Jin-Soo Kim
(*jinsoo.kim@snu.ac.kr*)

Systems Software &
Architecture Lab.

Seoul National University

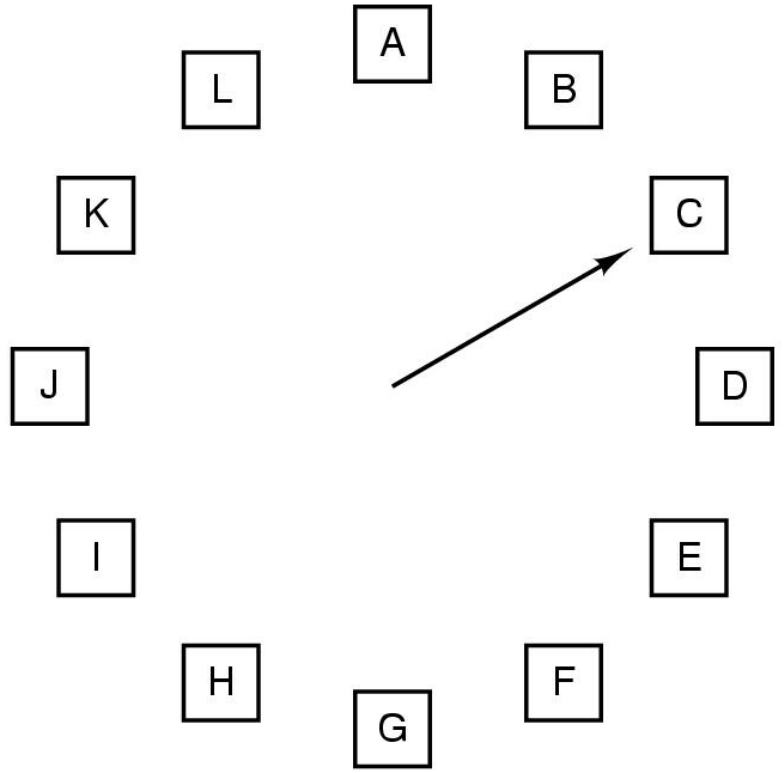
Fall 2020

Clock-Pro

(Song Jiang et al., USENIX ATC '05)

CLOCK

- An LRU approximation algorithm
- Uses R (Reference) bit in each PTE
- Arranges all of physical frames in a big circle
- A clock hand is used to select a victim
  - If (R == 1), turn in off and go to next page (second chance)
  - if (R == 0), evict the page
  - The hand moves quickly when pages are needed
- If memory is large, "accuracy" of information degrades



GCLOCK [ACM TODS '78]

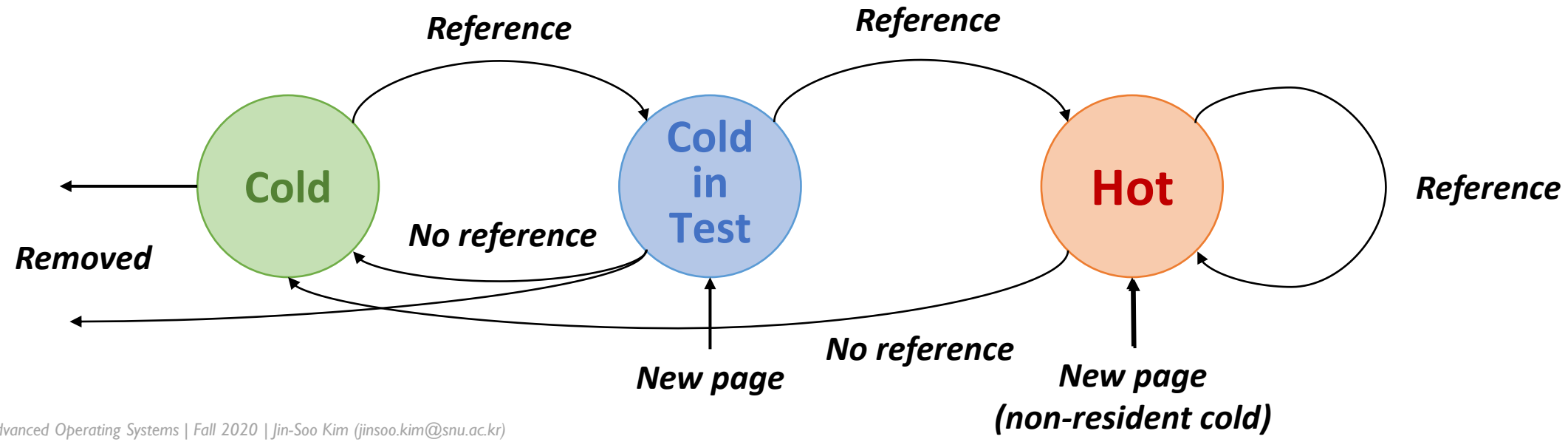
- Generalized CLOCK page replacement algorithm
- Associate a counter with each page frame
- GCLOCK(i)
  - Whenever a page is referenced, the counter is set to i
  - When a page fault occurs, a pointer that circles around the circular list of page frames is observed
  - If the counter is zero, the page is removed
  - Otherwise, the counter is decremented by 1, and the pointer is advanced to the next page
  - When a new page is placed due to demand fetch, the counter is set to i
- CLOCK == GCLOCK(1)

CLOCK-Pro

- An approximation of LIRS based on the CLOCK infrastructure
- Pages categorized into two groups: hot pages and cold pages based on their reuse distance (or inter-reference recency)
  - All hot pages (m_h pages) are resident
  - Some cold pages (m_c pages) are resident
  - Recently replaced m (= m_h + m_c) pages are tracked as non-resident cold pages
- Three hands are used:
  - **HAND$_{hot}$**: for hot pages
  - **HAND$_{cold}$**: for cold pages
  - **HAND$_{test}$**: for running a reuse distance test for a page
- The ratio between m_h and m_c are adaptively adjusted

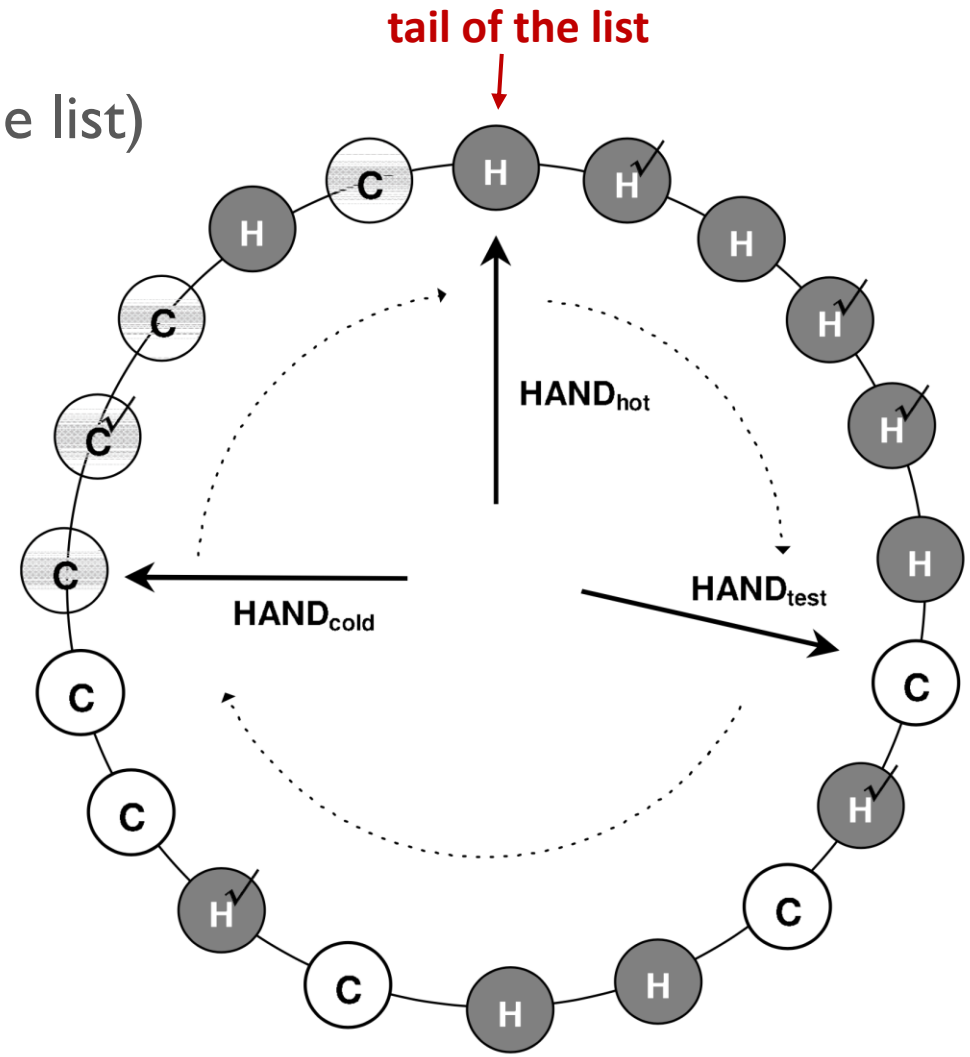
Basic Algorithm

- A cold page is granted a test period
  - A new page is inserted to the main memory as a cold page in test period
  - If the cold page in test period is referenced again, it is promoted to the hot page
  - If there is no re-reference to a cold page in test period, it is removed from the list
  - The hot page with the largest recency is demoted to cold page

