

# BLACK-BOX LOCALIZATION OF STORAGE PROBLEMS IN PARALLEL FILE SYSTEMS

Michael P. Kasick, Rajeev Gandhi, Priya Narasimhan (CMU)

## MOTIVATION

### Goals:

- To apply black-box problem diagnosis to GPFS
- To inject real storage-stack faults in a multi-tier storage cluster
- To localize faults to different storage layers: disk arrays, storage controllers, server nodes

### Insights:

- LUN behavior similar for most I/O requests, but degraded arrays exhibit sub-optimal performance
- Peer-similarity: Fault-free LUNs exhibit similar performance metrics while faulty LUNs exhibit dissimilarities
- Peer-comparison of metrics identifies anomalous LUNs, can localize to array, controller, or server by LUNs affected

## EXPERIMENTAL METHODS

### Instrumentation:

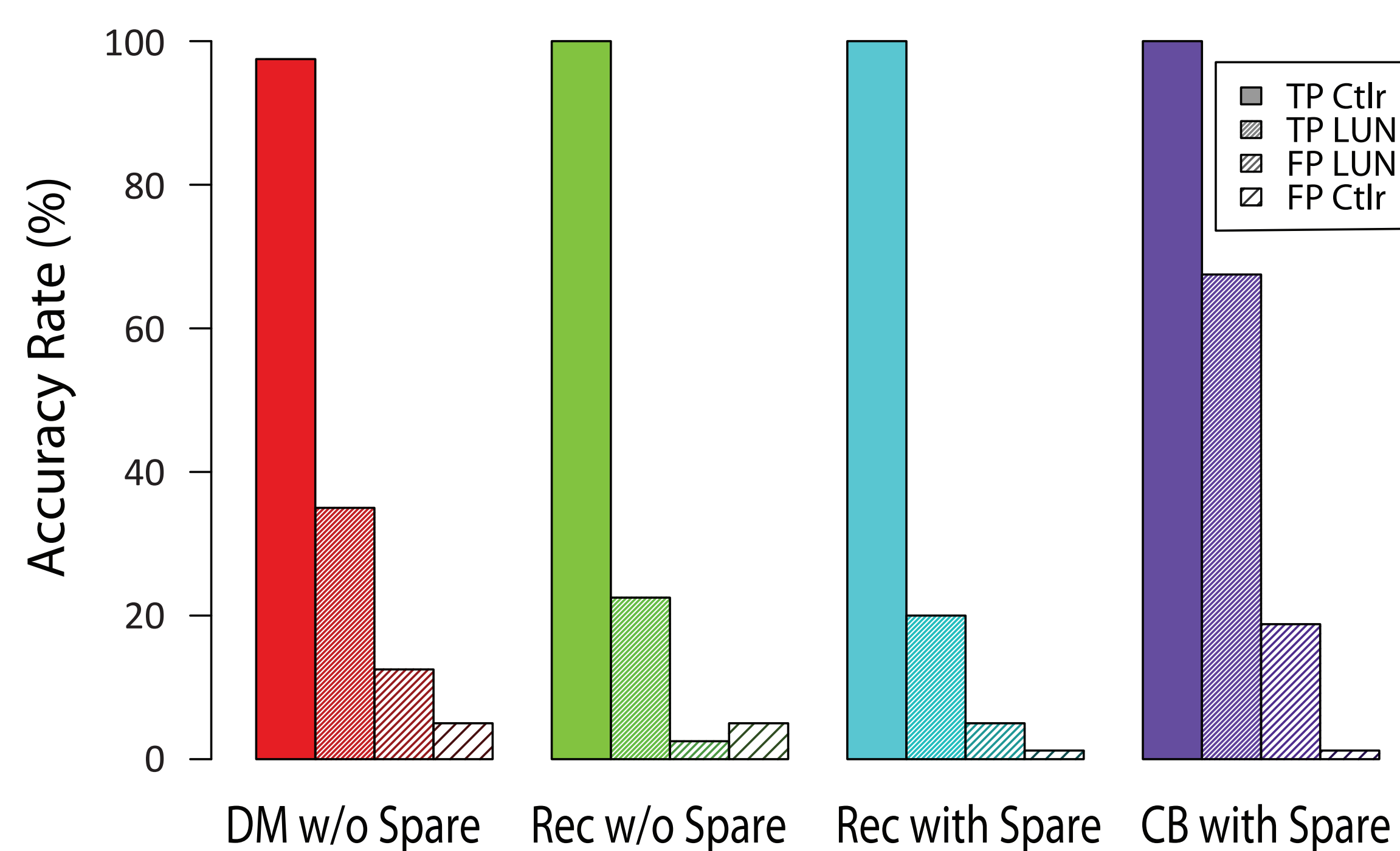
- LUN performance metrics gathered from all NSD servers
- Storage metrics of interest: throughput, latency
- Workloads:
  - ddw & ddr: Large I/O requests, saturating workload
  - iozone & iozone: Workload transitions, fsync
  - postmark: Metadata-heavy, small reads/writes
- Faults injected: Single disk failures with or without hot spares

