RobuSTore

Distributed Storage System Providing Robust, High Performance

Huaxia Xia and Andrew Chien

University of California, San Diego

Mar 16, 2006

Motivation

- **Large-Scale Scientific Applications**
  - **Large data sets: TB~PB**
  - **Large size per access: GB ~ 10s GB**
  - **Large number of distributed users: 100s**
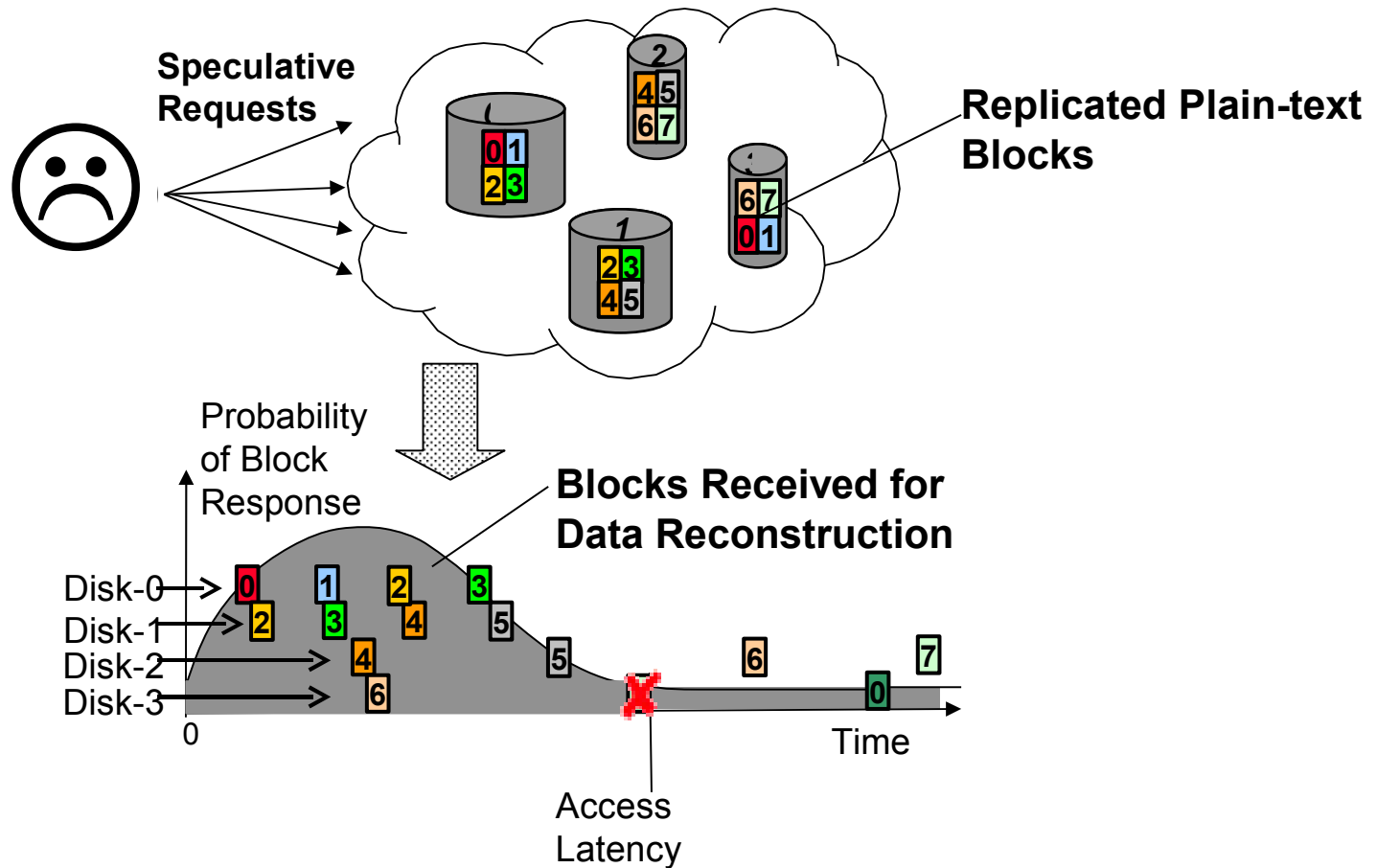
- **Want: Distributed Storage with**
  - **High Bandwidth**
    - **For Large-Size Data Accesses (Read-Dominated)**
  - **High Robustness**
    - **Robustness: Low Variability on Access Latency**

