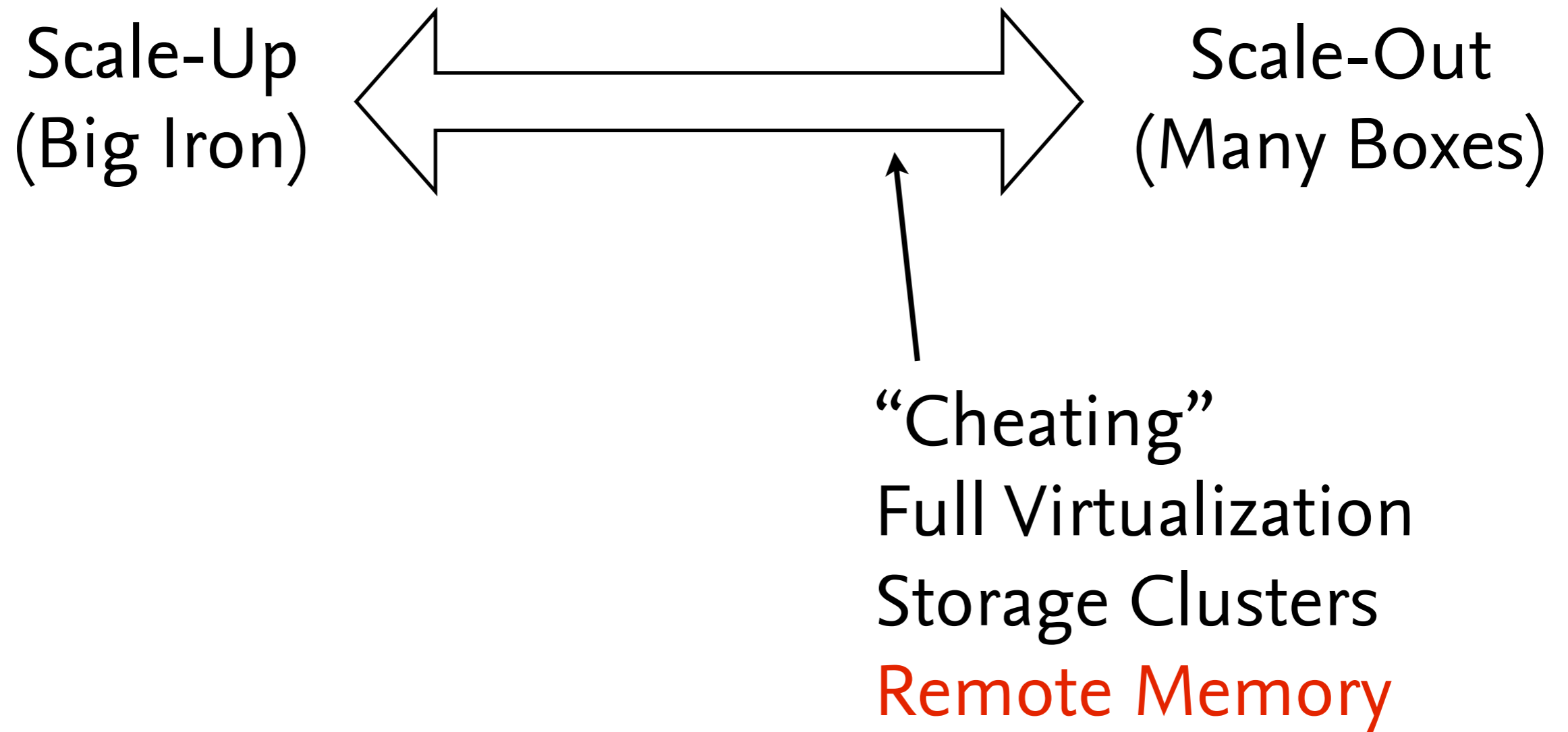


# Peak Performance

*Remote Memory Revisited*

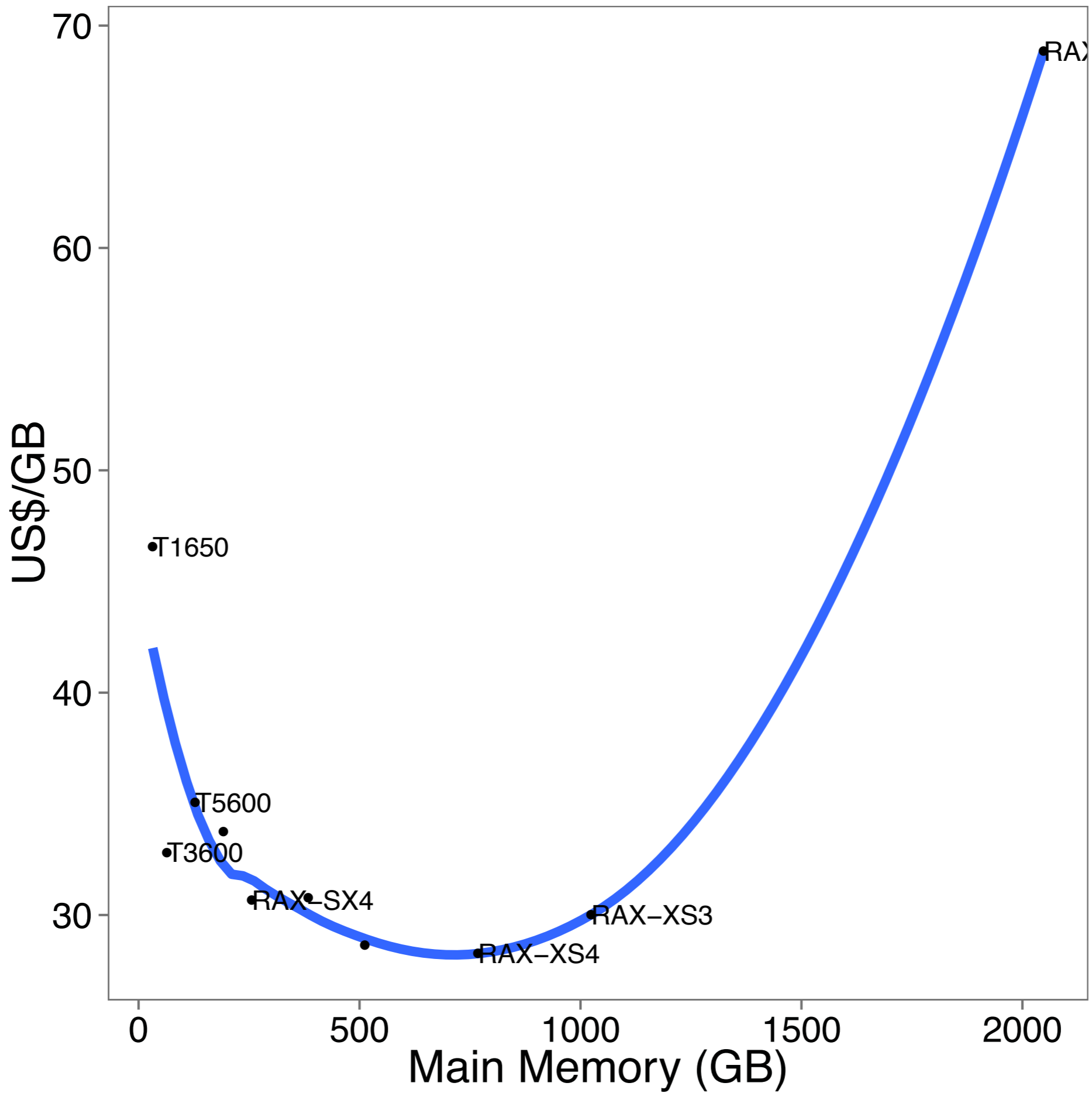
Hannes Mühleisen, Romulo Goncalves and Martin Kersten

# Database Scalability



# Why more memory?

- Memory is a critical resource, especially in OLAP use cases
  - Hash tables, intermediate results, ...
  - OS overcommits, leads to thrashing