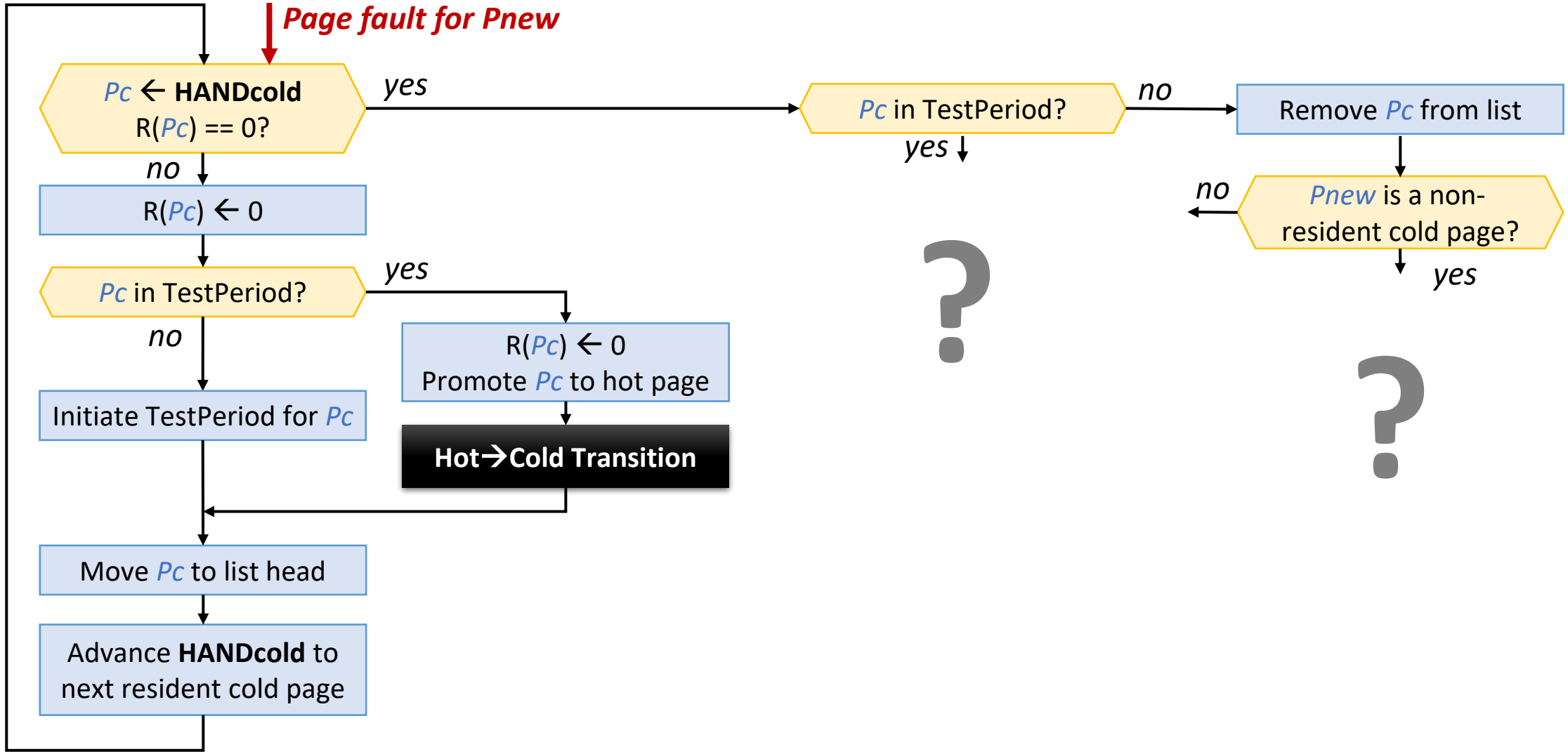


Clock Hands

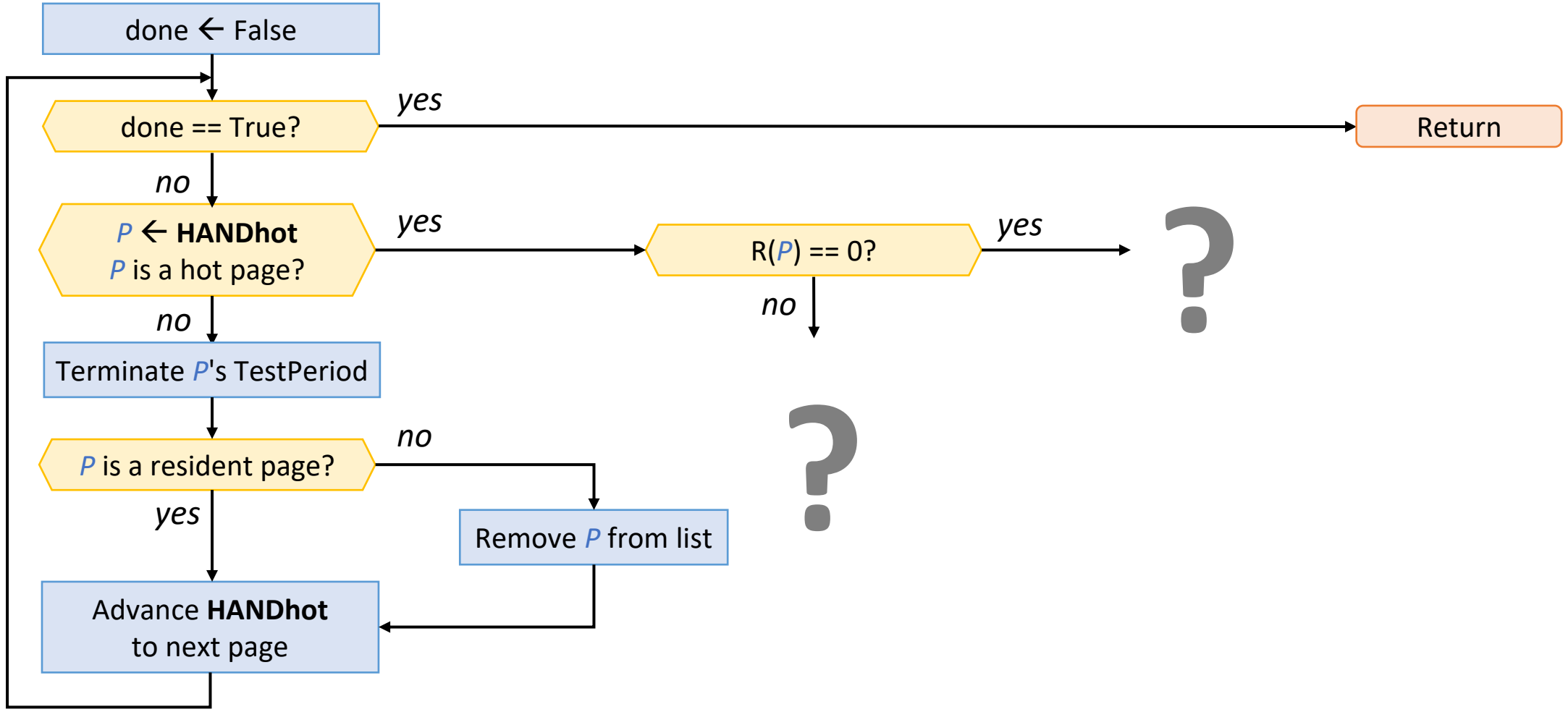
- **HAND$_{hot}$**
  - The hot page with the largest recency (tail of the list)
  - Used to turn hot pages into cold pages
- **HAND$_{cold}$**
  - The last resident page
  - Used to look for a victim page (same as the original clock hand)
- **HAND$_{test}$**
  - The last cold page in the test period
  - Used to remove non-resident cold pages



