

CS 755 – System and Network Architectures and Implementation

Module 8 – Storage and Replication

Martin Karsten

mkarsten@uwaterloo.ca

Notice

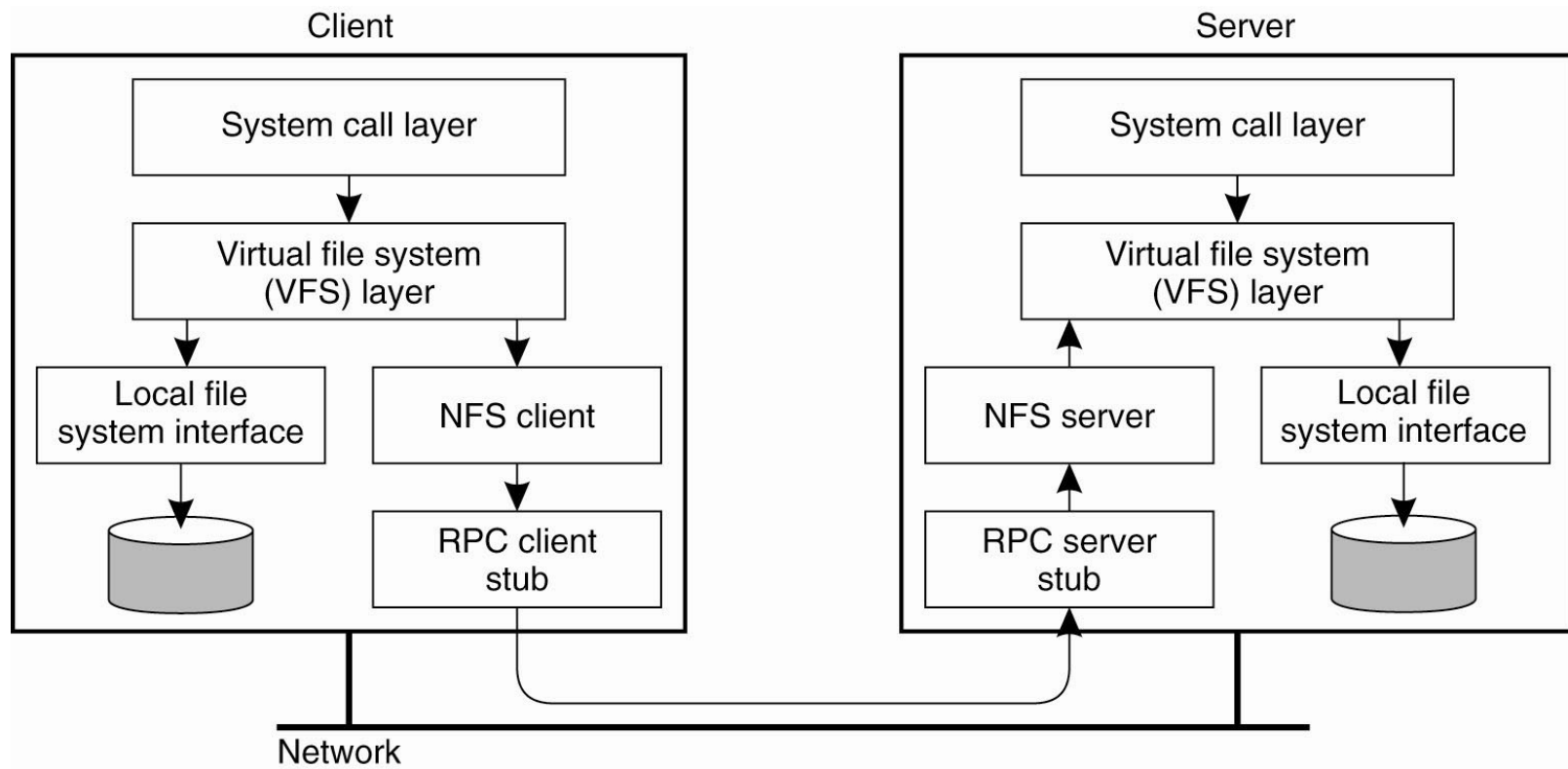
Some figures are taken from third-party slide sets. In this module, figures are taken from the Tanenbaum/van Steen slide set:

Tanenbaum & Van Steen, Distributed Systems: Principles and Paradigms, 2e, (c) 2007 Prentice-Hall, Inc. All rights reserved. 0-13-239227-5

