Global Scheduling of Multi-Mode Real-Time Applications upon Multiprocessor Platforms

Vincent Nélis, Patrick Meumeu Yomsi, Björn Andersson, and Joël Goossens

Abstract—Multi-mode real-time systems are those which support applications with different modes of operation, where each mode is characterized by a specific set of tasks. At run-time, such systems can, at any time, be requested to switch from its current operating mode to another mode (called “new mode”) by replacing the current set of tasks with that of the new-mode. Thereby, ensuring that all the timing requirements are met not only requires that a schedulability test is performed on the tasks of each mode but also that (i) a protocol for transitioning from one mode to another is specified and (ii) a schedulability test for each transition is performed. We propose two distinct protocols that manage the mode transitions upon uniform and identical multiprocessor platforms at run-time, each specific to distinct task requirements. For each protocol, we formally establish schedulability analyses that indicate beforehand whether all the timing requirements will be met during any mode transition of the system. This is performed assuming both Fixed-Task-Priority and Fixed-Job-Priority schedulers.

Index Terms—Multi-Mode Systems, Multiprocessor Scheduling, Real-Time Scheduling.

I. INTRODUCTION

Hard real-time systems require both functionally correct executions and results that are produced on time. Control of the traffic (ground or air), control of engines, control of chemical and nuclear power plants are just some examples of such systems. Currently, numerous techniques exist that enable engineers to design real-time systems while guaranteeing that all the temporal requirements are met. These techniques generally model each functionality of the application by a recurrent task, characterized by a computing requirement, a temporal deadline and an activation rate. Commonly, real-time applications are simply modeled by a single and finite set of such tasks. However, practical applications often exhibit multiple behaviors issued from several operating modes (e.g., an initialization mode, an emergency mode, a fault recovery mode, etc.), where each mode is characterized by its own set of functionalities, i.e., its set of tasks. During the execution of such multi-mode real-time applications, switching from the current mode (called the old-mode) to any other mode (called the new-mode) requires to substitute the currently executing task set with the set of tasks of the new-mode. This substitution introduces a transient phase, where tasks of both the old- and new-mode may be scheduled simultaneously, thereby leading to a possible overload that can compromise the system schedulability—indeed it can be the case that both the old- and new-mode have been asserted schedulable by the schedulability analysis but the transition between them fails at run-time.

The scheduling problem during a transition between two modes has multiple aspects, depending on the behavior and requirements of the old- and new-mode tasks when a mode change is initiated. Upon a mode change request:

- an old-mode task may be allowed to be immediately aborted or, on the contrary, can be required to complete the execution of its current active job (so that it preserves data consistency for instance). Using scheduling algorithms such as the one considered in this study, we will prove in Section V that aborting tasks upon a mode change request does not jeopardize the schedulability of the mode transitions. Hence, we assume in this paper the most problematic scenario in which every old-mode task must complete its current active job (if any) when a mode change is requested.
- a new-mode task either requires to be activated as soon as possible when a mode change is requested or requires to be activated only when all the active jobs issued from the old-mode have totally completed their execution.

Finally, there may be some tasks (called mode-independent tasks in the literature) that belong to more than one mode and such that their activation pattern must not be jeopardized during the transition between those modes. However this paper will consider only systems that do not include such tasks.

Transition scheduling protocols for tasks without mode-independent tasks are often classified with respect to the way they schedule the old- and new-mode tasks during the transitions. In the literature (see for instance [1] which considers uniprocessor systems), the following definitions are used.

Definition 1 (Synchronous/asynchronous protocol [1]): A transition protocol is said to be synchronous if it schedules new-mode tasks only when all the old-mode tasks have completed. Otherwise, it is said to be asynchronous.

Definition 2 (Protocol with/without periodicity [1]): A transition protocol is said to be “with periodicity” if and only if it is able to deal with mode-independent tasks. Otherwise, it is said to be “without periodicity”.

1In practice, mode-independent tasks typically allow to model daemon functionalities.
A. Related work

Numerous transition protocols have been proposed for uniprocessor platforms (a survey about this concern is presented in [1]). In such environments, existing researches [1–3] have shown that even if two modes of the application have been proven feasible, the transition between the two modes can cause violation of timing constraints, hence needing explicit analyses. Such analyses have been proposed in [4], considering the popular Rate Monotonic Algorithm. Unfortunately three years later, this analysis was shown optimistic [5] in the sense that some unfeasible task sets could be asserted schedulable. In the same paper [5], the authors improved the previous analysis and proposed a new one which considers the popular Deadline Monotonic Algorithm. An analysis of sporadic tasks scheduled on EDF is known as well [6]. In [7], the authors proposed an analysis which considers Fixed-Task-Priority scheduling (FTP), Earliest-Deadline-First [8] scheduling and arbitrary task activation pattern. Furthermore, for applications that were initially proven not schedulable during the transition phases, they derived the required offsets for delaying the initialization of transition between two modes in order to make the application schedulable.

Among the uniprocessor synchronous protocols, the authors of [1], [9], [10] proposed the following protocols.

- The Minimum Single Offset Protocol (MSO) [1] where the last activation of each old-mode task completes and then, the new-mode tasks are released.
- The Idle Time Protocol (IT) [10] where the periodic activations of the old-mode tasks are suspended at the first idle time-instant occurring during the transition and then, the new-mode tasks are released.
- The Maximum-Period Offset Protocol (MPO) [9] where the delays of first activation of each new-mode task is equal to the period of the less frequent task in both modes.

Among the uniprocessor asynchronous protocols, the authors of [5], [6], [11] proposed the following protocols.

- A protocol without periodicity [11] where tasks are assigned priorities according to the Deadline Monotonic Scheduling algorithm and are scheduled with time offsets during the mode change only.
- A protocol with periodicity has been introduced by Sha et al. in [12], assuming Fixed-Task-Priority scheduling. Then, the authors of [6] extended this protocol to the Earliest Deadline First [8] scheduling algorithm.
- The authors of [5] introduced a particular protocol which allows tasks to modify their parameters (period, execution time, etc.) during the mode changes. As in [11], this study assumes that the tasks are scheduled according to the Deadline Monotonic scheduling algorithm.

