The Impossibility of Ensuring SI in Genuine Replicated STMs

Masoud SAEIDA ARDEKANI (Lip6)
Pierre SUTRA (INRIA & Lip6)
Marc SHAPIRO (INRIA & Lip6)
Nuno PREGUIÇA (Universidade Nova de Lisboa)

WTTM 2011
September 23, Rome, Italy

Goal

- Build a highly scalable distributed STM

Scalability Objectives

- Partial Replication
  - A given object is replicated only on a sub-set of processes
- Genuine Partial Replication
  - Only processes that hold objects read/written by a transaction exchange messages

Snapshot Isolation

- Snapshot Read: read a consistent snapshot
- Snapshot Write: first-committer-wins behavior

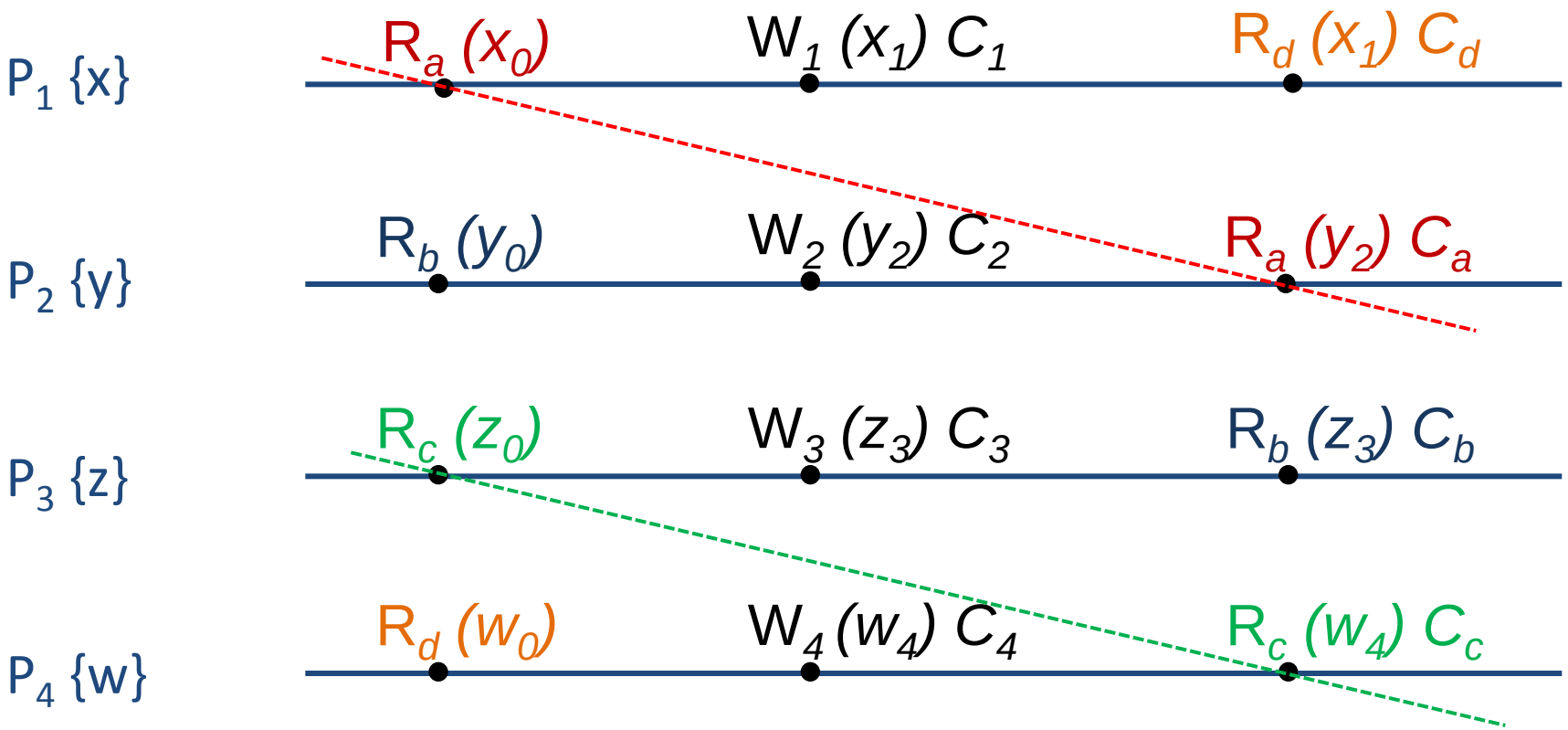
- Progress
  - Read-only transactions always commit
  - Non conflicting concurrent updates always commit