## DIAGNOSTIC ALGORITHM

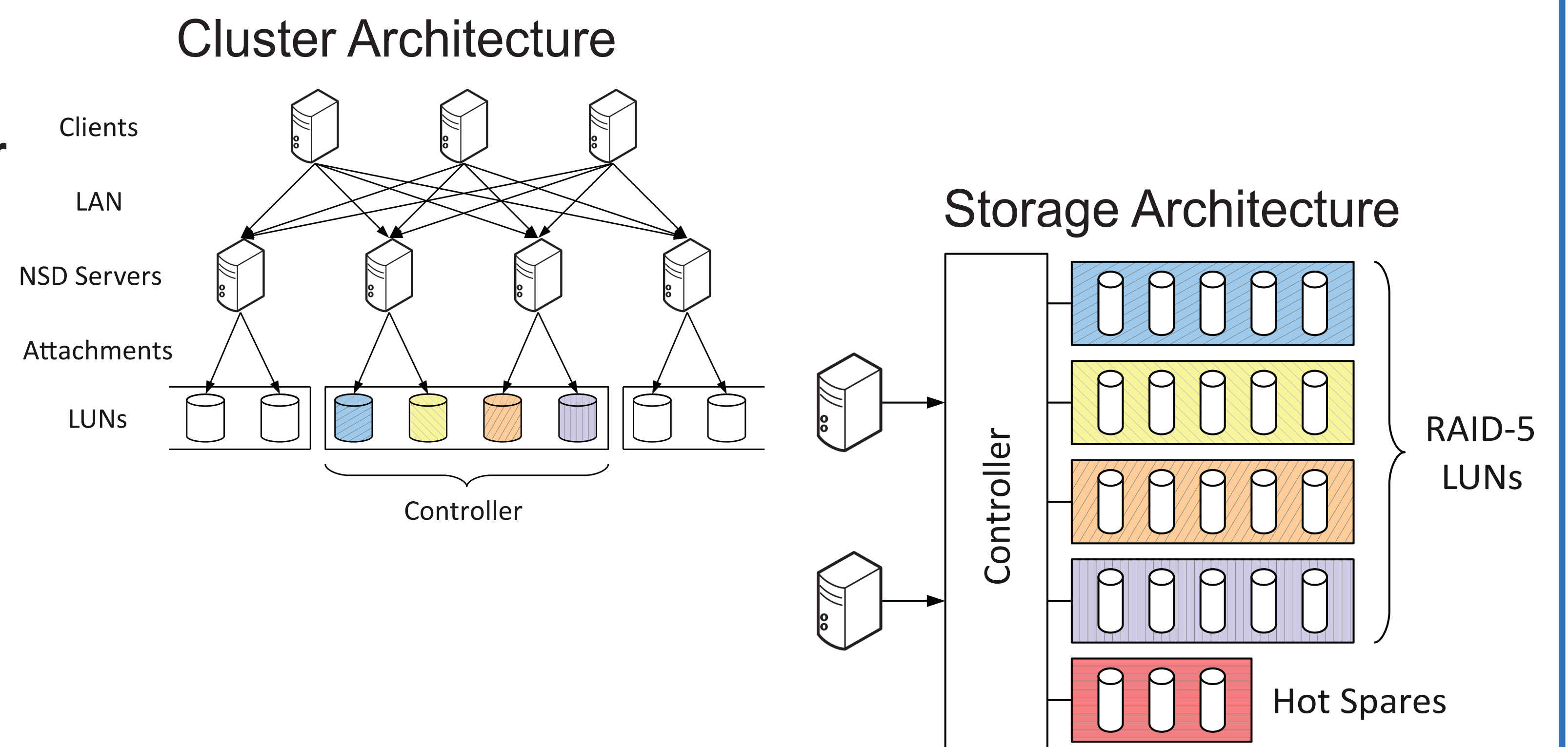
### Detect anomalous LUNs

- Peer-compare performance metric CDFs across LUNs
- Automatically identifies which LUNs are anomalous
- Threshold selection: Fault-free training session
- Localize fault to LUN, controller, or node
  - Identifies misbehaving resources common with LUNs
  - Enables manual inspection of faulty resources

