



# Temporal Protection in Real-Time Systems

Software Engineering Institute  
Carnegie Mellon University  
Pittsburgh, PA 15213

November 2016



Copyright 2016 Carnegie Mellon University

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

**NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.**

[Distribution Statement A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

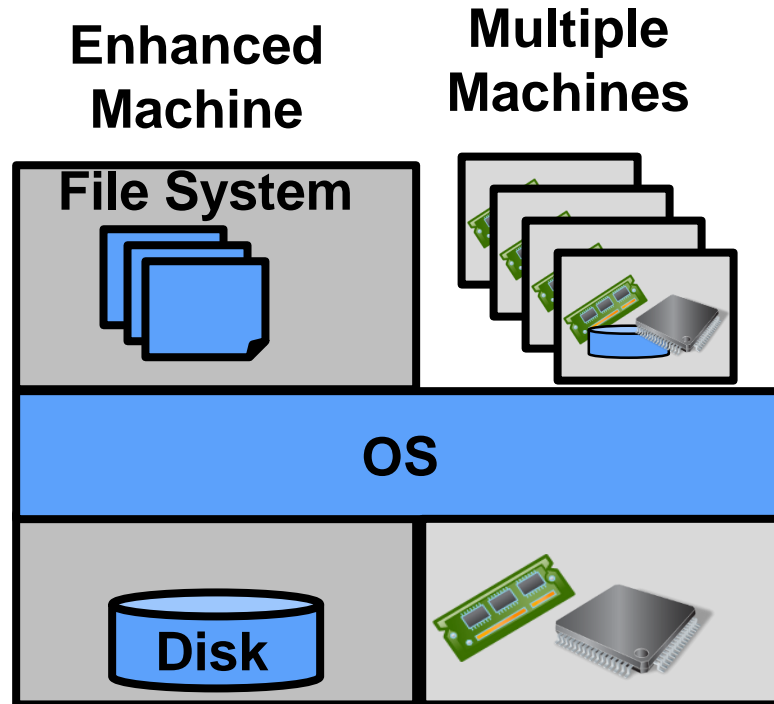
This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at [permission@sei.cmu.edu](mailto:permission@sei.cmu.edu).

Carnegie Mellon® is registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM-0004174



# OS Dual Objective



Icons credit: <http://www.doublejdesign.co.uk>



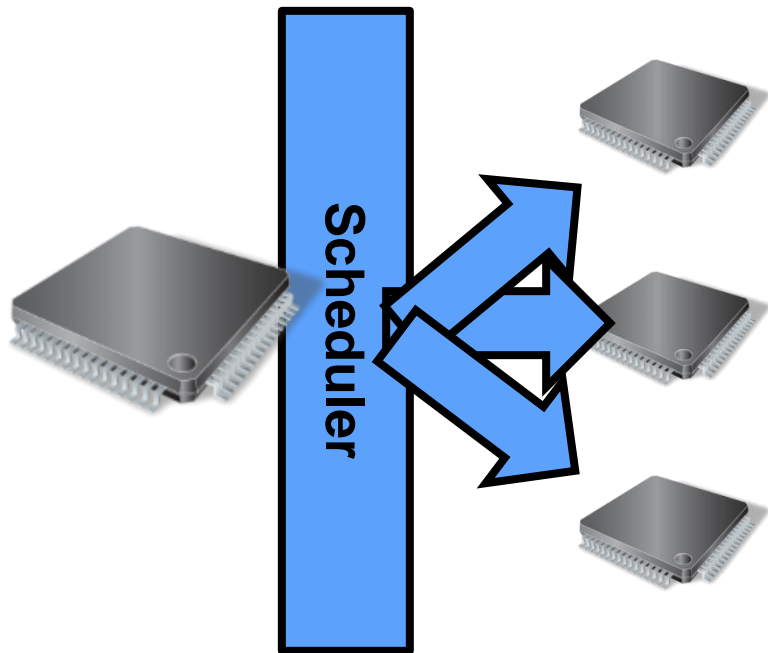
**Software Engineering Institute**

**Carnegie Mellon**

Temporal Protection RT Systems  
de Niz, November 2016

© 2016 Carnegie Mellon University

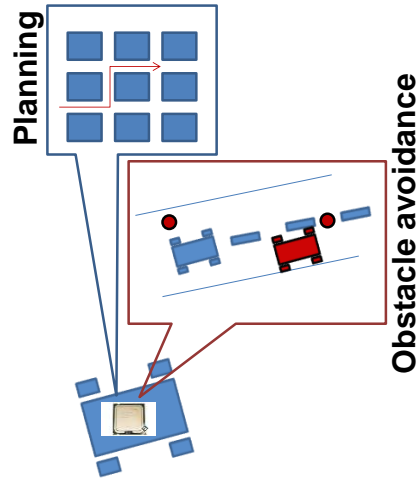
# Time-Sharing CPU – Round robin



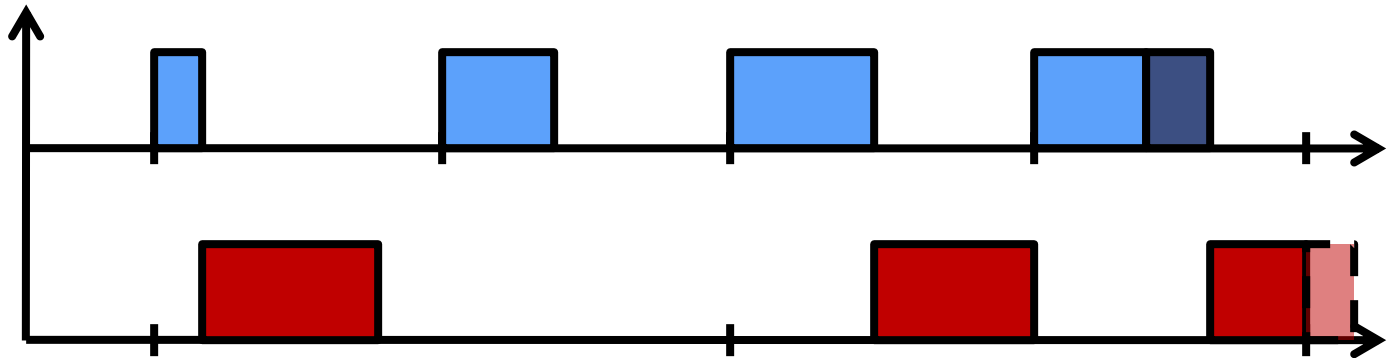
**Same time requirement – Fair Scheduling**



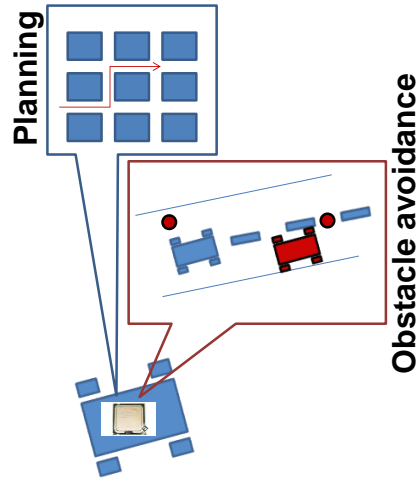
# Consolidation of Mixed-Criticality Tasks



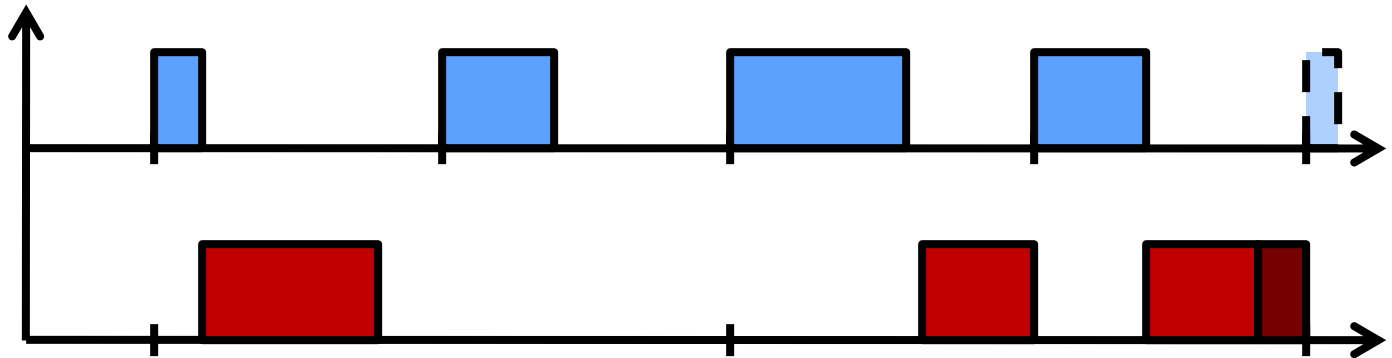
Shared Hardware  
Can lead to cycle stealing



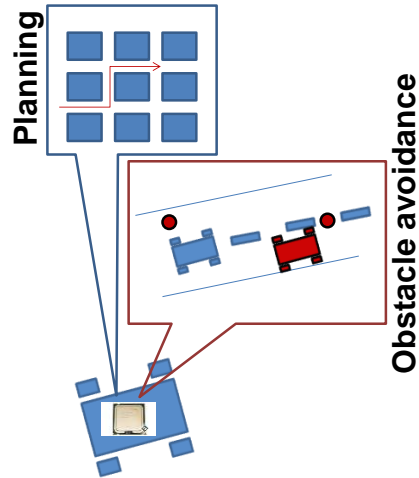
# Consolidation of Mixed-Criticality Tasks



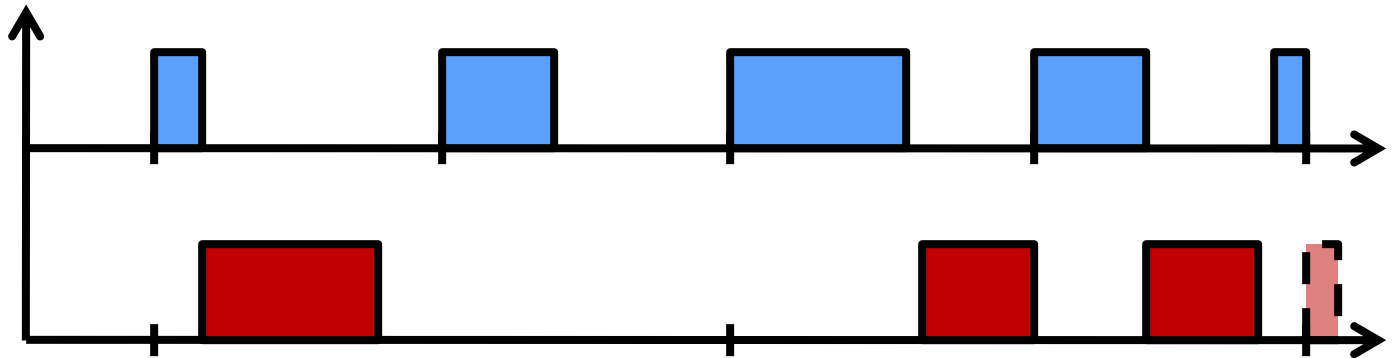
To avoid interference  
add temporal protection



# Consolidation of Mixed-Criticality Tasks



BUT  
Symmetric protection  
leads to *criticality inversion*



# Criticality Inversion

A higher-criticality task waits for a lower-criticality task to release a resource

- Symmetric temporal protection
- Scheduling policy is aimed at maximizing utilization (RMS/DMS/EDF)



