure 2: *separation of compute and storage*

The challenge of current serverless database service

➢ CPU and memory scale simultaneously

➢ Long resumption time after auto-pause

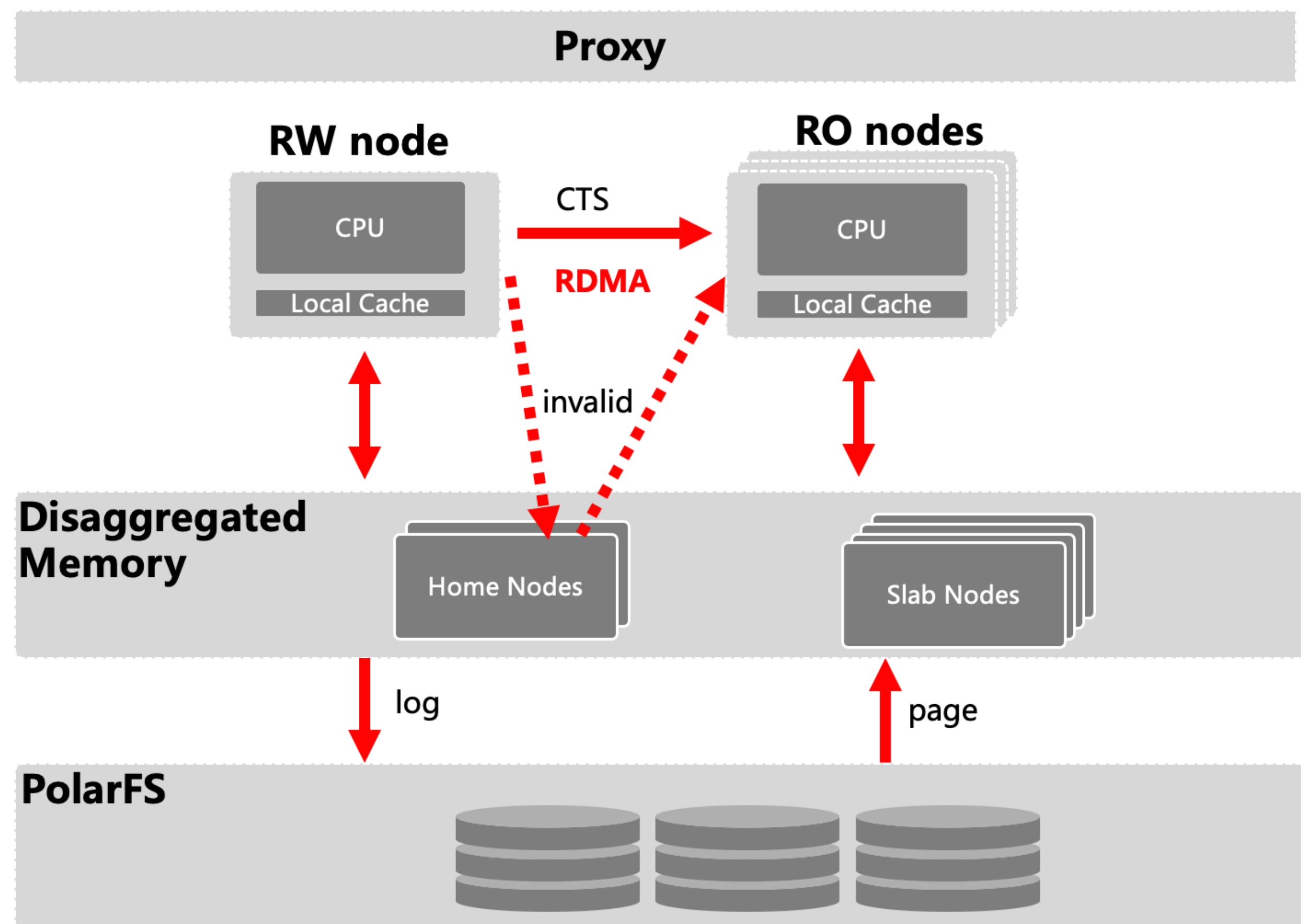➢ Scaling transparency

➢ Speed of cross-node Scaling

Outline

The current challenges of cloud database

The design of PolarDB Serverless and optimization

Benchmark Result

PolarDB Serverless architecture

- ✓ Fully disaggregation architecture

