Distributed Computing III

Murphy was an optimist

CSSE6400

Richard Thomas

April 11, 2022

Question

What communication faults may occur?

Question

What communication faults may occur?

Answer

- Message not delivered

Question

What communication faults may occur?

Answer

- Message not delivered
- Message delayed

Question

What communication faults may occur?

Answer

- Message not delivered
- Message delayed
- Receiver failed

Question

What communication faults may occur?

Answer

- Message not delivered
- Message delayed
- Receiver failed
- Receiver busy

Question

What communication faults may occur?

Answer

- Message not delivered
- Message delayed
- Receiver failed
- Receiver busy
- Reply not received

Question

What communication faults may occur?

Answer

- Message not delivered
- Message delayed
- Receiver failed
- Receiver busy
- Reply not received
- Reply delayed

Question

How do we detect faults?

Question

How do we detect faults?

Answer

- No listener on port – RST or FIN packet

Question

How do we detect faults?

Answer

- No listener on port – RST or FIN packet
- Process crashes – Monitor report failure

Question

How do we detect faults?

Answer

- No listener on port – RST or FIN packet
- Process crashes – Monitor report failure
- IP address not reachable – unreachable packet

Question

How do we detect faults?

Answer

- No listener on port – RST or FIN packet
- Process crashes – Monitor report failure
- IP address not reachable – unreachable packet
- Query switches

Question

How do we detect faults?

Answer

- No listener on port – RST or FIN packet
- Process crashes – Monitor report failure
- IP address not reachable – unreachable packet
- Query switches
- Timeout

Question

What to do if fault is detected?

Question

What to do if fault is detected?

Answer

- Retry
- Restart

Definition 1. Idempotency

Repeating an operation does not change receiver's state.

Byzantine Generals Problem

- *n* generals need to agree on plan
- Can only communicate via messenger
- Messenger may be delayed or lost
- Some generals are traitors
  - Send dishonest messages
  - Pretend to have not received message

Question

How to order asynchronous messages?

Question

How to order asynchronous messages?

Answer

- Timestamps?
  - Can’t keep clocks in sync
  - Limited clock precision

Consistency

Eventual Consistency weak guarantee

Linearisability strong guarantee

Causal Ordering strong guarantee

Linearisability

- Once value is written, all reads see same value
  - Regardless of replica read from

Linearisability

- Once value is written, all reads see same value
  - Regardless of replica read from
- Single-leader replication
  - Read from leader
  - Read from synchronous follower

Linearisability

- Once value is written, all reads see same value
  - Regardless of replica read from
- Single-leader replication
  - Read from leader
  - Read from synchronous follower
- Multi-leader replication can't be linearised

Linearisability

- Once value is written, all reads see same value
  - Regardless of replica read from
- Single-leader replication
  - Read from leader
  - Read from synchronous follower
- Multi-leader replication can't be linearised
- Leaderless replication
  - Lock value on quorum before writing

Causal Order

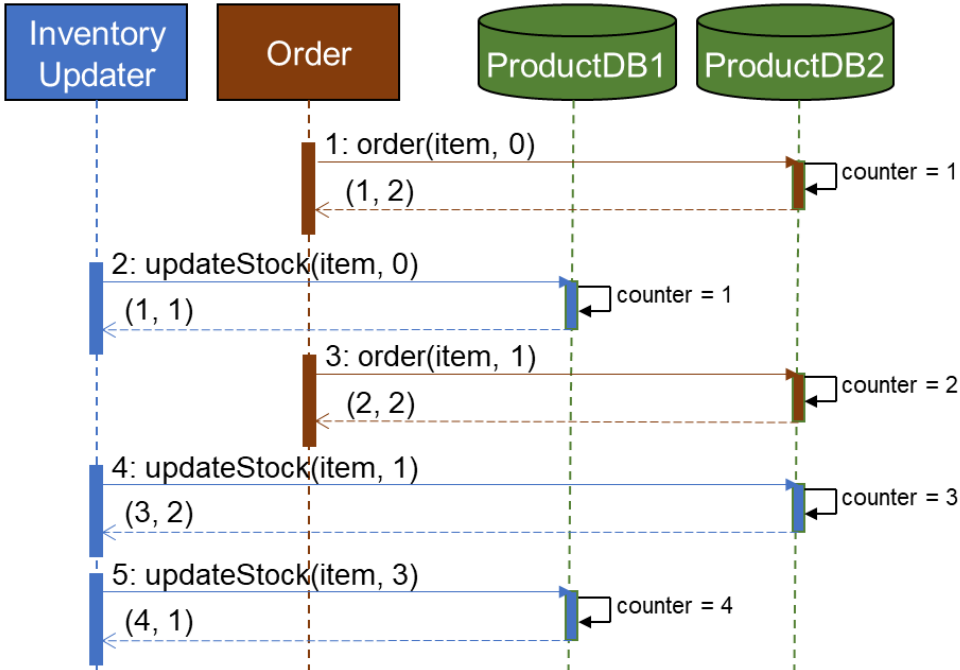
- Order is based on causality
  - What event needs to happen before another
  - Allows concurrent events