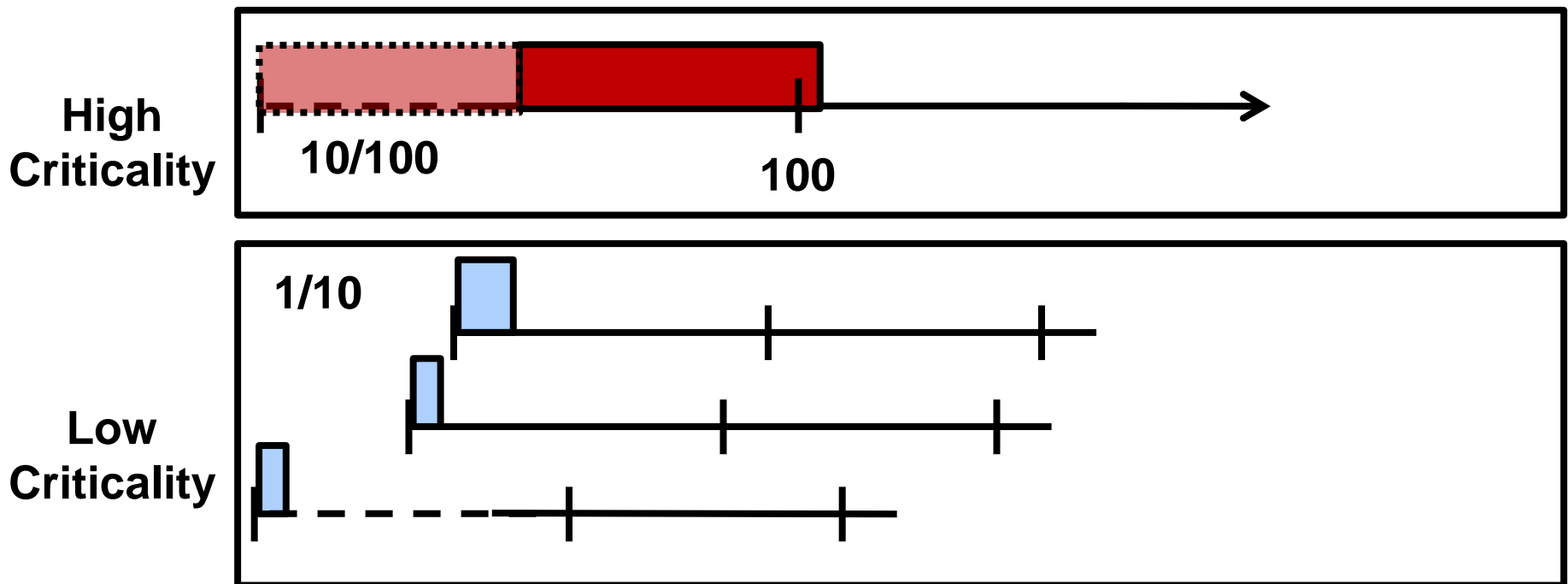
# Rate-Monotonic Priority

Shorter Period  $\rightarrow$  Higher Priority

- Ideal utilization

BUT: Poor Criticality Protection Due to **Criticality Inversion**

- If criticality order is opposite to rate-monotonic priority order



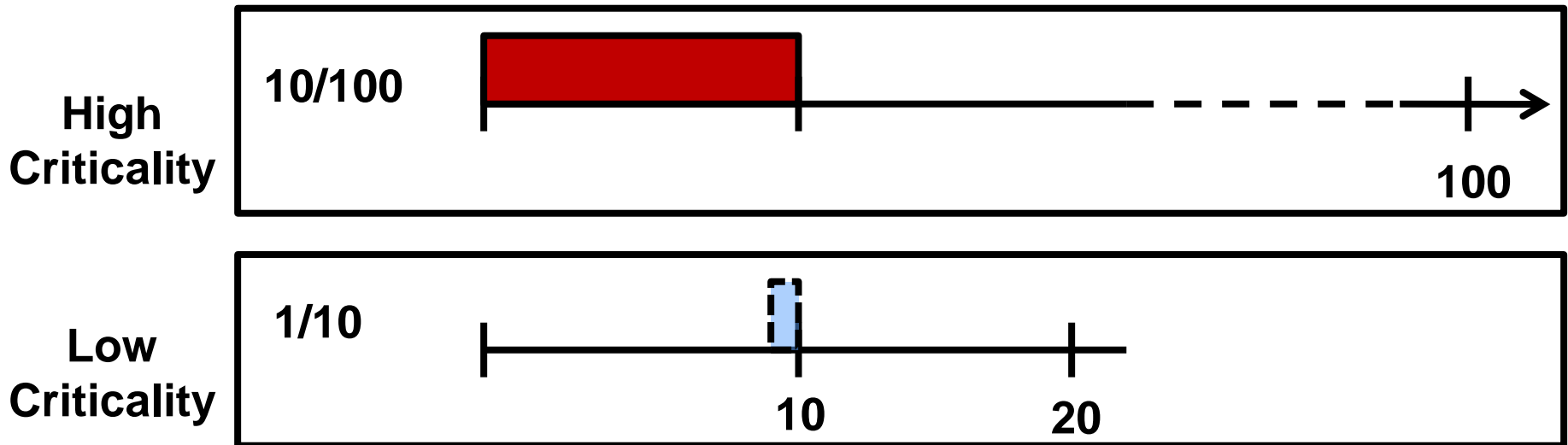
# Criticality As Priority Assignment (CAPA)

Higher Criticality → Higher Priority

- Ideal criticality protection:
  - lower criticality cannot interfere with higher criticality

BUT: Poor Utilization Due to **Priority Inversion**

- If criticality order is opposite to rate-monotonic priority order



# Task Model

$$\tau_i = (C_i, C_i^o, T_i, D_i, \zeta_i)$$

$C_i$  Normal Execution Budget of task  $i$

$C_i^o$  Overload Execution Budget of task  $i$

$T_i$  Period of task  $i$

$D_i$  Deadline of task  $i$   $D_i \leq T_i$

$\zeta_i$  Criticality of task  $i$



# Zero-Slack Scheduling

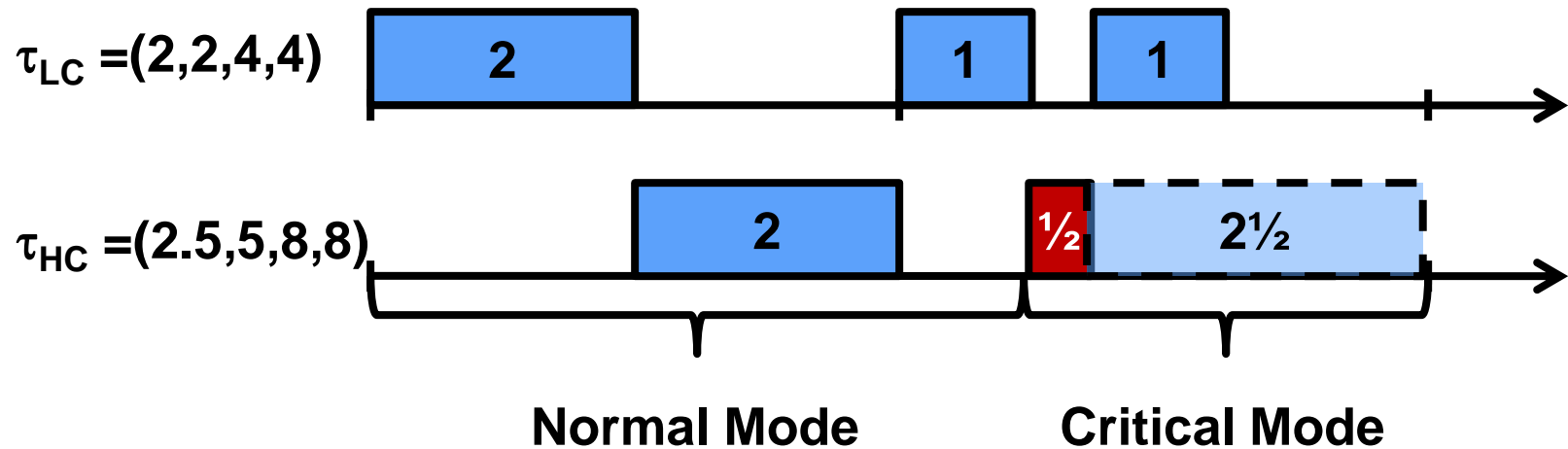
Start with RM

Calculate the last instant before  $\tau_{HC}$  misses its deadline

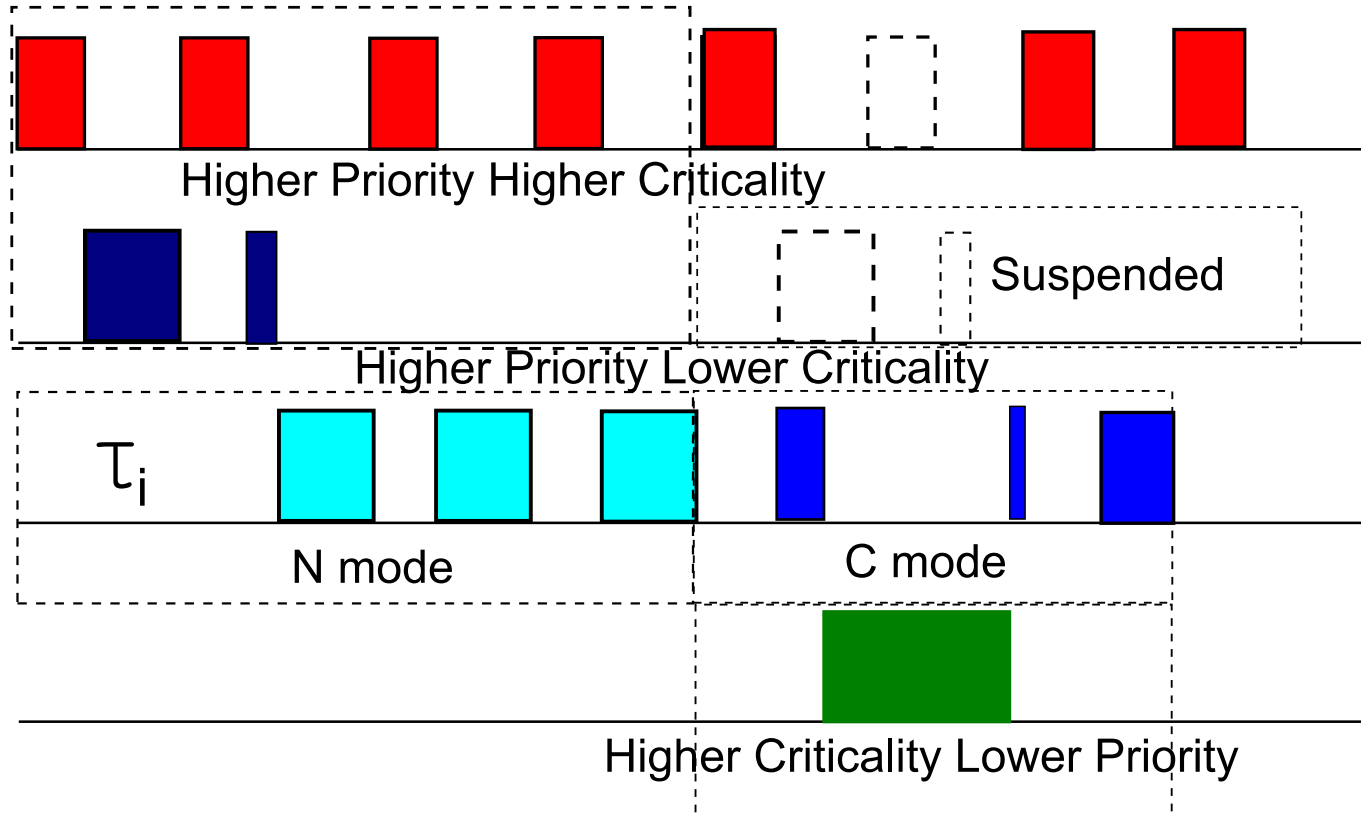
- this is called the **zero-slack** instant

Switch to criticality-as-priority

- Splits the execution window into
  - Normal mode (RM)
  - Critical mode (CAPA)