- ✓ Horizontal scaling disaggregated memory

- ✓ Multiple nodes share same data copy

- ✓ Cross nodes cache coherency

PolarDB Serverless —— Disaggregated Memory

✓ Consist of local caches, home nodes and slab nodes

✓ Support cache coherency between RW/RO node`s local caches coordinated by slab node and local caches
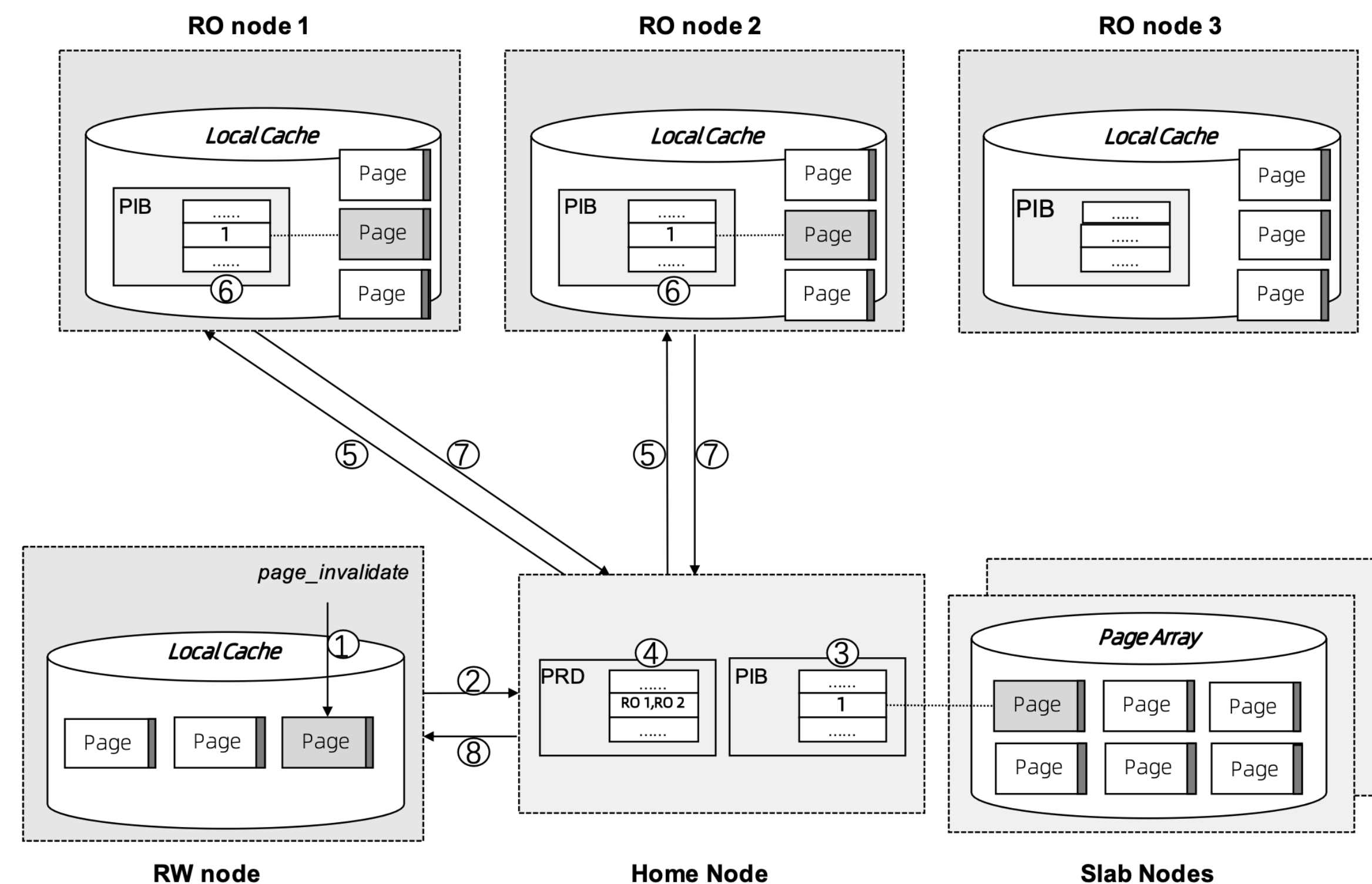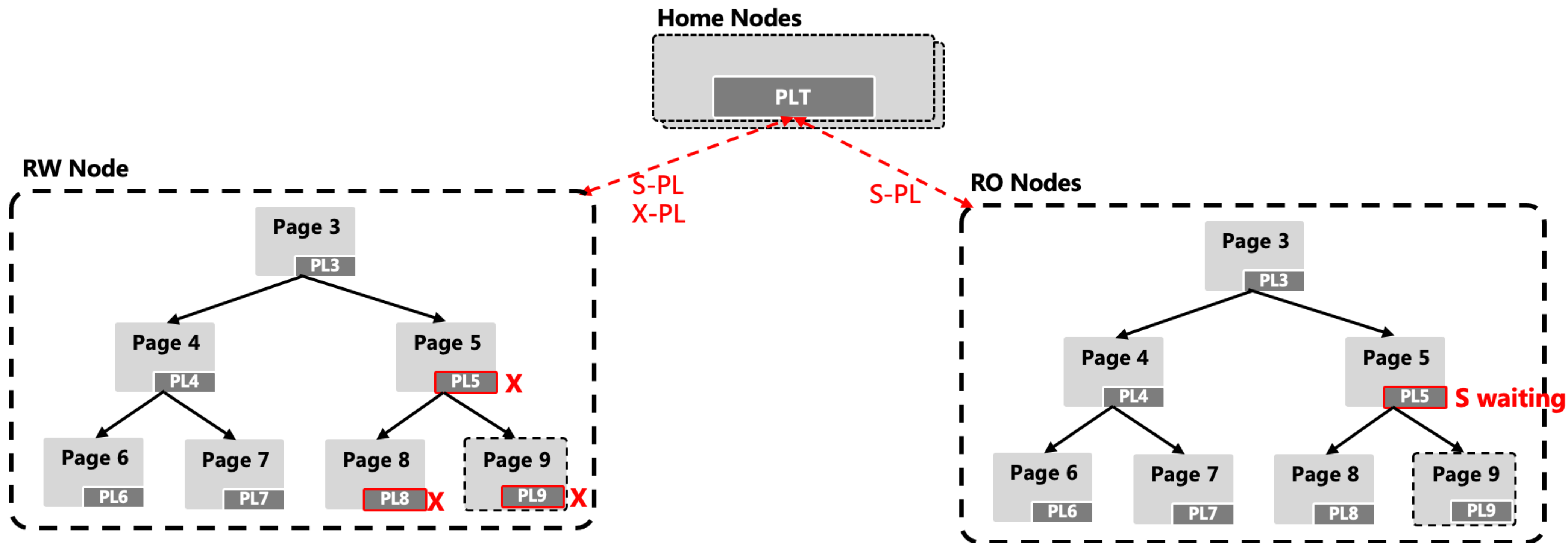
An example of cache invalidation