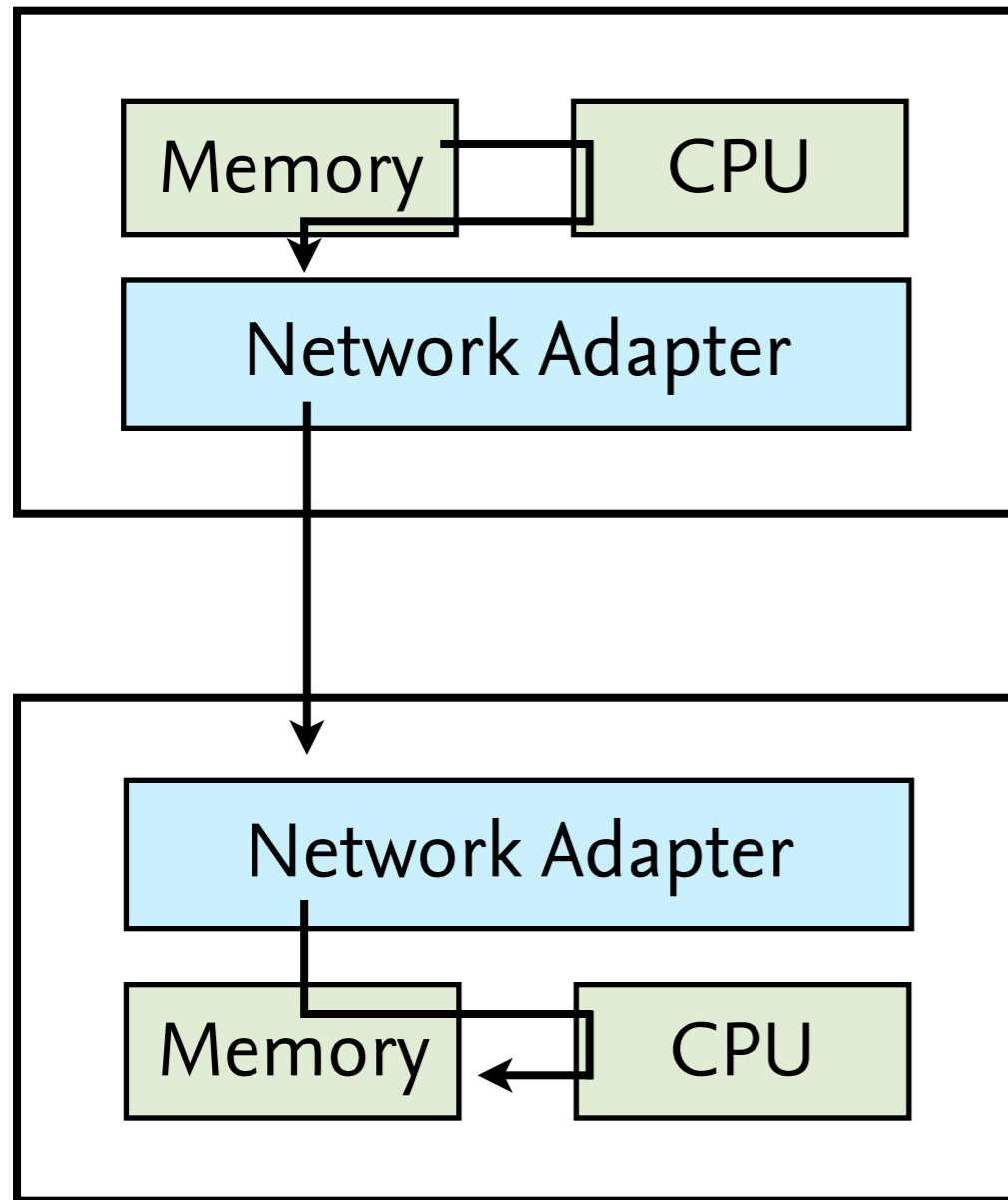
# Remote memory then

- Hack Kernel to page out to remote machines? [Tell et al. 2013]
- Store swapfile to remote file system?
- But DBs like to avoid swap anyway, so...
  - Store DB temporary files on remote system!

