REPLICATION
FOR LONG-TERM
PERSISTENCE

JERRY SALZBERG
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
Client

new data

Repository

file name = MD5 checksum

Data

Append - Only Write Semantics
Stateless Commit: Majority of R's accept
di: reading daemon
d2: repair order daemon
Repair list A

Delete X
Replace Y with copy from B

Repository A

Data

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
    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