

# Cache Related Preemption Delay for Set-Associative Caches

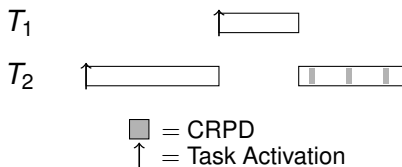
## Pitfalls and Solutions

Claire Burguière, Jan Reineke, Sebastian Altmeyer

Workshop on WCET Analysis, Dublin 2009



- Preemptive scheduling
- Cache related preemption delay (CRPD):
  - ▶ Impact of preemption on the cache content
  - ▶ Overall cost of additional reloads due to preemption



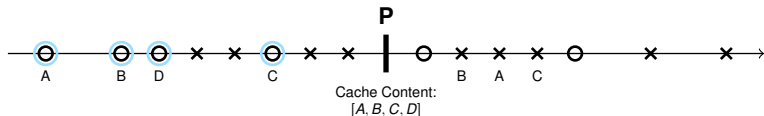
- CRPD computation:
  - ▶ Useful Cache Blocks (UCB)
  - ▶ Evicting Cache Blocks (ECB)
- CRPD for set-associative caches:
  - ⇒ **Pitfalls:**
    - ▶ LRU: CRPD not bounded by the number of ECBs
    - ▶ FIFO and PLRU: CRPD not bounded

## Definition (Useful Cache Block)

A memory block  $m$  at program point  $P$  is called a useful cache block, if

- $m$  may be cached at  $P$
- $m$  may be reused at program point  $P'$  that may be reached from  $P$  with no eviction of  $m$  on this path.

x = hit  
O = miss

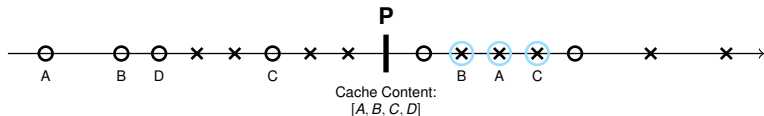


## Definition (Useful Cache Block)

A memory block  $m$  at program point  $P$  is called a useful cache block, if

- $m$  may be cached at  $P$
- $m$  may be reused at program point  $P'$  that may be reached from  $P$  with no eviction of  $m$  on this path.

x = hit  
O = miss

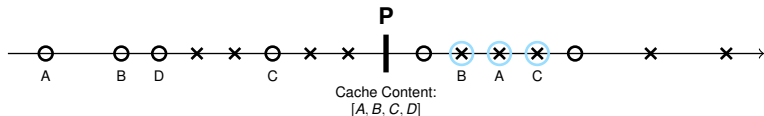


## Definition (Useful Cache Block)

A memory block  $m$  at program point  $P$  is called a useful cache block, if

- $m$  may be cached at  $P$
- $m$  may be reused at program point  $P'$  that may be reached from  $P$  with no eviction of  $m$  on this path.

x = hit  
O = miss



$$\text{CRPD}^{\text{LRU}} = \sum_{s=1}^c \text{CRPD}^{\text{LRU}}(s)$$

$$\text{CRPD}_{\text{UCB}}^{\text{LRU}}(s) = \text{BRT} \cdot \min(|\text{UCB}(s)|, n)$$

$n$  = associativity

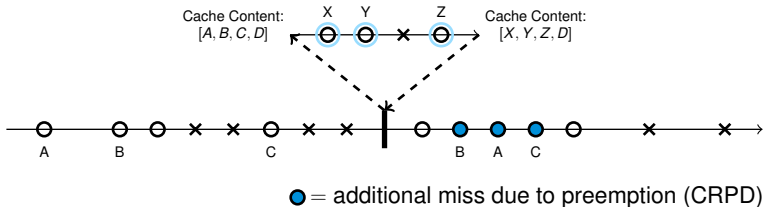
BRT = Block Reload Time

# Evicting Cache Blocks

## [Tomiyaama & Dutt, 2000]

### Definition (Evicting Cache Blocks (ECB))

*A memory block of the preempting task is called an evicting cache block, if it may be accessed during the execution of the preempting task.*

