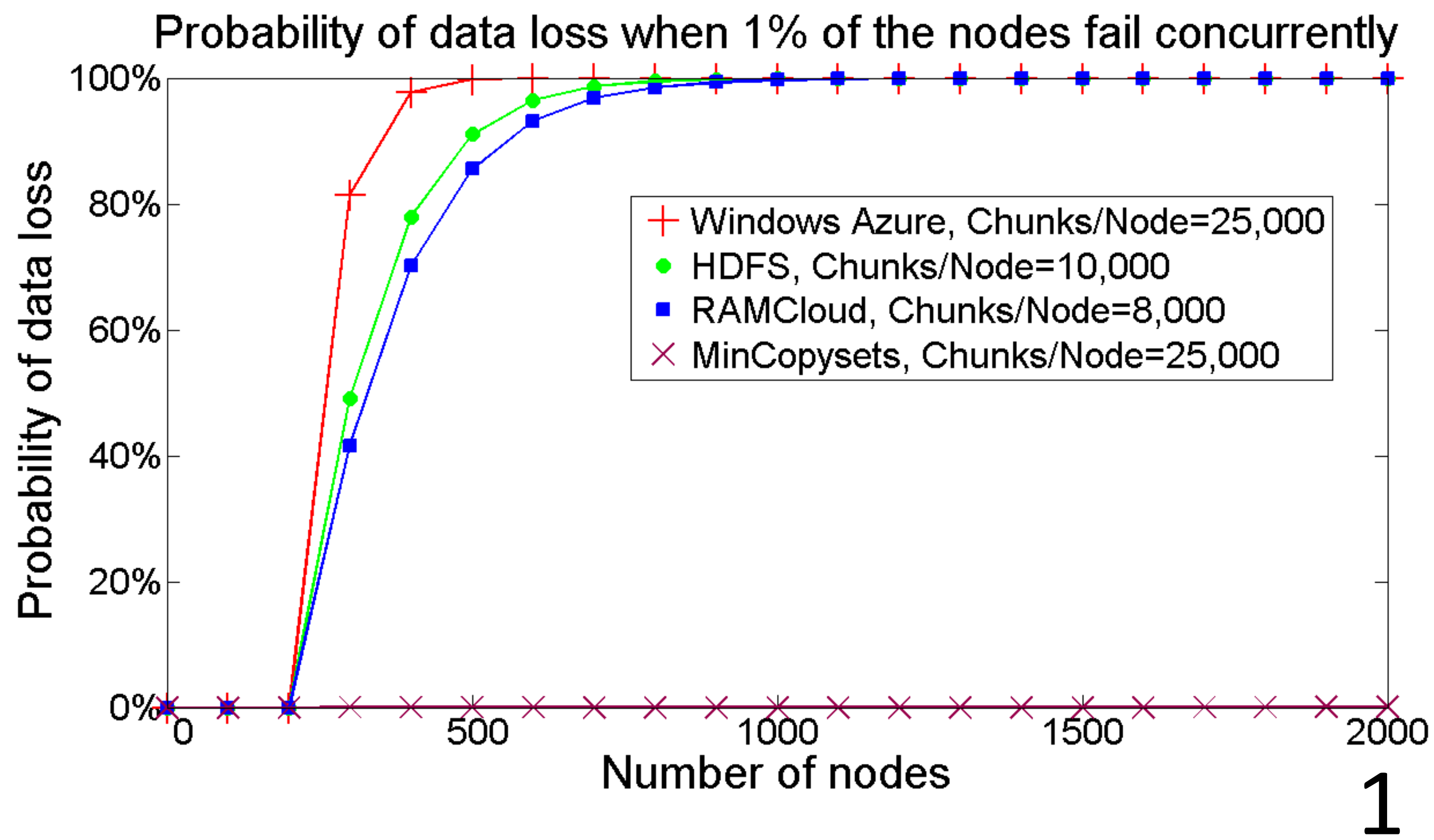
