

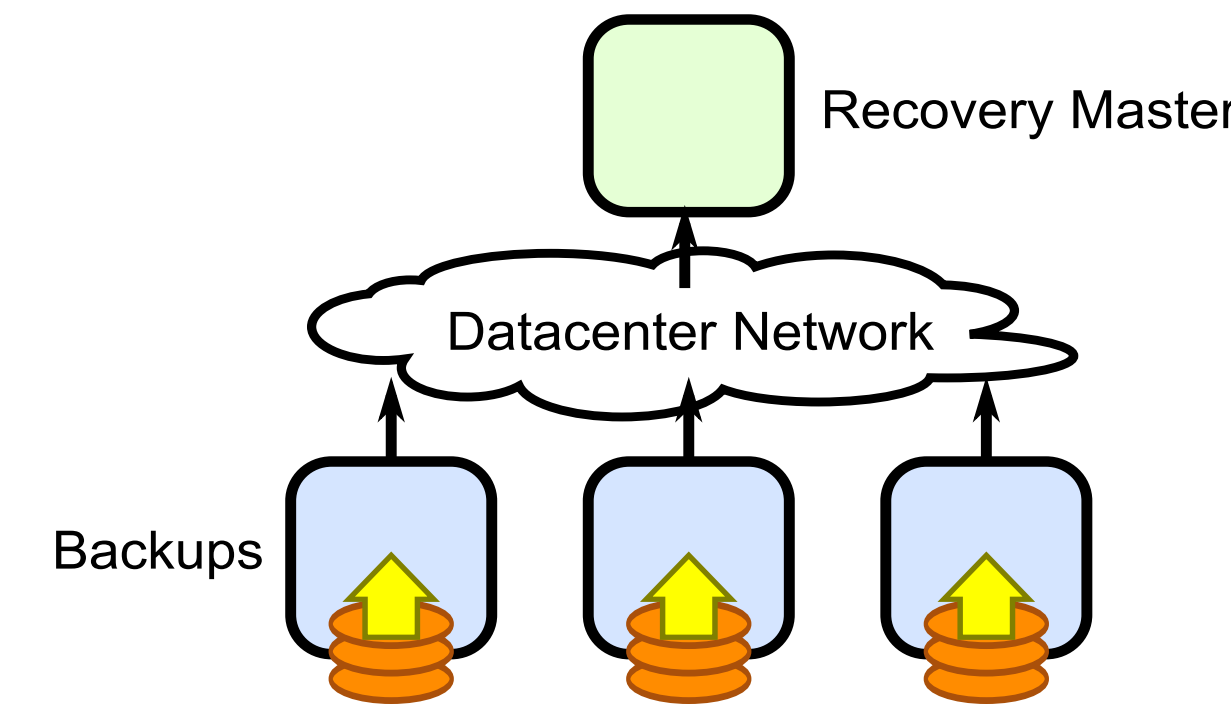
Motivation

- Building a storage system for future datacenters
 - Large scale: 1,000 - 100,000 storage servers
 - Low latency: 5-10 microsecond RTT across datacenter
- All data always in RAM
 - 64+ GB DRAM/server
- Durability goals:
 - Small impact on performance
 - Minimum cost and energy
- Why not keep replicas in DRAM of other servers?
 - Triples cost and energy usage
 - Power failures are still a problem
- RAMCloud's approach: **fast recovery**
 - 1 copy in DRAM, backup copies on disk/flash
 - Hypothesis: failures will not be noticed if fast enough

