# BulkSC: Bulk Enforcement of Sequential Consistency

Luis Ceze, James Tuck, Pablo Montesinos, Josep Torrellas

University of Illinois at Urbana-Champaign

http://iacoma.cs.uiuc.edu



ISCA 2007, San Diego, CA



Ceze, Tuck, Montesinos, Torrellas BulkSC: Bulk Enforcement of Sequential Consistency



• Defines the values that a read can return





- Defines the values that a read can return
- Supported by the hardware





- Defines the values that a read can return
- Supported by the hardware
- Has major implications on programmability





- Defines the values that a read can return
- Supported by the hardware
- Has major implications on programmability





- Defines the values that a read can return
- Supported by the hardware
- Has major implications on programmability



- Defines the values that a read can return
- Supported by the hardware
- Has major implications on programmability



- Defines the values that a read can return
- Supported by the hardware
- Has majo Affects the whole software stack:
  applications
  - •OS, libraries, drivers
  - •compilers

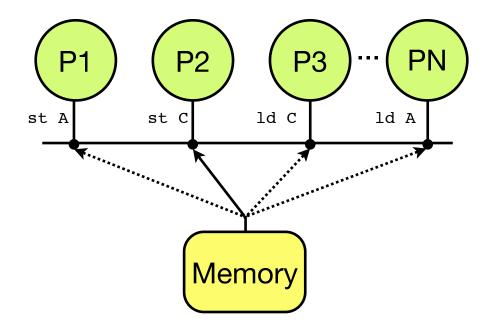
Data

language semantics

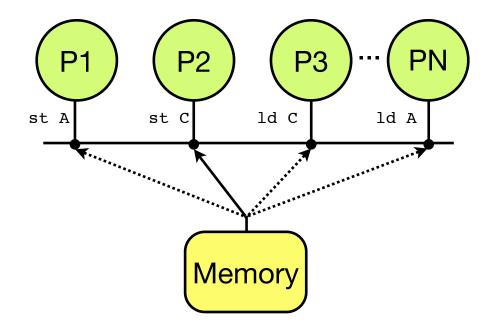
; }





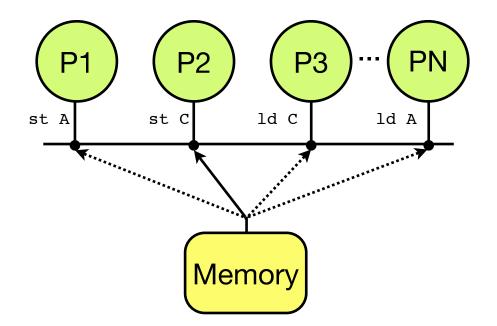








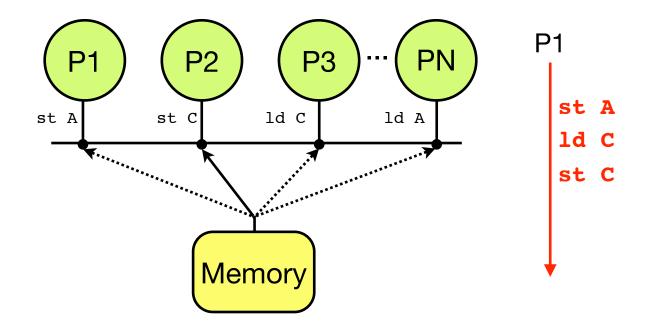




**Per-processor program order**: memory operations from individual processors maintain program order





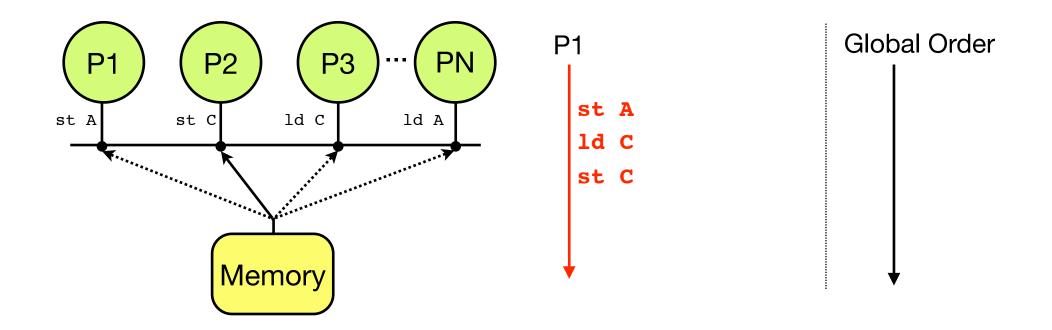


**Per-processor program order**: memory operations from individual processors maintain program order





Ceze, Tuck, Montesinos, Torrellas

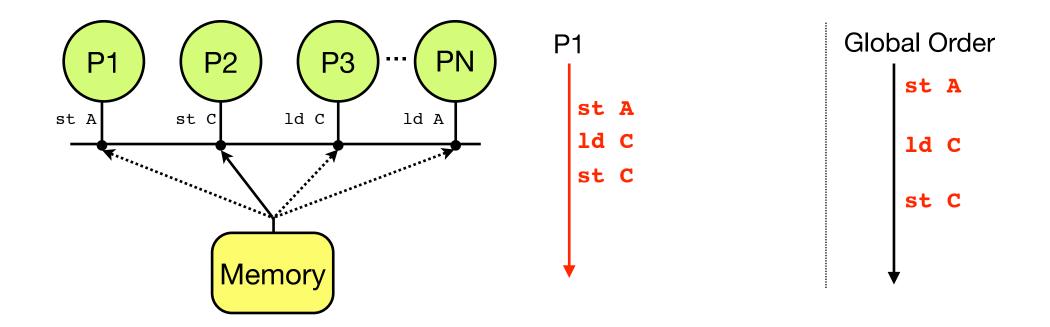


**Per-processor program order**: memory operations from individual processors maintain program order

[Lamport'79]



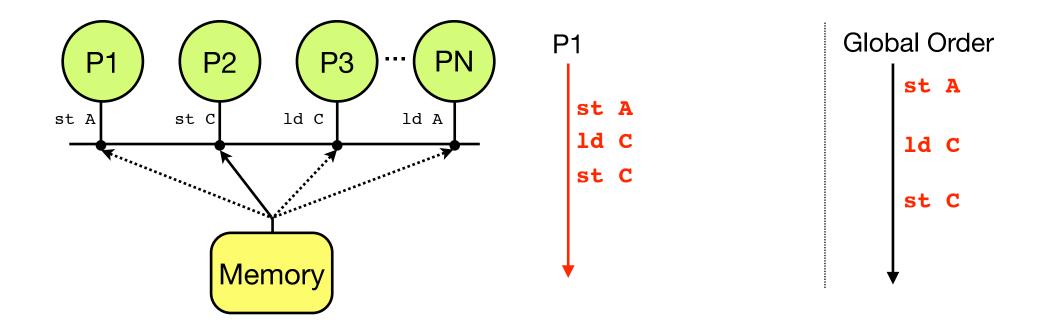
COMC 3



**Per-processor program order**: memory operations from individual processors maintain program order



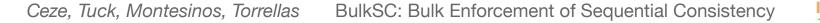




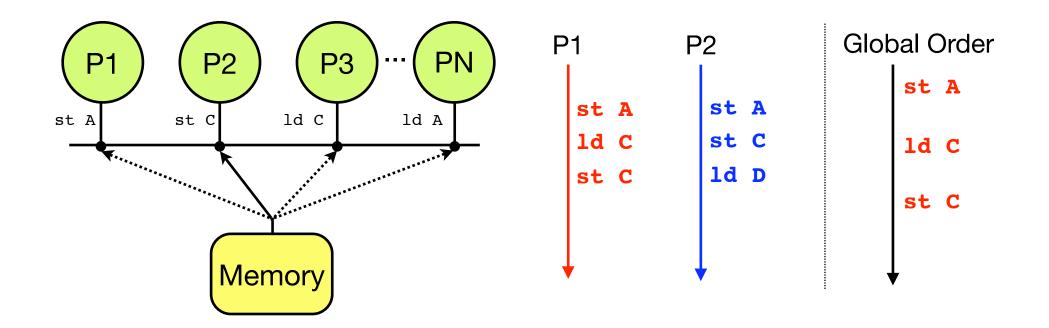
**Per-processor program order**: memory operations from individual processors maintain program order

**Single sequential order**: the memory operations from all processors maintain a single sequential order

[Lamport'79]



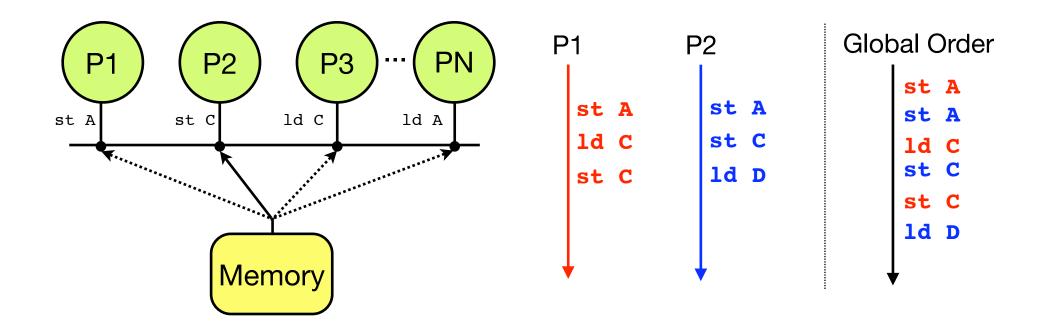
na a



**Per-processor program order**: memory operations from individual processors maintain program order

**Single sequential order**: the memory operations from all processors maintain a single sequential order



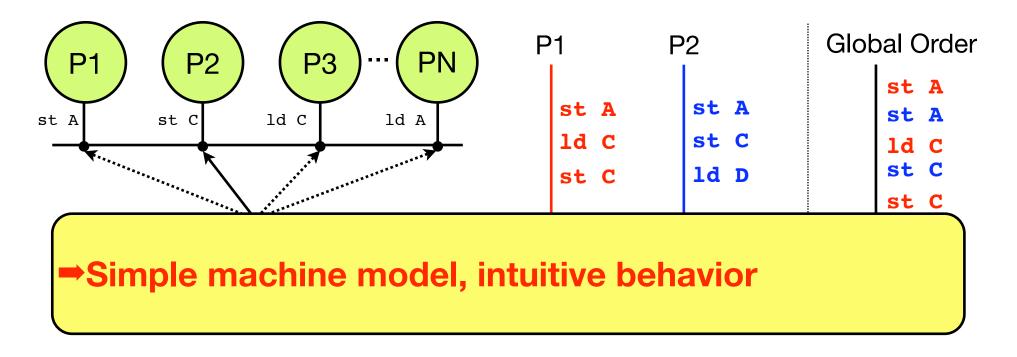


**Per-processor program order**: memory operations from individual processors maintain program order

**Single sequential order**: the memory operations from all processors maintain a single sequential order







**Per-processor program order**: memory operations from individual processors maintain program order

**Single sequential order**: the memory operations from all processors maintain a single sequential order

[Lamport'79]



oma 3





Low Performance





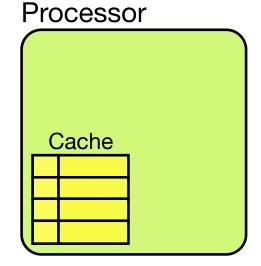
- Low Performance
  - restrictions on performance-enhancing reordering of memory operations





#### Low Performance

- restrictions on performance-enhancing reordering of memory operations
- Or Complex Implementation





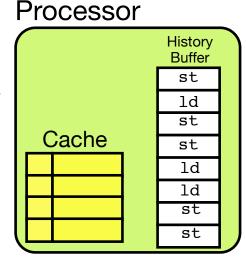


#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

#### Or Complex Implementation

• buffer long history of speculative memory accesses





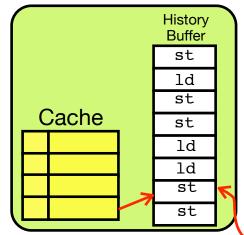
#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

## Or Complex Implementation

- buffer long history of speculative memory accesses
- check this history against coherence events and cache displacements

#### Processor



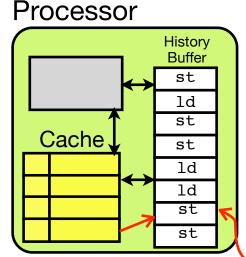


#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

## Or Complex Implementation

- buffer long history of speculative memory accesses
- check this history against coherence events and cache displacements
- coupled with key structures (LSQ, ROB, reg file, \$)



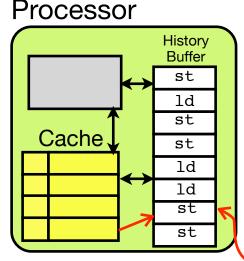


#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

## Or Complex Implementation

- buffer long history of speculative memory accesses
- check this history against coherence events and cache displacements
- coupled with key structures (LSQ, ROB, reg file, \$)
- typically fine-grain (instruction-level) undo



[Gharachorloo'91] [Ranganathan'96] [Gniady'97, SC++]





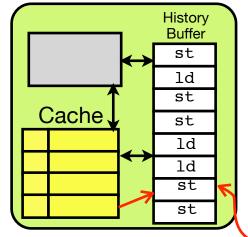
#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

## Or Complex Implementation

- buffer long history of speculative memory accesses
- check this history against coherence events and cache displacements
- coupled with key structures (LSQ, ROB, reg file, \$)
- typically fine-grain (instruction-level) undo
- Most current systems do not support SC

#### Processor

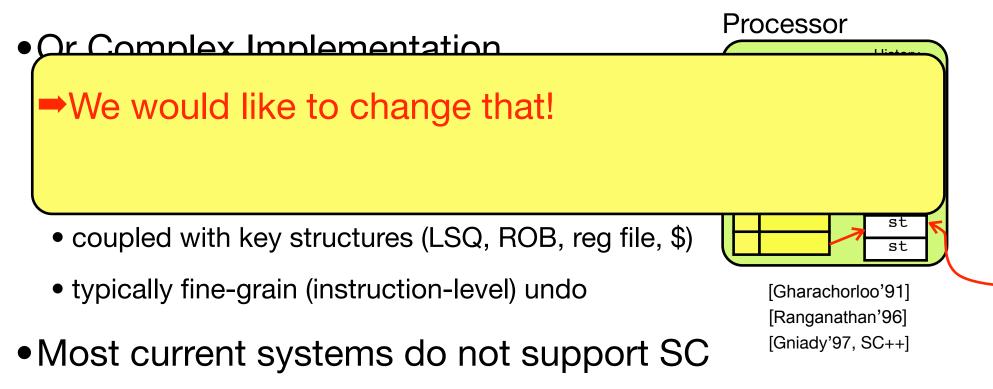


<sup>[</sup>Gharachorloo'91] [Ranganathan'96] [Gniady'97, SC++]





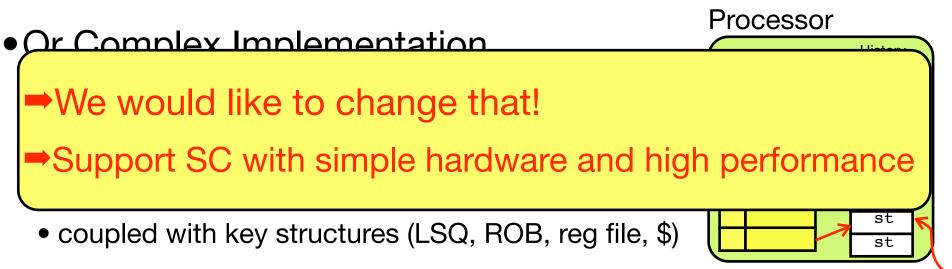
- Low Performance
  - restrictions on performance-enhancing reordering of memory operations





#### Low Performance

 restrictions on performance-enhancing reordering of memory operations

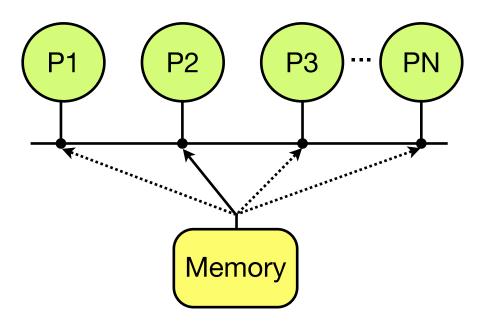


- typically fine-grain (instruction-level) undo
- Most current systems do not support SC