Causal Order

- Order is based on causality
  - What event needs to happen before another
  - Allows concurrent events
- Single-leader replication
  - Record sequence number of writes in log
  - Followers read log to execute writes

Causal Order

- Order is based on causality
  - What event needs to happen before another
  - Allows concurrent events
- Single-leader replication
  - Record sequence number of writes in log
  - Followers read log to execute writes
- Lamport timestamps


Inventory Updater
Order
ProductDB1
ProductDB2
1: order(item, 0)
(1, 2)
counter = 1
2: updateStock(item, 0)
(1, 1)
counter = 1
3: order(item, 1)
(2, 2)
counter = 2
4: updateStock(item, 1)
(3, 2)
counter = 3
5: updateStock(item, 3)
(4, 1)
counter = 4

Definition 2. Consensus

A set of nodes in the system agree on some aspect of the system's state.

Consensus Properties

Uniform Agreement All nodes must agree on the decision

Integrity Nodes can only vote once

Validity Result must have been proposed by a node

Termination Every node that doesn't crash must decide

Definition 3. Atomic Commit

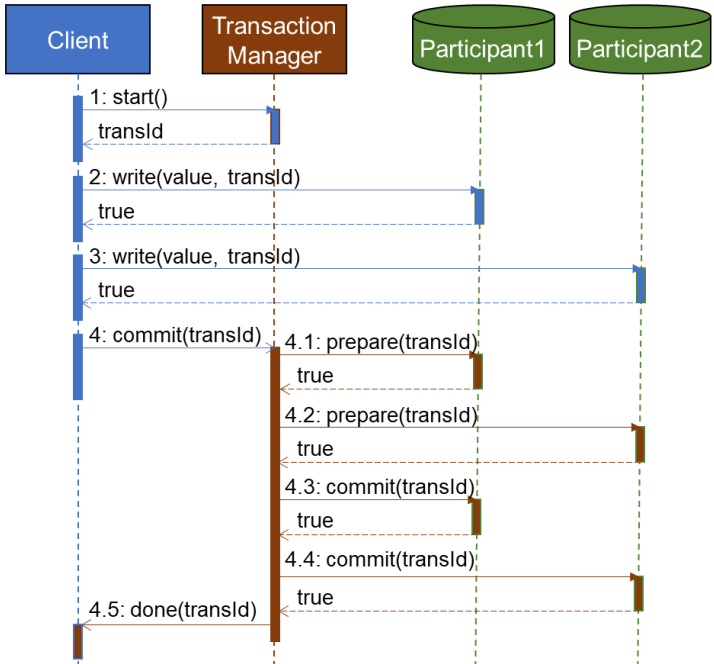
All nodes participating in a distributed transaction need to form consensus to complete the transaction.

Two-Phase Commit

Prepare Confirm nodes can commit transaction

Commit Finalise commit once consensus is reached

- Abort if consensus can't be reached


Client
Transaction Manager
Participant1
Participant2
1: start()
transId
2: write(value, transId)
true
3: write(value, transId)
true
4: commit(transId)
4.1: prepare(transId)
true
4.2: prepare(transId)
true
4.3: commit(transId)
true
4.4: commit(transId)
true
4.5: done(transId)