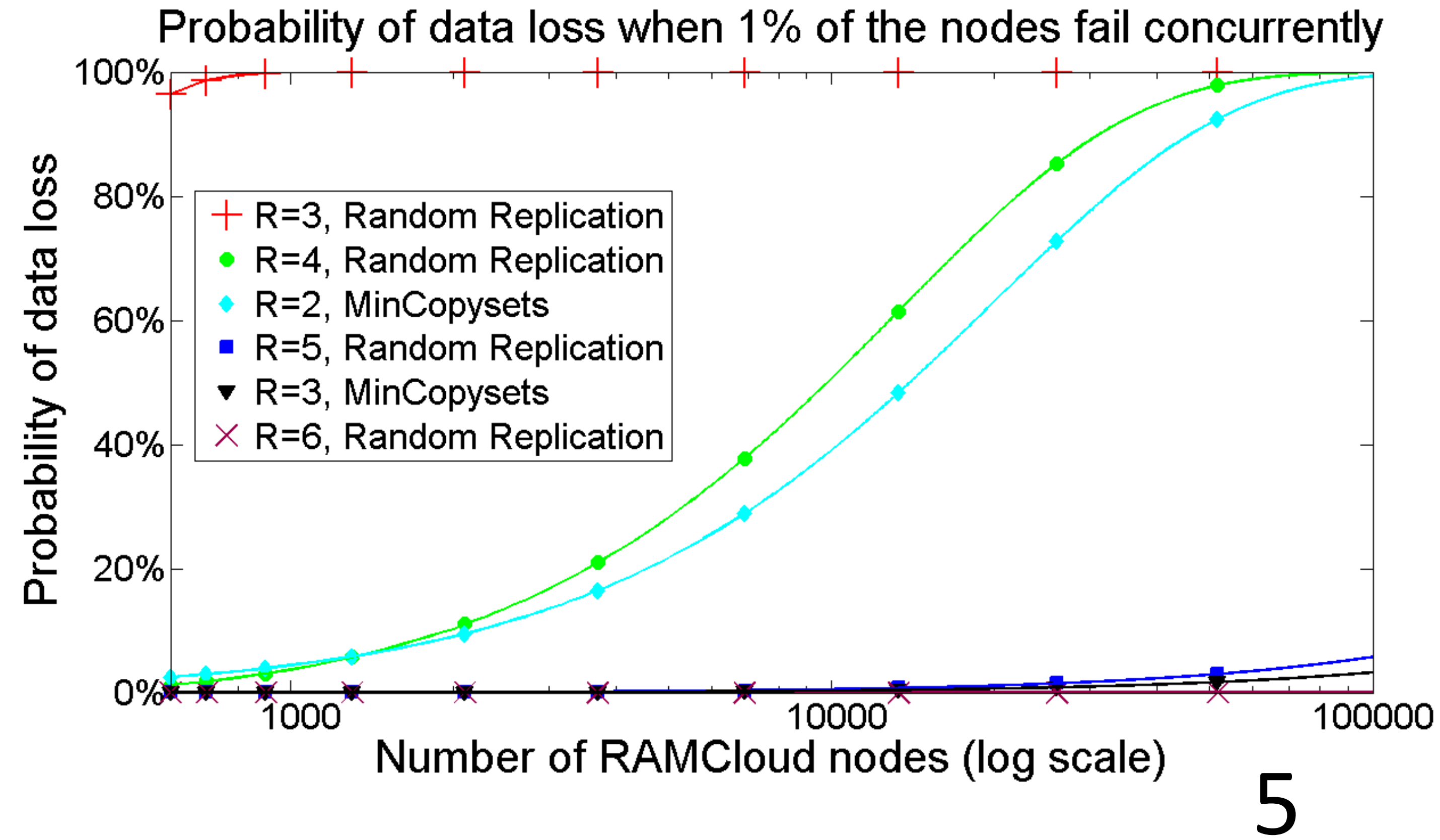


