REPLICATION FOR LONG-TERM PERSISTENCE

JERRY SALTZER
M.I.T. Lab. for CS

LIBRARY 2000

DESIRED TIME OF PERSISTENCE

>>

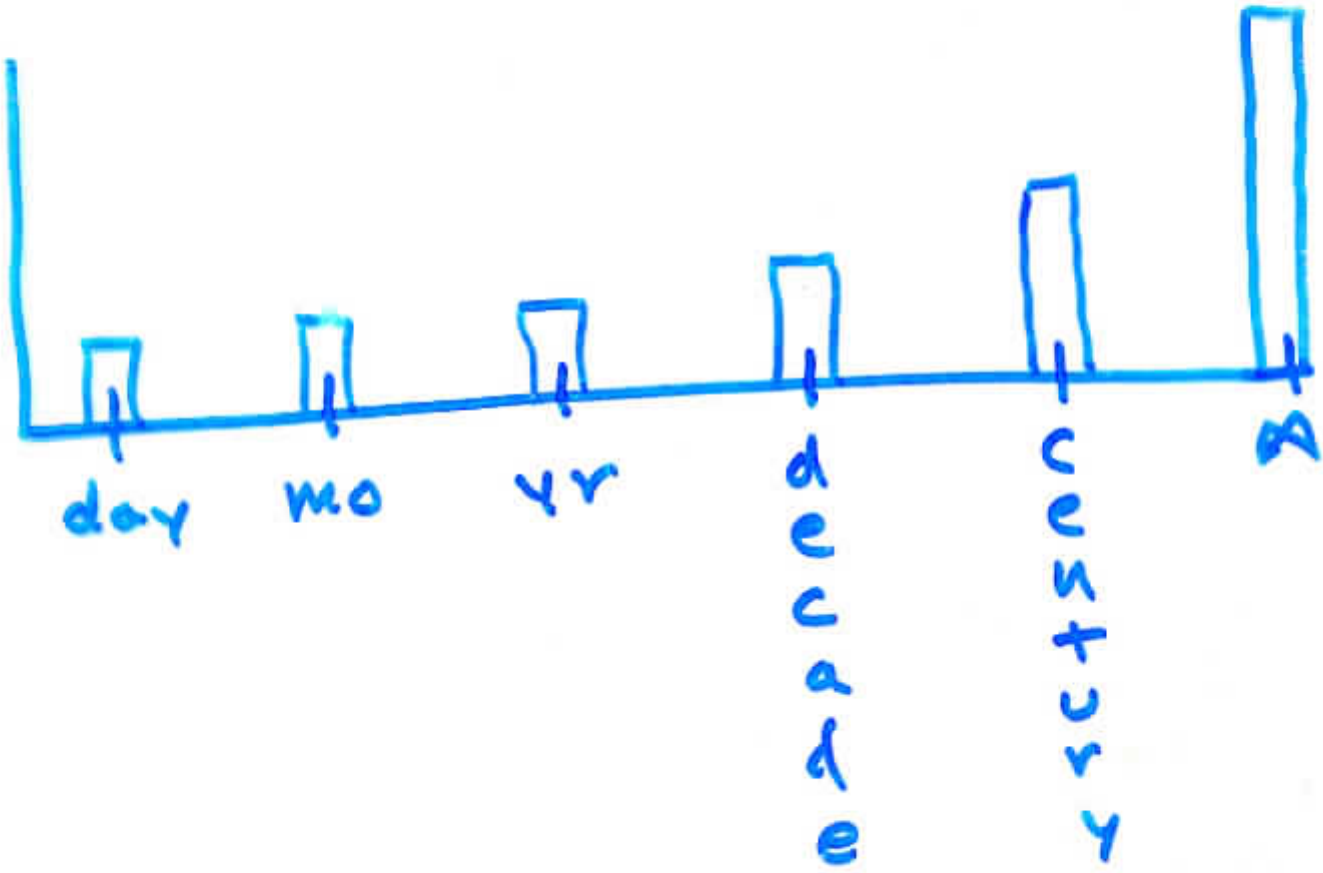
MTB SITE FAILURE

- Earthquake
- Hurricane
- Flood
- Meteor Strike
- City Fire
- Revolution

Strategy:

FULL REPLICATION, WITH GEOGRAPHIC DIVERSITY

TIME BETWEEN READS

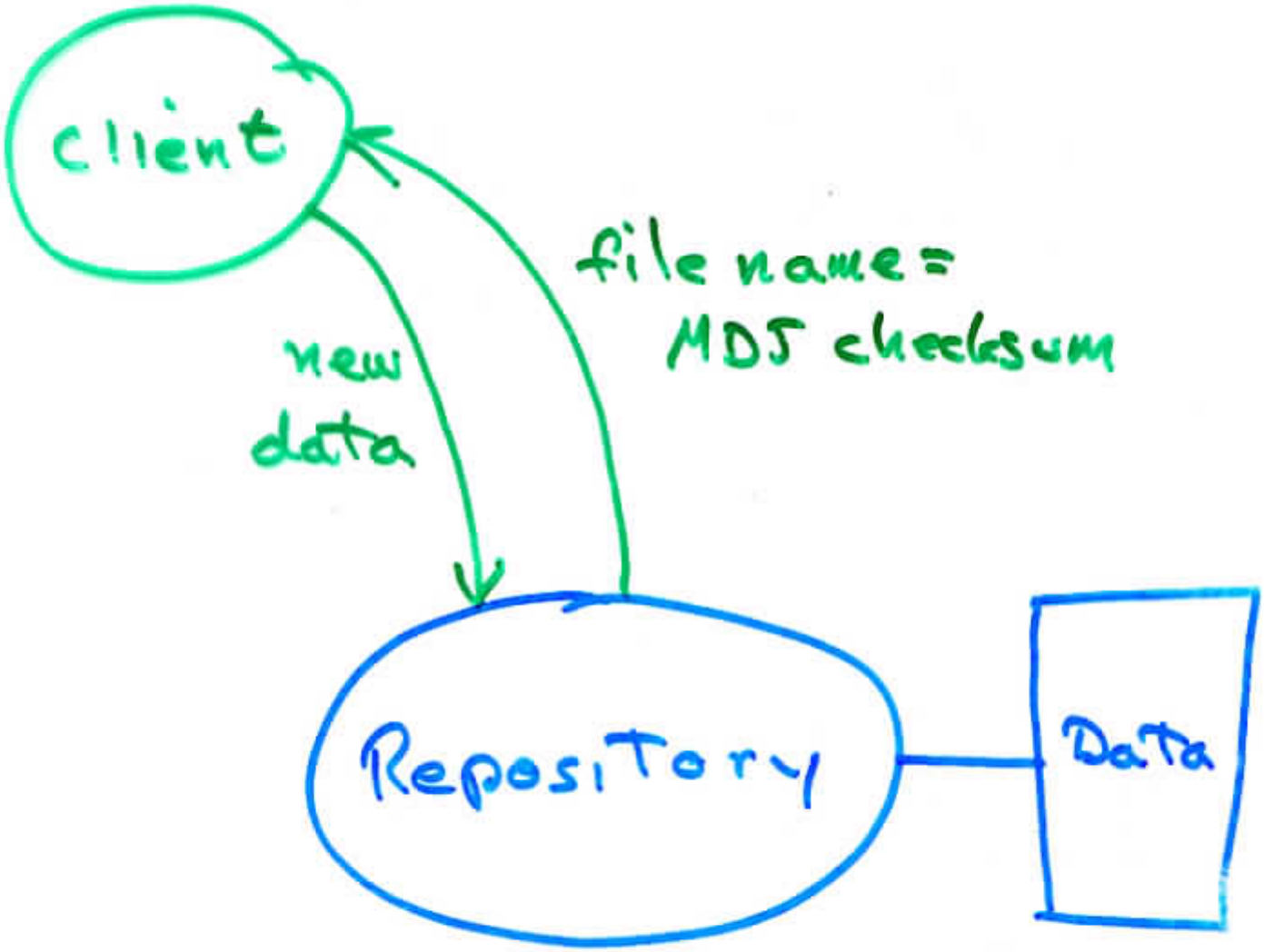


