

# SUSTAINABLE PREDICTIVE STORAGE MANAGEMENT

DAVID ESSARY  
DEPT. OF COMPUTER SCIENCE  
UNIVERSITY OF PITTSBURGH





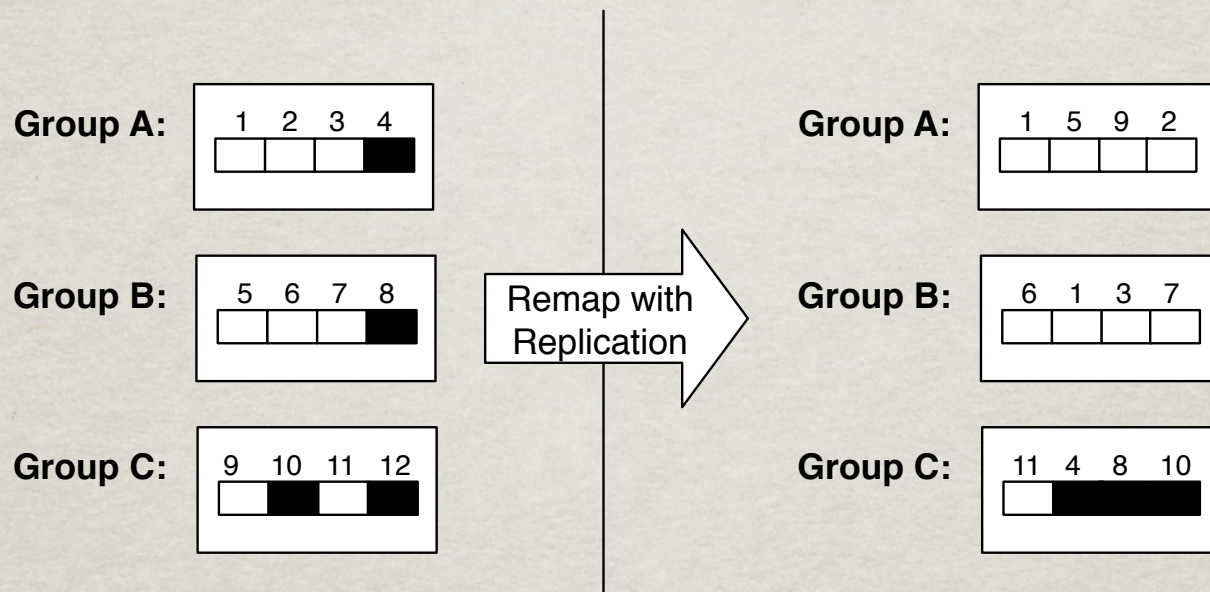
# INTRODUCTION

- ✻ Disk vs CPU
- ✻ Storage systems' power becoming more critical
- ✻ Rate of data generation is alarming
- ✻ No “silver bullet”
  
- ✻ Goal: Dynamic (and *adaptive*), sustainable predictive grouping engine
- ✻ Group = Disk track