# MinCopysets: Derandomizing Replication in Cloud Storage

Asaf Cidon, Ryan Stutsman, Stephen Rumble, Sachin Katti, John Ousterhout and Mendel Rosenblum

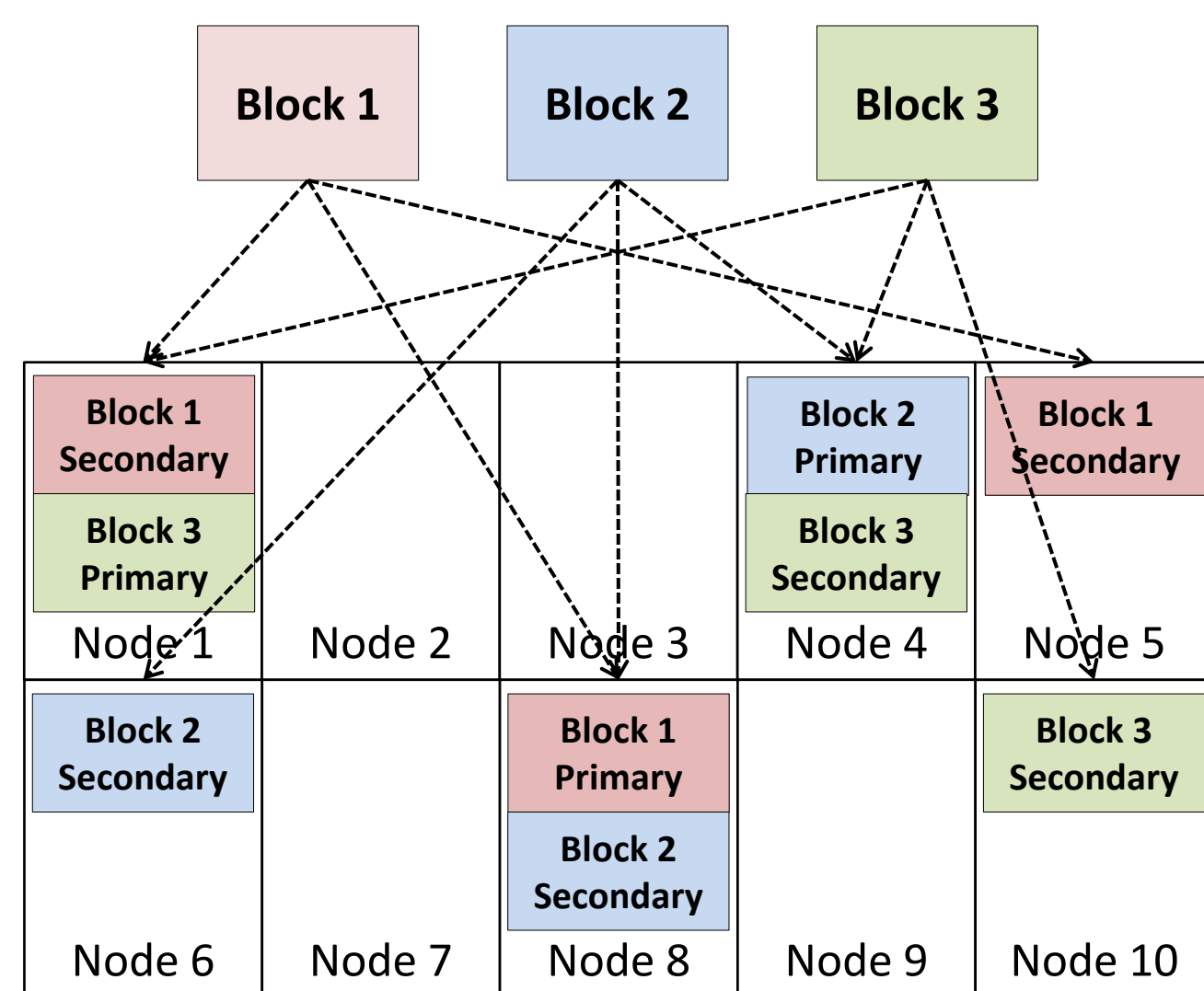


1



5

## Random Replication (HDFS, RAMCloud, GFS)



2

## MinCopysets' Trade Off

- Trade off between the frequency and magnitude of failures
- When failures occur, it's very unlikely that the failed nodes will store all copies of a block
  - Data loss occurs very rarely (once every 500 years)
- If MinCopysets loses data it will lose all data on a node
  - The magnitude of data loss is greater