Distributed Storage

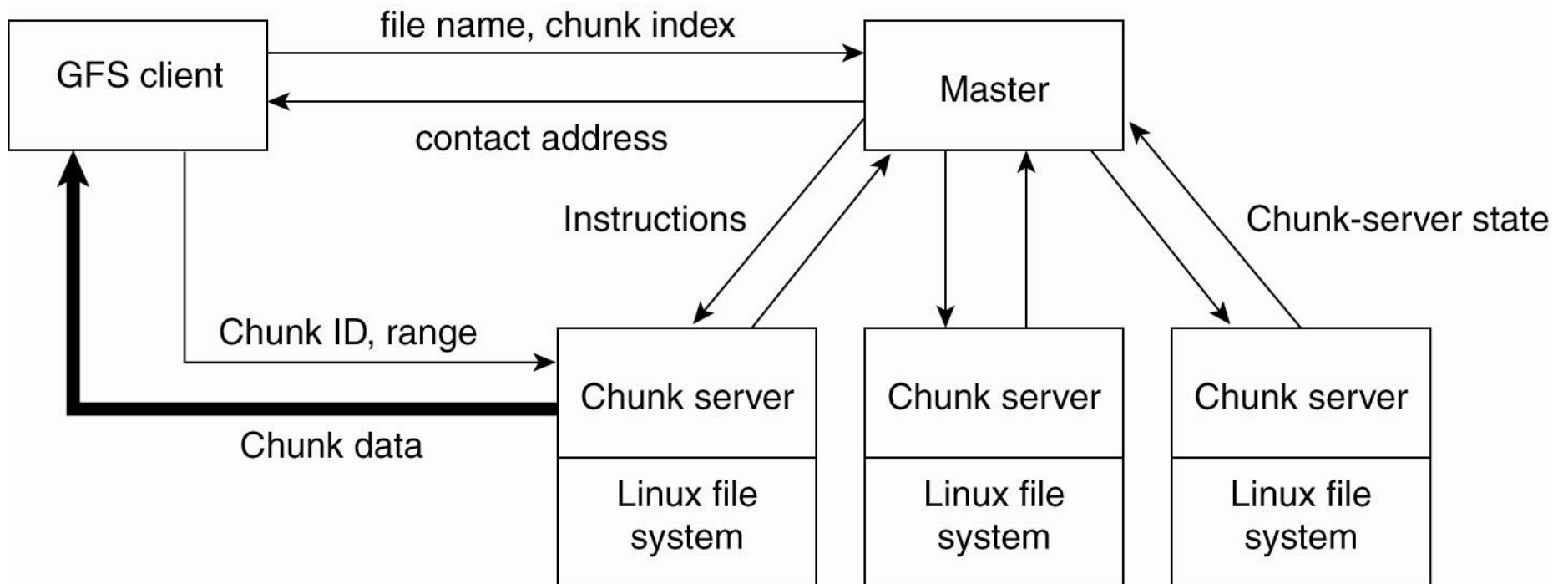
- remote file system
 - dedicated server
 - simplicity & management
- partition data across servers
 - load balancing
 - storage size