# New Toys

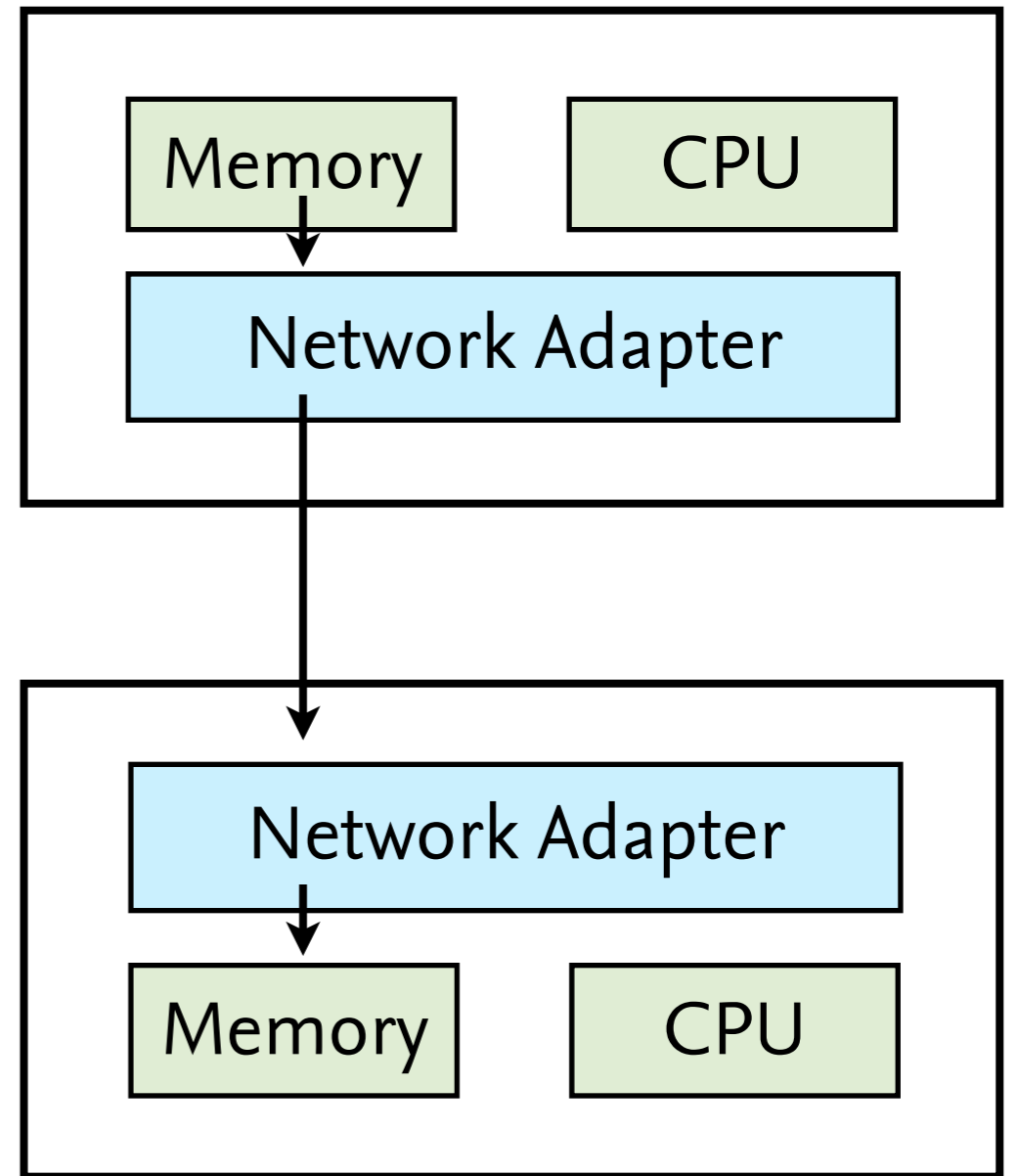


# The way it was

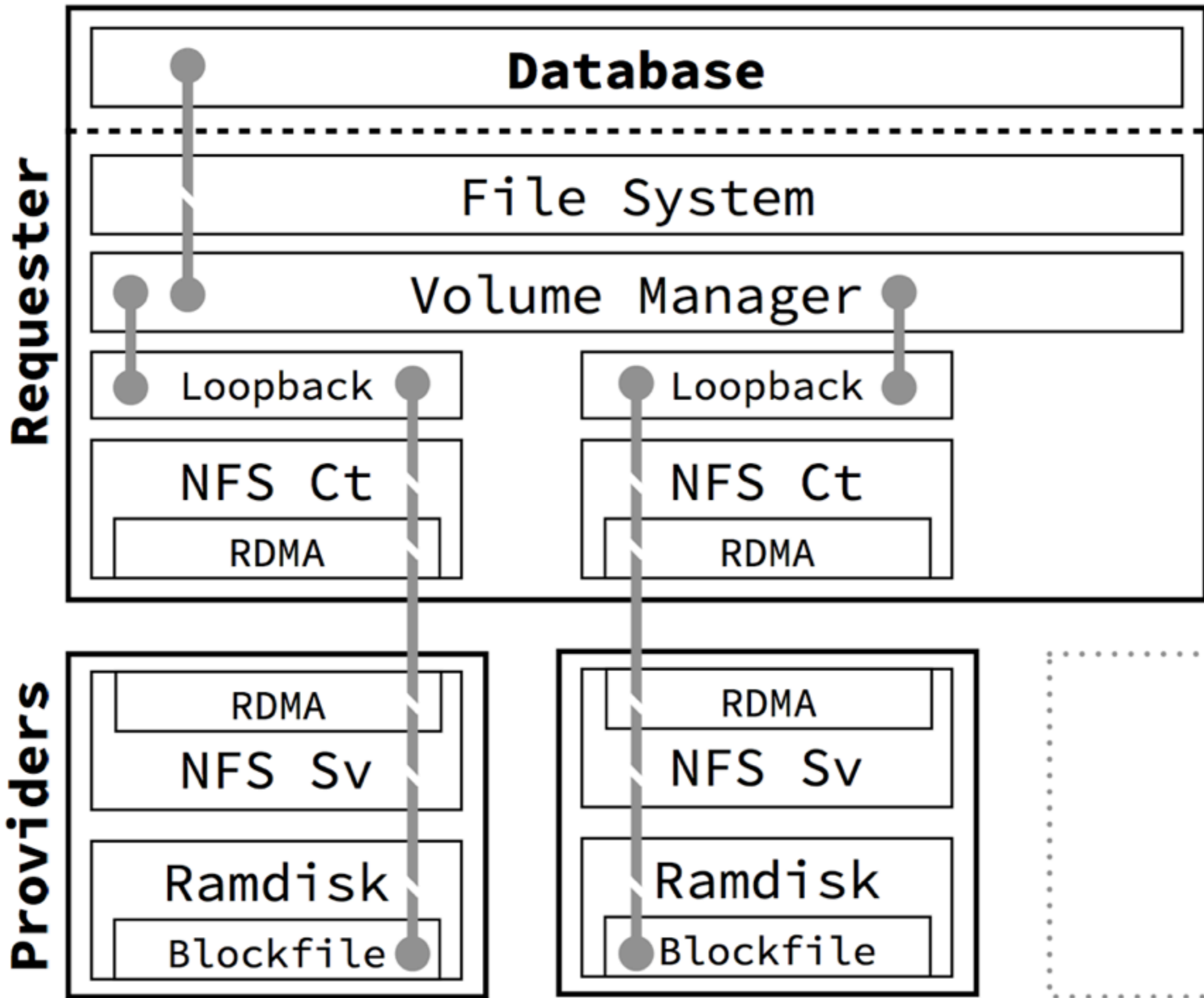