# MOTIVATING EXAMPLE



**Block Access Pattern:** 1,5,9,2,6,1,3,7,11

 Used block  
Free block

**Group Access Pattern**  
Prior to Remapping: A,B,C,A,B,A,B,C

**Group Access Pattern with**  
Remapping and Replication: A,B,C



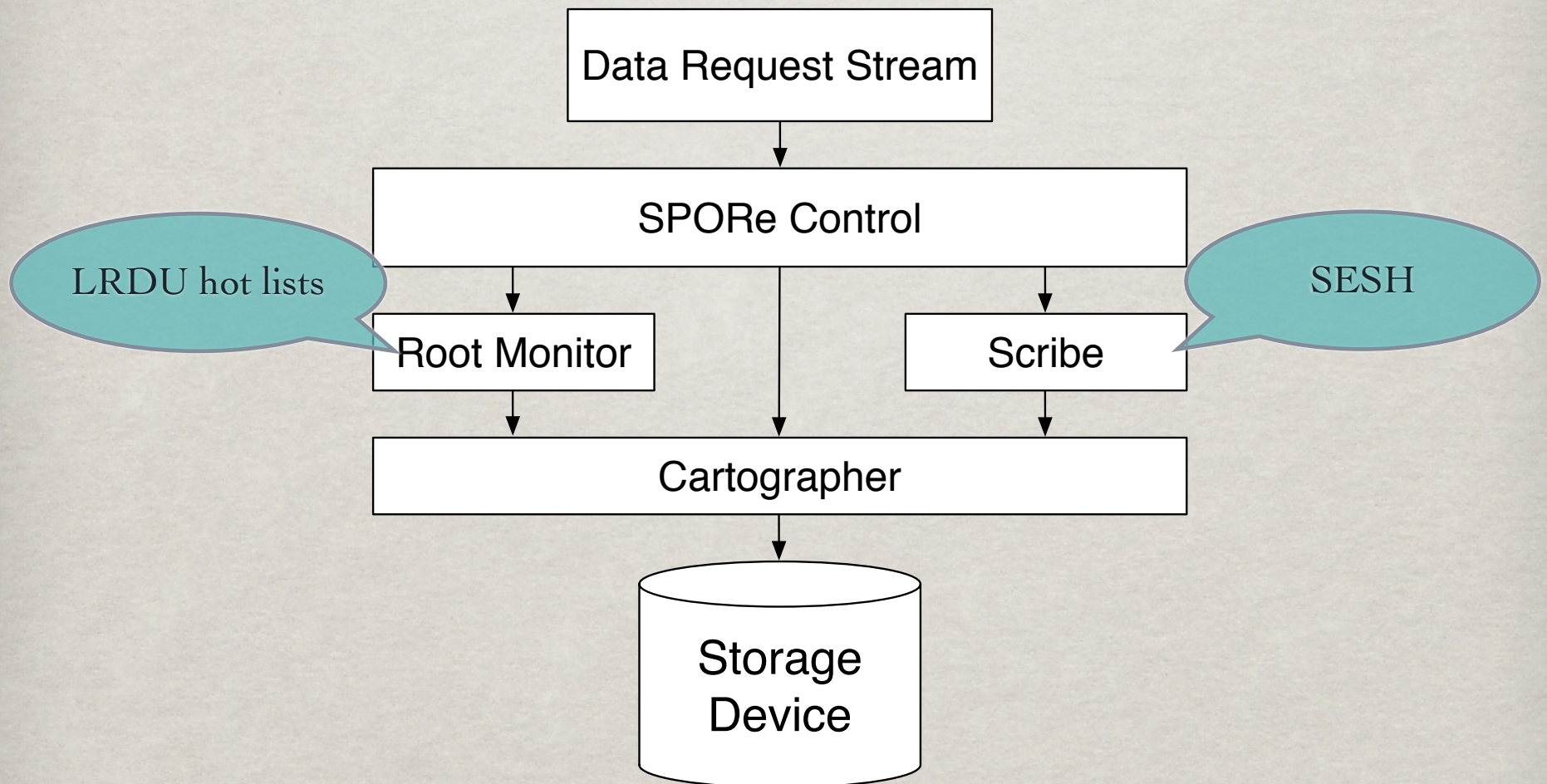
# CHALLENGES

- ✻ When are predictions performed?
- ✻ How are predictions made?
- ✻ How is predictive metadata gathered?
- ✻ Where are predictions to be located?
- ✻ How are predictions used?





# SPORE





# SUPERGROUPS

**Group 1 (Root: A)**

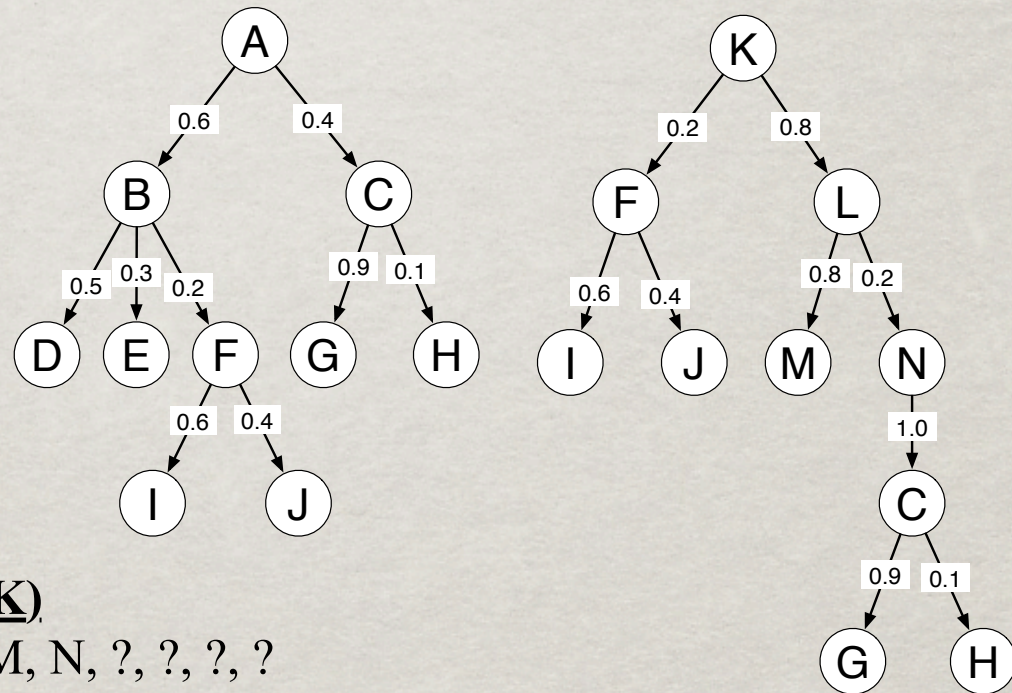
A, B, C, G, D, E, F, I

**Group 2 (Root: K)**

K, L, M, F, N, C, G, I

**Supergroup (Roots: A and K)**

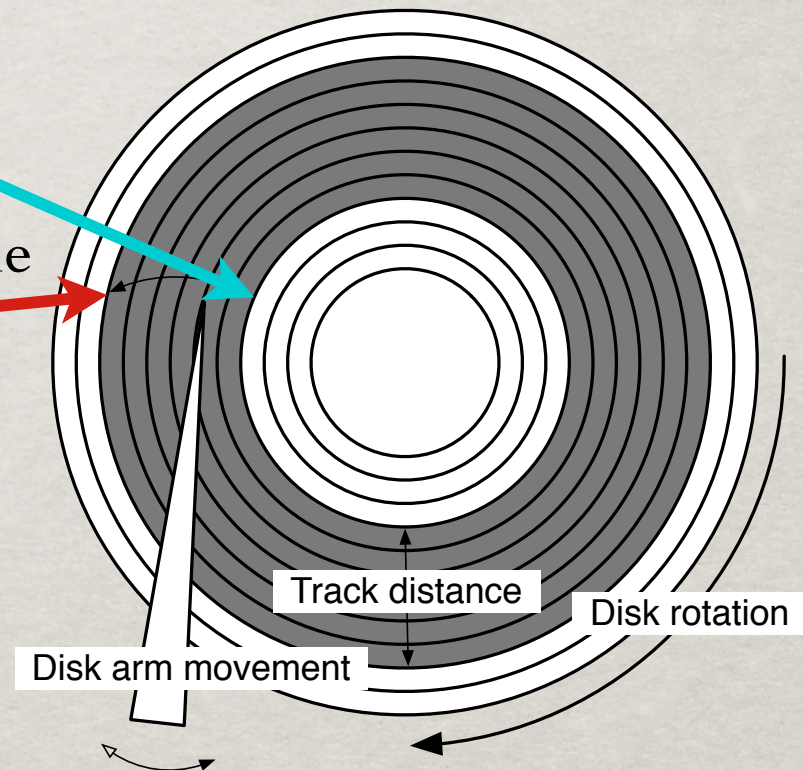
A, B, C, G, D, E, F, I, K, L, M, N, ?, ?, ?, ?





