The Atomic Commit Problem

R. Guerraoui - EPFL
An Agreement Problem
Atomic Commit

**Agreement:** No two processes decide differently

**Termination:** Every correct process eventually decides

**Commit-Validity:** 1 is only decided if all propose 1

**Abort-Validity:** 0 is only decided if some process proposes 0 or there is a failure
Distributed Transaction
• 70’s: Lampson/Gray (1st protocol)

• 80’s: Skeen/Dwork (1st result)

• 90’s: Hadzilacos/Guerraoui (problem)

• 2000’s: Kuznetsov (computability)

• 2017: Wang (complexity)
2-Phase Commit (2PC)
2PC

propose(1)

p1

propose(1)

p2

propose(1)

p3

crash

decide(0)

decide(0)
2PC is blocking

propose(1)

p1

propose(1)

p2

propose(1)

p3

crash
3PC

- Skeen 81
- Mohan – Strong – Finkelstein 83
- Guerraoui – Larrea - Schiper 96
- Keidar – Dolev 98
- Gray – Lamport 2004
Consensus

**Agreement:** No two processes decide differently

**Termination:** Every correct process eventually decides

**Validity:** The value decided is a value proposed
Commit with Consensus

propose(1)  
Cons(1,1)  
propose(1)  
Cons(1,1)  
propose(1)  
Cons(1,1)  

p1  

p2  

p3  

decide(1)

decide(1)

decide(1)
Commit with Consensus

propose(1)

p1

propose(1)
crash

propose(1)

p2

Cons(0,0)

decide(0)

propose(1)

p3

Cons(0,0)

decide(0)
Commit with Consensus

propose(1)

p1

propose(1)

p2

propose(1)

p3

Cons(0,0-1)

decide(0-1)

Cons(1,0-1)

decide(0-1)

crash
Weak Consensus

**Agreement**: No two processes decide differently

**Termination**: Every correct process eventually decides

**Weak consensus**: 0 and 1 are both possible values
• 70’s : Lampson/Gray (1st protocol)

• 80’s : Skeen/Dwork (1st result)

• 90’s: Hadzilacos/Guerraoui (problem)

• 2000’s: Kuznetsov (computability)

• 2017: Wang (complexity)
Computability (Weakest FD)

1. $<>P$ is not enough

2. $P$ is needed if one process can crash

3. The weakest FD is $(FS, FS$ or $(\Omega$ and $\xi)$)
1. $<>P$ is not enough (Gue’95)

$p1$  
propose(1)  

decide(0)  

$p2$  
propose(1)  

decide(0)  

$p3$  
propose(1)  

decide(0)  

$<>P$ becomes $P$
2. P is needed with one crash (FRT’99)
3. The WFD for Atomic Commit

- GK 02: \((FS, \Omega)\)

- DFGHTK 04: \((FS \land (\leftrightarrow FS \lor (\Omega \land \xi))))\)
Consensus

**Agreement:** No two processes decide differently

**Termination:** Every correct process eventually decides

**Validity:** The value decided is a value proposed

**Quittable consensus:** Q can be decided if there is a failure
• 70’s : Lampson/Gray (1st protocol)

• 80’s : Skeen/Dwork (1st result)

• 90’s: Hadzilacos/Guerraoui (problem)

• 2000’s: Kuznetsov (computability)

• 2017: Wang (complexity)
How fast can a transaction commit in a nice run?

Skeen/Dwork 83: 2n-2 messages assuming n-1 failures in a synchronous system
Complexity (Delays)

- 1 if synchrony
- 2 if asynchronous agreement (indulgent)
Complexity (Messages)

- $n - 1 + f$ if $f$ failures and synchrony
- 0 if validity only in nice executions
- $2n - 2$ if validity despite asynchrony
- $2n - 2 + f$ if agreement despite asynchrony
Today

- Sinfonia, Percolator, Clock-SI, Yesquel use 2PC
  - 2 message delays / 2n-2 messages
  - No termination + synchrony assumption

- INBAC
  - 2 message delays / 2n messages
  - Termination + agreement in asynchrony

- 0NBAC
  - 1 message delay / 0 messages
  - Validity only in nice executions
Netys 2017

- Abstract Dec 2 / Paper Dec 9
- Conference May 17/19