B. Contribution and paper organization

In this paper we propose two protocols without periodicity (SM-MSO which is synchronous and AM-MSO which is asynchronous) for managing mode transitions during the execution of multi-mode real-time applications on multiprocessor platforms. Both protocols can be considered as a generalization to multiprocessors of the MSO protocol proposed in [1]. We assume that every operating mode of the application is scheduled by a global, work-conserving, preemptive and Fixed-Job-Priority (FJP) scheduling algorithm (formal definitions are given in Section II-D). Some of the results presented here have already been published (see [13]–[16]). It is worth noticing that the problem of scheduling multi-mode applications upon multiprocessor platforms is much more complex than upon uniprocessor platforms, especially due to the presence of scheduling anomalies (see Chapter 5 of [17] for a definition) and it is now well known that real-time multiprocessor scheduling problems are typically not solved by applying straightforward extensions of techniques used for solving similar uniprocessor problems.

The paper is organized as follows. Section II defines the computational model used throughout the paper. Sections III and IV describe the synchronous and asynchronous protocols SM-MSO and AM-MSO, respectively. Section V introduces some definitions and basic results necessary for the establishment of our schedulability analyses. These four first Sections II–V are a common base of the paper, in the sense that these 14 pages describe both the models of computation and protocols independently of the platform and scheduler characteristics. Then, the four next Sections VI–IX are each specific to a platform and scheduler model. More precisely, they provide a schedulability analysis for both SM-MSO and AM-MSO, assuming in turn identical platforms and Fixed-Job-Priority schedulers (in Section VI), identical platforms and Fixed-Task-Priority schedulers (in Section VII), uniform platforms and Fixed-Job-Priority schedulers (in Section VIII) and uniform platforms and Fixed-Task-Priority schedulers (in Section IX)2. Finally, Section X gives our conclusions and future work, together with some remaining open problems.

II. MODELS OF COMPUTATION AND SPECIFICATIONS

A. Application specifications

We define a multi-mode real-time application $\tau$ as a set of $x$ operating modes denoted by $M^1, M^2, \ldots, M^x$ where each mode $M^k$ has to execute its associated task set $\tau^k = \{\tau^k_1, \tau^k_2, \ldots, \tau^k_n\}$ composed of $n$ tasks by following the scheduler $S^k$. At run-time, the application is either running in one and only one mode, i.e., it is executing only the set of tasks associated to that mode, or it is switching from one mode to another one. Since we do not consider mode-independent tasks in this study, it holds that $\tau^k \cap \tau^j = \emptyset$, $\forall k \neq j$.

Each task $\tau^k_i$ is modeled by a sporadic and constrained-deadline task characterized by three parameters $\langle C^k_i, D^k_i, T^k_i \rangle$—a worst-case execution time $C^k_i$, a minimum inter-arrival time $T^k_i$ and a relative deadline $D^k_i \leq T^k_i$—with the interpretation that, during the execution in mode $M^k$, task $\tau^k_i$ generates successive jobs $\tau^k_{ij}$ (with $j = 1, \ldots, \infty$) released at times $a^k_{ij}$ such that $a^k_{ij} \geq a^k_{i,j-1} + T^k_i$ (with $a^k_{1,1} \geq 0$), each such job has an execution requirement of at most $C^k_i$, and must be completed at (or before) its absolute

2Even though Fixed-Job-Priority schedulers encompass the family of Fixed-Task-Priority schedulers, the particular case of Fixed-Task-Priority schedulers is treated separately so that the schedulability analyses are more specific and therefore more accurate.
deadline noted $d^k_{i,j} \overset{\text{def}}{=} a^k_{i,j} + D^k_{i,j}$. In the particular case where $a^k_{i,j} = a^k_{i,j-1} + T^k_{i,j}$, $\forall j > 1$, the task $\tau^k_{i,j}$ is said to be periodic instead of sporadic. In the same vein, if $D^k_{i,j} = T^k_{i,j}$ then the task $\tau^k_{i,j}$ is said to be implicit-deadline instead of constrained-deadline.

Definition 3 (Active job): We say that a job $\tau^k_{i,j}$ is active at time $t$ if it has been already released (i.e., $t \leq a^k_{i,j}$) and it is not completed yet.

Since we assume $D^k_{i,j} \leq T^k_{i,j}$, there cannot be two jobs of a same task $\tau^k_{i,j}$ active at the same time in any feasible schedule. All the tasks are assumed to be independent, i.e., there is no communication, no precedence constraint and no shared resource (except the processors) between them. In [15], we introduced the following concept of enabled/disabled tasks.

Definition 4 (Enabled/disabled tasks [15]): At run-time, any task $\tau^k_{i}$ of the application can generate jobs if and only if $\tau^k_{i}$ is enabled. Symmetrically, a disabled task cannot generate jobs.

As such, disabling a task $\tau^k_{i}$ prevents future job releases from $\tau^k_{i}$. When all the tasks of any mode $\tau^k_{j}$ are enabled and all the tasks of all the other modes are disabled, the application is said to be running in mode $M^j$ (since only the tasks of mode $\tau^k_{j}$ can release jobs). We denote by $\text{enabled}(\tau^j_{i}, t)$ and $\text{disabled}(\tau^j_{i}, t)$ the subsets of enabled and disabled tasks of $\tau^j_{i}$ at time $t$, respectively.

B. Platform specifications

Many recent embedded systems are built upon multiprocessor platforms in order to fulfill the high computational requirements of applications. As pointed out in [18], [19], another advantage of such a choice is the fact that multiprocessor systems are more energy efficient than equally powerful uniprocessor platforms. Indeed, raising the frequency of a single CPU results in a multiplicative increase of the consumption while adding CPUs results in an additive increase. Two distinct multiprocessor architectures are commonly used in the industrial world and thus, are considered in this paper: identical and uniform platforms.

Identical platform. In such multiprocessor platforms, all the CPUs have the same computational capabilities, with the interpretation that in any interval of time two CPUs execute the same amount of work (assuming that none of them is idling). In the remainder of this paper, any platform composed of $m$ identical CPUs will be modeled by $\pi \overset{\text{def}}{=} \{\pi_1, \pi_2, \ldots, \pi_m\}$ where $\pi_i$ denotes the $i$th CPU of the platform.

Uniform platform. In such multiprocessor platforms, the CPUs are allowed to have different computational capabilities. That is, a parameter $s_i$ is associated to every CPU $\pi_i$ with the interpretation that in any time interval of length $t$, CPU $\pi_i$ executes $s_i \cdot t$ units of execution (if it is not idling). This parameter can be seen as the execution speed of the CPU. In the remainder of this paper, any platform composed of $m$ uniform CPUs is modeled by $\pi \overset{\text{def}}{=} \{s_1, s_2, \ldots, s_m\}$, where $s_i$ is the execution speed of CPU $\pi_i$. Without loss of generality, we assume that $s_1 \geq s_{i-1}$ $\forall i = 2, 3, \ldots, m$, meaning that CPU $\pi_m$ is the fastest CPU while $\pi_1$ is the slowest one. For all $k \in [1, m]$, we denote by $s(k)$ the cumulated speed of the $(m-k+1)$ fastest CPUs, i.e.,

$$s(k) \overset{\text{def}}{=} \sum_{i=k}^{m} s_i \quad (1)$$

Notice that identical platforms are a particular case of uniform platforms where $s_i = s_j$ $\forall i, j \in [1, m]$. In this particular case we assume without any loss of generality that $\forall i: s_1 = 1$.

