

Today

Synch + Comm. in Shared
memory systems

* MCS paper

* Lecture videos

✓ * Parallel Systems Lessons

✓ * Spin lock algorithms

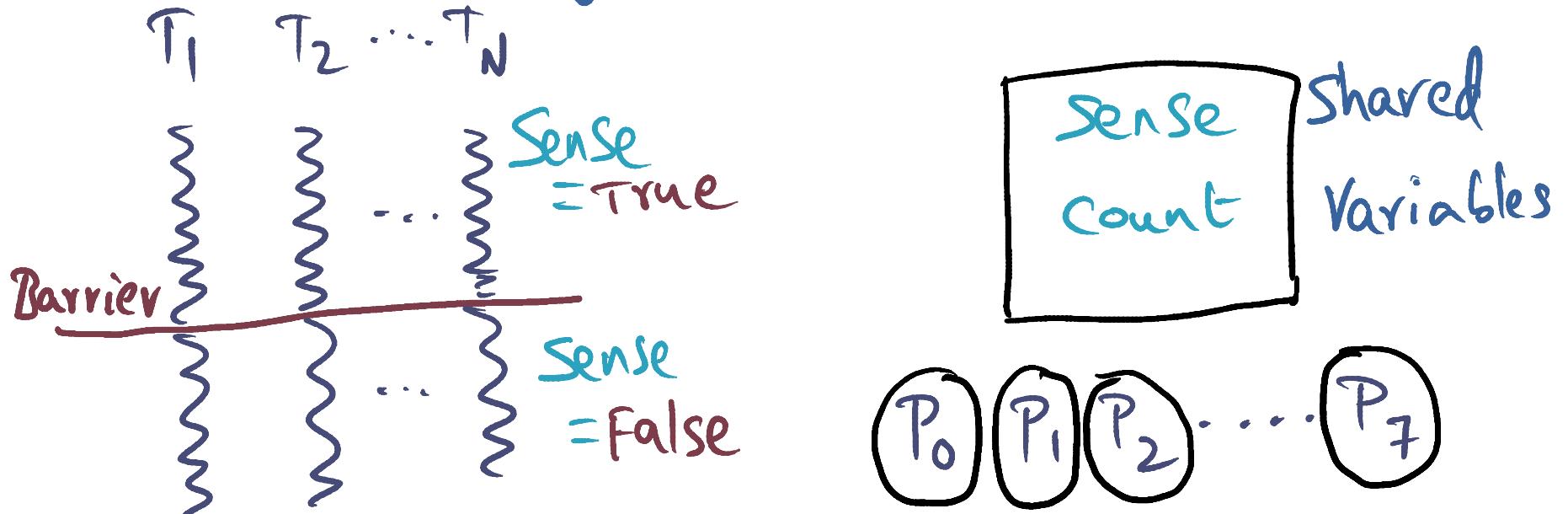
→ * Barrier algorithms

* Next week

Please
watch
videos $\Rightarrow \{$ * Communication in SM systems
* Sched in SM systems

Centralized Barrier

Sense reversing barrier



All except last :

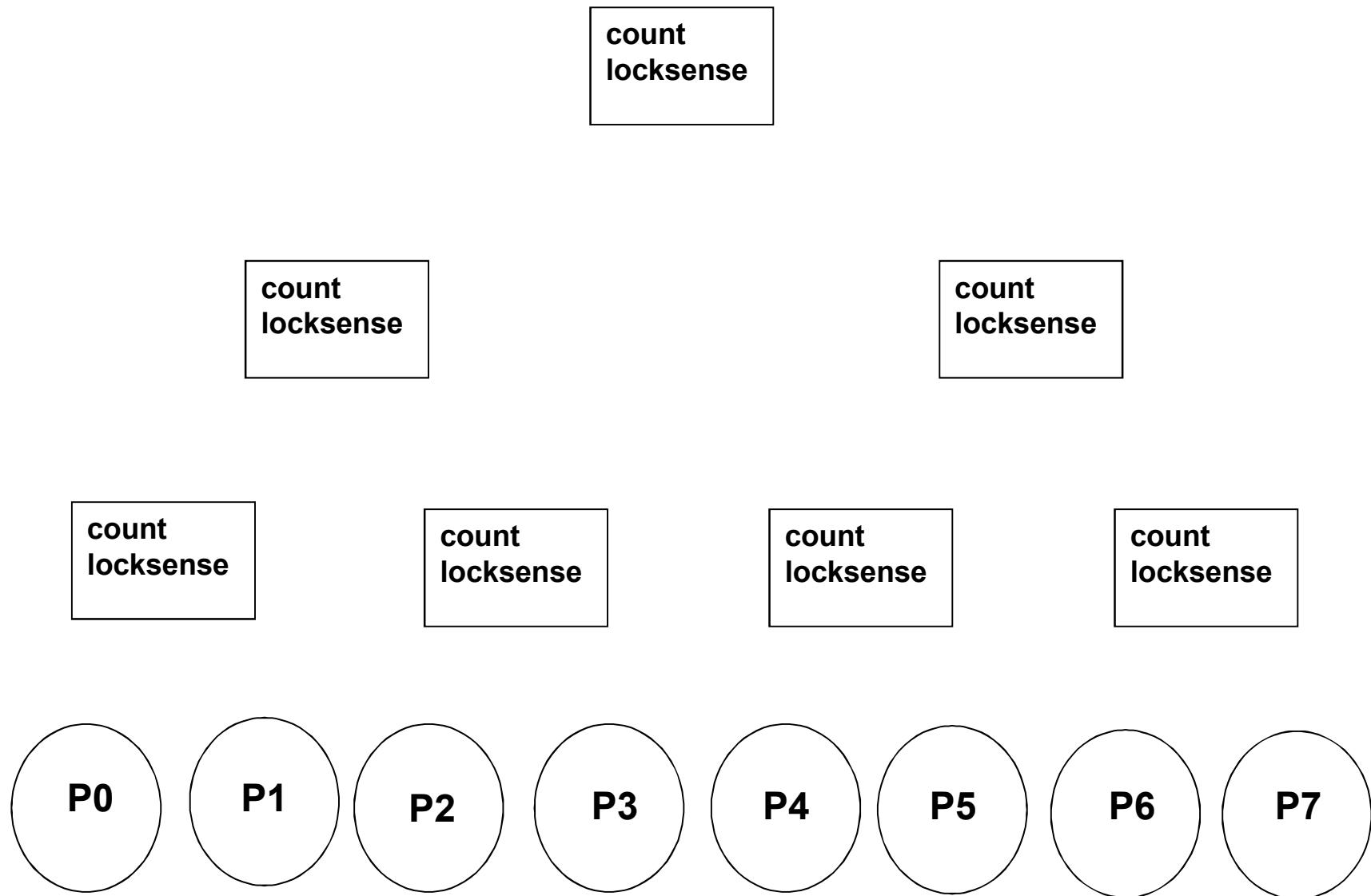
- Decrement count

- Spin on Sense reversal

Last:

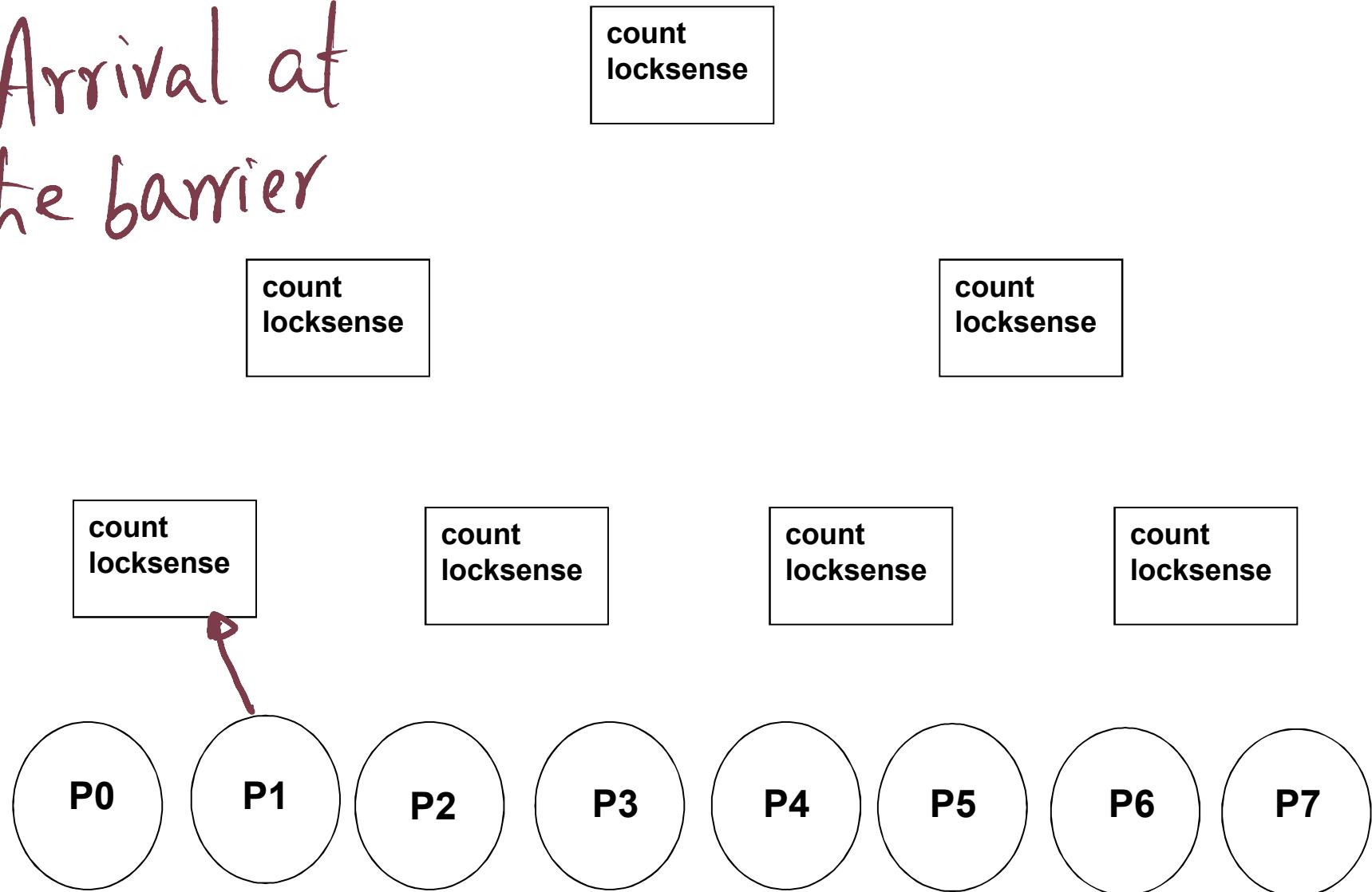
- Reset count to N ; reverse sense

Tree Barrier



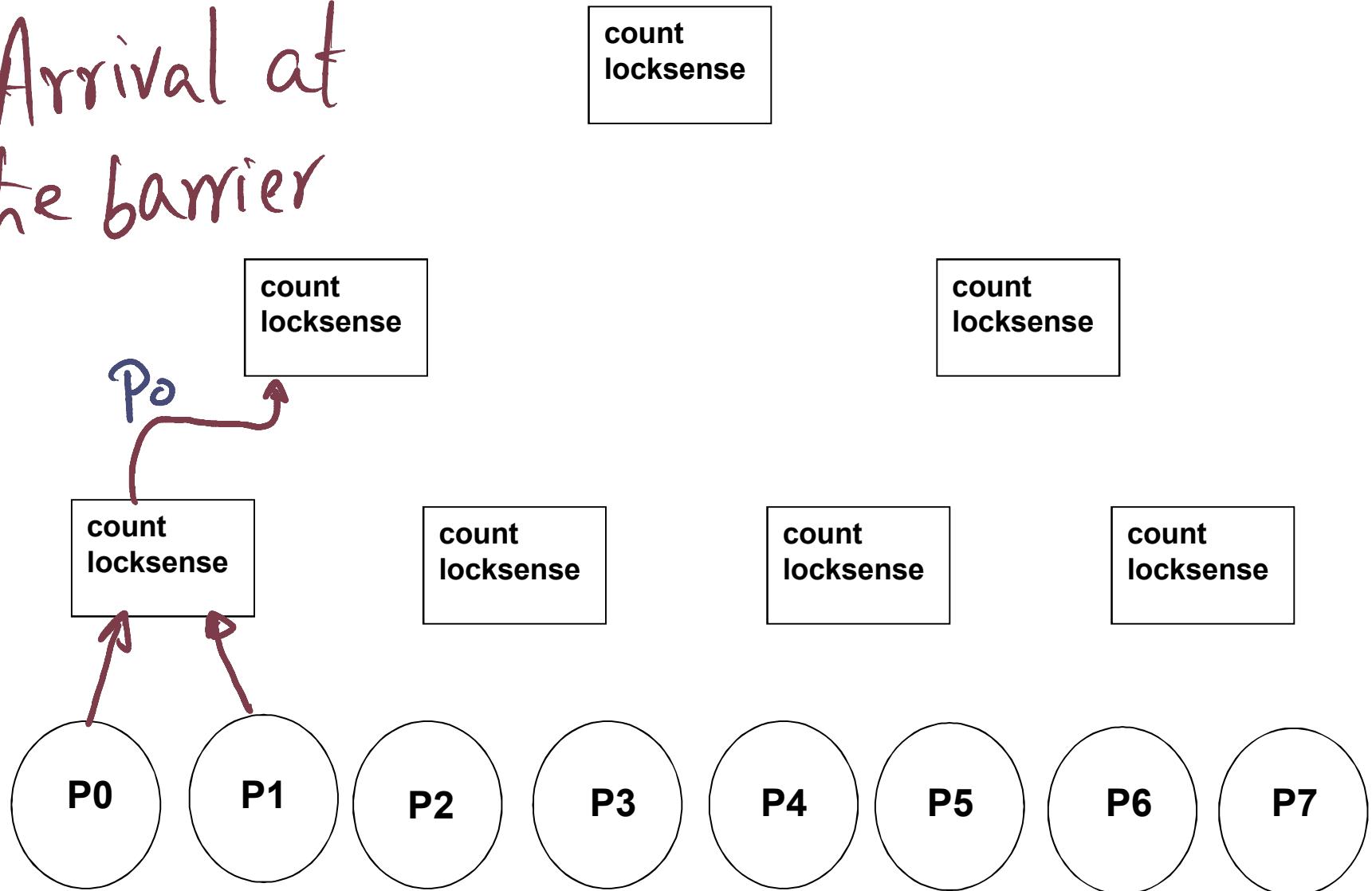
Tree Barrier

Arrival at
the barrier



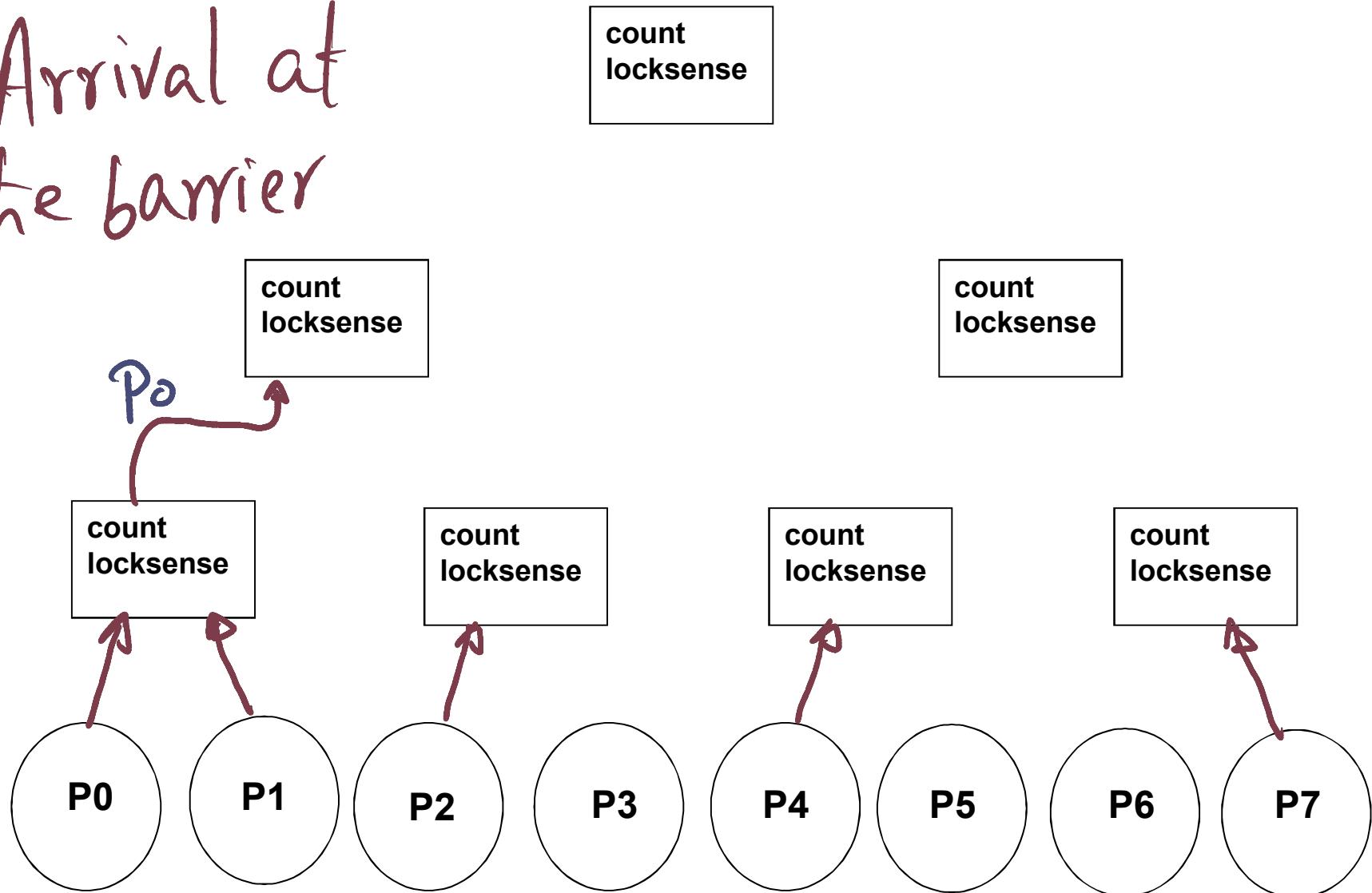
Tree Barrier

Arrival at
the barrier



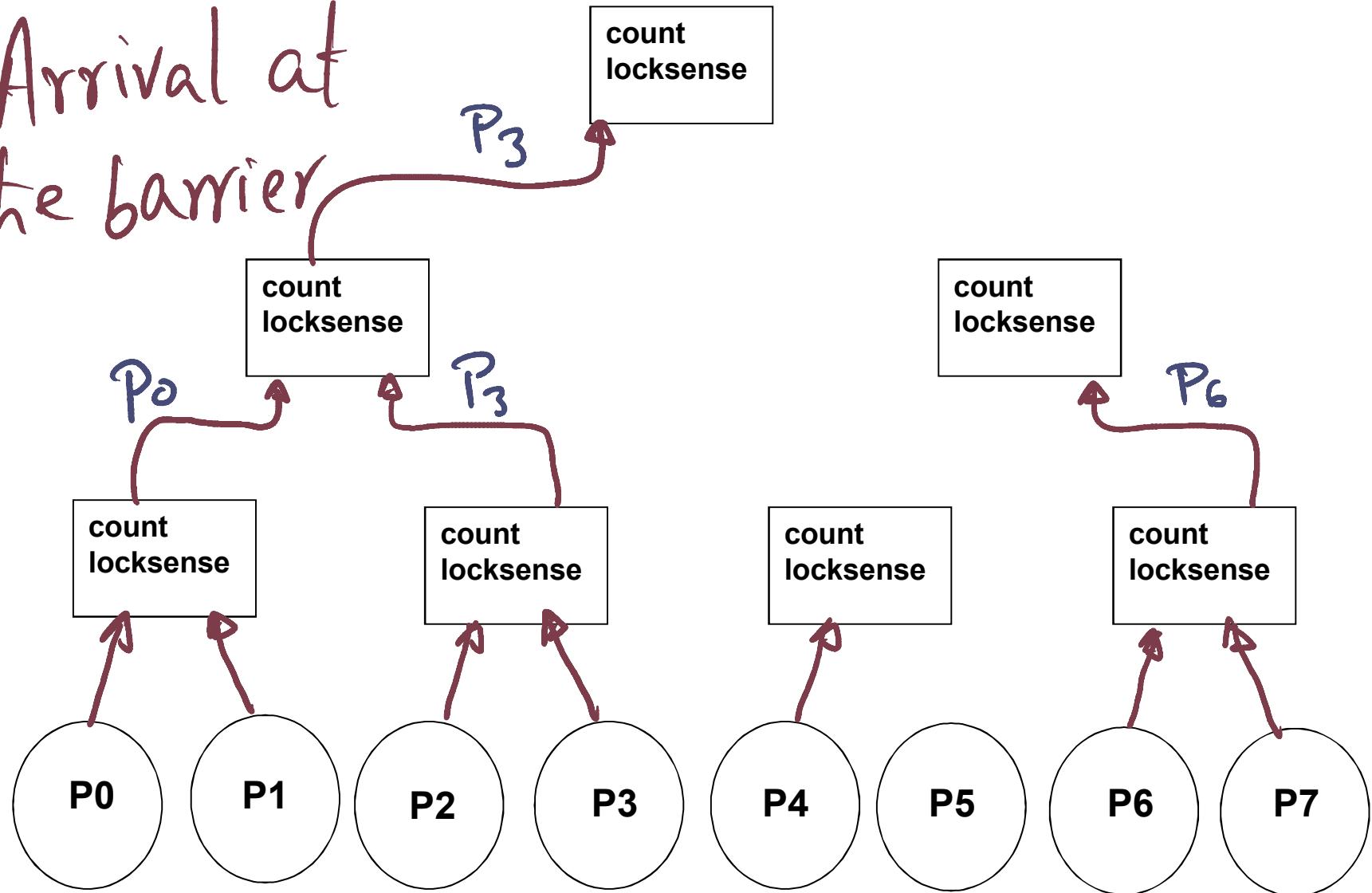
Tree Barrier

Arrival at
the barrier



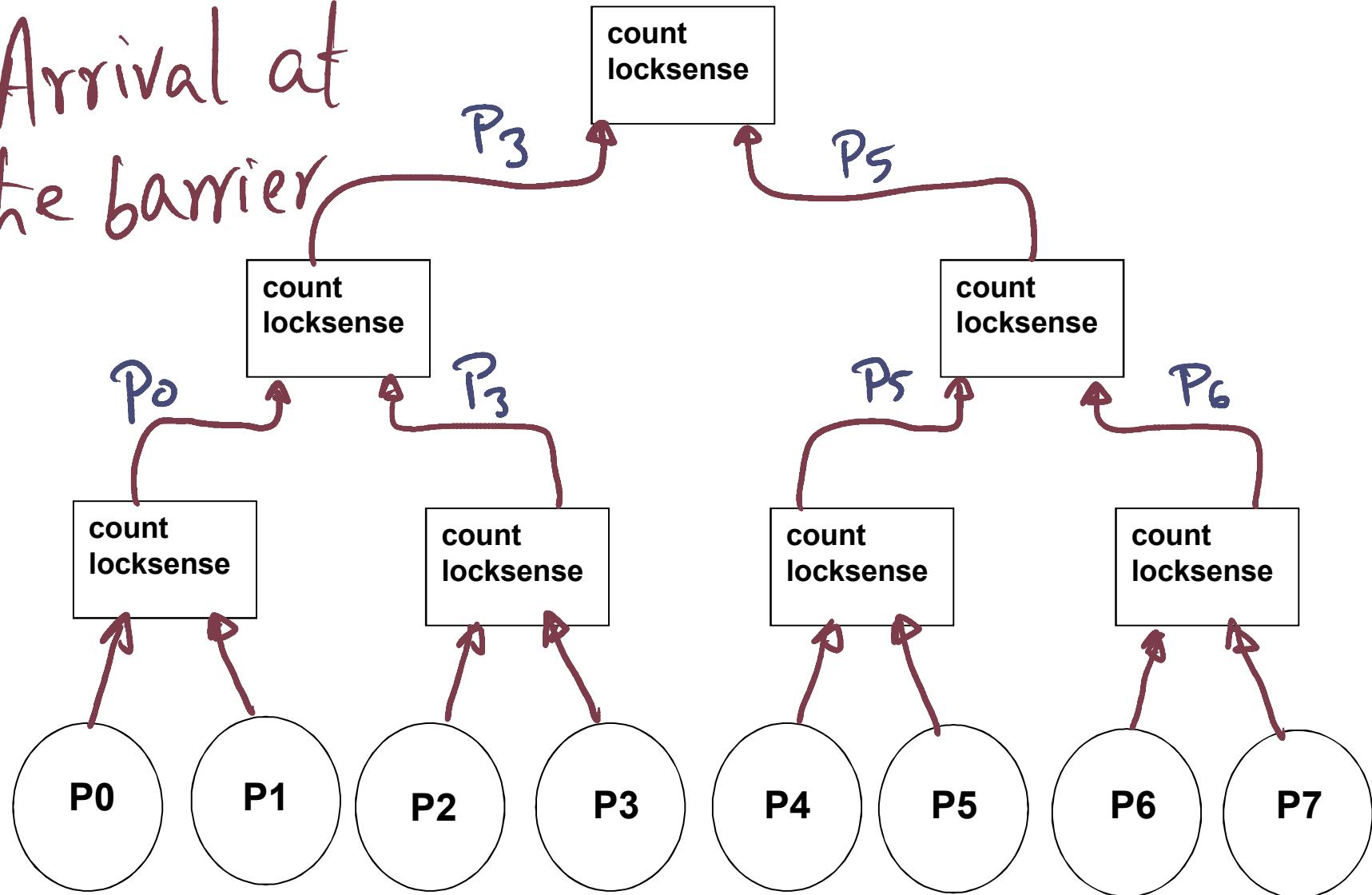
Tree Barrier

Arrival at
the barrier



Tree Barrier

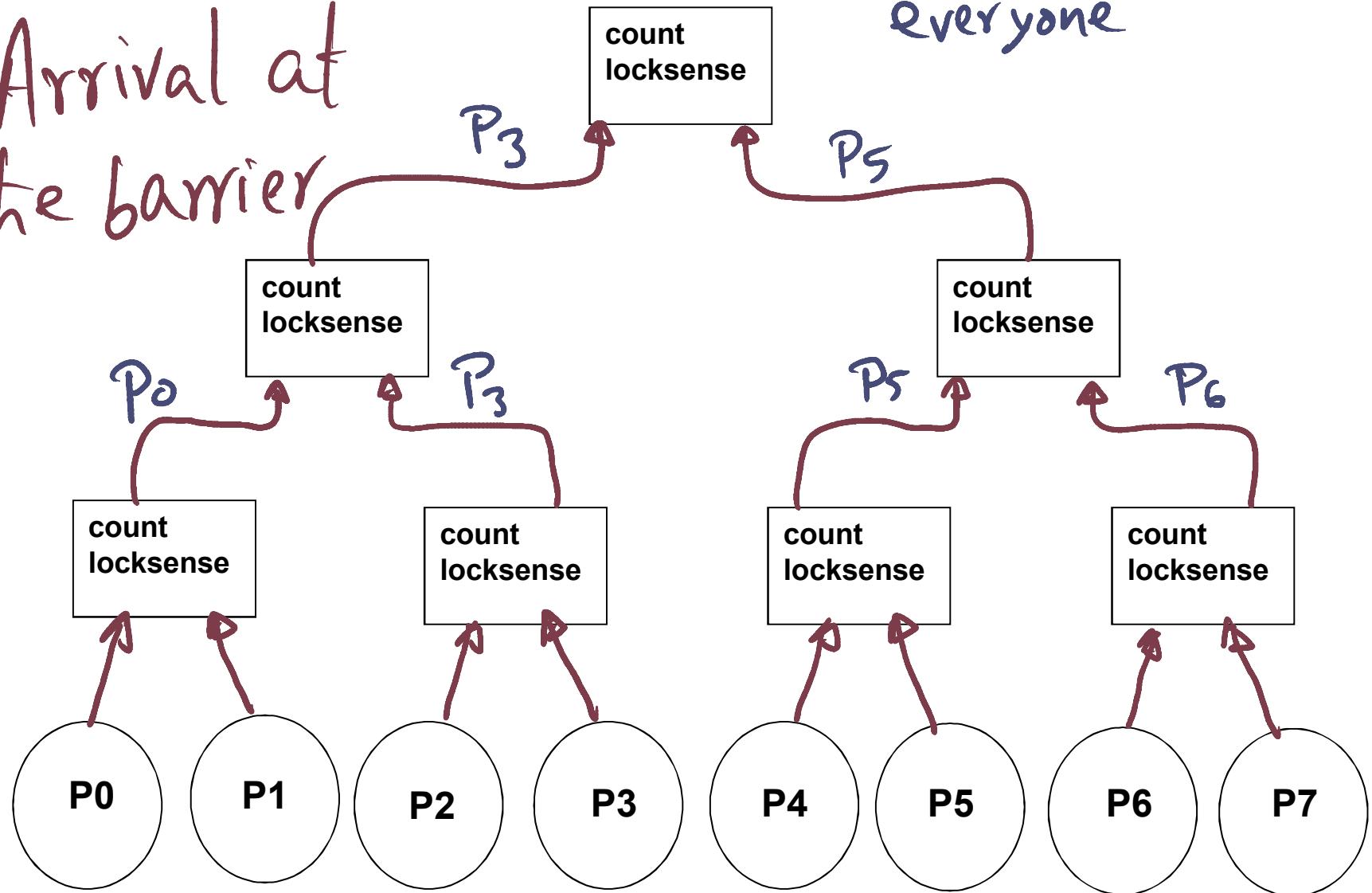
Arrival at
the barrier



Tree Barrier

Arrival at
the barrier

Time to Wakeup
everyone

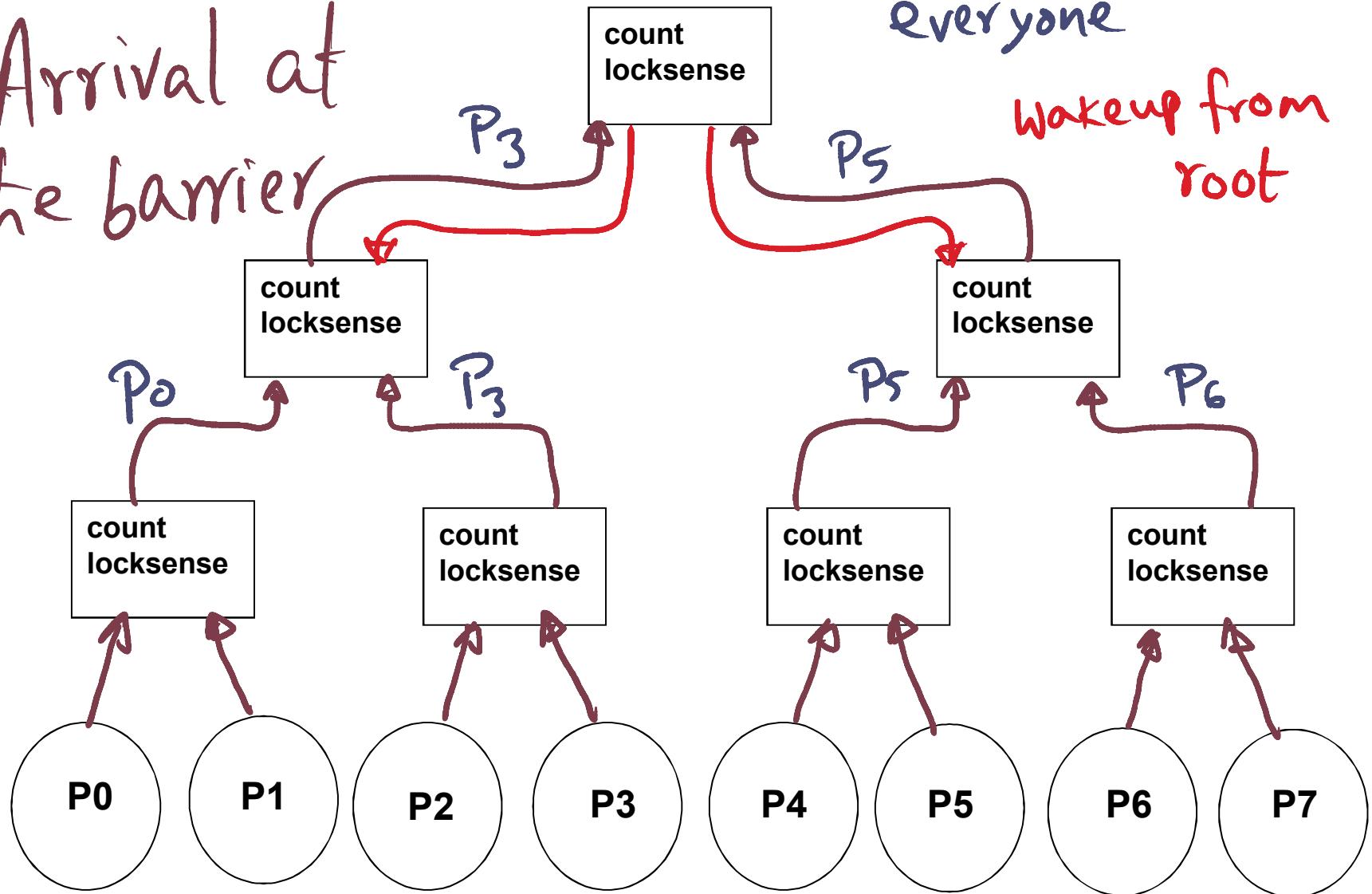


Tree Barrier

Arrival at
the barrier

Time to Wakeup
everyone

wakeup from
root

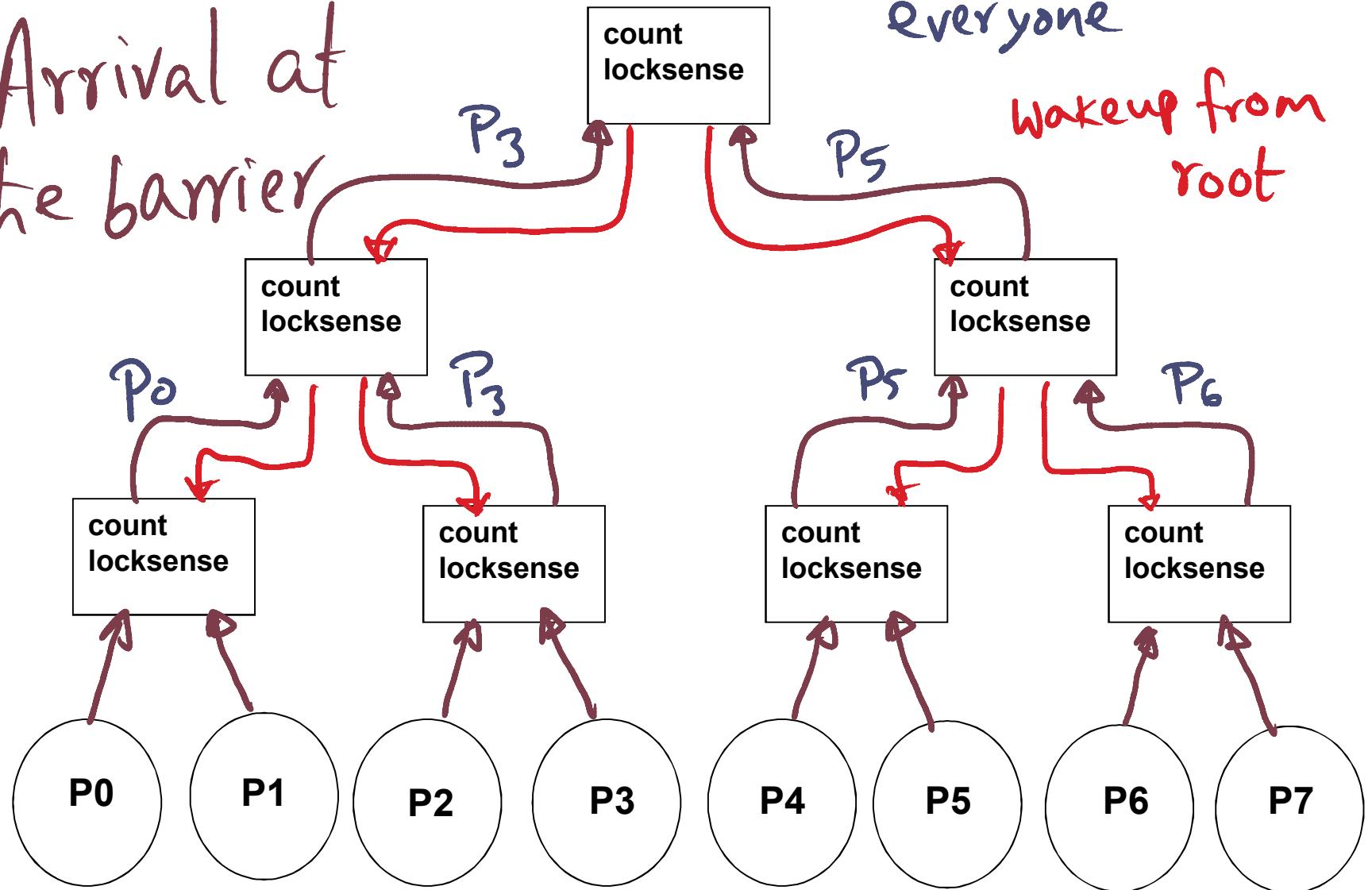


Tree Barrier

Arrival at
the barrier

Time to Wakeup
everyone

wakeup from
root

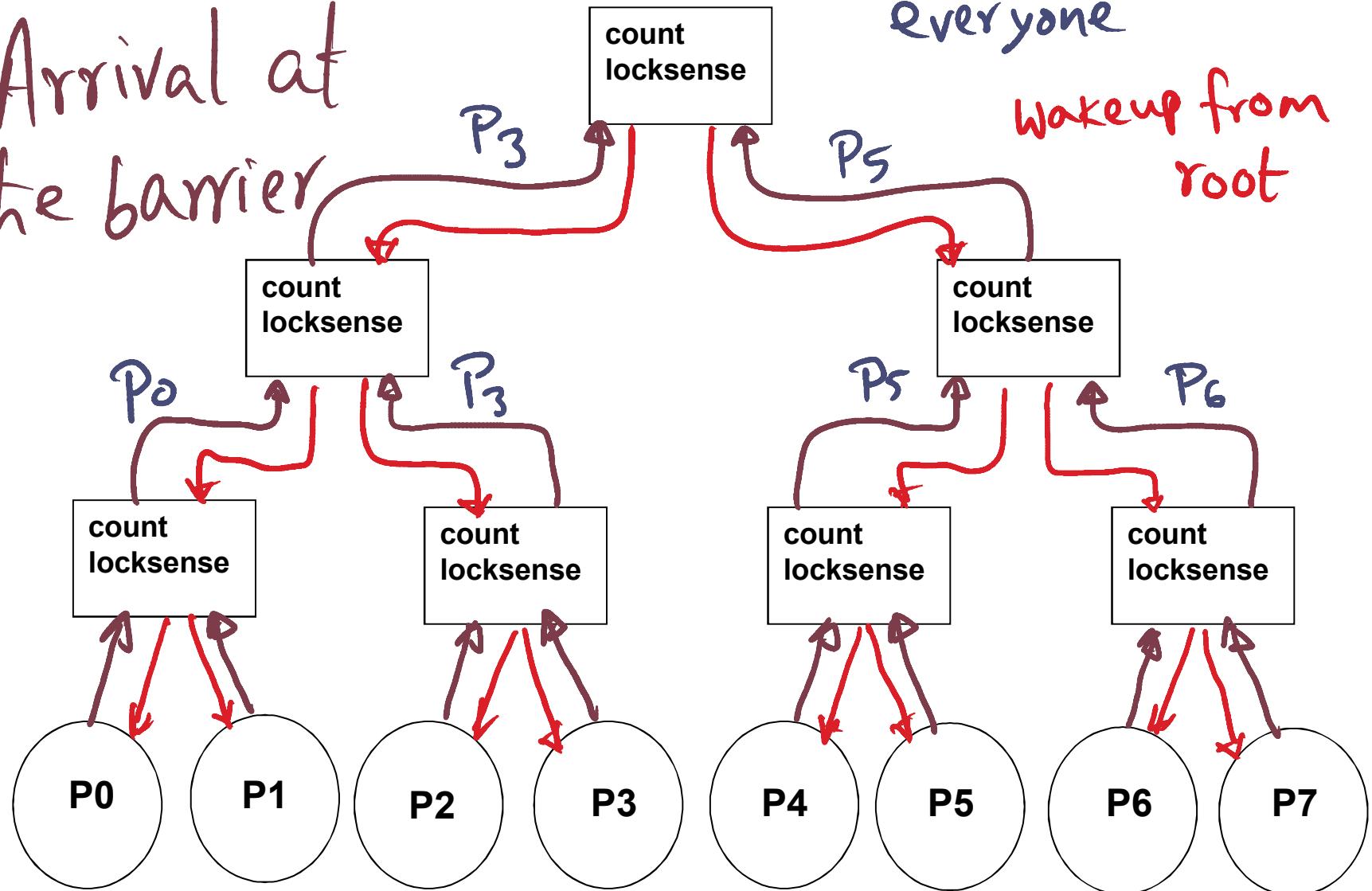


Tree Barrier

Arrival at
the barrier

Time to Wakeup
everyone

wakeup from
root

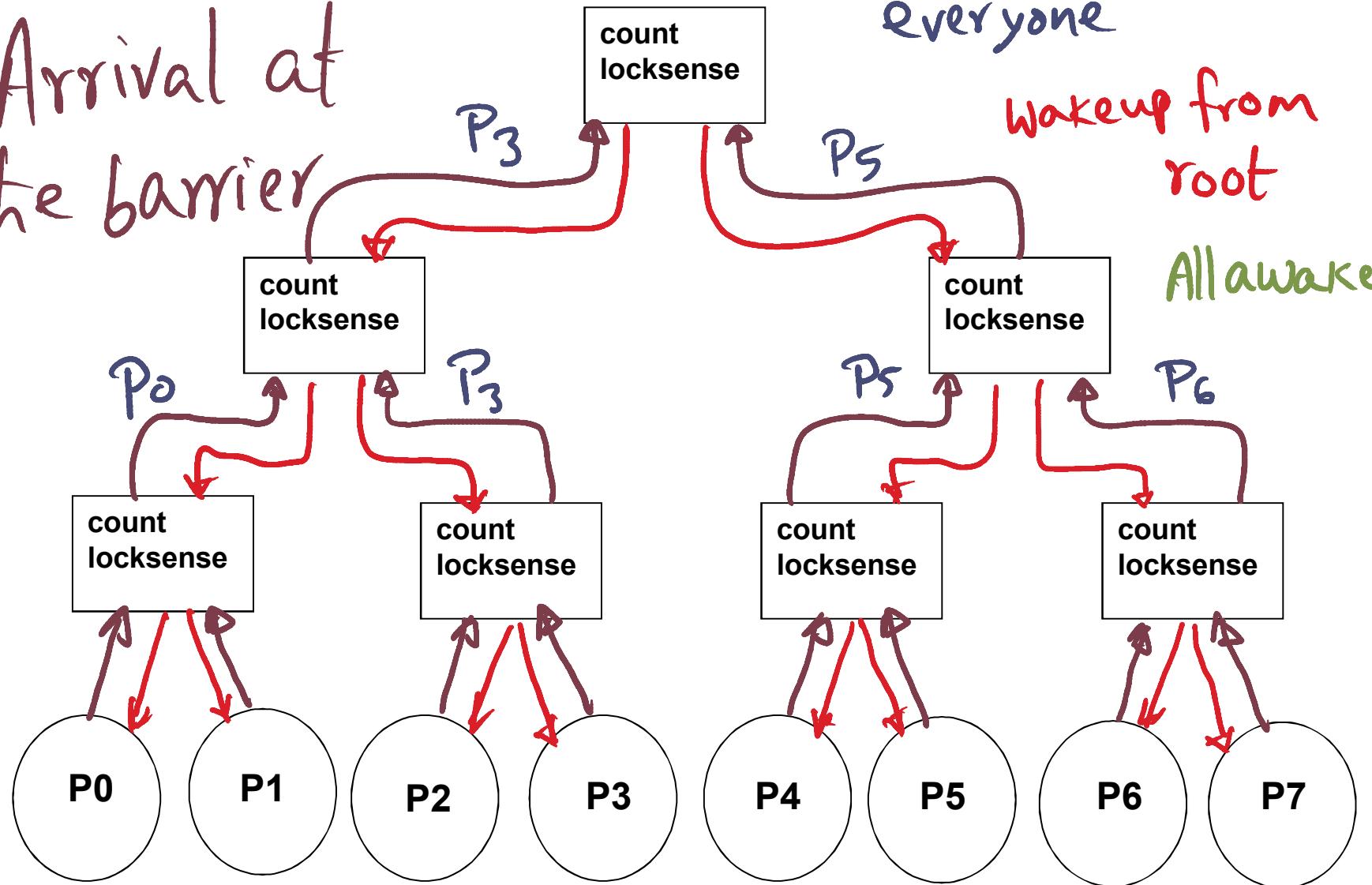


Tree Barrier

Arrival at
the barrier

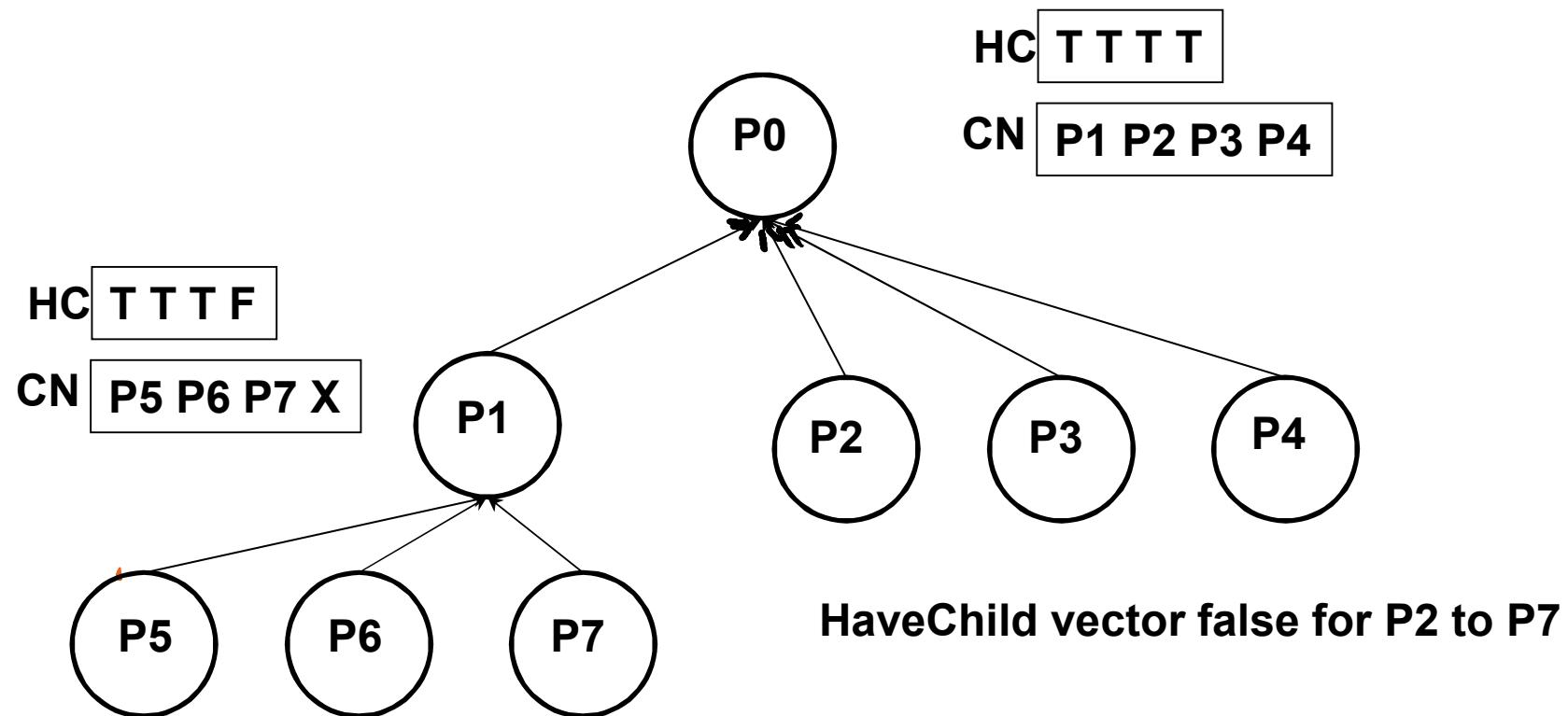
Time to Wakeup
everyone

wakeup from
root
All awake!