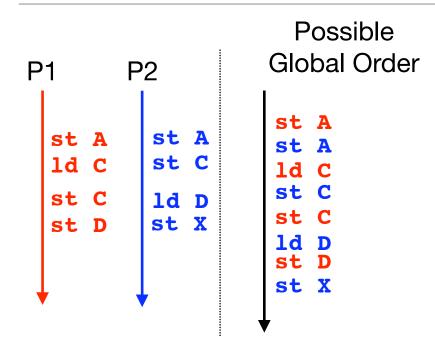


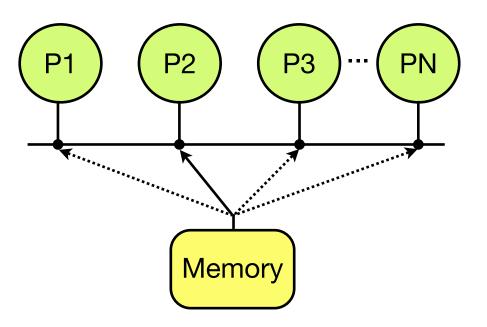




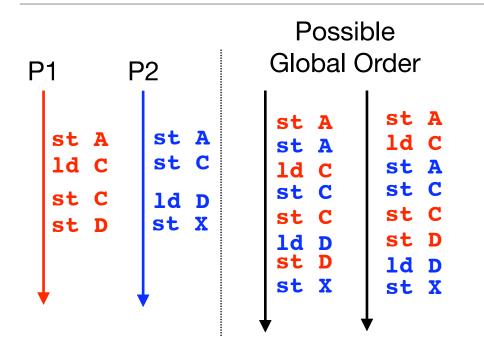


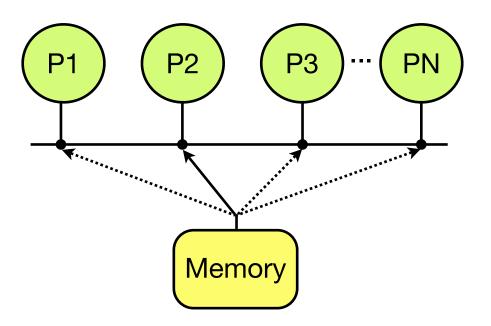




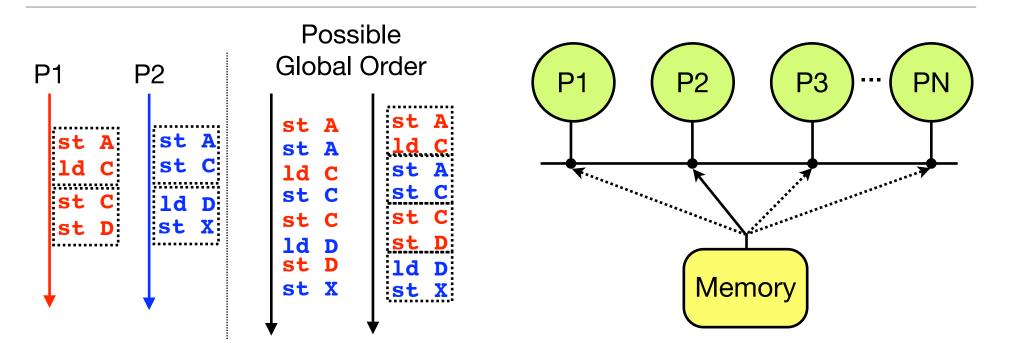




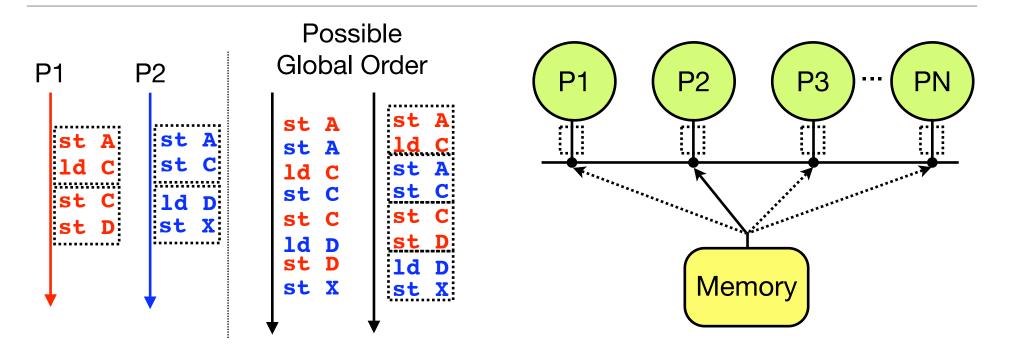




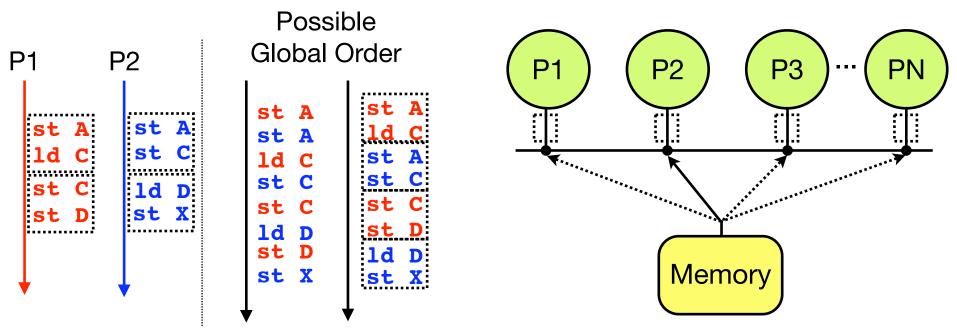






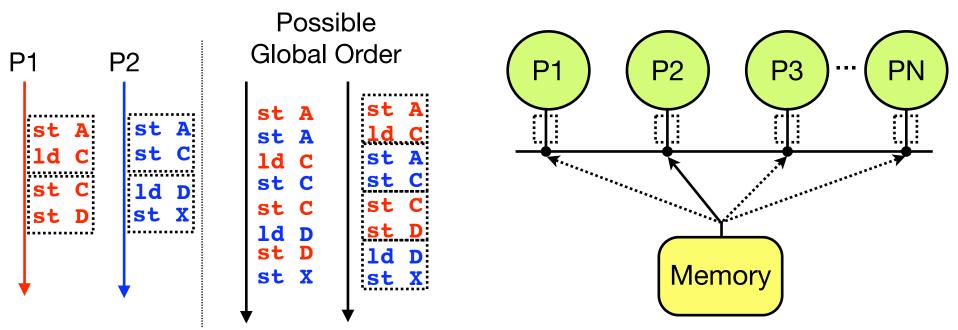




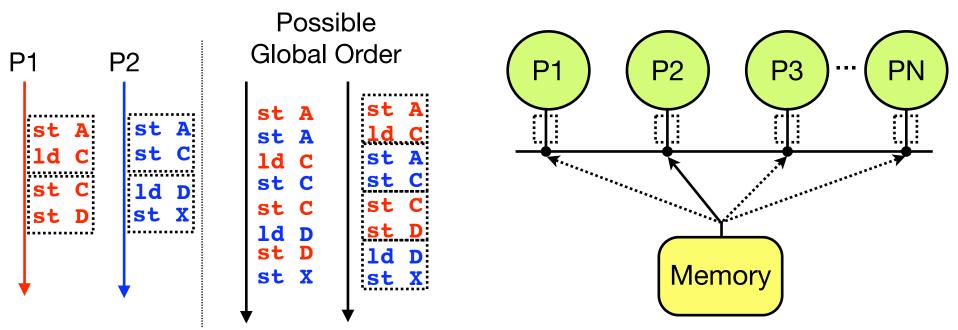


➡Group instructions into Chunks, enforce SC only at Chunk granularity

ma 5

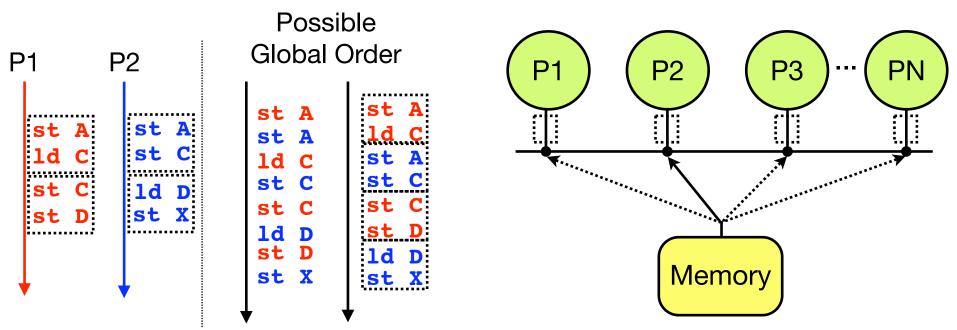


- ➡Group instructions into Chunks, enforce SC only at Chunk granularity
  - 1. substantially reduce hardware complexity



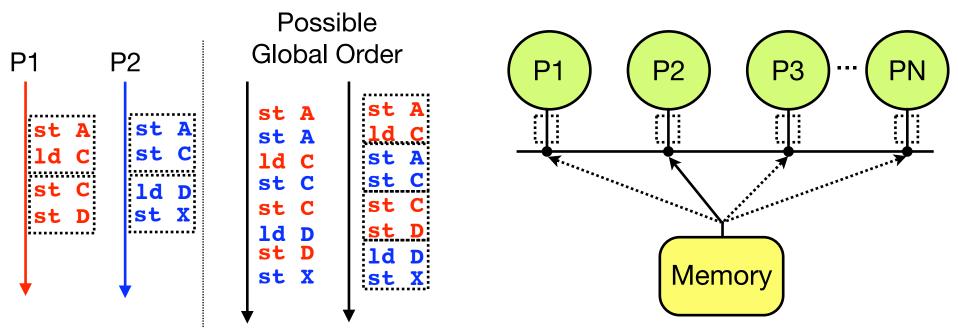
➡Group instructions into Chunks, enforce SC only at Chunk granularity

- 1. substantially reduce hardware complexity
- 2. enable high performance



➡Group instructions into Chunks, enforce SC only at Chunk granularity

- 1. substantially reduce hardware complexity
- 2. enable high performance
- 3. retain programmability



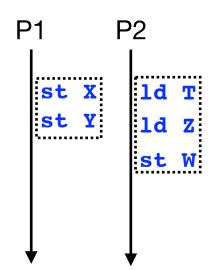
➡Group instructions into Chunks, enforce SC only at Chunk granularity

- 1. substantially reduce hardware complexity
- 2. enable high performance
- 3. retain programmability
- Execute a chunk atomically and in isolation, like a single instruction

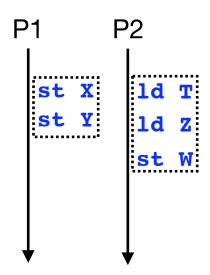








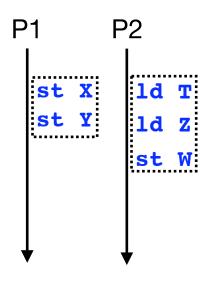




Speculative Execution



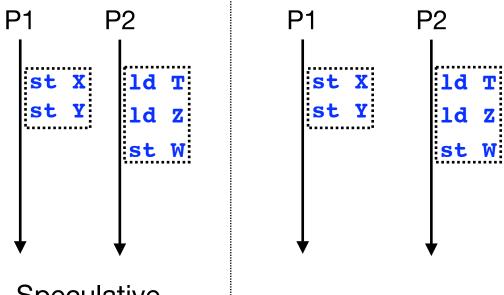




Speculative Execution

Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)

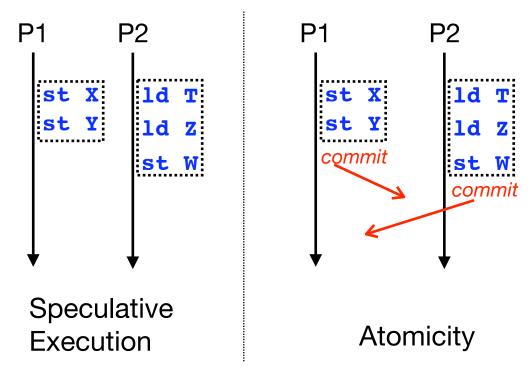




Speculative Execution

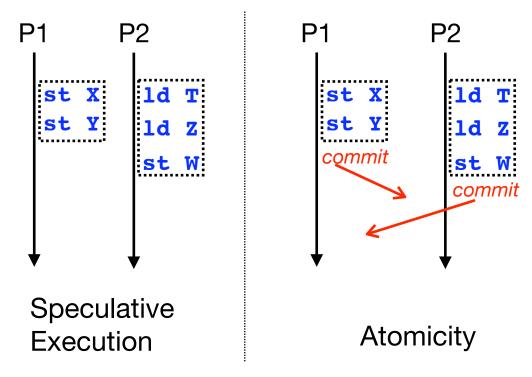
Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)





Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)

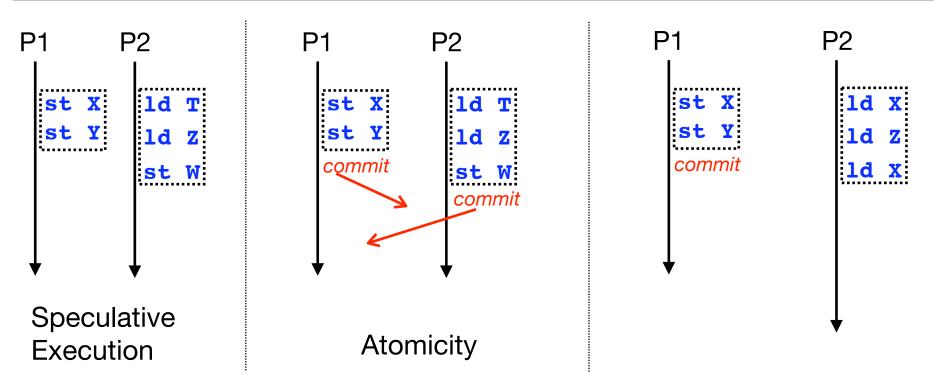




Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing) Isolation: a chunk should not see "outside" state

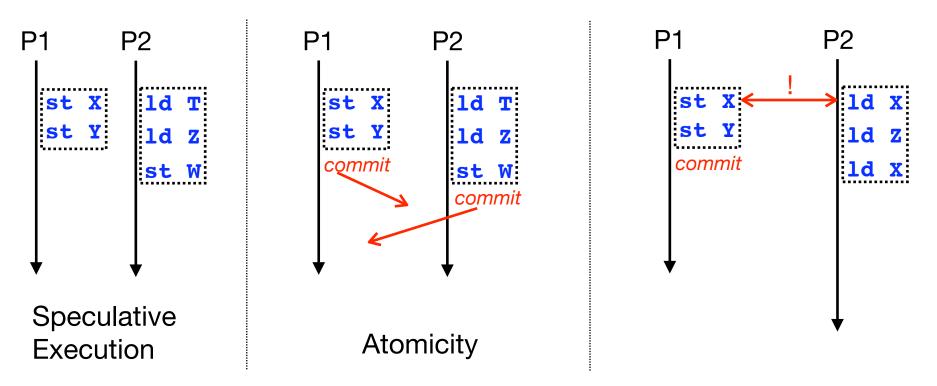
changing during its execution





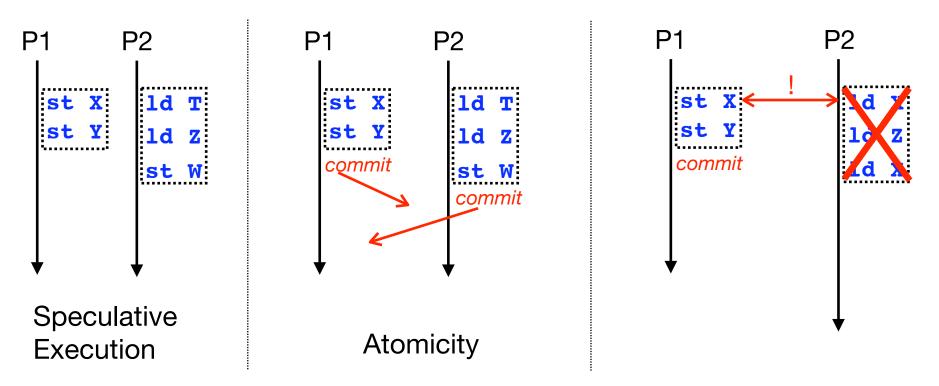
Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)





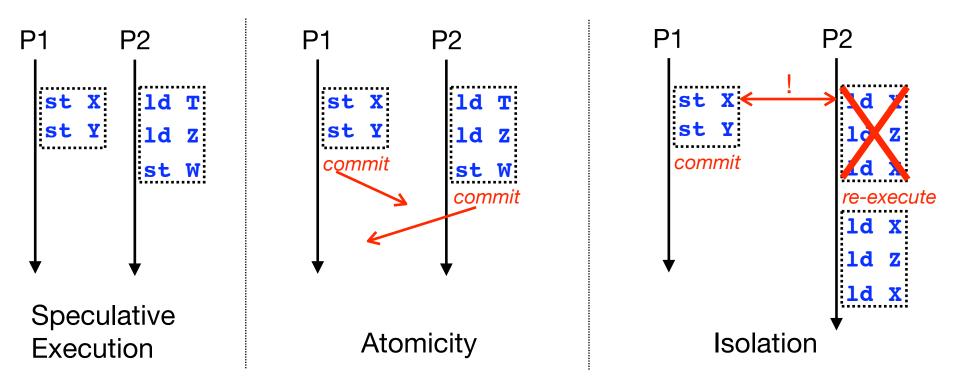
Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)





Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)





