Fast Crash Recovery in RAMCloud

Ryan Stutsman

SEDCL Retreat
June 3, 2011

Durability and Availability

- **Goals:**
  - No impact on performance
  - Minimum cost, energy

- **Replicate in DRAM of other masters?**
  - 3x system cost, energy
  - Still have to handle power failures
  - Replicas unnecessary for performance

Durability and Availability

- **RAMCloud approach:**
  - 1 copy in DRAM
  - Backup copies on disk/flash: durability ~ free!

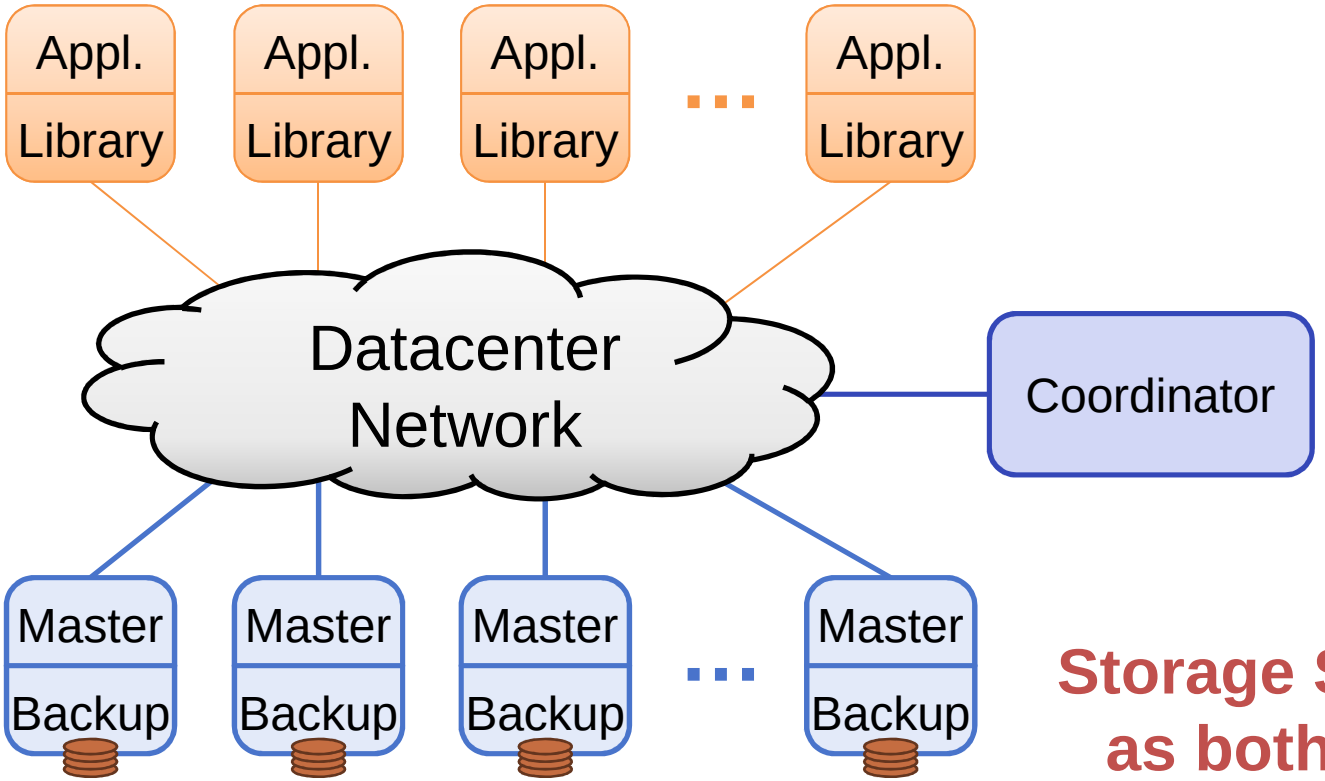
Problems

- **Synchronous disk I/O's during writes?**
  - **Buffered logging**

- **Data unavailable after crashes?**
  - **Fast recovery**

RAMCloud Architecture