# Critical Instant of a Task $\tau_i$



# Interference in Zero-Slack Scheduling

Task Set Divided into

- $H^{lc}$  : Higher priority, lower criticality
- $H^{hc}$  : Higher priority, higher criticality
- $L^{lc}$  : Lower priority, lower criticality
- $L^{hc}$  : Lower priority, higher criticality

Interfering tasks in normal mode (Normal mode)

- $H^{lc} + H^{hc} + L^{hc}$

Interfering tasks in critical mode (C mode)

- $H^{hc} + L^{hc}$



# Scheduling Guarantee

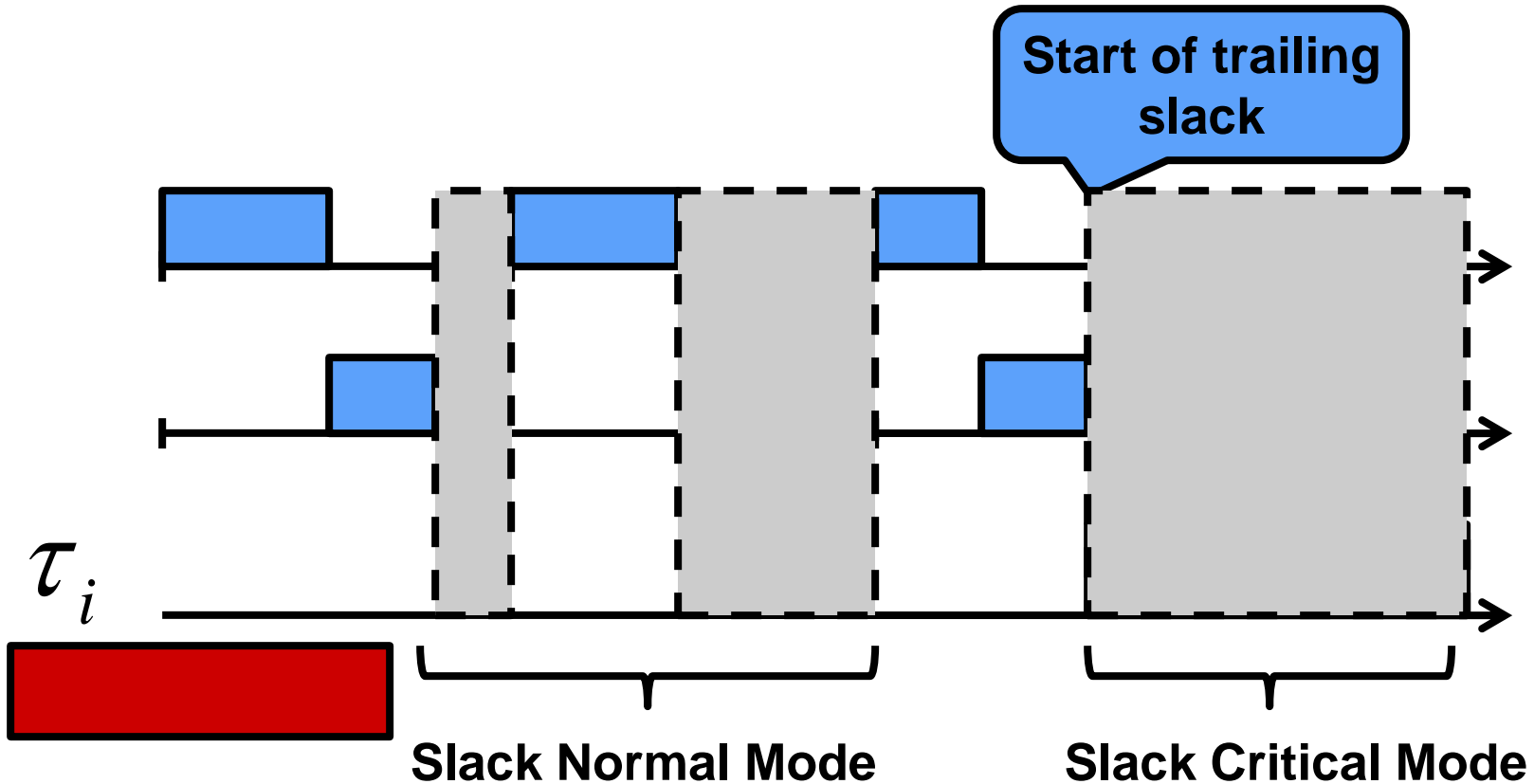
A task  $\tau_i$  is guaranteed  $C_i^0$  before  $D_i$

if no  $\tau_j \mid \zeta_j < \zeta_i$

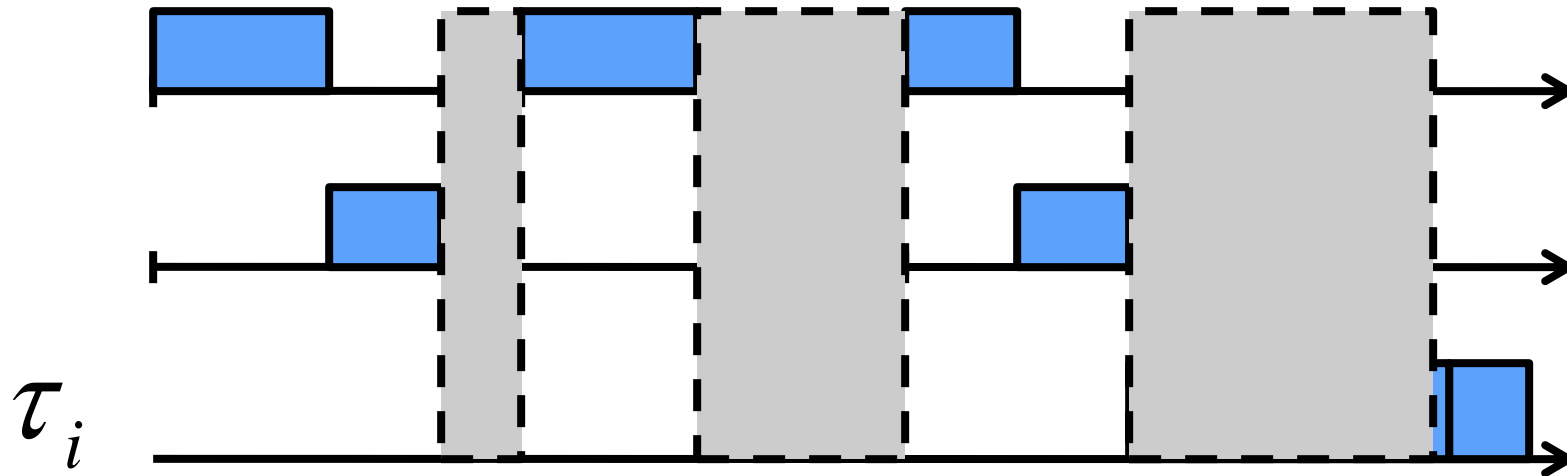
executes beyond its  $C_j$



# Calculating The Zero-Slack Instant



# Calculating The Zero-Slack Instant



New slack can open after each iteration

Needs to repeat until no new slack opens



# ZSRM Properties

## Subsumes RM

- If criticalities are aligned to priorities
- No critical mode

## Subsumes CAPA

- If not enough slack, only critical mode

## Graceful Degradation

- In overloads, deadlines are missed in reverse criticality order



# Implementation

## ZSRM

Scheduling algorithm calculates zero-slack instants offline

## Linux/ RK

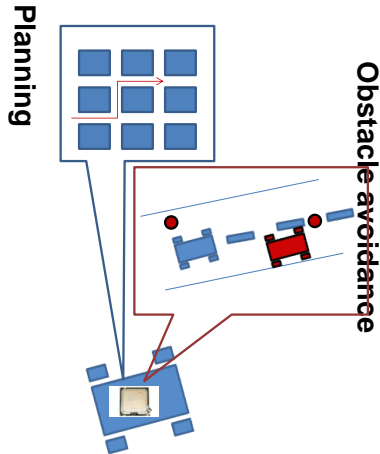
- Resource reservation in Linux
  - CPU, Net, Mem, Disk
- Bundled into resource sets that provide a form of virtual machine
- Multiple implementations
  - Nano/RK for sensor networks

## Special Zero-Slack Reserves

- Switch to critical mode
  - Stop lower-criticality tasks on zero-slack instant
- Tasks in critical mode in stack



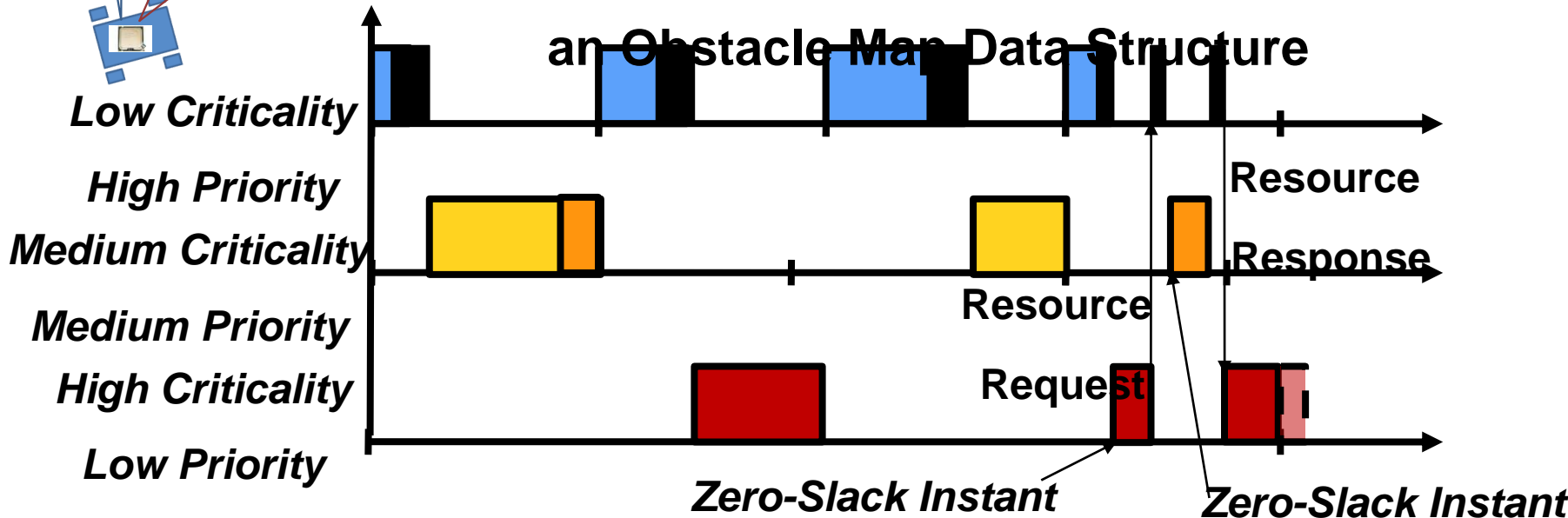
# What about Shared Resources?