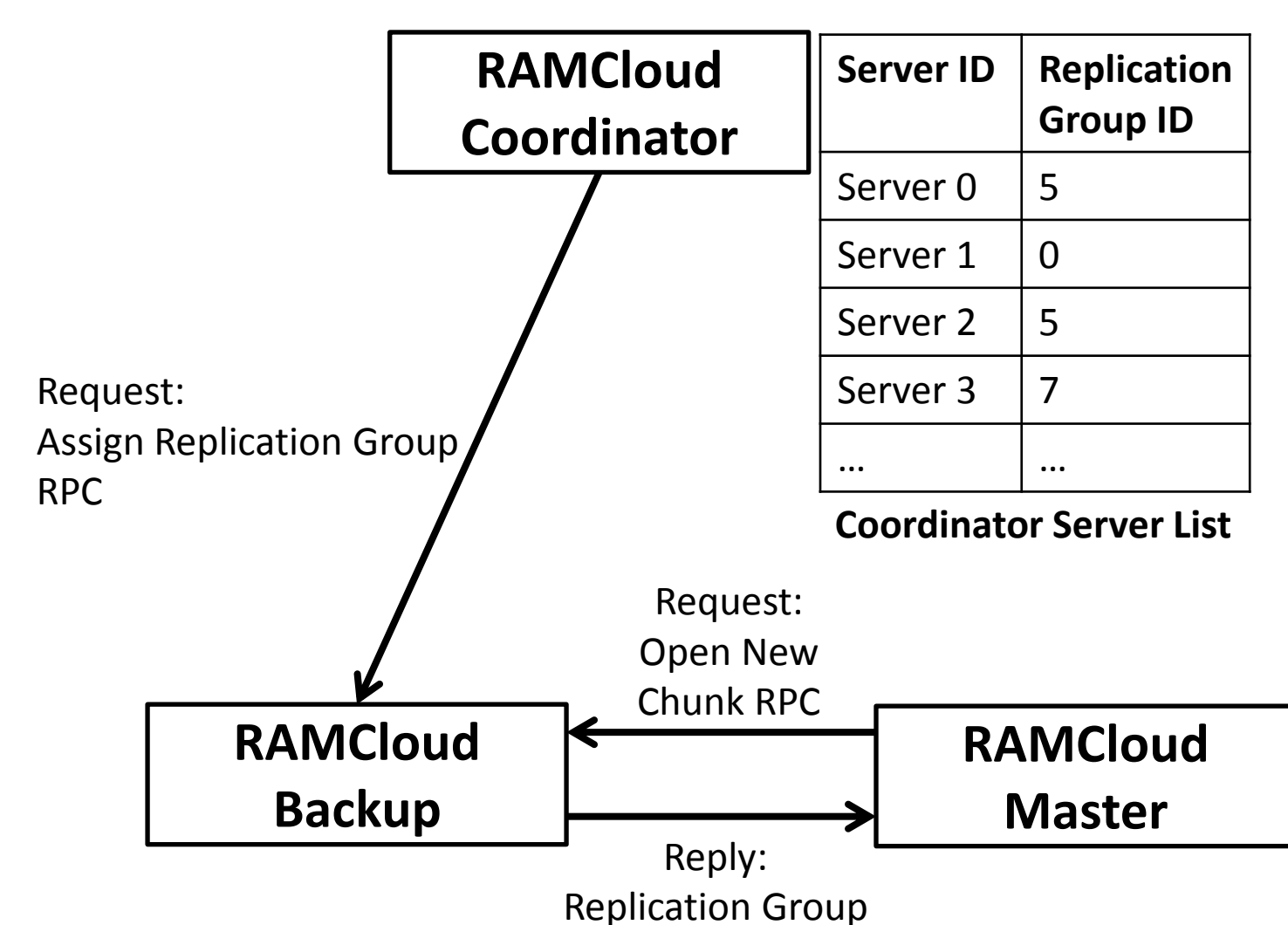
6

## The Problem

- Cloud storage uses randomization for load balancing and replication
  - Load balancing: data partitioned into blocks, randomly distributed across cluster
  - Replication: blocks randomly replicated on different machines
- Randomized replication loses data in power outages
  - Scenario: 0.5-1% of the nodes fail to reboot
  - Result: 5-10 blocks are lost (e.g., Yahoo '09, LinkedIn '12)