Many-Copy

# RDMA



Zero-Copy



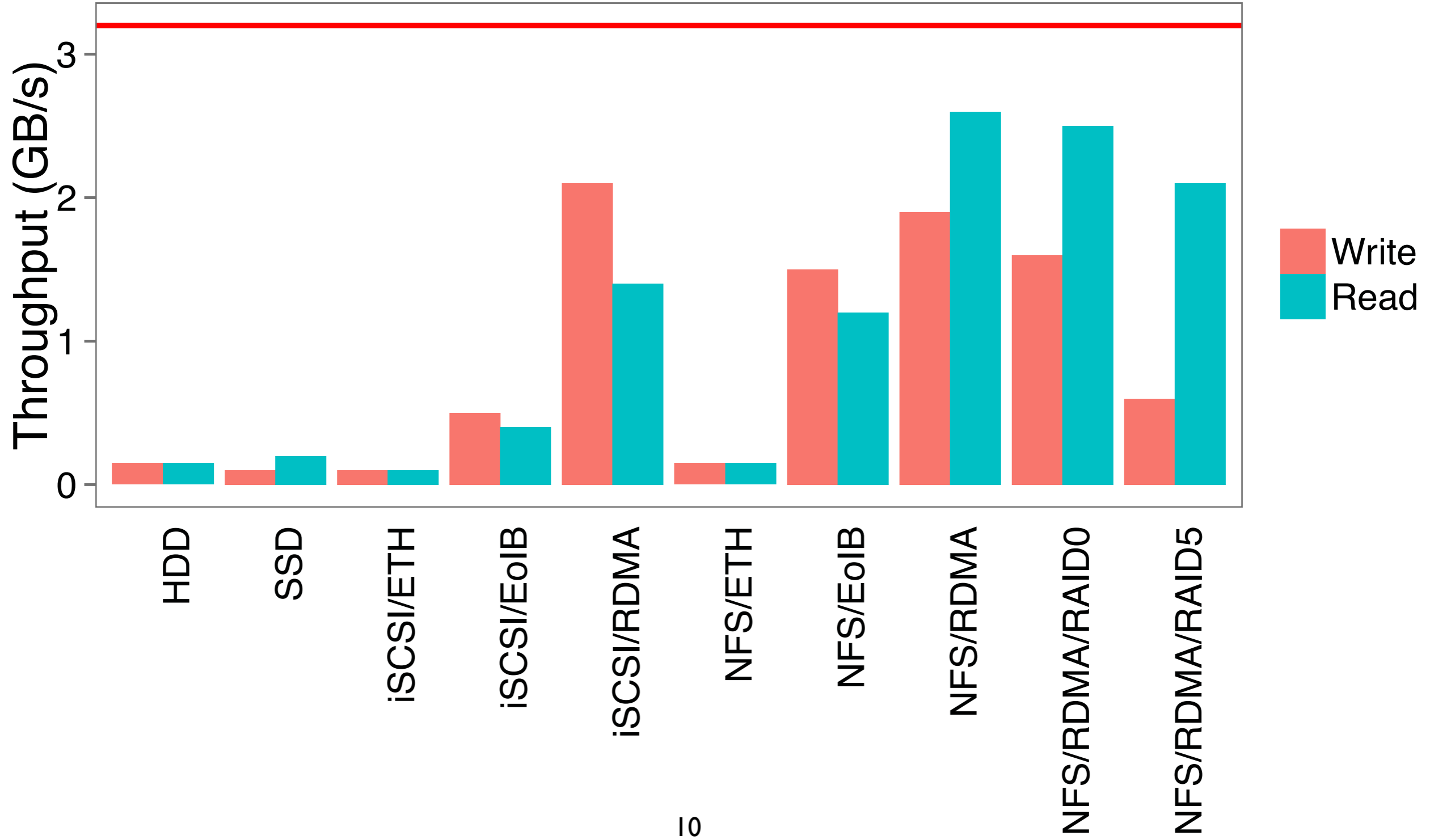


# Experimental Setup