MTB Media Failure

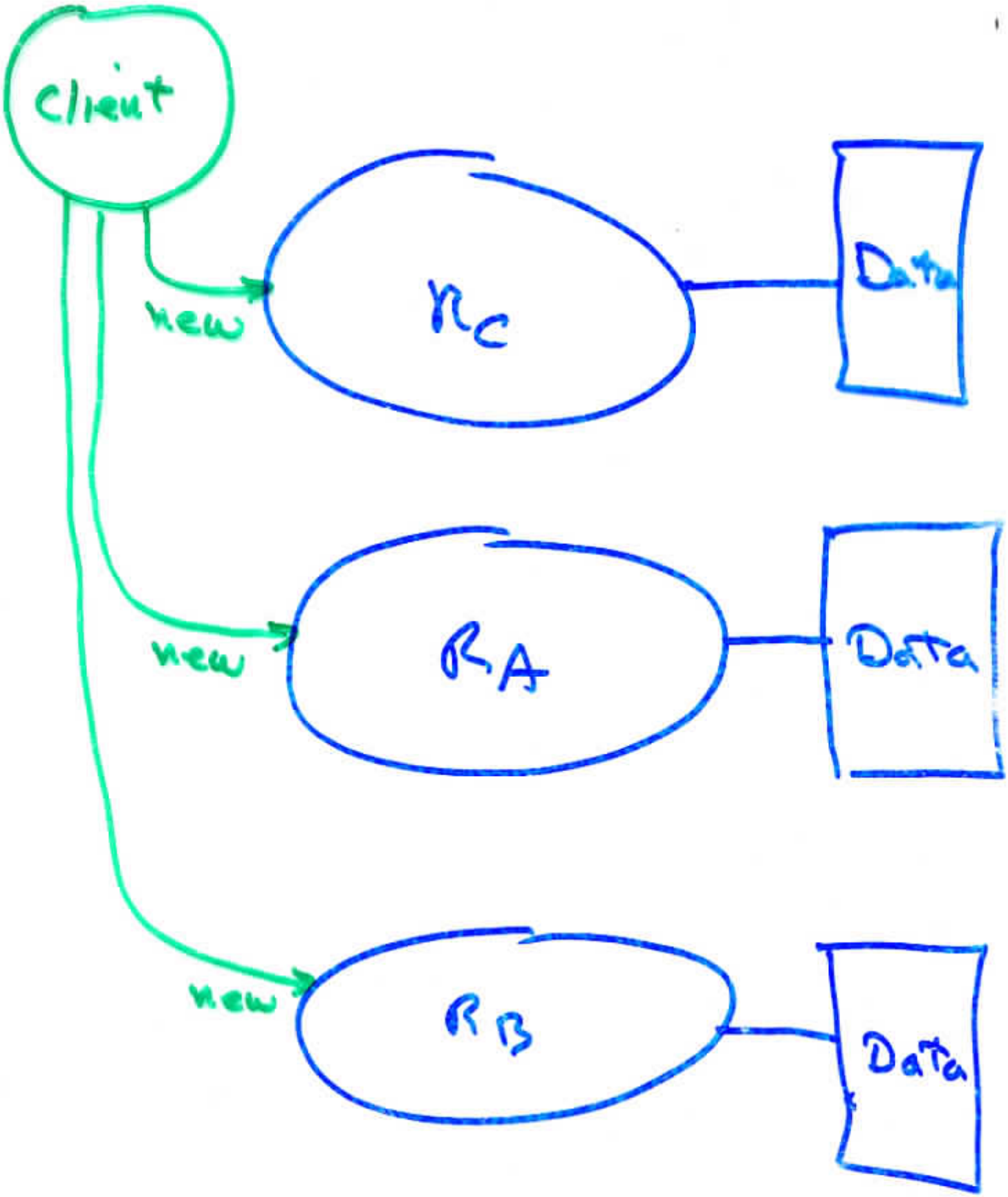
---

strategy:

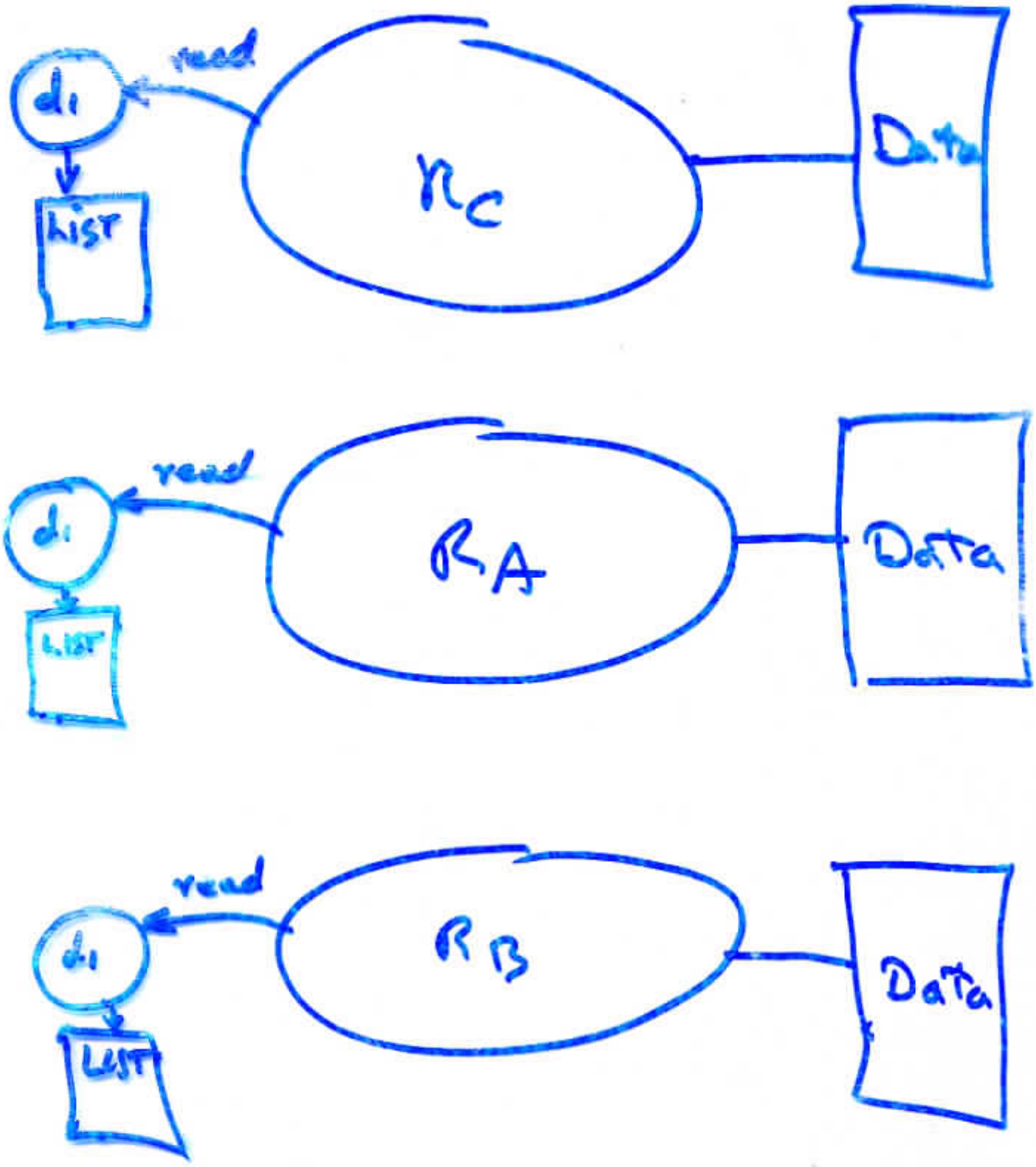
READING DAEMON



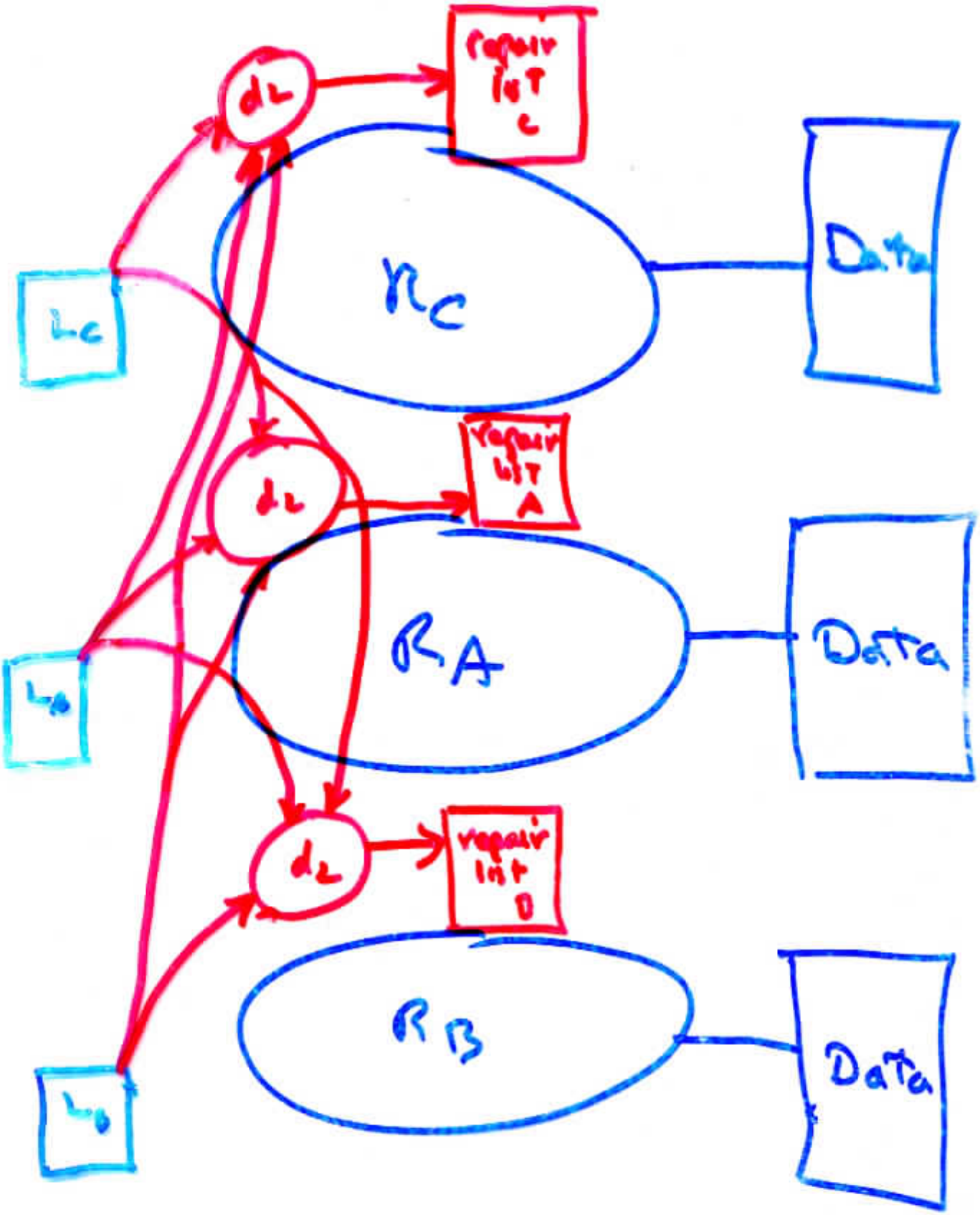
Append-Only Write
Semantics



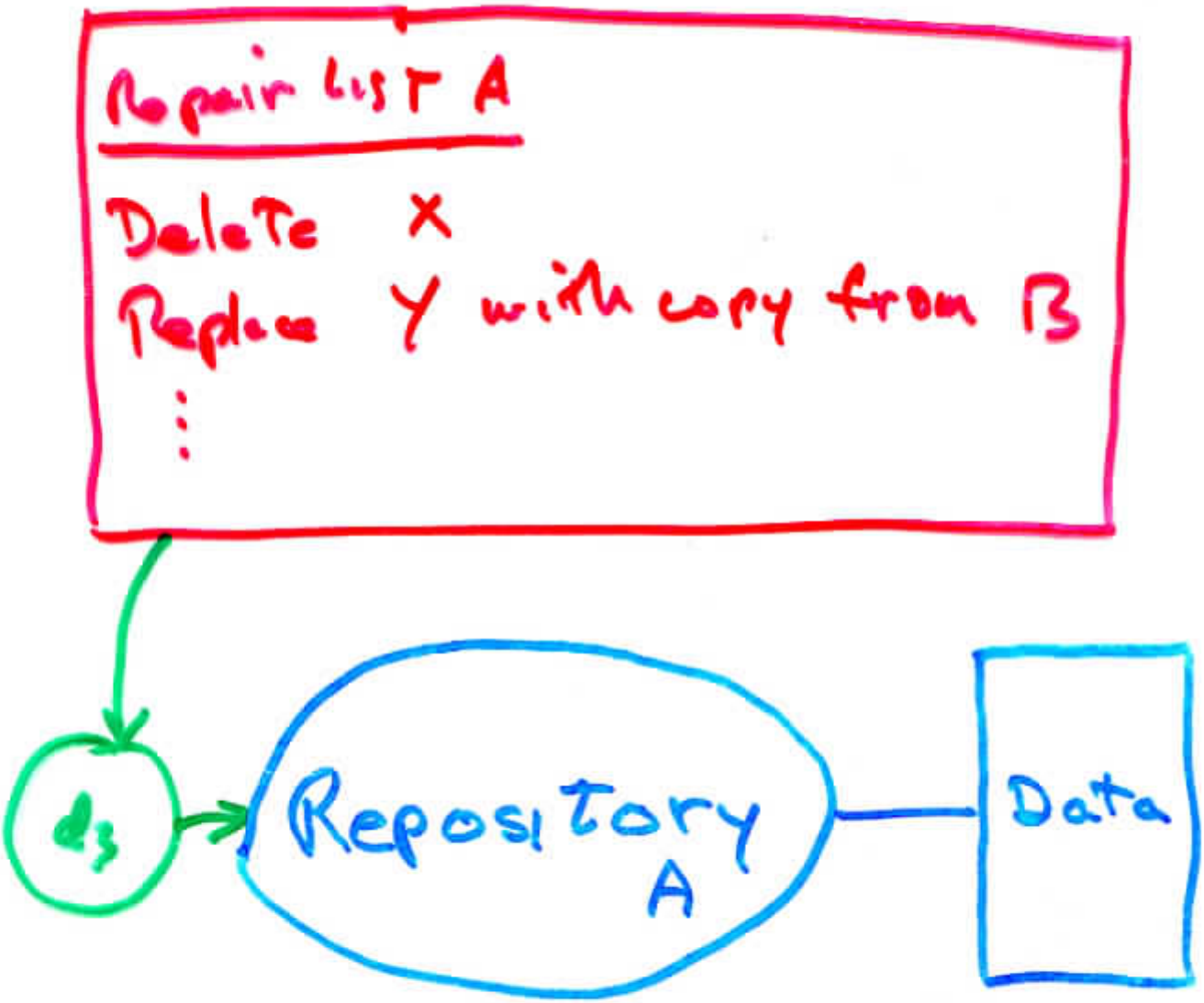
Stateless Commit: Majority of R's accept



d1 : reading daemon



d_2: repair order daemon



d3: repair execution daemon
(one/replica)

Construct repair list for replica J:

```
acquire list_k, k = 1...n
let U = union ( list_k, k = 1...n )
for each entry in U
  if entry checked in majority of k
    if entry is unchecked or
                    missing in list_j
      issue repair order:
                    "copy entry to j"
    endif
  otherwise
    If entry in list_j
      issue repair order:
                  "delete entry at j"
    endif
  endif
endfor
```

Simplicity

- Relaxed Consistency
- Stateless Protocols
- Append-Only Semantics
- Idempotent Repair
- Unique file name
  = MD5 checksum