# GROUP SCANNING

- ❁ Does the *offending block* exist between the *current* disk head location and the *target* location
- ❁ If so, we instead switch to the predictive group containing the offending block





# REDUCING COMMIT OVERHEAD

- ✻ Commit predictions to device *opportunistically*
- ✻ Use only items already in main memory
  - ✻ Avoid additional seeks
- ✻ Avoid *updating* a group if it contains 75% of objects that it “should” contain
  - ✻ We call this percentage the *overlap threshold*





# OVERLAP THRESHOLD

## Group $\alpha$

A, B, C, D, E, F, G

## Group $\alpha'$

A, B, C, D, H, I, J

Overlap: 62.5%

Result: Replace group  $\alpha$  with  $\alpha'$

---

## Group $\alpha$

A, B, C, D, E, F, G

## Group $\alpha''$

A, B, C, D, E, F, J

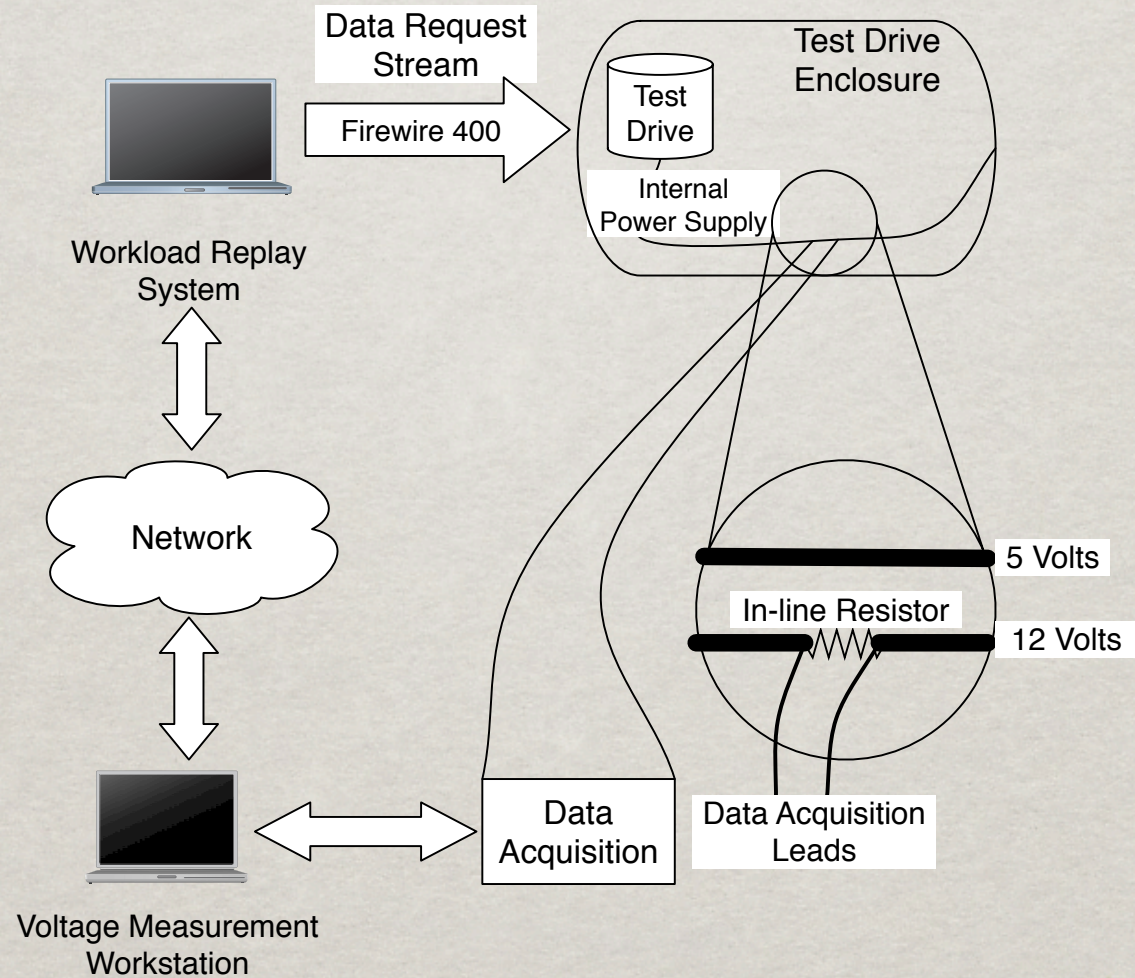
Overlap: 87.5%

Result: Abort update and keep group  $\alpha$



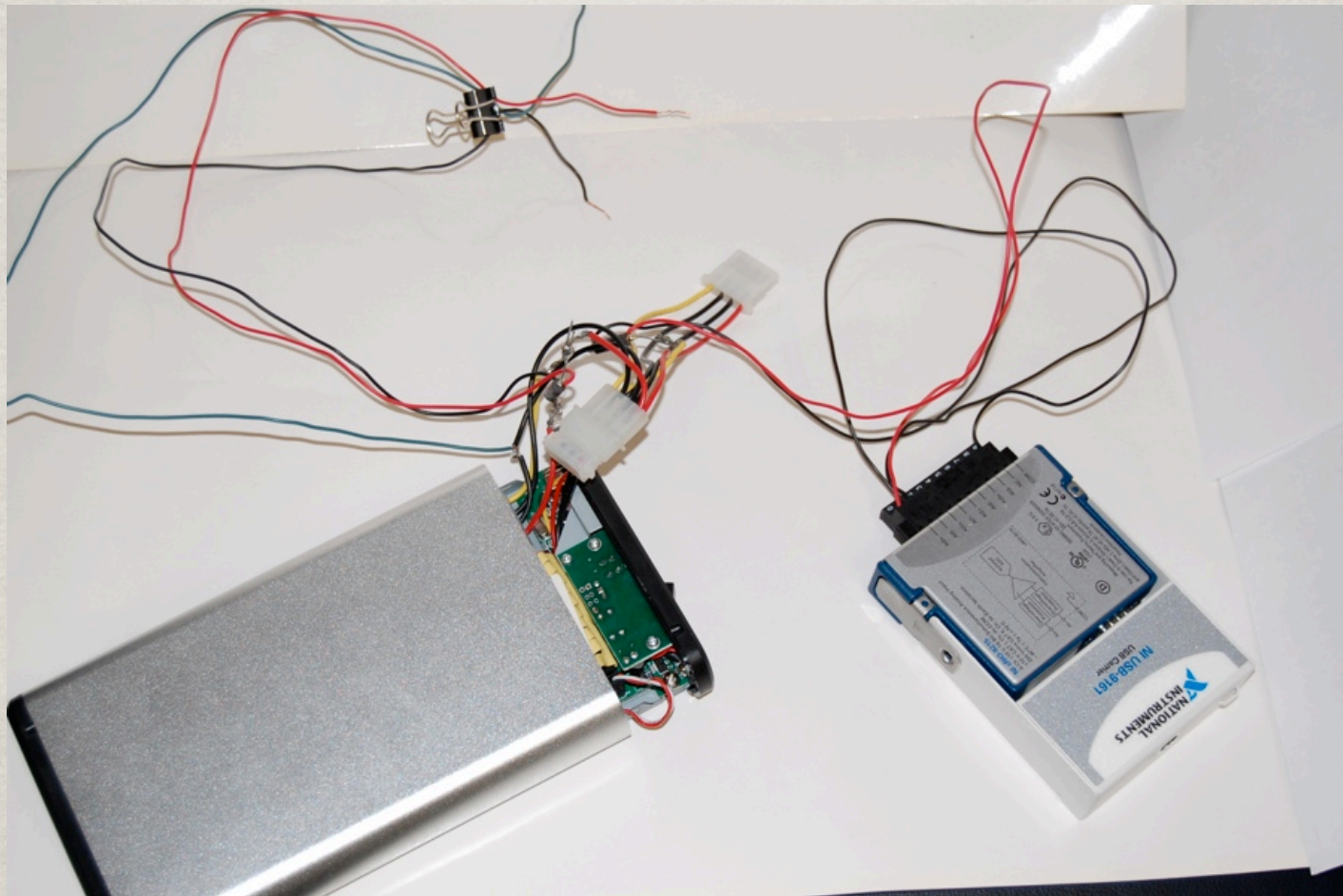


# HARDWARE-BASED VALIDATION





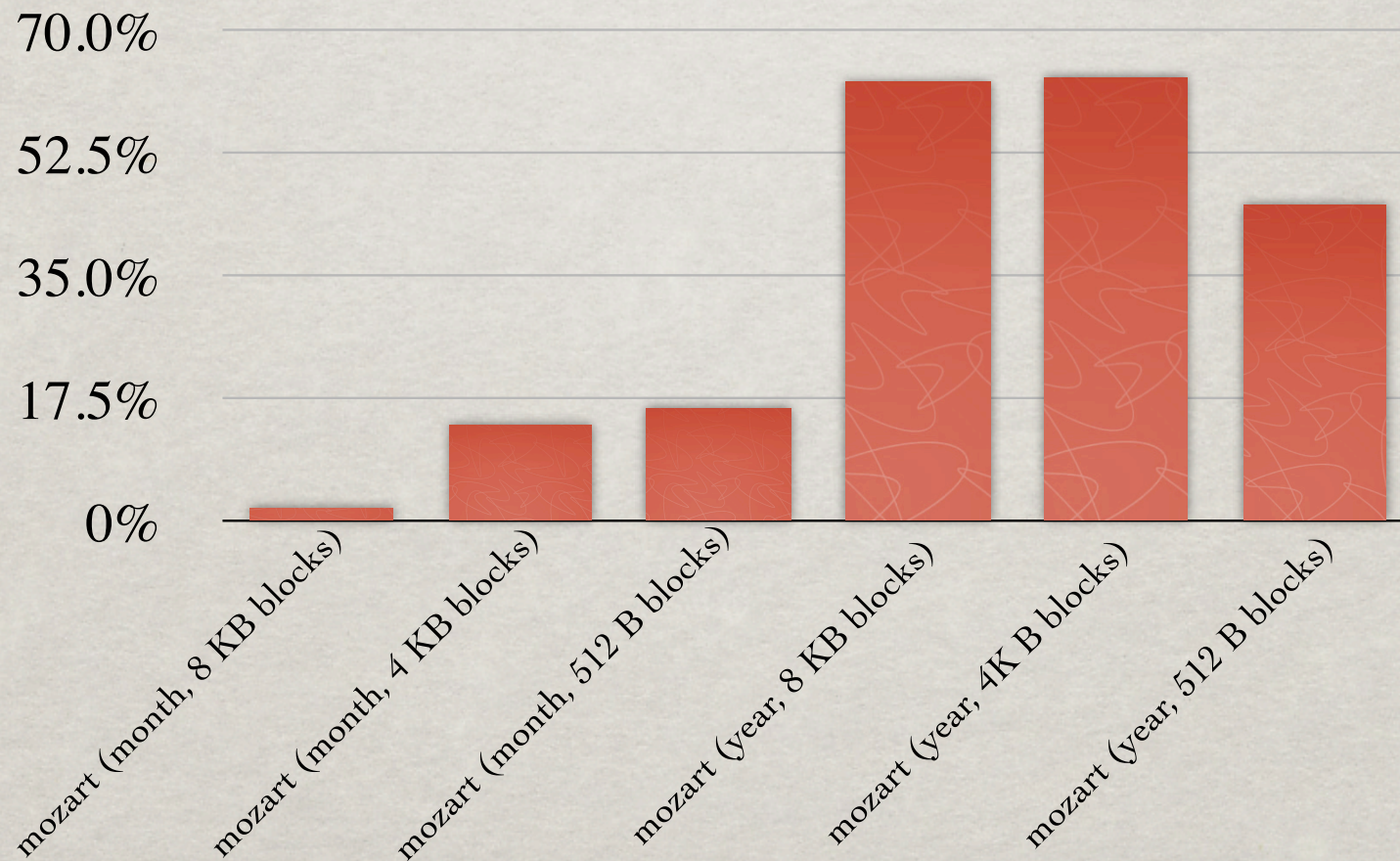
# HARDWARE VALIDATION DAQ AND EXTERNAL HD





# VALIDATION RESULTS WD DRIVE

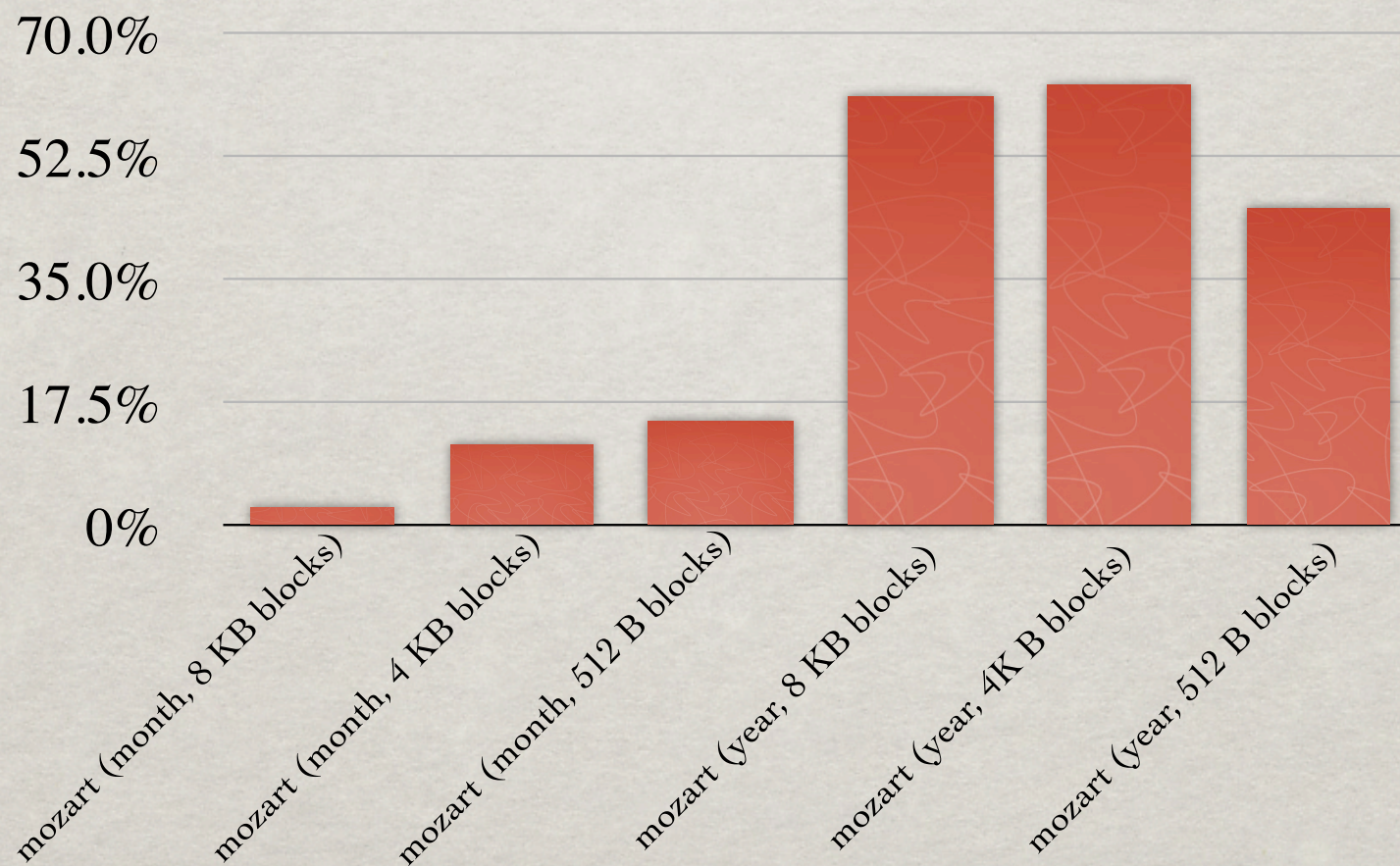
Percentage Latency Reduction (Group Size: 8K)