Atomicity: all updates in the chunk are made visible to other processors at once (all or nothing)



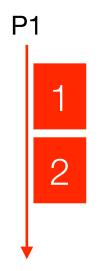


#### Per-processor program order: *chunks* from individual processors maintain program order





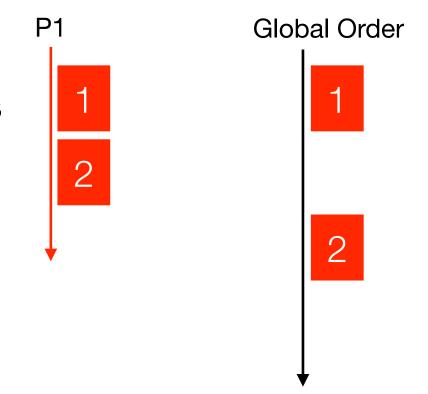
#### **Per-processor program order**: *chunks* from individual processors maintain program order







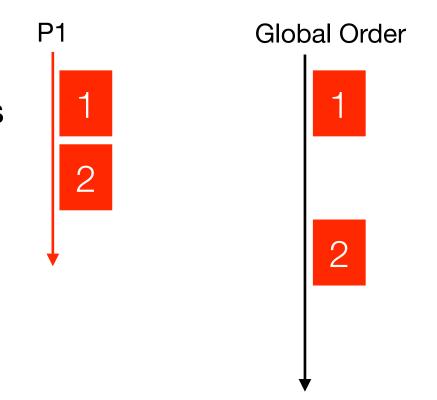
#### **Per-processor program order**: *chunks* from individual processors maintain program order





#### **Per-processor program order**: *chunks* from individual processors maintain program order

### **Single sequential order**: *chunks* from all processors maintain a single sequential order

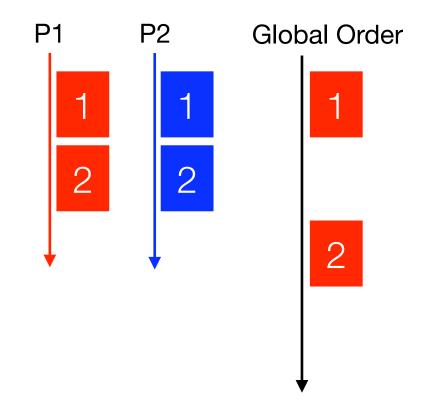






#### **Per-processor program order**: *chunks* from individual processors maintain program order

### **Single sequential order**: *chunks* from all processors maintain a single sequential order

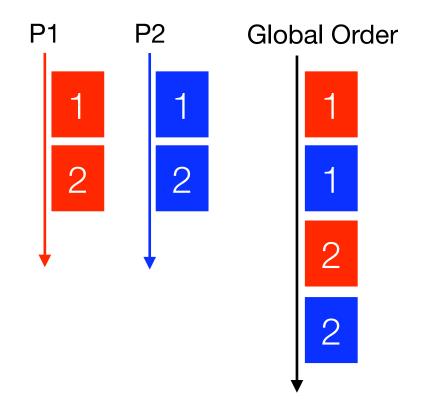






#### **Per-processor program order**: *chunks* from individual processors maintain program order

### **Single sequential order**: *chunks* from all processors maintain a single sequential order

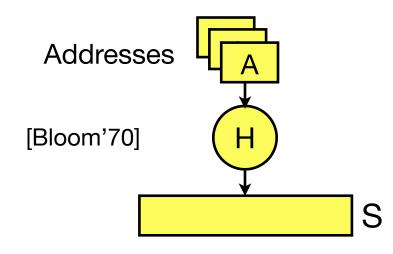


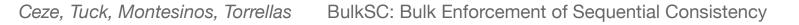




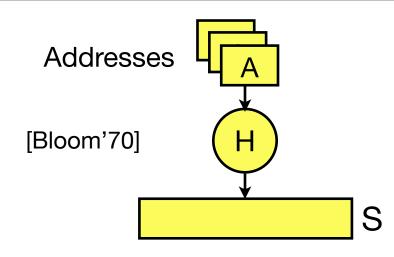


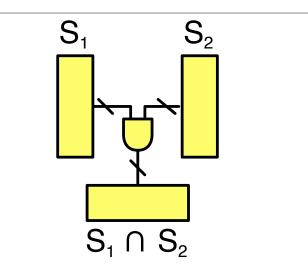






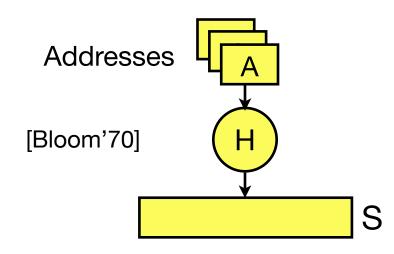


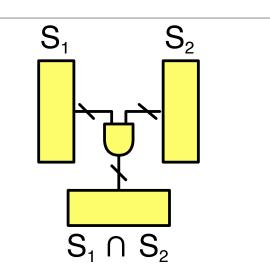








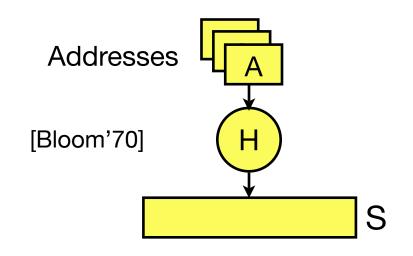


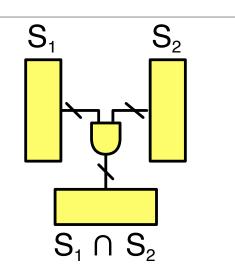


#### **Signature Operations**



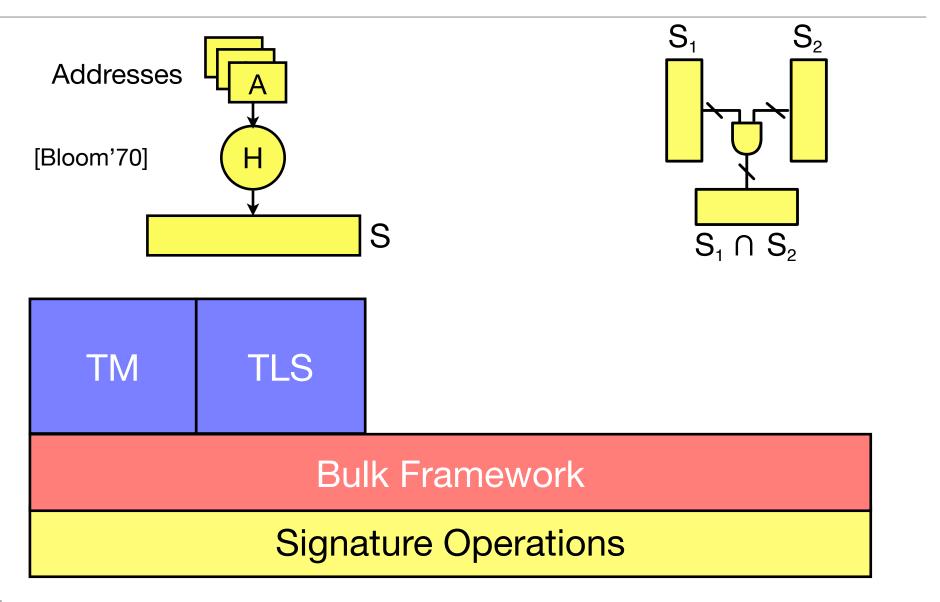




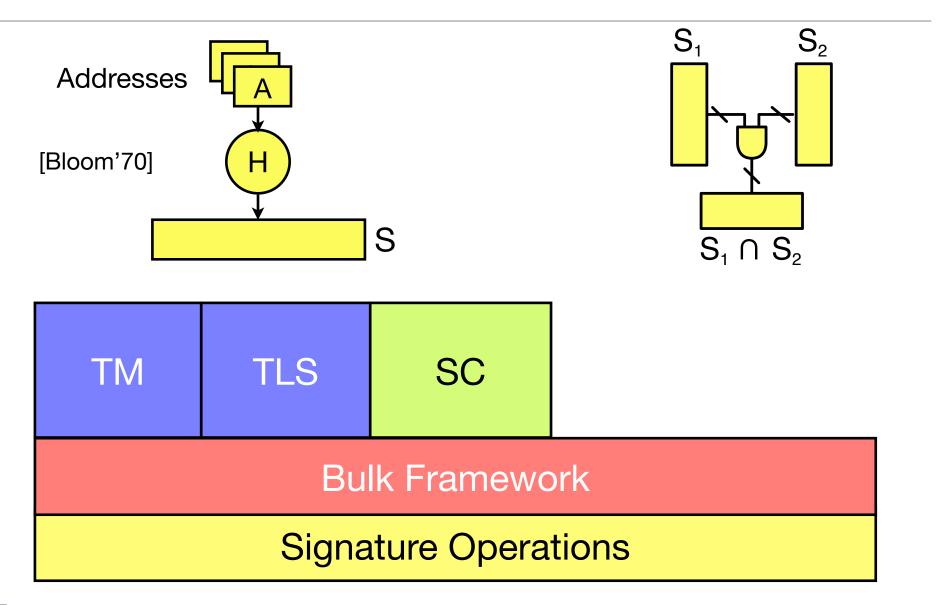




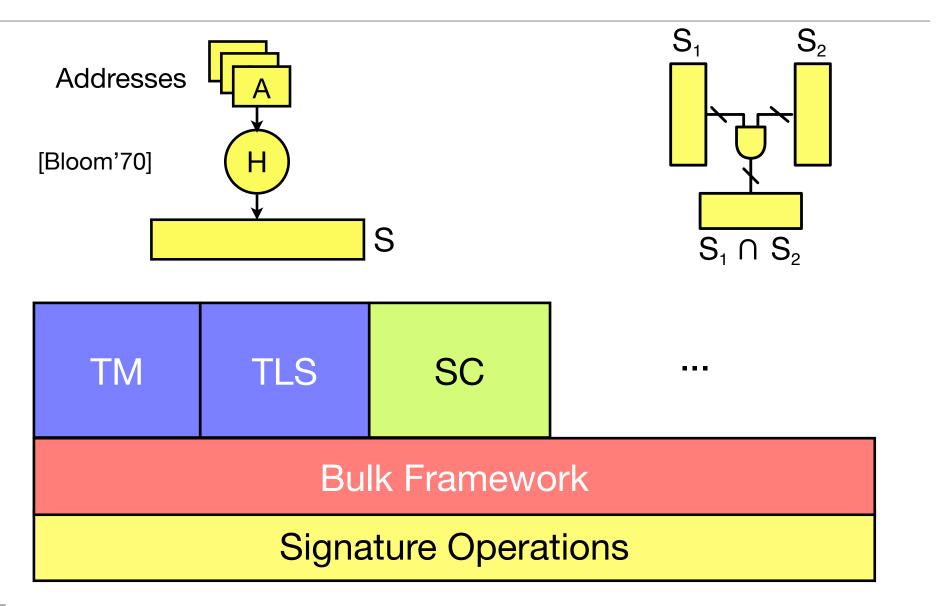














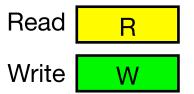




 Key idea: Summarize in hardware addresses accessed by a chunk into a pair of signatures



•Key idea: Summarize in hardware addresses accessed by a chunk into a pair of signatures





•Key idea: Summarize in hardware addresses accessed by a chunk into a pair of signatures



- Support signature operations with simple hardware
  - intersection, union, is-empty?, ...
  - much of the system operation is based on signatures



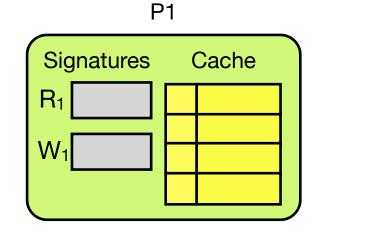
# **Efficiently Operating with Chunks**

•Key idea: Summarize in hardware addresses accessed by a chunk into a pair of signatures

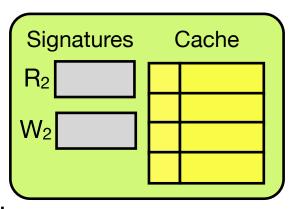


- Support signature operations with simple hardware
  - intersection, union, is-empty?, ...
  - much of the system operation is based on signatures
- •At chunk commit
  - W signature is sent to other processors for disambiguation
  - R, W signatures are cleared