MCS Tree Barrier (4-ary arrival)

$N=8$

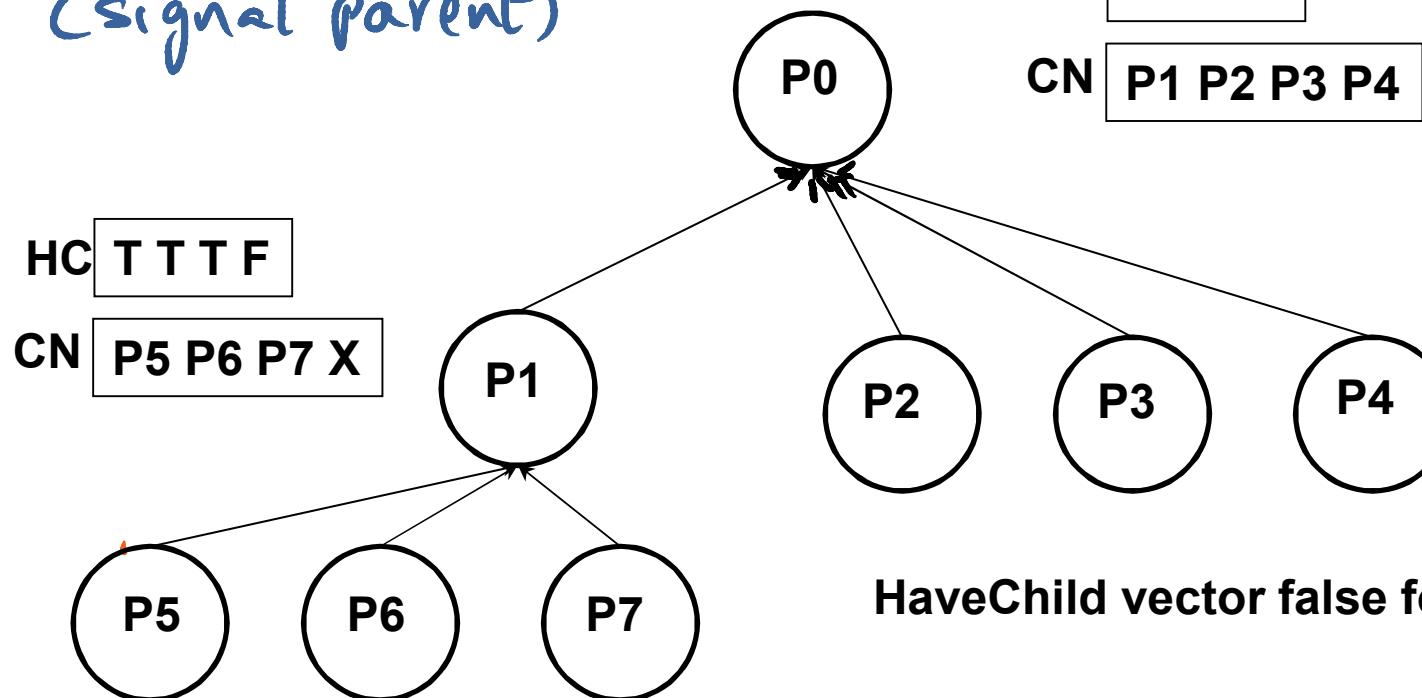


MCS Tree Barrier (4-ary arrival)

$N=8$

HC - Have children

CN - child not ready
(signal parent)

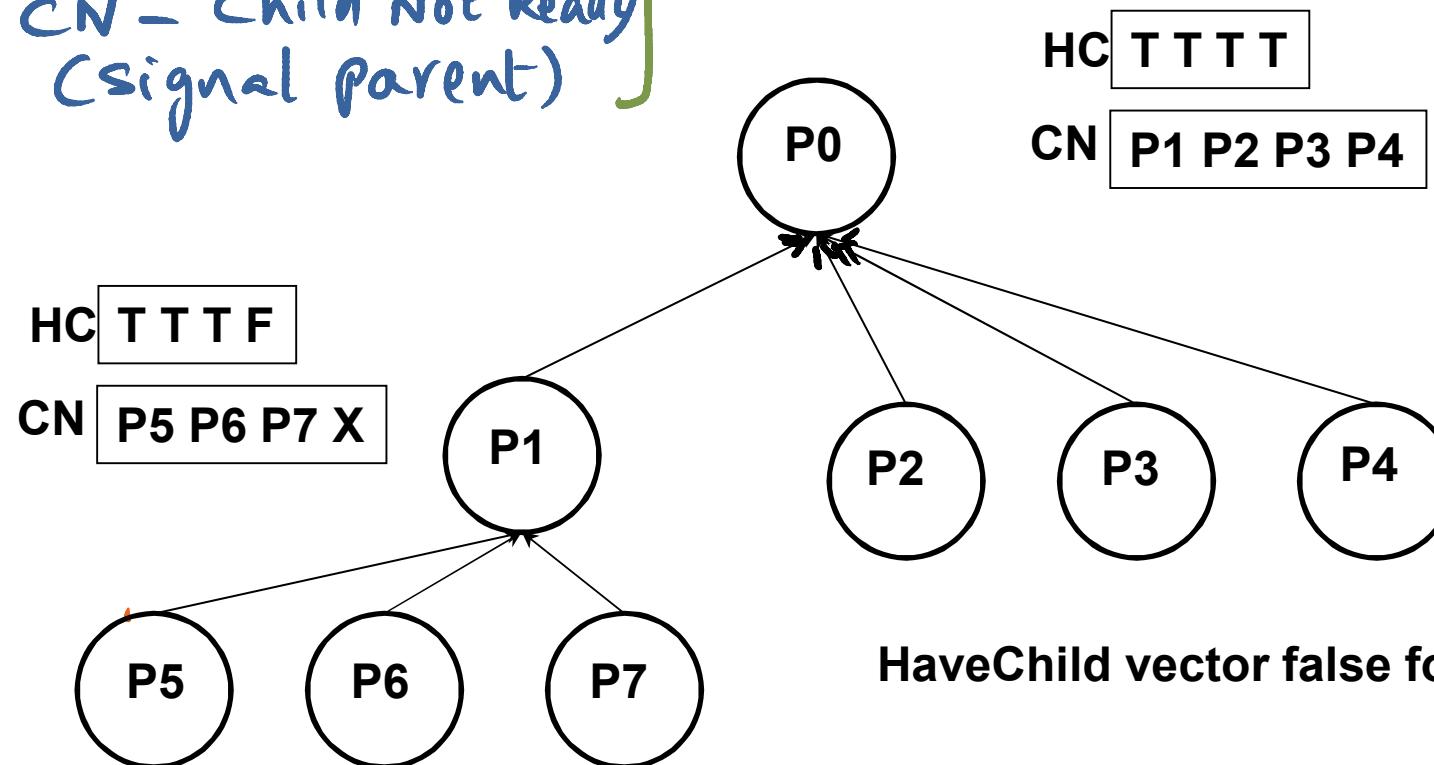


MCS Tree Barrier (4-ary arrival)

$N=8$

HC - Have children
 CN - child not Ready
(signal parent)}

Associated with every Parent



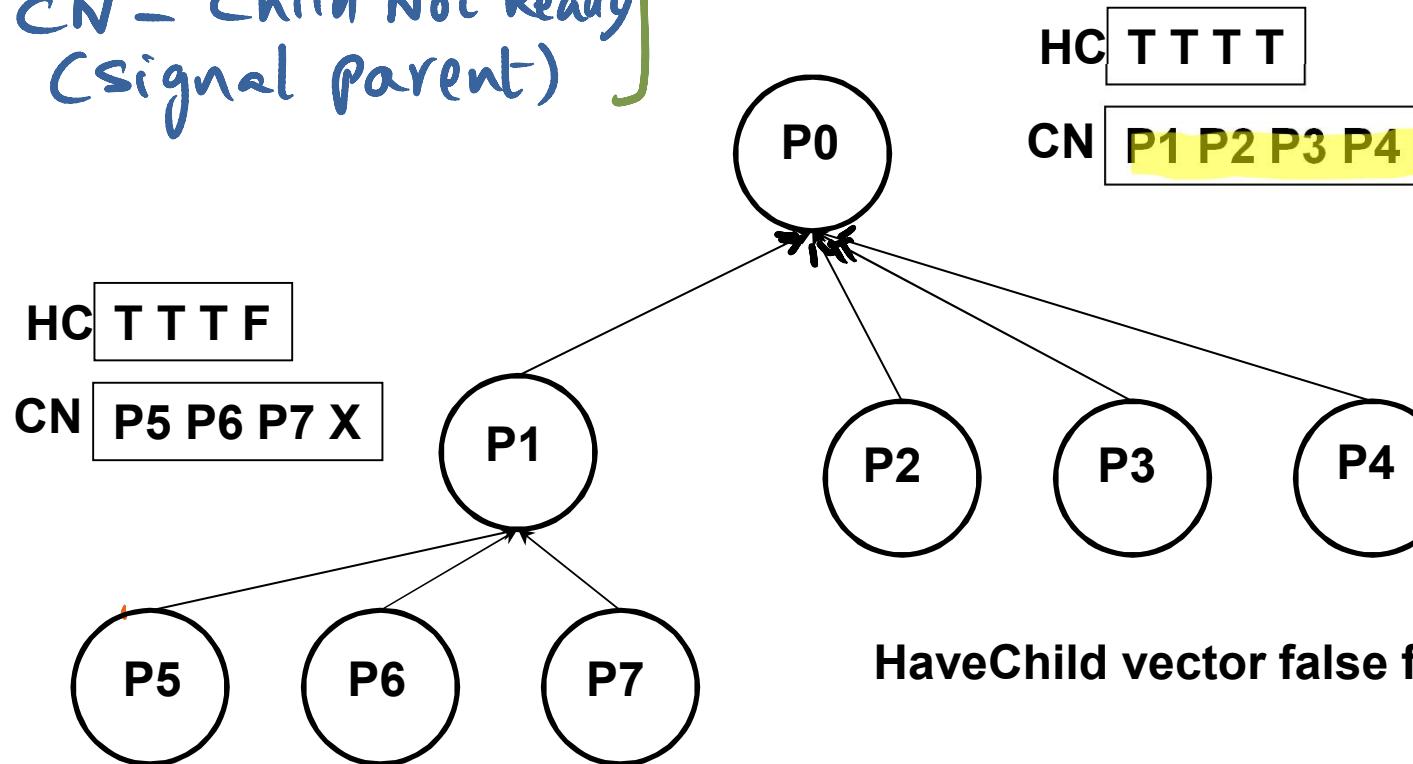
MCS Tree Barrier (4-ary arrival)

$N=8$

HC - Have children

CN - child not ready
(signal parent)

Associated with every Parent



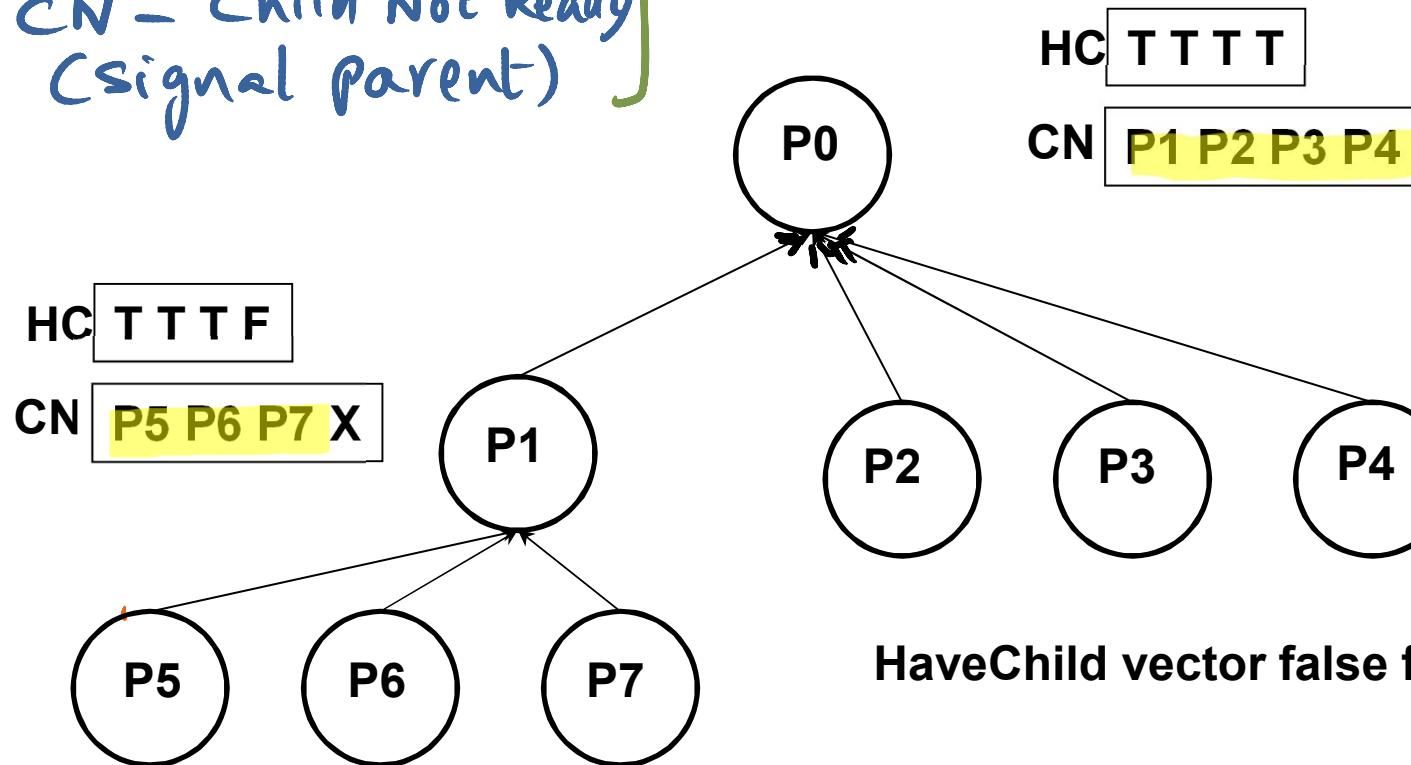
MCS Tree Barrier (4-ary arrival)

$N=8$

HC - Have children

CN - child not ready
(signal parent)

Associated with every Parent

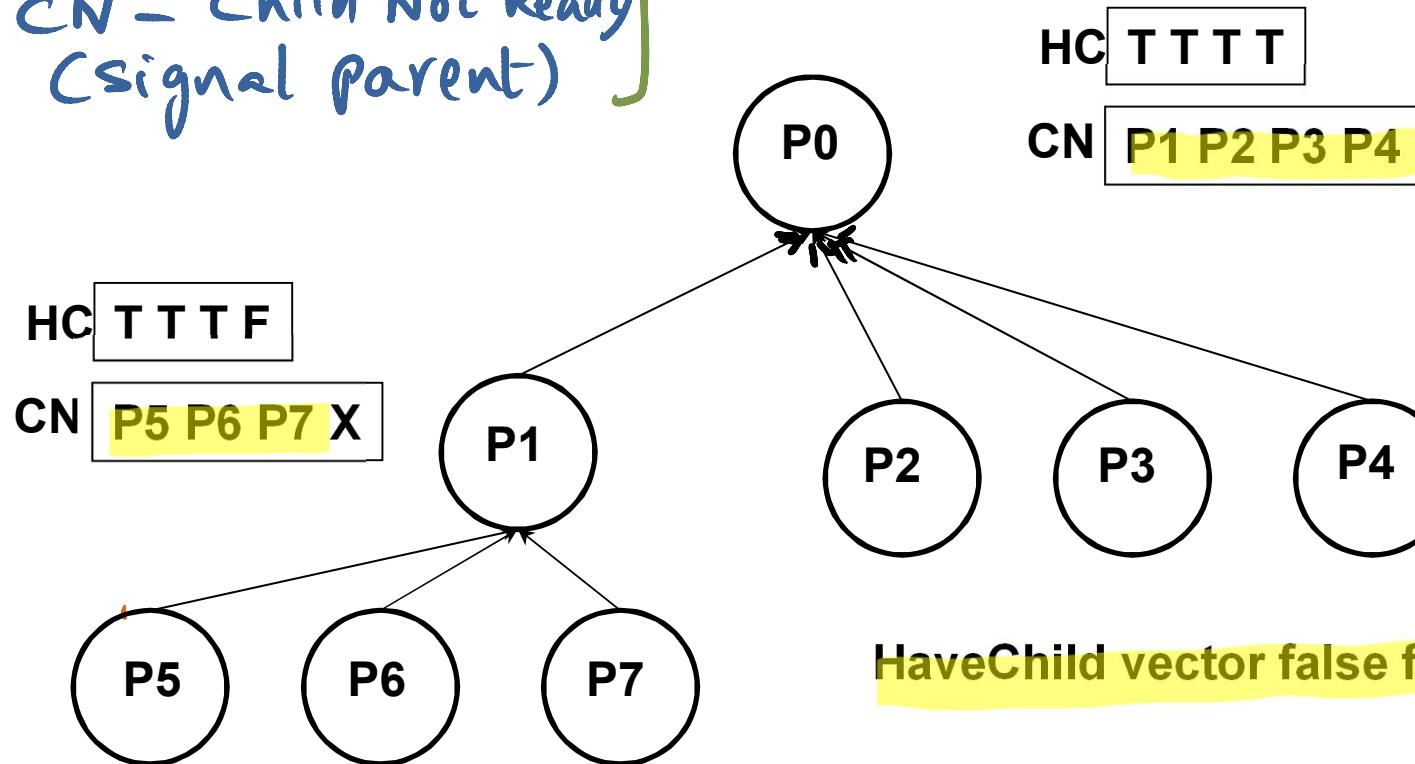


MCS Tree Barrier (4-ary arrival)

$N=8$

HC - Have children
CN - child not Ready
(signal parent)

Associated with every Parent

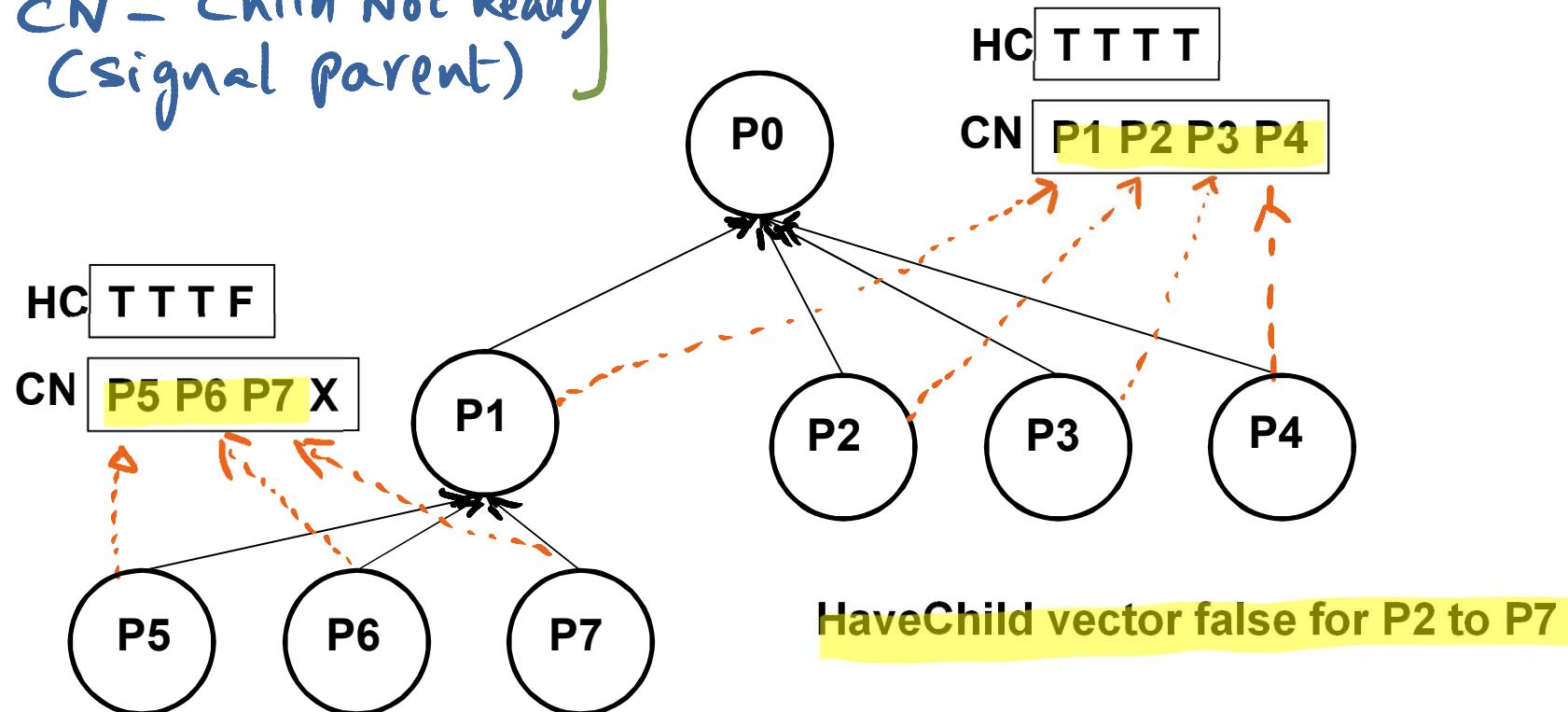


MCS Tree Barrier (4-ary arrival)