- **Challenge:**
  - **High Variability of Individual Disk Performance**
    - **Disk Rotation**
    - **Disk Heterogeneity**
    - **Network Heterogeneity**
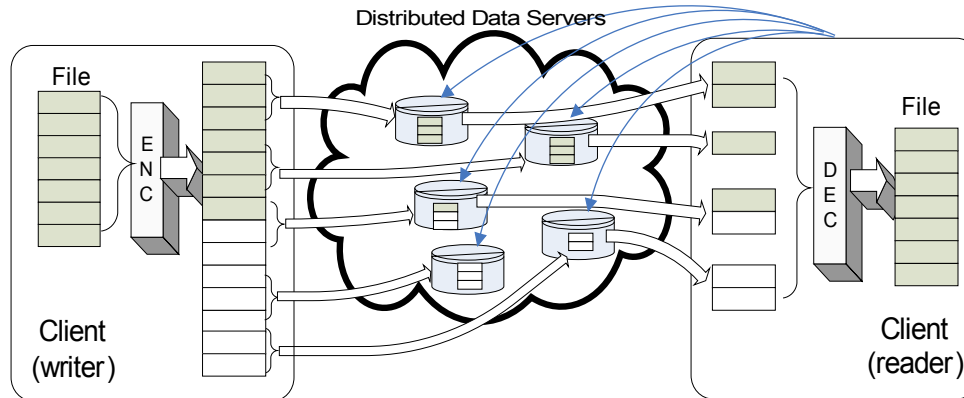   - **Dynamic Shared Accesses**
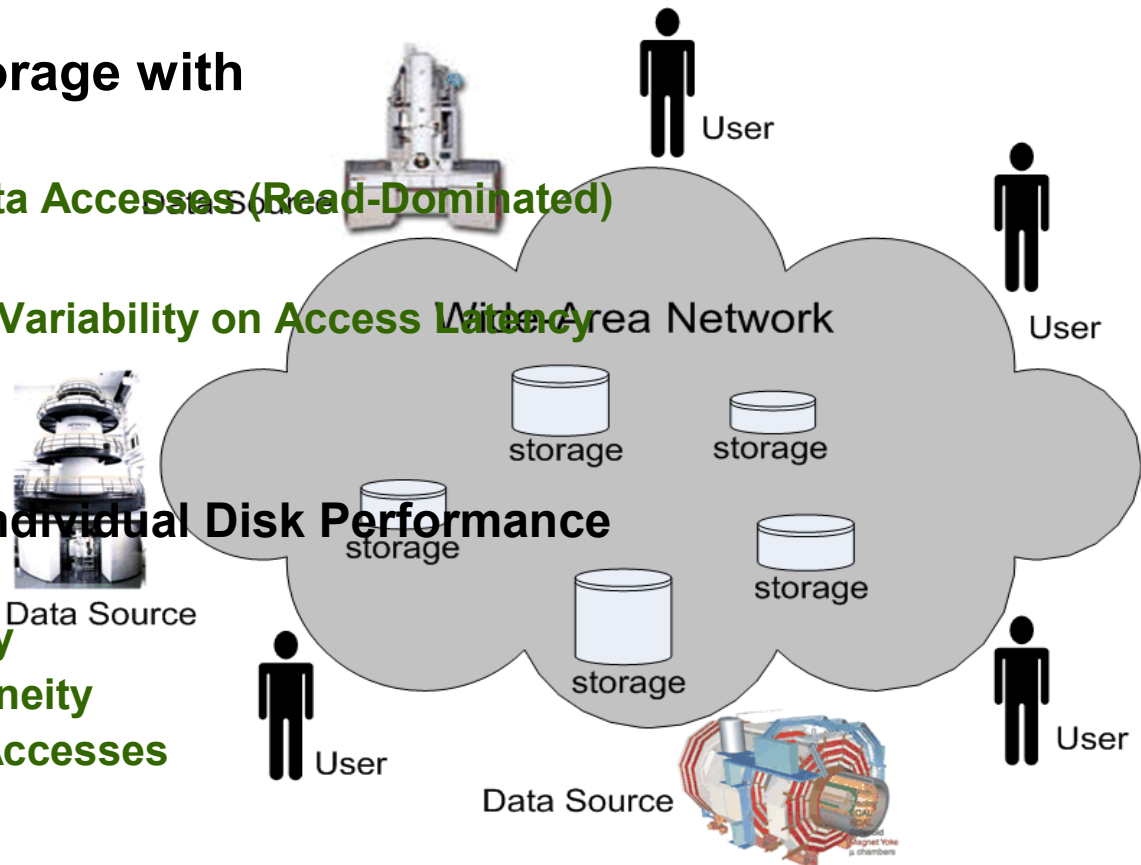
Why Existing Systems Cannot?