# VALIDATION RESULTS WD DRIVE

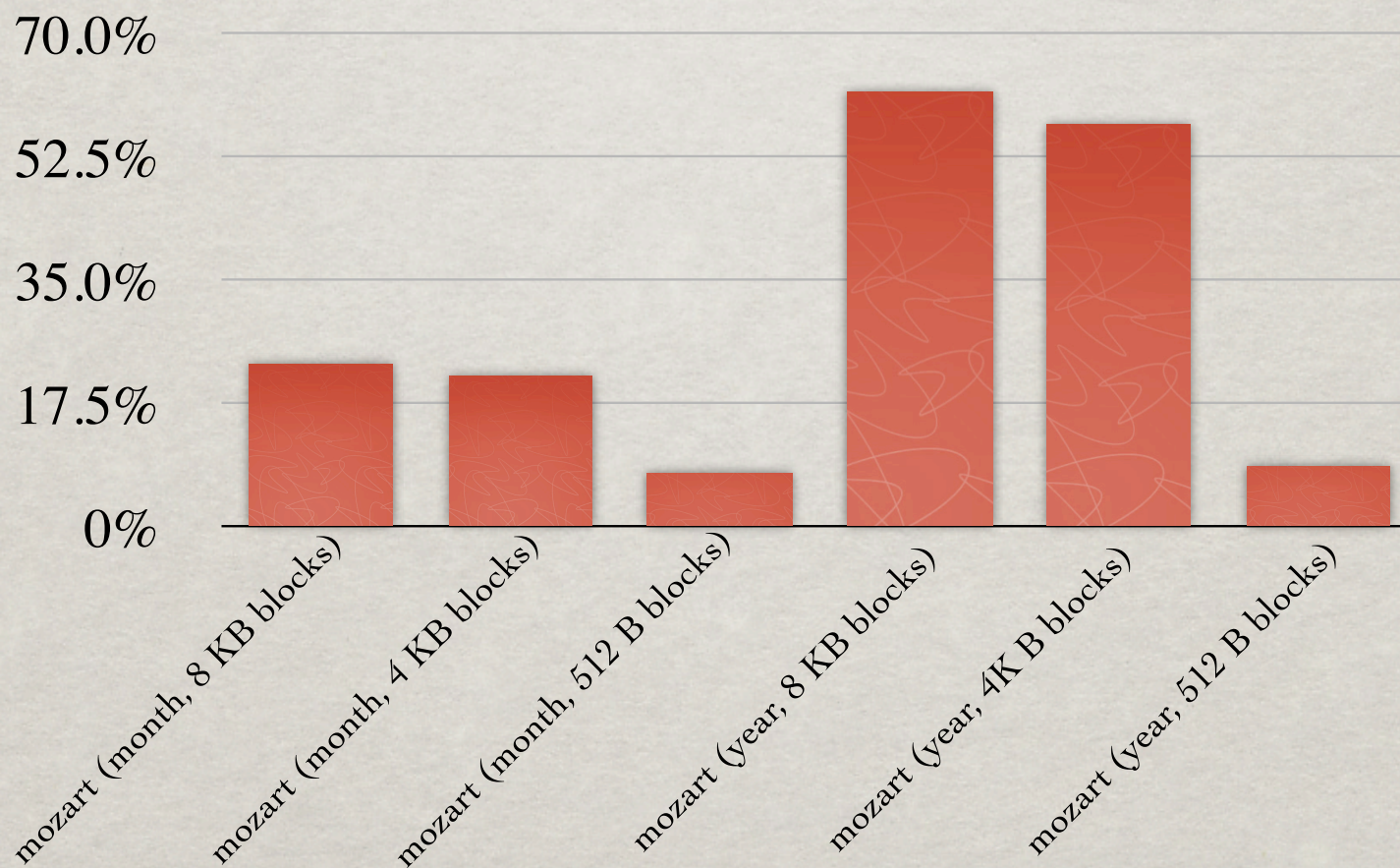
Percentage Energy Reduction (Group Size: 8K)