- Consider a new medium priority and Medium criticality task (say v2v task)

- Let Planning and Obstacle Avoidance share

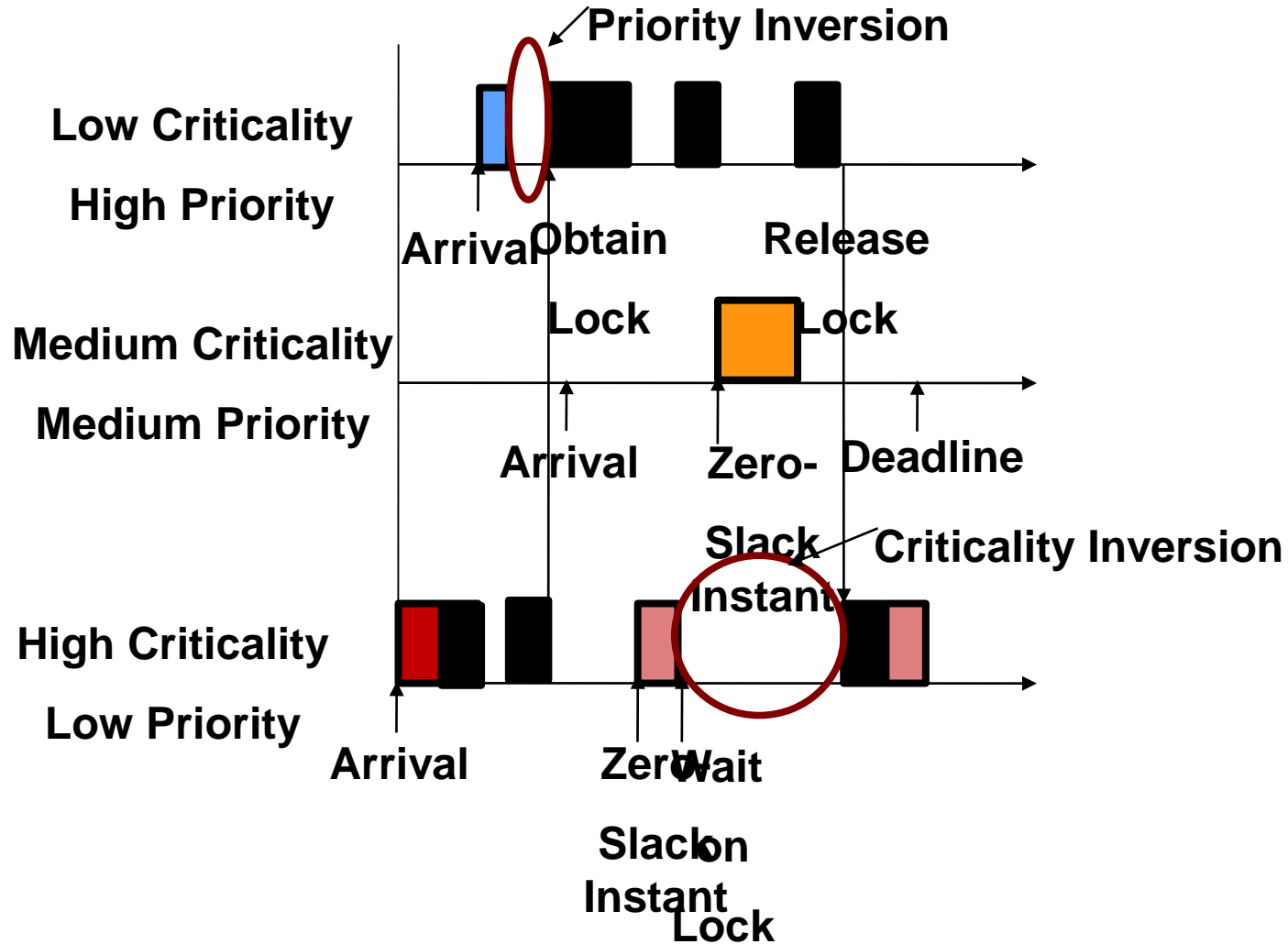
an Obstacle Map Data Structure



Potentially leads to unbounded criticality inversion of V2V Task



# Priority and Criticality Inversion



# Blocking in Zero-Slack Scheduling

A job  $J_h$  waiting for a job  $J_l$  to exit critical section  $Z_{l,k}$  is considered to be ***blocked*** at time  $t$ , *if and only if* one of the following conditions is satisfied at  $t$ :

- 1) The priority of  $J_l$  is lower than  $J_h$ 's priority and  $J_l$  is running in its ***normal*** mode.
- 2) The criticality of  $J_l$  is lower than  $J_h$ 's criticality and  $J_h$  is running in its ***critical*** mode.



# Priority and Criticality Inheritance Protocol (PCIP)



# PCIP Definition

A task  $\tau_i$  that holds a lock to a resource can inherit the priority from a task  $\tau_j$  and the criticality from a task  $\tau_k$  ( $\tau_k$  can be the same as  $\tau_j$ ), both requesting a lock to the resource held by  $\tau_i$  as follows:

- $\tau_i$  *inherits the priority* of  $\tau_j$  if  $\tau_j$ 's priority is higher.
  - This inherited priority has an immediate effect on the scheduling of  $\tau_i$
- $\tau_i$  *inherits the criticality* of  $\tau_k$  if  $\tau_k$ 's criticality is higher.
  - This criticality is used by  $\tau_i$  immediately as soon as  $\tau_k$  requests the lock held by  $\tau_i$



# PCIP Possible Blocking

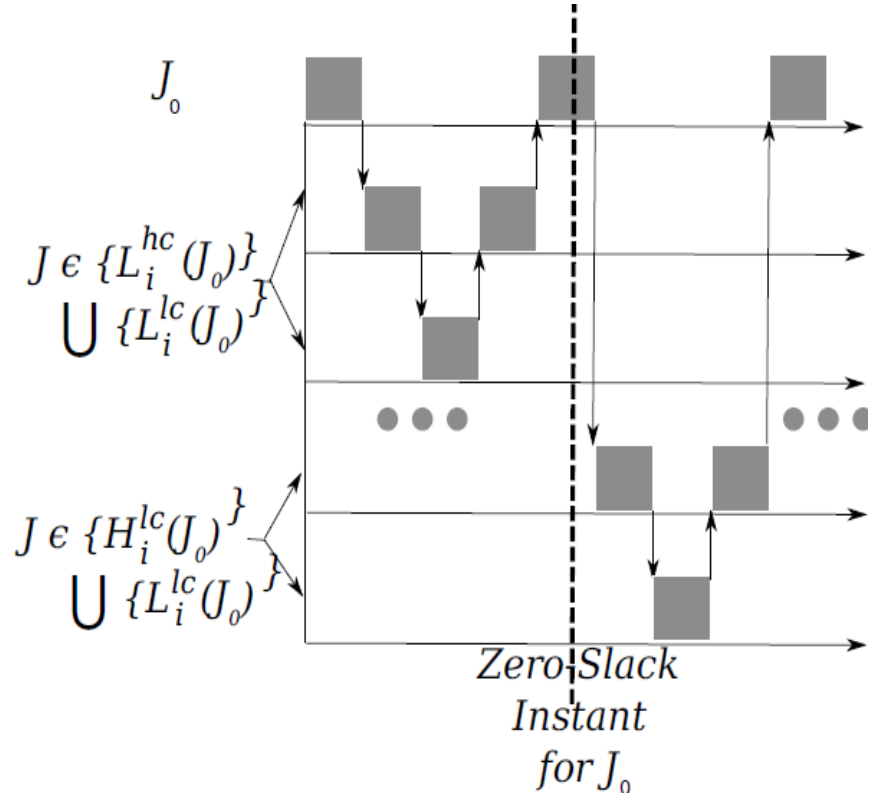
Consider a Job  $J_0$

•  $L_i^{hc}(J_0)$  is the set of jobs with lower priority and higher criticality

•  $L_i^{lc}(J_0)$  is the set of jobs with lower priority and lower criticality

•  $H_i^{lc}(J_0)$  is the set of jobs with

higher priority and lower criticality and other sets can lead to priority or criticality inversion

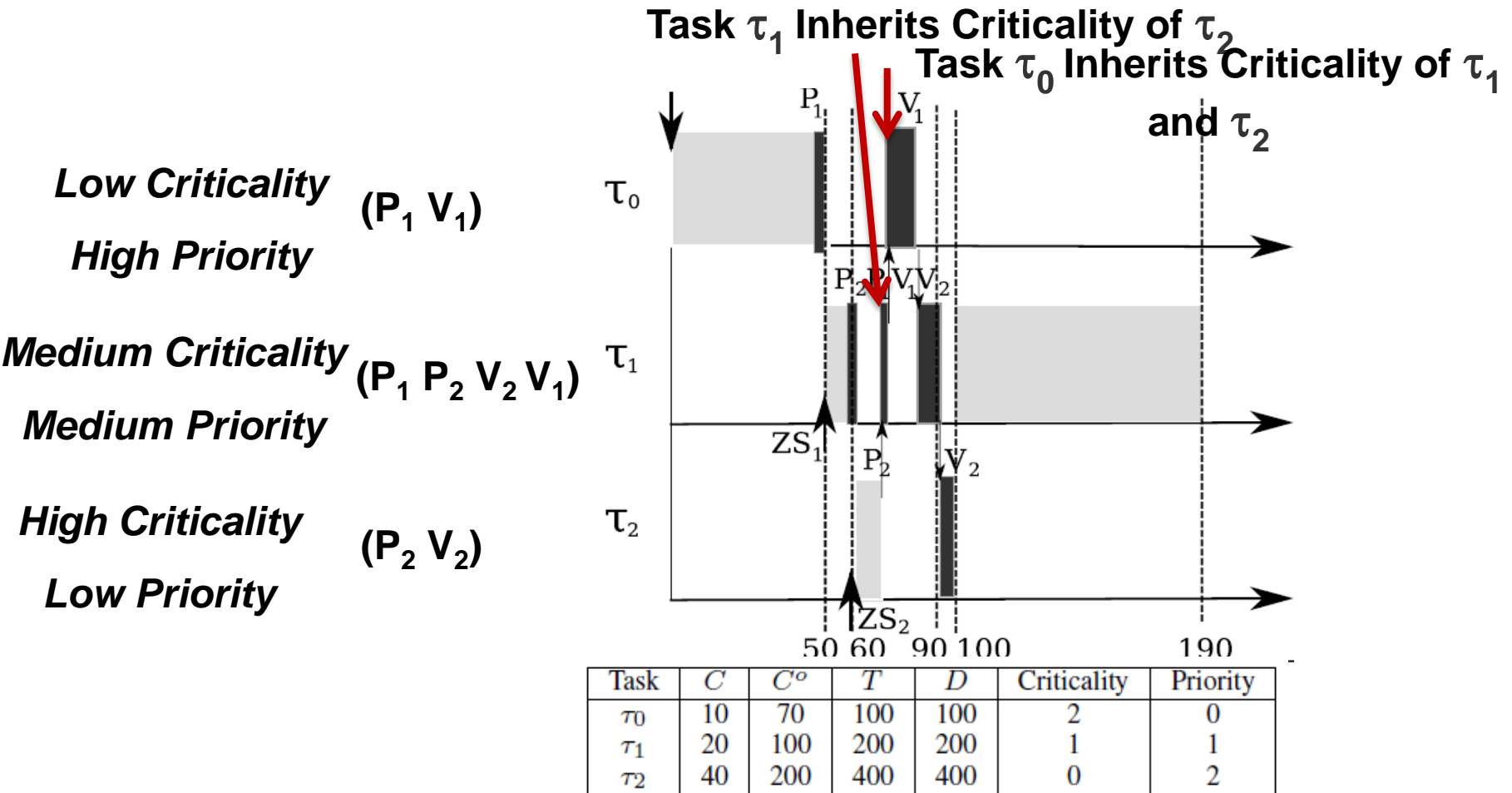


# PCIP Properties

- Under PCIP, given a job  $J_0$  for which there are  $n$  jobs  $\{J_1, J_2, \dots, J_n\}$ , with  $J_i$  in  $\{L_i^{hc}(J_0) \cup L_i^{lc}(J_0) \cup H_i^{lc}(J_0)\}$ , job  $J_0$  can be blocked for at most the duration of one critical section in each of  $\beta_{0,i}^*$ .
  - where,  
 $\beta_{0,i}^*$  is the set of critical sections of  $J_i$  that can block  $J_0$
- Under PCIP, if there are “ $m$ ” locks which can block job  $J$ , then  $J$  can be blocked at most “ $m$ ” times in its *normal* mode and blocked at most “ $m$ ” times in its *critical* mode.