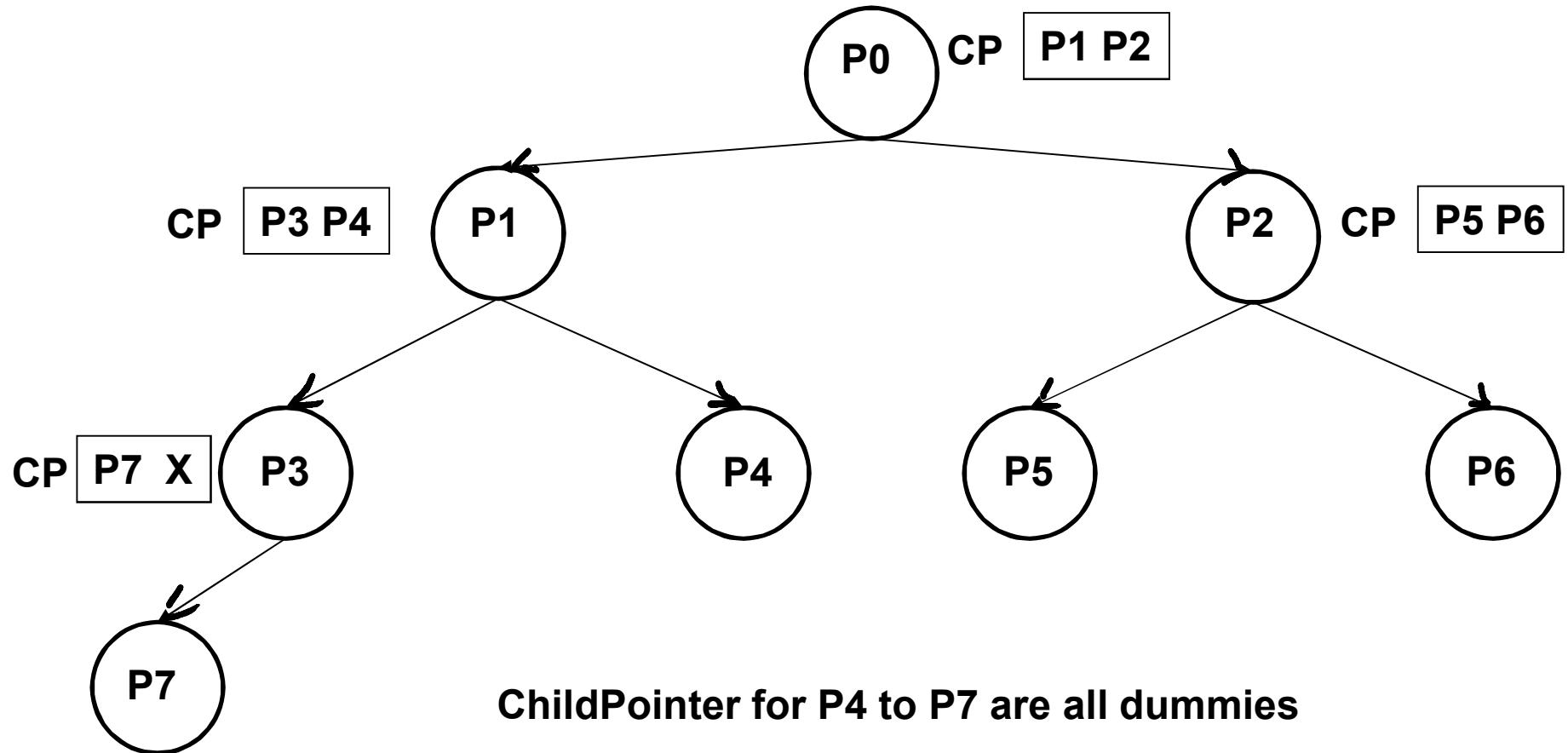
$N=8$

HC - Have children
 CN - child not Ready
(signal parent)}

Associated with every Parent

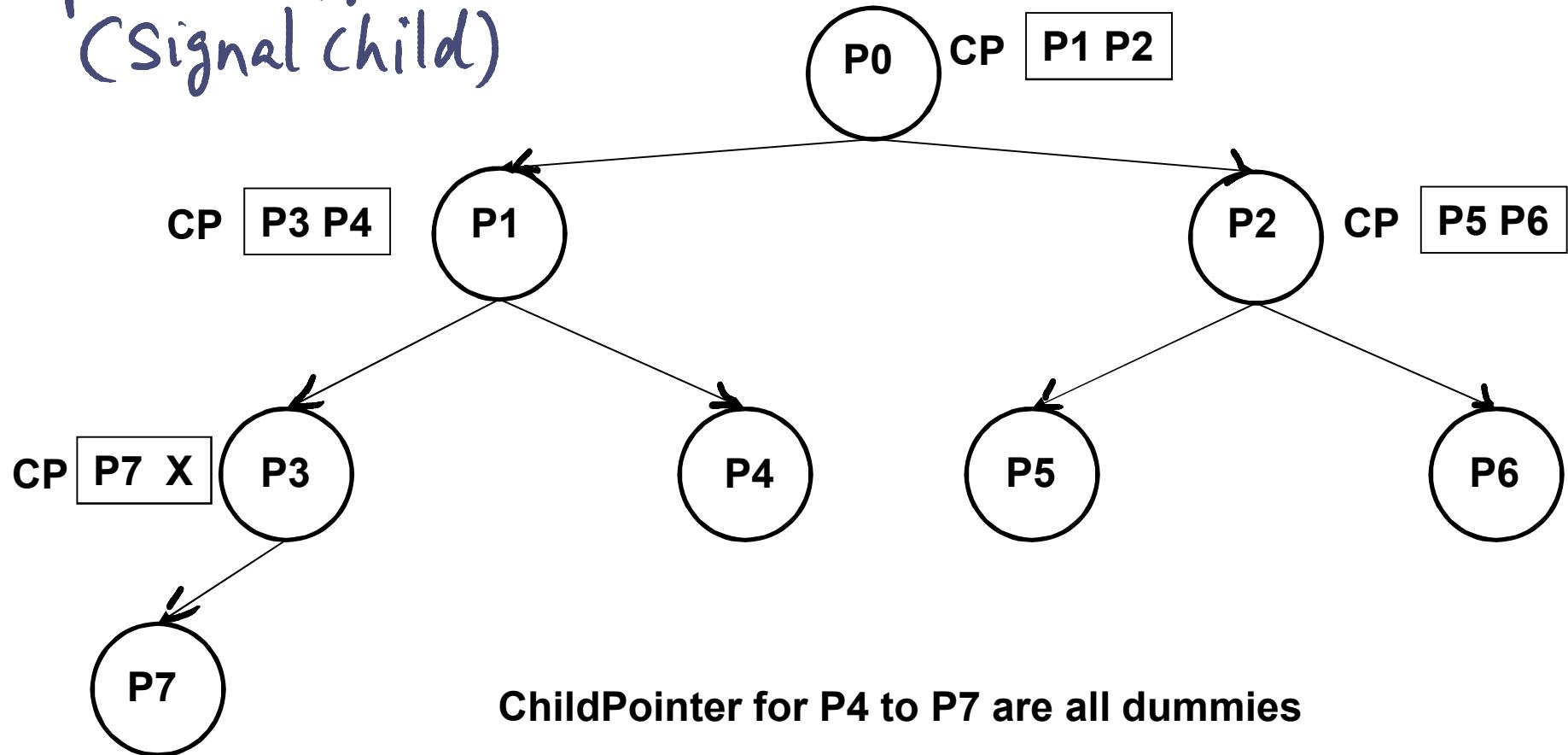


MCS Tree Barrier (binary wakeup)



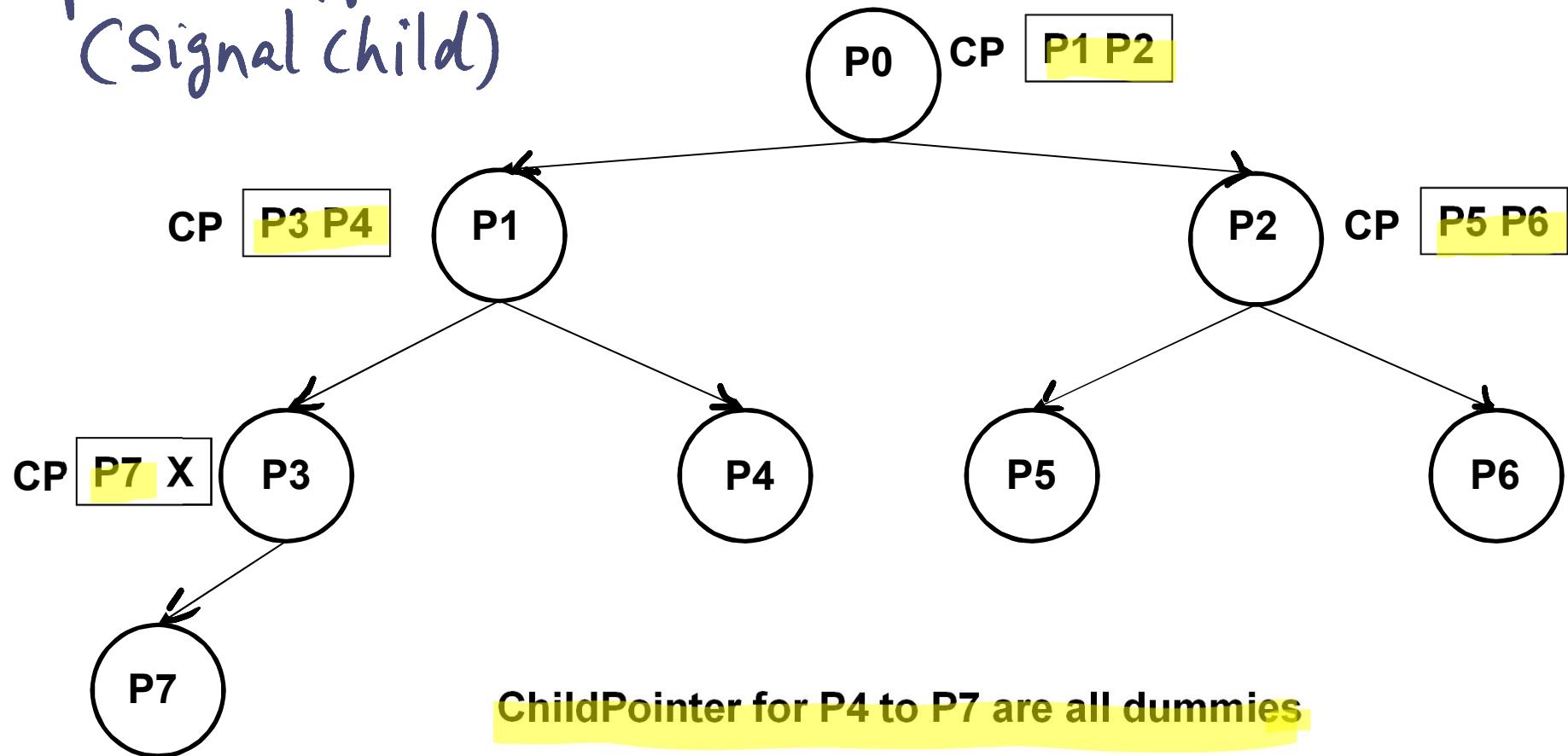
MCS Tree Barrier (binary wakeup)

CP - child pointer
(Signal child)



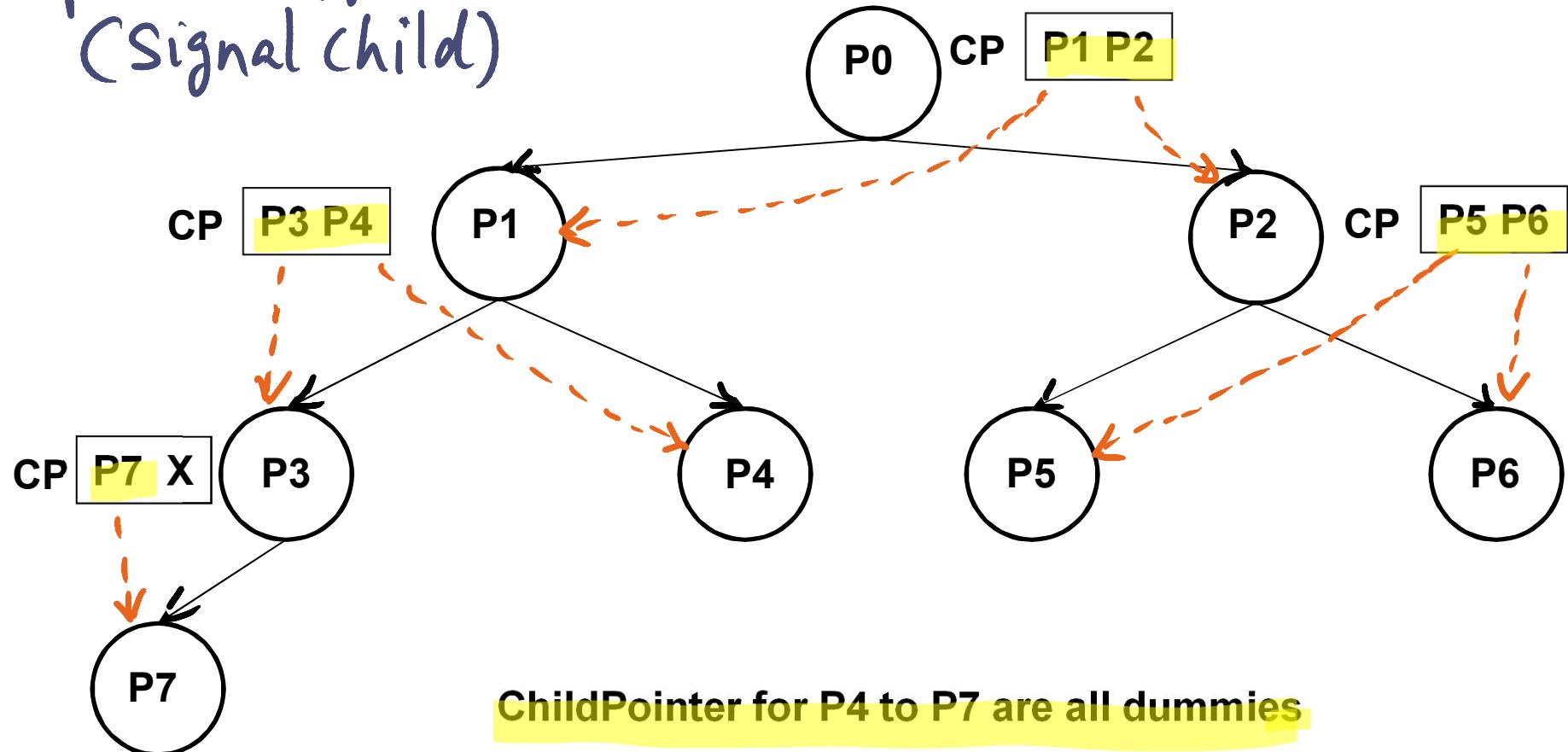
MCS Tree Barrier (binary wakeup)

CP - child pointer
(Signal child)



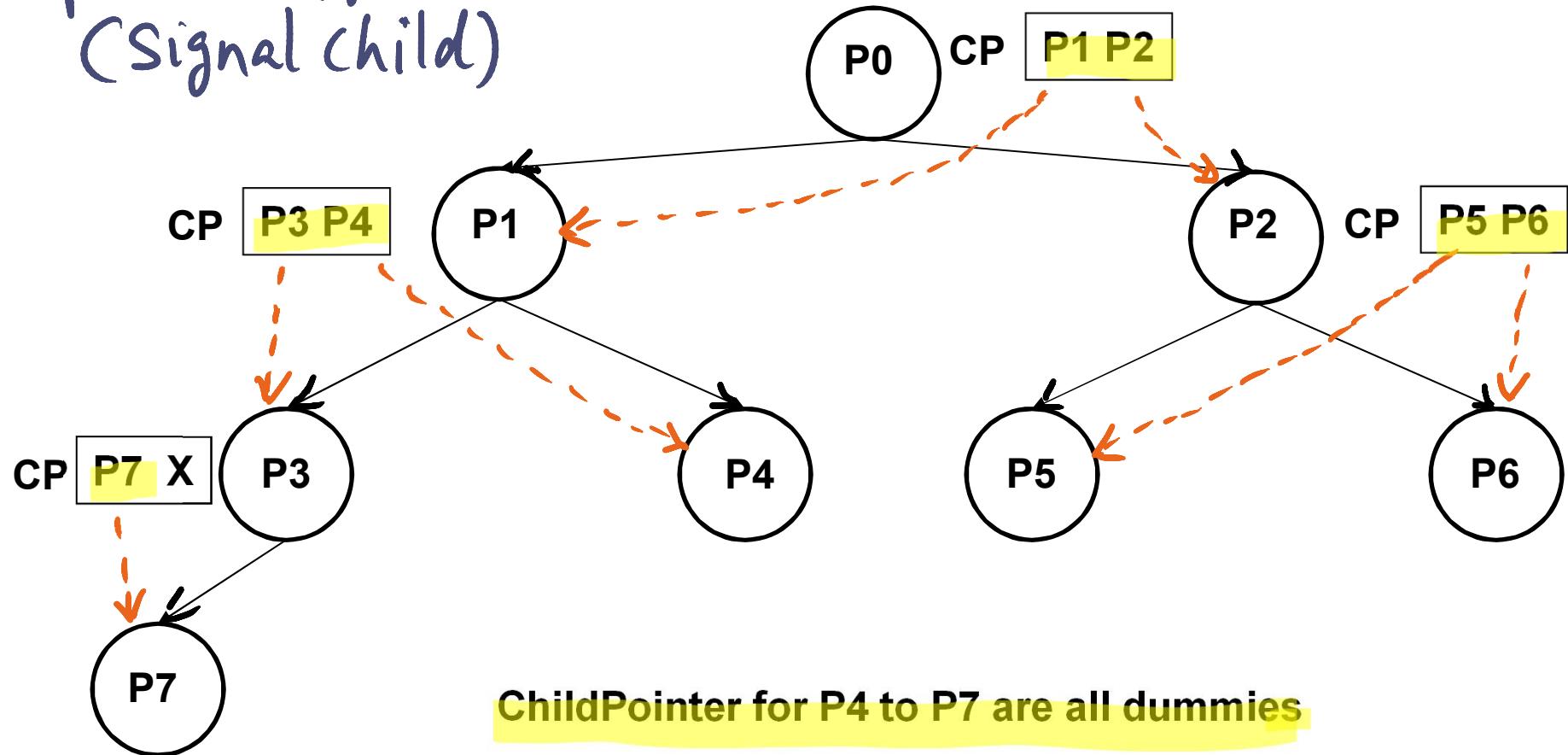
MCS Tree Barrier (binary wakeup)

CP - child pointer
(Signal child)



MCS Tree Barrier (binary wakeup)

CP - child pointer
(Signal child)



All Awake!

Virtues of MCS Barrier

- * No need for Atomic RMW
- * $O(n)$ Space
- * $O(\log n)$ network trans.
on critical path
- * Works for NCC + CC NUMA

Virtues of MCS Barrier

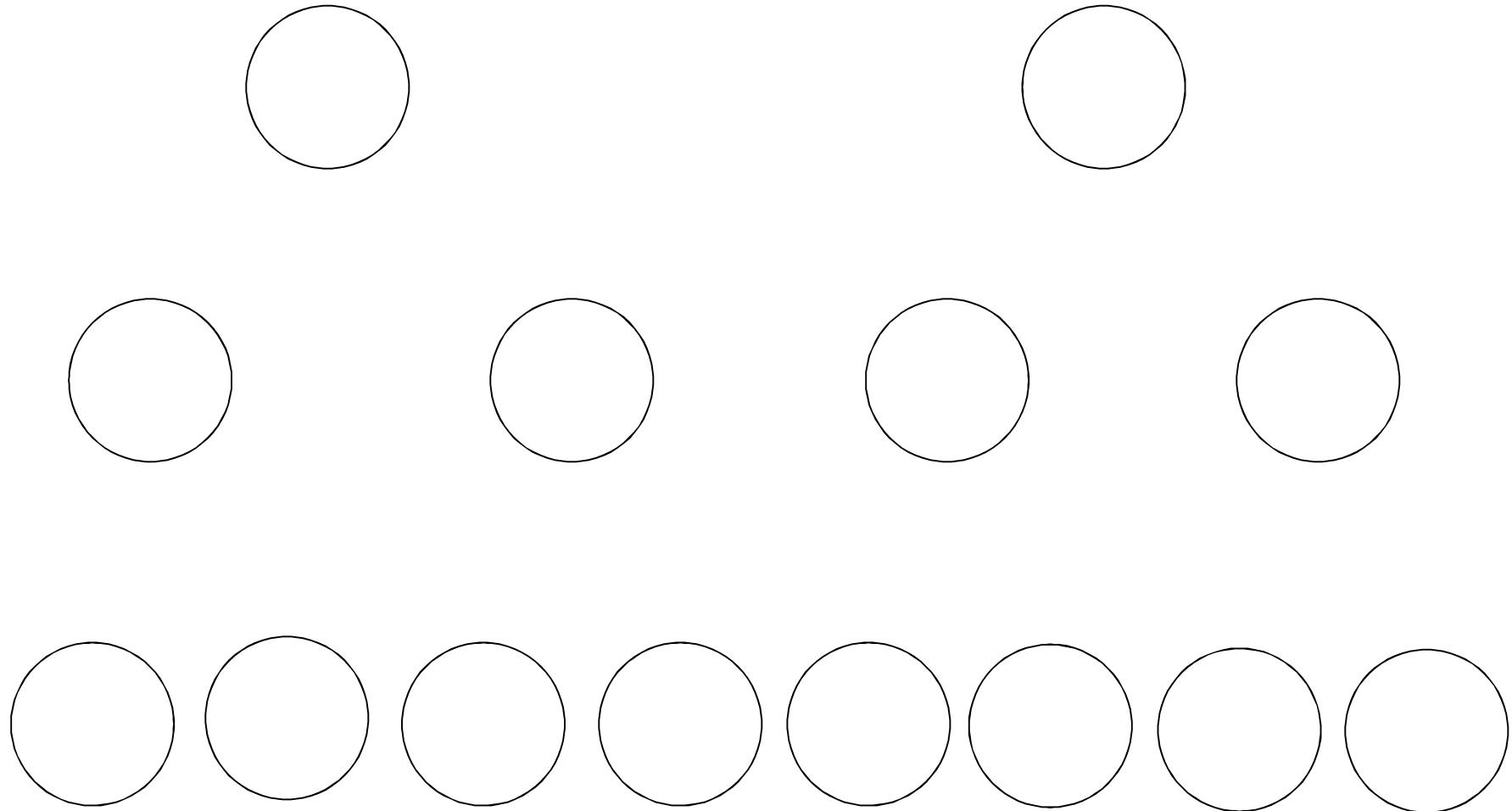
- * No need for Atomic RMW
- * $O(n)$ Space
- * $O(\log n)$ network trans.
on critical path
- * Works for NCC + CC NUMA

(Warning: Bug in Barrier Alg
Pseudo code in MCS paper
P - 19)