UNIX / NFS



Distributed File Service

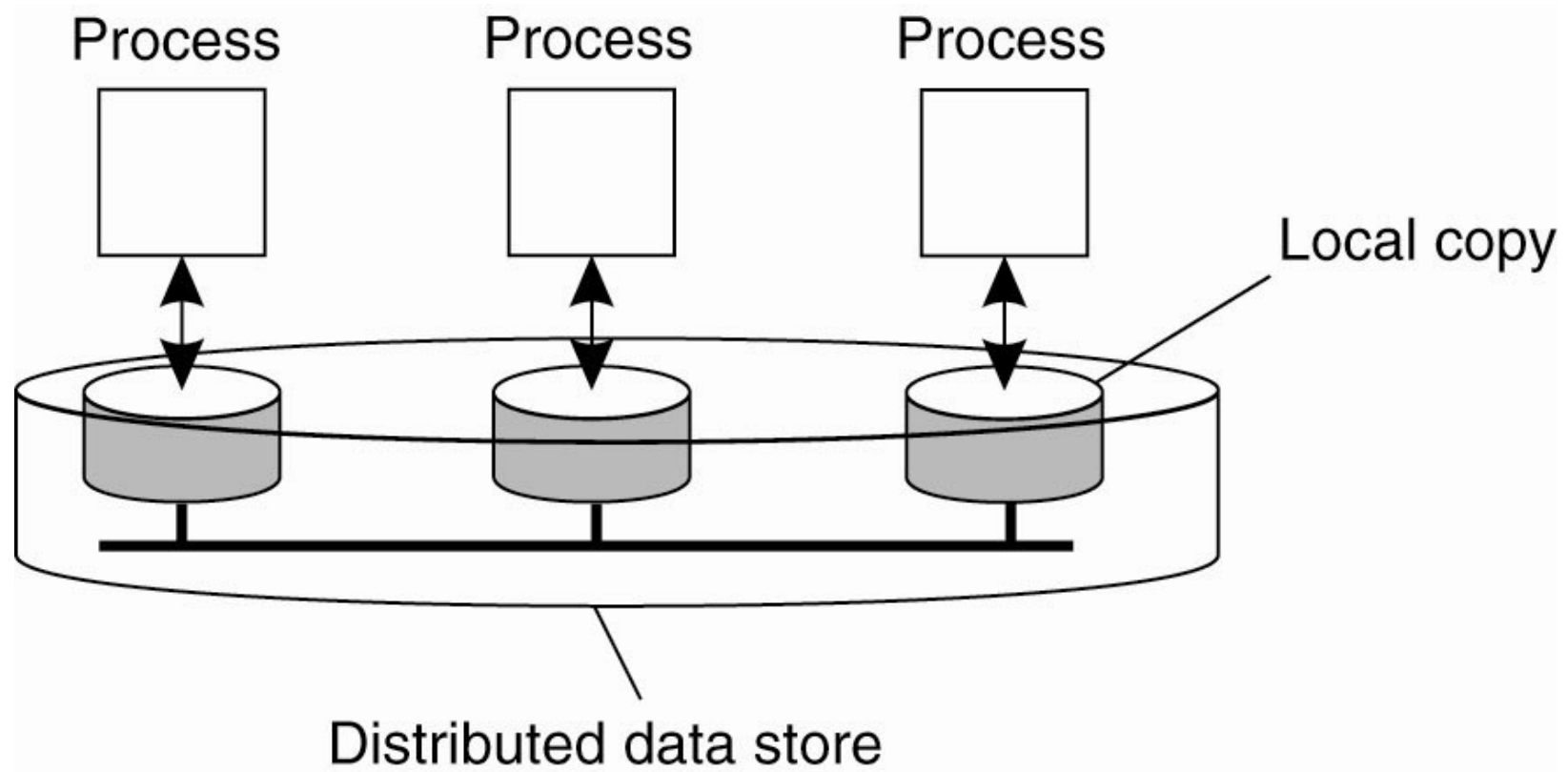
- Google File System - overview