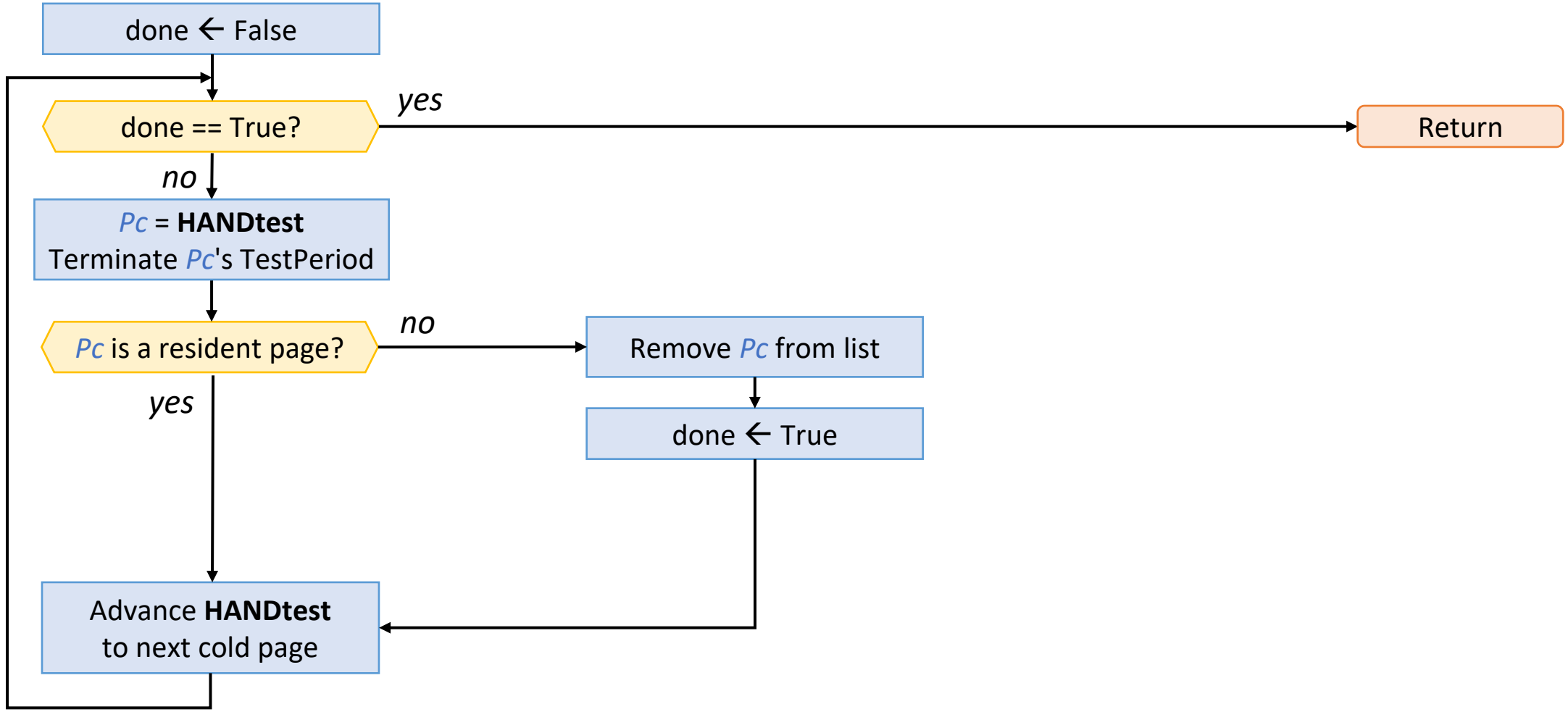
HANDcold: Handling Page Fault



HANDhot: Hot→Cold Transition



HANDtest: Cold Page Removal

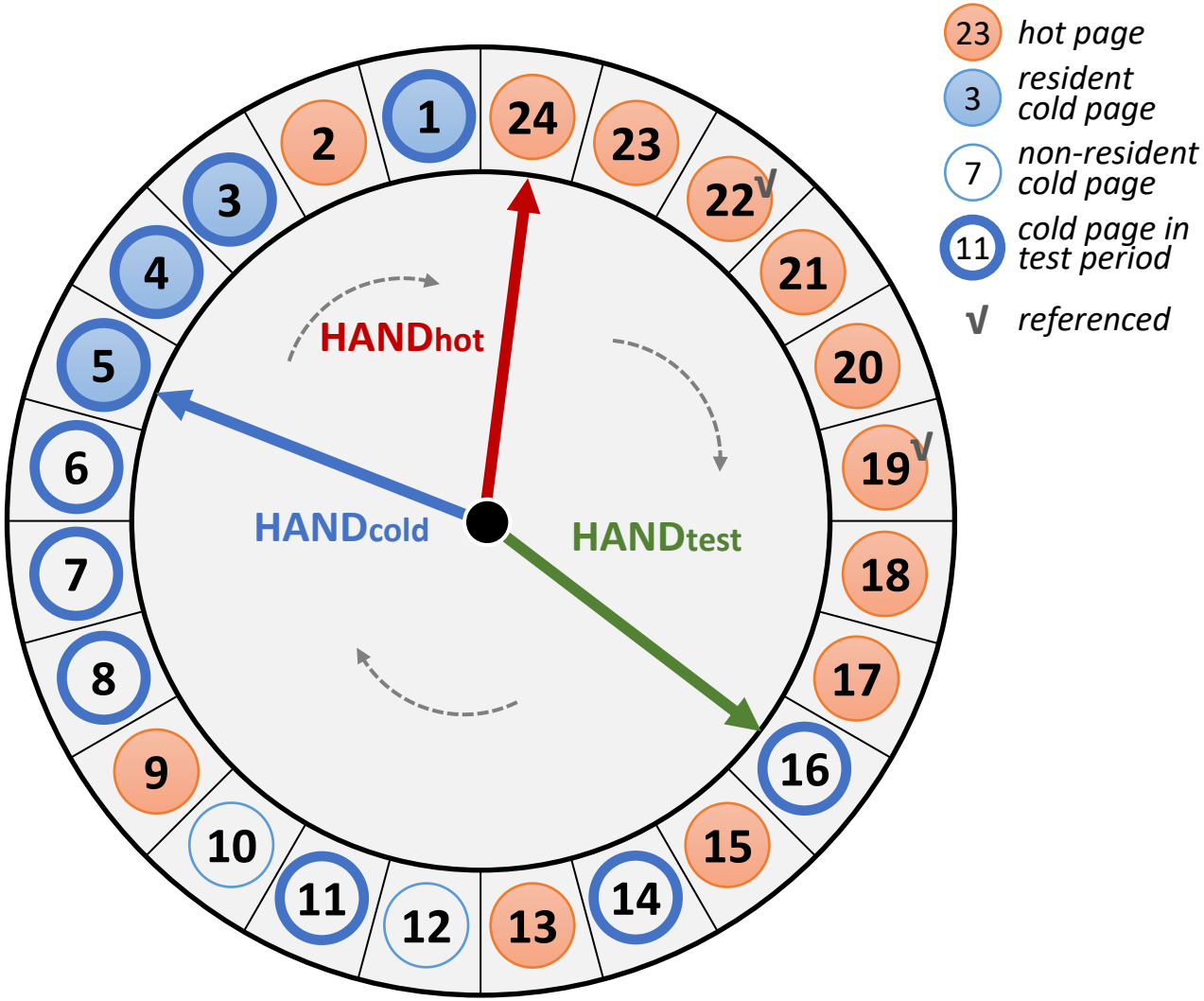


Example:

- 12 hot pages (resident)
- 4 resident cold pages
- 8 non-resident cold pages (metadata only)

- Reference bits of page 22 and 19 are set to 1

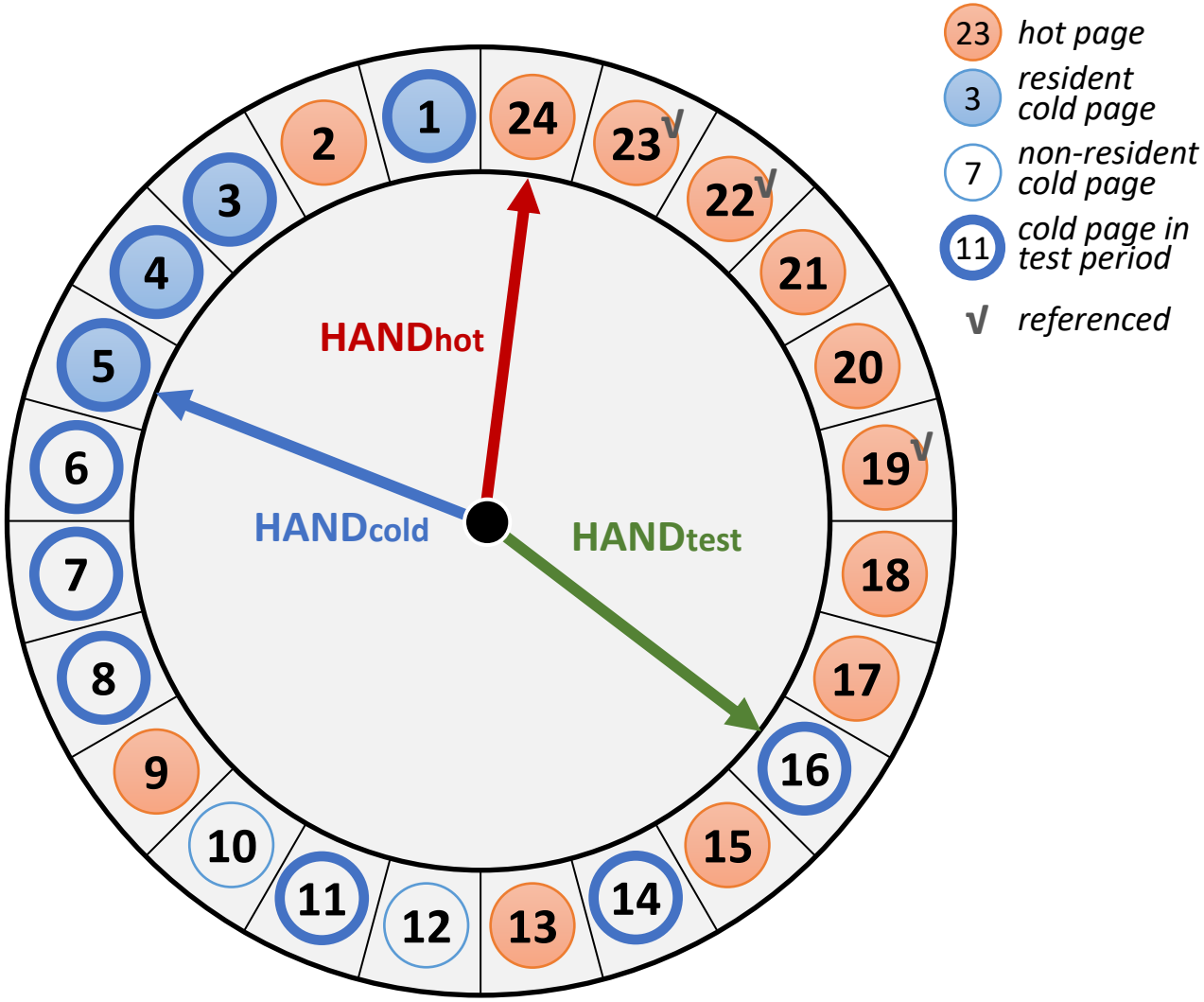
- Consider the access sequence: 23, 4, 25, 26, 7, 27, …



This example is borrowed from "Clock-Pro: An Effective Cache Replacement in OS Kernel," by Prof. Xiaodong Zhang

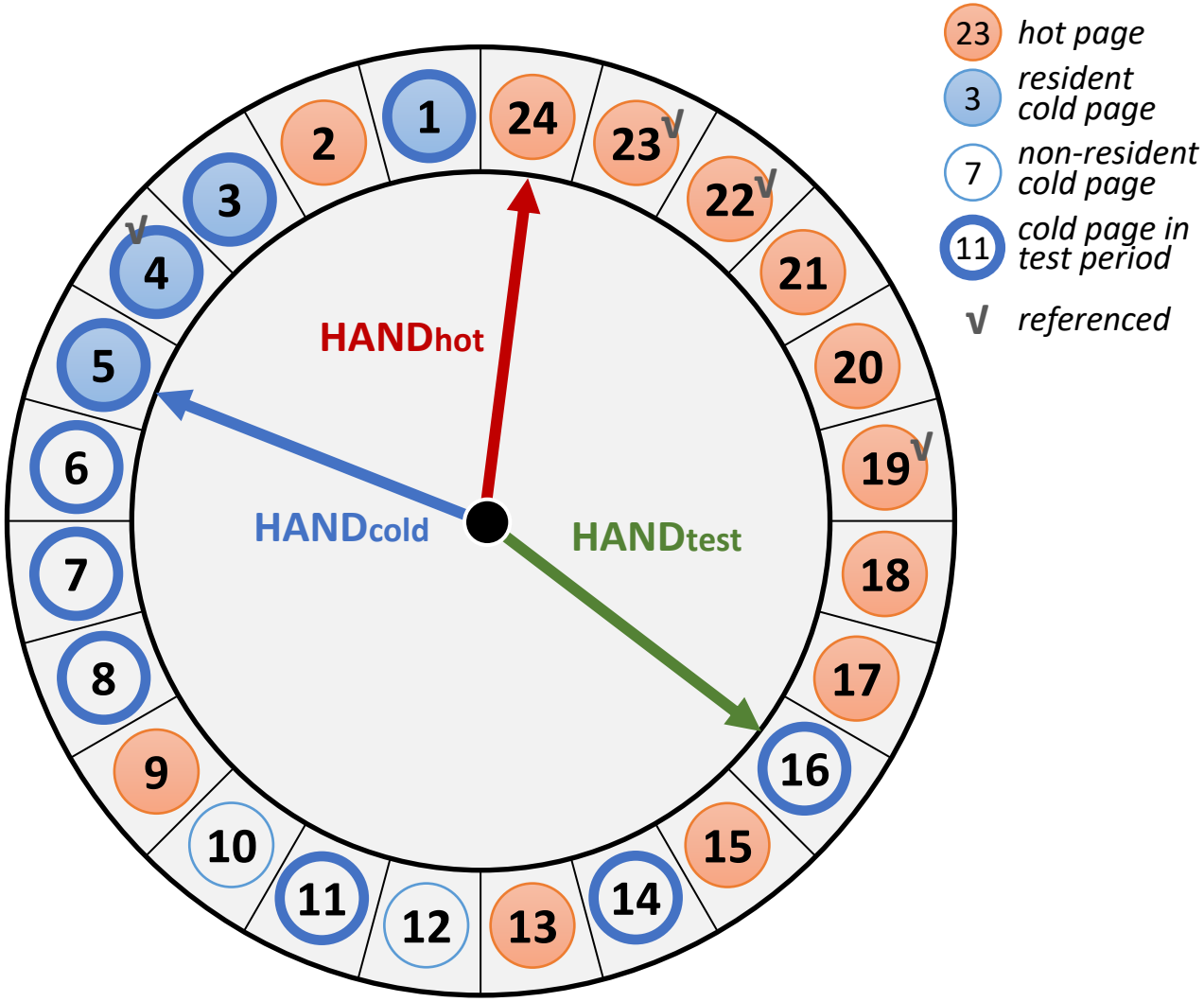
Example: Access to Page 23

- No page fault!
- Set the reference bit of page 23 to 1 (done by hardware)
- No other operation required