Tournament Barrier

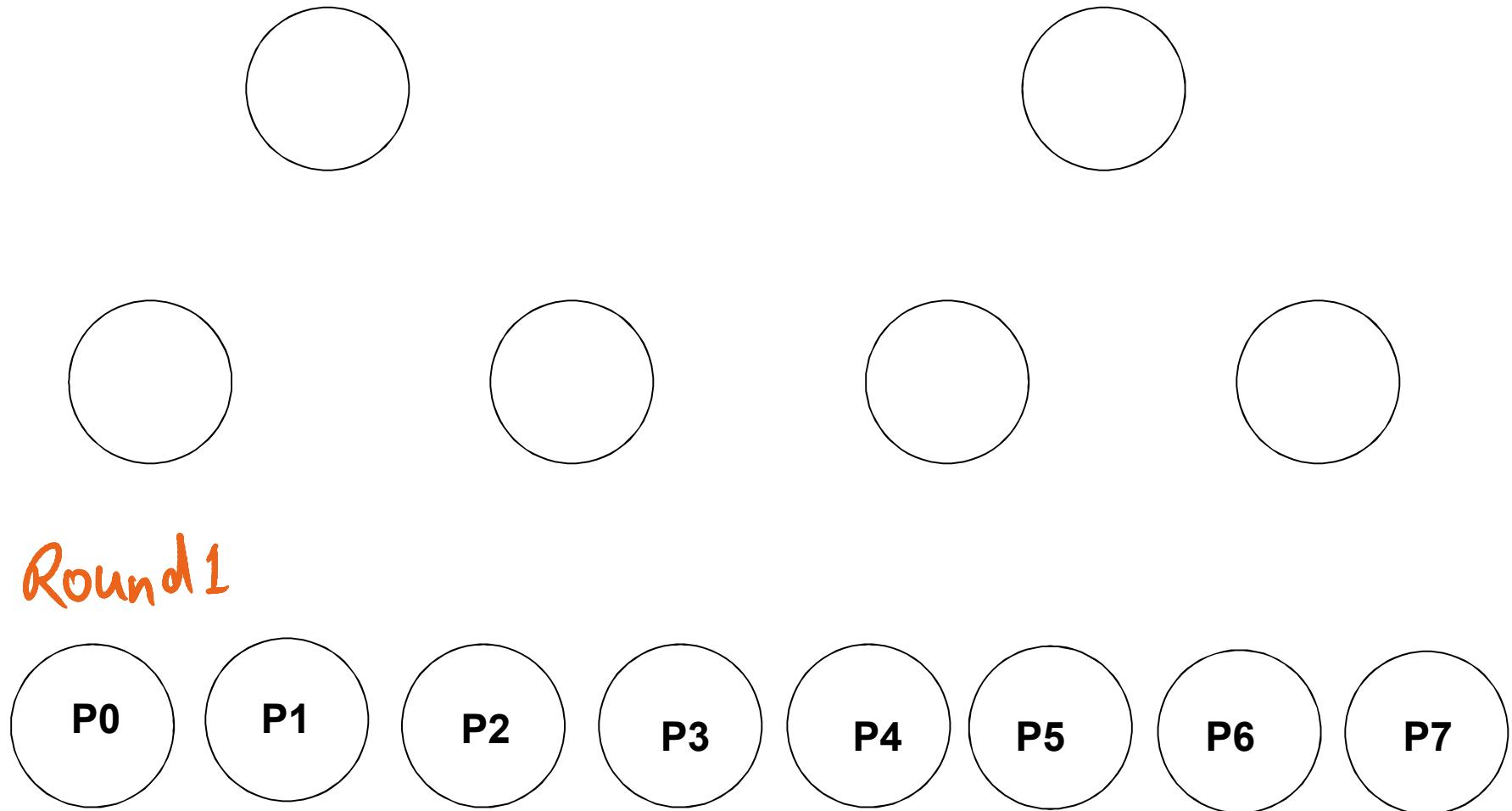
N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



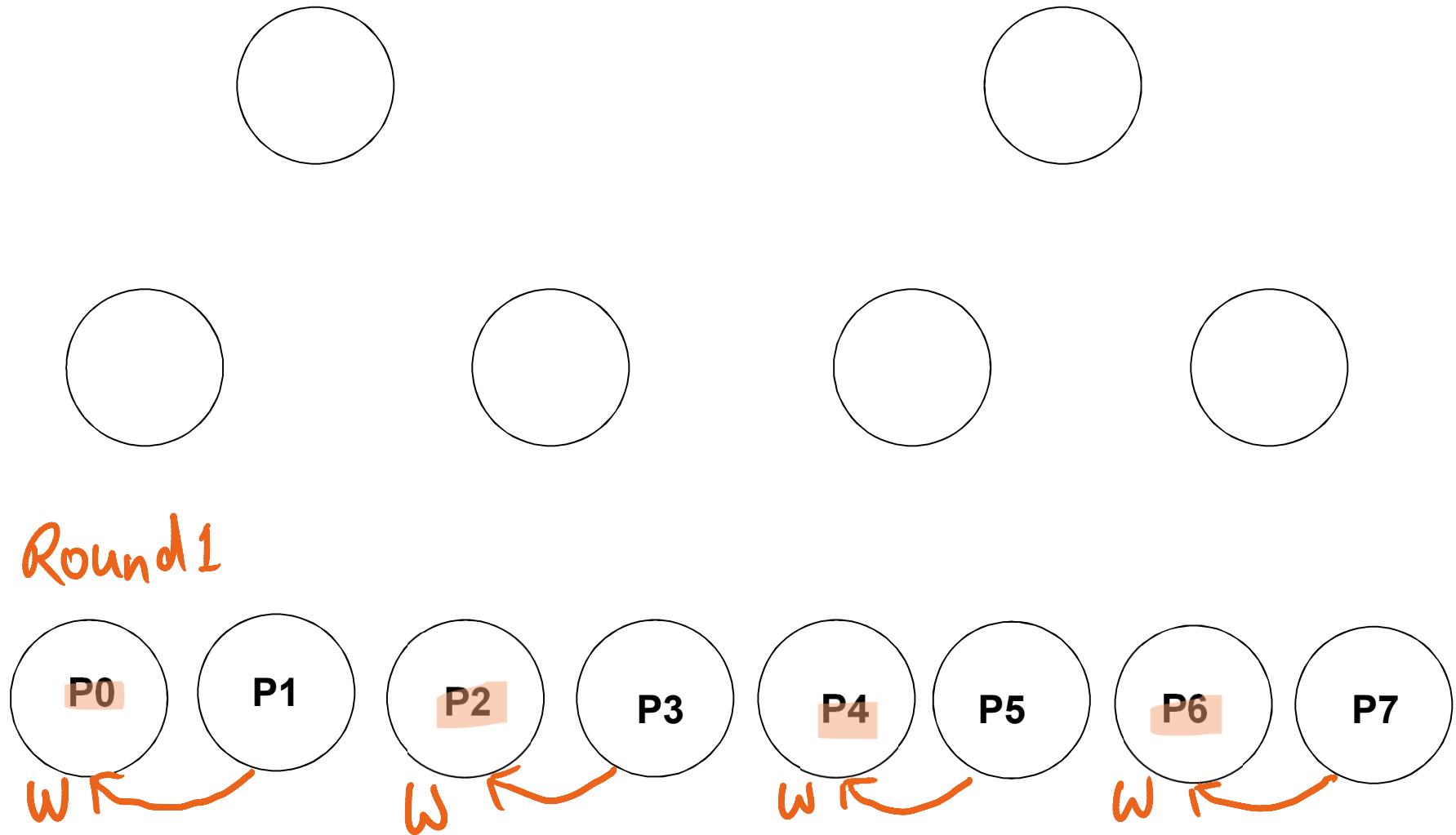
N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



N Players $\Rightarrow \log_2 N$ rounds

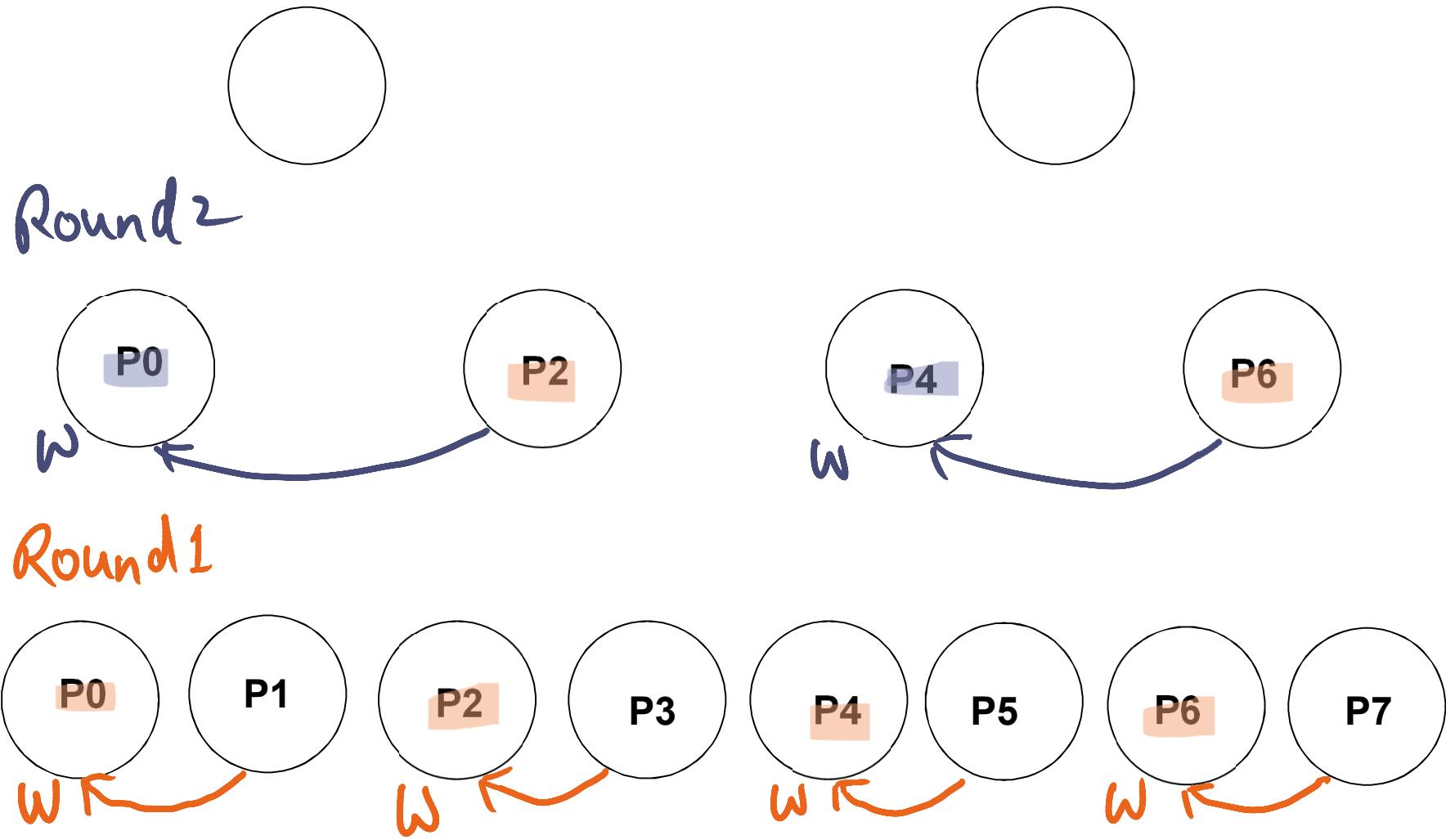
Tournament Barrier



Round 1

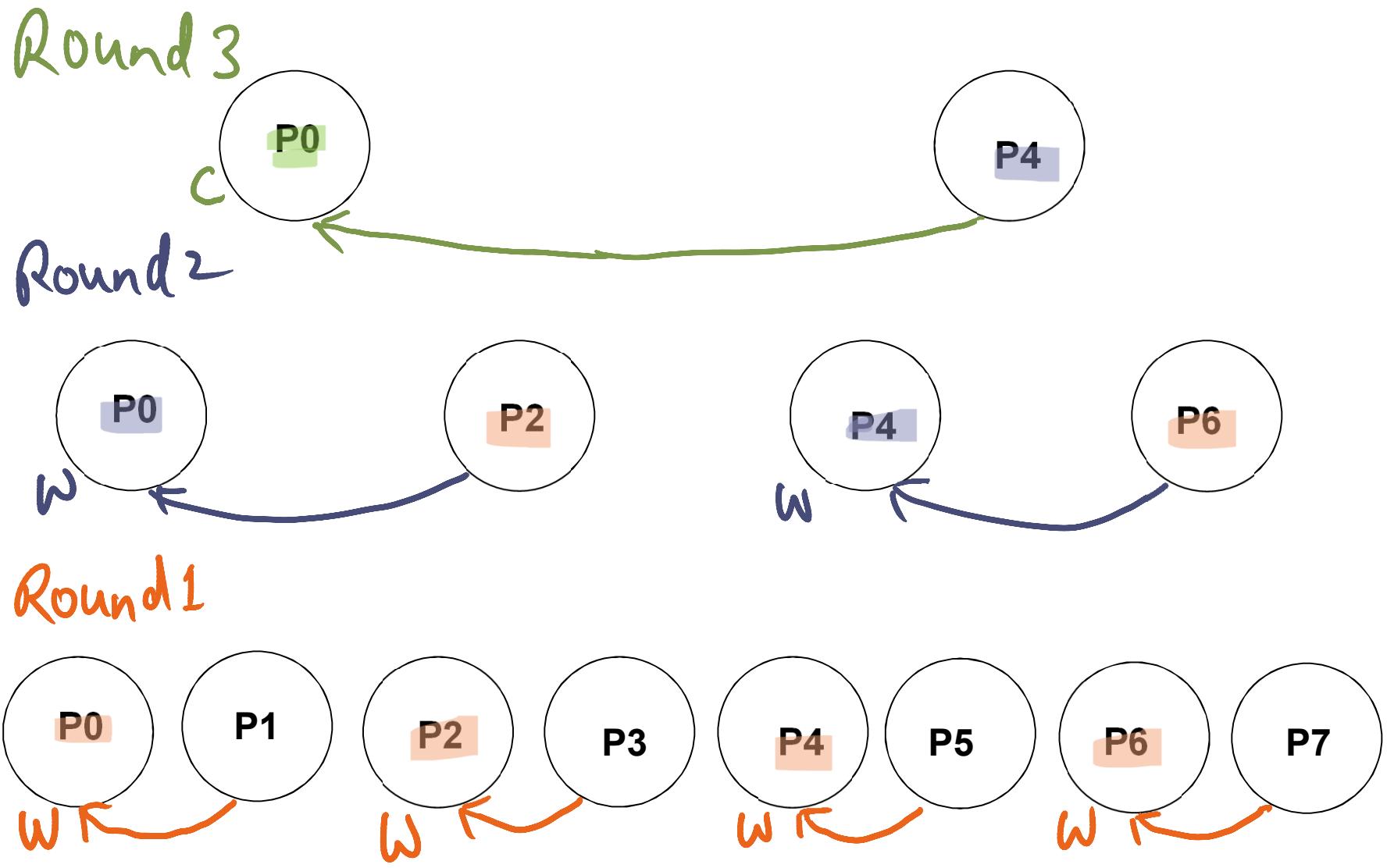
N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



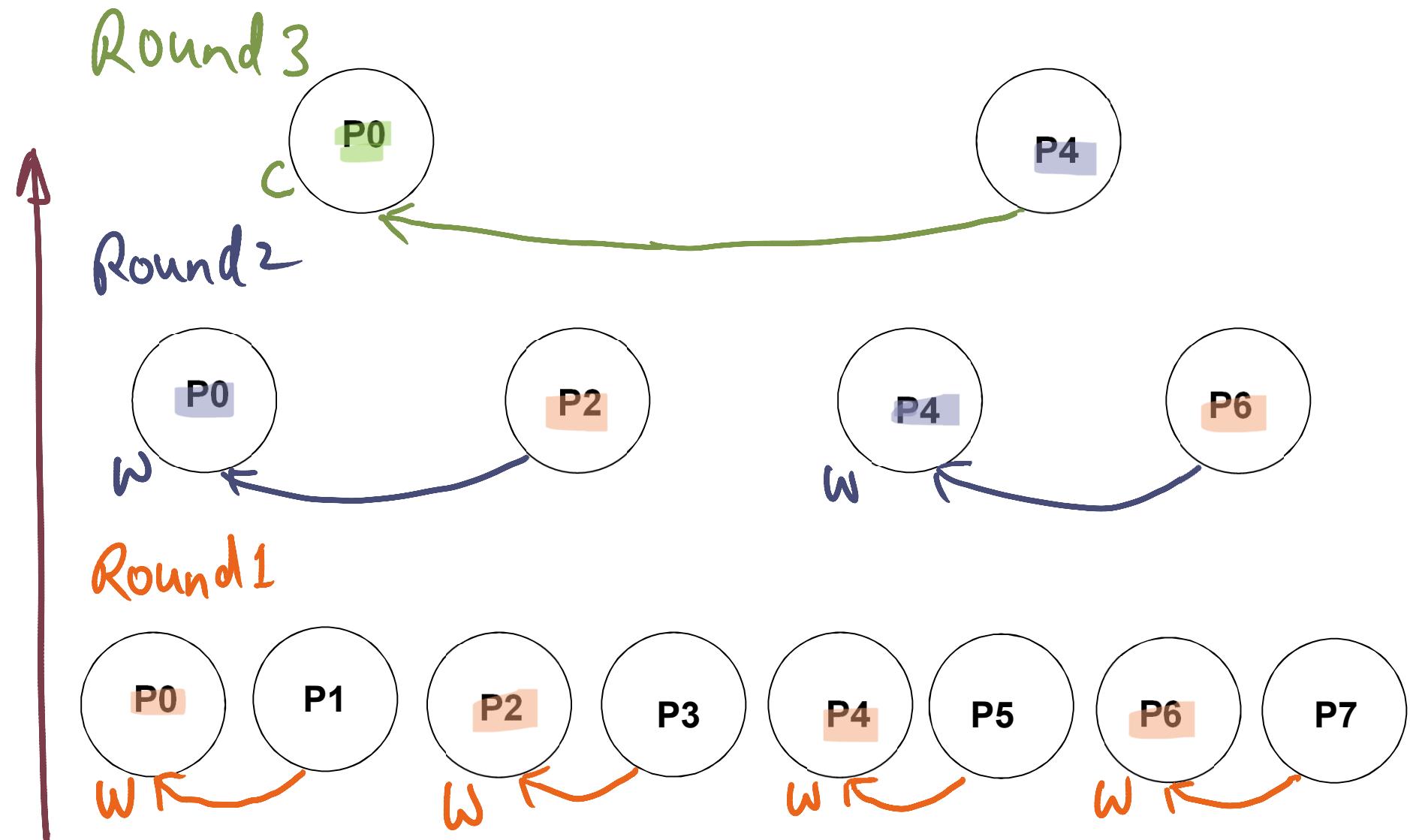
N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



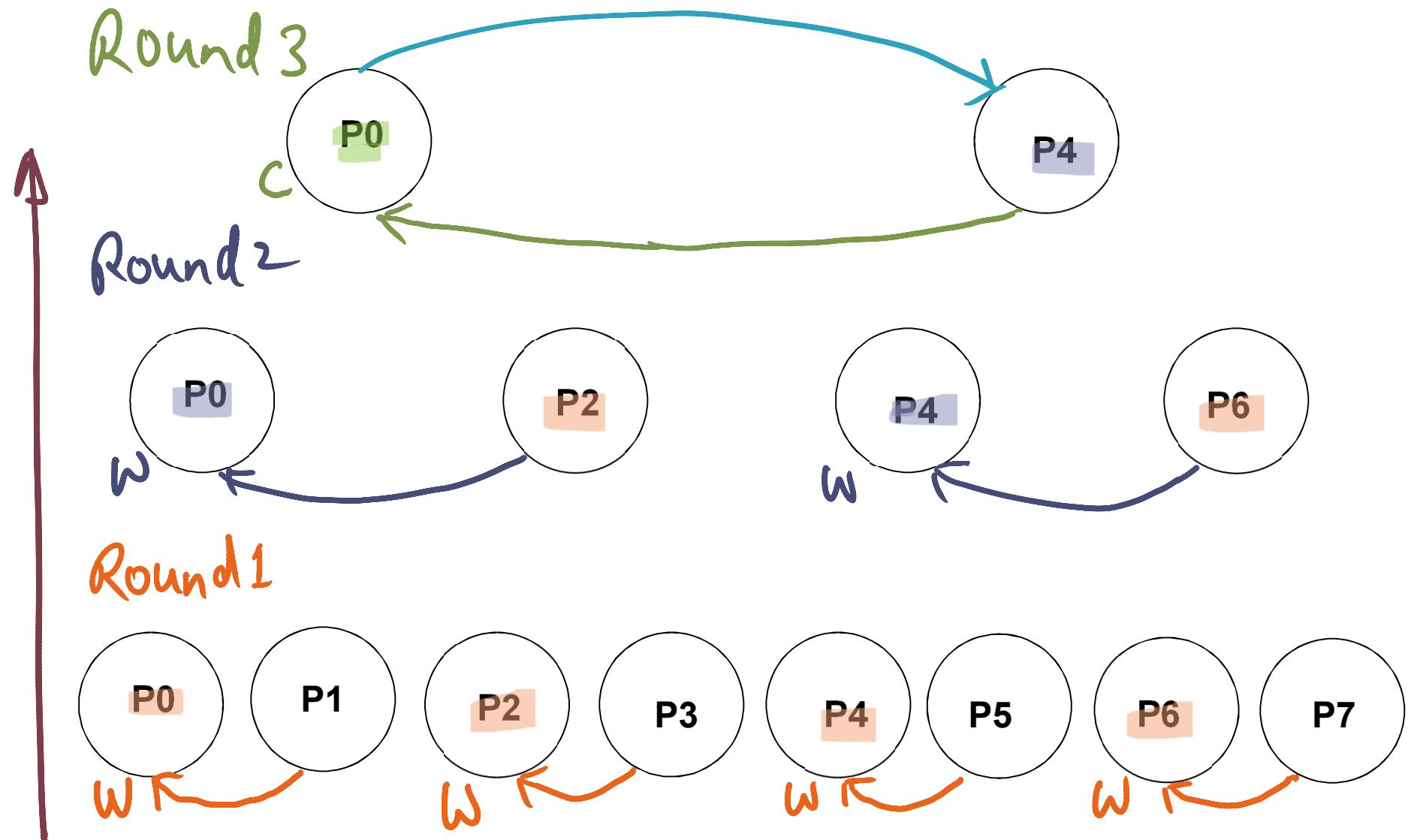
N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



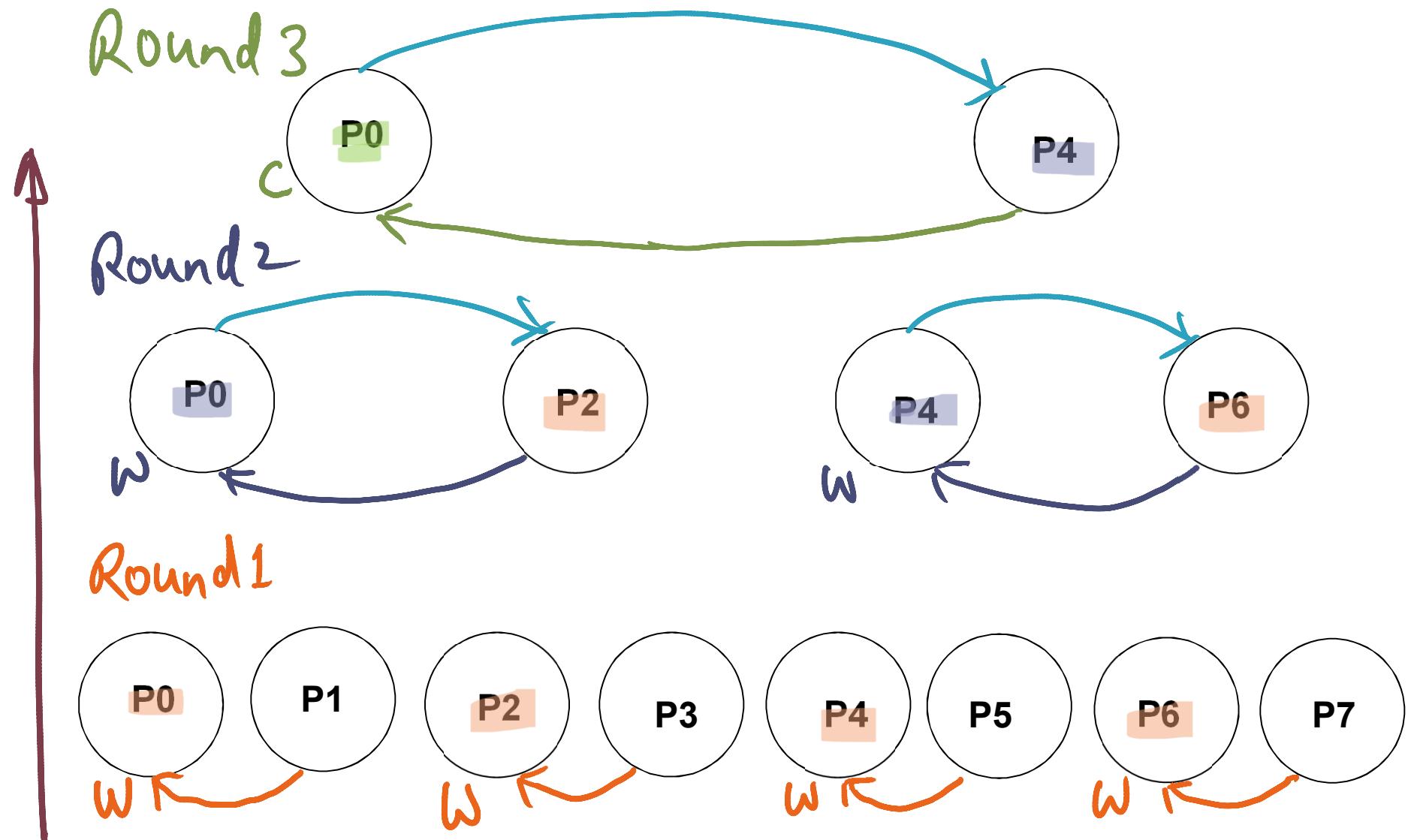
Arrival N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