C. Mode transition specifications

While the application is running in any mode $M^j$, a mode change can be initiated by any task of $\tau^i$ or by the system itself, whenever it detects a change in the environment or in its internal state for instance. This is performed by invoking a MCR($j$) (i.e., a Mode Change Request), where $M^j$ is the destination mode. We denote by $t_{\text{MCR}(j)}$ the invoking time of the last MCR($j$). From the time at which a mode change is requested to the time at which the transition phase ends, $M^j$ and $M^j$ are referred to as the old- and new-mode, respectively.

At run-time, mode transitions are managed as follows. Suppose that the application is running in mode $M^j$ and the system (or any task of $\tau^j$) comes to request a mode change to mode $M^j$, with $j \neq i$. At time $t_{\text{MCR}(j)}$, the system entrusts the scheduling decisions to a transition protocol which immediately disables all the old-mode tasks, thus preventing them from releasing new jobs. At this time, the active jobs issued from these disabled tasks, henceforth called the rem-jobs (for “remaining jobs”), may have two distinct behaviors: either they can be aborted upon the MCR($j$) or they can complete their execution. From the schedulability point of view, we will show that aborting some (or all) rem-jobs upon a mode change request does not jeopardize the system schedulability during the transition phase. Consequently, we assume the worst-case scenario for every mode transition, i.e., the scenario in which every old-mode task has to complete its last released job (if any) during every mode transition\(^3\). The fact that the rem-jobs have to complete their execution upon the MCR($j$) brings the following problem. Even if both tasks sets $\tau^i$ and $\tau^j$ (from the old- and new-mode, respectively) have been asserted to be schedulable upon the $m$ CPUs at system design-time, the presence of the rem-jobs may cause an overload during the transition phase (at run-time) if all the new-mode tasks of $\tau^j$ are enabled immediately upon the mode change request. Indeed, the schedulability analysis performed beforehand on $\tau^j$ did not take into account the additional work generated by the rem-jobs. To solve this problem, transition protocols usually delay the enabling of each new-mode task until it is safe to do so. However, these delays are also subject to hard constraints. More precisely, we denote by $D^j_{\text{k}}(M^j)$ the relative transition deadline of task $\tau^j_{k}$ during every transition from mode $M^i$ to mode $M^j$, with the following interpretation: the transition protocol must ensure that $\tau^j_{k}$ is enabled not later

\(^3\)Abortting a job consists in suddenly stopping its execution and removing it from the system memory. But in the real world, suddenly killing a process may cause system failures and the rem-jobs often have to complete their execution.
than time $t_{tas} + D_{j}^k(M^i)$. Finally, when all the rem-jobs are completed and all the new-mode tasks of $\tau_j^i$ are enabled, the system entrusts the scheduling decisions to the scheduler $S_j^i$ of the new-mode $M^j$ and the transition phase ends.

In short, the goal of any transition protocol is to fulfill the following requirements during every mode change:
1) Complete each rem-job $\tau_{a,b}^j$ by its absolute deadline $d_{a,b}^j$.
2) Enable each new-mode task $\tau_k^i$ by its absolute transition deadline $t_{tas} + D_{k}^i(M^i)$.
3) Complete each new-mode job $\tau_{a,b}^j$ by its absolute deadline $d_{a,b}^j$.

Definition 5 (Valid protocol [15]): A transition protocol $\mathcal{A}$ is said to be valid for a given application $\tau$ and platform $\pi$ if and only if $\mathcal{A}$ meets all the job and transition deadlines during every transition between every pair of operating modes of $\tau$.

This notion of “valid protocol” is directly related to that of a “validity test” defined as follows.
Definition 6 (Validity test [15]): For a given transition protocol $\mathcal{A}$, a validity test is a condition based on the tasks and platform characteristics that indicates a priori whether $\mathcal{A}$ is valid for a given application $\tau$ and platform $\pi$.

D. Scheduler specifications

We consider the global preemptive scheduling problem of sporadic constrained-deadline tasks upon multiprocessor platforms. “Global” schedulers, in contrast to partitioned ones, allow different tasks and different jobs of the same task to be executed upon different CPUs. When preemptive, global schedulers allow any job to be interrupted at any time prior to completion on any CPU and resumed (possibly later) on any other CPU. We consider that every mode $M^k$ uses its own scheduler denoted by $S^k$ which can be either Fixed-Task-Priority (FTP) or Fixed-Job-Priority (FJP) according to the following interpretations.

• FTP schedulers assign a priority to each task at system design-time (i.e., before the execution of the application) and then at run-time, every released job uses the priority of its task and the priority of a job is kept constant until it completes.
• FJP schedulers assign a priority to each job at run-time (i.e., as soon as it arrives in the system) and every job keeps its priority constant until it completes. As such, different jobs issued from the same task may have different priorities $^5$.

Without loss of generality we assume that, at any time, two active jobs cannot have the same priority. Furthermore, we consider work-conserving schedulers according to the following definition.

Definition 7 (Work-conserving global scheduler): A CPU cannot be idle if there is a job awaiting execution. Usually, priority-based schedulers assign at each instant in time the $m$ highest priority active jobs (if any) to the $m$ CPUs.

$^5$This requirement is automatically fulfilled for synchronous protocols since no new-mode jobs are scheduled during the mode transitions.

$^6$According to these interpretations, FTP schedulers are a particular case of FJP schedulers in which all the jobs issued from a same task receive the same priority determined beforehand.

The above definition of work-conserving schedulers encompasses a large family of schedulers, but suffers from an important lack of determinism. Indeed for a given set of jobs, multiple (and different) schedules can sometimes be derived from the same work-conserving scheduler (and thus from the same job priority assignment). The following example illustrates this drawback.

Example 1: Let us consider the set $J$ of 5 jobs with respective processing time 4, 8, 4, 4 and 6. Suppose that $J$ is scheduled on a 2-processors identical platform $\pi$ by an FTP, global, preemptive and work-conserving scheduler such that $J_1 > J_2 > J_3 > J_4 > J_5$. According to Definition 7, Figures 1 and 2 depict two possible different schedules corresponding to this priority assignment.

