Replication and Consistency Management

Kai Shen
Dept. of Computer Science, University of Rochester

Why replication?
- important to provide highly reliable services.
- allow load sharing on read-only requests.

Challenges:
- scalability, reliability.
- consistency management.

Viabilities:
- replica interconnects
- replication consistency requirements

Replica Interconnect

- RAID – redundant array of inexpensive disks
  - replication in the RAID controller
  - (often) block-level replication
  - inflexible in replication degree, or types of consistencies
  - can’t deal with host failure

- Local-area networks
  - fast, reliable connections

- Wide-area or wireless networks
  - slow, unreliable connections

Sample Network Services

Terms:
- read requests: those that don’t change the data
- write requests: those that do change the data

Email servers:
- replicated data: mailboxes
  - SMTP receiver will write; POP3/IMAP reader will write

Online discussion group:
- replicated data: discussion messages
  - posting a message will write

Generic data store (or hashtable)
- store key/value pairs
  - retrieve value with the key
  - delete key/value pairs
Replication Consistency for Network Services

- Must-haves:
  - convergence: if allowed a little extra time; all replicas will converge to the same state
  - no lost writes: this state corresponds to a particular serial execution of all writes
  - recovery: things can be recovered at the event of failures

- Optionals:
  - how much each replica deviates from each other at any particular moment?

Replication Consistency Maintenance

- synchronous replication
  - two-phase commit (standard method used in transactional databases)
  - problem: completion of write requests is constrained by the slowest node (low capacity or high runtime load)

- asynchronous replication
  - based case
  - primary-secondary replication
  - mobile wireless environments: conflict resolution through pairwise communications

Asynchronous Replication: Base Case

- out-of-order writes.
- fine with services with commutative writes (e.g., online discussion groups), but not others.

Asynchronous Replication: Primary-secondary

- Primary may not be always available in unreliable networks: mobile wireless networks.
Asynchronous Replication in Mobile or Wireless Networks

- Conflict-resolution through pairwise communications
- Nodes exchange writes at each session; resolve conflicts through undo, redo, application-specific rules
- Bayou paper