1. Set RW local cache PIB

2. Send invalidate instrument

3. Set home node PIB

4. Get RO info from PRD

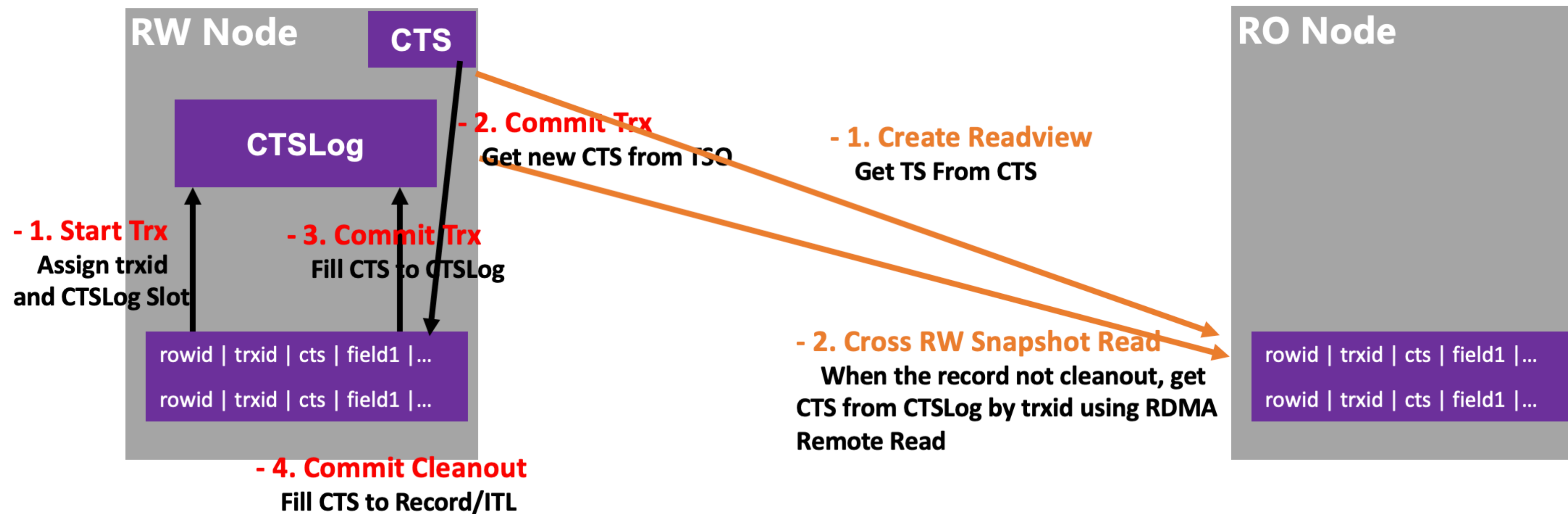56. Remote set RO local cache PIB