Approach

Disk Bottleneck

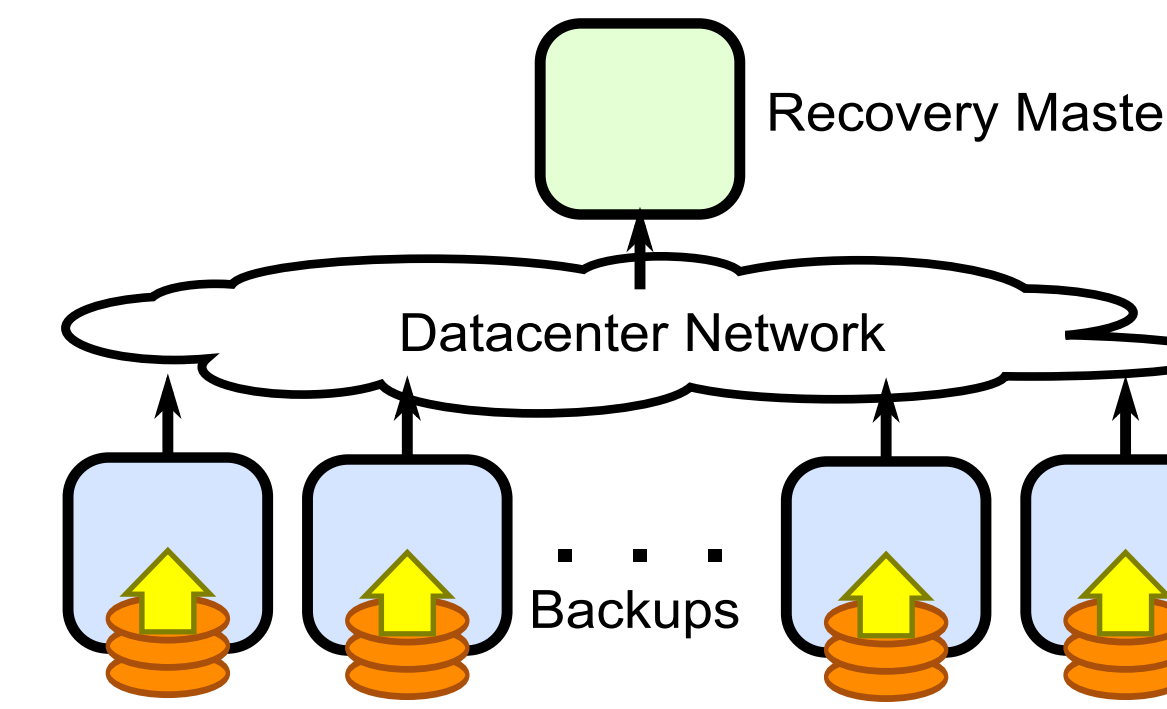
64 GB / 3 disks / 100 MB/s/disk = 3.5 minutes



- Static set of backups is insufficient
 - **Harness scale:** Use many disks
 - Read from all 1,000+ machines
 - Scatter data throughout the cluster
 - 64 GB / 1000 disks / 100 MB/s/disk = **0.6 seconds**