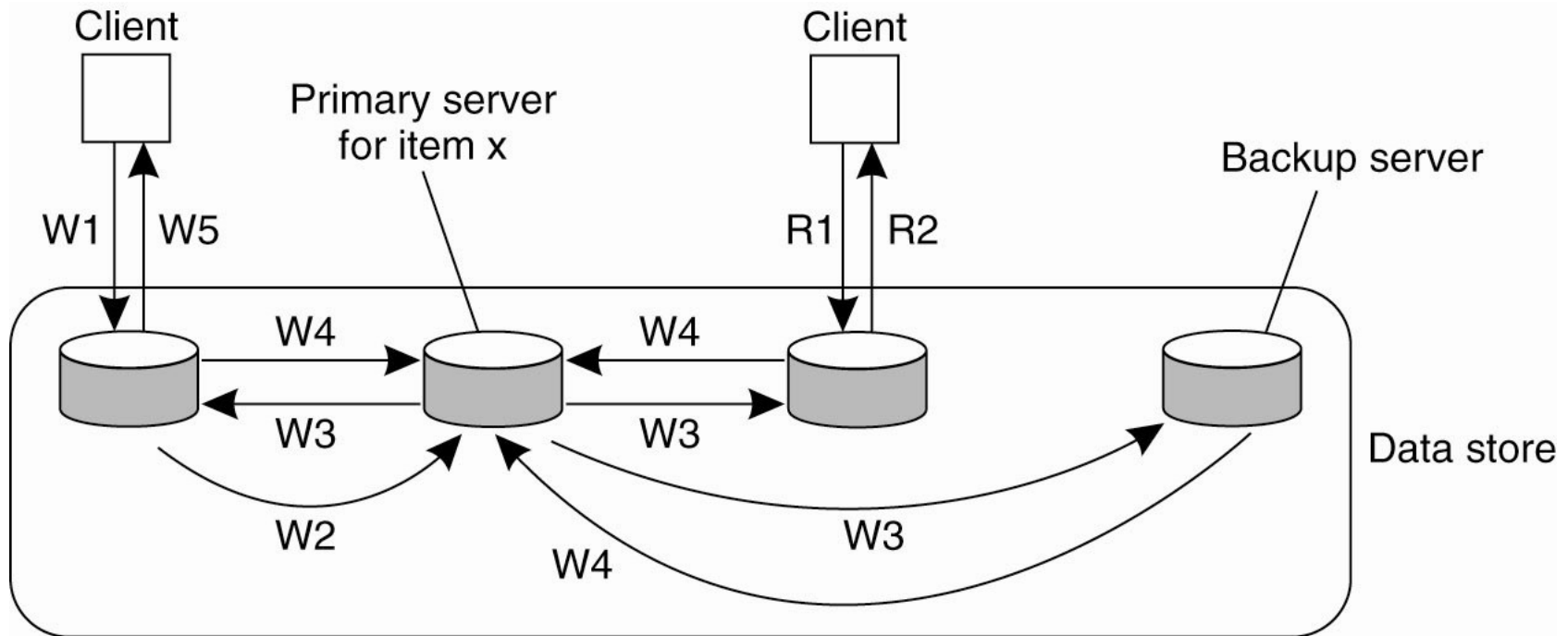
Replication

- Why?
 - availability
 - performance
- Challenges
 - transparency
 - consistency