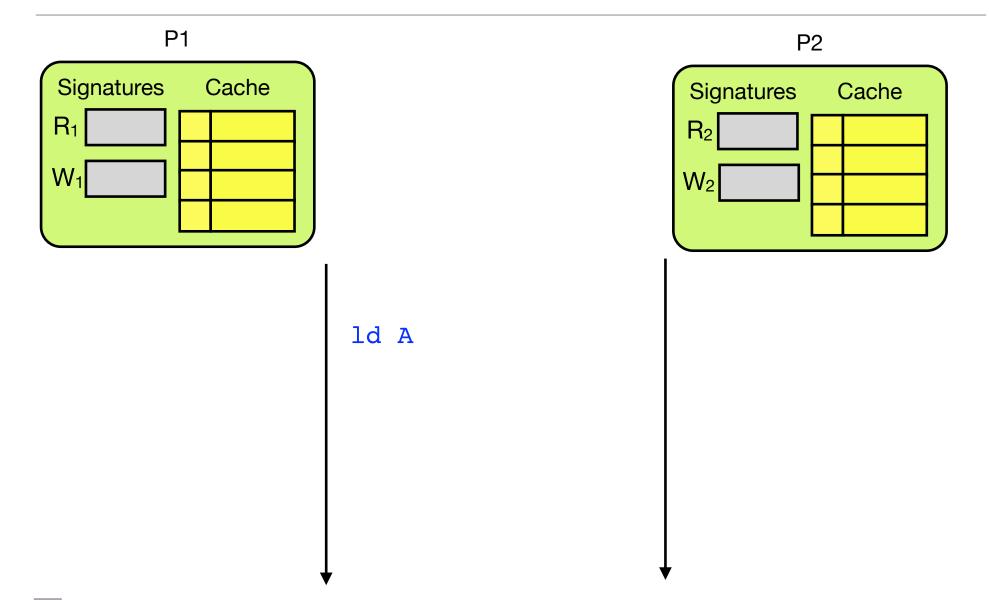




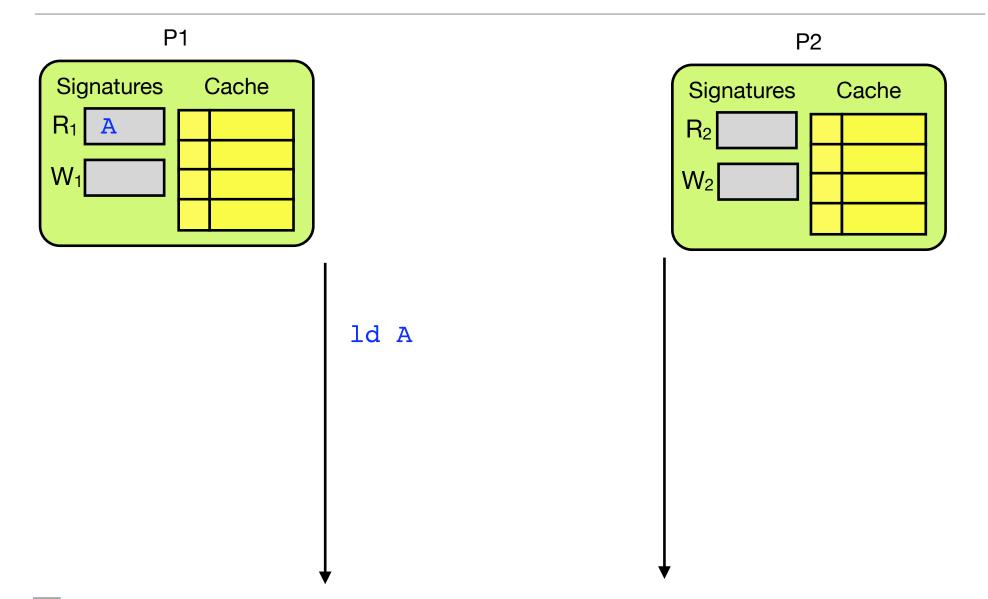




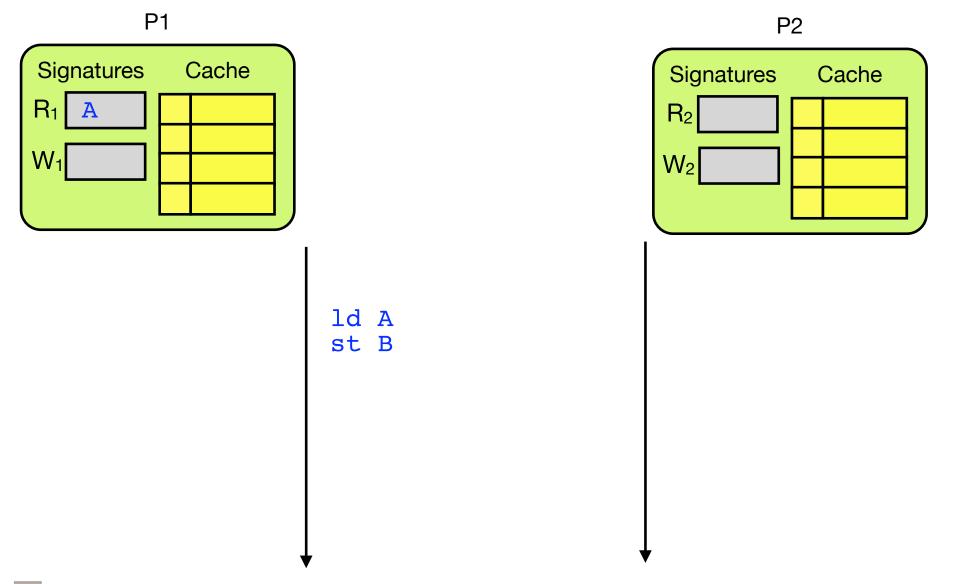




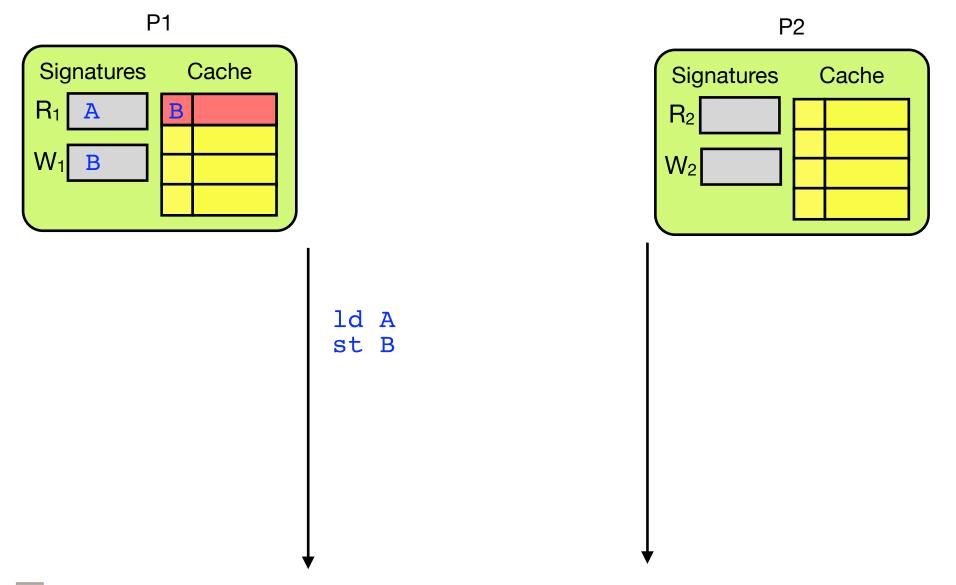




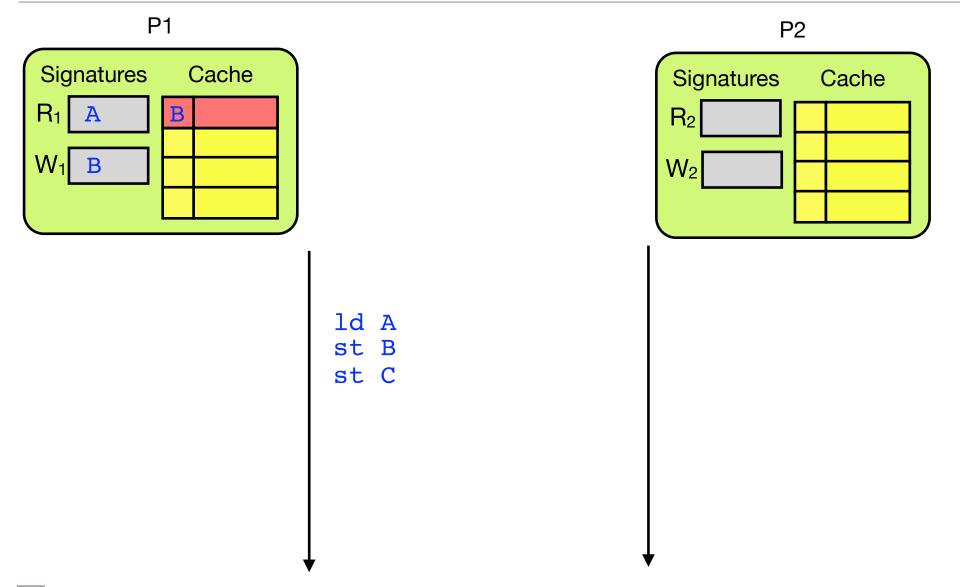




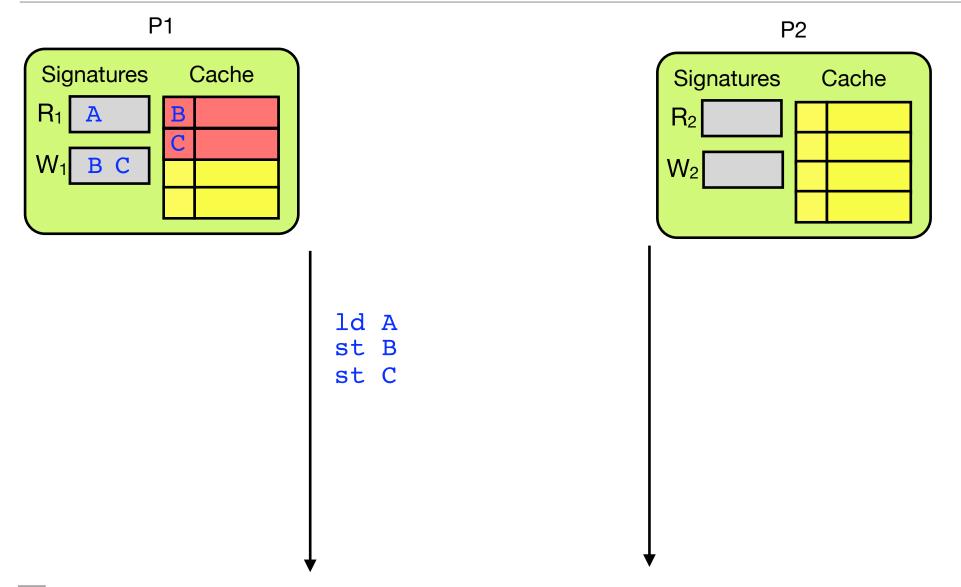




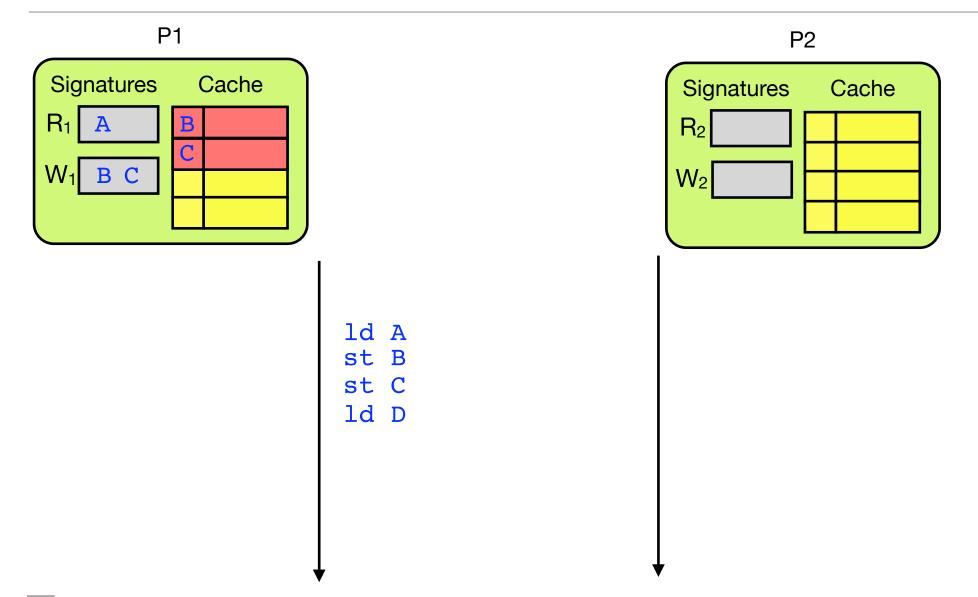




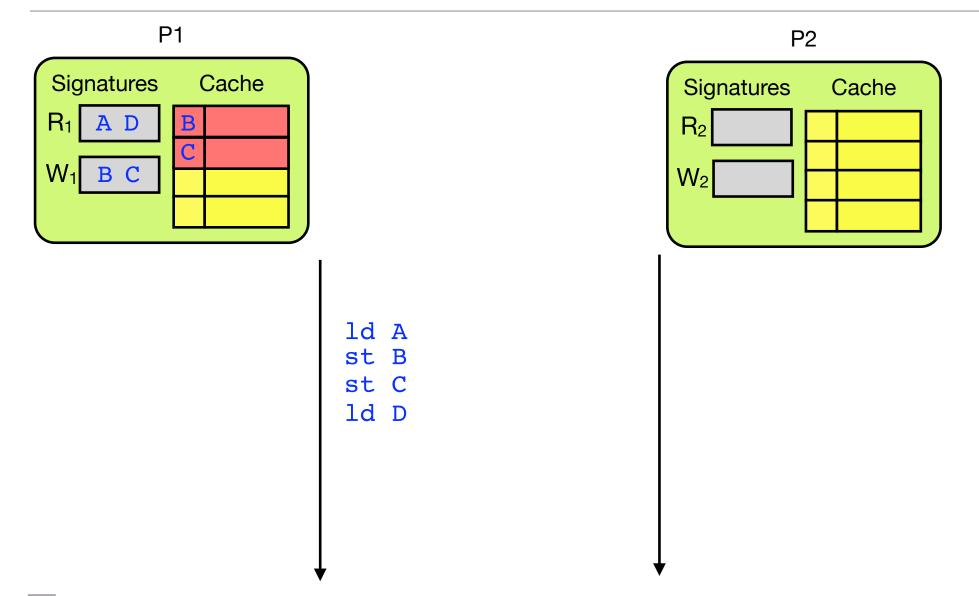




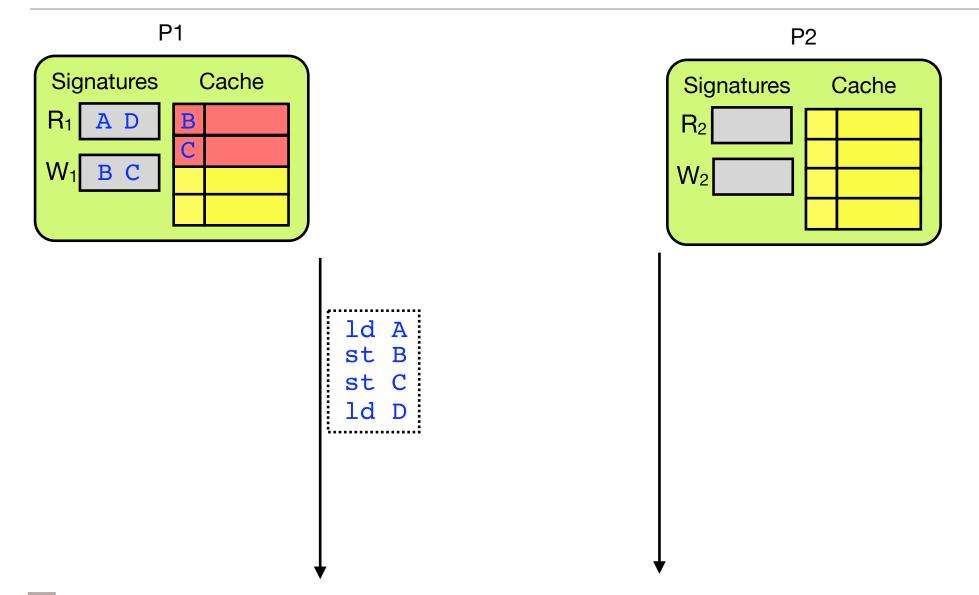




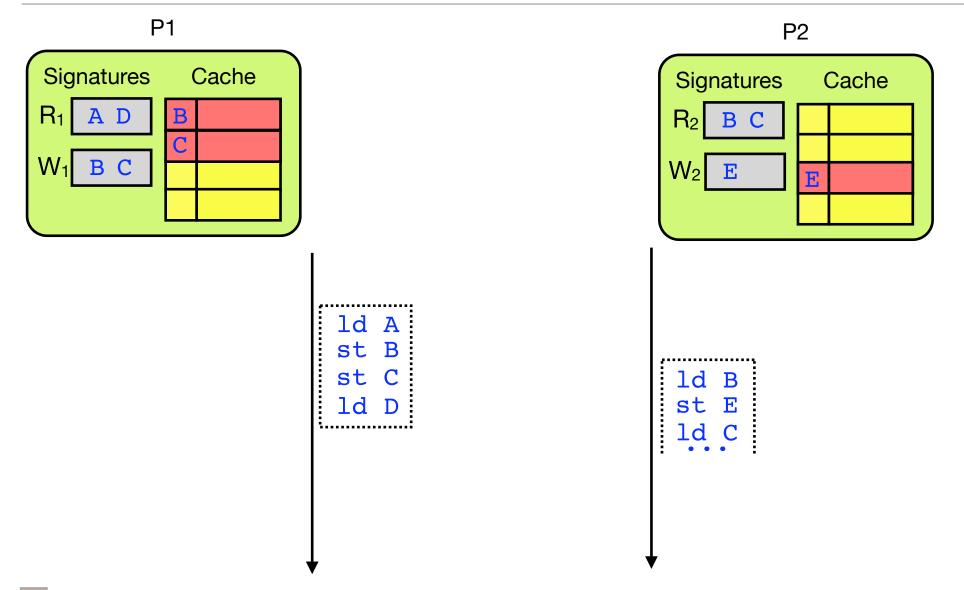




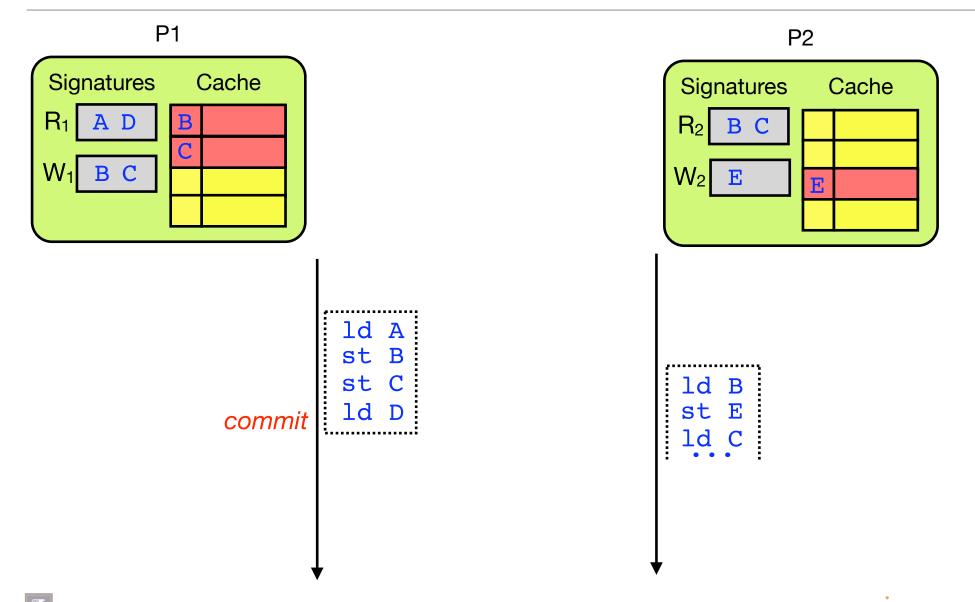






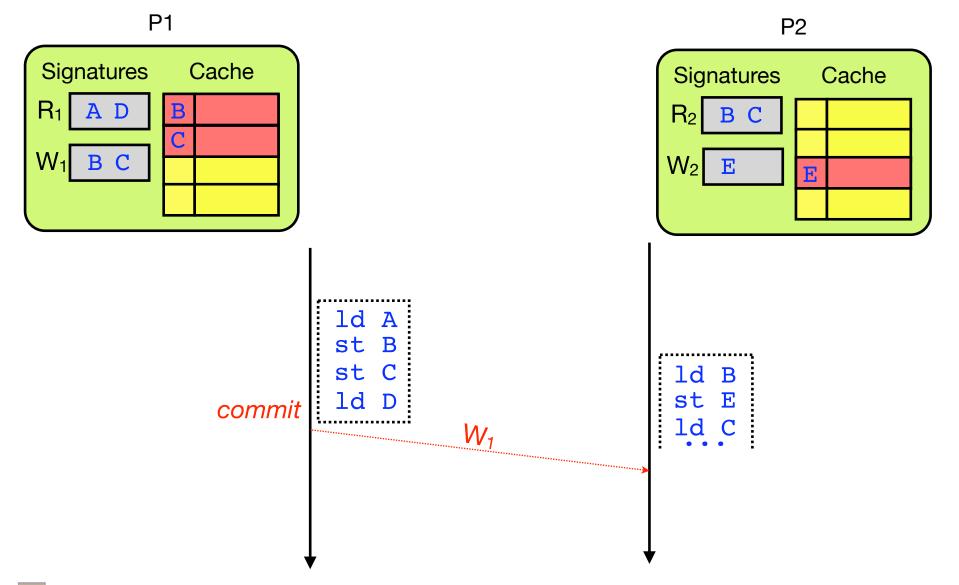


i-acoma 10

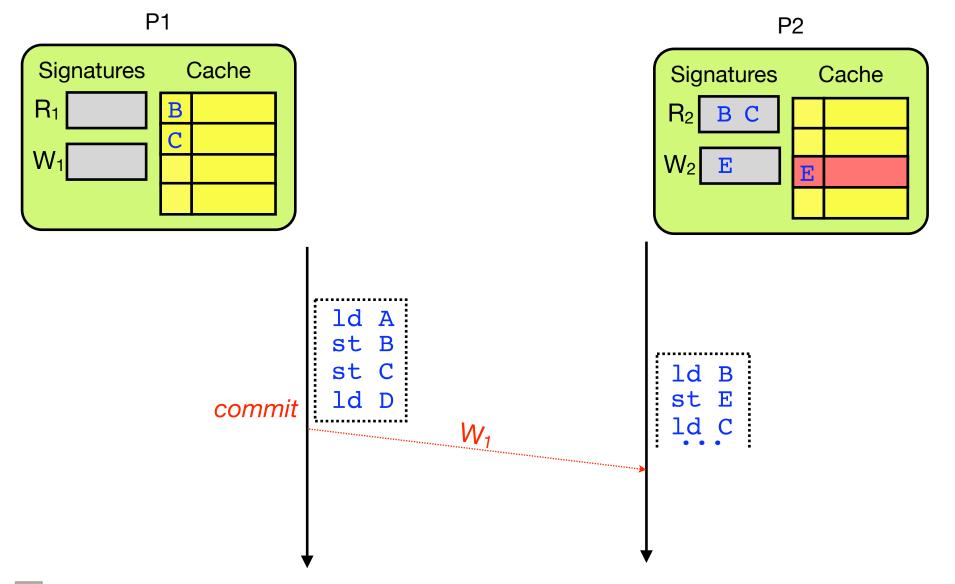


aroup

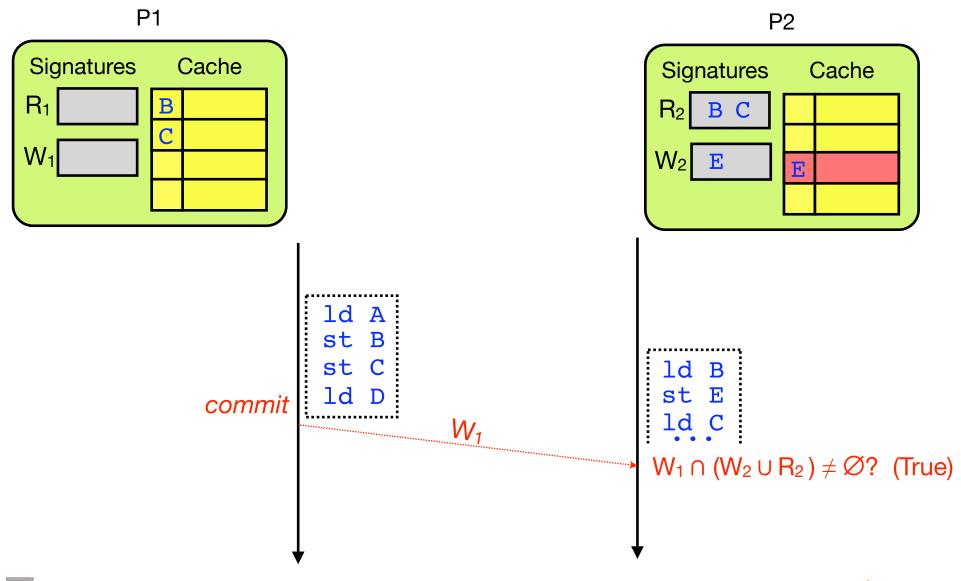
10



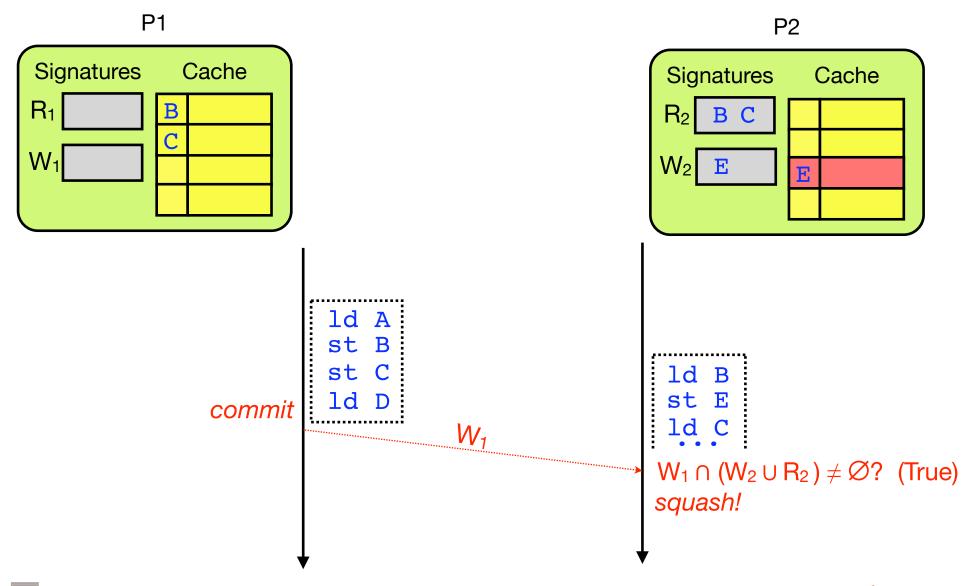
i-acoma 10



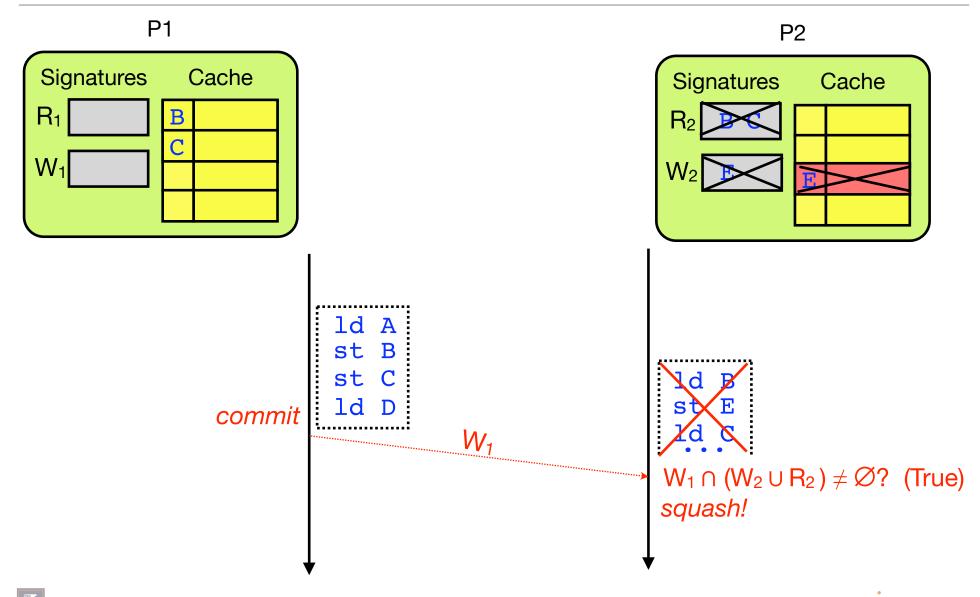
i-acoma 10



-acoma 10



-acoma 10



10

group

Ceze, Tuck, Montesinos, Torrellas BulkSC: Bulk Enforcement of Sequential Consistency



• Chunks are arbitrarily defined by the hardware





- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions





- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*





- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:



- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:
  - Instructions inside a chunk arbitrarily reordered

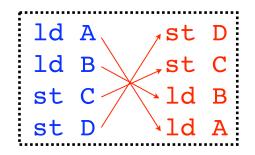


- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in program order
- High-performance:
  - Instructions inside a chunk arbitrarily reordered

ld	A
ld	В
st	С
st	D

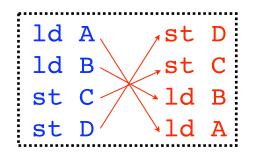


- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:
  - Instructions inside a chunk arbitrarily reordered





- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:
  - Instructions inside a chunk arbitrarily reordered



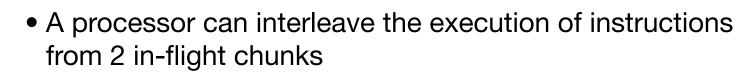
 A processor can interleave the execution of instructions from 2 in-flight chunks