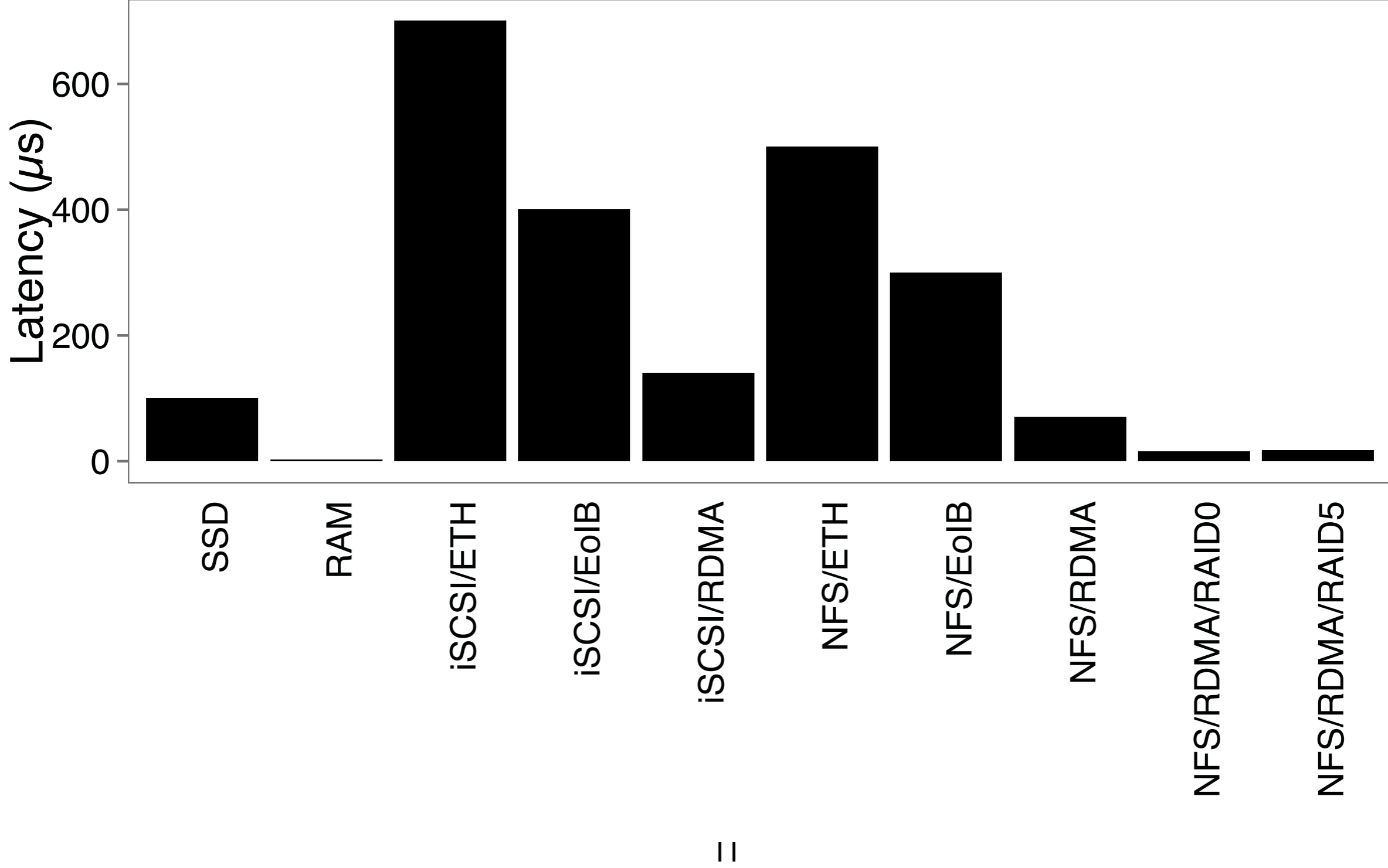
- 14 Linux COTS Boxes
  - 16 GB RAM
  - InfiniBand QDR
- 182 GB Memory total (and usable!)