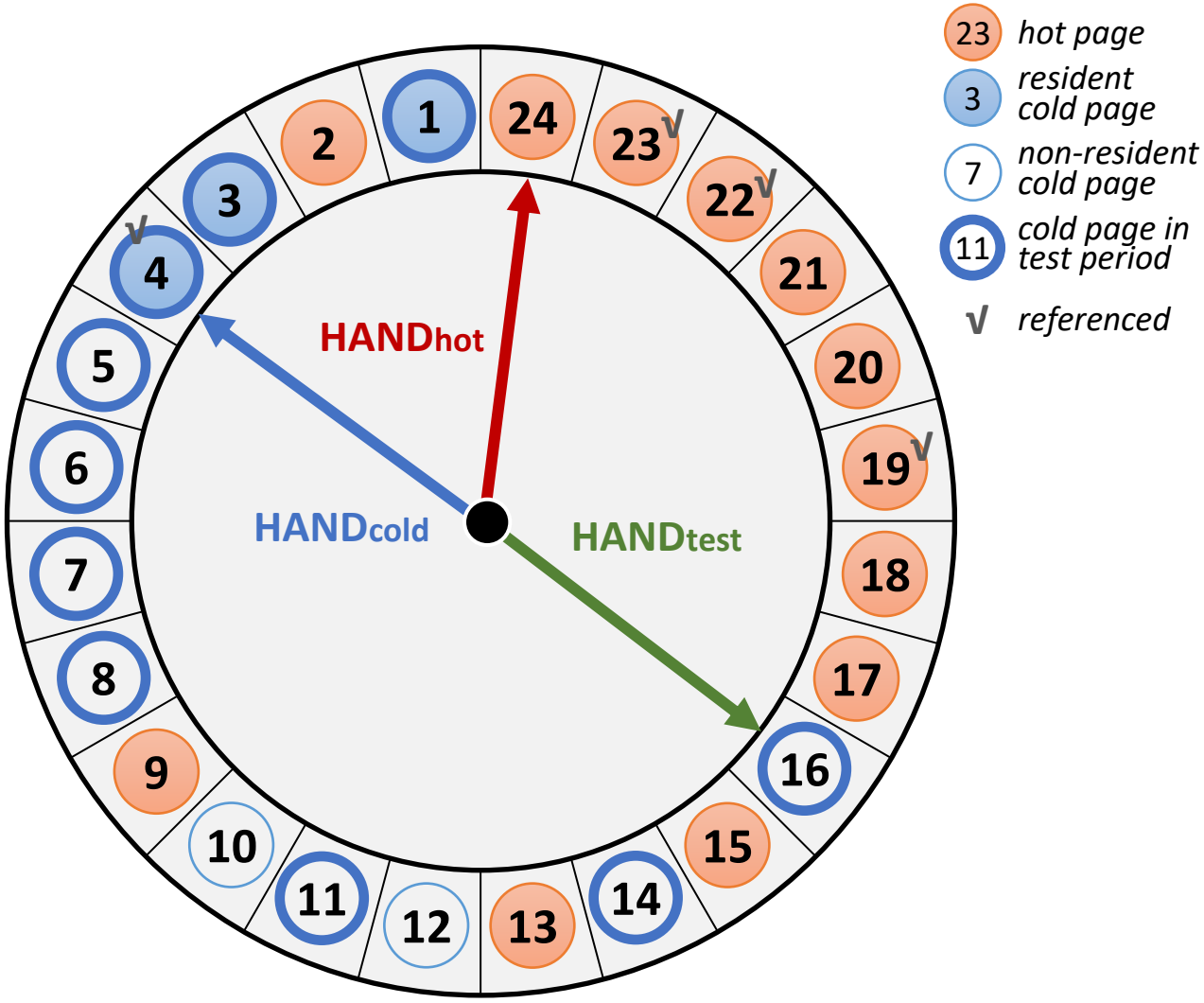
Example: Access to Page 4

- No page fault!
- Set the reference bit of page 4 to 1
- No other operation required



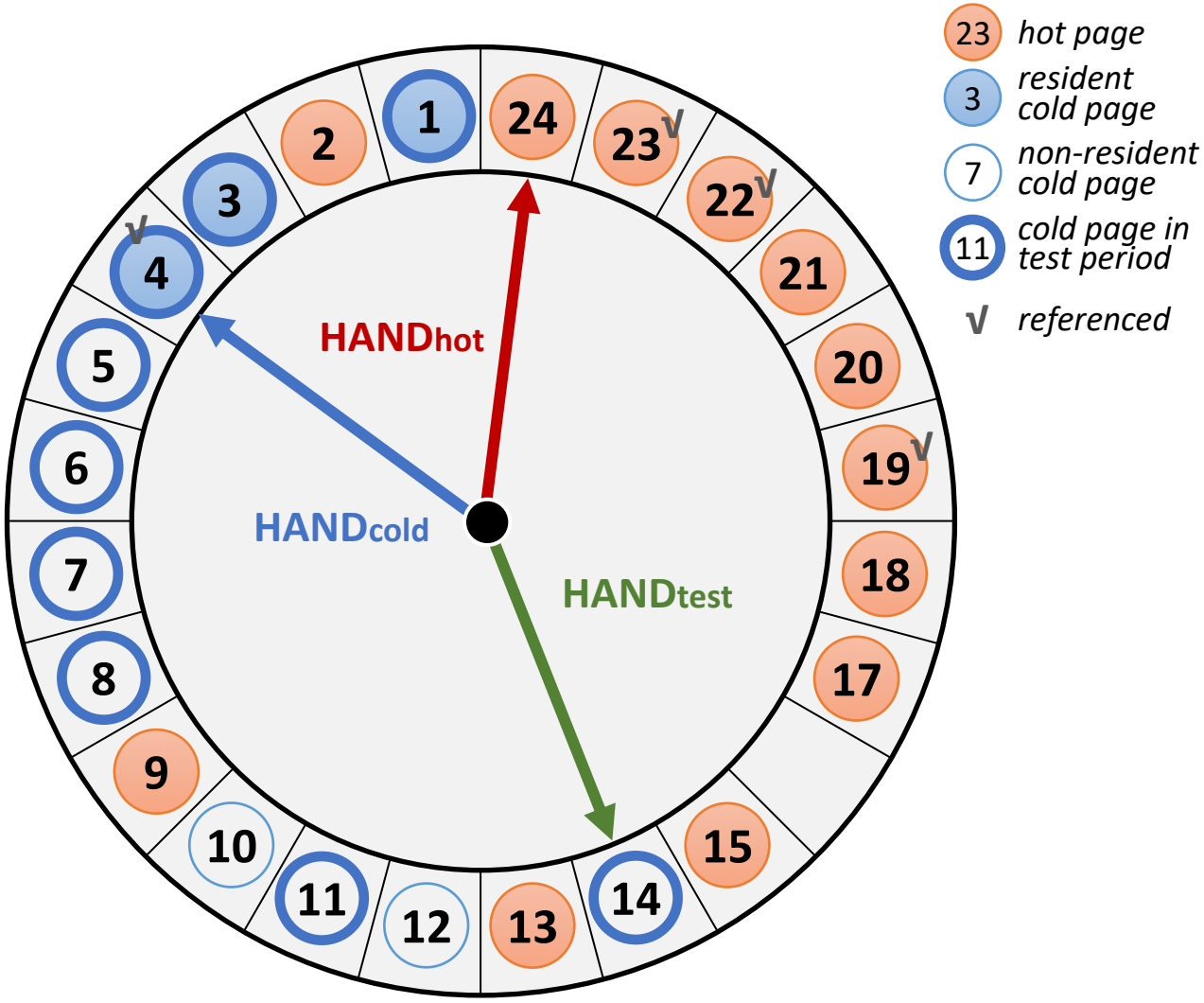
Example: Access to Page 25

- Page fault!
- Run **HAND$_{cold}$**
  - Page 5 is not referenced and in test period
  - Demote page 5 to non-resident cold page and reclaim the space
  - Move HANDcold to the next resident cold page



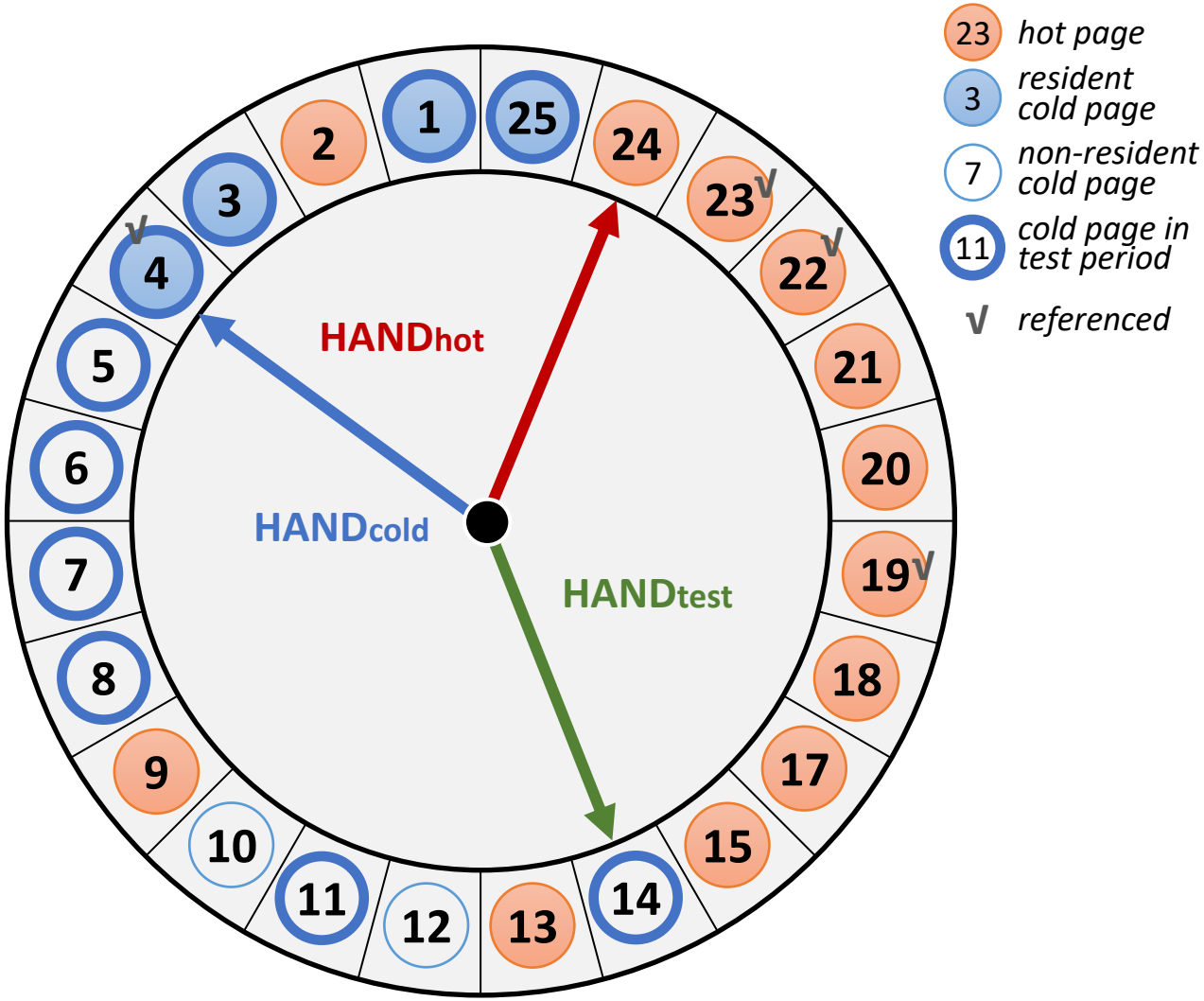
Example: Access to Page 25

- Run **HAND**$_{test}$
  - Now, # non-resident cold pages > *m*
  - Terminate the test period of page 16
  - Remove page 16 from the list
  - Move HANDtest to the next cold page 14



