

~~Hypothesis~~

RAMCloud ~~Assumptions~~

Hopes

Mendel Rosenblum
Stanford University



Talk Agenda

- **Present assumptions we are making in RAMCloud**
 - Infrastructure for data centers 5 years in future
 - Hardware and Workload assumptions
- **Elicit feedback**
 - Unconscious assumptions
 - Hopelessly wrong assumptions
- **Disclaimer: Goal #1 of RAMCloud is research**
 - Research agenda is pushing latency and scale
 - Significant risk is OK
 - Version 1.0 design decisions (OK to leave stuff out, oversimplify)

Network Assumptions

- **Assumption: Network latency going down**
 - Microsecond latencies require low latency switches & NICs
 - Encouraging: Arista: 0.6 μ s/switch
 - Lots of discouraging
 - Chicken and egg problem
- **Assumption: No significant oversubscription**
 - Scale-based solutions require high bisection bandwidth
 - Example: Recovery
- **Actively engaging networking researchers and industry partners**

Vote: Are we crazy?

Workload Assumptions

- **Less than 2 second recovery time is fast enough**
- **Applications will tolerate 2 second hiccups**

Vote: Will enough important applications tolerate 2 second pauses when failures happen?

Server Assumptions

- **Assumption: Simultaneous failures of all servers unlikely (i.e. all DRAMs can't lose contents)**
 - Require Uninterruptible power supply
 - Example: Google's battery backup per machine
 - Need only enough to transfer write buffers to non-volatile storage
- **Want to avoid non-volatile latency in write request path**

Vote: What do you think?

Assumption: Latency is goodness

- **High latency is a major application inhibitor**
 - Increases complexity, decreases innovation
 - RAMCloud will enable new applications
- **Low latency is underappreciated**
 - Users won't ask for it but will love it

Vote: Are we right?

Low Latency: Stronger Consistency?

- **Might be able to help with scaling challenges**
 - Example: Durability, update multiple in-memory replicas
- **Cost of consistency rises with transaction overlap:**
 - $O \sim R * D$**
 - O = # overlapping transactions**
 - R = arrival rate of new transactions**
 - D = duration of each transaction**
- **R increases with system scale**
 - Eventually, scale makes consistency unaffordable
- **But, D decreases with lower latency**
 - Stronger consistency affordable at larger scale
 - Is this phenomenon strong enough to matter?

Locality Assumptions

- **Locality is getting harder to find and exploit**
 - Example: Facebook usage
- **Flat model rather than tiered storage**

Vote: Right?

Read/Write Ratio Assumption

- **Read request dominate writes**
- **Postulate: Fast reads will cause applications to use more information leading to more reads than writes**

Vote: What do you think?

Replication Assumptions

- **Replication isn't needed for performance**
- **High throughput and load balancing eliminates the need for object replication**
- **Few applications need more than 1M ops/second to a single object**

Vote: Right?

WAN Replication Assumption

- **There are interesting applications that don't require cross data center replication**
 - Speed of light not our friend

Transaction Assumption

- **Multi-object update support is required**
 - Distributed applications need support for concurrent accesses

Security assumptions

- **Threat model**
 - RAMCloud servers and network physically secure
 - Don't need encryption on RPCs, DoS defense, etc.

Data Model Assumption

- **Simple table/object with indexing sufficient**
 - Dictated by speed
 - SQL/relational model can be built on top

Thank you