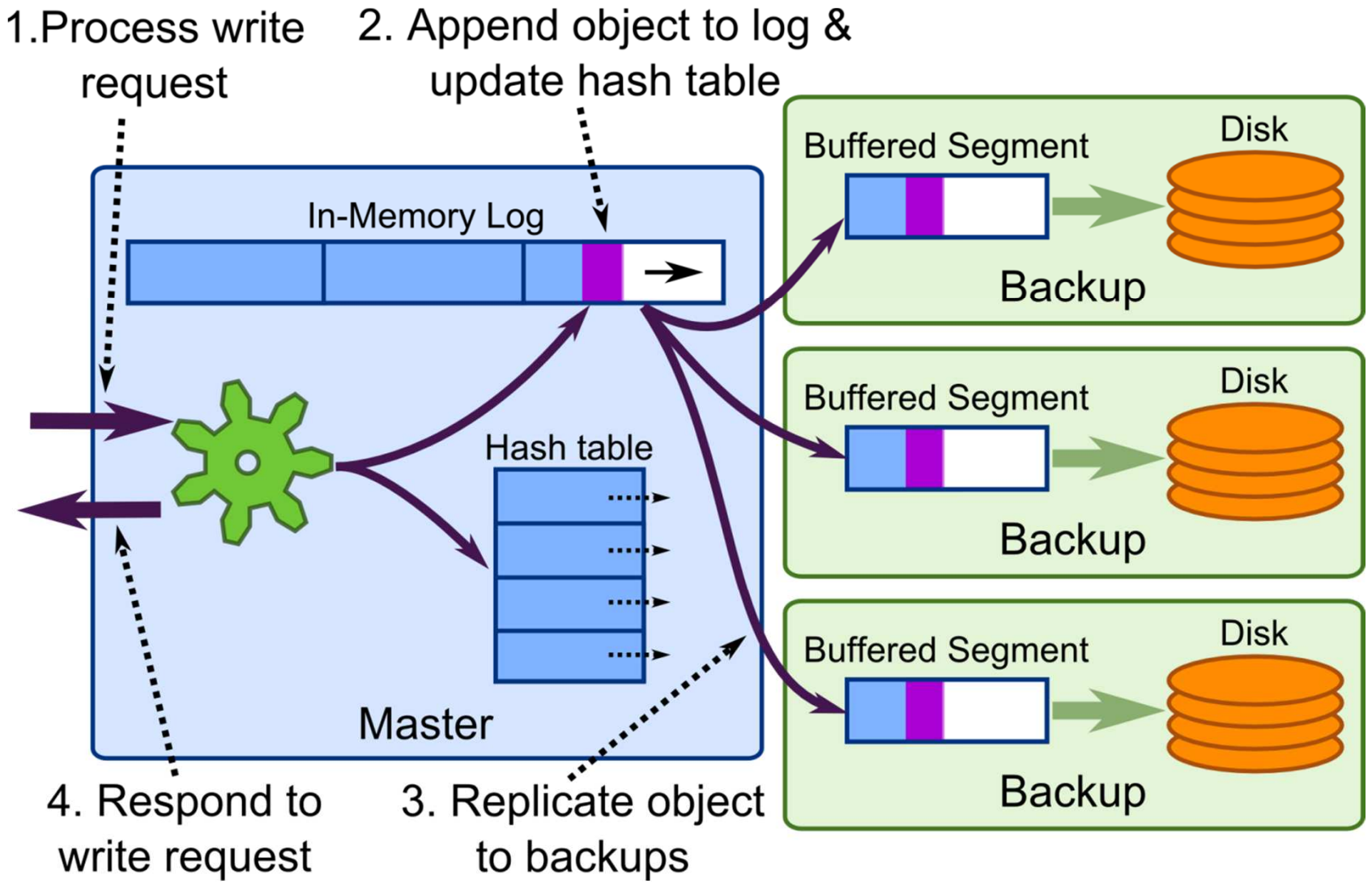
1,000 – 100,000 Application Servers



1,000 – 10,000 Storage Servers

Storage Servers act as both a Master and a Backup

Buffered Logging



Buffered Logging

- **No disk I/O during write requests**
  - But must guarantee buffered data durability:
    - DIMMs with built-in flash backup?
    - Caches on enterprise disk controllers?
    - Per-server battery backups?
- **8 MB I/Os minimize disk latency overhead**
  - Achieves 90% disk bandwidth
  - Helpful for reading during fast recovery as well
- **Master's memory also log-structured**
  - Log cleaning ~ generational garbage collection