Fig. 1: A possible schedule of $J_1, J_2, J_3, J_4$ and $J_5$.

Fig. 2: Another possible schedule of $J_1, J_2, J_3, J_4$ and $J_5$.

In order to get around this lack of determinism, we introduce two refinements of Definition 7 that we name weakly and strongly work-conserving schedulers, respectively. Weakly work-conserving schedulers concern only identical platforms whereas strongly work-conserving schedulers concern only in uniform (and non-identical) platforms. The rationale for introducing these two refinements is to have one and only one possible schedule for any given set of synchronous jobs, multiprocessor platform and job priority assignment.

Definition 8 (Weakly work-conserving scheduler): A scheduler $\mathcal{S}$ is weakly work-conserving if and only if:
• no CPU idles while there are active jobs awaiting execution, and
• if there are more than one job awaiting execution and more than one CPU available for the execution of those jobs then $\mathcal{S}$ assigns the highest priority waiting job to the available CPU with the highest index.

Property 1 (Unique schedule): For any given finite set $J$ of jobs, any weakly work-conserving scheduler $\mathcal{S}$ and any identical multiprocessor platform $\pi$, there exists one and only one possible schedule of $J$ upon $\pi$ following $\mathcal{S}$.

$^6$The term “synchronous” jobs is commonly used in the literature to refer to jobs that are all ready for execution at the same time.
In order to illustrate this property, let us consider the set of 5 jobs used in Example 1, a 2-processors identical platform π and any weakly work-conserving scheduler assigning priorities such that J₁ > J₂ > J₃ > J₄ > J₅. The unique possible schedule of J upon π is the one depicted in Figure 1. Indeed at time 0, CPUs π₁ and π₂ are idle and the second condition of Definition 8 imposes J₁ to execute on π₂. From the same rule, J₄ must execute on π₂ at time 8. Notice that the refinement of “weakly” work-conserving scheduler clarifies the job-to-CPU assignment rule when the highest-priority waiting job has to be dispatched to a CPU.

Definition 9 (Strongly work-conserving scheduler): A scheduler S is strongly work-conserving if and only if:

- no CPU idles while there are active jobs awaiting execution, and
- at every time during the system execution, the job-to-CPU assignment uses the rule: highest priority active job upon highest indexed CPU.

In contrast to the refinement of “weakly” work-conserving schedulers, the “strongly”-refinement clarifies the job-to-CPU assignment rule at each time-instant during the system execution. It is essential to keep in mind that in our study weakly work-conserving schedulers will be used only on identical platforms whereas strongly work-conserving schedulers will be used only on uniform and non-identical platforms. For strongly work-conserving schedulers, the concept of migrating jobs to faster CPUs as soon as possible (as specified by the second condition of Definition 9) has been widely used over the years on uniform platforms (see [20]–[25]). This refinement is extremely important, especially because it yields the following property.

Property 2 (Staircase property): Let J denote any finite set of synchronous jobs, π any uniform multiprocessor platform and S any strongly work-conserving scheduler. In the schedule of J upon π by S, CPU πᵋ idles before or at the same time-instant as CPU πᵋ₊₁ for all ℓ < m.

Informally speaking, the schedule of J upon π by S forms a staircase (see Figure 3).

This property stems from the fact that the CPUs are indexed in such a manner that sᵋ ≥ sᵋ₊₁ ∀i > j. Thus, it holds from the second condition of Definition 9 that at any instant t, if S idles the ℓth-slowest CPU then S also idles the jth-slowest CPUs for all j < i. Also, it results from the same condition that the i th CPU that starts idling is always πᵋ.

The following definition introduces the fundamental notion of predictability, and Lemmas 1 and 2 are essential for the rest of the paper.

Definition 10 (Predictability [26]): Let A denote a scheduler, and let J = {J₁, J₂, J₃, ...} be a potentially infinite set of jobs, where each job Jᵋ = (aᵋ, cᵋ, dᵋ) is characterized by an arrival time aᵋ, a computing requirement cᵋ and an absolute deadline dᵋ. Let rᵋ and fᵋ denote the time at which job Jᵋ starts and completes its execution (respectively) when J is scheduled by A. Now, consider any set J’ = {J’₁, J’₂, J’₃, ...} of jobs obtained from J as follows. Job J’ᵋ has an arrival time aᵋ, an execution requirement cᵋ ≤ cᵋ and a deadline dᵋ. Let rᵋ and fᵋ denote the time at which job J’ᵋ starts and completes its execution (respectively) when J’ is scheduled by A. Algorithm A is said to be predictable if and only if for any set of jobs J and for any such J’ obtained from J, it is the case that rᵋ ≤ rᵋ and fᵋ ≤ fᵋ ∀i.

Informally speaking, Definition 10 claims that an upper-bound on the starting time and on the completion time of each job can be determined by analyzing the situation under the assumption that each job executes for its WCET. The result from [26]–[28] that we will be using can be stated as follows.

Lemma 1 (See [26]–[28]): On identical multiprocessor platforms, any FJP, global, preemptive and weakly work-conserving scheduler is predictable.

In the same vein, the result from [24] that we will be using can be stated as follows.

Lemma 2 (See [24]): On uniform multiprocessor platforms, any FJP, global, preemptive and strongly work-conserving scheduler is predictable.

We use the notation P to refer to a specific job priority assignment. A job priority assignment can be seen as a key component of any scheduler, but the definition of a scheduler is more general since, in addition to a job priority assignment, a scheduler must also provide specifications like “global or partitioned”, “preemptive or non-preemptive”, etc. For any job priority assignment P, we denote by Jᵋ >ₚ Jᵋ the fact that job Jᵋ has a higher priority than Jᵋ according to P, and we assume that every assigned priority is distinct from the others. That is, ∃P, i, j such that i ≠ j we have either Jᵋ >ₚ Jᵋ or Jᵋ <ₚ Jᵋ.

Similarly, and without any distinction with the interpretation given above, we will sometimes use the notations Jᵋ >ₛ Jᵋ and Jᵋ <ₛ Jᵋ where S is the scheduler of mode Mᵋ, and we will sometimes use the notations Jᵋ > Jᵋ and Jᵋ < Jᵋ when the job priority assignment has no label (for instance, when we will depict some examples of schedules, we will just say “Jᵋ > Jᵋ” without giving a name to the job priority assignment). Finally, the problems and solutions presented in this paper are addressed under the following assumptions:

- Assumption 1. The set τᵋ of tasks of every mode Mᵋ can be scheduled by Sᵋ on m CPUs without missing any
**IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS**

Input: $M^i$: the old mode
Input: $M^j$: the new-mode
Input: the rem-jobs

1: while true do
2: Schedule the rem-jobs according to $S^i$
3: if (any rem-job $J_k$ completes at time $t$) then
4: if (active($\tau^i, t) = \phi$) then
5: enable all the new-mode tasks of $\tau^j$
6: enter the new-mode $M^j$
7: end if
8: end if
9: end while

Fig. 4: SM-MSO protocol

deadline.