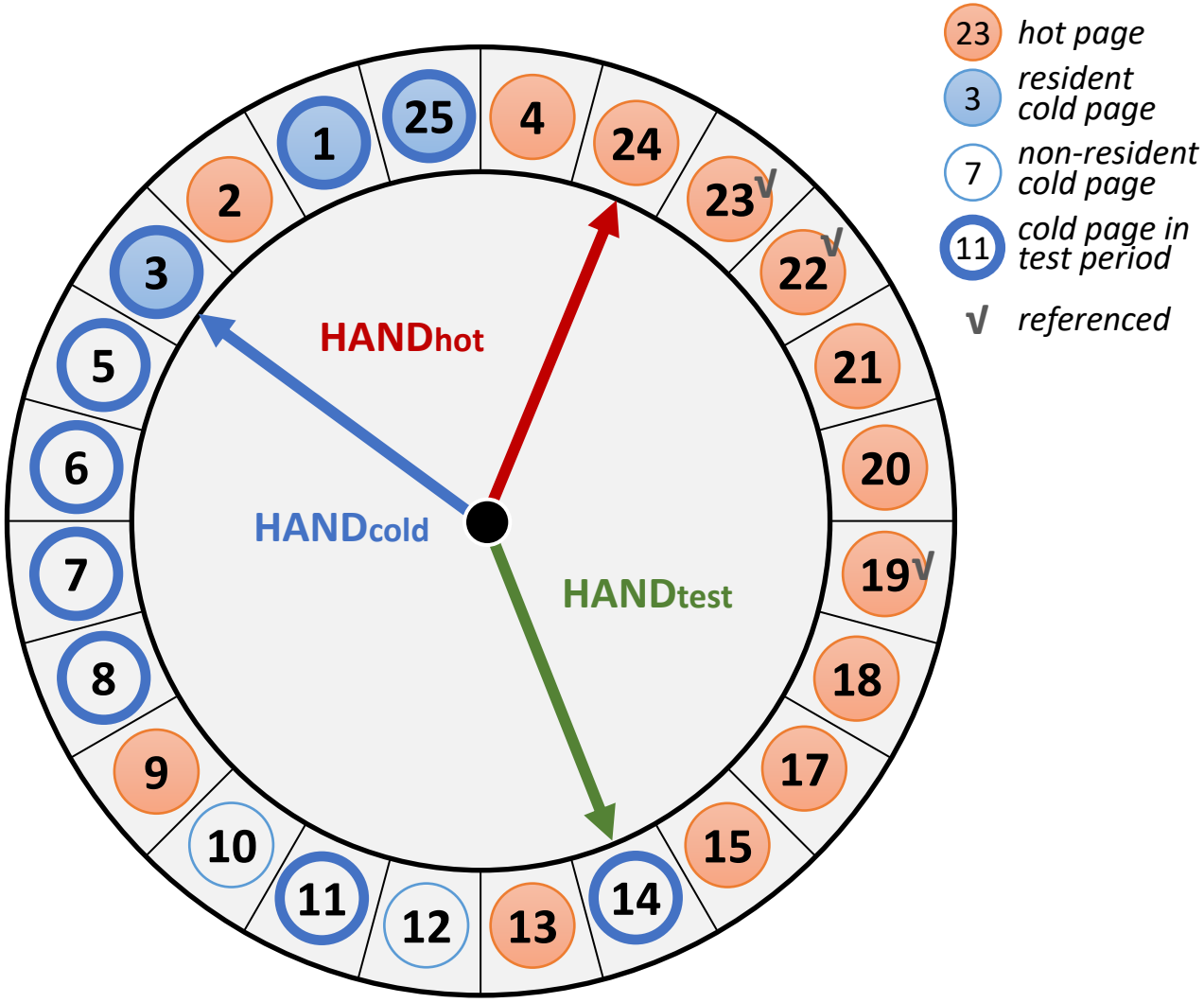
Example: Access to Page 25

- Load page 25
  - Page 25 is not a non-resident cold page
  - Set page 25 as resident cold page with test period
  - Insert page 25 at the head of the list



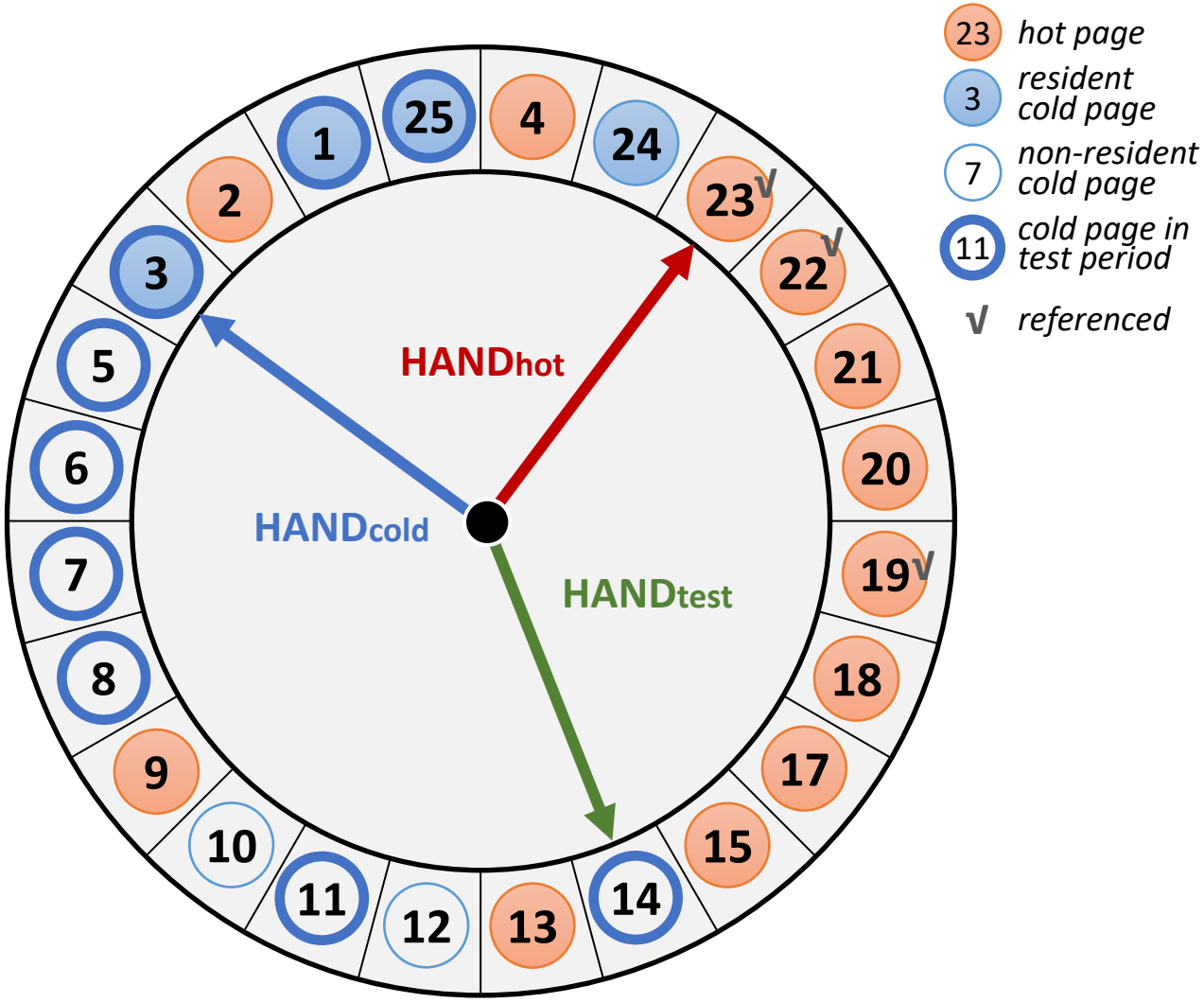
Example: Access to Page 26

- Page fault!
- Run **HAND$_{cold}$**
  - Page 4 is referenced and in test period
  - Clear the reference bit
  - Promote page 4 to hot page
  - Insert page 4 at the head of the list
  - Move HANDcold to the next resident cold page



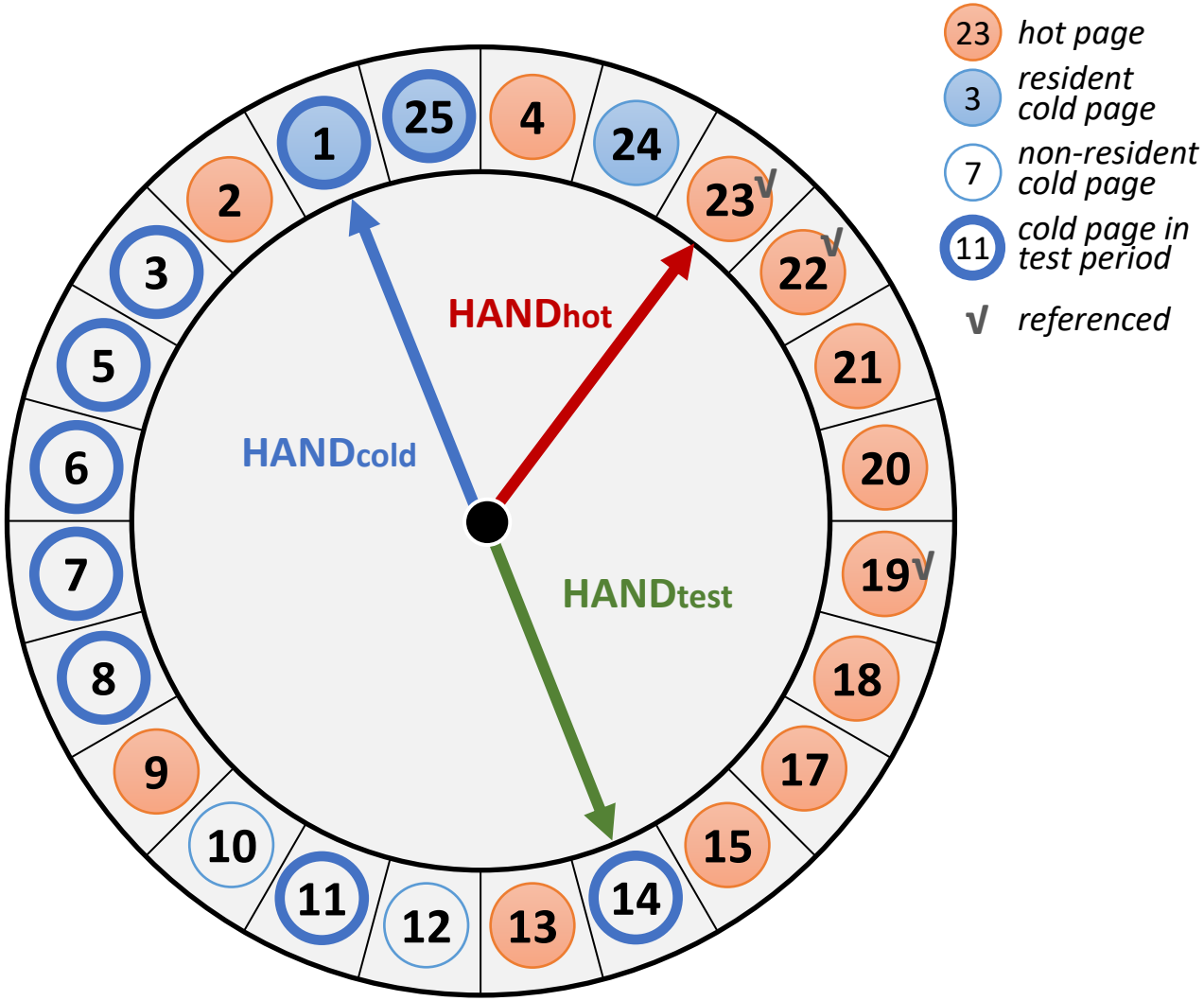
Example: Access to Page 26

- Run **HAND**$_{hot}$
  - Page 24 is a hot page with no reference
  - Demote page 24 to resident cold page
  - Move HANDhot to next page