Master Crash Recovery

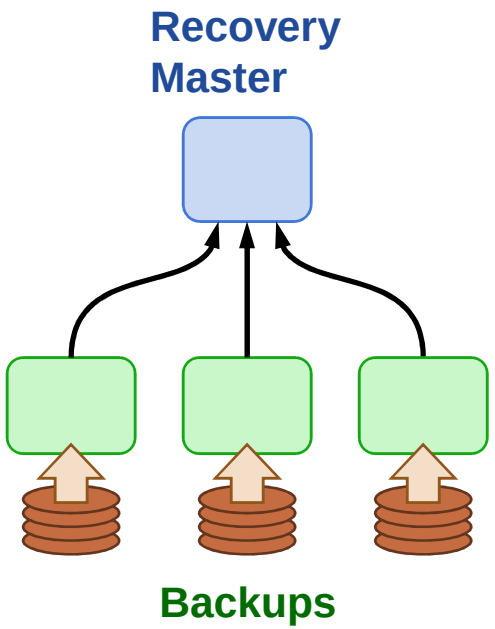
- **Problem: Data unavailable after a master crash**

- **Goal: 1 – 2 second recovery**
  - Applications just see “hiccups”
  - Good enough for “continuous availability?”

- **Solution: Harness system scale**

Recovery, First Try

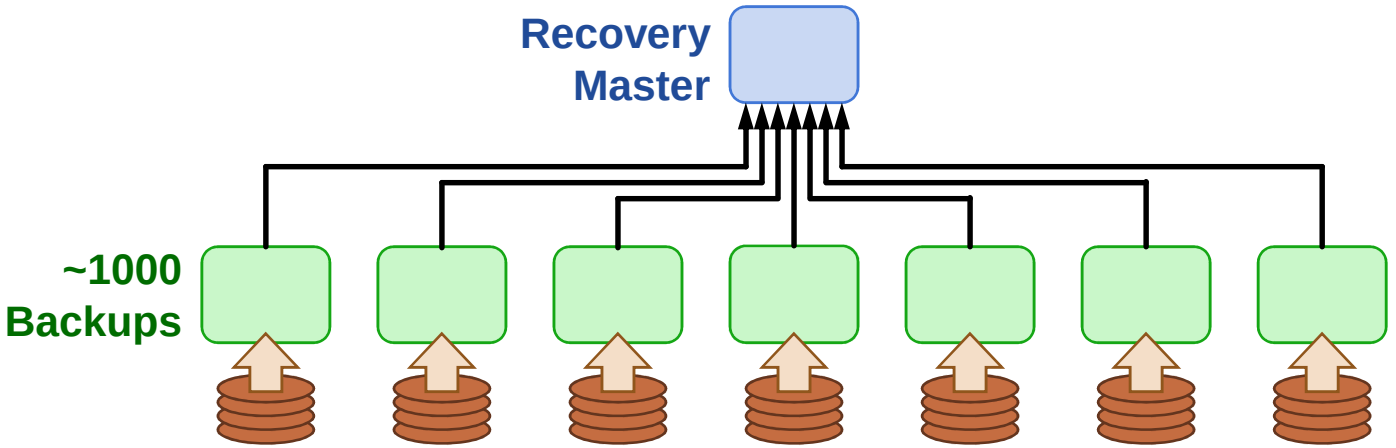
- **Master chooses backups statically**
  - Each backup stores entire log for master
- **Crash recovery:**
  - Choose recovery master
  - Backups read log info from disk
  - Transfer logs to recovery master
  - Recovery master replays log
- **First bottleneck: disk bandwidth:**
  - 64 GB / 3 backups / 100 MB/sec/disk ≈ 210 seconds
- **Solution: more disks (more backups)**



Recovery, Second Try

- **Scatter logs:**
  - Each log divided into 8MB segments
  - Master chooses different backups for each segment (randomly)
  - Segments scattered across all servers in the cluster

- **Crash recovery:**
  - All backups read from disk in parallel
  - Transmit data over network to recovery master