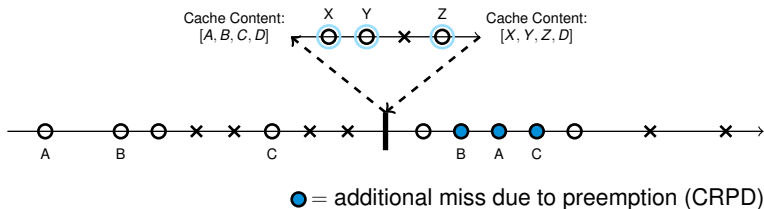


# Evicting Cache Blocks

## [Tomiya & Dutt, 2000]

### Definition (Evicting Cache Blocks (ECB))

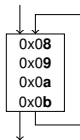
*A memory block of the preempting task is called an evicting cache block, if it may be accessed during the execution of the preempting task.*



$$\text{CRPD}_{\text{ECB}}^{\text{LRU}}(s) \stackrel{?}{=} \text{BRT} \cdot \min(|\text{ECB}(s)|, n)$$

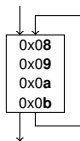


# CRPD computation for LRU using ECB: Pitfall



$[b, a, 9, 8] \xrightarrow{8} [8, b, a, 9] \xrightarrow{9} [9, 8, b, a] \xrightarrow{a} [a, 9, 8, b] \xrightarrow{b} [b, a, 9, 8]$  0 misses

# CRPD computation for LRU using ECB: Pitfall



$$\begin{array}{l} \text{ECBs} \\ = \{e\} \end{array} \left( \begin{array}{l} [b, a, 9, 8] \xrightarrow{8} [8, b, a, 9] \xrightarrow{9} [9, 8, b, a] \xrightarrow{a} [a, 9, 8, b] \xrightarrow{b} [b, a, 9, 8] \quad 0 \text{ misses} \\ [e, b, a, 9] \xrightarrow{8^*} [8, e, b, a] \xrightarrow{9^*} [9, 8, e, b] \xrightarrow{a^*} [a, 9, 8, e] \xrightarrow{b^*} [b, a, 9, 8] \quad 4 \text{ misses} \end{array} \right.$$

- $|\text{UCB}(s)| = 4$
- $|\text{ECB}(s)| = 1$
- $n = 4$
- number of additional misses = 4

- ECB derivation used only to know if the set is used by the preempting task:

$$\text{CRPD}_{\text{ECB}}^{\text{LRU}}(s) = \begin{cases} 0 & \text{if } \text{ECB}(s) = \emptyset \\ \text{BRT} \cdot n & \text{otherwise} \end{cases}$$

$$[b, a] \xrightarrow{a} [b, a] \xrightarrow{e^*} [e, b] \xrightarrow{b} [e, b] \xrightarrow{c^*} [c, e] \xrightarrow{e} [c, e] \quad 2 \text{ misses}$$

$$\begin{array}{l} \text{ECBs} \\ = \{\mathbf{x}\} \end{array} \left( \begin{array}{l} [b, a] \xrightarrow{a} [b, a] \xrightarrow{e^*} [e, b] \xrightarrow{b} [e, b] \xrightarrow{c^*} [c, e] \xrightarrow{e} [c, e] \quad 2 \text{ misses} \\ [x, b] \xrightarrow{a^*} [a, x] \xrightarrow{e^*} [e, a] \xrightarrow{b^*} [b, e] \xrightarrow{c^*} [c, b] \xrightarrow{e^*} [e, c] \quad 5 \text{ misses} \end{array} \right)$$

- $|\text{UCB}(s)| = 2$
- $|\text{ECB}(s)| = 1$
- $n = 2$
- **But: number of additional misses = 3**

## Notation

$m_P(p, s)$  = *number of misses that policy  $P$  incurs on access sequence  $s \in M^*$  starting in state  $p$*

## Definition (Relative miss competitiveness)

Policy  $P$  is  $(k, c)$ -miss-competitive relative to policy  $Q$ , if

$$m_P(p, s) \leq k \cdot m_Q(q, s) + c$$

for all access sequences  $s \in M^*$  and compatible cache-set states  $p, q$ .