- **Assumption 2.** Job migrations and preemptions are permitted and are carried out at no loss or penalty.
- **Assumption 3.** Job parallelism is forbidden, i.e., jobs execute on at most one CPU at any instant in time.
- **Assumption 4.** For every mode $M^i$ it holds that $m \leq n_i$, where $n_i$ is the number of tasks in mode $M^i$.

Regarding Assumption 1, it allows us to focus only on the schedulability of the application during the **transient phases** corresponding to mode transitions, rather than on the schedulability of the application during the execution in a given mode.

Regarding Assumption 4, it is worth noticing that since job parallelism is forbidden and tasks are assumed to be constrained-deadline, there are at most $n_i$ jobs active at the same time during the execution of any mode $M^i$. As a result, it holds for each mode $M^i$ that in every schedulable application where $m > n_i$, there are always $m - n_i$ CPUs that constantly idle. We will see later that these $m - n_i$ idling CPUs are the slowest ones and the problem in that case thereby reduces to the same problem upon the subset of the $n_i$ fastest CPUs among these $m$ CPUs.

**III. THE SYNCHRONOUS PROTOCOL SM-MSO**

**A. Description of the protocol**

The protocol SM-MSO (which stands for “Synchronous Multiprocessor Minimum Single Offset protocol”) is an extension to multiprocessor platforms of the protocol MSO defined in [1] for uniprocessor platforms. This protocol supports both uniform and identical platforms. The main idea of SM-MSO is the following: upon a MCR($j$), $\forall j$, all the tasks of the old-mode (say $M^i$) are disabled and the rem-jobs continue to be scheduled by the old-mode scheduler $S^i$ upon the $m$ CPUs. Once all the rem-jobs are completed, all the new-mode tasks (i.e., the tasks of $\tau^j$) are simultaneously enabled. Algorithm 4 gives the pseudo-code of this protocol and Example 2 illustrates how SM-MSO handles the mode transitions.

**Example 2:** Let us consider a platform $\pi$ composed of only 2 identical CPUs and an application composed of 2 modes $M^i$ and $M^j$ depicted in blue and red, respectively. We assume that these two modes contain only synchronous implicit-deadline periodic tasks. The old-mode $M^i$ contains 4 tasks with characteristics given in Table I and uses an FTP scheduler $S^i$ such that $\tau^i_1 > S^i_1$, $\tau^i_2 > S^i_2$, $\tau^i_3 > S^i_3$, $\tau^i_4$.

![Fig. 5: Illustration of a mode transition handled by SM-MSO.](Image)

<table>
<thead>
<tr>
<th>Tasks</th>
<th>$C^i_k$</th>
<th>$D^i_k$ = $T^i_k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau^i_1$</td>
<td>40</td>
<td>120</td>
</tr>
<tr>
<td>$\tau^i_2$</td>
<td>20</td>
<td>120</td>
</tr>
<tr>
<td>$\tau^i_3$</td>
<td>40</td>
<td>120</td>
</tr>
<tr>
<td>$\tau^i_4$</td>
<td>60</td>
<td>120</td>
</tr>
</tbody>
</table>

**TABLE I: Characteristics of the tasks in $M^i$.**

The new-mode $M^j$ contains 3 tasks $\tau^j_1$, $\tau^j_2$, $\tau^j_3$ and uses an FTP scheduler $S^j$ such that $\tau^j_1 > S^j_1$, $\tau^j_2 > S^j_2$, $\tau^j_3$. The characteristics of these tasks are: $C^j_1 = 100$ and $C^j_2 = C^j_3 = 40$. The deadline and period of these new-mode tasks do not have any importance in this example and we intentionally omitted to specify them. Figure 5 illustrates the SM-MSO transition protocol between these two modes.

At time 120, every task of $M^i$ releases its second job and the scheduler $S^i$ starts the execution of $\tau^i_1$, $\tau^i_2$, $\tau^i_3$, and $\tau^i_4$ on CPU $\pi_2$ and $\pi_1$, respectively. Then, suppose that the system requests a mode change at time 130. Here starts the transition phase from mode $M^i$ to mode $M^j$. As specified by the protocol SM-MSO, all the old-mode tasks are immediately disabled and the remaining active jobs $\tau^i_1$, $\tau^i_2$, $\tau^i_3$, and $\tau^i_4$ (named the rem-jobs from this point forward) continue to be scheduled according to the old-mode scheduler $S^i$. These rem-jobs execute until time 220, time at which they are all completed. At this instant 220, the condition at line 4 of Algorithm 4 is verified. Thus, SM-MSO enables all the new-mode tasks and starts scheduling the incoming new-mode jobs according to the new-mode scheduler $S^j$. Notice that at any time during every transition phase, our protocol SM-MSO allows the system (or any task) to request any other mode change. At the very end of the current transition phase (at time 220 in this example), SM-MSO enables all the tasks of the mode $M^j$ assuming that MCR($z$) is the last mode change that has been requested.

**B. Design of a validity test**

In order to establish a **validity test** for the protocol SM-MSO, two key results are required:
1) It must be proved for every mode transition that disabling the old-mode tasks upon a MCR does not jeopardize the schedulability of the rem-jobs when they continue to be scheduled by the old-mode scheduler. That is, it must be guaranteed that the absolute deadline \( d_{a,b} \) of every rem-job \( \tau_{a,b} \) is met during every mode transition from every mode \( M^i \).

2) It must be proved for every mode transition that the length of the transition phase can never be larger than the minimum transition deadline of all new-mode tasks. Indeed, it follows from this statement and the definition of SM-MSO that all the transition deadlines would be met during every mode transition.

We provided a proof for the first key result in [15] (the proof is replicated in Section V, page 12), and this result holds for any uniform platform (including identical platforms). About the second key result, it is worth noticing that there is no job release (and therefore no preemption) during every transition phase since we consider only FJP schedulers and all the old-mode tasks are disabled upon any mode change request. As a consequence, the length of any transition phase corresponds to the time needed to complete all the rem-jobs (this clearly appears in Figure 5). In the literature (and hereafter as well), the time needed to complete a given set of synchronous jobs upon a given platform is called the makespan defined as follows.