Scattered Logs

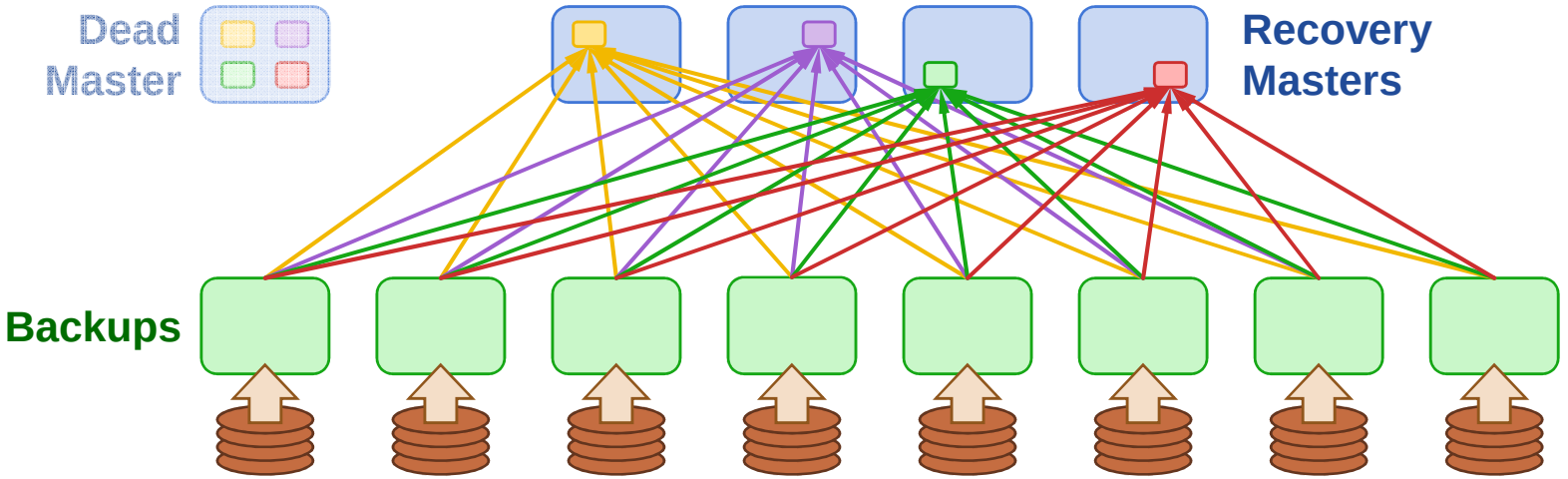
- **Disk no longer a bottleneck:**
  - 64 GB / 8 MB/segment / 1000 backups ≈ 8 segments/backup
  - 100 ms/segment to read from disk
  - 0.8 seconds to read all segments in parallel

- **Second bottleneck: NIC on recovery master**
  - 64 GB / 10 Gbits/second ≈ 60 seconds
  - Recovery master CPU is also a bottleneck

- **Solution: more recovery masters**
  - Spread work over 100 recovery masters
  - 64 GB / 10 Gbits/second / 100 masters ≈ 0.6 seconds

Recovery, Third Try

- **Divide each master's data into partitions**
  - Recover each partition on a separate recovery master
  - Partitions based on tables & key ranges, *not log segment*
  - Each backup divides its log data among recovery masters