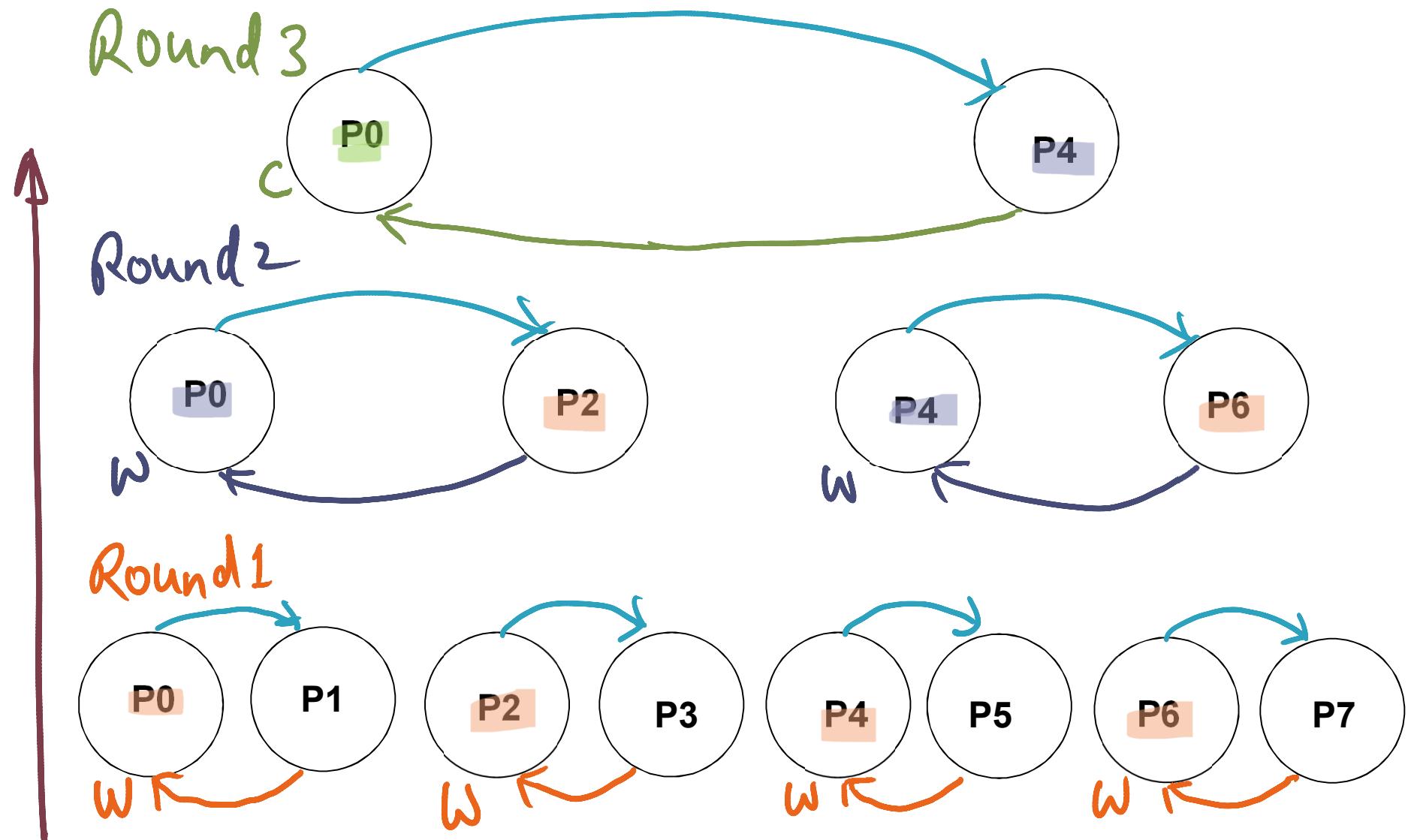
Arrival N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



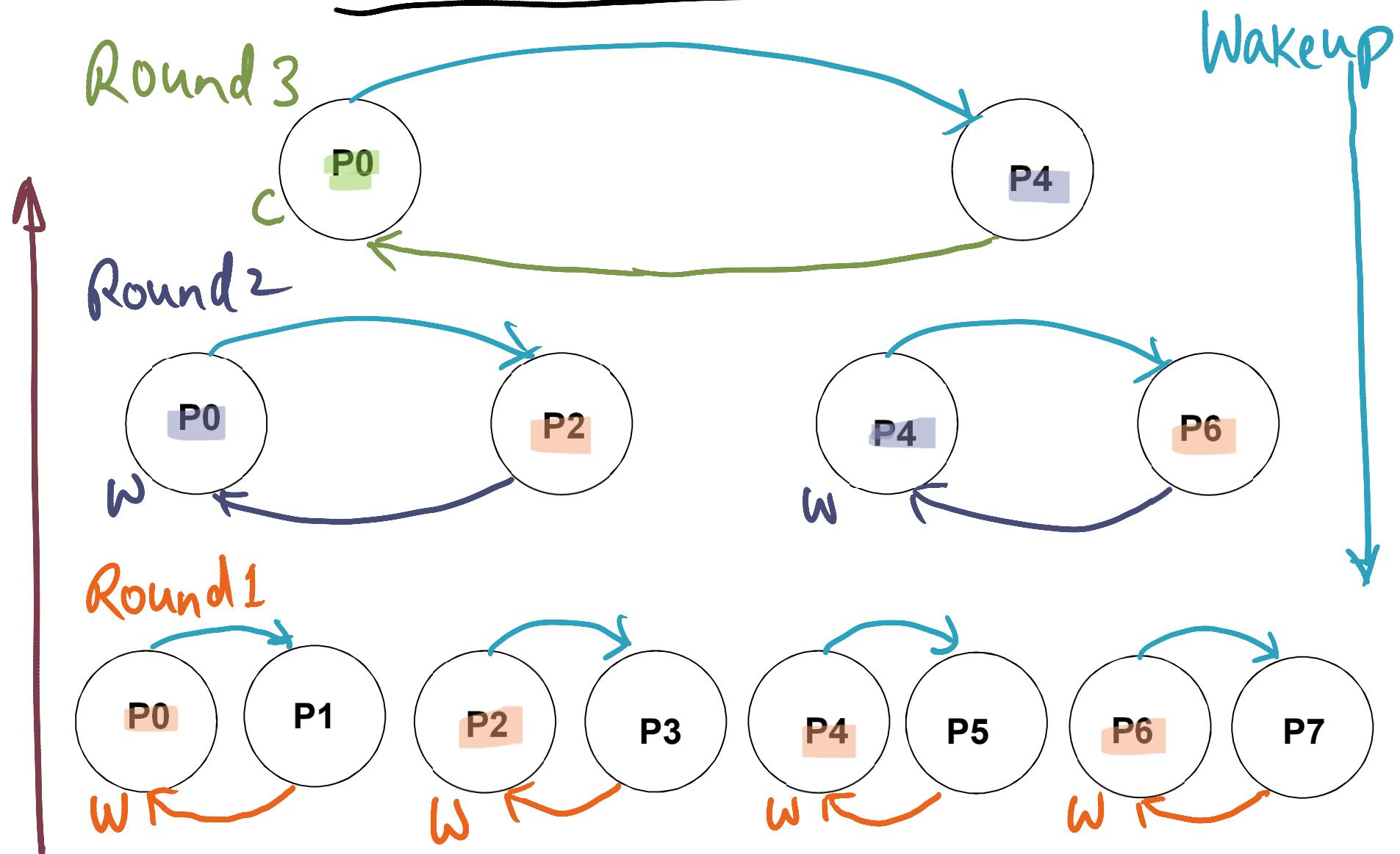
Arrival N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



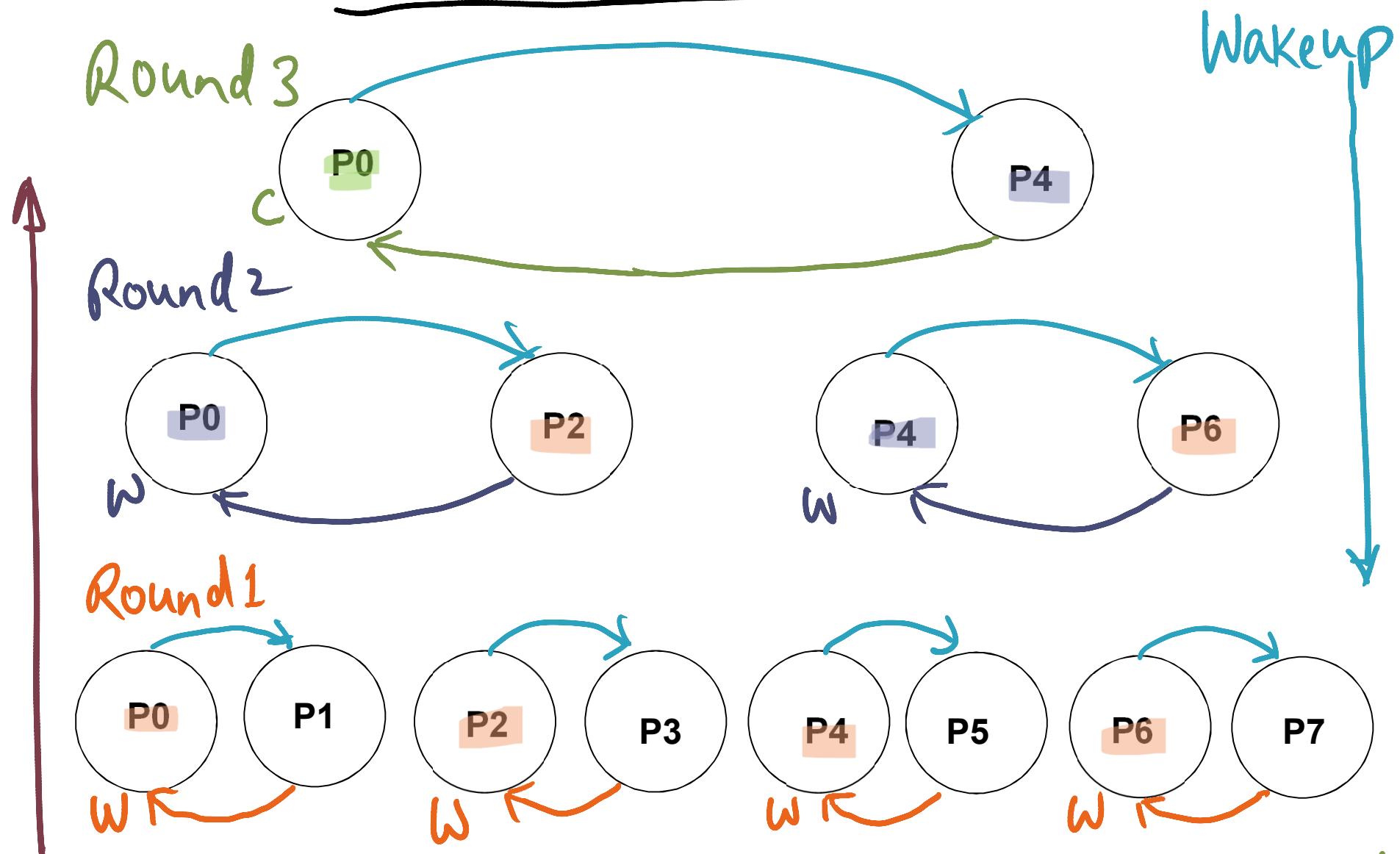
Arrival N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



Arrival N Players $\Rightarrow \log_2 N$ rounds

Tournament Barrier



Arrival N Players $\Rightarrow \log_2 N$ rounds All Awake!

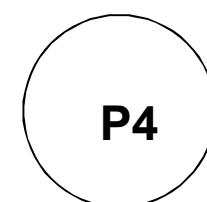
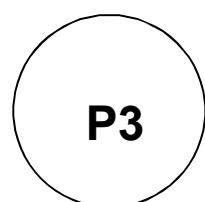
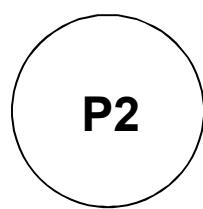
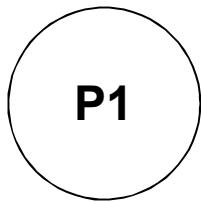
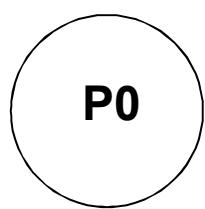
Virtues of Tournament

- Static determination of the spin variable (compared to Tree barrier)
 - => allocate the spin variable close to the winner in each of the matches.
- No need for any fetch-and-phi operation!
 - => Shared atomic R/W instructions are sufficient to implement this barrier algorithm.
- Total amount of comm. needed is similar to binary tree barrier => O(N)
- Comm. can exploit parallel paths in the ICN if available
- Alg works in a cluster machine as well
 - arrows are just message sent from one processor to the other.

Comparison to MCS

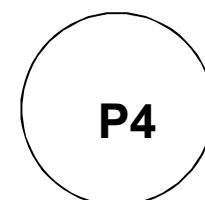
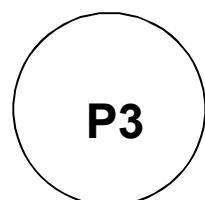
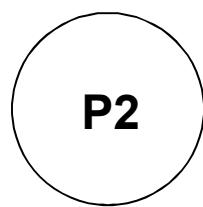
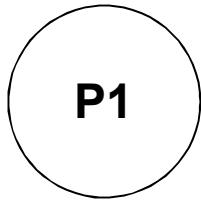
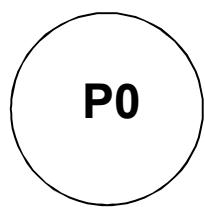
- Does not exploit spatial locality of caches due to the binary nature of the algorithm => MCS allocates multiple spin variables in the same cache line, so that the spinning is more efficient and the communication is lesser.
- Like MCS no need for fetch-and-phi
- But tournament is good on clusters as well while MCS is not

Dissemination Barrier



Dissemination Barrier

N need not be power of 2

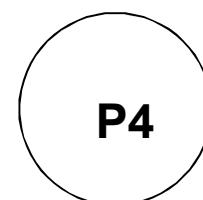
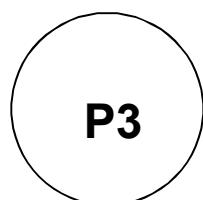
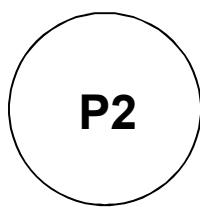
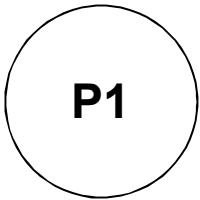
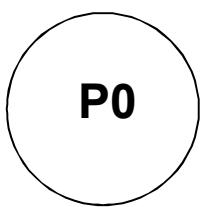


Dissemination Barrier

N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^k) \bmod N}$$



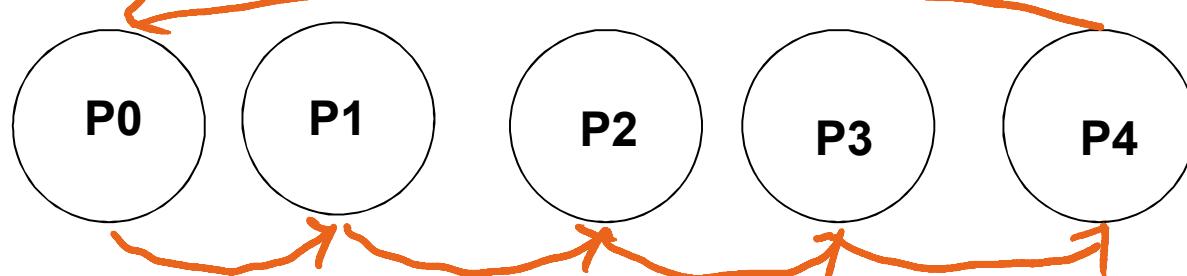
Dissemination Barrier

N need not be power of 2

Each round(K):

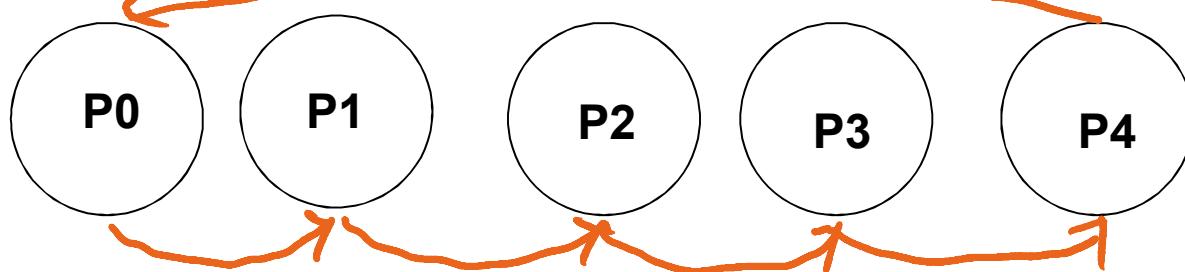
$$P_i \rightarrow P_{(i+2^K) \bmod N}$$

Round \varnothing



Dissemination Barrier

Round \varnothing



N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^k) \bmod N}$$

$O(N)$ comm. events
per round

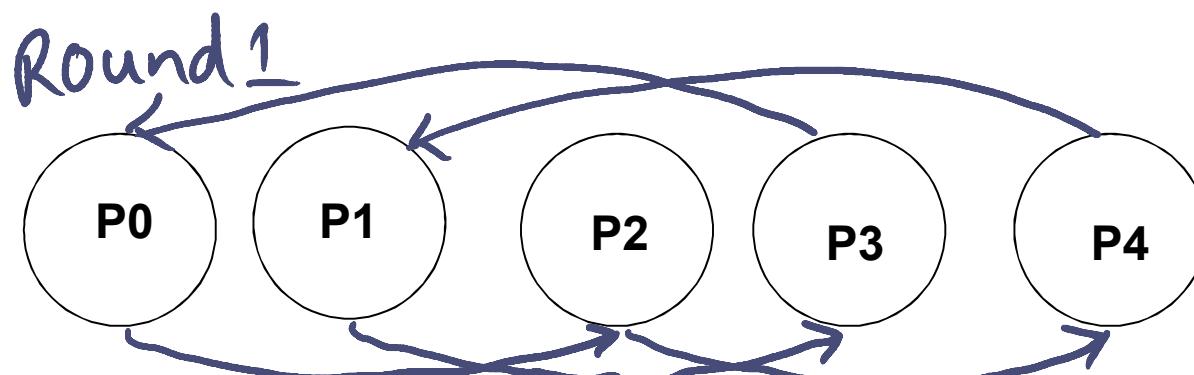
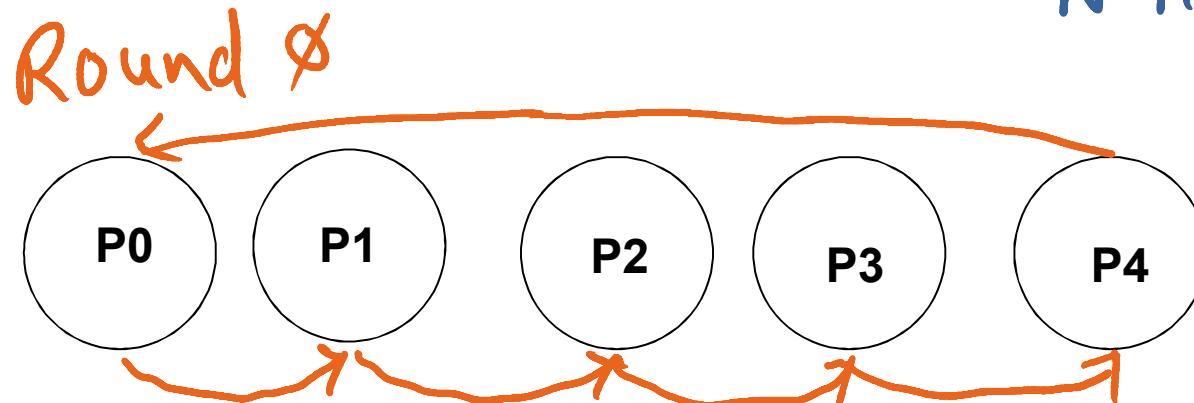
Dissemination Barrier

N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^K) \bmod N}$$

$O(N)$ comm. events
per round



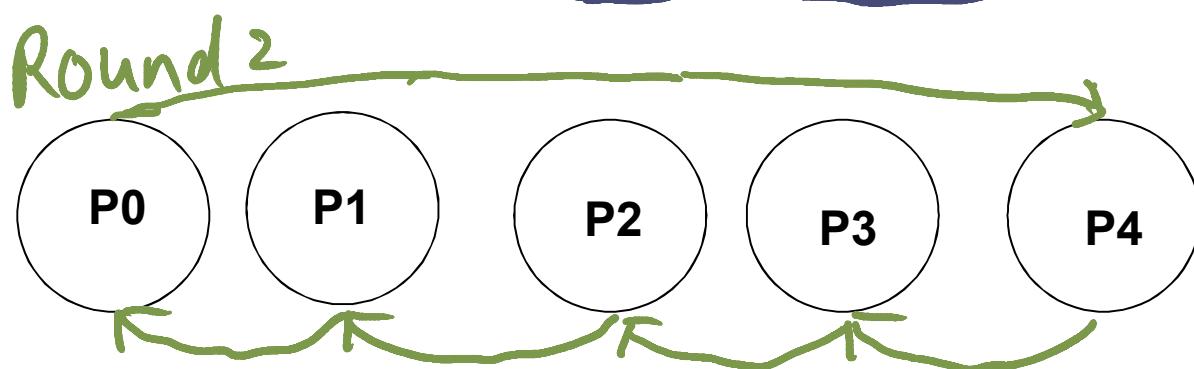
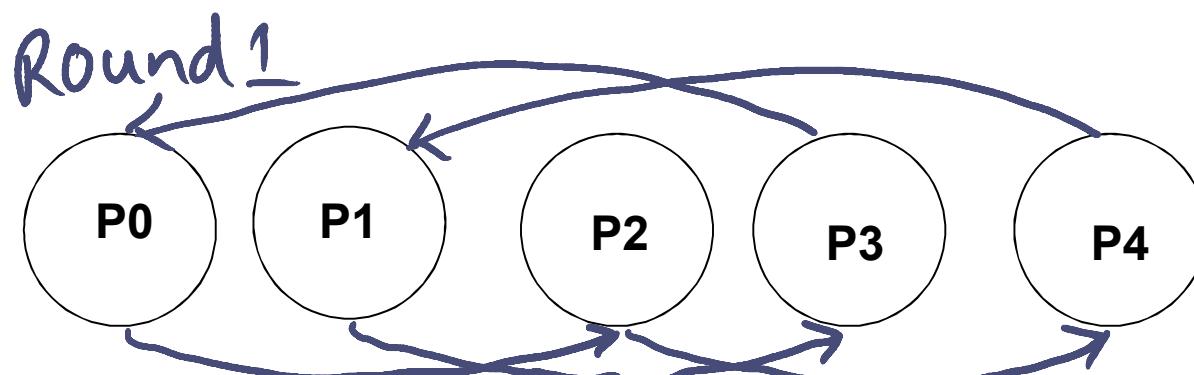
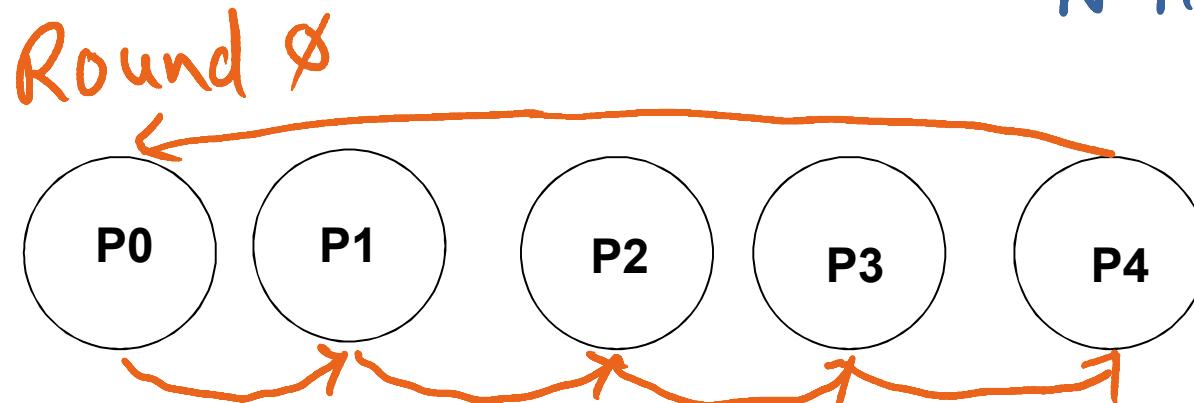
Dissemination Barrier

N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^K) \bmod N}$$

$O(N)$ comm. events
per round



Question

How many rounds for carrier completion?