## Notation

$m_P(p, s)$  = *number of misses that policy  $P$  incurs on access sequence  $s \in M^*$  starting in state  $p$*

## Definition (Relative miss competitiveness)

Policy  $P$  is  $(k, c)$ -miss-competitive relative to policy  $Q$ , if

$$m_P(p, s) \leq k \cdot m_Q(q, s) + c$$

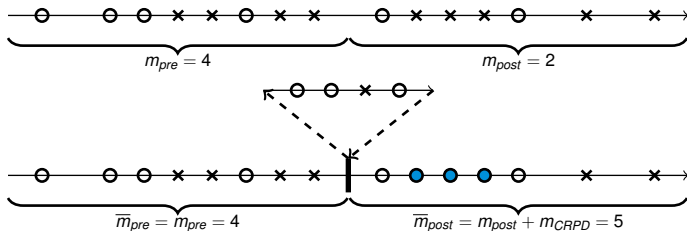
for all access sequences  $s \in M^*$  and compatible cache-set states  $p, q$ .

- PLRU( $n$ ) is  $(1, 0)$ -miss-competitive relative to LRU( $1 + \log_2 n$ ).
- FIFO( $n$ ) is  $(\frac{n}{n-r+1}, r)$ -miss-competitive relative to LRU( $r$ ).

# A sequence of memory accesses

## ■ Notation:

- ▶  $m$  = number of misses
- ▶  $\bar{m}$  = number of misses in the case of preemption

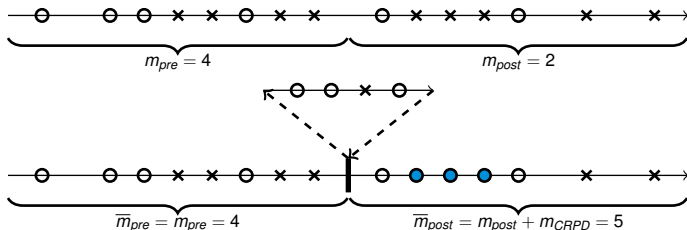




# A sequence of memory accesses

## ■ Notation:

- ▶  $m$  = number of misses
- ▶  $\bar{m}$  = number of misses in the case of preemption



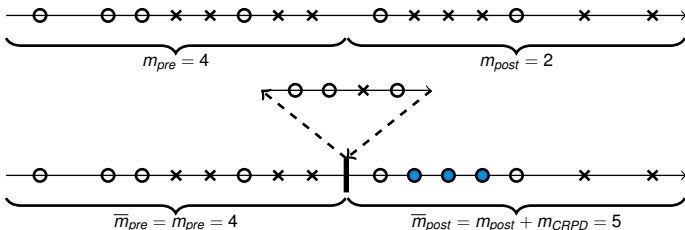
## ■ Relative miss competitiveness:

$$\bar{m}^{P(t)} = \bar{m}_{pre}^{P(t)} + \bar{m}_{post}^{P(t)}$$

# A sequence of memory accesses

## ■ Notation:

- ▶  $m$  = number of misses
- ▶  $\bar{m}$  = number of misses in the case of preemption



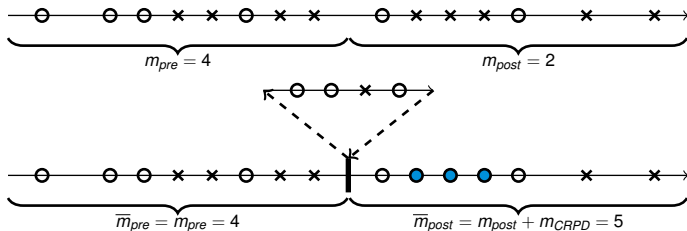
## ■ Relative miss competitiveness:

$$\begin{aligned} \bar{m}^{P(t)} &= \bar{m}_{pre}^{P(t)} + \bar{m}_{post}^{P(t)} \\ &\leq [k \cdot m_{pre}^{LRU(s)} + c] + [k \cdot (m_{post}^{LRU(s)} + m_{CRPD}^{LRU(s)}) + c] \end{aligned}$$

# A sequence of memory accesses

## ■ Notation:

- ▶  $m$  = number of misses
- ▶  $\bar{m}$  = number of misses in the case of preemption