78. Return result
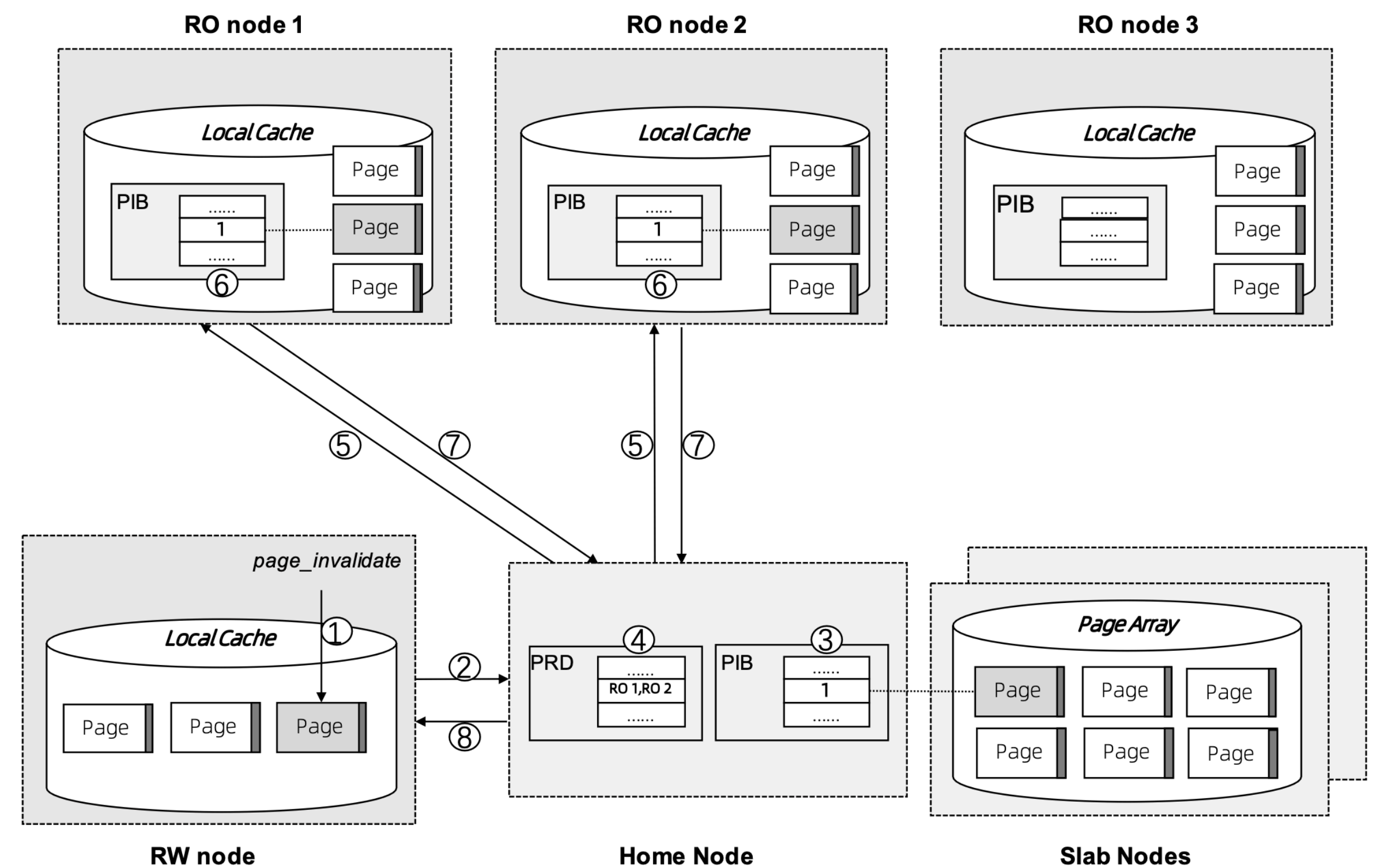
Cross-node B+Tree consistency

- ✓ Global physical latch
- ✓ Stickiness policy
- ✓ Fast path by RDMA CAS