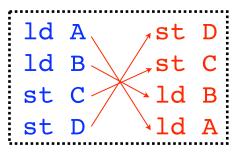


- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:

Ceze. Tuck. Montesinos. Torrellas

• Instructions inside a chunk arbitrarily reordered

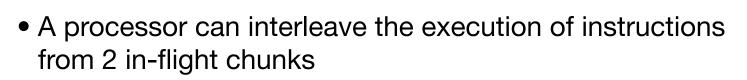


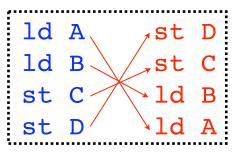






- Chunks are arbitrarily defined by the hardware
  - e.g. every 2000 dynamic instructions
- A processor commits chunks in *program order*
- High-performance:
  - Instructions inside a chunk arbitrarily reordered



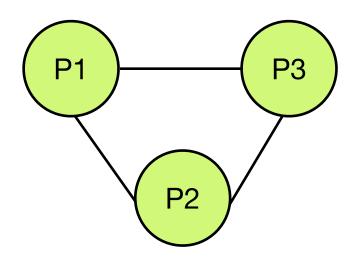






Ceze, Tuck, Montesinos, Torrellas BulkSC: Bulk Enforcement of Sequential Consistency

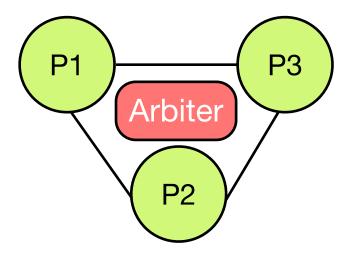








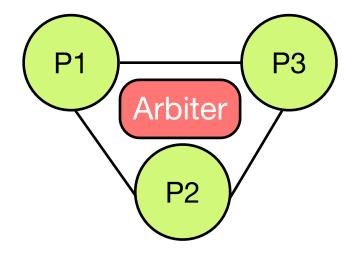
Need an Arbiter for chunk commits







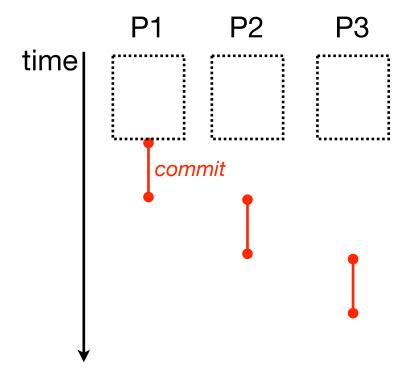
- Need an Arbiter for chunk commits
- •Naive:
  - allow a single commit at a time

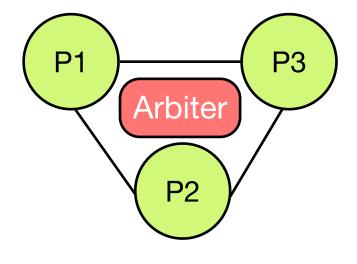






- Need an Arbiter for chunk commits
- •Naive:
  - allow a single commit at a time





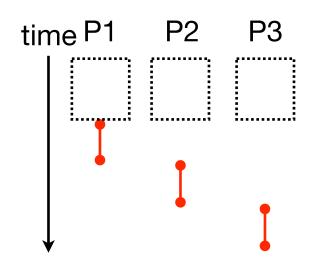


## **Allowing Simultaneous Chunk Commits**



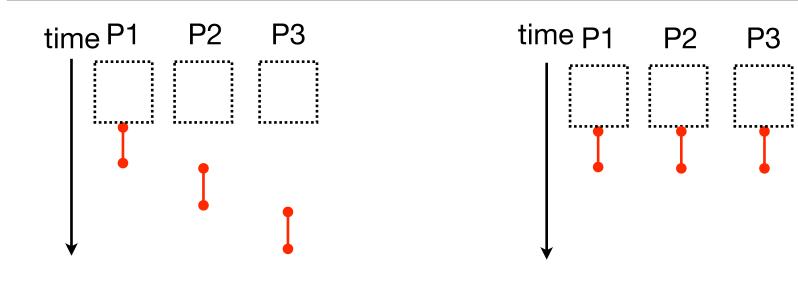


## **Allowing Simultaneous Chunk Commits**

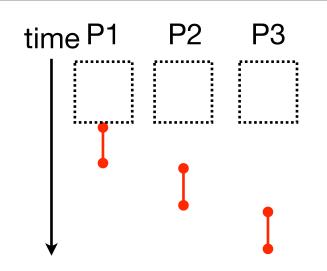


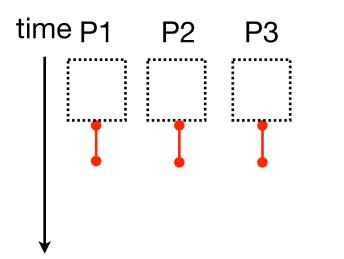






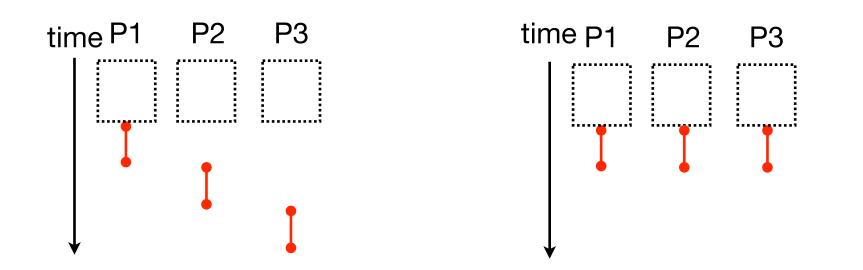






• Conditions:

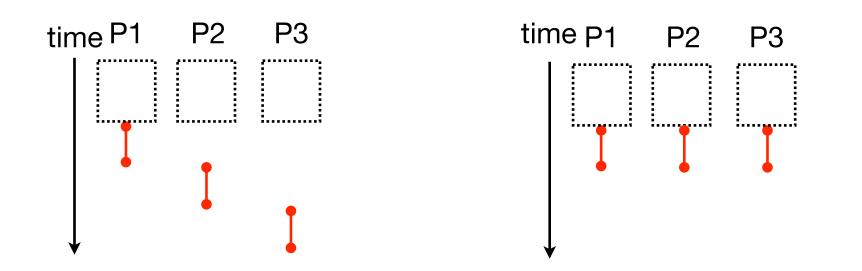




#### • Conditions:

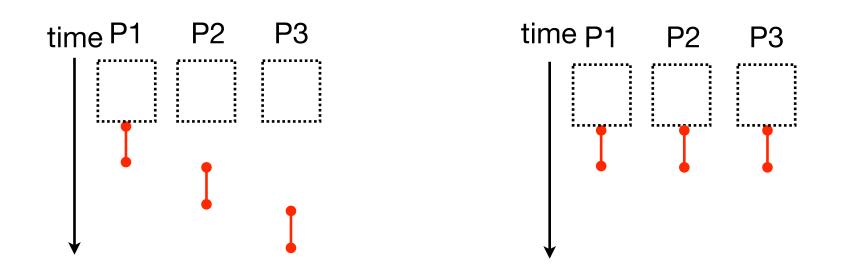
• committing chunks' memory operations should not intersect





#### • Conditions:

- committing chunks' memory operations should not intersect
  - both reads and writes



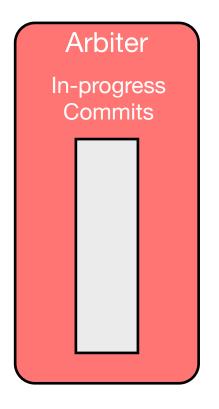
#### • Conditions:

- committing chunks' memory operations should not intersect
  - both reads and writes
- this can be efficiently enforced with chunk signatures