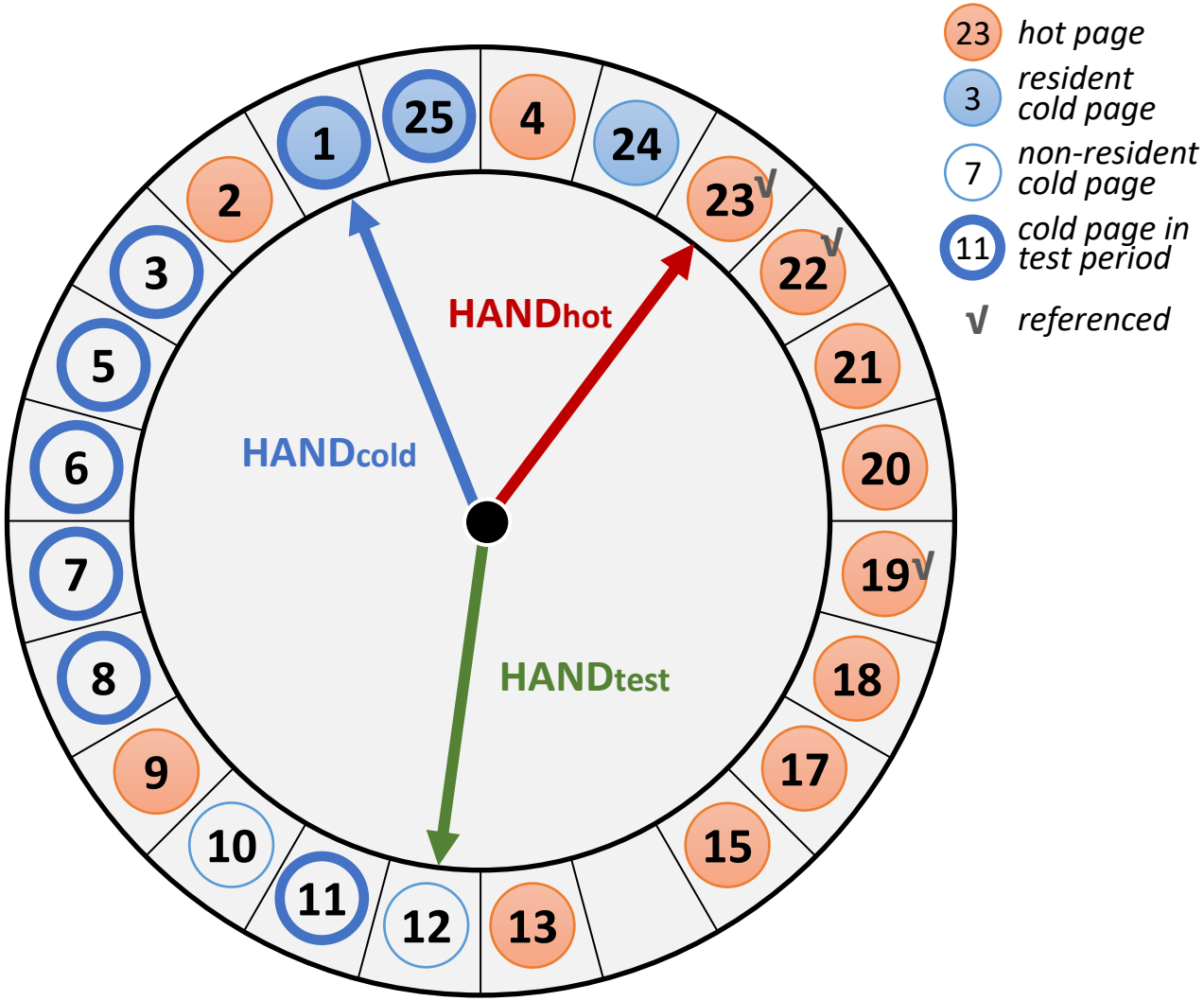
Example: Access to Page 26

- Run **HAND**$_{cold}$
  - Page 3 is not referenced and in test period
  - Demote page 3 to non-resident cold page and reclaim the space
  - Move HANDcold to the next resident cold page



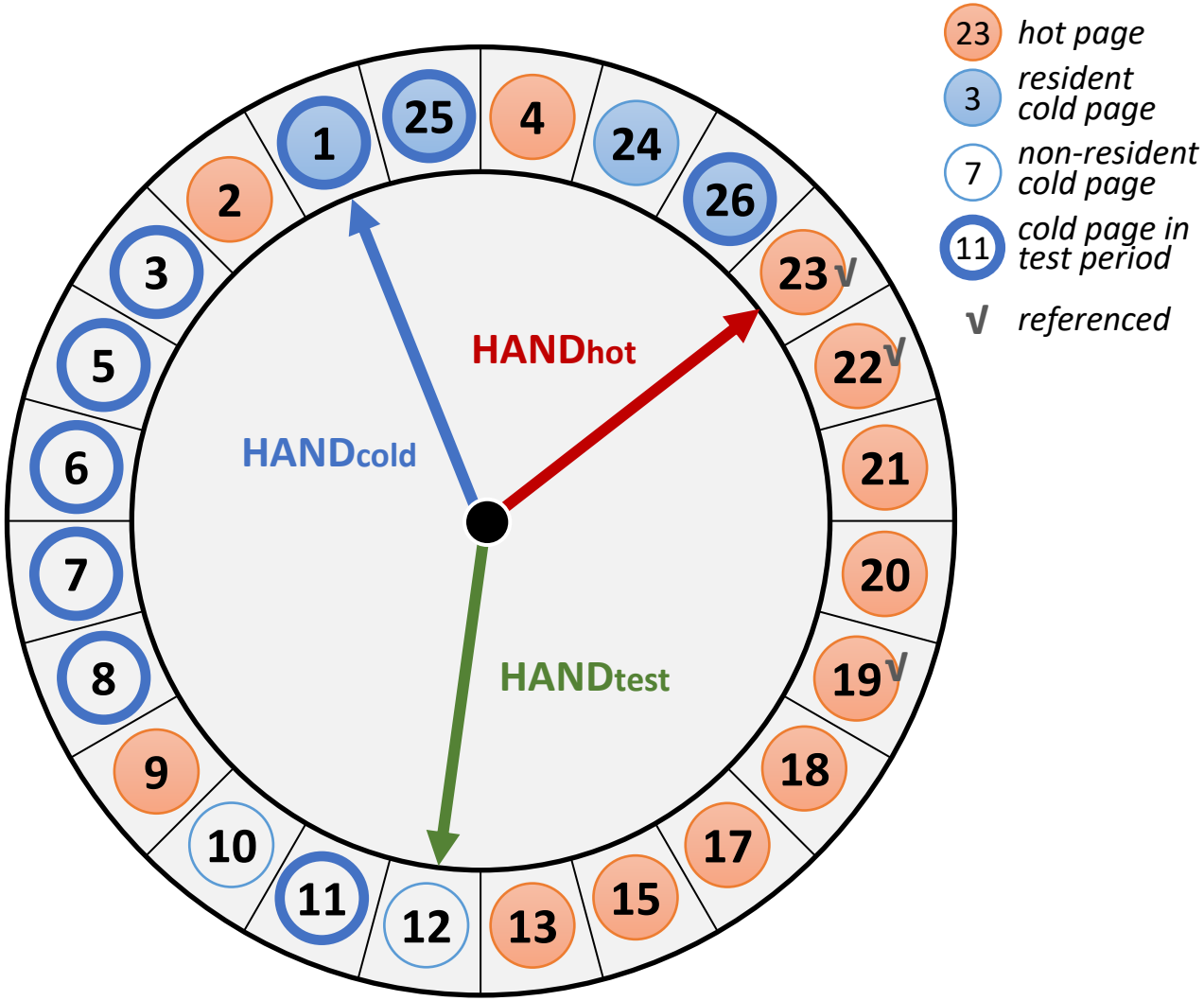
Example: Access to Page 26

- Run **HAND**$_{test}$
  - Now, # non-resident cold pages > *m*
  - Terminate the test period of page 14
  - Remove page 14 from the list
  - Move HANDtest to the next cold page 12