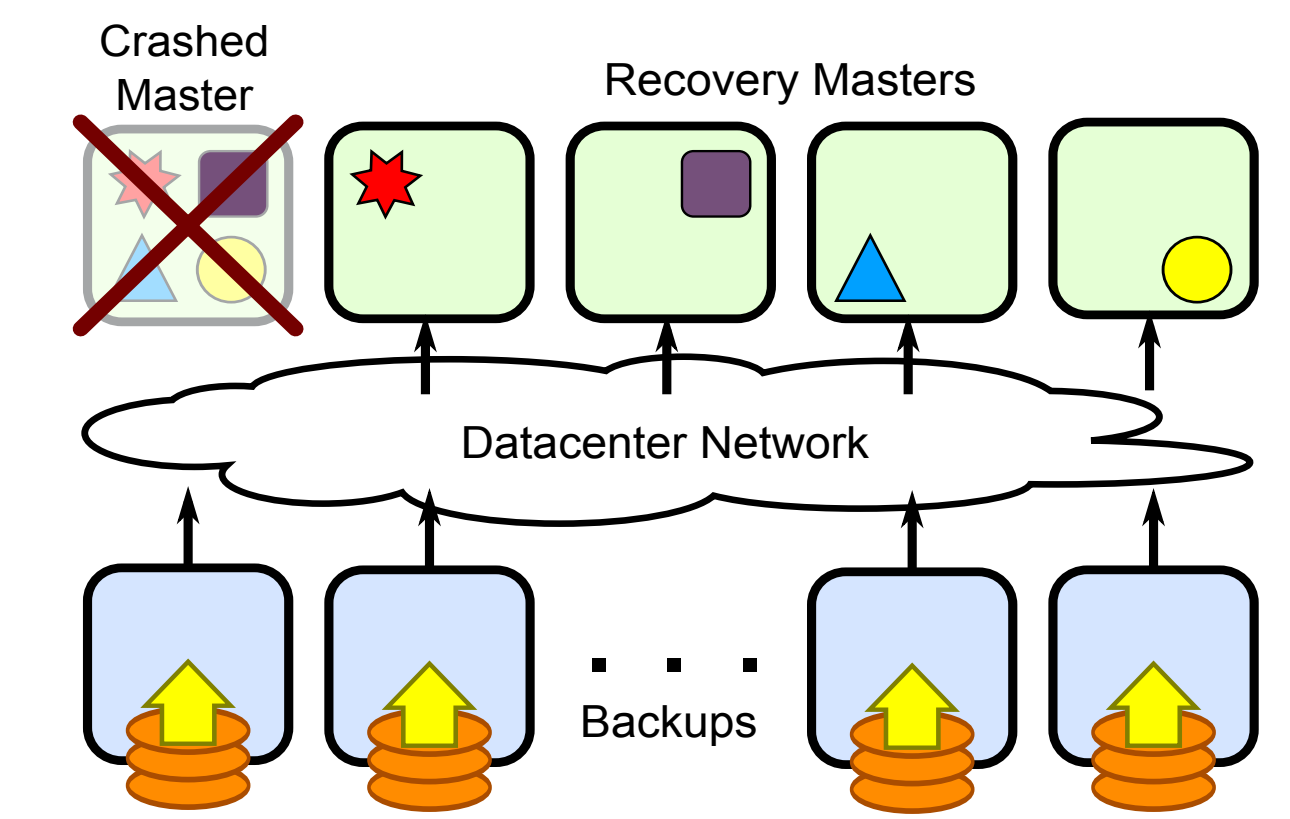
Network Bottleneck

64 GB / 10 Gbps = 1 minute



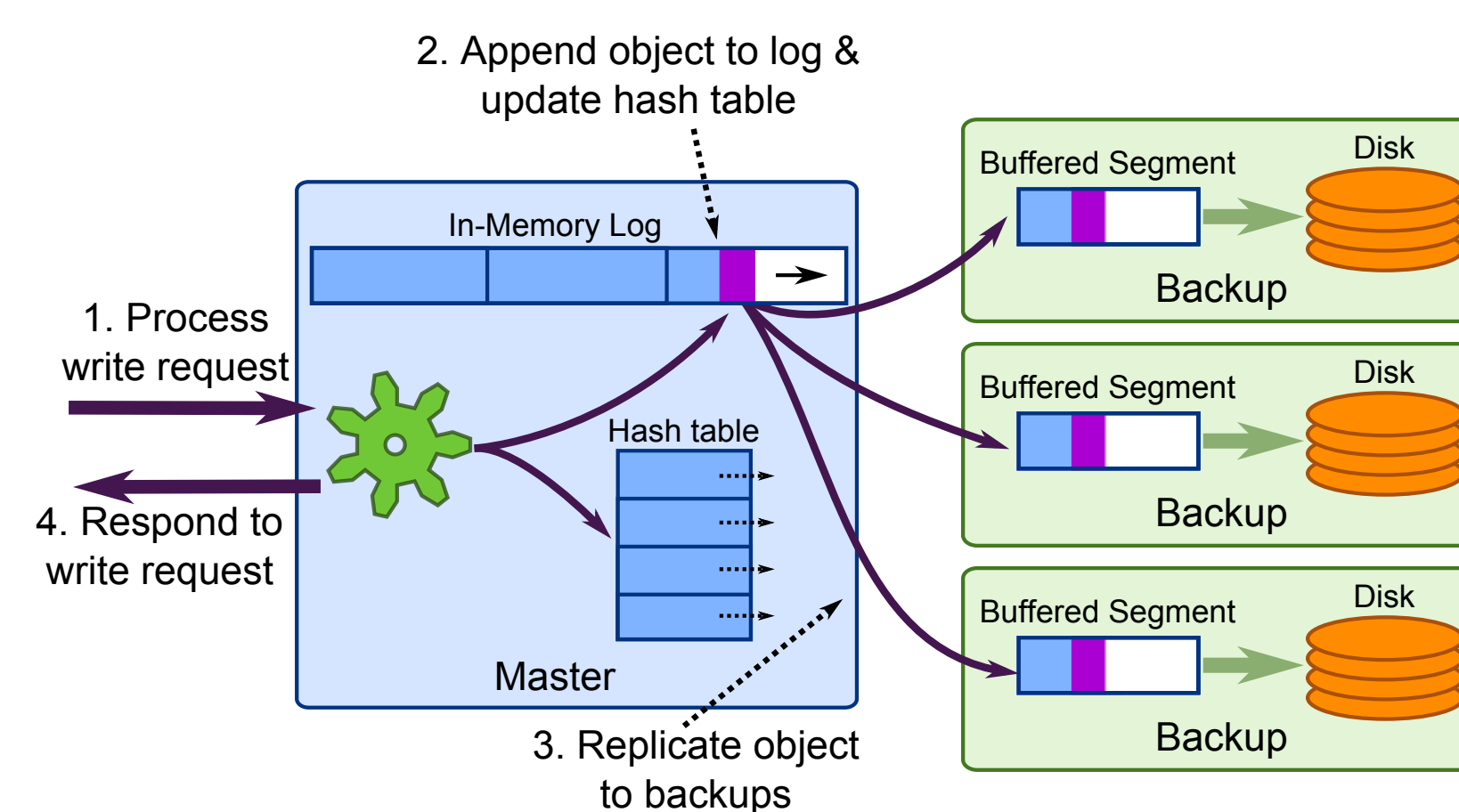
- Incast is too much for a single NIC
 - **Harness scale:** Use many hosts (NICs)
 - About 100 recovery masters will do
 - Each recovery master can recover about 400-800 MB/s
 - Need a ratio of about 6 disks to each recovery master