# PCIP Illustration



# Priority and Criticality Ceiling Protocol (PCCP)



# PCCP Definition

- Each lock is assigned both a *priority ceiling* and a *criticality ceiling*
  - *Priority ceiling* is the highest possible priority of any locker of the lock
  - *Criticality ceiling* is the highest possible criticality of any locker of the lock
- Both the *priority ceiling* and the *criticality ceiling* of a lock are acquired by task whenever it holds the lock



# PCCP – Maximum Blocking

- Each job  $J$  can only be blocked twice
  - At most once in **Normal** execution mode
  - At most once in **Critical** execution mode
- Each job  $J_w$  can block job  $J$  only once
  - Otherwise,  $J_w$  is  $L_i^{lc}(J_o)$
  - And, Job  $J$  has to be blocked by  $J_w$  once in Normal mode
  - However,  $J_w$  cannot obtain the processor again as it is in  $L_i^{lc}(J_o)$  !!!



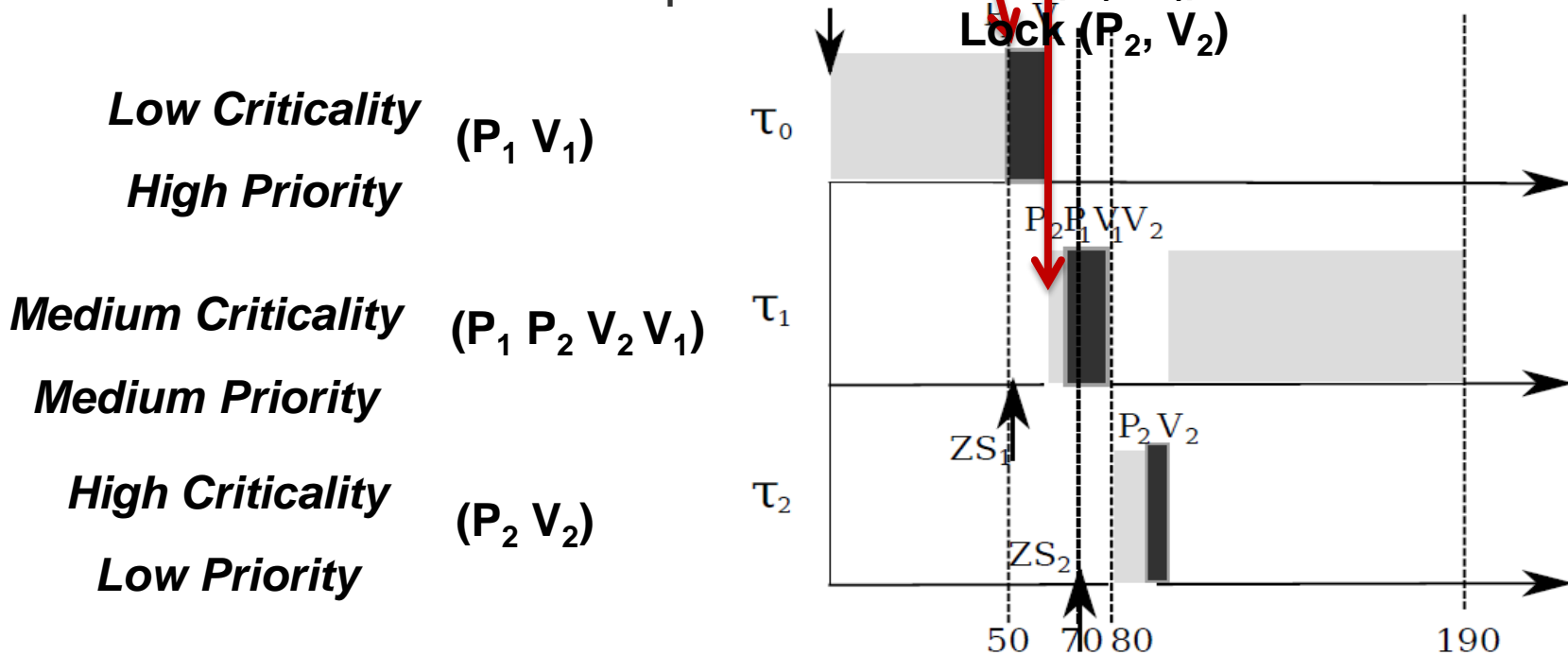
# PCCP - No Deadlocks

- Under PCCP, **no job**  $J_k$  can preempt another job  $J_i$  while  $J_i$  holds a lock (i.e. is inside the critical section) that is also accessed by  $J_k$ .
- PCCP prevents *Transitive Blocking*
- PCCP prevents *Deadlocks*



# PCCP Illustration

Task  $\tau_0$  acquires the Priority and Criticality Ceiling of Task  $\tau_1$   
 Task  $\tau_1$  acquires the Priority and Criticality Ceiling of Task  $\tau_2$



Task	$C$	$C^o$	$T$	$D$	Criticality	Priority
$\tau_0$	10	70	100	100	2	0
$\tau_1$	20	100	200	200	1	1
$\tau_2$	40	200	400	400	0	2



# PCIP Blocking Term Analysis

- PCIP Blocking Term  $B_i$  for Task  $\tau_i$

$$B_i = \min\left(\sum_{\tau_j \in \{H_i^{lc} \cup L_i^{lc} \cup L_i^{hc}\}} \lambda(\beta_{i,j}^*), \sum_{\Psi_{i,k} \in \Psi_i} 2\Lambda(\Psi_{i,k})\right)$$

where,

- $\beta_{i,j}^*$  is the set of critical sections of  $\tau_j$  that can block  $\tau_i$
- $\lambda(\beta_{i,j}^*)$  is the **length of the longest critical sections** of  $\beta_{i,j}^*$  that can block task  $\tau_i$
- $\Lambda(\Psi_{i,j})$  is the **length of the critical section** protected by lock  $\Psi_{i,j}$



# PCCP Blocking Term Analysis

- PCCP Blocking Term  $B_i$  for Task  $\tau_i$

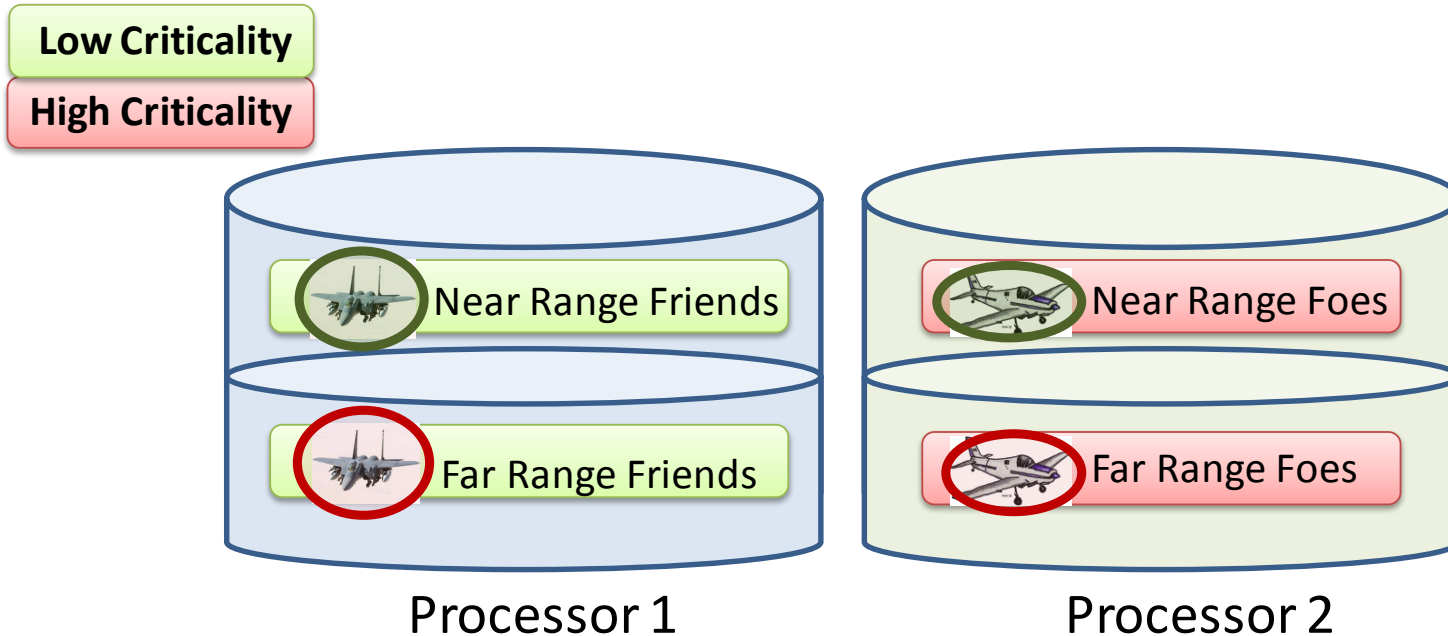
$$B_i = \max_{\tau_j \in \{H_i^{lc} \cup L_i^{lc} \cup L_i^{hc}\}} 2\lambda(\beta_{i,j}^*)$$

where,

- $\beta_{i,j}^*$  is the set of critical sections of  $\tau_j$  that can block task  $\tau_i$
- $\lambda(\beta_{i,j}^*)$  is the **length of the longest critical sections** of  $\beta_{i,j}^*$  that can block task  $\tau_i$



# Criticality isolation strategy (S)



	H	L
Both overload	0	0
High criticality overload	0	1
Low criticality overload	1	0
No overload	1	1

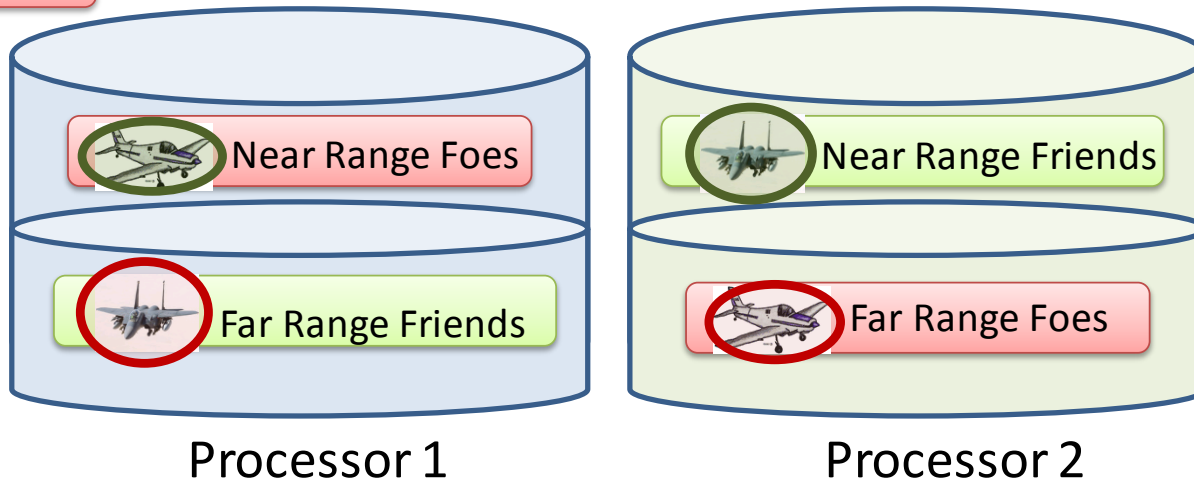
Assume that the system is schedulable without overloads