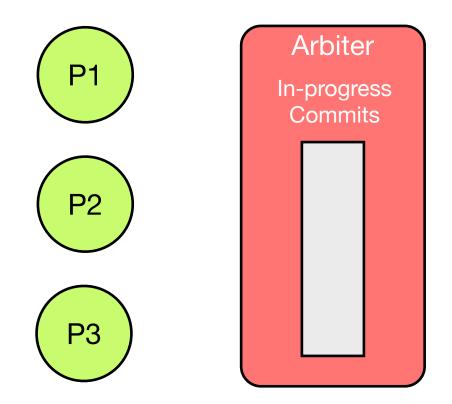




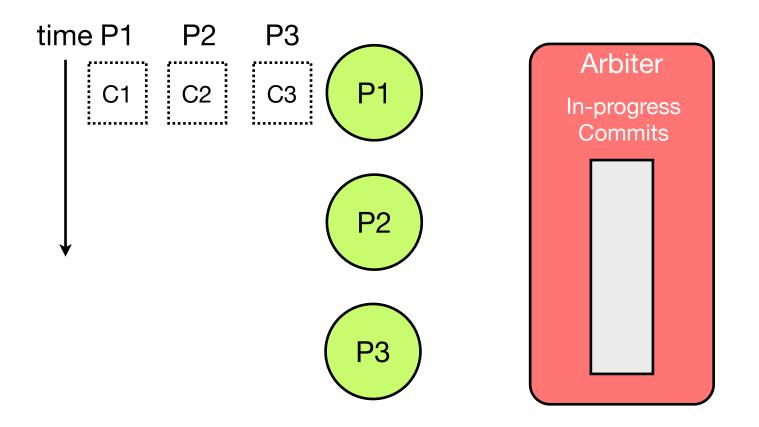




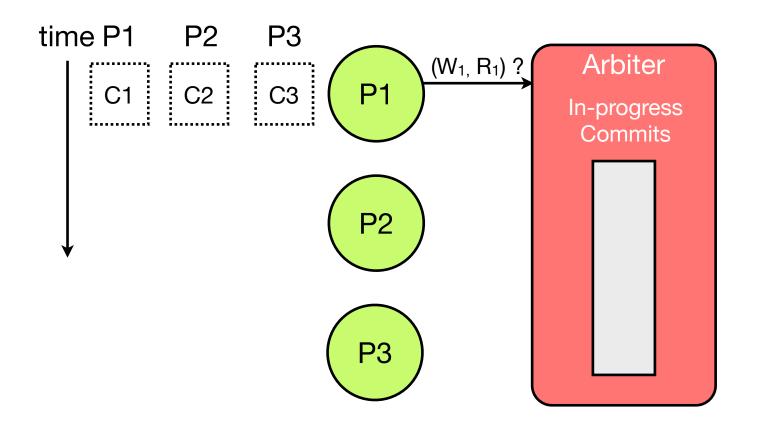




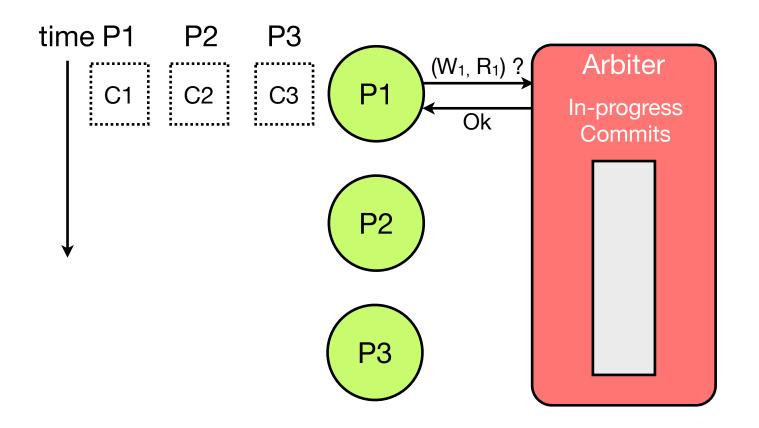




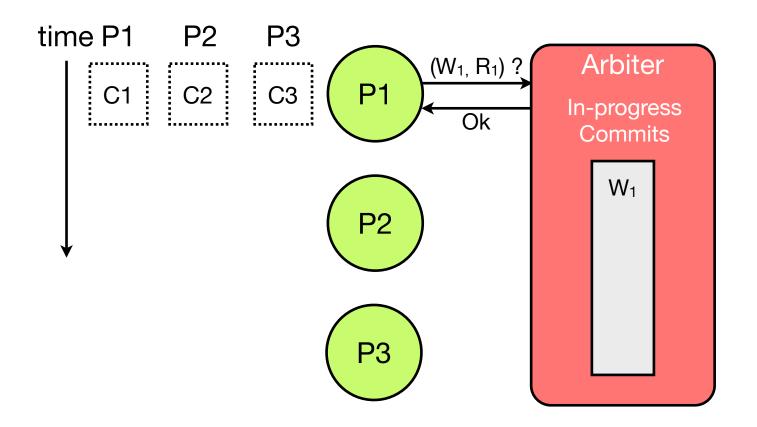




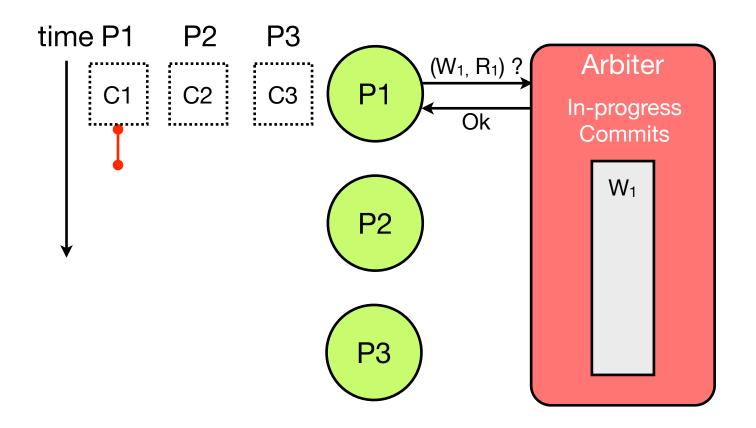




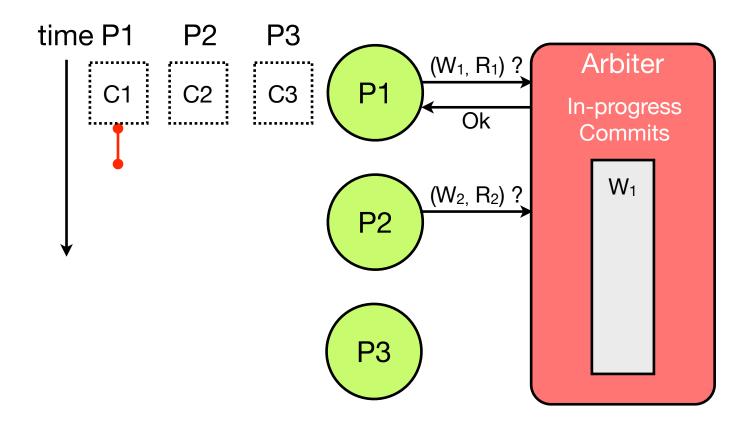




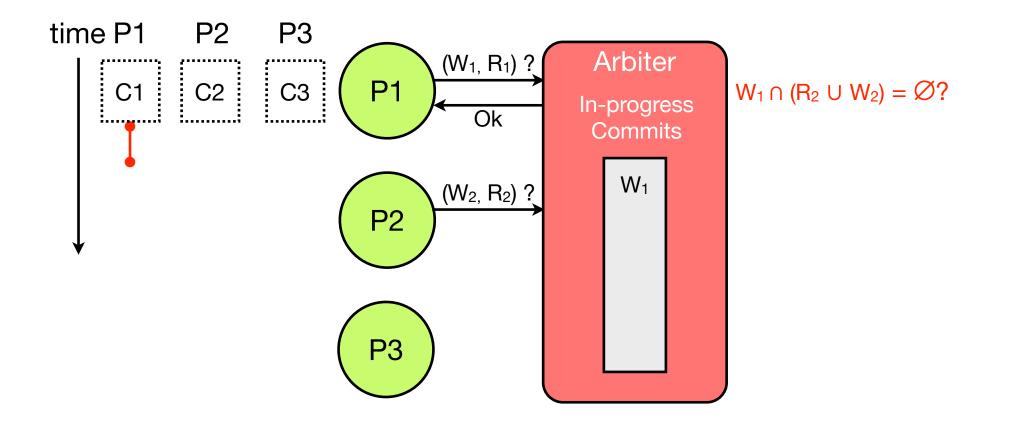




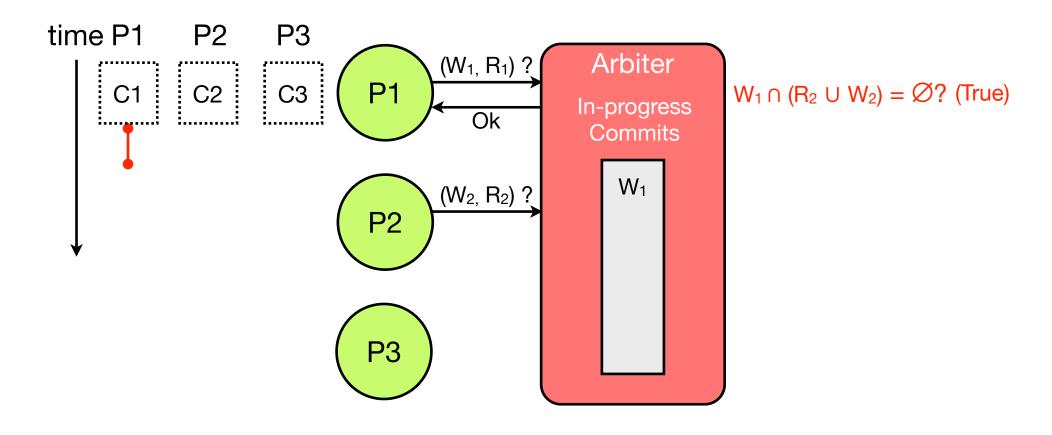




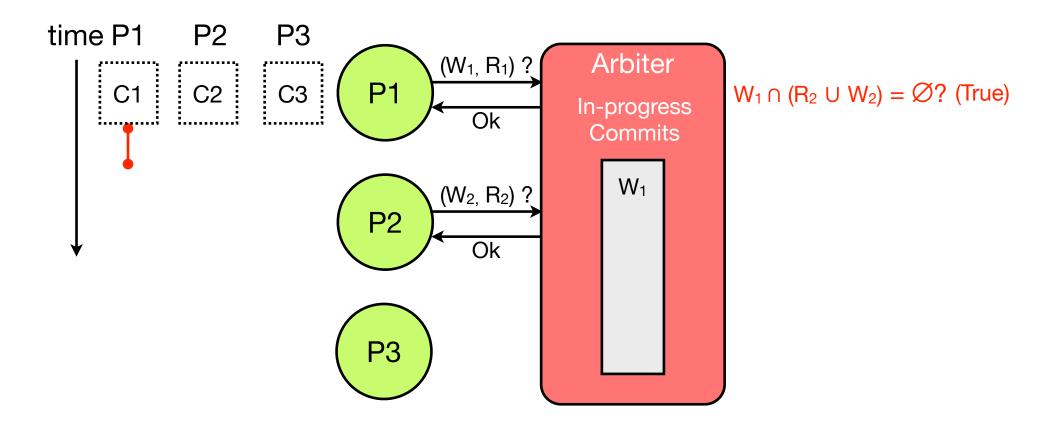




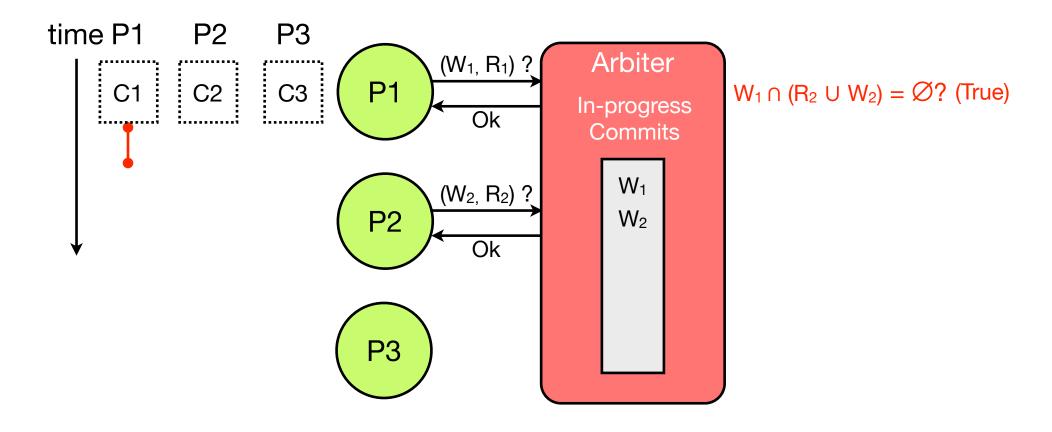




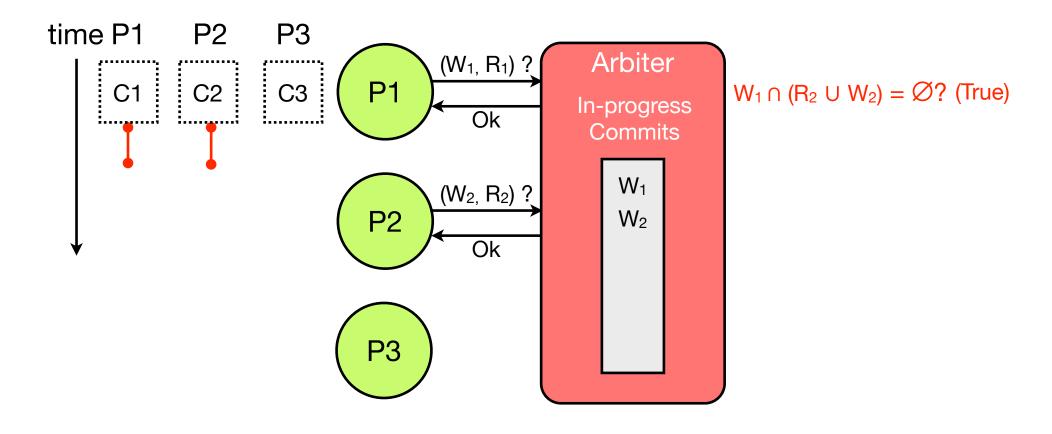




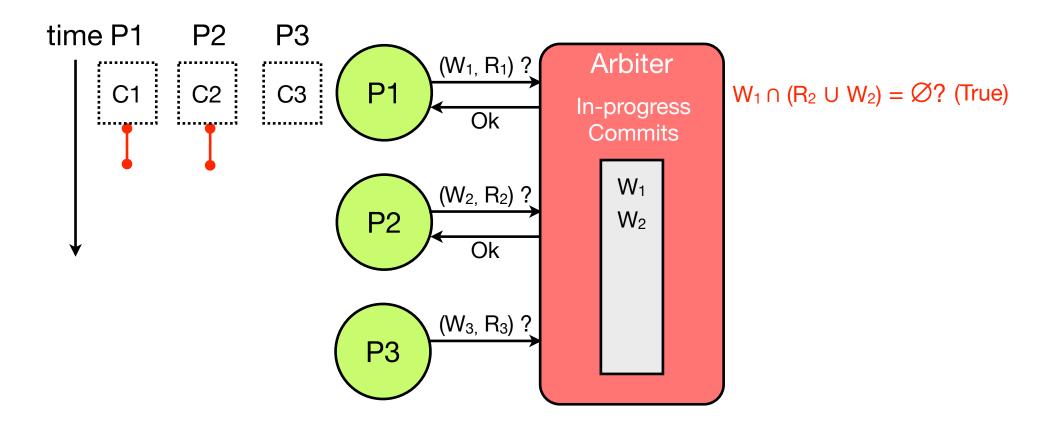