Impossibility Result

- No message-passing transactional system can achieve:
  - genuine partial replication
  - snapshot isolation
  - read-set not known in advance

Counter Example

Snapshot-Read is violated!
It cannot be detected without breaking Genuineness!

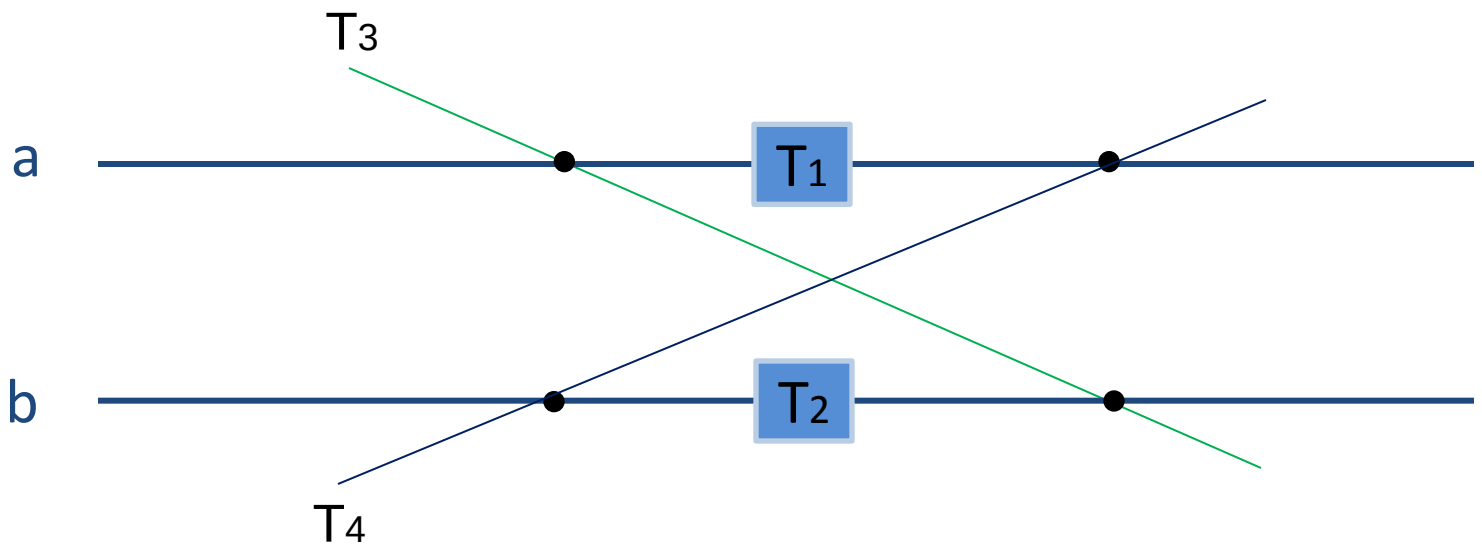


To circumvent: relax Snapshot Isolation

SI Decomposition

- Avoid Cascading Aborts
- Consistent Snapshot
- Write Conflict Freedom
- Snapshot Monotonicity
  - cause the impossibility result !