**Definition 11 (Makespan):** Let \( J = \{J_1, J_2, \ldots, J_n\} \) denote any set of \( n \) jobs of processing times \( c_1, c_2, \ldots, c_n \). Let \( \pi \) denotes any uniform multiprocessor platform composed of \( m \) CPUs. Let \( P \) denote any job priority assignment and \( S \) denotes the schedule of \( J \) upon \( \pi \) by any work-conserving scheduler (including weakly and strongly work-conserving schedulers) using the priority assignment \( P \). The makespan denoted by \( \text{ms}(J, \pi, P) \) is the earliest instant in \( S \) such that the \( n \) jobs of \( J \) are completed.

According to Definition 11, the length of any transition phase corresponds to the makespan generated by the set of jobs that are active in the system when the mode change is requested, i.e., the set of rem-jobs. Since the value of the makespan obviously depends on the number and processing times of the jobs (as well as on the CPU speeds), then the length of any transition phase from any mode \( M^i \) to any other mode \( M^j \) depends on both the number of rem-jobs and their remaining processing time at time \( t_{MCR(j)} \). From this observation, determining an upper-bound on the makespan requires to consider the worst-case scenario, i.e., the scenario in which the number and the remaining processing time of the rem-jobs at time \( t_{MCR(j)} \) is such that the generated makespan is maximum. This worst-case scenario is thus entirely defined by a specific set of rem-jobs that we name the critical rem-job set defined as follows.

**Definition 12 (Critical rem-job set \( J^{wc} \)):** Assuming any transition from a specific mode \( M^i \) to any other mode \( M^j \), the critical rem-job set \( J^{wc} \) is the set of jobs issued from the tasks of \( \tau^i \) that leads to the largest makespan.

For any work-conserving FJP scheduler (including FTP schedulers) and uniform platform (including identical platform), we will show that the critical rem-job set \( J^{wc} \) of every transition from mode \( M^i \) to mode \( M^j \) is the one where each task \( \tau^i_k \) has a rem-job at time \( t_{MCR(j)} \) with a remaining processing time equals to \( C^j_k \) (i.e., the WCET of \( \tau^i_k \)). This result is very intuitive: the makespan is as large as the number and processing times of the rem-jobs are large.

In this paper we address the problem of establishing mathematical expressions that provide the maximum makespan for any given set of synchronous jobs and especially for the critical rem-job set during each mode transition. This intention stems from the fact that the knowledge of the maximum makespan allows us to assert (or refute) that every new-mode task will meet its transition deadline during any mode transition using SM-MSO, thus ensuring the validity of SM-MSO for a given application \( \tau \) and platform \( \pi \) as follows.

**Validity Test 1 (For protocol SM-MSO):** For any multi-mode real-time application \( \tau \) and any uniform multiprocessor platform \( \pi \), protocol SM-MSO is valid provided that, for every mode \( M^j \),

\[
\text{ms}(J^{wc}, \pi, P^j) \leq \min_{1 \leq k \leq n, j} \left\{ D^j_k(M^i) \right\}
\]

where \( P^j \) is the job priority assignment derived from the old-mode scheduler \( S^j \) and \( \text{ms}(J^{wc}, \pi, P^j) \) is an upper-bound on the makespan, considering the set \( J^{wc} \) of jobs, the platform \( \pi \) and the job priority assignment \( P^j \).

The above expression can be interpreted as follows: all the transition deadlines will be met during the execution of the system if, for every mode \( M^j \), the maximum makespan (i.e., the maximum transition latency) generated by the rem-jobs issued from the tasks of \( \tau^i \) cannot be larger than the minimum transition deadline of every task of mode \( M^j \).

This validity test is a sufficient condition that indicates, a priori, if all the deadlines will be met during all possible mode changes using the protocol SM-MSO. Unfortunately, to the best of our knowledge, the problem of determining the maximum makespan has never been studied in the literature. Rather, authors usually address the problem of determining a job priority assignment that minimizes the makespan [29], [30]. The goal in that framework being to ultimately reduce the completion times of the jobs as much as possible. This problem of finding priorities that minimize the makespan can be cast as a strongly NP-hard bin-packing problem [29], [30] for which numerous heuristics have been proposed in the literature. On the contrary, we provide in Sections VI-IX different upper-bounds on the makespan, assuming in turn identical platforms and FJP schedulers, identical platforms and FTP schedulers, uniform platforms and FJP schedulers and finally, uniform platforms and FTP schedulers.

C. FTP schedulers vs. FJP schedulers

As mentioned in Section II-D, FTP schedulers are a particular case of FJP schedulers. However the remainder of this study distinguishes between these two scheduler families because FTP schedulers allow to determining a more precise
upper-bound $\mathbb{m}(J_i^{\text{wc}}, \pi, P_i^1)$ than FJP schedulers. The reason of this stems from the fact that the priority of each task (and thus the priority of every job) is known at system design-time for FTP schedulers whereas it is unknown beforehand for FJP schedulers. At first blush, assuming that the job priority assignment $P_i$ is unknown for FJP schedulers can seem inconsistent since during every mode transition, we consider the critical rem job set in the computation of $\mathbb{m}(J_i^{\text{wc}}, \pi, P_i^1)$ (and this critical rem-job set is determined at system design-time). Therefore, it could be thought that $P_i$ can simply be derived from $J_i^{\text{wc}}$. But this intuition is erroneous because for a given FJP scheduler, several job priority assignments can be derived from the same critical rem-job set as shown in the following example. Actually, given set of jobs, we are not aware of any job priority assignment leading to the maximum makespan.

Example 3: Let us consider a platform $\pi$ composed of only 2 identical CPUs and an application $\tau$ composed of 2 modes $M_i$ and $M_i^1$. Suppose that a mode change is requested from $M_i$ to $M_i^1$ and the old-mode scheduler $S_i^1$ is EDF. The old-mode $M_i^1$ contains 3 tasks with characteristics given in Table II.

<table>
<thead>
<tr>
<th>Tasks</th>
<th>$C_i^1$</th>
<th>$D_i^1 = T_i^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tau_1^1$</td>
<td>5</td>
<td>15</td>
</tr>
<tr>
<td>$\tau_2^1$</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>$\tau_3^1$</td>
<td>7</td>
<td>18</td>
</tr>
</tbody>
</table>

**TABLE II: Characteristics of the tasks in $M_i^1$.**

As introduced earlier, the critical rem-job set for this mode transition is given by $J_i^{\text{wc}} = \{J_1, J_2, J_3\}$ with processing time $C_i^1, C_i^2$ and $C_i^3$, respectively. This will be formally proved in Corollary 1 (on page 14), assuming any FJP scheduler and any uniform platform. Actually, this critical rem-job set specifies only the processing time of the jobs, not the release time, neither the absolute deadline. Consequently, different job priority assignments can be derived from $J_i^{\text{wc}}$. We depict two of them in Figures 6 and 7. In both figures the time is relative to the instant $t_{\text{MCR}(j)}$ (i.e., $t_{\text{MCR}(j)} = 0$). The release time and the absolute deadline of each job $J_k$ is denoted by $a_k$ and $d_k$, respectively. These two job priority assignments are obtained as follows.