# VALIDATION RESULTS WD DRIVE

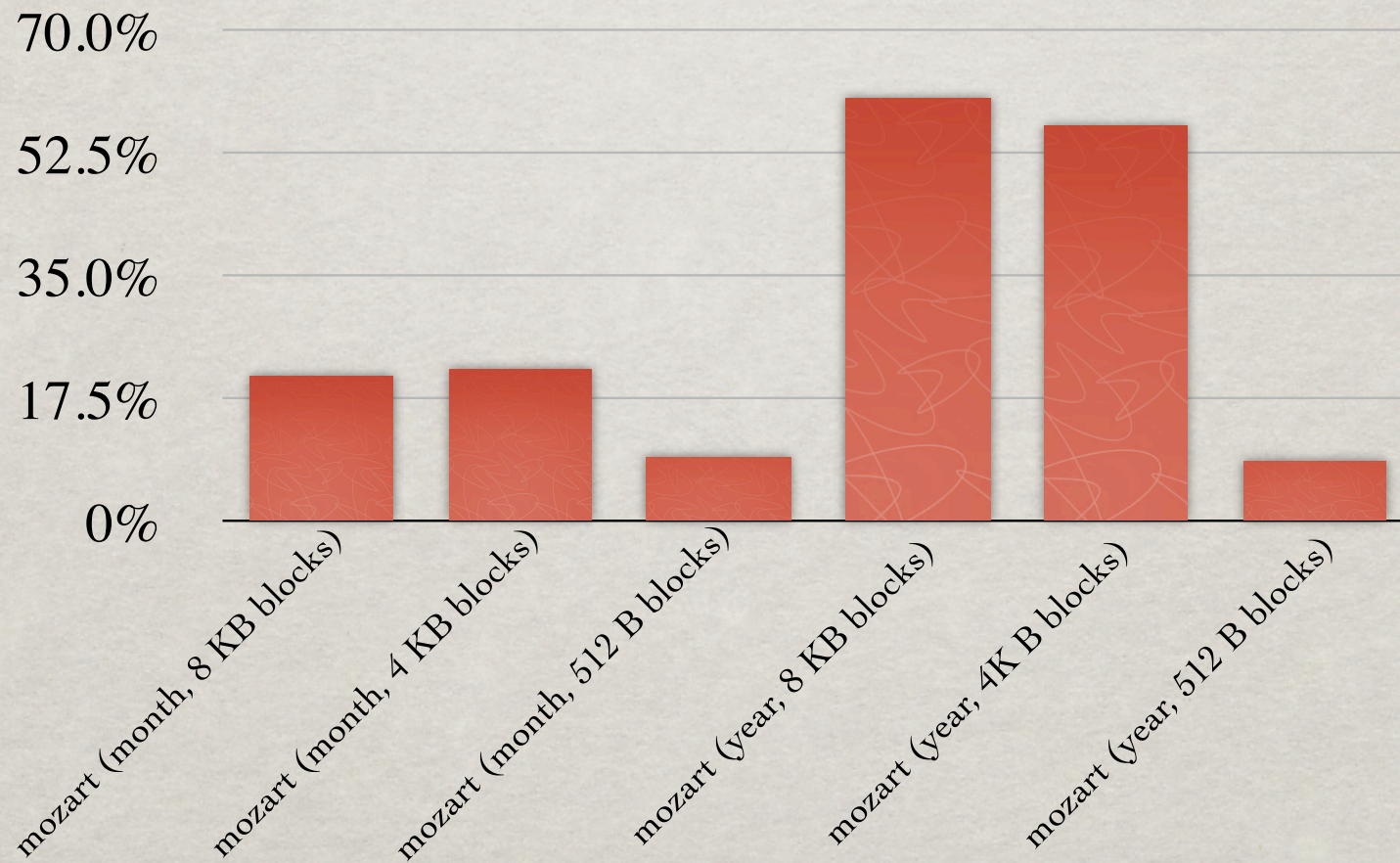
Percentage Latency Reduction (Group Size: 1K)





# VALIDATION RESULTS WD DRIVE

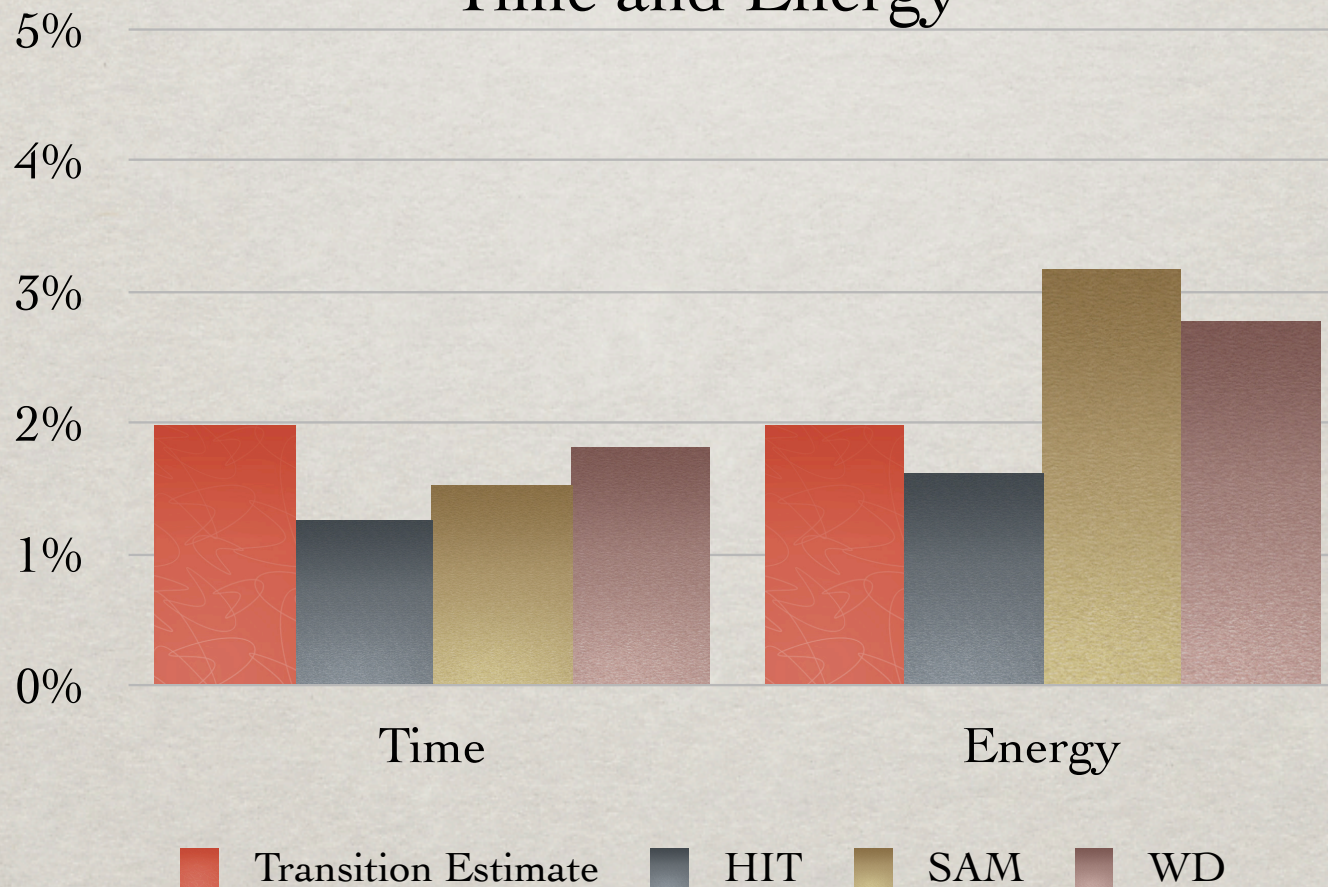
Percentage Energy Reduction (Group Size: 1K)