Under overloads only one task can meet its deadline



# Criticality mixture strategy (T)

Low Criticality  
High Criticality



Processor 1

Processor 2

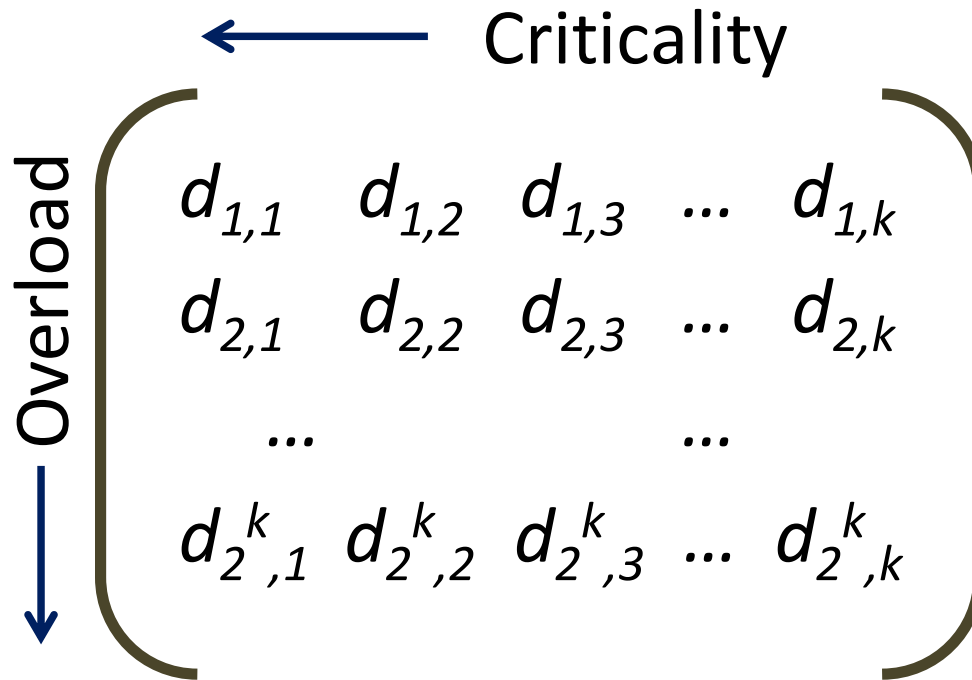
	H	L
Both overload	1	0
High criticality overload	1	0
Low criticality overload	1	0
No overload	1	1

***T is better than S***

Assume that the system is schedulable without overloads  
Under overloads only one task can meet its deadline  
*Assume that a uniprocessor mixed-criticality scheduling algorithm like ZSRM is used within each processor*



# Generalization: Ductility Matrix



Say we have 'k' criticality levels

$2^k$  possible overload scenarios

- *All criticality levels overload* to *No overload*



# Quantification of Ductility

$$P_d(D) = \sum_{c=1}^k \left\{ \frac{1}{2^c} \frac{\sum_{r=1}^{2^k} d_{r,c}}{2^k} \right\}$$

<i>Scheme S</i>	H	L
Both overload	0	0
High criticality overload	0	1
Low criticality overload	1	0
No overload	1	1

For *S*,  $P_d(D) = 0.375$

<i>Scheme T</i>	H	L
Both overload	1	0
High criticality overload	1	0
Low criticality overload	1	0
No overload	1	1

For *T*,  $P_d(D) = 0.5625$

Shows that **T** is better than **S**

Other Projection functions can be used

$P_d(D)$  favors the more critical tasks **exponentially** over the lower criticality tasks



# Outline

Mixed-criticality task scheduling problem

Zero-slack scheduling for uni-processors

- Zero-slack metrics & properties

Generalizing resource allocation to distributed mixed criticality tasks

- Generalized metric: Ductility matrix

Compress-on-Overload Packing (COP)

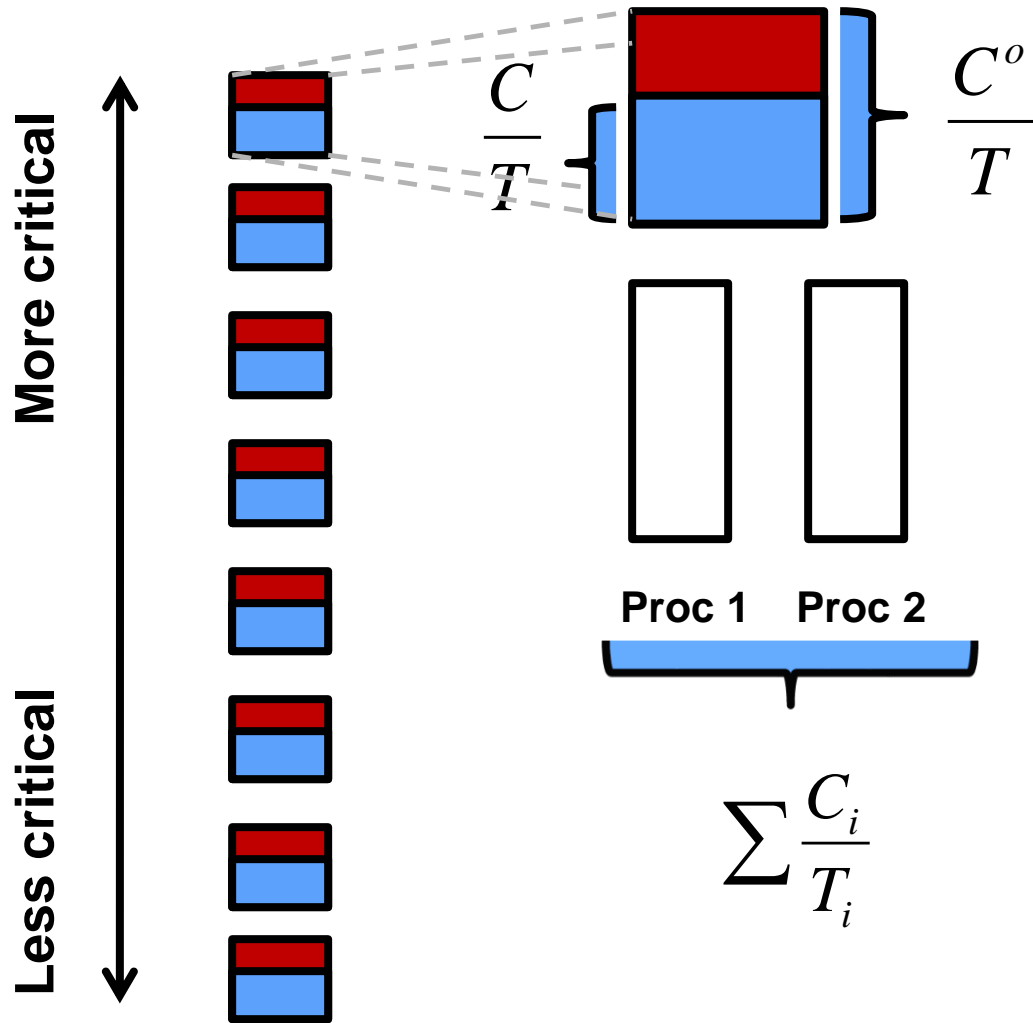
- COP Performance

Radar surveillance case study

Conclusions

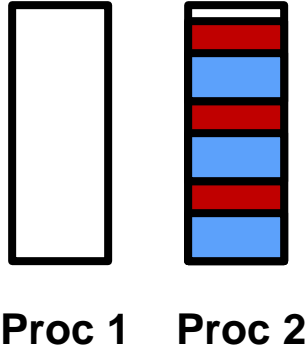
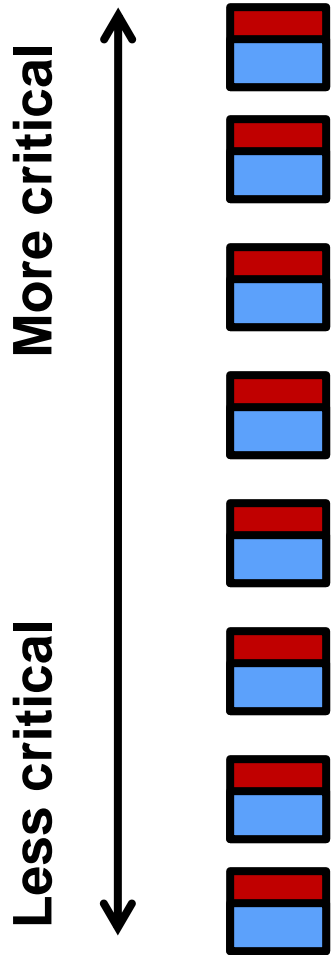


# Compress-on-Overload Packing (COP)



# Compress-on-Overload Packing (COP)

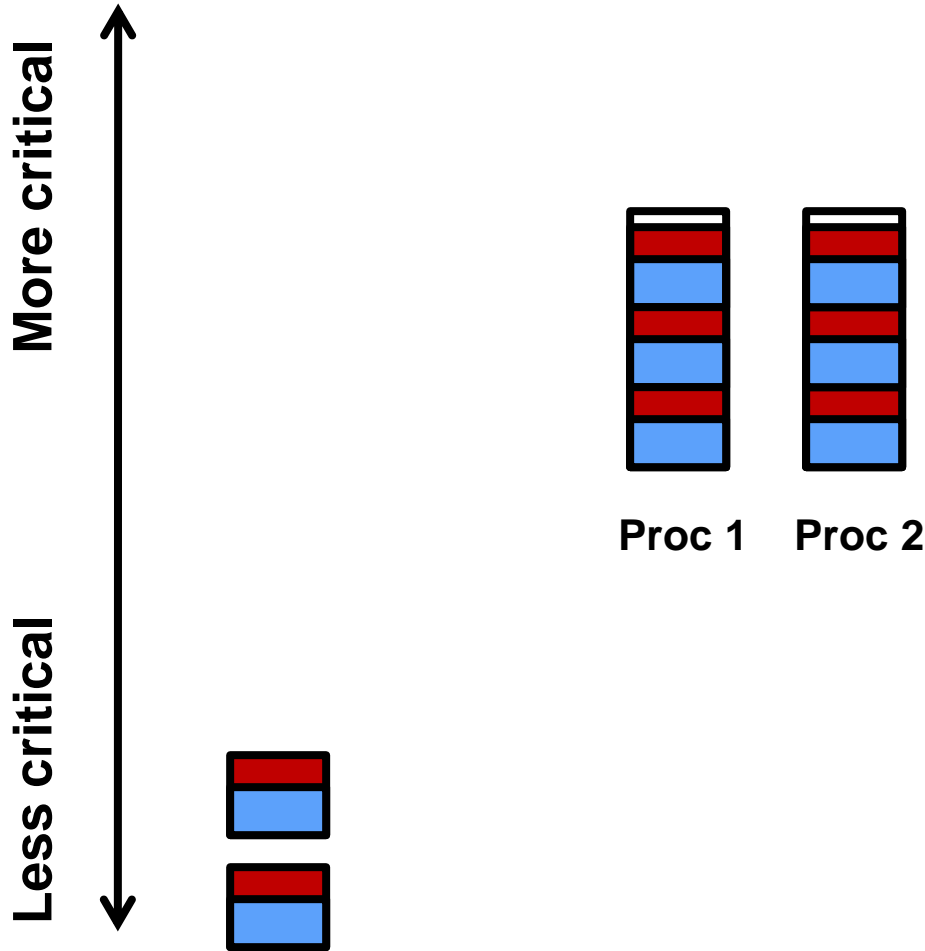
Phase 1: Pack by criticality then size  
object size =  $\frac{C_i^o}{T_i}$



