Group Communication

Phuong Hoai Ha & Yi Zhang

Introduction to Lab. assignments
March 24th, 2004

Today’s schedule
• Introduction to group communication
• Desired group communication
• Multicast communication
• Group membership service

Coordination in distributed systems
• Coordination is needed by distributed systems but hard to achieve:
  – Events happen concurrently
  – Communication links are not reliable
  – Computers can crash
  – New nodes can join the systems
  – Asynchronous environments
  ⇒ respect an efficient way to coordinate a group of processes

Group communication
• What is a group?
  – A number of processes which cooperate to provide a service.
  – An abstract identity to name a collection of processes.
• Group Communication
  – For coordination among processes of a group.

Who Needs Group Communication?
• Highly available servers (client-server)
• Database Replication
• Multimedia Conferencing
• Online Game
• Cluster management
• …

Distributed Web Server
• High availability
Online Game

• Fault-tolerance, Order

Different Comm. Methods

• Unicast
  – Point-to-Point Communication
  – Multiple copies are sent.

• Broadcast
  – One-to-All Communication
  – Abuse of Network Bandwidth

• Multicast
  – One-to-multiple Communication

Today’s schedule

• Introduction to group communication

Group Comm. Properties

• Name Abstraction

• Efficiency ⇒ Multicast

• Delivery Guarantee
  – Ordering
  – Failure behavior
  – Reliability
  – ...

• Dynamic Membership ⇒ Group membership service

Properties of Communication

• Ordering
  – Total order, causal order

• Failure behavior
  – Failure atomicity

• Reliability
  – Validity, integrity, agreement

Group Properties

• Name of group

• Addresses of group members

• Dynamic group membership

• Options:
  – Peer group or client-server group
  – Closed or Open Group
Peer Group

- All the members are equal.
- All the members send messages to the group.
- All the members receive all the messages.

Client-Server Group

- Replicated servers.
- Clients do not care which server answers.

Desired Group Communication

- Name Abstraction
- Efficiency \(\Rightarrow\) Multicast
- Delivery Guarantees \(\Rightarrow\) Reliability, Ordering
- Dynamic Membership \(\Rightarrow\) Group membership service

Today's schedule

- Introduction to group communication
- Desired group communication
- Multicast communication
- Group membership service

Multicast communication

- Send message over a distribution tree.
- Use network hardware support for broadcast or multicast when it is available.
- Minimize the time and bandwidth utilization

Reliability

Correct processes: those that never fail.

- Integrity
  A correct process delivers a message at most once.
- Validity
  A message from a correct process will be delivered by the process Eventually.
- Agreement
  A message delivered by a correct process will be delivered by all other correct processes in the group.

\(\Rightarrow\) Validity + Agreement = Liveness
Ordering
Assumptions: a process belongs to at most one group.
- FIFO
  - If \( m \rightarrow m' \), all correct processes that deliver \( m' \) will deliver \( m \) before \( m' \).
- Causal
  - If \( m \rightarrow m' \), all correct processes that deliver \( m' \) will deliver \( m \) before \( m' \).
- Total
  - If a correct process delivers \( m \) before \( m' \), all other correct processes that deliver \( m' \) will deliver \( m \) before \( m' \).

Examples
- Assumption:
  - Reliable one-to-one send operation (e.g. TCP)
- Basic multicast
  - Requirement:
    - All correct processes will eventually deliver the message from the correct multicaster.
  - Implementation:
    - \( B\text{-multicast}(g, m) \): \( \forall p \in g: \text{send}(p, m) \);
    - On receive(\( m \)) at \( p \): \( B\text{-deliver}(m) \) at \( p \).
    - Properties: integrity, validity.

Examples (cont.)
- Reliable multicast
  - Requirements: integrity, validity, agreement
  - Implementation:
    - \( R\text{-multicast}(g, m) \):
      - \( B\text{-multicast}(g, m) \);
      - On B-deliver(\( m \)) at process \( q \):
        - received := \{m\};
        - if( \( m \notin \text{received} \) )
          - \( \text{received} := \text{received} \cup \{ m \} \);
          - if( \( q \neq p \) ) \( B\text{-multicast}(g, m) \);
          - \( R\text{-deliver}(m) \);
    - Properties: integrity, validity.
  - Encourage to implement in more efficient ways (e.g. IP-multicast, etc.)

Examples (cont.)
- FIFO-ordered multicast:
  - Assumption:
    - A process belongs to at most one group.
  - Implementation:
    - \( \text{local variables at } p \):
      - \( S_p = 1, R_p[g] = 0 \);
    - \( F\text{-multicast}(g, m) \): B-multicast(g, m);
    - On B-deliver(\( m \)) at process \( q \):
      - \( m \rightarrow \text{received} \)
      - if( \( S_q < R_p[g] + 1 \) )
        - \( \text{F-deliver}(m) \);
      - \( R_p[g] := S_q \);
    - \( \text{Encourage to implement causally ordered, totally ordered multicasts.} \)

Today's schedule
- Introduction to group communication
- Desired group communication
- Multicast communication
- Group membership service

Group membership service
- Four tasks:
  - Interface for group membership changes
  - Failure detector
  - Group address expansion
  - Membership change notification
- Group partition:
  - Primary-partition
  - Partitionable
Group views

- Group views:
  - Lists of the current ordered group members
  - A new one is generated when processes join or leave/fail.
- View delivery
  - when the membership changes & a member is notified of it.
  - Requirements
    - Order
      - If process $p$ delivers $v(g) \rightarrow v'(g)$, no other process delivers $v'(g) \rightarrow v(g)$.
    - Integrity
      - If process $p$ delivers $v(g) \rightarrow v'(g)$, $p \in v'(g)$.
    - Non-triviality
      - If $q$ joins a group and becomes indefinitely reachable from $p$, eventually $q$ is always in the view $p$ delivers.
- View-synchronous group communication
  - Extend the reliable multicast semantics with group views.

Examples

- Ensemble: reliable group communication toolkit
  - Next talk

IP-multicast

- Multicast:
  - Yes:
    - efficiency
  - No:
    - Reliability
    - Ordering

- Group membership service:
  - Yes:
    - Interface for group membership change
    - Group address expansion
  - No:
    - Failure detector
    - Membership change notification

References

  - Section 4.5 Group Communication
  - Section 11.4 Multicast Communication
  - Section 14.2.2 Group Communication
  - ...