Conceptual View



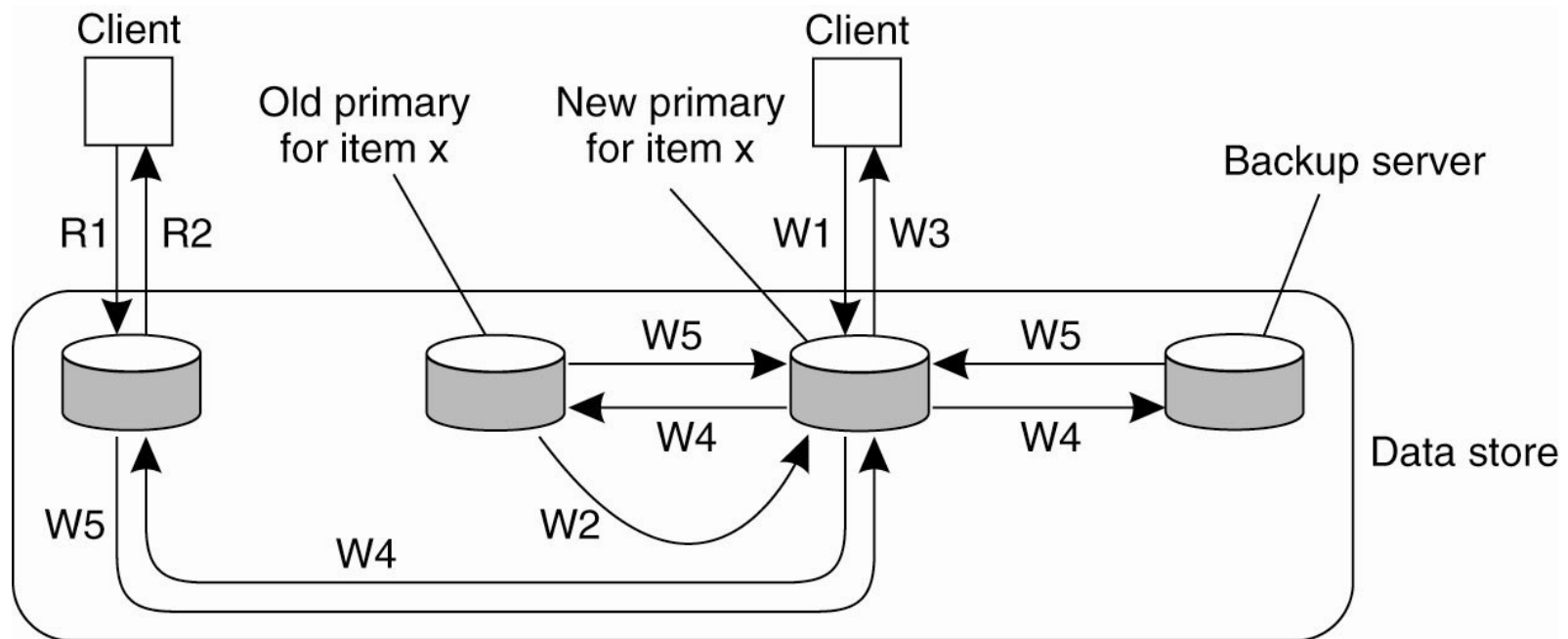
Remote-Write Protocol



- W1. Write request
- W2. Forward request to primary
- W3. Tell backups to update
- W4. Acknowledge update
- W5. Acknowledge write completed

- R1. Read request
- R2. Response to read

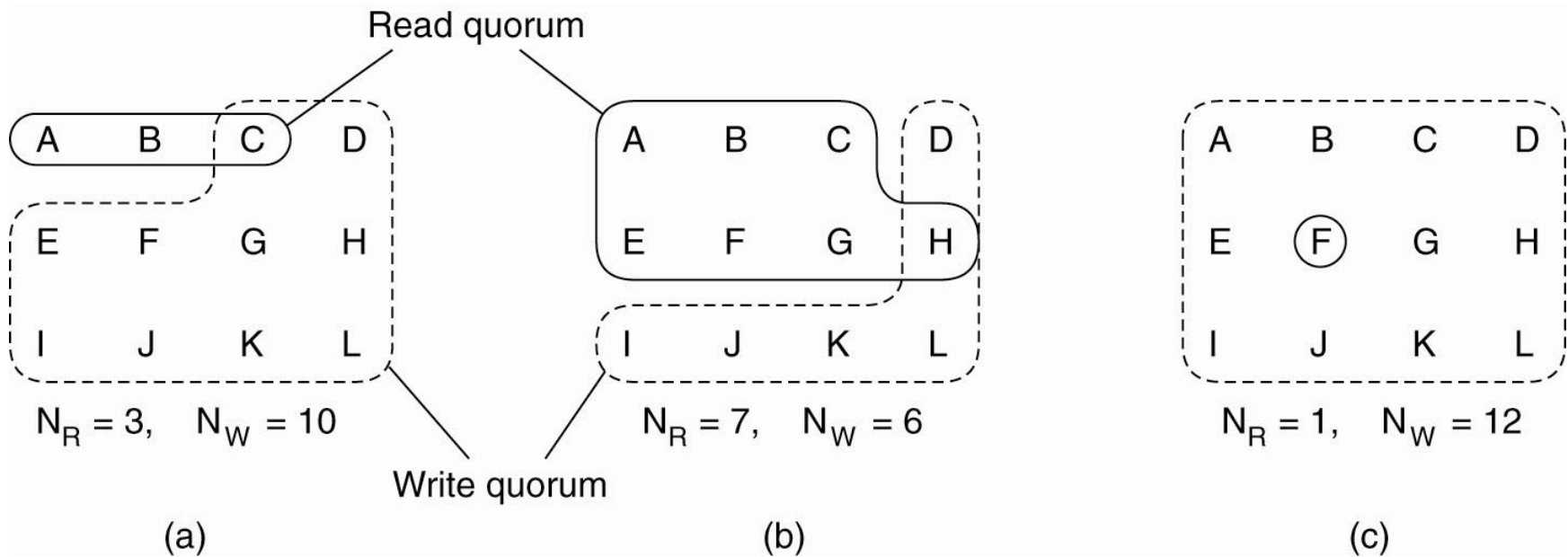
Local-Write Protocol



W1. Write request
W2. Move item x to new primary
W3. Acknowledge write completed
W4. Tell backups to update
W5. Acknowledge update

R1. Read request
R2. Response to read

Quorum-Based Protocols



Consistency Models

- strict consistency
 - magically “correct” global order
- sequential consistency
 - some global order (and local order)
- causal consistency – as before
- PRAM / processor consistency
 - local order, refinements

Access Semantics

- flat file
 - read, write, append – block-based
 - position-based access
- structured file
 - e.g., row/column store with keys
 - search-based access
- transaction
 - atomicity across multiple files