

两阶段提交协议

陈宗志

背景

- 两阶段提交协议是一种历史悠久的分布式协议,最早用在分布式的数据库中,实现分布式事务
- 为了满足数据库中事务的Atomicity, 必须要么都提交, 要么都失败.

协议内容

1. 协调者发送一个VOTE-REQ消息给所有的参与者
2. 当参与者接收到VOTE-REQ消息后, 它会发送一个包含参与者投票结果的消息(Yes or Not)给协调者. 如果参与者投的是No, 它会决定Abort事务并停止运行
3. 协调者收集来自所有参与者的投票信息. 如果都是Yes, 同时协调者也是投Yes. 那么协调者就决定Commit, 并向所有参与者发送Commit 消息. 否则协调者决定Abort, 向所有参与者发送Abort消息
4. 每个参与者等待来自协调者的Commit 或者 Abort 消息. 收到以后执行相应的动作, 然后停止

协议内容

- 步骤 1, 2 指投票阶段
- 步骤 3, 4 指决定阶段

异常处理

- 超时
- 宕机回复

优化

- 超时处理优化
- cooperative termination protocol

协调者

```
write START_2PC to local log;
multicast VOTE_REQUEST to all participants;
while not all votes have been collected {
    wait for any incoming vote;
    if timeout {
        write GLOBAL_ABORT to local log;
        multicast GLOBAL_ABORT to all participants;
        exit;
    }
    record vote;
}

if all participants sent VOTE_COMMIT and coordinator votes COMMIT {
    write GLOBAL_COMMIT to local log;
    multicast GLOBAL_COMMIT to all participants;
}else{
    write GLOBAL_ABORT to local log;
    multicast GLOBAL_ABORT to all participants;
}
wait for GLOBAL_ABORT or GLOBAL_COMMIT response
write END_TRANSACTION to local log;
```

参与者

```
write INIT to local log;
wait for VOTE_REQUEST from coordinator;
if timeout {
    write VOTE_ABORT to local log;
    exit;
}
if participant votes COMMIT {
    write VOTE_COMMIT to local log;
    send VOTE_COMMIT to coordinator;
    wait for DECISION from coordinator;
    if timeout {
        multicast DECISION_REQUEST to other participants;
        wait until DECISION is received; /* remain blocked*/
        write DECISION to local log;
    }
    if DECISION == GLOBAL_COMMIT
        write GLOBAL_COMMIT to local log;
        send GLOBAL_COMMIT confirm;

    else if DECISION == GLOBAL_ABORT
        write GLOBAL_ABORT to local log;
        send GLOBAL_ABORT confirm;
} else {
    write VOTE_ABORT to local log;
    send VOTE_ABORT to coordinator;
}
```


参与者

```
while true {  
    wait until any incoming DECISION_REQUEST is received;  
    read most recently recorded STATE from the local log;  
    if STATE == GLOBAL_COMMIT  
        send GLOBAL_COMMIT to requesting participant;  
    else if STATE == INIT or STATE == GLOBAL_ABORT;  
        send GLOBAL_ABORT to requesting participant;  
    else  
        skip; /* participant remains blocked */  
}
```

协议缺点

- 容错能力较差
- 性能较差



Thanks