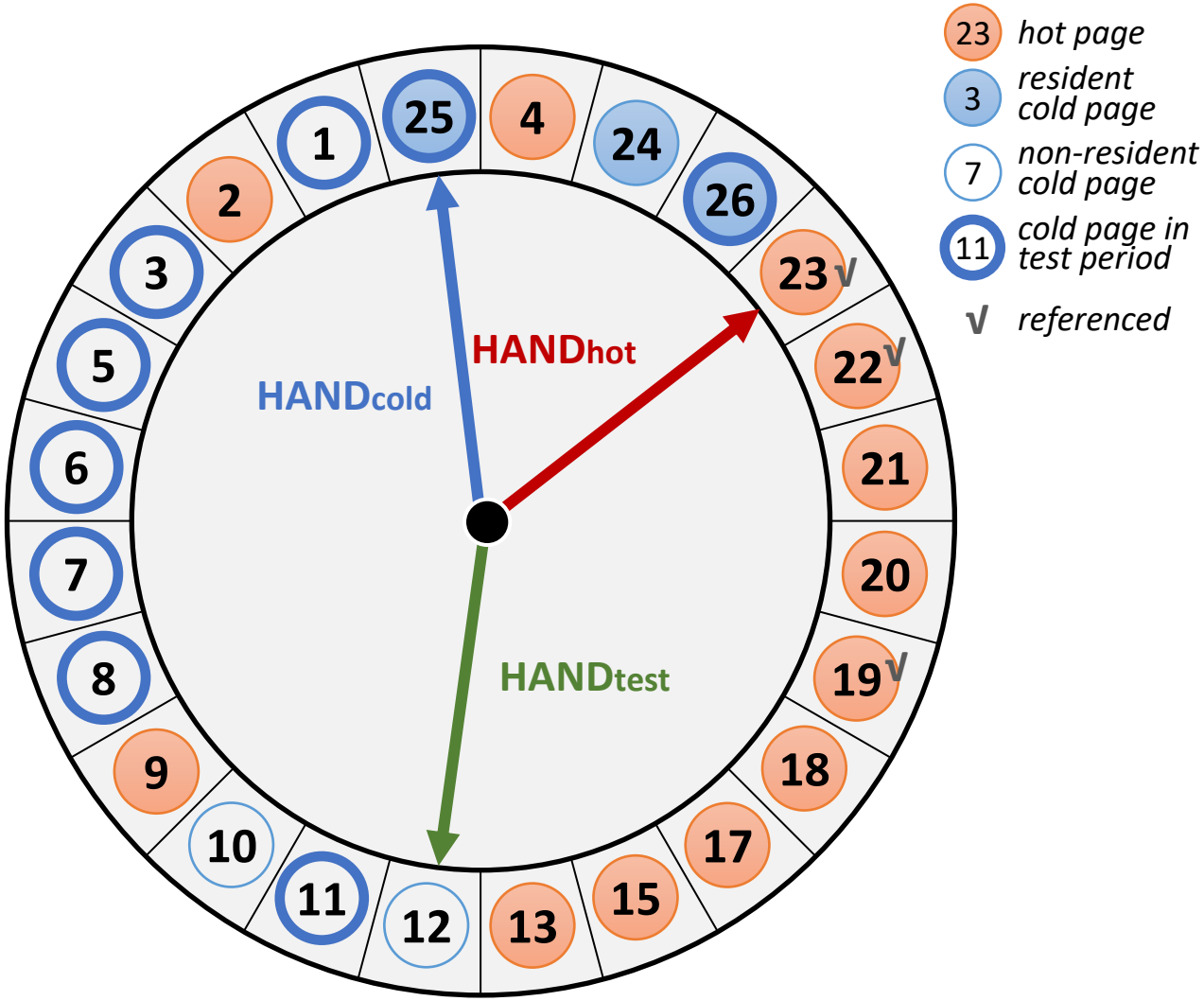
Example: Access to Page 26

- Load page 26
  - Page 26 is not a non-resident cold page
  - Set page 26 as resident cold page with test period
  - Insert page 26 at the head of the list



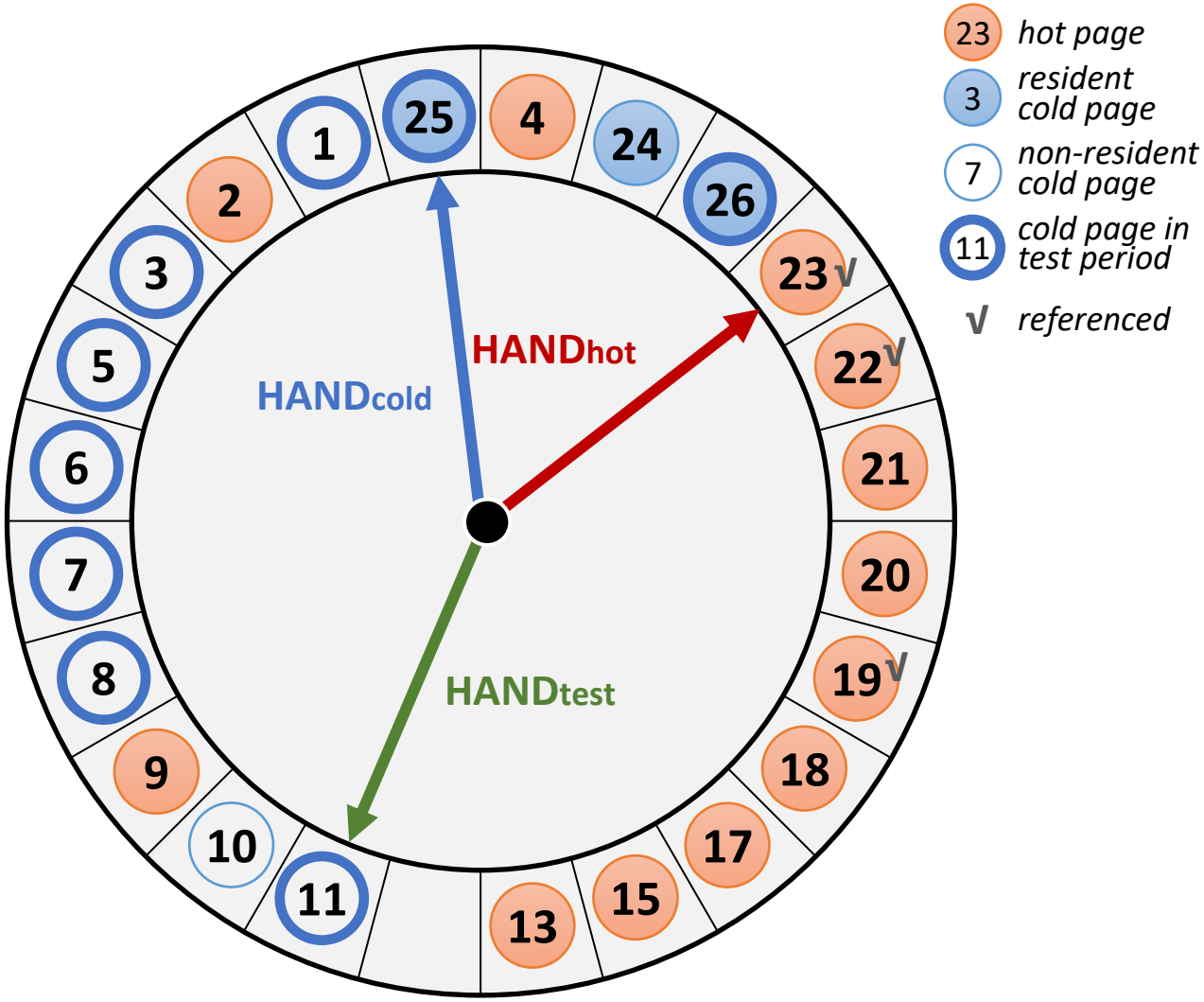
Example: Access to Page 7

- Page fault!
- Run **HANDcold**
  - Page 1 is not referenced and in test period
  - Demote page 1 to non-resident cold page and reclaim the space
  - Move HANDcold to the next resident cold page



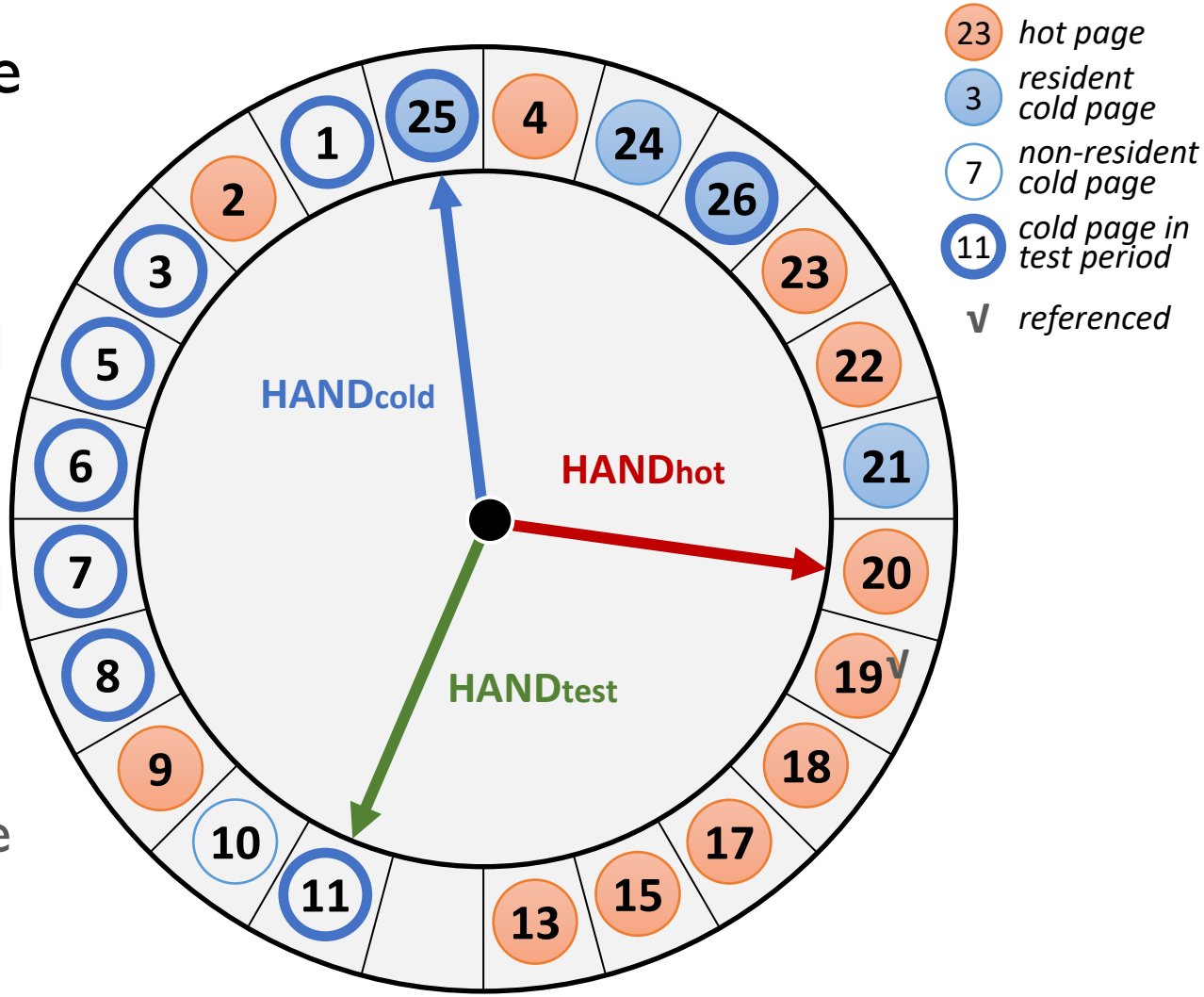
Example: Access to Page 7

- Run **HAND**$_{test}$
  - Now, # non-resident cold pages > *m*
  - Remove page 12 from the list
  - Move HANDtest to the next cold page 14