Cross-node Snapshot Isolation with RDMA

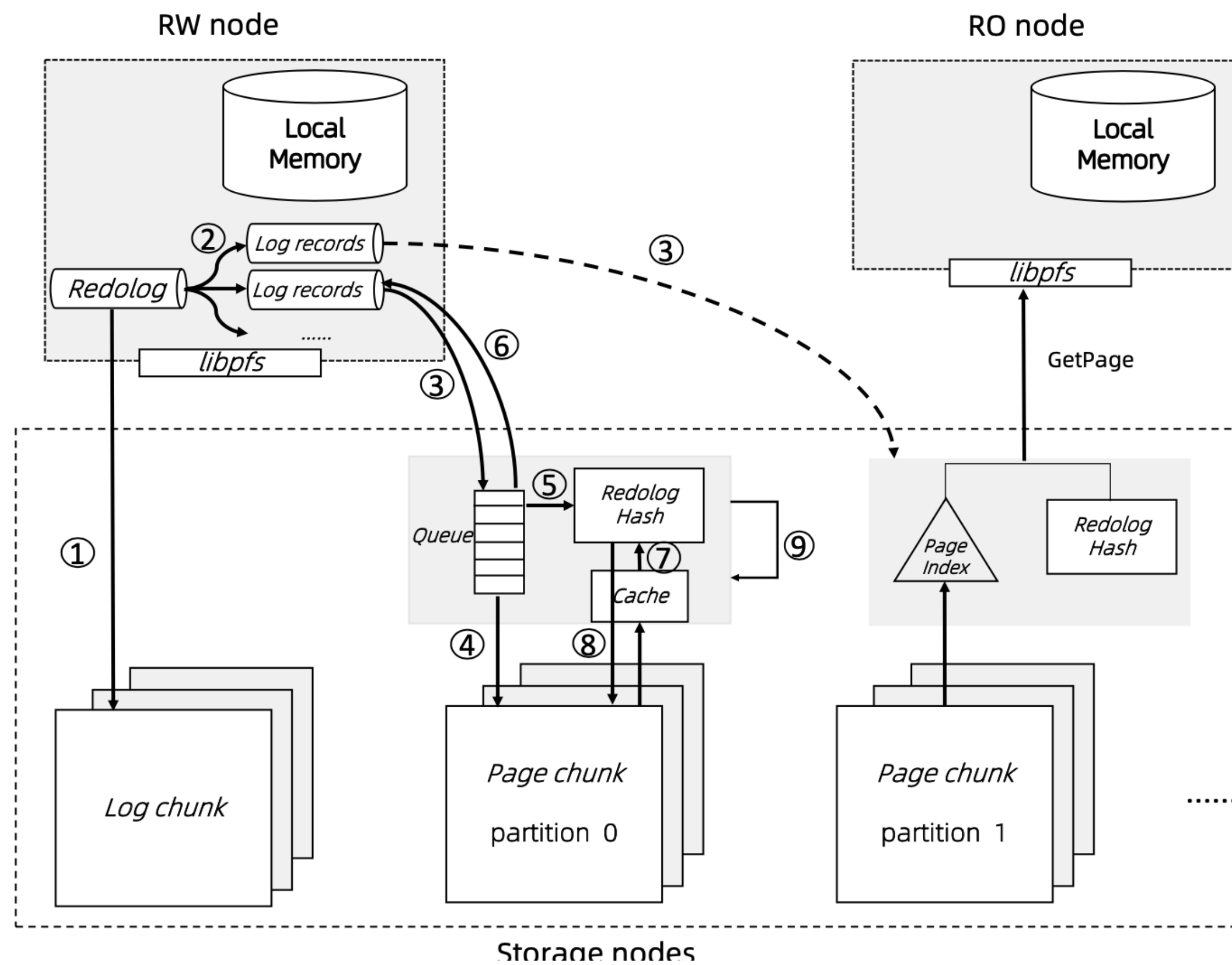# Page Materialization Offloading
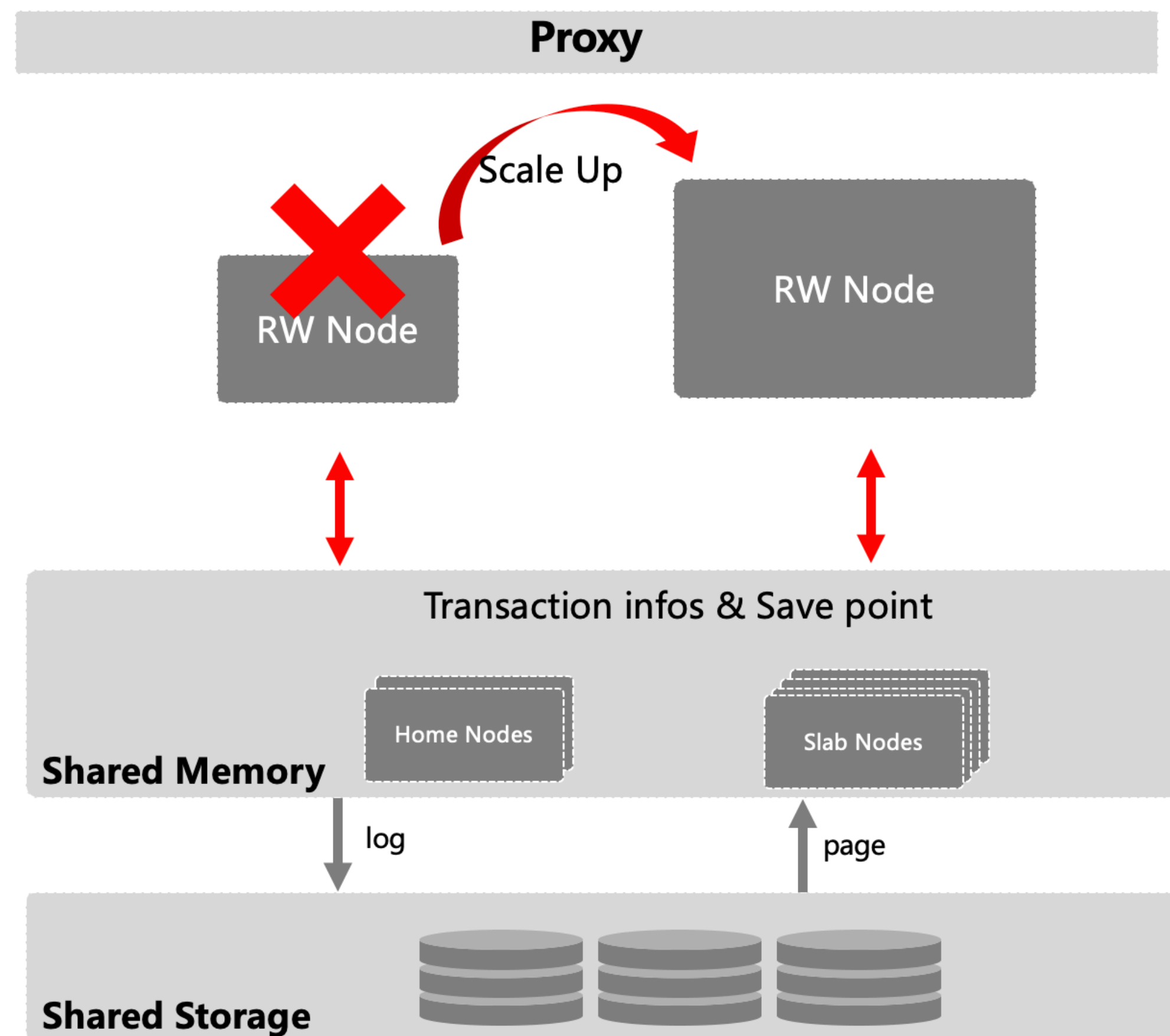
✓ log is database