Start of Recovery

1. Coordinator C gets “M is down” report

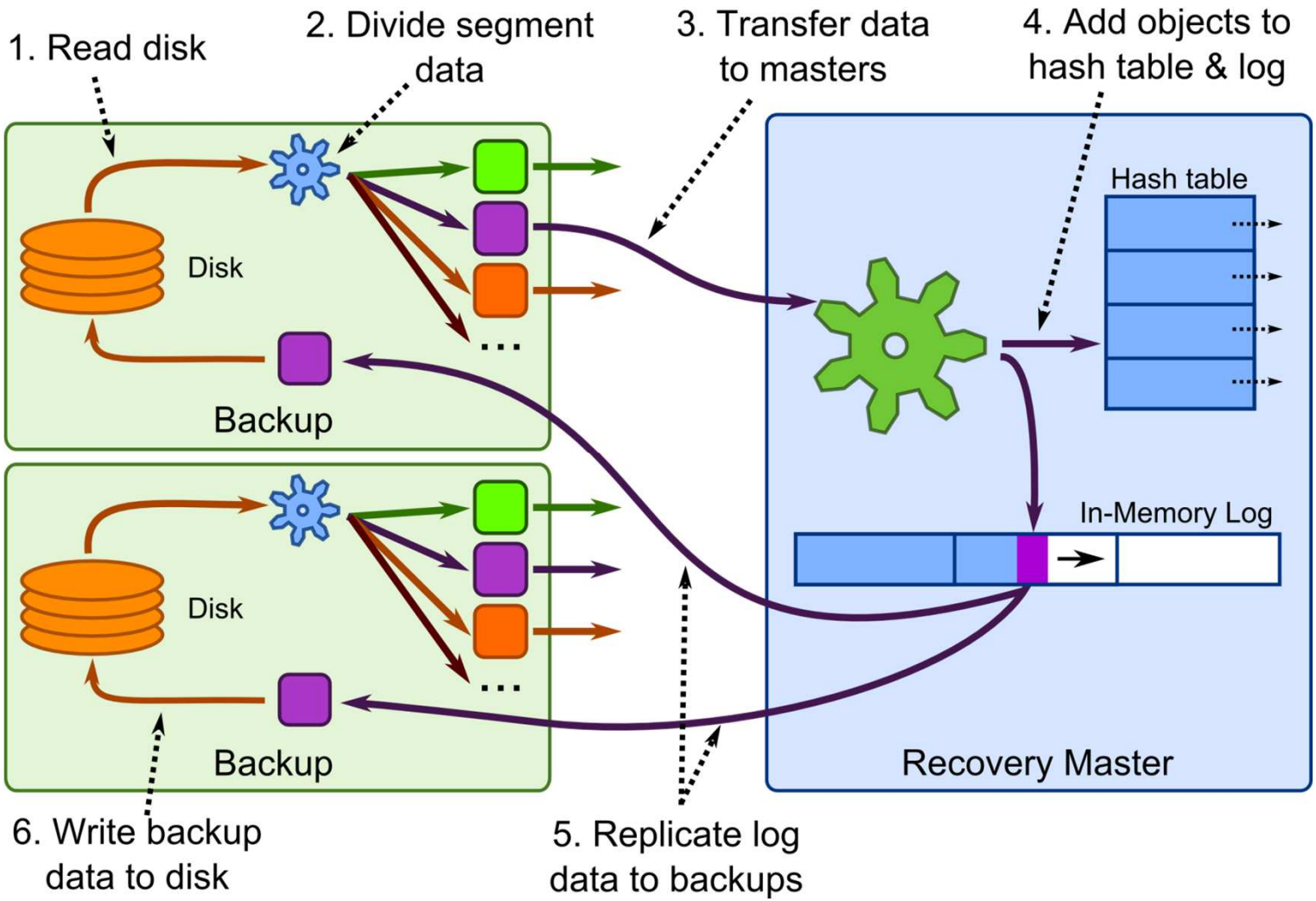
2. C verifies M is down

3. C broadcasts to cluster:
   “which segments do you have for M?”
   “begin reading and partitioning them”

4. C verifies no segments are missing

5. C notifies selected recovery masters:
   “recover this partition of M from these locations”

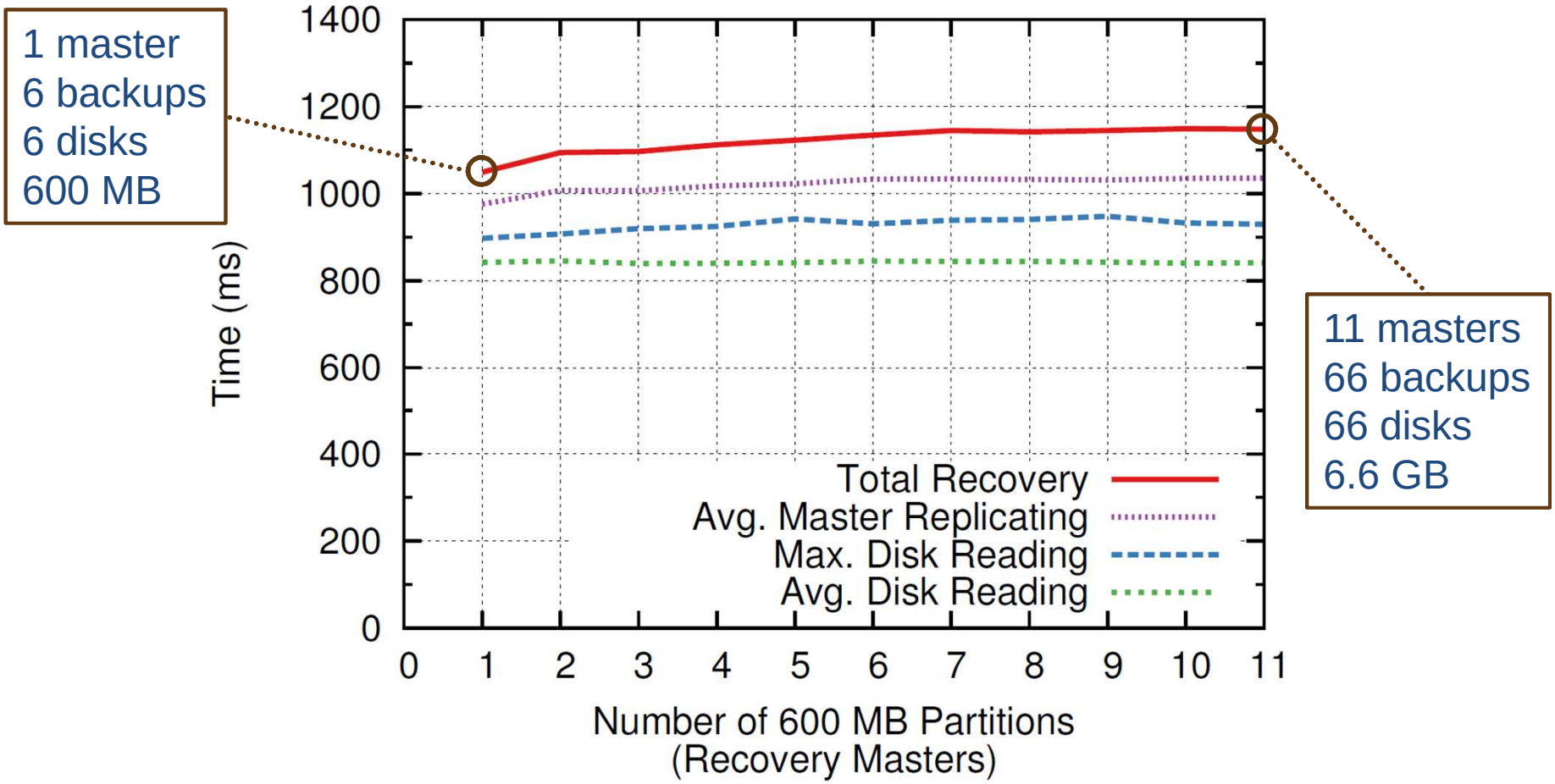
Recovery


1. Read disk
2. Divide segment data
3. Transfer data to masters
4. Add objects to hash table & log
Disk
Backup
Disk
Backup
Hash table
In-Memory Log
Recovery Master
6. Write backup data to disk
5. Replicate log data to backups

Parallelism During Recovery

- **Tightly overlapped and pipelined**
  - Each recovery master can recover 400-800 MB of log per second
  - All steps are all being performed by all hosts simultaneously
- **Segments can be replayed in any order**
  - Possible due to version number on objects
  - Recovery masters can work independently and distribute load over backups
  - Recovery masters can replay any data when it becomes available

Recovery Scalability



TODO

- **Restoring locality**
- **Recovering Backups**
- **Cold boot**

Recovering Backups

- **In parallel with master recovery**
- **Each master recreates any segments it has stored on that backup elsewhere**
- **Expect to re-replicate 8 segments**
  - 1,000 machines, 64 GB per master
  - Just need to buffer on a new backup
  - Fast compared to master recovery

Recovery Summary

Before Crash

- **Scatter log data**
- **Balance partitions**
  - Steve's talk

On Crash

- **Detect failure**
- **Find log data**
- **Check log integrity**
- **Select new masters**
- **Replay log data**

- **Recreate backups**