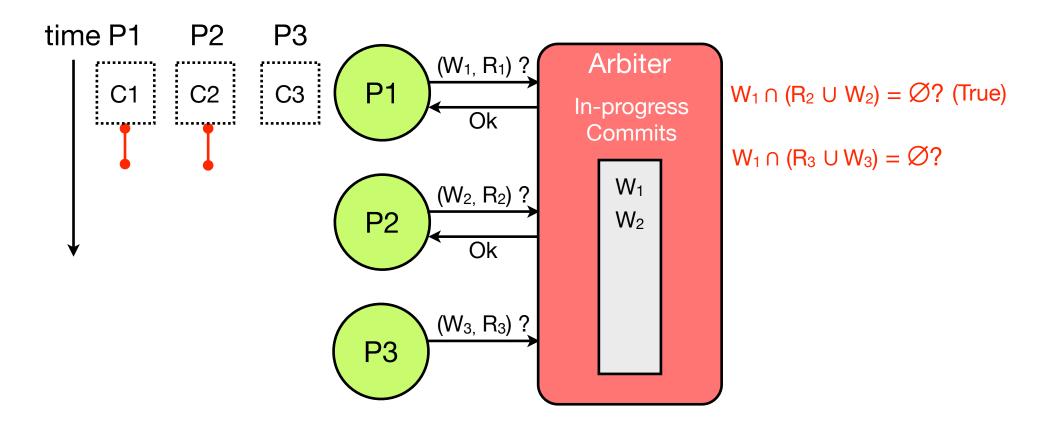




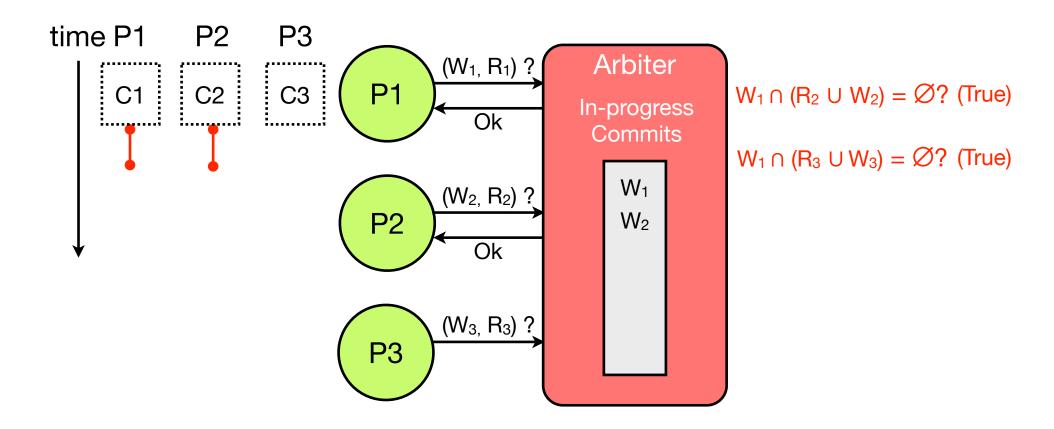




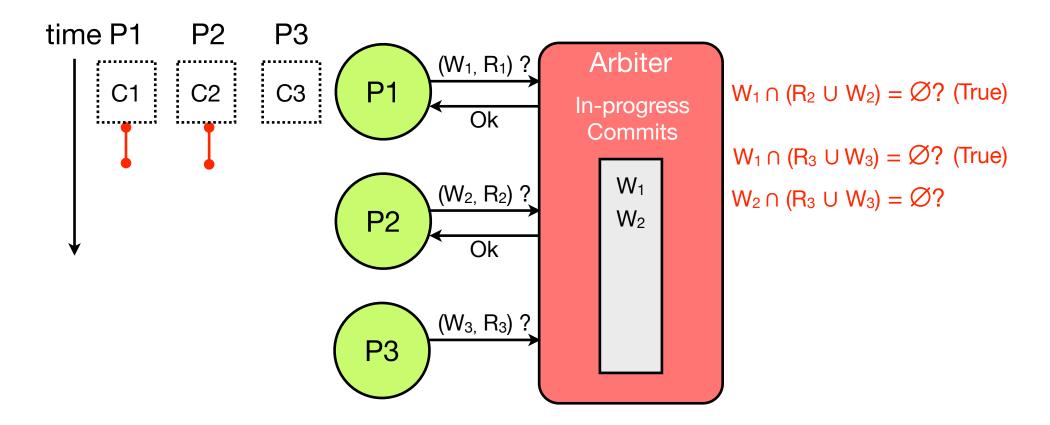




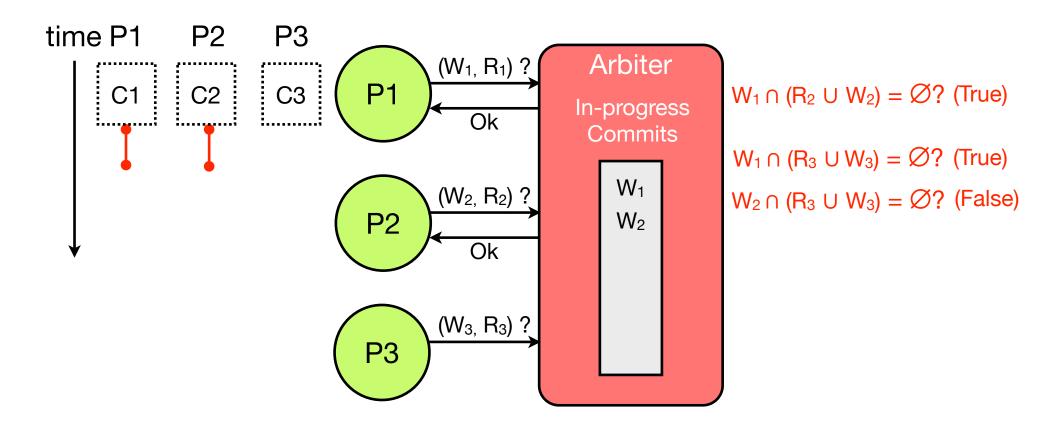




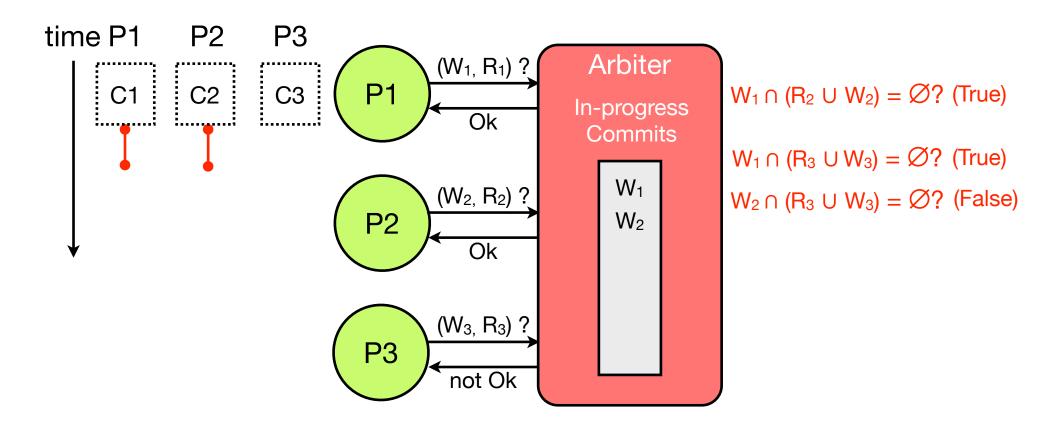




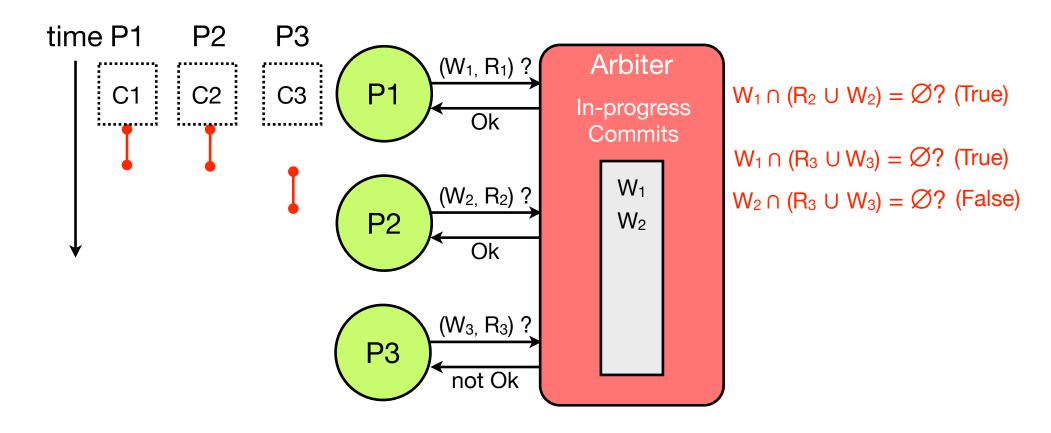




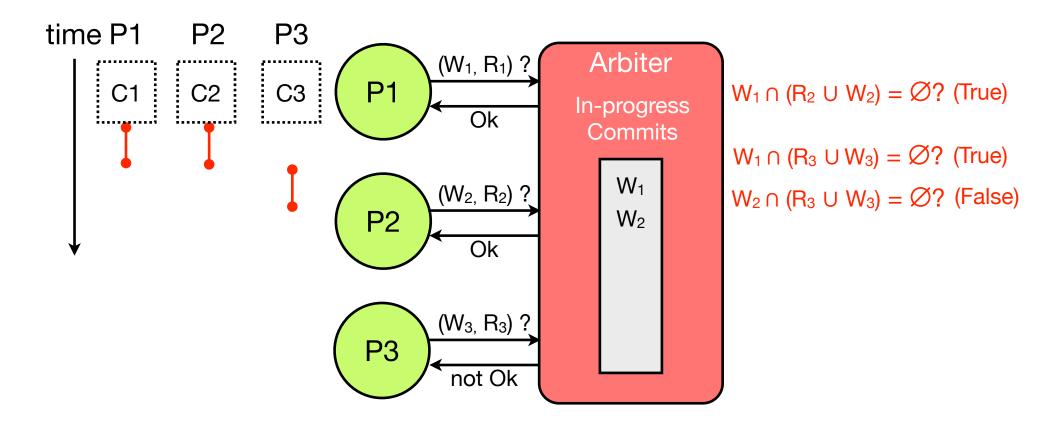












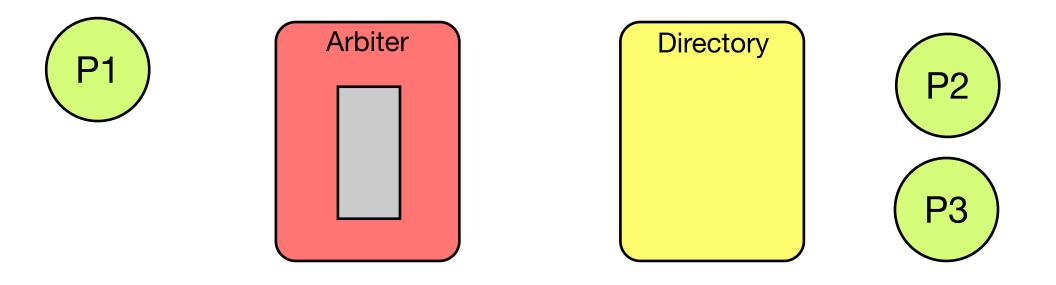
Write signatures stay in the arbiter until commit completes



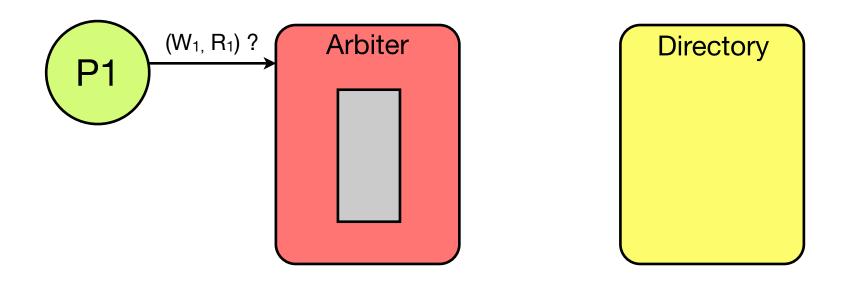


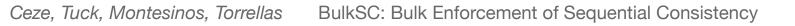
Ceze, Tuck, Montesinos, Torrellas BulkSC: Bulk Enforcement of Sequential Consistency