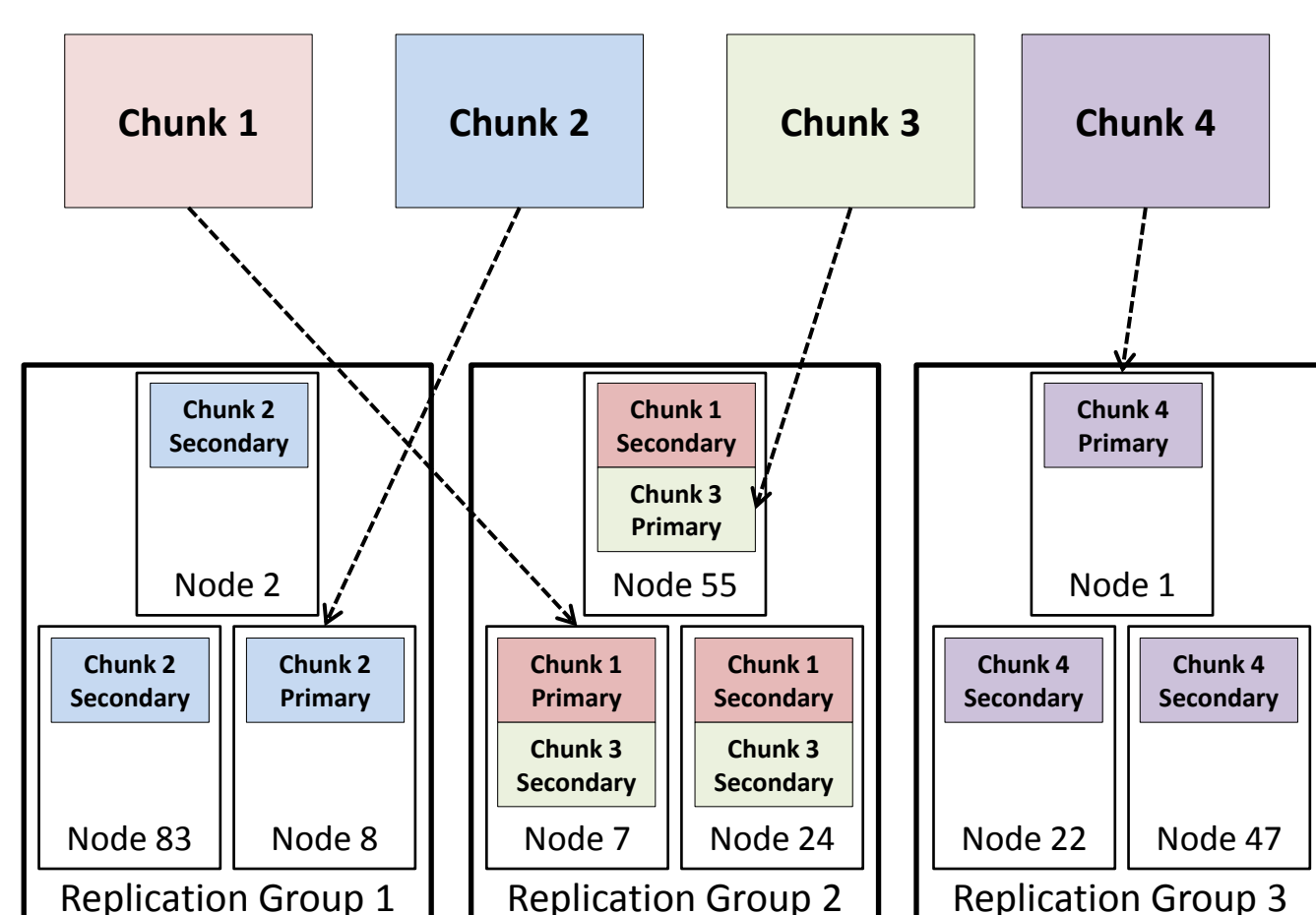
3

## RAMCloud Implementation



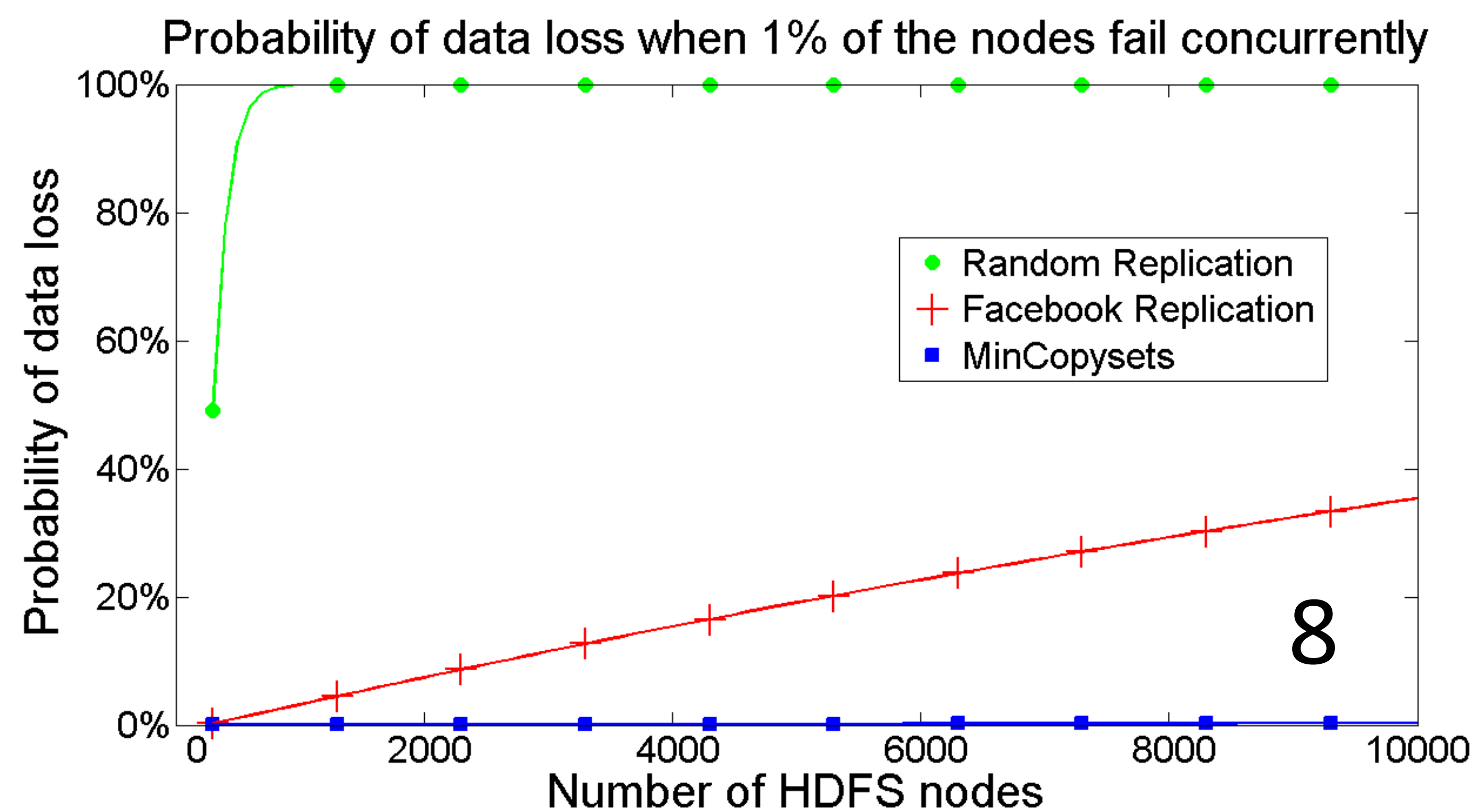
7

## Solution: MinCopysets



4

## Facebook's Replication is not General Purpose



8