Atomic Commitment Protocols

Hans Philippi

March 21, 2017
Examples:

- A large bank managing accounts at different locations
- A money transfer between different banks
- A travelling agency, offering a combination of a hotel reservation and a flight reservation
- A deal in an MMOG
- Choosing a date and time for a meeting of several people
Finishing distributed transactions

- Goal: reach unanimity concerning termination
- either *commit* or *abort*
- Atomic Commitment Protocol (ACP)
- complications by failures
Kinds of failures

- site failures
  site is operational or down (fail/stop behaviour)
- partial failure in network
  some sites operational, some sites down
- total failure
  all sites down
- communication failure
  network partition, components

assumptions

- undeliverable messages dropped
- failure detection by timeout
Two phase commit

Every site maintains a DT-log (Distributed Transaction Log)
Every site will vote with respect to termination
([ready T] or [abort T])
Decision by COORD: Commit or Abort
ACP: correctness requirements

- All processes that reach a decision, reach the same one.
- A process cannot reverse its decision after it has reached one.
- The Commit decision can only be reached if all processes vote [ready T].
- If there are no failures and all processes vote [ready T] then the decision will be to Commit.
Two phase commit: phase 1

1. The coordinator logs a `<prepare-T>` record.
2. The coordinator sends a `[prepare-T]` message to all participants.
3. When a participant receives a `[prepare-T]`, it responds by sending to the coordinator a message containing that participant’s vote to the coordinator. If the participant votes `[abort T]`, it logs `<abort T>` and does a local rollback. Otherwise, it logs a record `<Ready T>` and sends a message `[Ready T]`, in that order.
Two phase commit: phase 2

1. The coordinator collects the vote messages from all participants. If all of them are positive and the coordinator’s vote is also positive, then the coordinator decides Commit, logs \langle \text{Commit } T \rangle \text{ and sends } [\text{Commit } T] \text{ messages to all participants. Otherwise the coordinator logs } \langle \text{Abort } T \rangle \text{ and sends } [\text{Abort } T] \text{ messages to all participants that voted } [\text{Ready } T].

2. Each participant that voted [\text{Ready } T] waits for a [\text{Commit } T] \text{ or } [\text{Abort } T] \text{ message from the coordinator. When it receives it, it logs the decision and acts accordingly.
Acting on time-out / recovery

So far, so good, but what to do when things go wrong? Our point of view: the position of a participant (or coord).

- Apparently, some error occurred (time-out, recovery)
- I am aware of the transactions I am involved in
- I need to be aware of my status in the protocol
- The DT-log provides me with this status
- I act according to this status, with forced termination as primary goal
Acting on time-out / recovery

What do I see in my DT log?

- A <Commit T> record
- An <Abort T> record
- A <Prepare T> record
- No <Ready T> record
- A <Ready T> record

Which are the corresponding actions?
Acting on time-out / recovery

What do I see in my DT log?
- A <Commit T> record, or
- An <Abort T> record

Action:
Applyingly, the decision has already been made.
Act according to this decision.
Acting on time-out / recovery

What do I see in my DT log?

- A <Prepare T> record, no decision record

Action:
?

Atomic Commitment Protocols
Acting on time-out / recovery

What do I see in my DT log?

- A <Prepare T> record, no decision record

Action:

My role was coordinator.
I started the protocol, but apparently no decision has been made yet.
I have the possibility to enforce an Abort.
Acting on time-out / recovery

What do I see in my DT log?

- No <Ready T> record, no decision record

Action:

?
What do I see in my DT log?
- No `<Ready T>` record, no decision record

Action:

My role was participant.
Apparently, I did not vote yet, so no Commit-decision has been made.
I have the possibility to enforce an Abort.
What do I see in my DT log?

- A <Ready T> record, no decision record

Action:

?
What do I see in my DT log?

- A <Ready T> record, no decision record

Action:

My role was participant. I voted positive, but I am not sure about the decision.

Try to contact the Coord and ask for the decision ...
What do I see in my DT log?
- A <Ready T> record, no decision record

Action:

My role was participant. I voted positive, but I am not sure about the decision.

1. Try to contact the Coord and ask for the decision
2. If Coord does not respond, start a Cooperative Termination Protocol
Run CTP (Cooperative Termination Protocol)

- Choose a Leader among the remaining participants
- This Leader acts as a new coordinator and requests the status from all remaining participants
- If one of the remaining participants knows the original decision, the Leader will broadcast this decision
- If one of the remaining participants has not yet voted, the Leader will broadcast an Abort decision
- If all of the remaining participants have voted positive and no one knows the decision, the protocol is *blocked*!
Blocking protocols

- 2PC is a blocking protocol
- blocking requires human intervention (from DBA) ...
- ... and/or correction protocols
- 3PC is an attempt to prevent blocking ...
- .. but cannot avoid blocking in the case of communication failures
- There is a proof of the claim that a non-blocking commitment protocol cannot exist