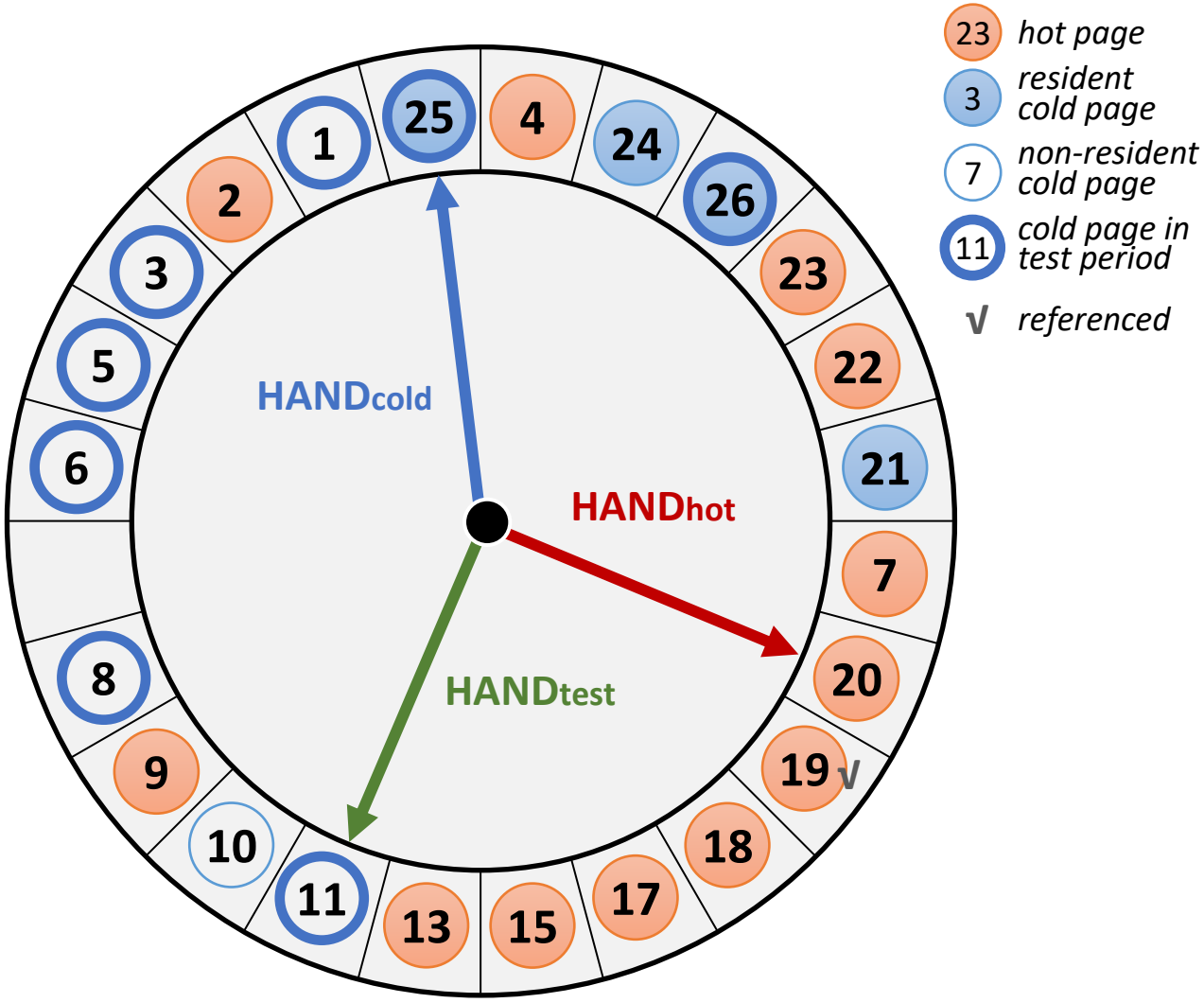
Example: Access to Page 7

- Page 7 is a non-resident cold page
- Run **HAND**$_{hot}$
  - Page 23 is referenced
  - Clear the reference bit of page 23 and move to next page
  - Page 22 is referenced
  - Clear the reference bit of page 22 and move to next page
  - Page 21 is not referenced
  - Demote page 21 to resident cold page
  - Move to next page



Example: Access to Page 7

- Load page 7
  - Promote page 7 to hot page
  - Insert page 7 at the head of the list



Observations

- A page becomes a resident cold page
  - On a first access
  - When it is demoted from a hot page
- The test period of a cold page is terminated
  - When it is promoted to the hot page
  - When it is swept by HANDhot or HANDtest
- Only a (resident) cold page is evicted
  - A hot page should be demoted to a cold page first to get evicted
- A newly inserted page should get another reference while it is in a test period to become a hot page

Adaptive CLOCK-Pro

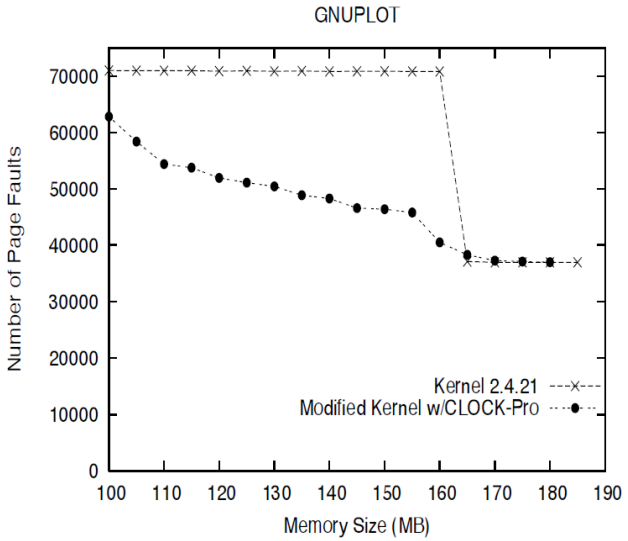
- m_c is dynamically adjusted to the current reuse distribution
- If a cold page (resident or not) is accessed during its test period, increment m_c by 1
  - For a page with a small reuse distance, retaining the page in memory for a longer period of time with a large m_c would save an additional page fault
- If a cold page (resident or not) passes its test period without a re-access, decrement m_c by 1
  - For a cold page with its reuse distance larger than its test period, retaining it in memory with a large m_c is a waste of space

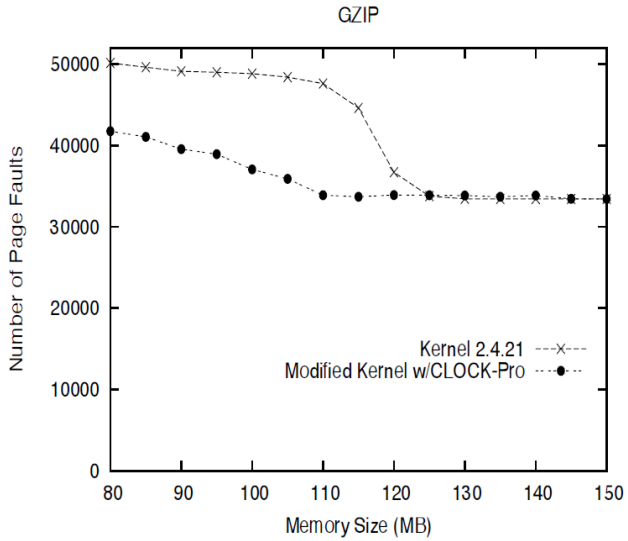
Evaluation Methodology

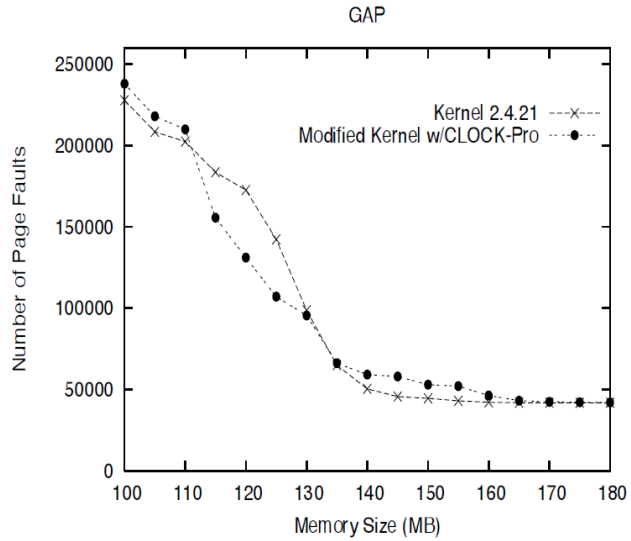
- Simulation with I/O buffer cache traces
- Simulation with memory access traces
- Simulation with memory access + I/O traces
- Evaluation on Linux
  - Clock-Pro implemented in Linux 2.4.21
  - All pages are placed in a single clock list with three hands
  - Compare the modified kernel with the original

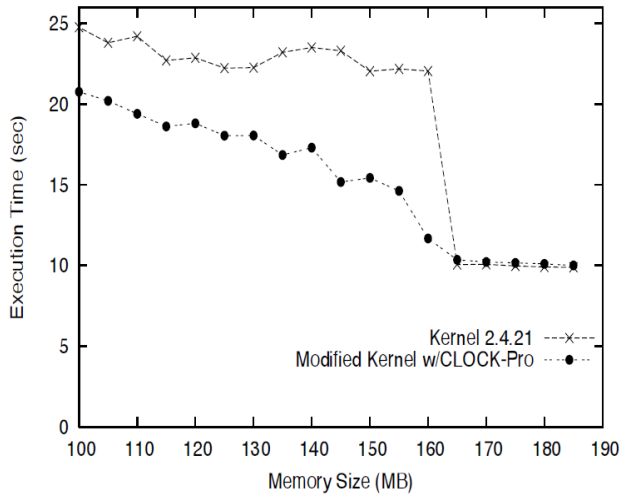
- SPEC CPU2000 and memory intensive software tools are used as benchmarks

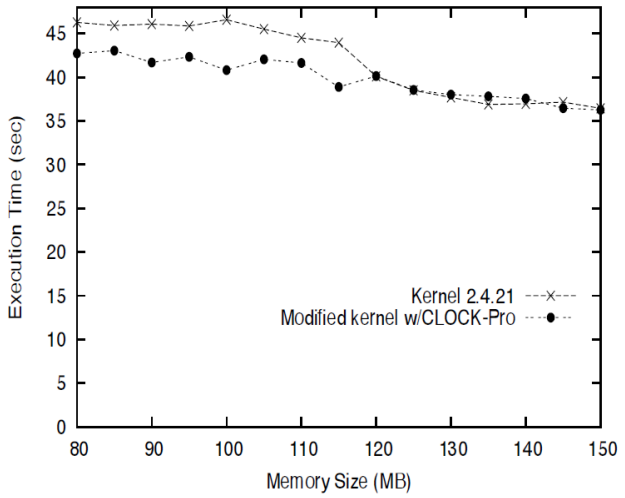
Results on Linux

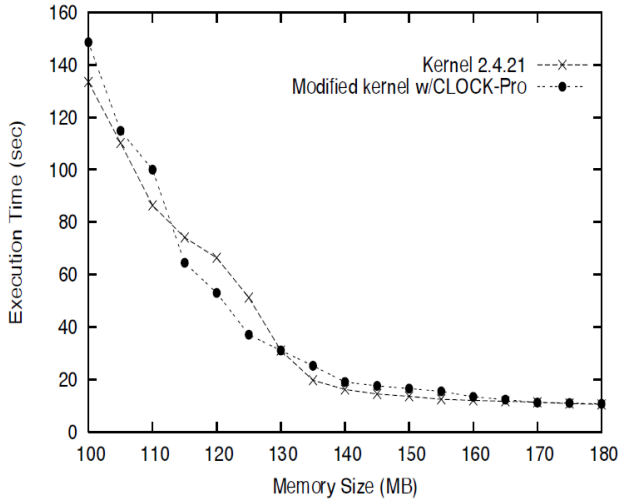












Summary

- A variation of LIRS for VM page replacement
- Based on reuse distance
- A cold page is granted a test period

- Officially adopted in NetBSD
- Also affected Linux page replacement design
- A patch is available for OpenLDAP