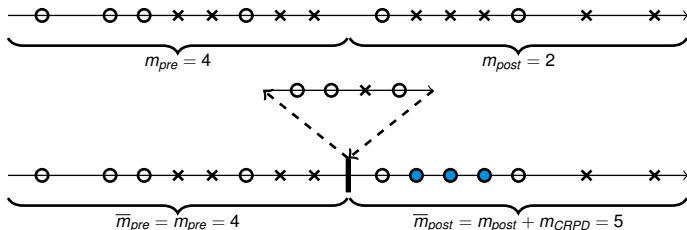
## ■ Relative miss competitiveness:

$$\begin{aligned}
 \bar{m}^P(t) &= \bar{m}_{pre}^P(t) + \bar{m}_{post}^P(t) \\
 &\leq [k \cdot m_{pre}^{LRU(s)} + c] + [k \cdot (m_{post}^{LRU(s)} + m_{CRPD}^{LRU(s)}) + c] \\
 &= [k \cdot (m_{pre}^{LRU(s)} + m_{post}^{LRU(s)}) + c] + [k \cdot m_{CRPD}^{LRU(s)} + c]
 \end{aligned}$$

# A sequence of memory accesses

## ■ Notation:

- ▶  $m$  = number of misses
- ▶  $\bar{m}$  = number of misses in the case of preemption



## ■ Relative miss competitiveness:

$$\begin{aligned}
 \bar{m}^{P(t)} &= \bar{m}_{pre}^{P(t)} + \bar{m}_{post}^{P(t)} \\
 &\leq [k \cdot m_{pre}^{LRU(s)} + c] + [k \cdot (m_{post}^{LRU(s)} + m_{CRPD}^{LRU(s)}) + c] \\
 &= [k \cdot (m_{pre}^{LRU(s)} + m_{post}^{LRU(s)}) + c] + [k \cdot m_{CRPD}^{LRU(s)} + c] \\
 &= [k \cdot m^{LRU(s)} + c] + [k \cdot m_{CRPD}^{LRU(s)} + c]
 \end{aligned}$$

- PLRU(8) using LRU(4):

$$\bar{m}^{\text{PLRU}(8)} \leq m^{\text{LRU}(4)} + m_{\text{CRPD}}^{\text{LRU}(4)}$$

- FIFO(8) using LRU(5):








$$\bar{m}^{\text{FIFO}(8)} \leq (2 \cdot m^{\text{LRU}(5)} + 5) + (2 \cdot m_{\text{CRPD}}^{\text{LRU}(5)} + 5)$$

## Pitfalls

- *LRU:*  
 *$|ECBs|$  is not an upper-bound*
- *FIFO and PLRU:*  
 *$|UCBs|$ ,  $|ECBs|$  and  $n$  do not bound the number of additional misses*

## Solutions

- *LRU:*  
 *$|ECBs| = 0$   
 $\Rightarrow$  the set is not used*
- *FIFO and PLRU:*  
*using relative competitiveness and LRU bounds*

-  [Altmeyer, S. & Burguière, C. \(2009\).](#)  
In ECRTS'09.
-  [Lee, C.-G., Hahn, J., Min, S. L., Ha, R., Hong, S., Park, C. Y., Lee, M. & Kim, C. S. \(1996\).](#)  
In RTSS'96 p. 264, IEEE Computer Society.
-  [Negi, H. S., Mitra, T. & Roychoudhury, A. \(2003\).](#)  
In CODES+ISSS'03 ACM.
-  [Reineke, J. \(2008\).](#)  
Caches in WCET Analysis.  
PhD thesis, Universität des Saarlandes, Saarbrücken.
-  [Staschulat, J. & Ernst, R. \(2007\).](#)  
ACM Trans. on Embedded Computing Sys. 6, 25.
-  [Tan, Y. & Mooney, V. \(2004\).](#)  
In SCOPES'04 pp. 182–199,.
-  [Tomiyama, H. & Dutt, N. D. \(2000\).](#)  
In CODES'00 ACM.