$N \log_2 N$

$\log_2 N$

$\lceil \log_2 N \rceil$

N

Question
How many rounds for carrier completion?

$N \log_2 N$

$\log_2 N$

$\lceil \log_2 N \rceil$

N

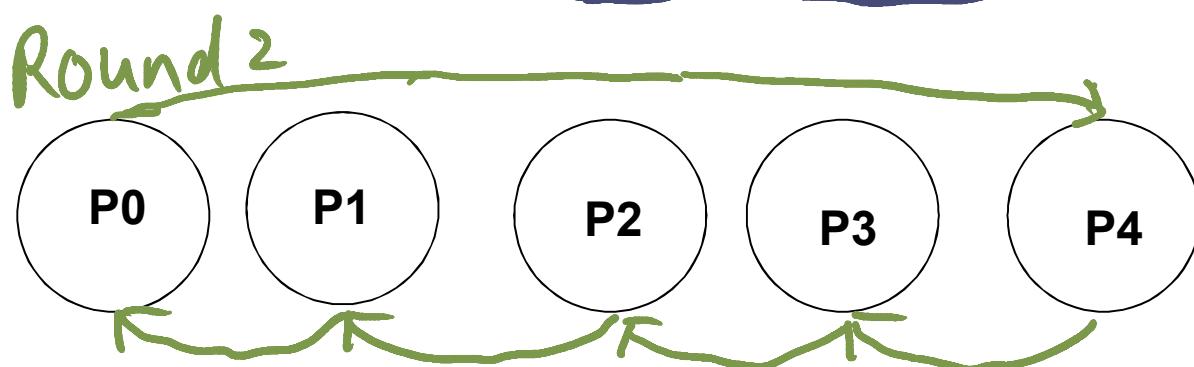
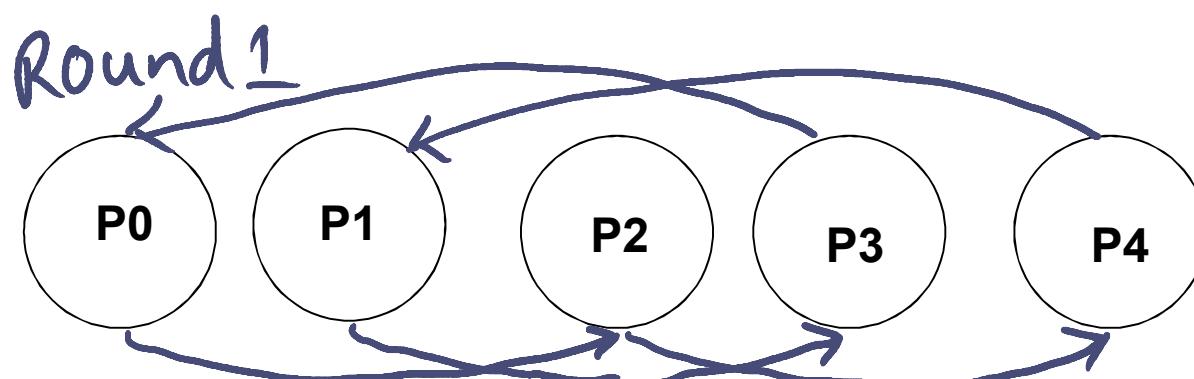
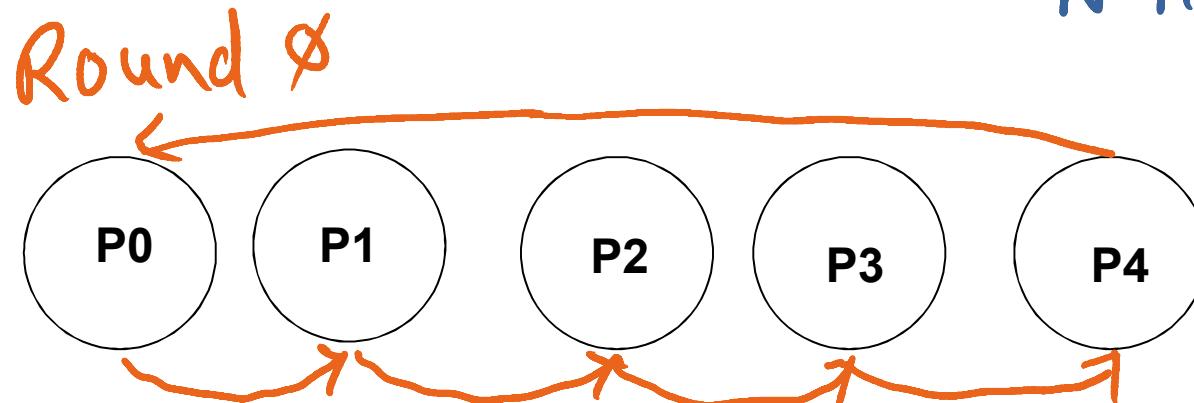
Dissemination Barrier

N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^K) \bmod N}$$

$O(N)$ comm. events
per round



Dissemination Barrier

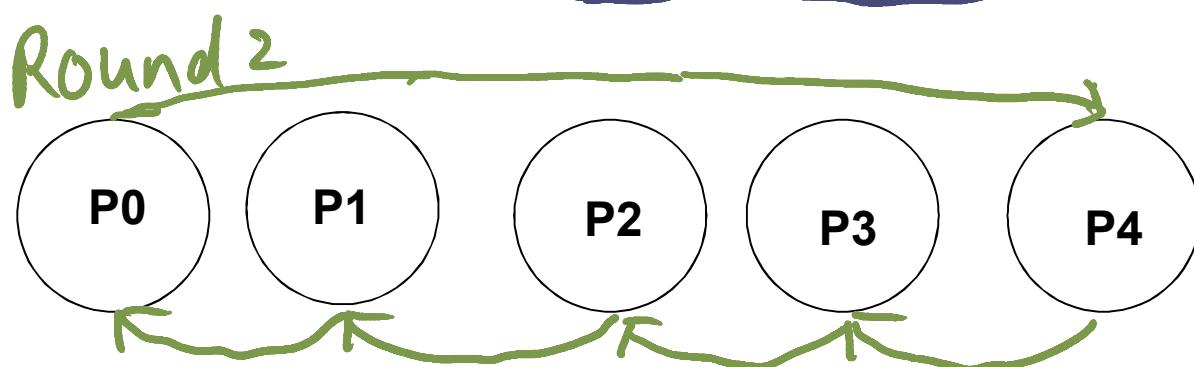
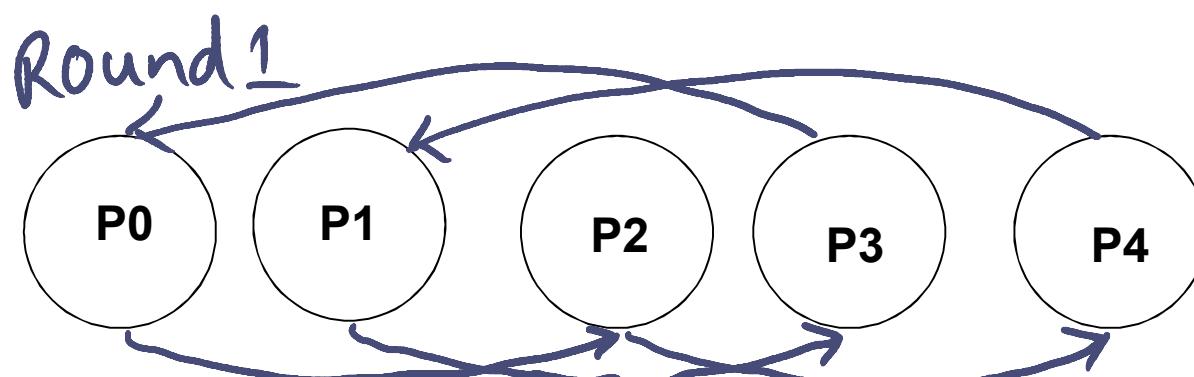
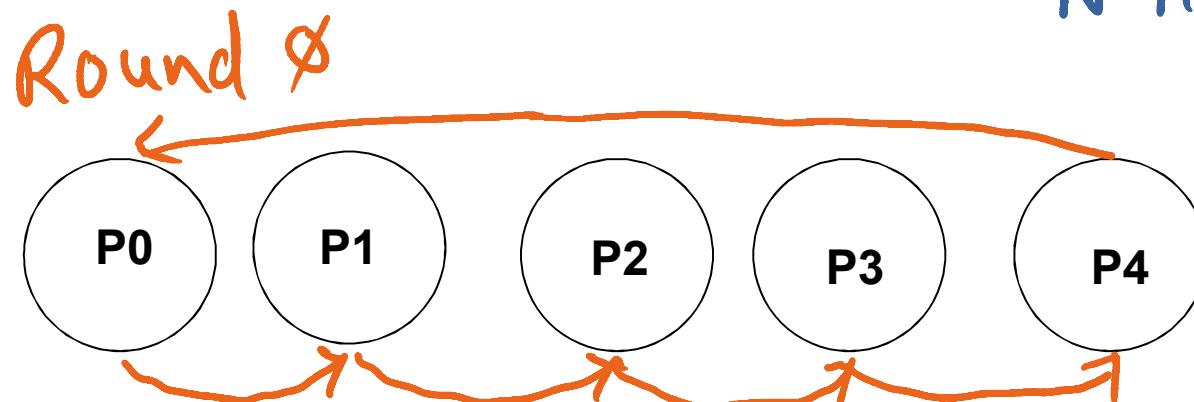
N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^k) \bmod N}$$

$O(N)$ comm. events
per round

- $\lceil \log_2 N \rceil$ rounds



Dissemination Barrier

N need not be power of 2

Each round(K):

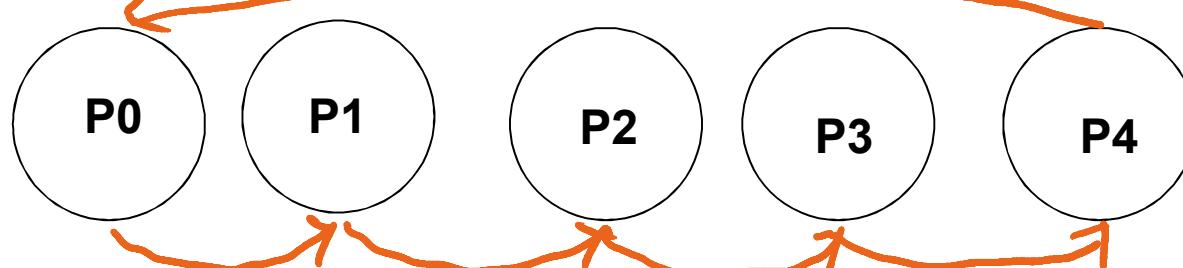
$$P_i \rightarrow P_{(i+2^K) \bmod N}$$

$O(N)$ comm. events
per round

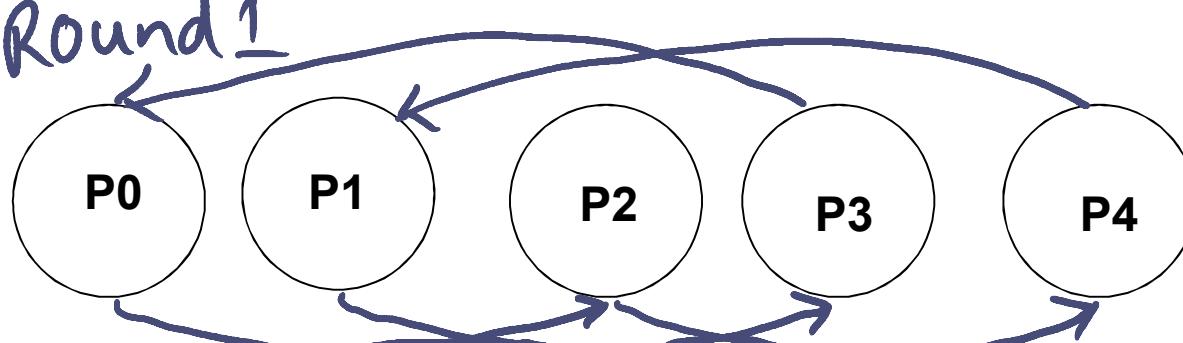
- $\lceil \log_2 N \rceil$ rounds

All Awake!

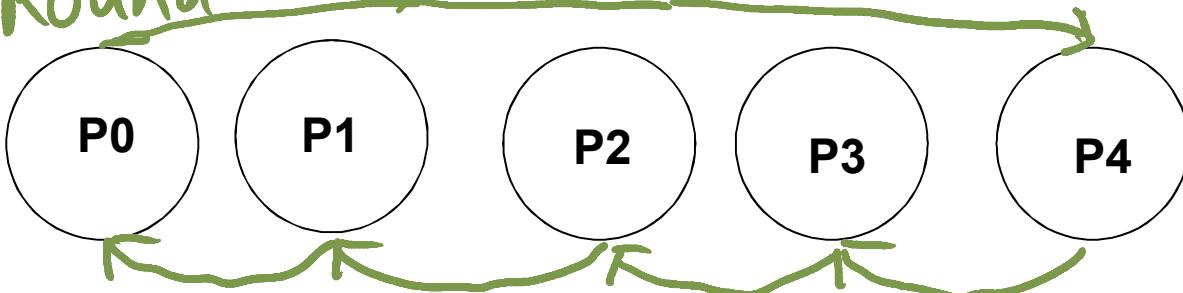
Round 0



Round 1

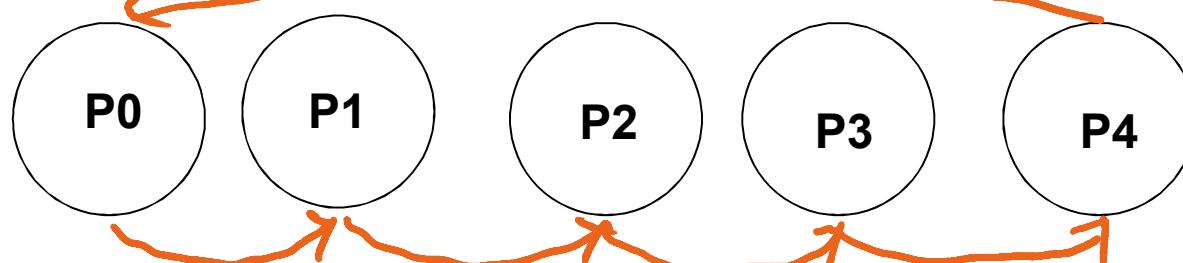


Round 2

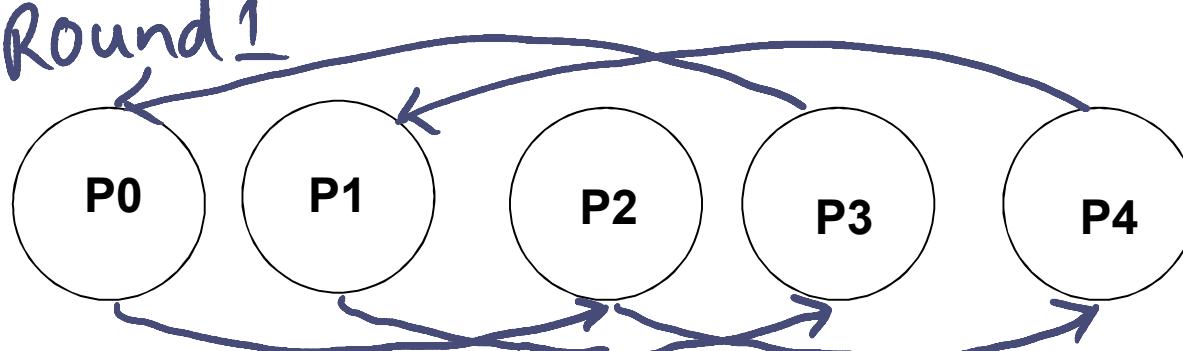


Dissemination Barrier

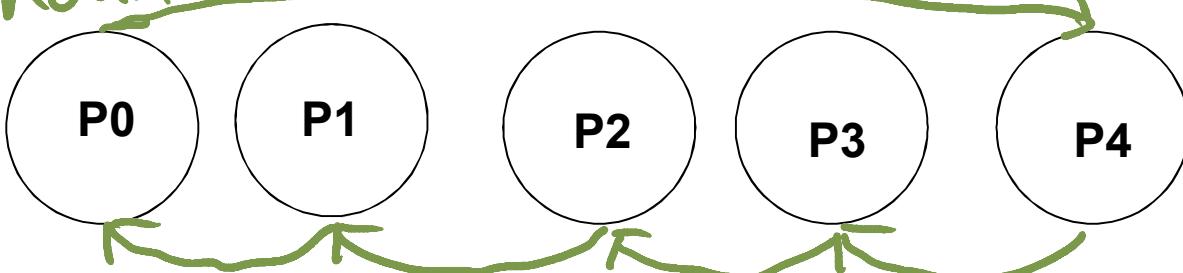
Round 0



Round 1



Round 2



N need not be power of 2

Each round(K):

$$P_i \rightarrow P_{(i+2^k) \bmod N}$$

$O(N)$ comm. events
per round

- $\lceil \log_2 N \rceil$ rounds

All Awake!