# Compress-on-Overload Packing (COP)

Phase 2: Pack by criticality then size

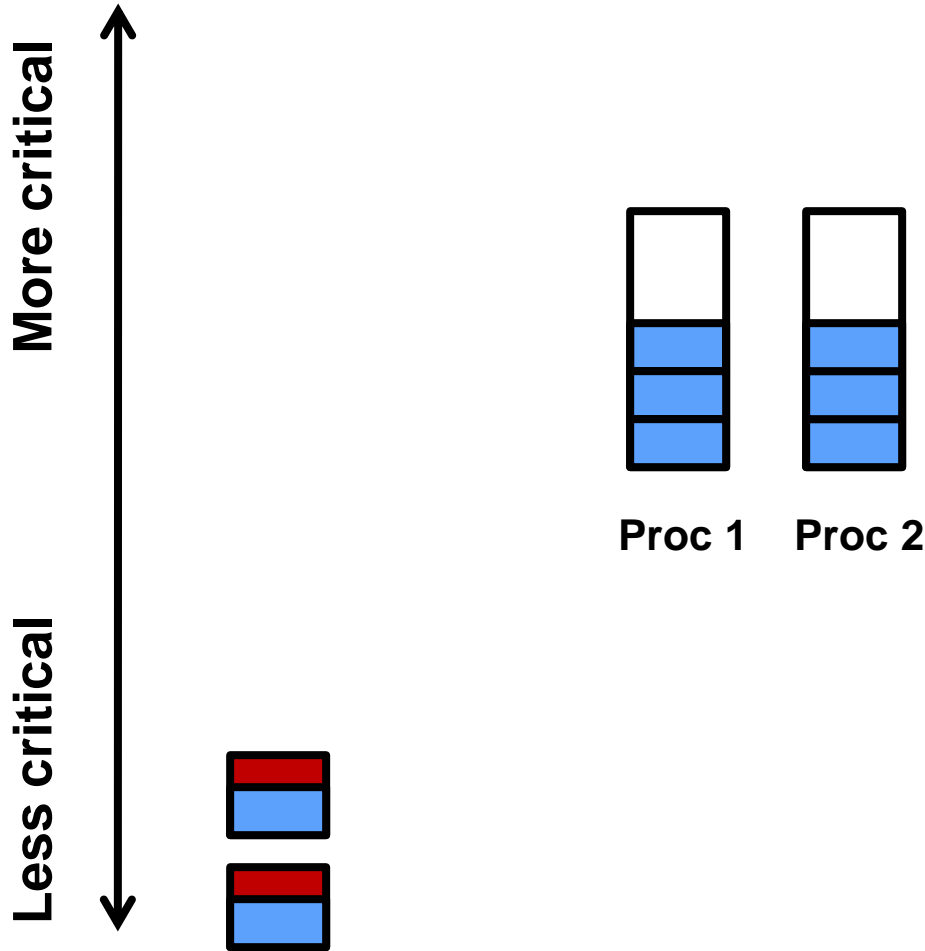
$$\text{object size} = \frac{C_i}{T_i}$$



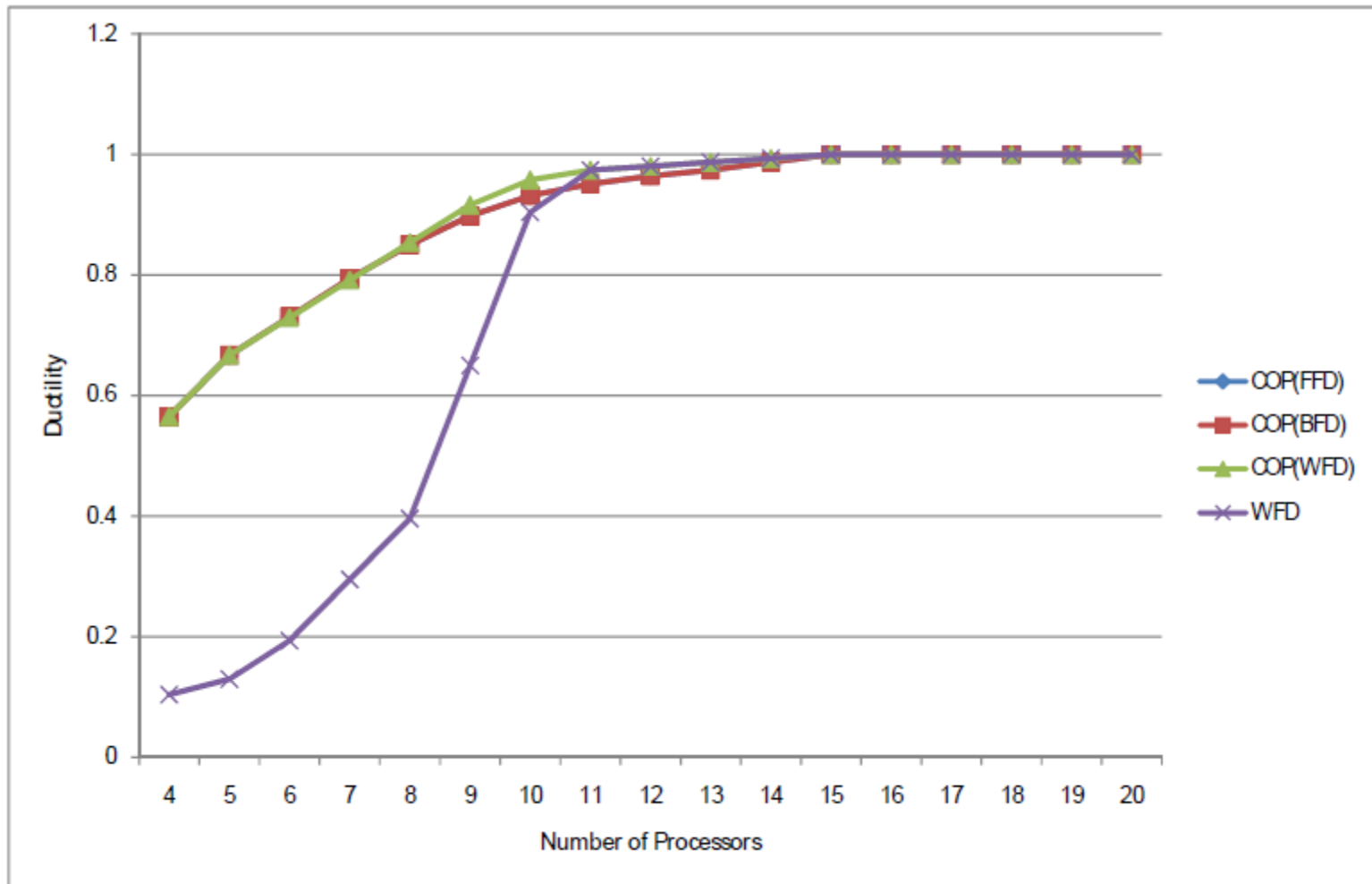
# Compress-on-Overload Packing (COP)

Phase 2: Pack by criticality then size

$$\text{object size} = \frac{C_i}{T_i}$$

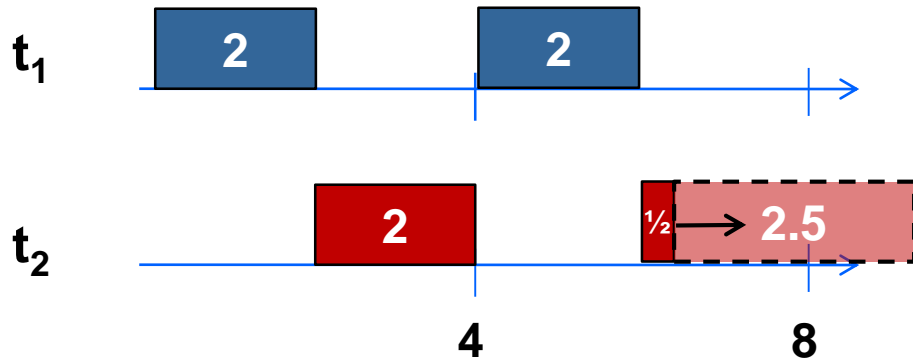


# COP Performance



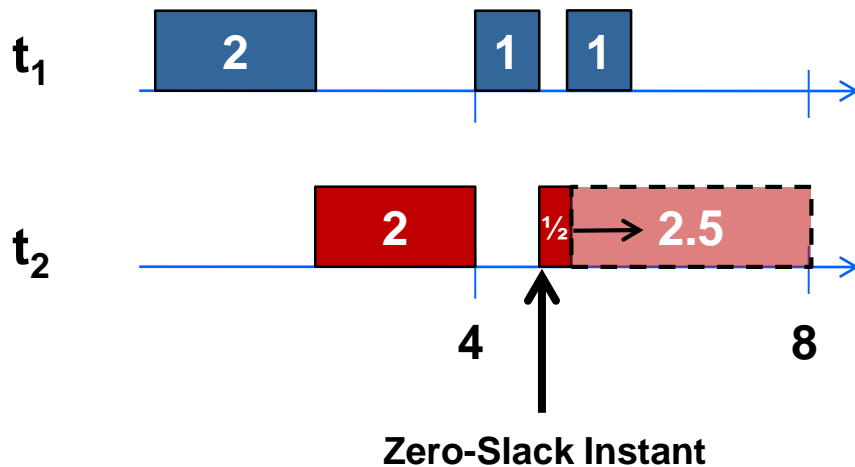
# Overloading in Mixed-Criticality Systems

Task	Period	Criticality	WCET	NCET
$t_1$ Surveillance Cov.	4	Mission	2	2
$t_2$ Collision Avoid.	8	Safety	5	2.5



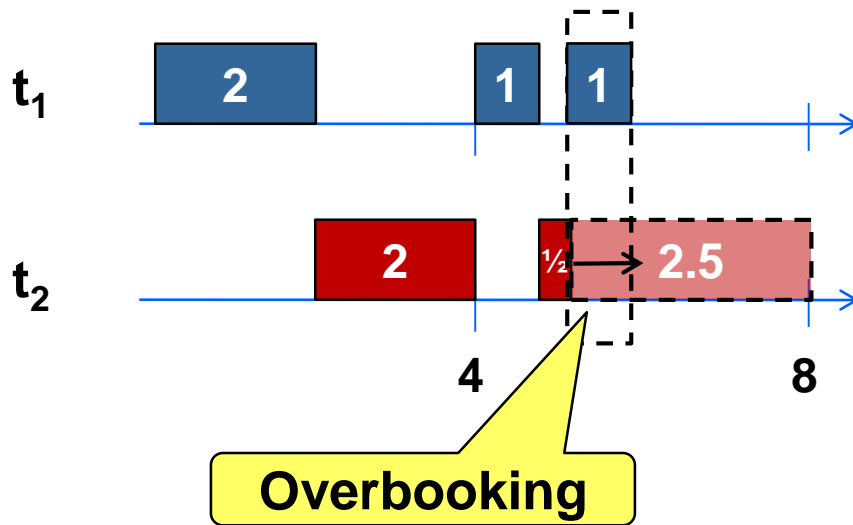
# Zero-Slack Rate Monotonic

Task	Period	Criticality	WCET	NCET
$t_1$ Surveillance Cov.	4	Mission	2	2
$t_2$ Collision Avoid.	8	Safety	5	2.5



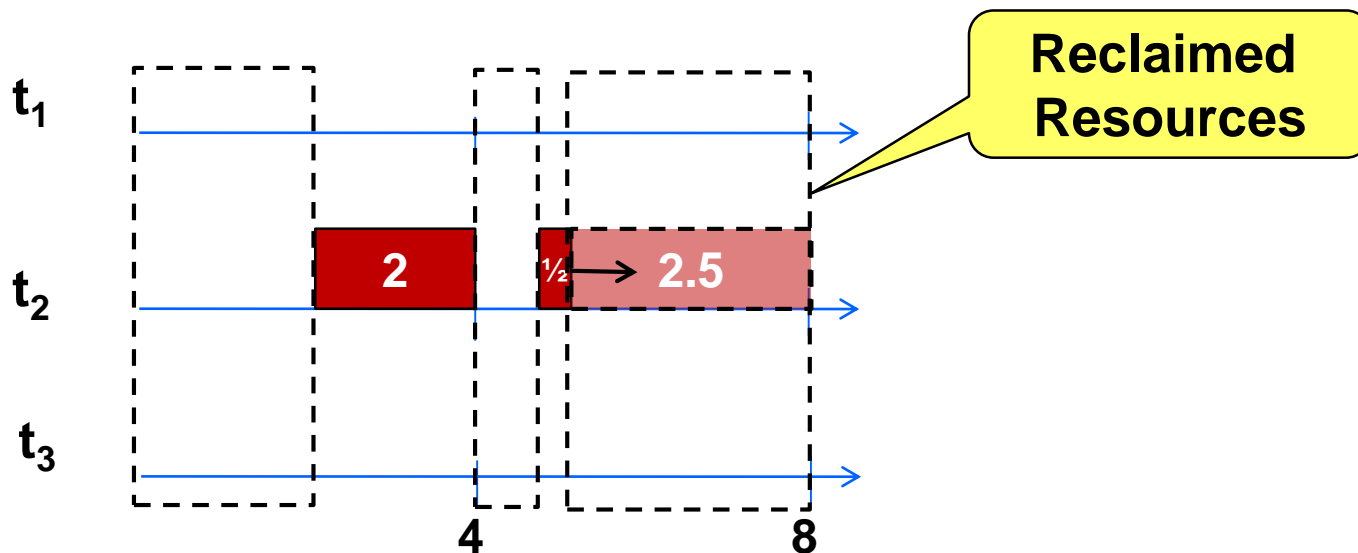
# Zero-Slack Rate Monotonic

Task	Period	Criticality	WCET	NCET
$t_1$ Surveillance Cov.	4	Mission	2	2
$t_2$ Collision Avoid.	8	Safety	5	2.5