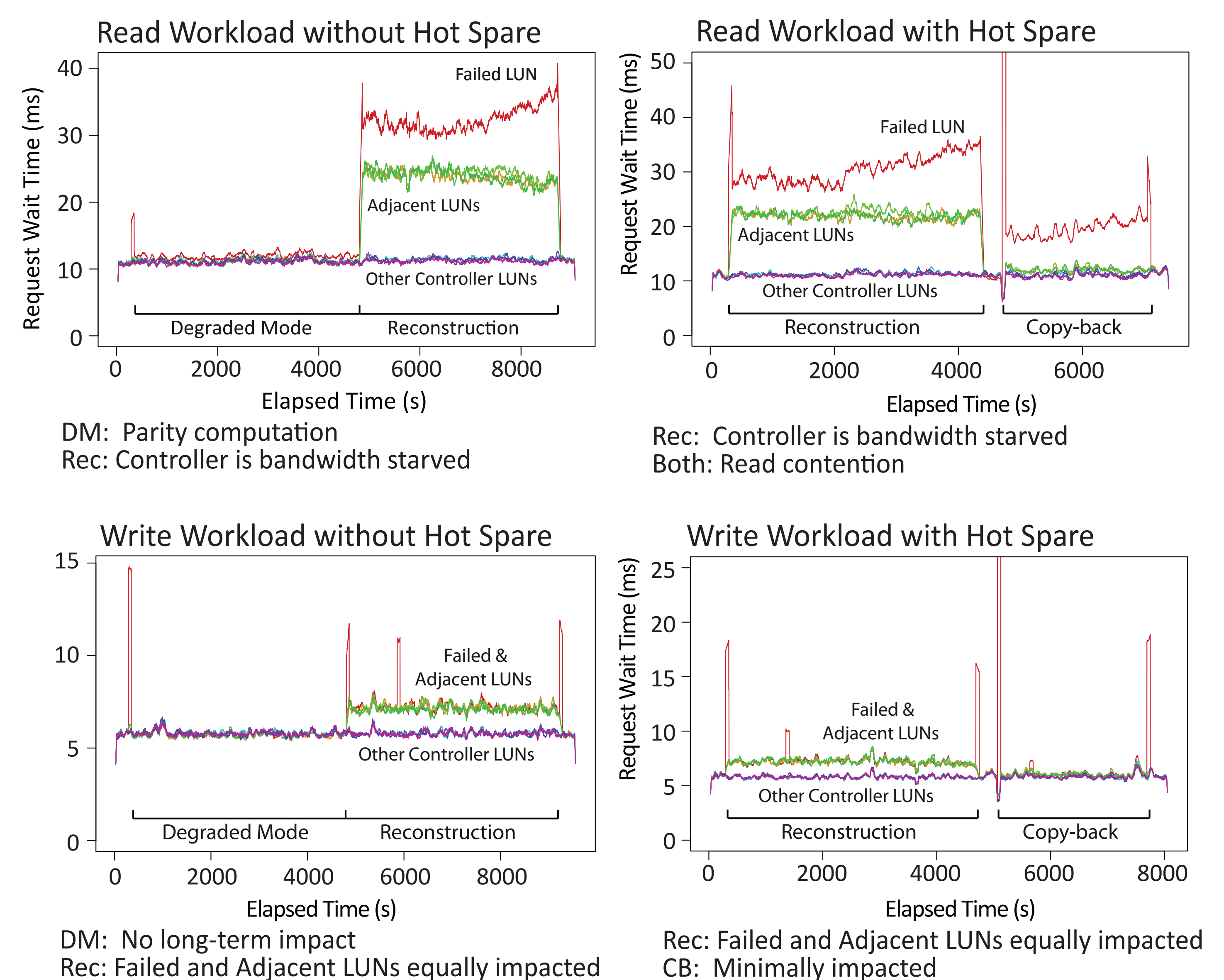
## RESULTS



## ARCHITECTURE



## BEHAVIOR OF DISK FAILURES



## PRODUCTION CLUSTER DIAGNOSIS

- Intrepid Storage Cluster (Argonne National Laboratory):**
  - 128 GPFS NSDs, 1152 LUNs across 16 controllers, 4.5 PB
  - Instrumented and analyzed over three months
- Diagnosed Problems:**
  - Eight independent disk failures
    - 5 controller-failed, 3 operator-failed
    - Diagnosed: 7 in degraded mode, 7 in reconstruction
  - Twelve enclosure reboots
    - All disks manually failed and reinserted
    - Detected disk failures in 10 of 12 events
  - Three lost attachments
    - Attachment between controller and server is lost
    - Diagnosed for all three events