Fast Recovery



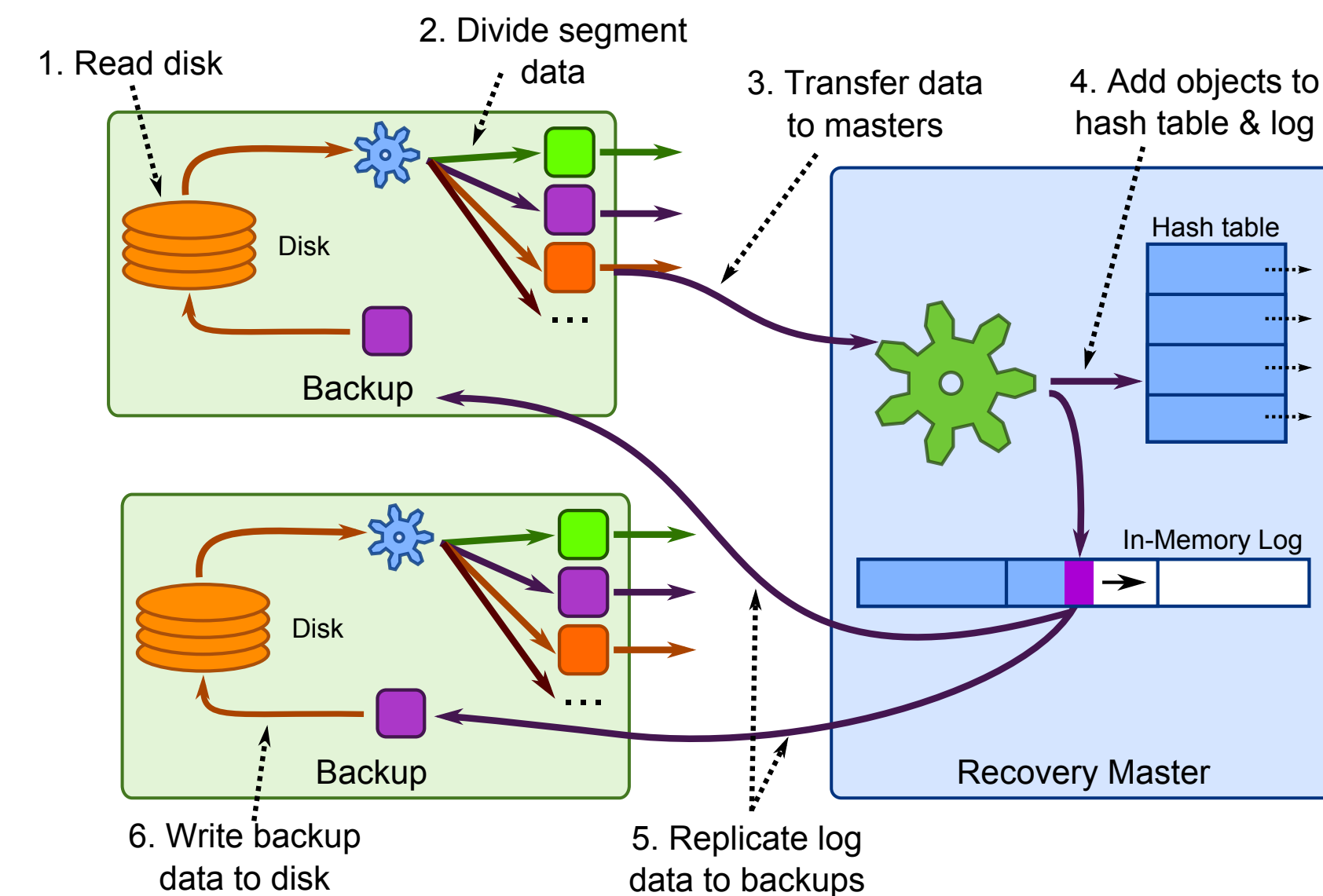
- **1-2 second recovery** with 1,000 nodes
 - Read backup data from 1,000 nodes and replay in parallel on 100 new master servers.
 - Leverage scale: avoid disk, memory, cpu, and network bottlenecks.