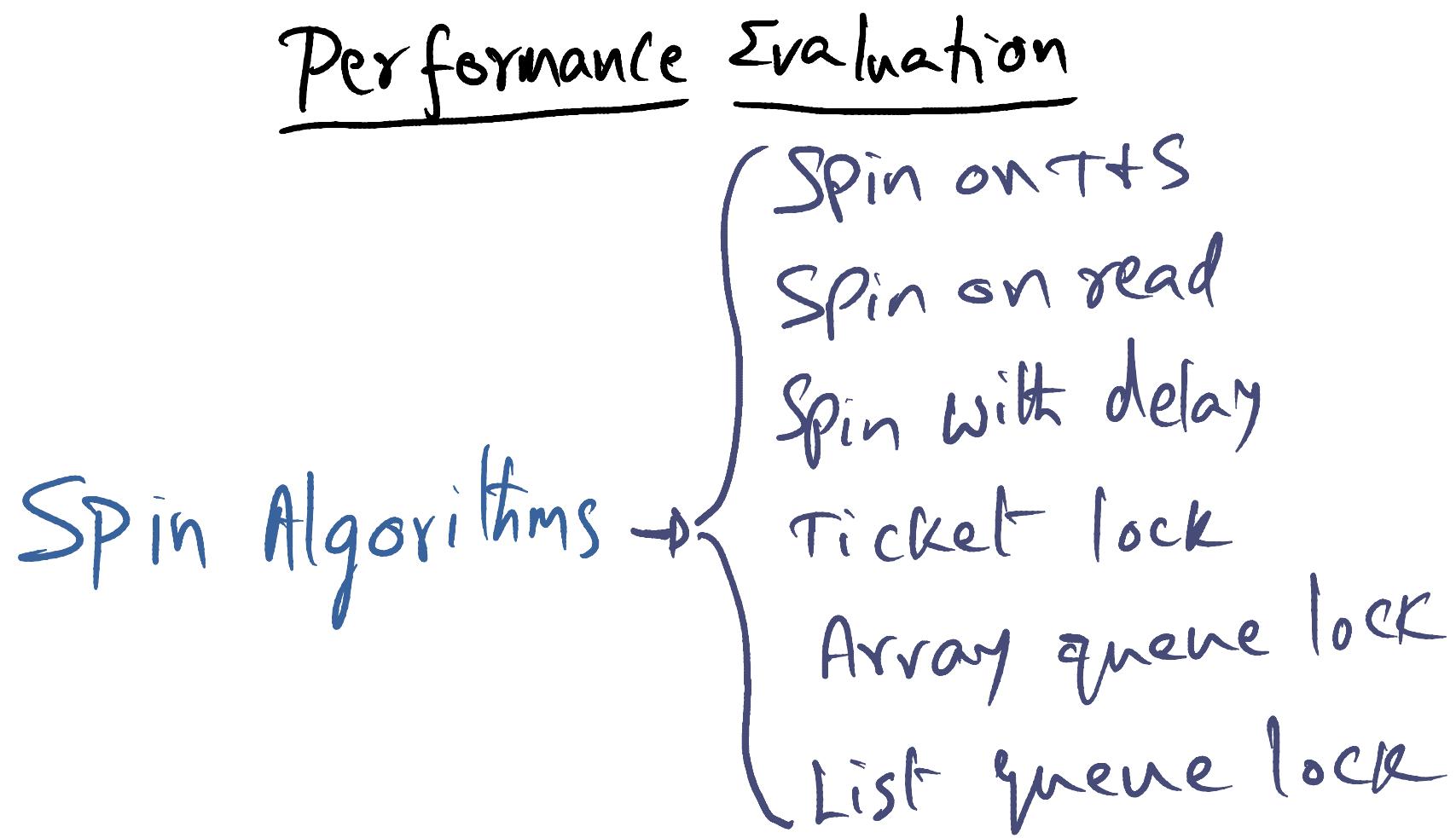
- No hierarchy

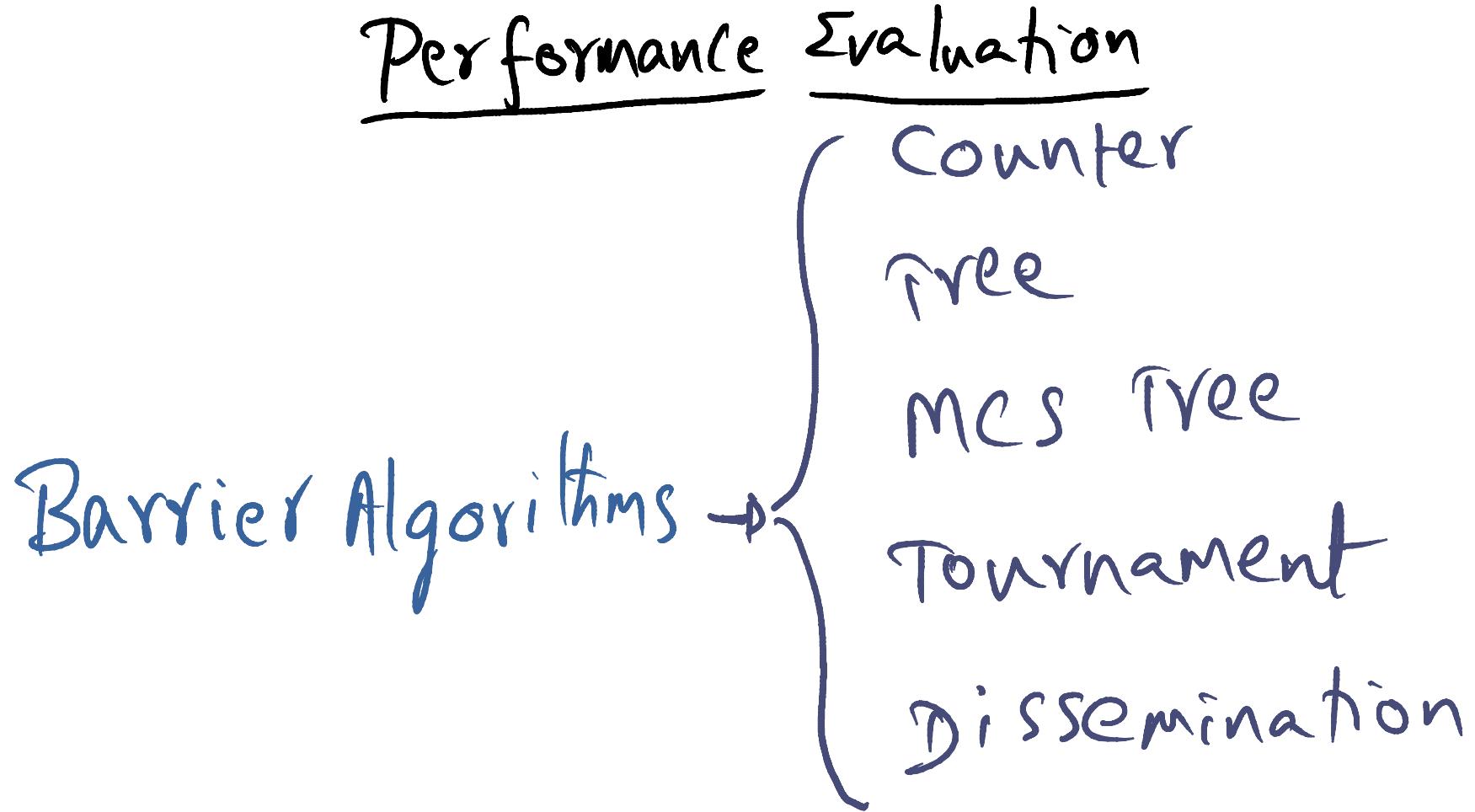
- works for NCC
and clusters

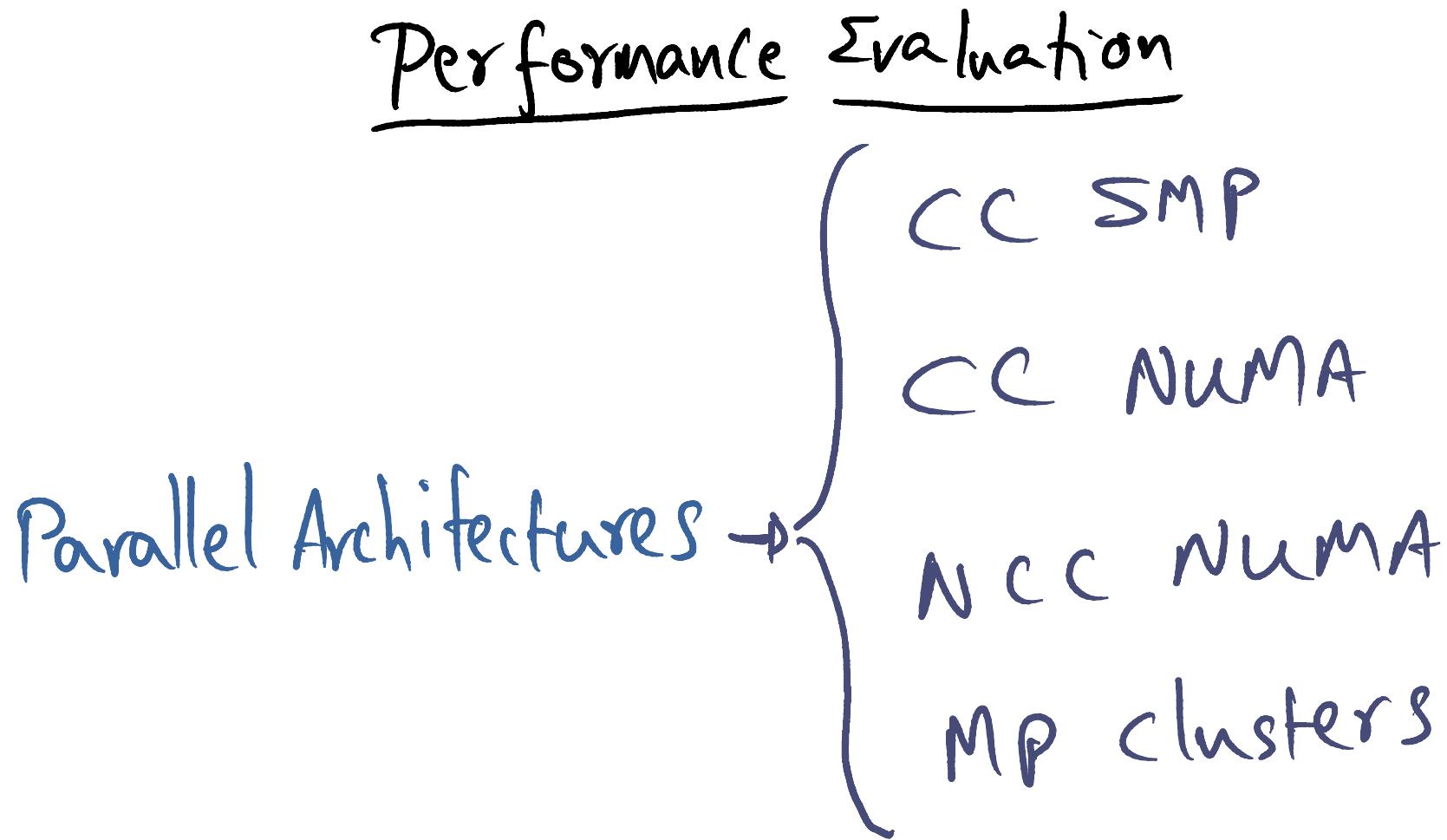
Virtues of Dissemination

- No hierarchy
- No pair wise sync
- Each node works independently to send out the messages per protocol
- All will realize the barrier is complete when they have received the $\text{ceil}(\log_2 N)$ messages from their peers
- Total amount of comm: $O(N \log_2 N)$
- In comparison, tournament incurs $O(\log_2 N)$ comm since in that set up the comm decreases as we go up the tree; since dissemination is based on gossip, the amount of comm in each round is constant and equal to N
- Works for NCC and MP machines in addition to SM

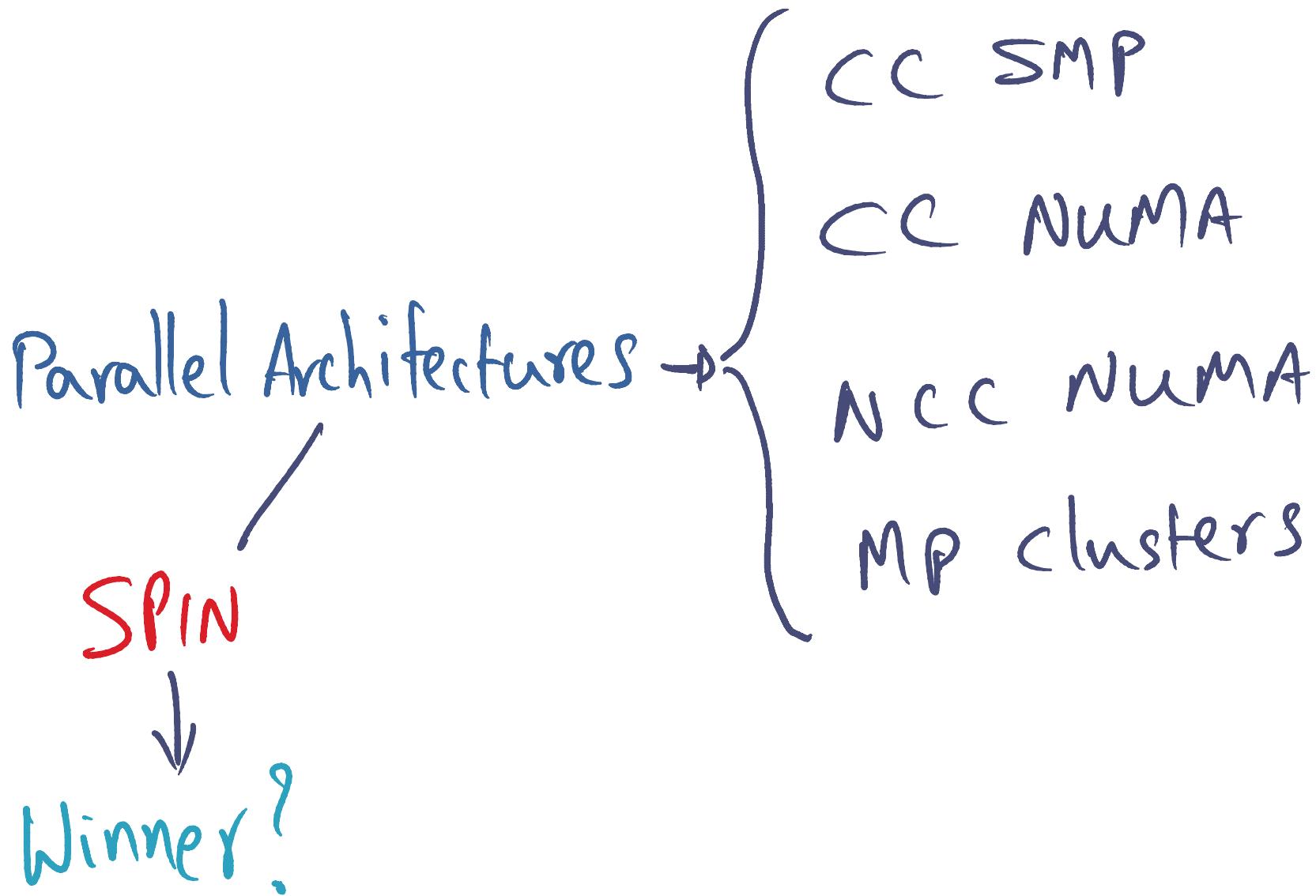
Performance Evaluation

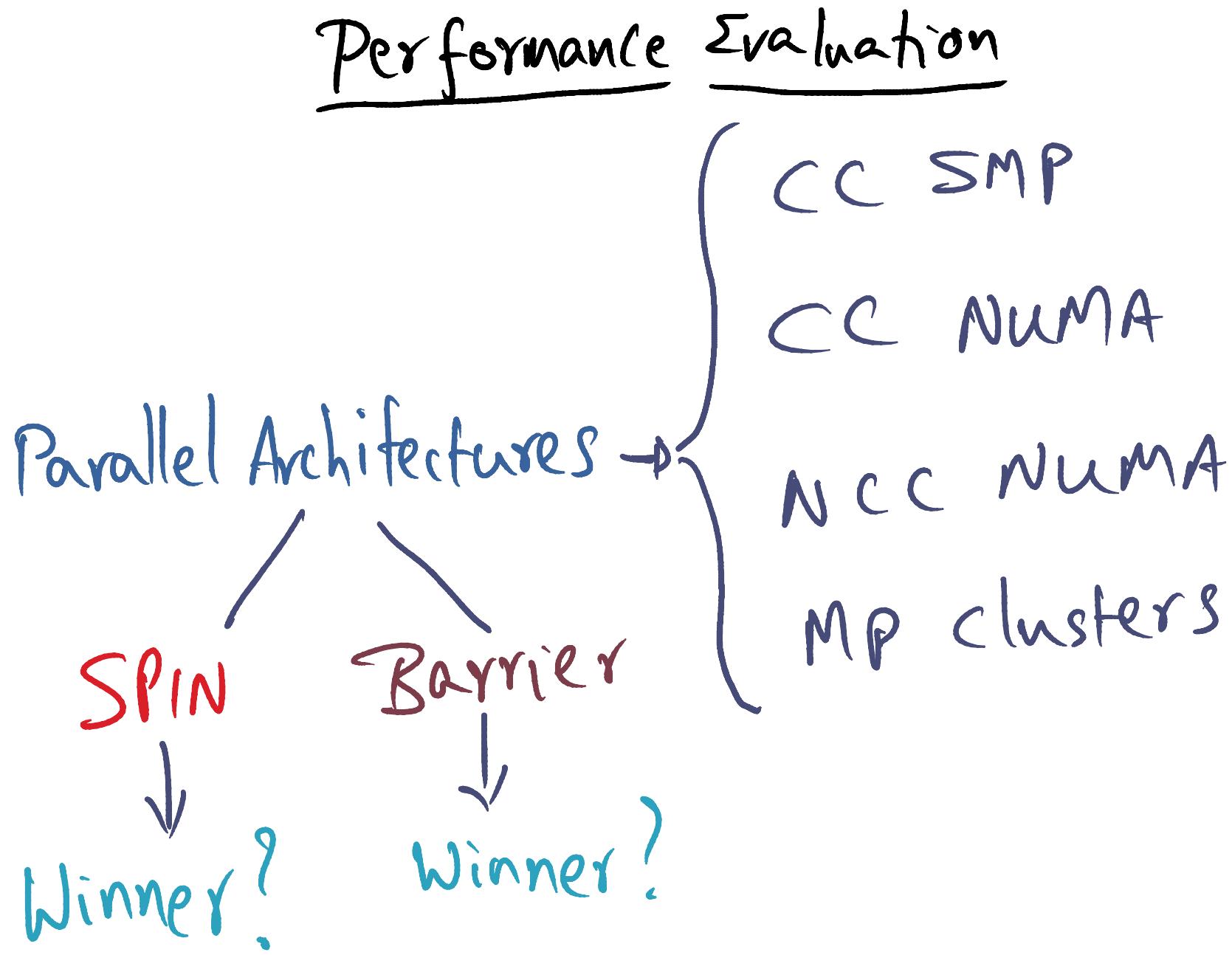






Performance Evaluation





BBN Butterfly

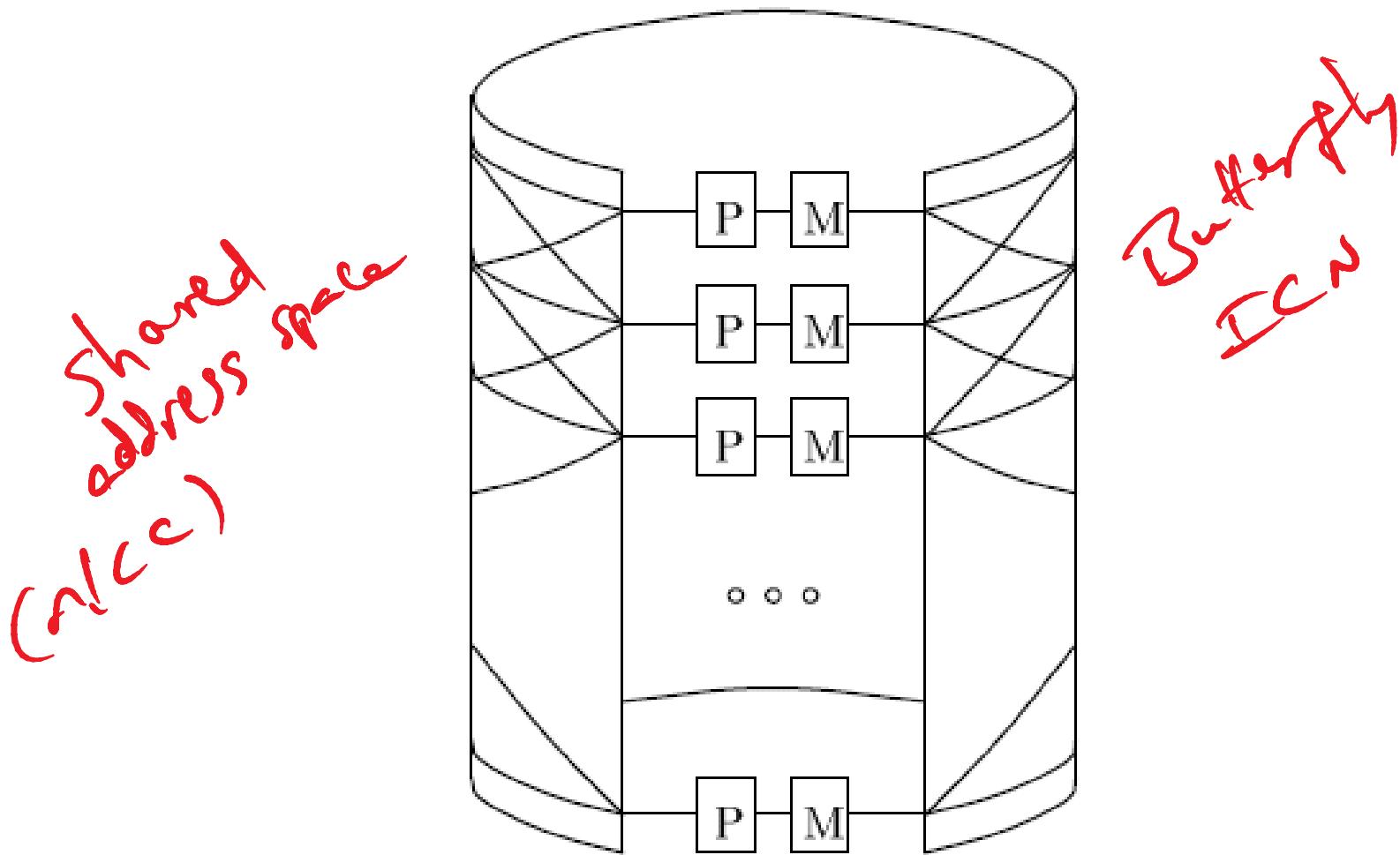


Figure 2: The BBN Butterfly 1.

Sequent Symmetry

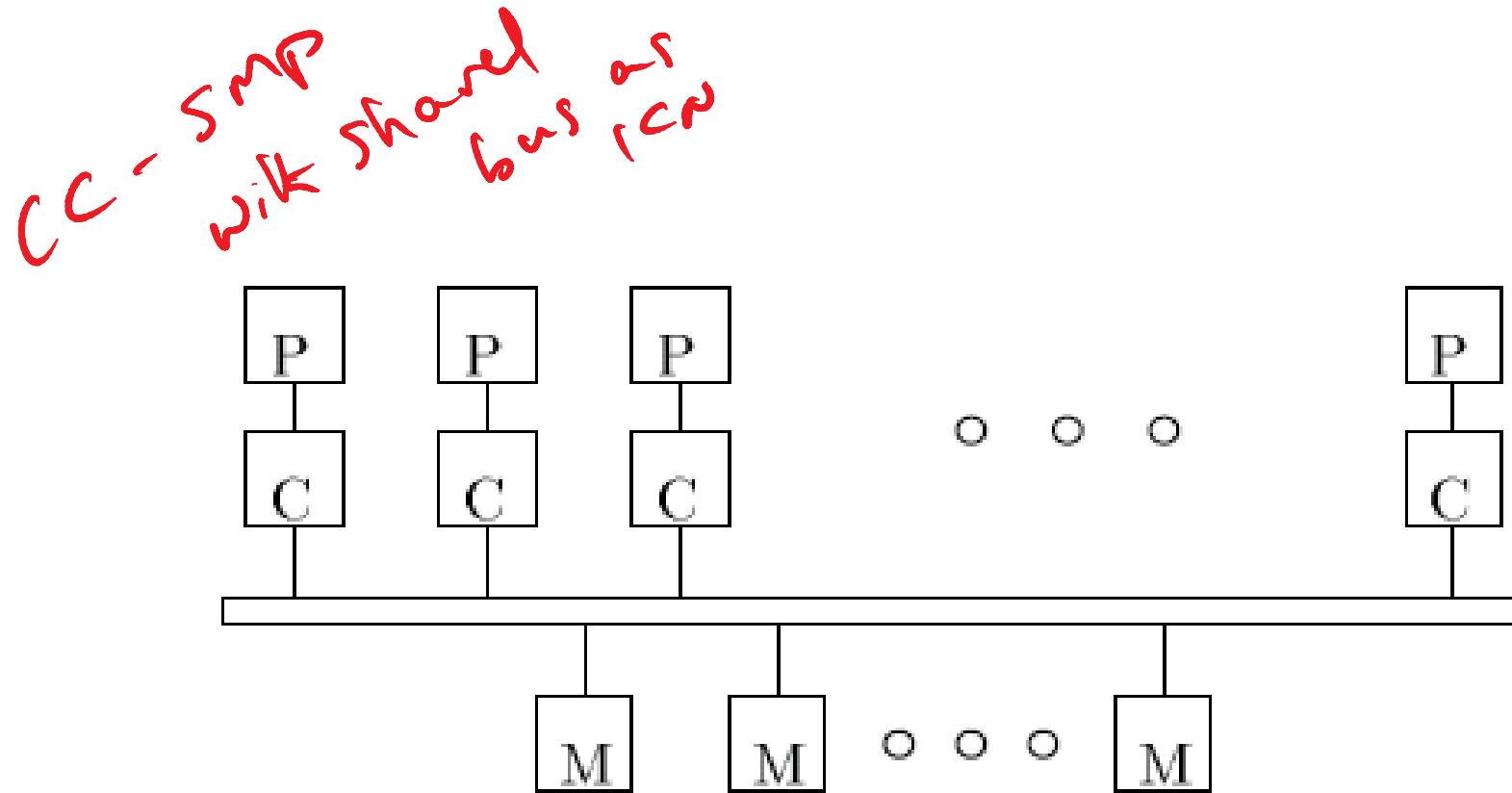


Figure 3: The Sequent Symmetry Model B.

Performance

- Spinlock
 - Which one do you think will win on symmetry?
 - Which one do you think will win on BBN?
- Barrier
 - Which one do you think will win on symmetry?
 - Which one do you think will win on BBN?

*- form an intuition
- verify whether reported
result agree with
your opinion*

Compare their barrier and lock results on symmetry and BBN against the results reported for the same algorithms on a different parallel machine, namely, **KSR-1** in the paper “Scalability study of KSR-1”
(See T-Square Modules: Section 10.10)