# Throughput



# Latency

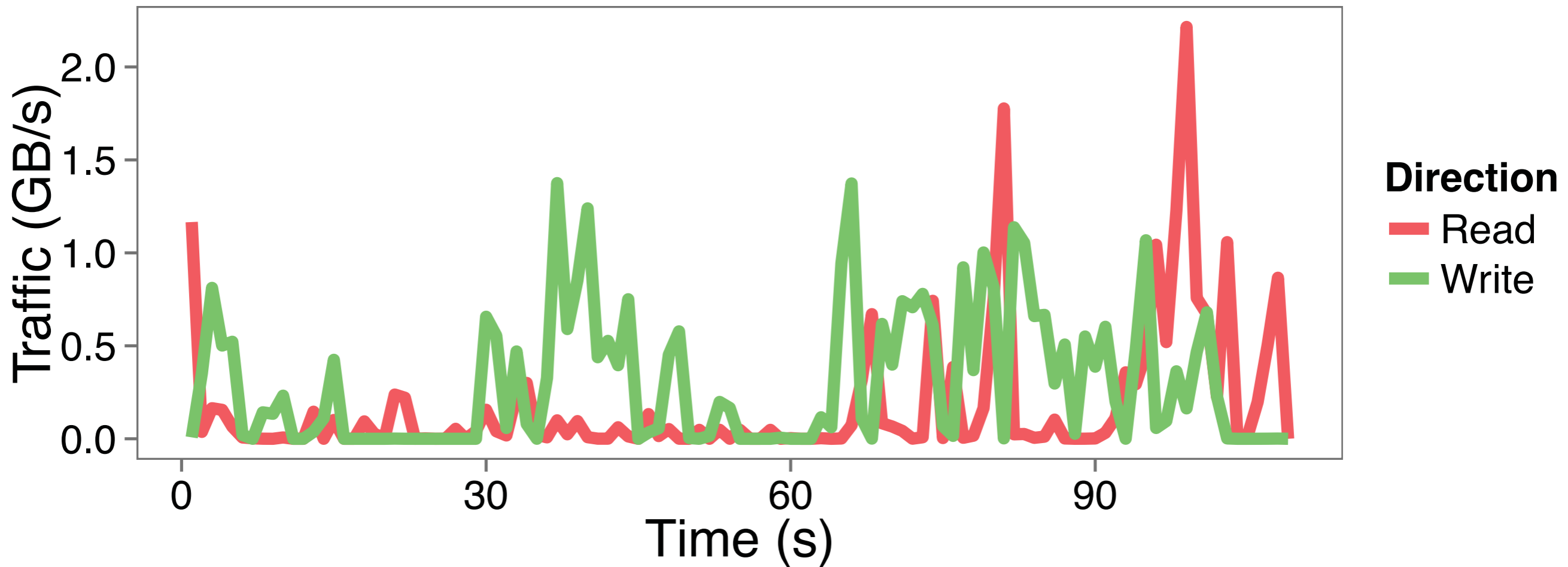


# OLAP DB (TPC-H)

- TPC-H: benchmark for relational databases focused on analytics (OLAP)
- Queries tend to have large intermediate results (SF=100):