Figure 7: Page Materialization Offloading

PolarDB Serverless —— Transparent Auto-Scaling

- ✓ Live transactions migration

- ✓ Pre-flush dirty pages

- ✓ Warm buffer pool

PolarDB Serverless —— Performance optimization

Optimistic Locking for Cross-node B+Tree Consistency

Index-Aware Prefetching

Outline

The current challenges of cloud database

The design of PolarDB Serverless and optimization
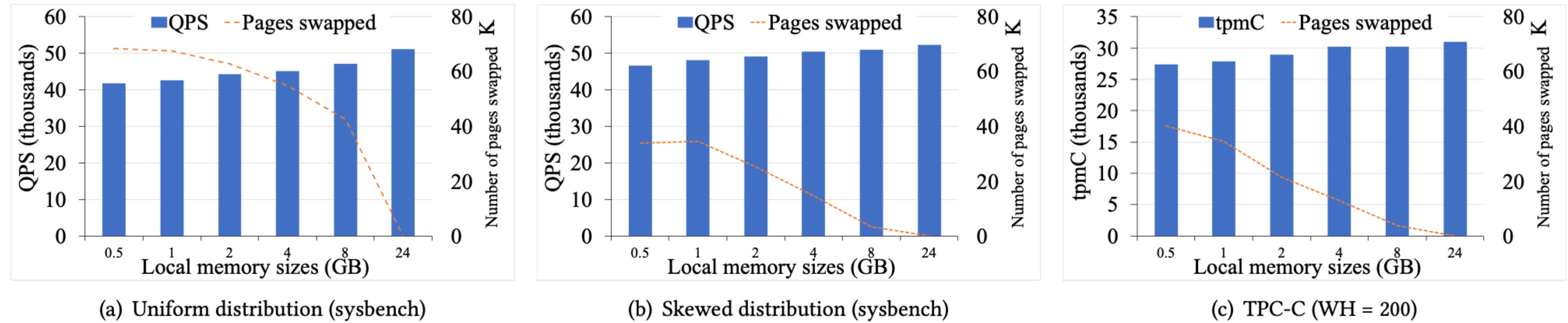
Benchmark Result

Sysbench & TPC-C & TPC-H

Figure 11: Performance of mixed reads and writes with varying local memory sizes.