**Job priority assignment 1.** If we assume that the three jobs are released exactly at the MCR invoking time $t_{\text{MCR}(j)}$, i.e., $a_1 = a_2 = a_3 = t_{\text{MCR}(j)}$, then the absolute deadline of each job $J_k$ is given by $d_k = t_{\text{MCR}(j)} + D_i^k$. In Figure 6, the deadline of each job is thus: $d_1 = 15, d_2 = 16$ and $d_3 = 18$ and according to EDF, this leads to the job priority assignment $J_1 >_{\text{EDF}} J_2 >_{\text{EDF}} J_3$ (and to a makespan of 12).

**Job priority assignment 2.** Starting from the previous release pattern in which all the jobs are released simultaneously at time $t_{\text{MCR}(j)}$, one can slightly move backward the release time of job $J_3$ (for instance) in such a manner that $J_3$ is released at time $t_{\text{MCR}(j)} - 5$ (see Figure 7). Its absolute deadline $d_3$ is thus shifted to time $t_{\text{MCR}(j)} + 13$ and since no assumption is made about the schedule before time $t_{\text{MCR}(j)}$, we can suppose that $J_3$ did not execute before $t_{\text{MCR}(j)}$. Therefore, the processing time of $J_3$ at time $t_{\text{MCR}(j)}$ is $C_i^3 = 5$ and the job priority assignment resulting from this new release pattern is $J_3 >_{\text{EDF}} J_2 >_{\text{EDF}} J_1$ (leading to a makespan of 10).

In the particular case of EDF, shifting the absolute deadline of these three jobs by distinct amplitudes can modify their relative priorities and a possibly large number of job priority assignments can be derived from the same critical rem-job set $J_i^{\text{wc}}$.

Because the prior knowledge of the critical rem-job set does not allow determining a unique job priority assignment, FJP schedulers require to consider every possible job priority assignment in order to determine an upper-bound on the makespan. Hence, we refine the notation of $\mathbb{m}(J, \pi, P^1)$ as follows: the upper-bound on the makespan is denoted by $\mathbb{m}(J, \pi, P)$ when $P$ is explicitly specified (in the context of FTP scheduler) and by $\mathbb{m}(J, \pi)$ otherwise (in the context of FJP scheduler), with the interpretation that for every job priority assignment $\mathcal{X}$:

$$\mathbb{m}(J, \pi) \geq \mathbb{m}(J, \pi, \mathcal{X})$$

It goes without saying that the prior knowledge of the jobs priority assignment allows for establishing tighter upper-bounds on the makespan, i.e., the upper-bound $\mathbb{m}(J, \pi, P)$ is tighter than $\mathbb{m}(J, \pi)$. From these refined notations, Expression 2 of Validity Test 1 can be rewritten as

$$\mathbb{m}(J_i^{\text{wc}}, \pi) \leq \min_{j \neq i} \left\{ \min_{1 \leq k \leq n_j} \left\{ D_k^j(M^1) \right\} \right\}$$

for FJP schedulers, and as

$$\mathbb{m}(J_i^{\text{wc}}, \pi, P_i^1) \leq \min_{j \neq i} \left\{ \min_{1 \leq k \leq n_j} \left\{ D_k^j(M_i^1) \right\} \right\}$$

for FTP schedulers, where $P_i^1$ is the job priority assignment derived from the old-mode FTP scheduler $S_i^1$. 

![Fig. 6: Assuming that the three jobs are released simultaneously upon the MCR(j) allows to derive a first job priority assignment.](image)

![Fig. 7: Another job priority assignment can be derived by slightly modifying the release pattern of the jobs. Note that this modification leads to another makespan.](image)
IV. THE ASYNCHRONOUS PROTOCOL AM-MSO

A. Description of the protocol

The protocol AM-MSO (which stands for “Asynchronous Multiprocessor Minimum Single Offset” protocol) is an asynchronous version of the protocol SM-MSO. This protocol supports both uniform and identical platforms. The main idea of this second protocol is to reduce the delay applied to the enablement of the new-mode tasks, by enabling them as soon as possible. In contrast to SM-MSO, rem-jobs and new-mode tasks can be scheduled simultaneously during the transition phases according to the scheduler $S^{\text{trans}}$ defined as follows: (i) the priorities of the rem-jobs are assigned according to the old-mode scheduler; (ii) the priorities of the new-mode jobs are assigned according to the new-mode scheduler, and (iii) the priority of each rem-job is higher than the priority of every new-mode job.

Formally, suppose that the system is transitioning from mode $M^{\text{old}}$ to mode $M^{\text{new}}$ and let $J_i$ and $J_j$ be two active jobs during this transition. According to these notations we have $J_j >_{S^{\text{trans}}} J_i$ if and only if one of the following conditions is satisfied:

\[(J_j \in M^{\text{old}} \text{ and } J_i \in M^{\text{new}})\]
\[\text{ or } (J_j \in M^{\text{old}} \text{ and } J_i \in M^{\text{old}} \text{ and } J_j >_{S^{\text{old}}} J_i)\]
\[\text{ or } (J_j \in M^{\text{new}} \text{ and } J_i \in M^{\text{new}} \text{ and } J_j >_{S^{\text{new}}} J_i)\]

AM-MSO proceeds as follows: upon a MCR(j), ∀j, all the old-mode tasks are disabled and the rem-jobs continue to be scheduled by $S^j$ (assuming that $M^j$ is the old-mode). Whenever any rem-job completes (say at time $t$), if there is no more waiting rem-jobs AM-MSO immediately enables some new-mode tasks, in contrast to SM-MSO which waits for the completion of all the rem-jobs. In order to select the new-mode tasks to enable at time $t$, AM-MSO uses the following heuristic: it considers every disabled new-mode task by non-decreasing order of transition deadline and enables those which can be scheduled by $S^j$ upon the current available CPUs, i.e., the CPUs that are not running a rem-job and are therefore available for executing some new-mode tasks. The following example illustrates how AM-MSO manages mode transitions.

Example 4: Let us consider the same task sets as in Example 2. Figure 8 illustrates the AM-MSO transition protocol on a 2-processors platform.