Query	Read (GB)	Write (GB)
1	14	50
18	5	28
21	7	9
3	6	6
13	2	7

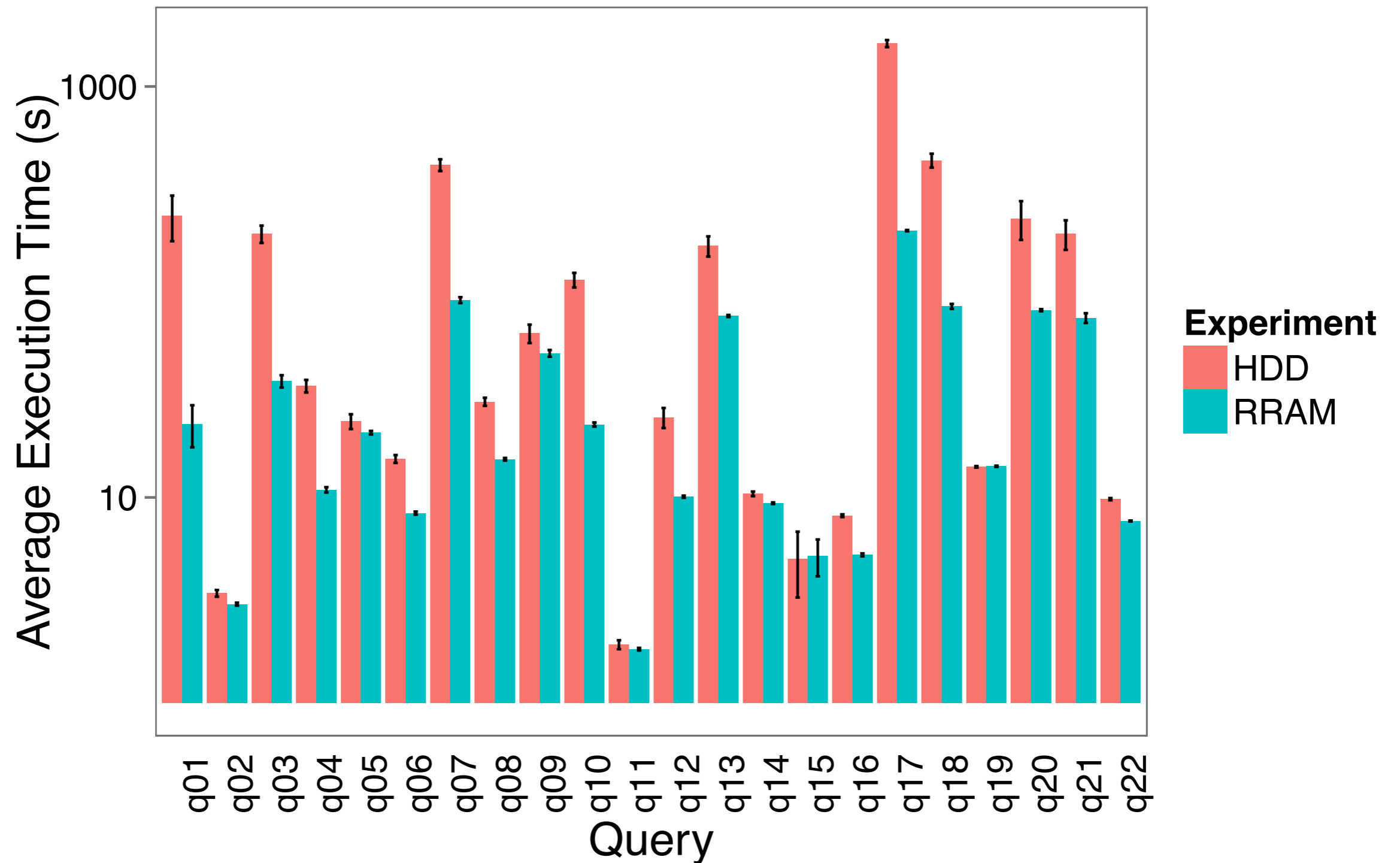
# Example: Query 18



# TPC-H Experiment

- Single node runs MonetDB with TPC-H database (SF=100)
- Gets remote memory from the 14 memory providers
- DB temporary partition resides either on disk or in remote memory
- Hot runs, 5 repetitions per query and setup

# TPC-H 100 Results



# Summary

- Remote Memory is interesting (...)
  - Lightweight technique
- RDMA allows for remote memory to make sense from a technical perspective
- OLAP database scenarios can benefit from this
- Open issue: Hardware pricing/TCO



# Thank You!

Questions?

<http://is.gd/remotemem>