# VALIDATION RESULTS

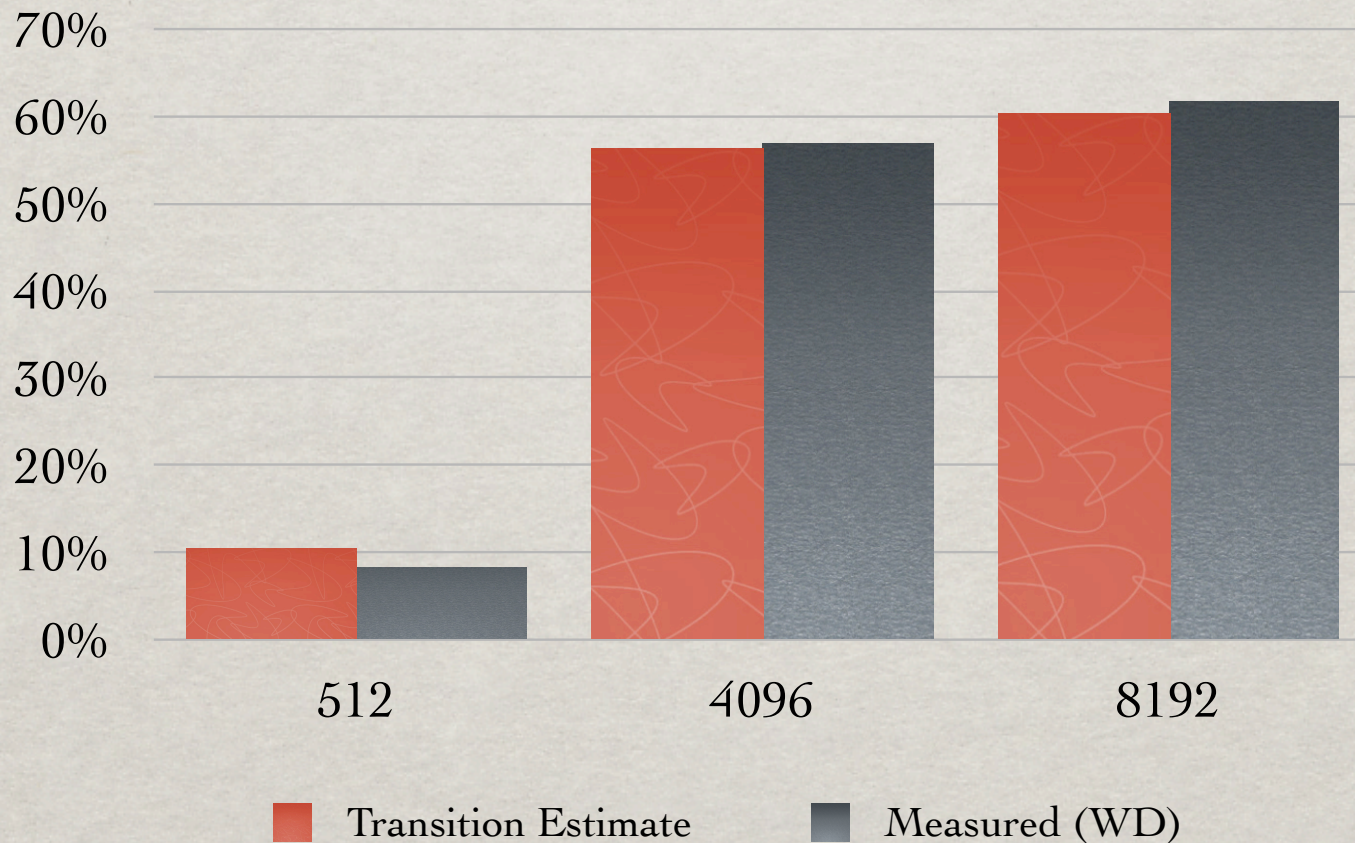
## Percentage Average Reduction for Time and Energy





# VALIDATION RESULTS

## Percentage Reduction of Energy by Block Size





# SPORE CONCLUSIONS

- ✱ Opportunistic, dynamic, sustainable
- ✱ Replicates data on the fly (no warm-up period)
- ✱ Simultaneously reduces
  - ✱ Track distance (up to 80% reduction)
  - ✱ Track seeks (up to 65%)
  - ✱ Latency due to mechanical movement (up to 63%)
  - ✱ Energy due to mechanical movement (up to 61%)
- ✱ Strong correlation between seek reduction and energy and latency reduction
- ✱ Latency and energy results validated by live hardware