- ✓ The performance losses in OLTP workload are at most 18.5%, 10.7% and 13.4%

- ✓ The performance losses in TPC-H is significant than OLTP workload
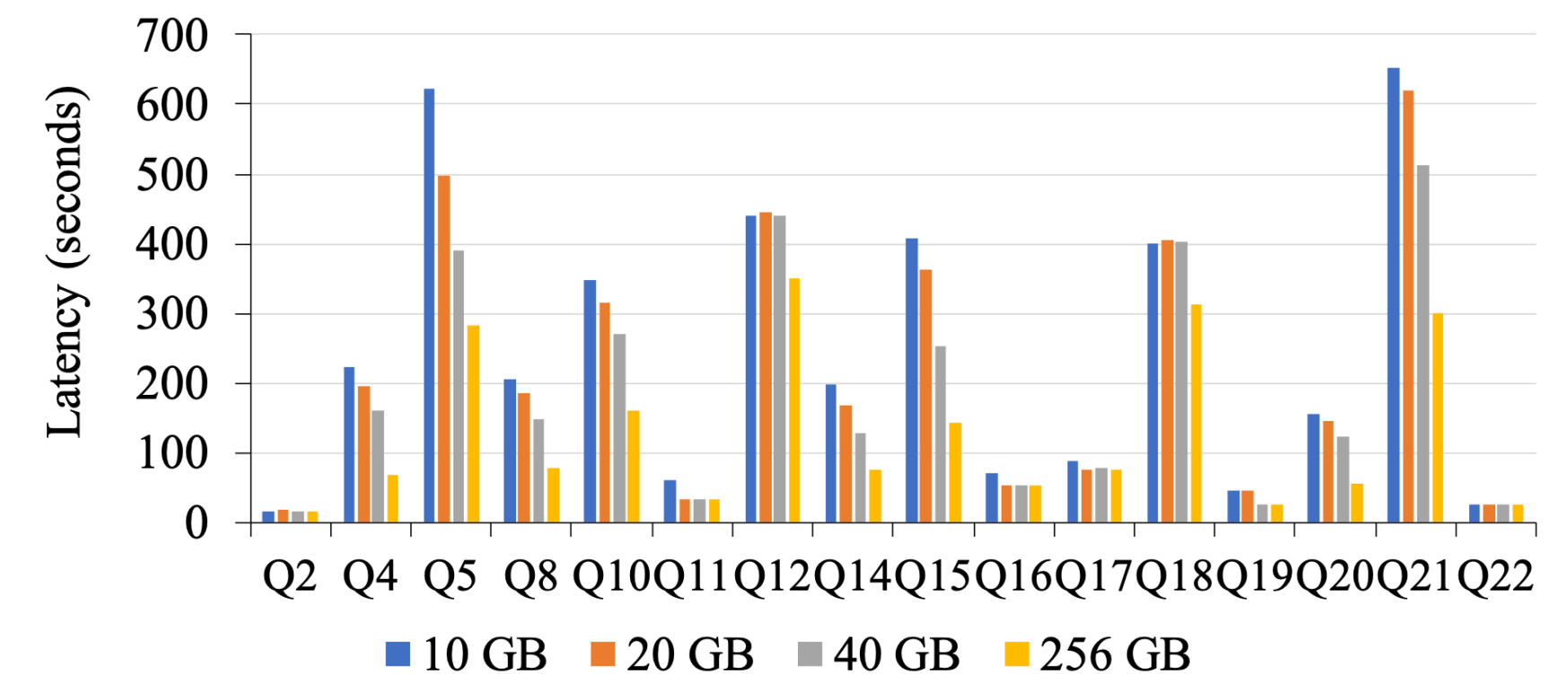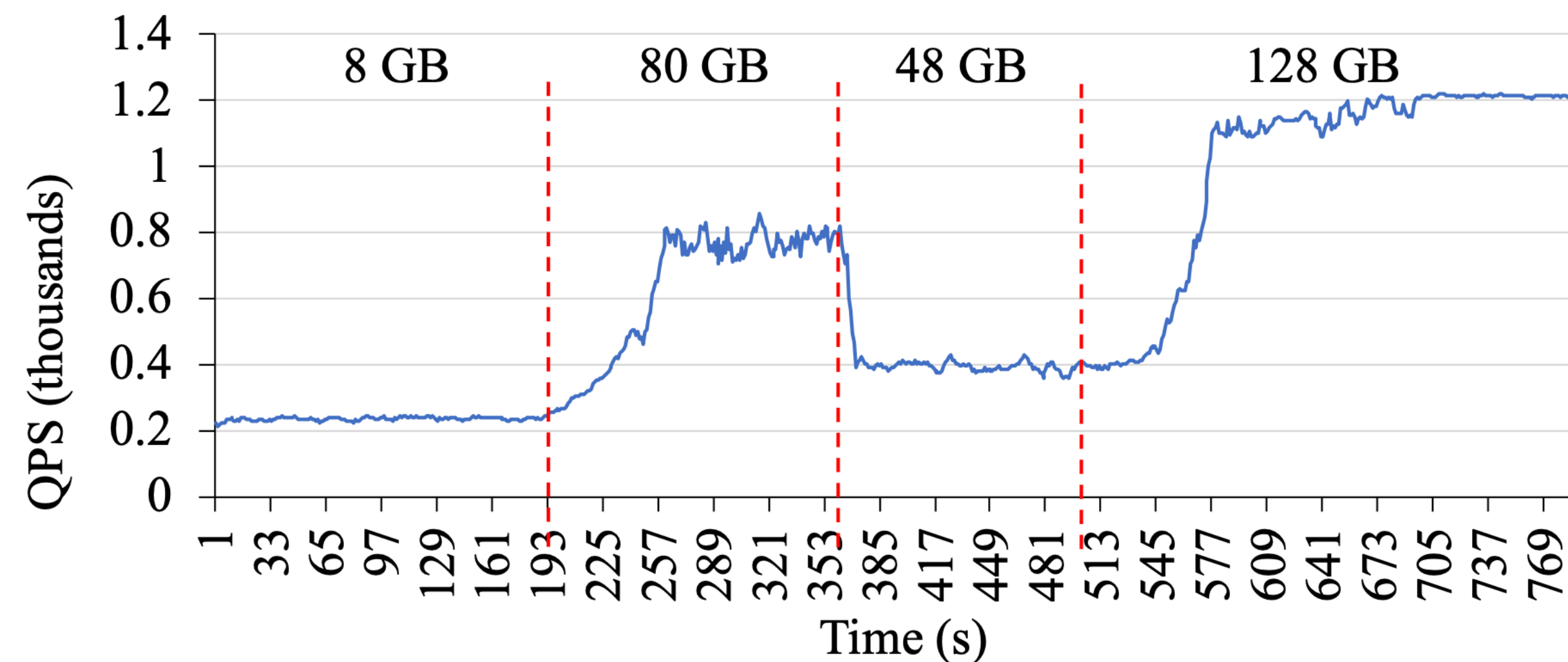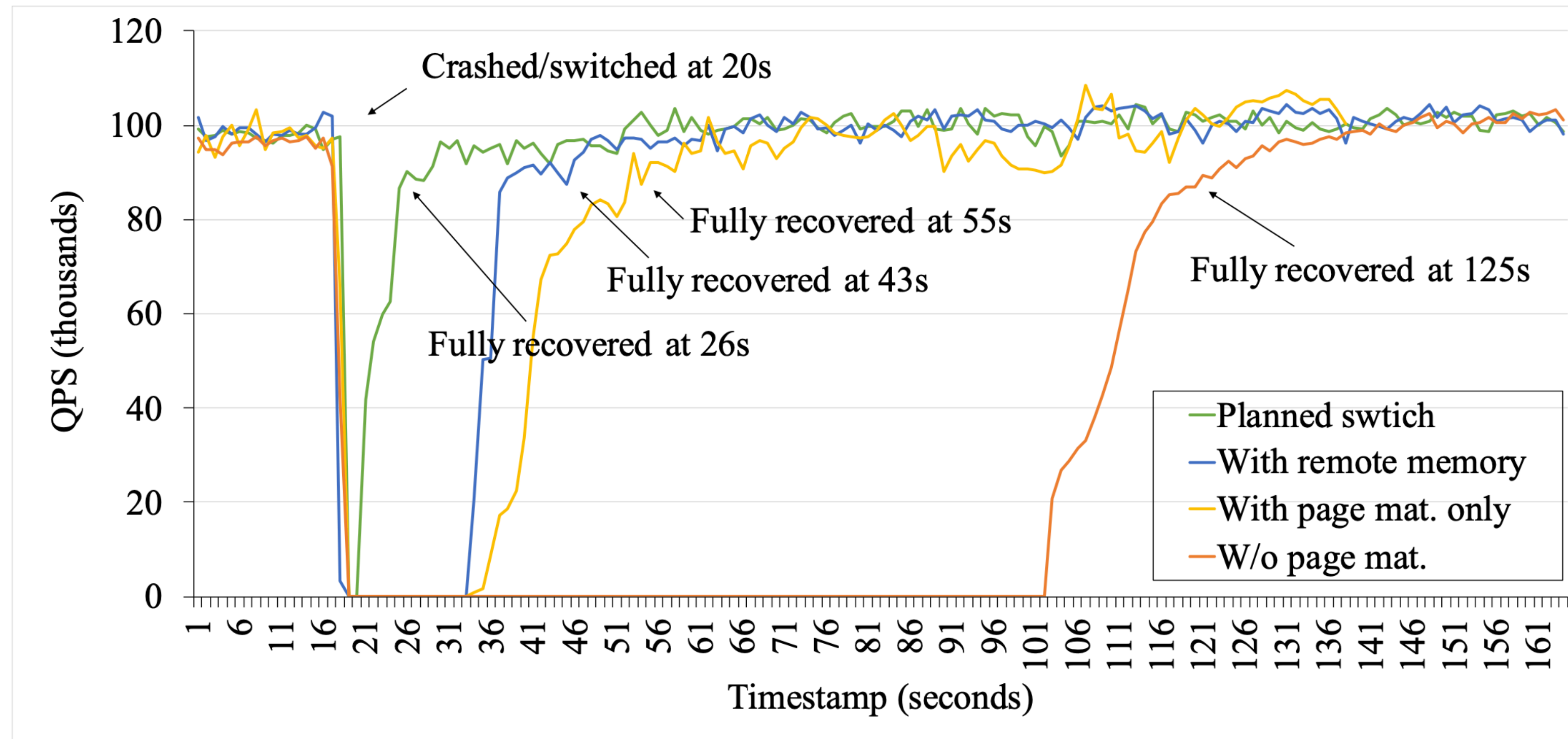

Figure 12: Latency of TPC-H queries (SF = 100) with varying capacities of the local cache (10 GB to 256 GB)

Memory on-demand provisioning

Throughput of PolarDB Serverless while scaling out-/in the remote memory (i.e., 8GB, 80GB, 48GB, 128GB) while processing range queries.

Figure 9: Recovery time for the RW with shared memory or local memory

Thanks

Yingqiang Zhang
yingqiang.zyq@alibaba-inc.com

阿里云 | 奥运会全球指定云服务商

WWW.ALIYUN.COM