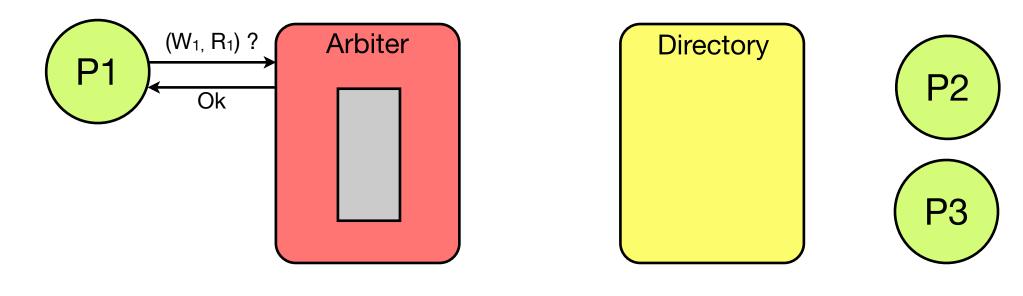




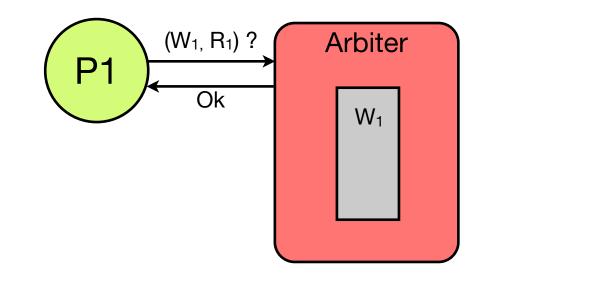


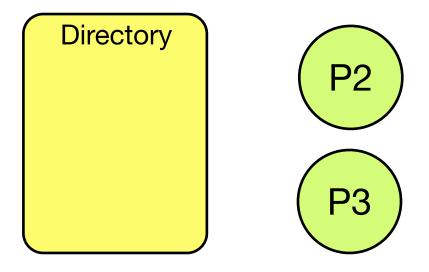
P2

**P**3

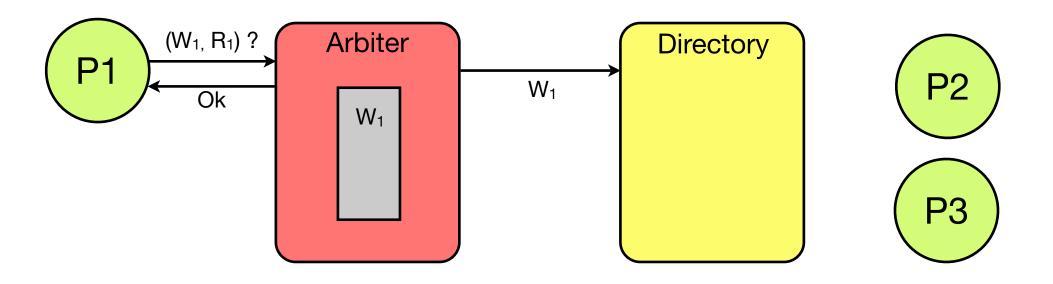




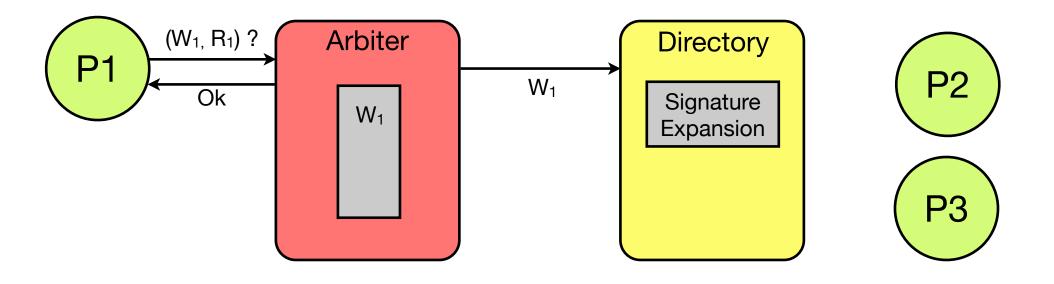




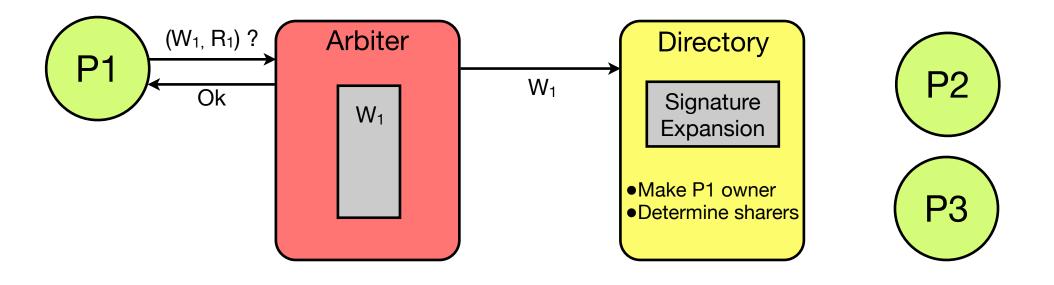
-acoma 15



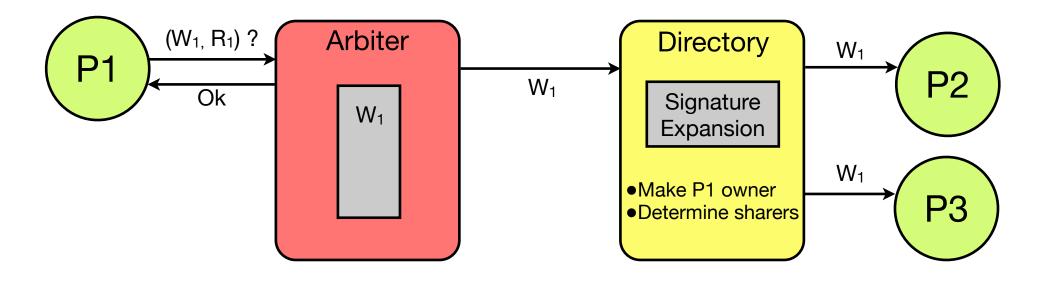




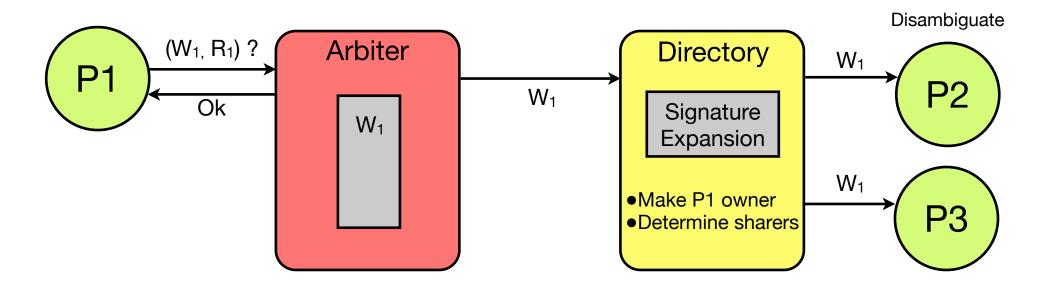
-acoma 15

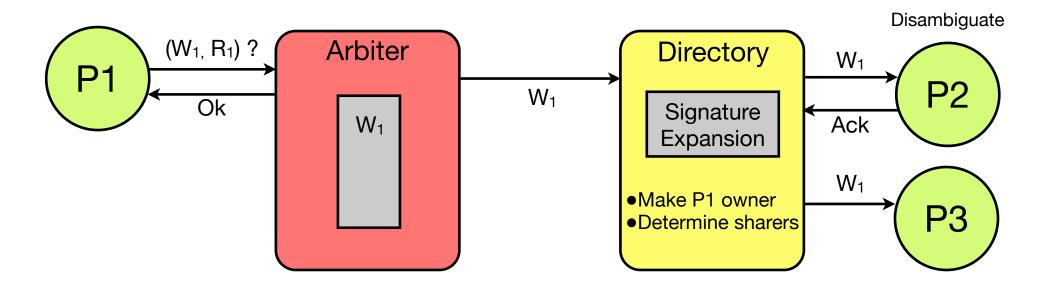




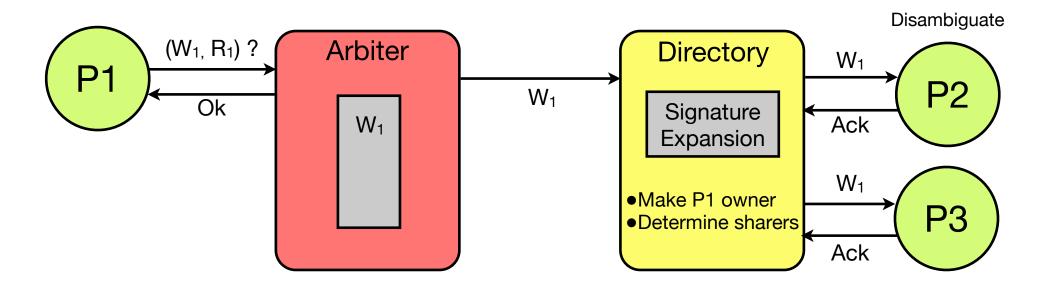


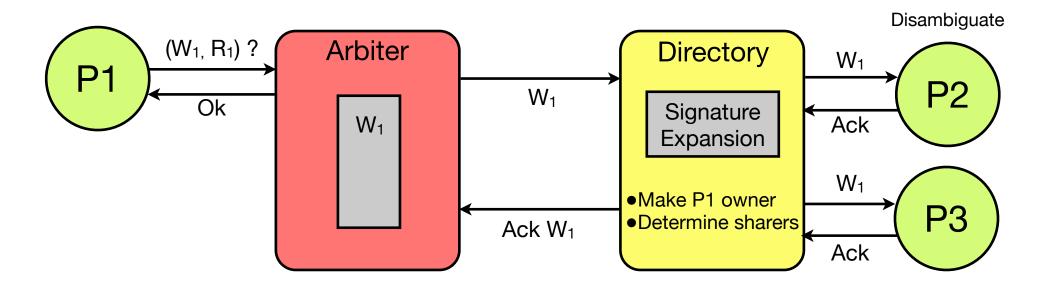
i-acoma 15

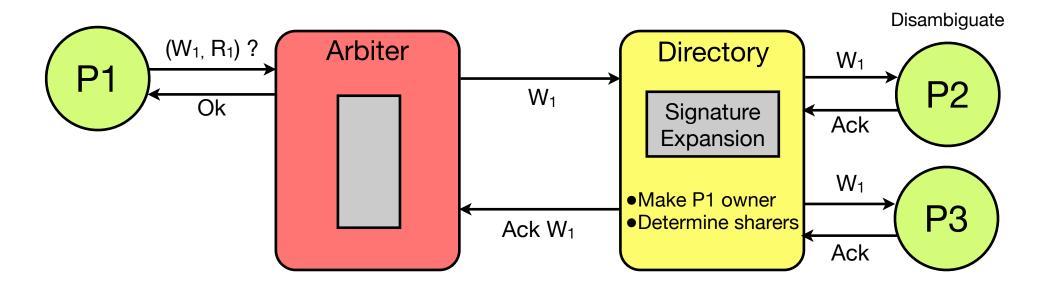


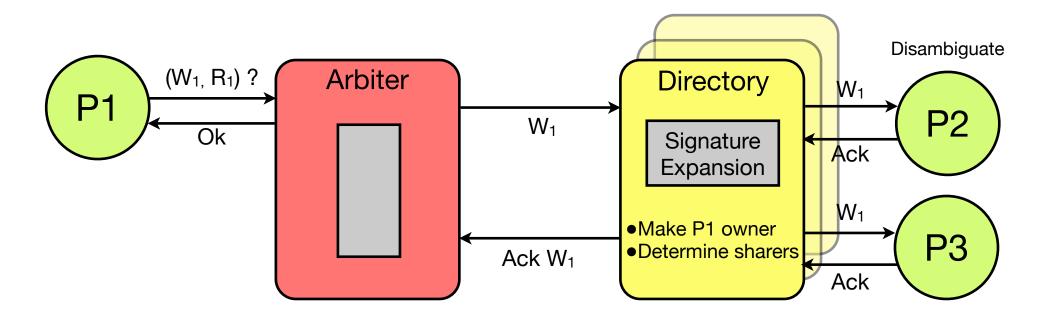


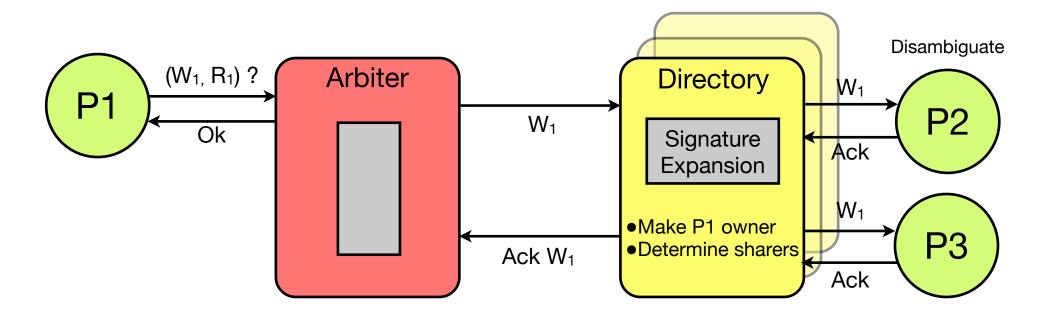








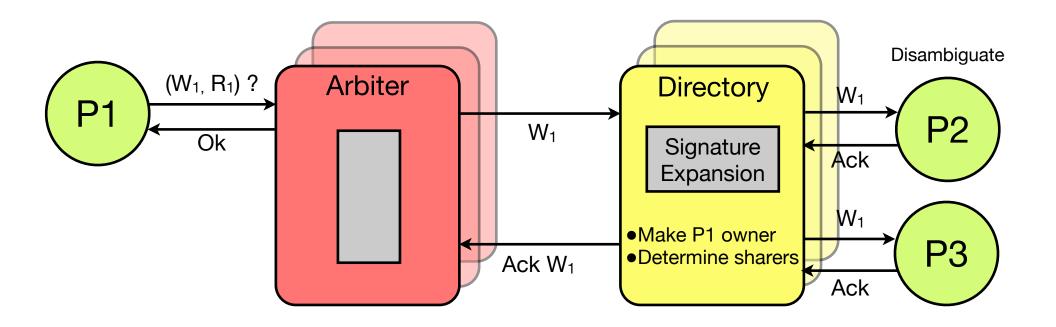




The whole process only uses signatures







The whole process only uses signatures Can support a distributed arbiter





Ceze, Tuck, Montesinos, Torrellas BulkSC: Bulk Enforcement of Sequential Consistency



• Simplifies the HW complexity of highperformance SC



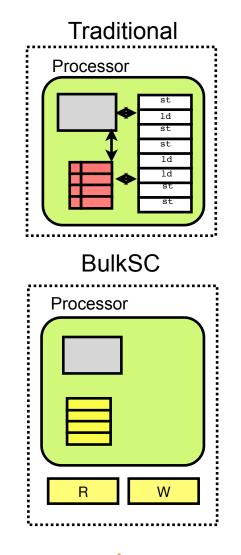


- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches





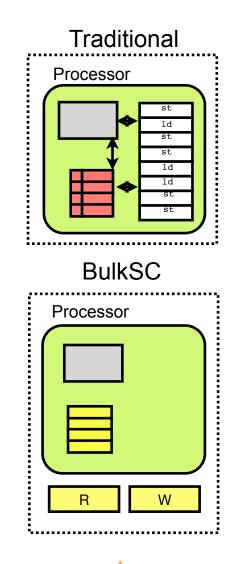
- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches







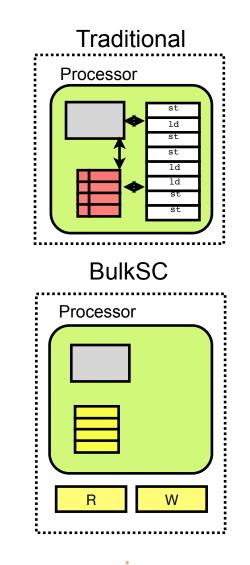
- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches
  - single-thread optimizations without worrying about multiprocessor issues





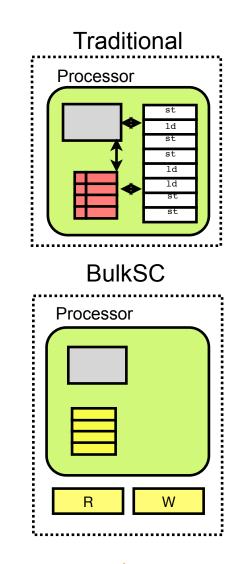


- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches
  - single-thread optimizations without worrying about multiprocessor issues
  - no associative structures in the processor



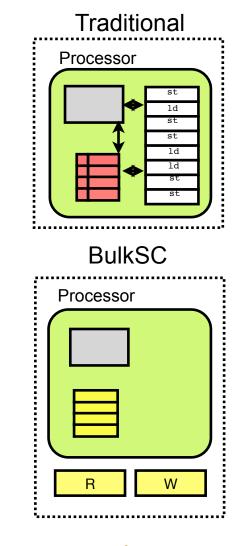
16

- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches
  - single-thread optimizations without worrying about multiprocessor issues
  - no associative structures in the processor
  - no extra bits in the cache for versioning [ISCA'06]





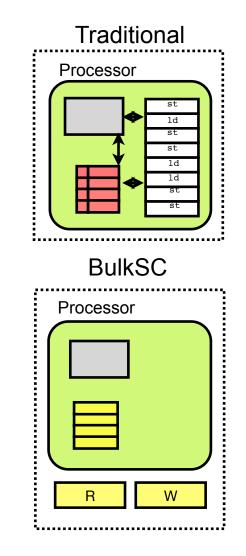
- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches
  - single-thread optimizations without worrying about multiprocessor issues
  - no associative structures in the processor
  - no extra bits in the cache for versioning [ISCA'06]
- Memory ordering framework for MPs







- Simplifies the HW complexity of highperformance SC
  - decouples consistency enforcement from core microarchitecture and caches
  - single-thread optimizations without worrying about multiprocessor issues
  - no associative structures in the processor
  - no extra bits in the cache for versioning [ISCA'06]
- Memory ordering framework for MPs
  - small extension to support TM, TLS, ...









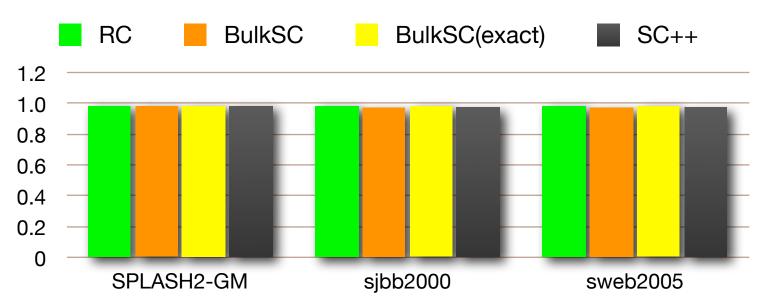


• Cycle-accurate simulations [SESC]



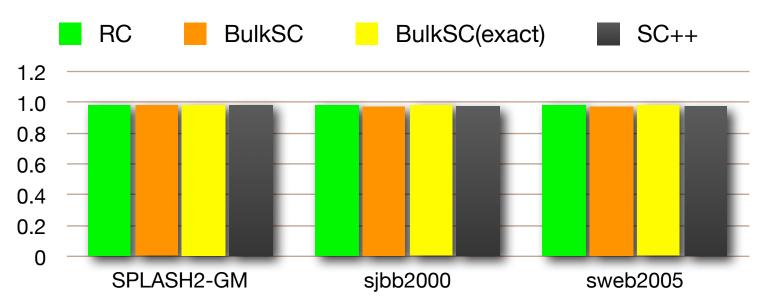


• Cycle-accurate simulations [SESC]



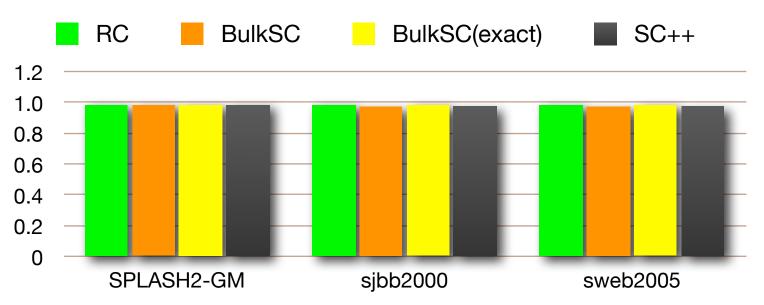


• Cycle-accurate simulations [SESC]



• Result: SC with performance comparable to RC (1%) with little BW cost (5-13%)

• Cycle-accurate simulations [SESC]



• Result: SC with performance comparable to RC (1%) with little BW cost (5-13%)

• Much simpler hardware than SC++





Interaction with explicit synchronization, TM





- Interaction with explicit synchronization, TM
- Forward progress



- Interaction with explicit synchronization, TM
- Forward progress
- •I/O





- Interaction with explicit synchronization, TM
- Forward progress
- •I/O
- Distributed arbiter





- Interaction with explicit synchronization, TM
- Forward progress
- •I/O
- Distributed arbiter
- Directory design for signatures



- Interaction with explicit synchronization, TM
- Forward progress
- •I/O
- Distributed arbiter
- Directory design for signatures
- Two optimizations for private data



- Interaction with explicit synchronization, TM
- Forward progress
- •I/O
- Distributed arbiter
- Directory design for signatures
- Two optimizations for private data
- Discussion on scalability







Presented BulkSC





- Presented BulkSC
  - coarse-grain, signature-based enforcement of SC





- Presented BulkSC
  - coarse-grain, signature-based enforcement of SC
- Performance comparable to RC





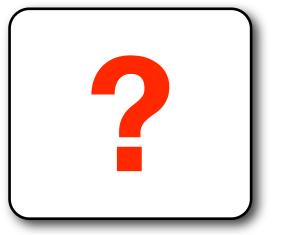
- Presented BulkSC
  - coarse-grain, signature-based enforcement of SC
- Performance comparable to RC
- Simplicity: decouple consistency enforcement from processor design



- Presented BulkSC
  - coarse-grain, signature-based enforcement of SC
- Performance comparable to RC
- Simplicity: decouple consistency enforcement from processor design
- Independent of network ordering properties



- Presented BulkSC
  - coarse-grain, signature-based enforcement of SC
- Performance comparable to RC
- Simplicity: decouple consistency enforcement from processor design
- Independent of network ordering properties
- Generic ordering framework for speculative multiprocessors



# BulkSC: Bulk Enforcement of Sequential Consistency

Luis Ceze, James Tuck, Pablo Montesinos, Josep Torrellas

University of Illinois at Urbana-Champaign http://iacoma.cs.uiuc.edu/



ISCA 2007, San Diego, CA