Data Scattering



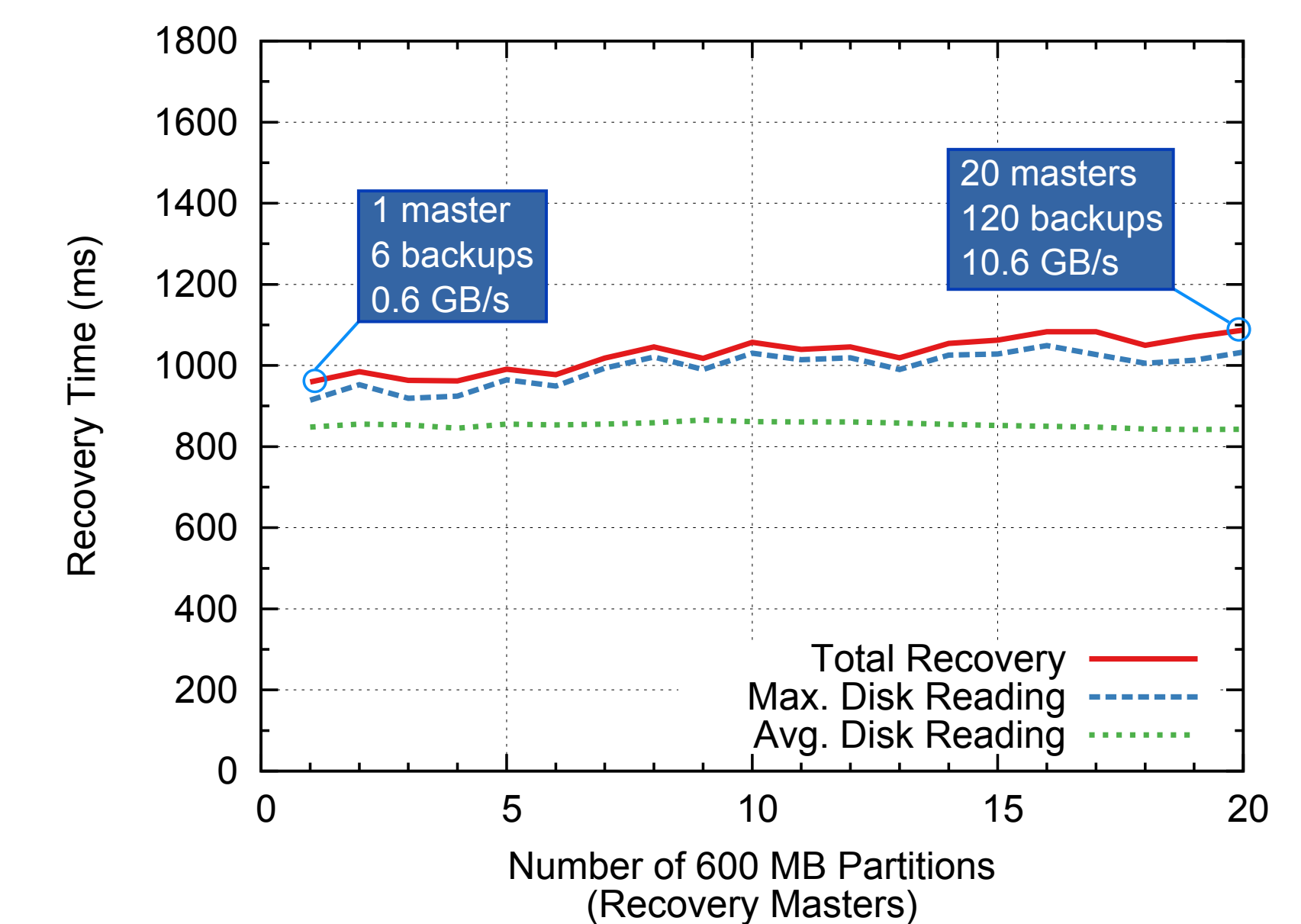
- Masters replicate writes to backups immediately
- Backups buffer it and flush to disk/flash in batch
 - Need auxiliary power source for buffers for power failure
- Backup locations chosen randomly to scatter segments
 - Constraints on placement due to correlated failures
 - Tweaked to balance expected read time
 - Provides the needed read bandwidth for recovery

Replay



- Every host is involved in recovery and they work in parallel
- Work on each host proceeds in parallel (steps are pipelined)
- Recovery masters make several parallel requests to backups
- Prevents pipeline stalls when backups are not ready with data
- New log segments are buffered until recovery is complete

Results



- 60-node cluster, 32 Gbps Infiniband network
- Recovered 11.7GB in 1.1 seconds (35GB in 1.6s with SSDs)
- Expect to recover entire master in 1 - 2 seconds at scale
- Time spent re-replicating is the current bottleneck
- Implementation hides disk speed variance well