- using UCB [Lee et al., 1996]:

$$\text{CRPD}_{\text{UCB}} = \text{BRT} \cdot |\{s_i \mid \exists m \in \text{UCB} : m \bmod c = s_i\}|$$

---

- using ECB [Tomiyaama & Dutt, 2000]:

$$\text{CRPD}_{\text{ECB}} = \text{BRT} \cdot |\{s_i \mid \exists m \in \text{ECB} : m \bmod c = s_i\}|$$

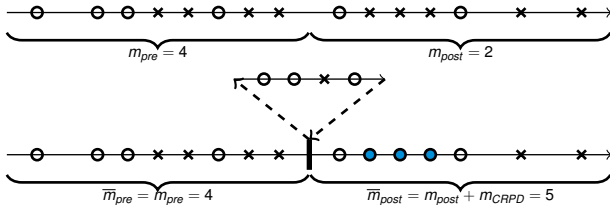
---

- using UCB and ECB [Negi et al., 2003, Tan & Mooney, 2004]:

$$\text{CRPD}_{\text{UCB\&ECB}} = \text{BRT} \cdot |\{s_i \mid \exists m \in \text{UCB} : m \bmod c = s_i \\ \wedge \exists m' \in \text{ECB} : m' \bmod c = s_i\}|$$



# Relative Competitiveness – Example



$$\begin{aligned} \bar{m}_{pre}^{P(t)} &\leq k \cdot \bar{m}_{pre}^{LRU(s)} + c \\ &= k \cdot m_{pre}^{LRU(s)} + c \end{aligned}$$

$$\begin{aligned} \bar{m}_{post}^{P(t)} &\leq k \cdot \bar{m}_{post}^{LRU(s)} + c \\ &= k \cdot (m_{post}^{LRU(s)} + m_{CRPD}^{LRU(s)}) + c \end{aligned}$$

$$\begin{aligned} \bar{m}^{P(t)} &= \bar{m}_{pre}^{P(t)} + \bar{m}_{post}^{P(t)} \\ &\leq k \cdot m_{pre}^{LRU(s)} + c + k \cdot (m_{post}^{LRU(s)} + m_{CRPD}^{LRU(s)}) + c \\ &= (k \cdot (m_{pre}^{LRU(s)} + m_{post}^{LRU(s)}) + c) + (k \cdot m_{CRPD}^{LRU(s)} + c) \\ &= (k \cdot m^{LRU(s)} + c) + (k \cdot m_{CRPD}^{LRU(s)} + c). \end{aligned}$$

## ■ PLRU( $n$ ) using LRU( $1 + \log_2 n$ ):

- ▶  $k = 1, c = 0$

$$\bar{m}^{\text{PLRU}(n)} \leq m^{\text{LRU}(1+\log_2 n)} + m_{\text{CRPD}}^{\text{LRU}(1+\log_2 n)}$$

- ▶ Example ( $n=8$ ):

$$\bar{m}^{\text{PLRU}(8)} \leq m^{\text{LRU}(4)} + m_{\text{CRPD}}^{\text{LRU}(4)}$$

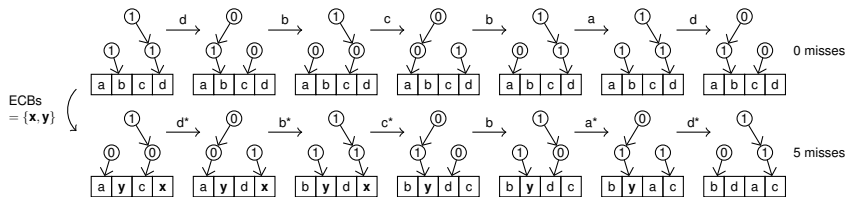
## ■ FIFO( $n$ ) using LRU( $r$ ):

- ▶  $k = \frac{n}{n-r+1}, c = r$

$$\bar{m}^{\text{FIFO}(n)} \leq \left( \frac{n}{n-r+1} \cdot m^{\text{LRU}(r)} + r \right) + \left( \frac{n}{n-r+1} \cdot m_{\text{CRPD}}^{\text{LRU}(r)} + r \right)$$

- ▶ Example ( $n=8, r=5$ ):

$$\bar{m}^{\text{FIFO}(8)} \leq (2 \cdot m^{\text{LRU}(5)} + 5) + (2 \cdot m_{\text{CRPD}}^{\text{LRU}(5)} + 5)$$



- $|\text{UCB}(s)| = 4$
- $|\text{ECB}(s)| = 2$
- $n = 4$
- But: number of additional misses = 5