Non-Monotonic Snapshots

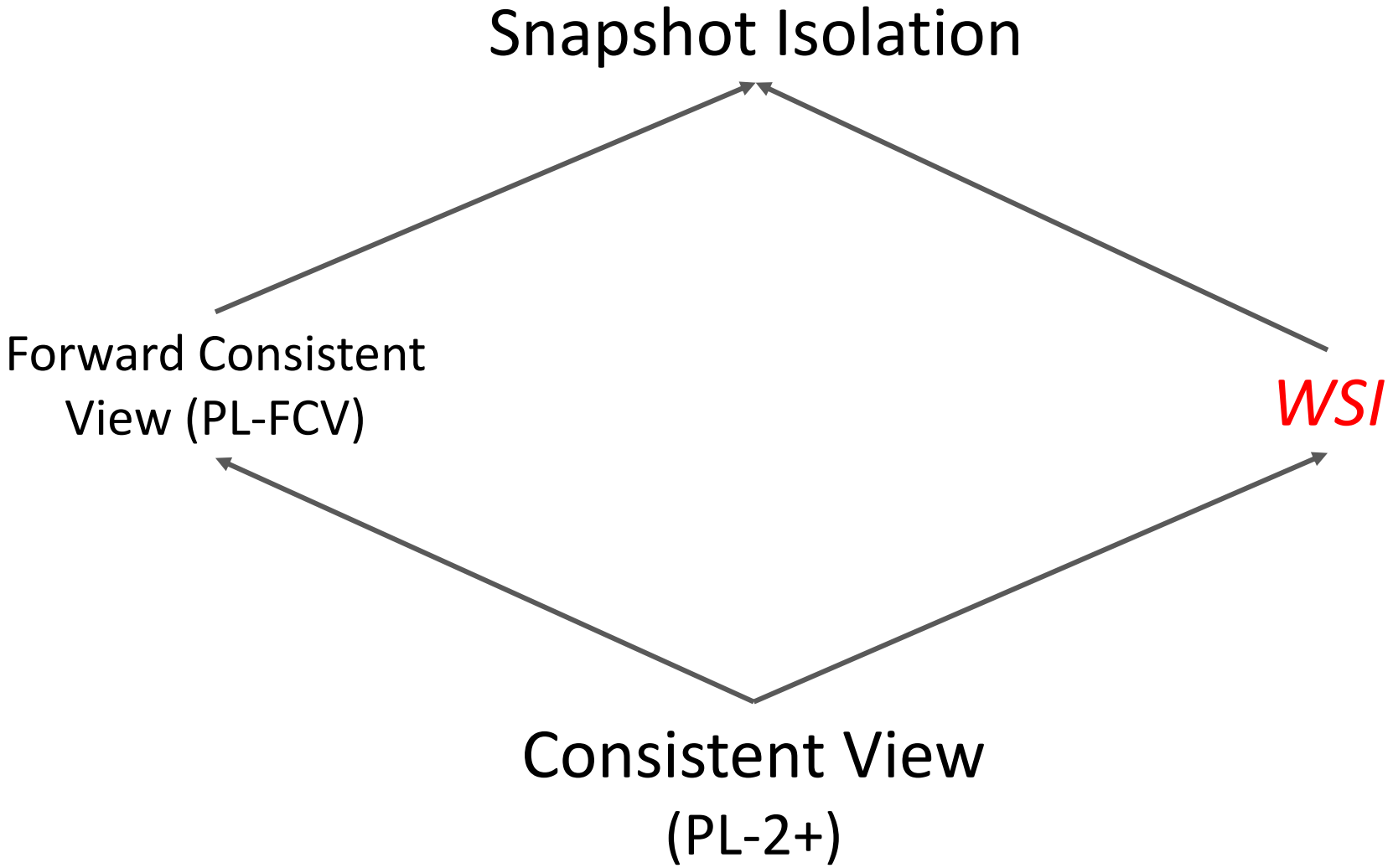


SI forbids this situation!

Weak Snapshot Isolation (WSI)

- Avoid Cascading Aborts
- Consistent Snapshot
- Write Conflict Freedom
- ~~Snapshot Monotonicity~~

Adya's Consistency Hierarchy



Summary

- Impossible to have genuine partial replication under SI if read-set is unknown!

- Drop snapshot monotonicity to circumvent impossibility result