Our Approach: RobuSTore

- **Key Idea: Erasure Codes + Speculative Accesses**

- **Erasure Codes**
  - **Tolerate the Late Arrived Blocks**
    - **Considering late arrived blocks as lost blocks**
  - **Allow to Reconstruct the Original Data from the First Returned Subset of Blocks**
    - →**Reduce the dependence of the request on any individual disk**
- **Speculative Accesses**
  - **Request Redundant Data Blocks**
  - **Once Received Enough Blocks, Cancel the Ongoing Requests**

4

RobuSTore Improves Bandwidth and Robustness

5
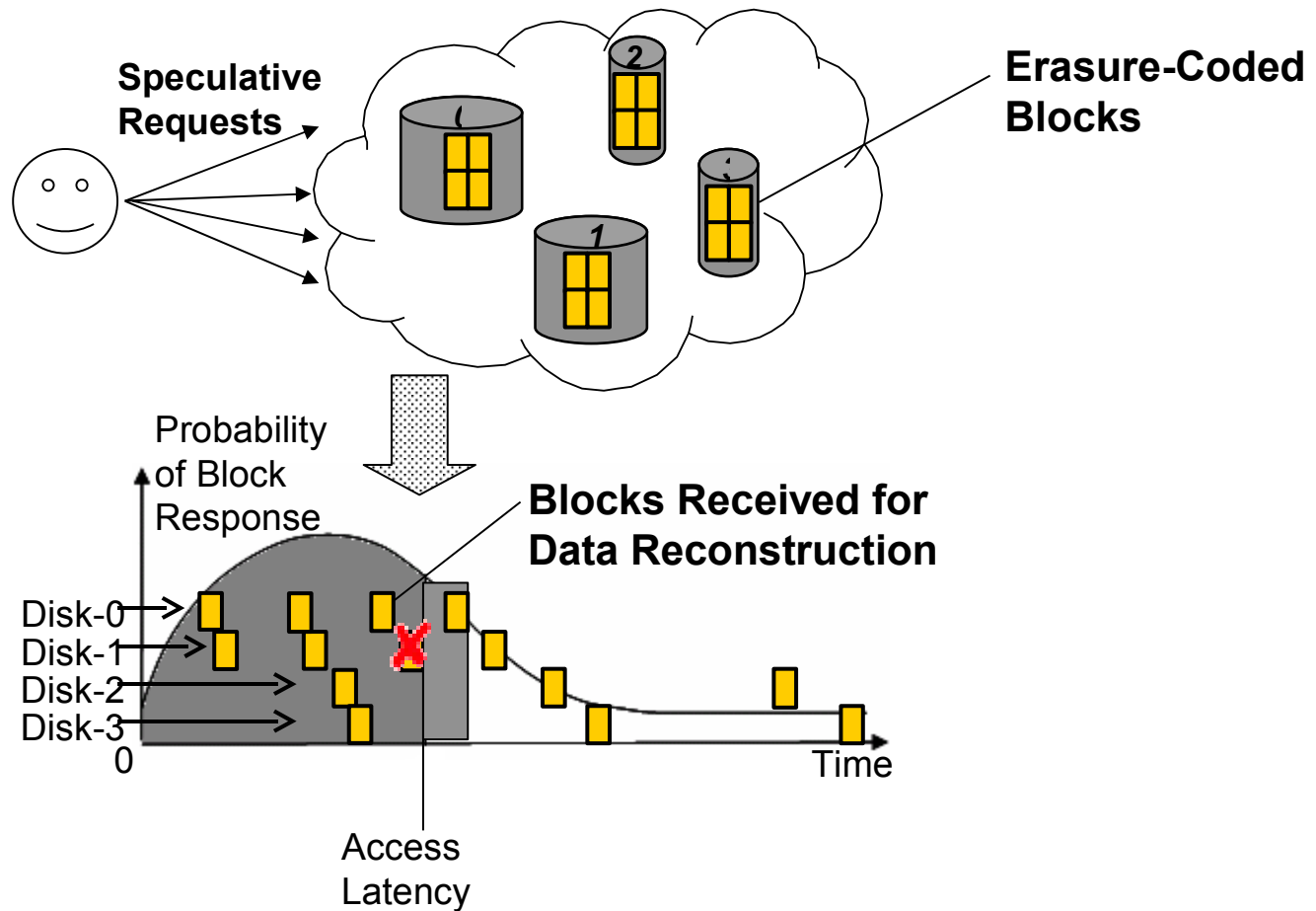
Compare RobuSTore and Traditional Parallel Storage
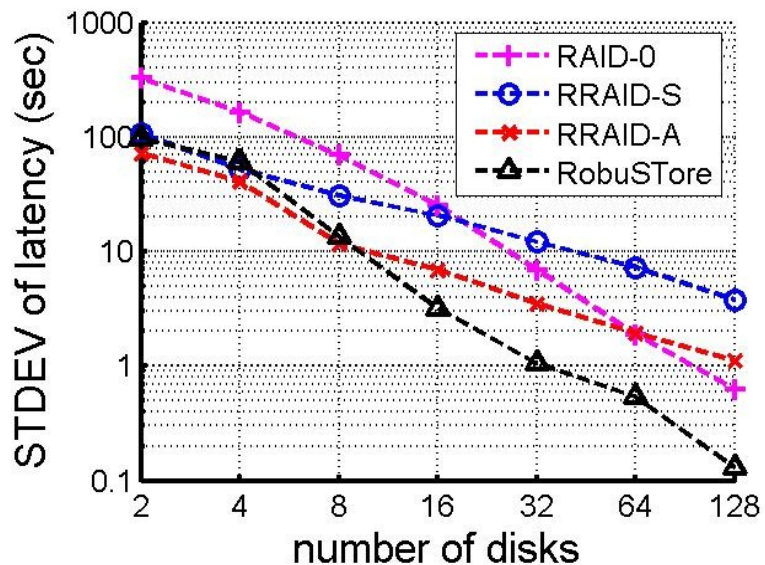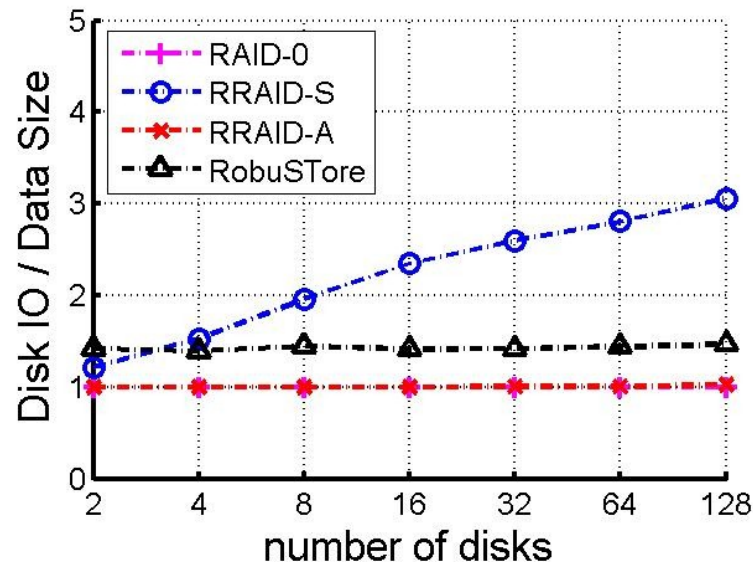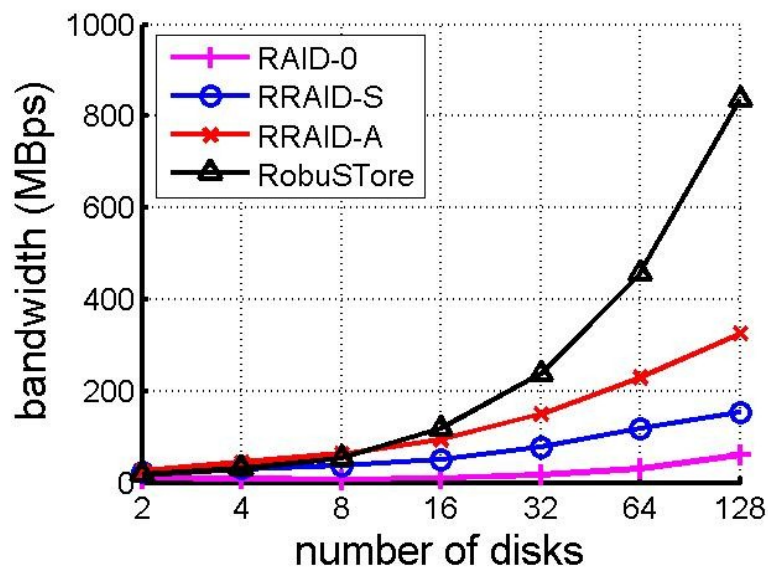
For > 8 Disks

- **RobuSTore Achieves Highest Bandwidth and Best Robustness**
  - **On 64 disks, RAID-0, RRAID-S, RRAID-A and RobuSTore deliver 31, 117, 228, 459 MBps bandwidth, with STDEV of latency of 1.9, 7.3, 1.9, and 0.5 seconds respectively**
- **RobuSTore has ~1.5 Disk IO Overhead**

Thank You!