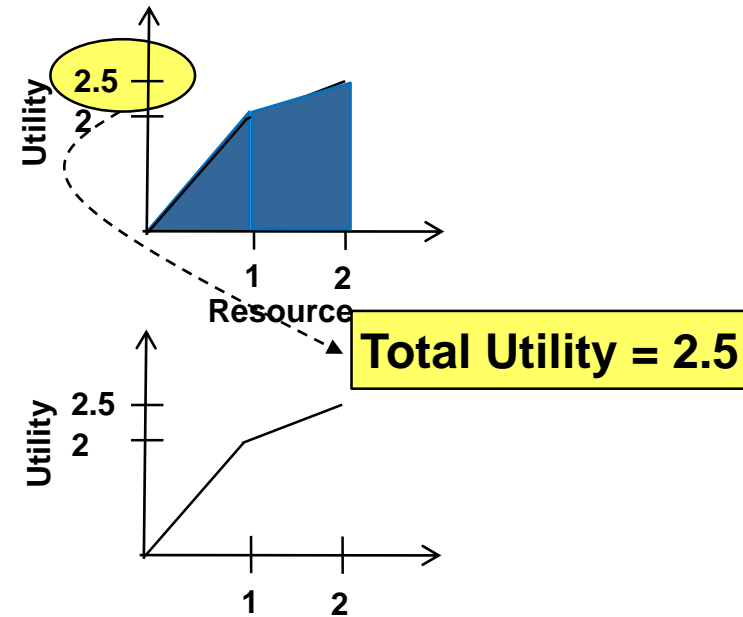
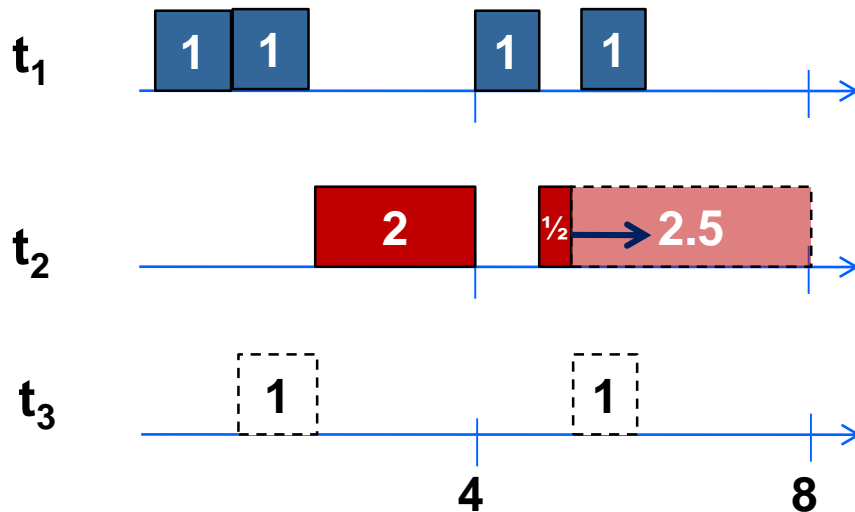
# Reclaiming Resources in Mixed-Criticality Systems

Task	Period	Criticality	WCET	NCET	Utility
$t_1$ Surveillance Cov.	4	Mission	2	2	{2,2.5}
$t_2$ Collision Avoid.	8	Safety	5	2.5	
$t_3$ Amount of Intelligence	4	Mission	2	2	{2,2.5}



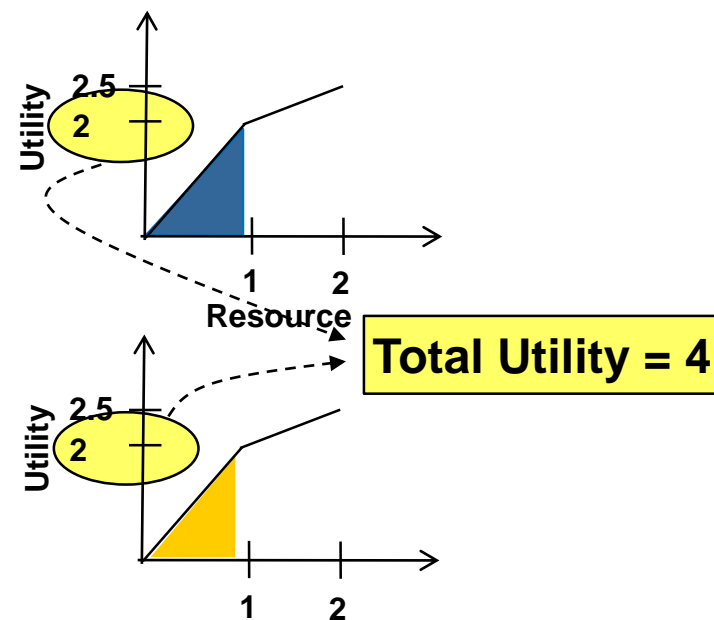
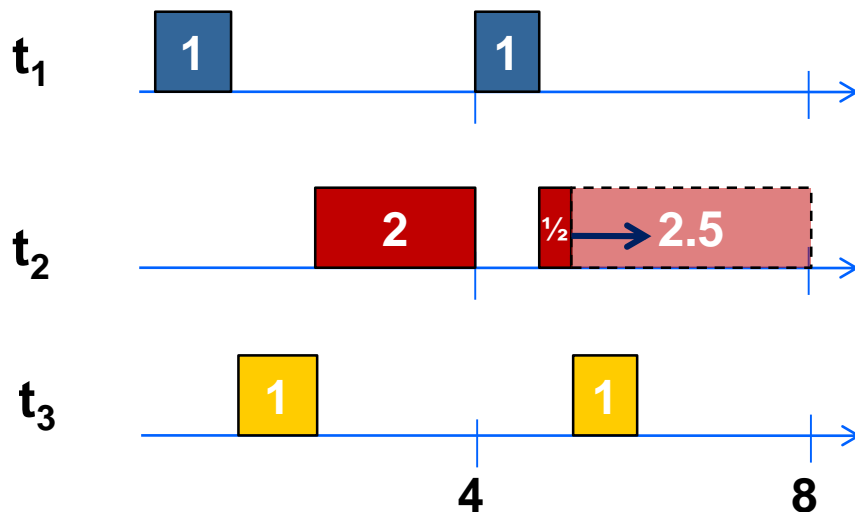
# Using Reclaimed Resources to Maximized Utility

Task	Period	Criticality	WCET	NCET	Utility Levels
$t_1$ Surveillance Cov.	4	Mission	2	2	{2,2.5}
$t_2$ Collision Avoid.	8	Safety	5	2.5	
$t_3$ Amount of Intelligence	4	Mission	2	2	{2,2.5}



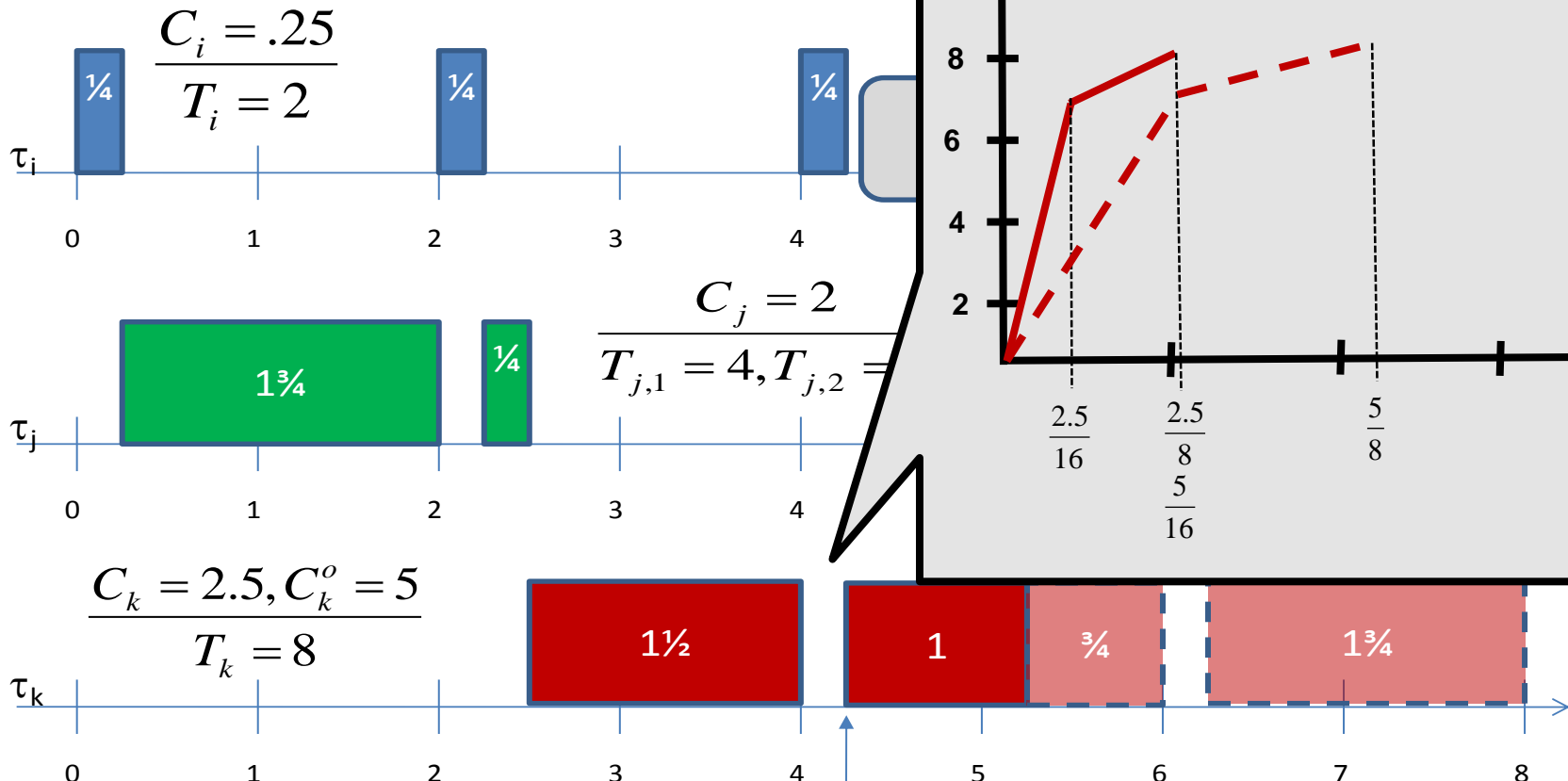
# Using Reclaimed Resources to Maximized Utility

Task	Period	Criticality	WCET	NCET	Utility Levels
$t_1$ Surveillance Cov.	4	Mission	2	2	{2,2.5}
$t_2$ Collision Avoid.	8	Safety	5	2.5	
$t_3$ Amount of Intelligence	4	Mission	2	2	{2,2.5}



**ZS-QRAM: More mission-critical utility from same resources**

# ZSQRAM: Period Degradation



ZS Instant



# ZSQRAM: Period Degradation