Similarly to protocol SM-MSO, AM-MSO schedules the rem-jobs according to the old-mode scheduler from time $t_{\text{MCR}(j)} = 130$ to time $t$. Then at time $t$, the rem-job $\tau^j_{2,2}$ completes on CPU $\pi_1$ and there is no more waiting rem-jobs. Here AM-MSO reacts differently from SM-MSO: it scans every disabled task of $\tau^j$ (in non-decreasing order of transition deadline) and enables some of them in such a manner that the resulting set of enabled new-mode tasks can be scheduled by $S^j$ upon 1 CPU (since at this time $t$, only the CPU $\pi_1$ is available for executing some new-mode tasks). We actually have no guarantee that scanning all the disabled tasks in non-decreasing order of transition deadline is optimal, but this heuristic appears as the most intuitive choice. At time 220, AM-MSO performs the same treatment as at time $t$. But since we assumed that every task set $\tau^k$, ∀$k$, is schedulable by $S^k$ on $\pi$, we know that all the remaining disabled new-mode tasks can be enabled at this time 220.

Notice that, in contrast to SM-MSO, the protocol AM-MSO allows mode changes to be requested during the mode transitions only until some new-mode tasks have been enabled (the instant $t$ in Figure 8). Indeed, if the system is transitioning from any mode $M^i$ to any other mode $M^j$ and a mode change is requested to any mode $M^z$ before time $t$, then AM-MSO can consider that the system is transitioning from mode $M^i$ to mode $M^z$ and the new-mode therefore becomes the mode $M^z$. However after time $t$, some tasks of mode $M^j$ have already been enabled and AM-MSO does not allow the system to request any other mode change until the end of the transition phase from $M^i$ to $M^j$, i.e., until all the tasks of mode $M^j$ are enabled.

In order to determine whether a task can be safely enabled, protocol AM-MSO uses a binary function $\text{sched} (\pi, S, \tau^j)$ that returns True if and only if the task set $\tau^j$ is schedulable by $S$ upon $\pi$. This function is essential as we must always guarantee that all the deadlines are met for all the jobs in the system, including the deadlines of all the new-mode jobs. Considering a specific scheduler $S$, such a function can be derived from schedulability tests proposed for $S$ in the literature\(^8\). Algorithm 9 provides a pseudo-code for protocol AM-MSO.

\(^8\)To the best of our knowledge, there is no efficient necessary and sufficient schedulability test for any multiprocessor scheduler that complies with the requirements specified in Section II-D. Theodore Baker has proposed in [31] a necessary and sufficient schedulability test for arbitrary-deadline sporadic tasks scheduled by Global-EDF but its time-complexity is very high so only small applications can be tested. Fortunately, many sufficient schedulability tests have been proposed for scheduler such as Global-EDF (see for instance [23], [32]–[35]) and Global-DM (see for instance [20], [36], [37]).
Input: $M^i$: the old mode
Input: $M^j$: the new-mode
Input: the rem-jobs
Input: $t$: the current time during the transition
Input: $\pi$: the platform (uniform or identical)

1: if ($t$ is the MCR invoking time) then
2: Disable all the tasks of $\tau^i$
3: Sort the task set “disabled($\tau^j$, $t$)” by non-decreasing order of transition deadlines
4: $\pi^{av\ell} \leftarrow \emptyset$
5: end if
6: Schedule the rem-jobs according to $S^{trans}$
7: if (any rem-job $J_k$ completes at $t$ on any CPU $\pi_k$) then
8: $r \leftarrow$ number of active rem-jobs at time $t$
9: if ($r < m$) then
10: /* Due to the completion of $J_k$, one CPU $\notin \pi^{av\ell}$ becomes available. */
11: if ($\pi$ is identical) then
12: /* The scheduler is weakly work-conserving. Thus, the CPU that becomes available is $\pi_k$ */
13: $\pi^{av\ell} \leftarrow \pi^{av\ell} \cup \{\pi_k\}$
14: else
15: /* The scheduler is strongly work-conserving. Thus, the CPU that becomes available is the $(m-r)^{th}$ slowest CPU. */
16: $\pi^{av\ell} \leftarrow \pi^{av\ell} \cup \{\pi_{m-r}\}$
17: end if
18: end if
19: for each $\tau^j \in$ disabled($\tau^j$, $t$) do
20: $\tau^{temp} \leftarrow$ enabled($\tau^j$, $t$) $\cup \{\tau^j\}$
21: if (sched($\pi^{av\ell}$, $S^j$, $\tau^{temp}$)) then
22: enable $\tau^j$
23: end if
24: end for
25: if ($r = 0$) then
26: enter the new-mode $M^j$
27: else
28: Schedule all the rem-jobs and new-mode jobs according to $S^{trans}$
29: end if
30: end if

Fig. 9: AM-MSO protocol

### B. Design of a validity test

For a given application $\tau$ and platform $\pi$, the main idea to determine whether AM-MSO allows to meet all the transition deadlines is to run Algorithm 9 for every possible mode transition, while considering the worst-case scenario for each one—the scenario in which the new-mode tasks are enabled as late as possible. From our definition of protocol AM-MSO, we know that every instant at which some new-mode tasks are enabled corresponds to an instant at which at least one CPU has no more rem-job to execute, i.e., an “idle-instant” defined as follows.

**Definition 13 (Idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$):** Let $J = \{J_1, J_2, \ldots, J_n\}$ be any finite set of $n$ synchronous jobs. Let $\pi$ be a uniform multiprocessor platform and let $\mathcal{P}$ be the job priority assignment used during the schedule of $J$ upon $\pi$. If $S$ denotes that schedule then the idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (with $k = 1, \ldots, m$) is the earliest instant in $S$ such that at least $k$ CPUs idle.

By definition of the protocol AM-MSO, and in particular from the definition of $S^{trans}$, a new-mode job never preempts a rem-job during the transition phases. Thereby, during every transition phase, new-mode tasks are enabled at each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ ($\forall k = 1, \ldots, m$) where $J$ is the set of rem-jobs at the MCR invoking time and $\mathcal{P}$ is the job priority assignment derived from the old-mode scheduler when the mode change is requested. For obvious reasons, the exact values of these idle-instants depend on both the number of jobs in $J$ and their actual execution times. Therefore, these exact value cannot be determined at system design-time and the main idea of our validity test is the following.

**First**, for every mode $M^j$ we determine the set $J$ of rem-jobs that leads to the largest idle-instants $\text{idle}_k(J, \pi, \mathcal{P})$ ($\forall k \in [1, m]$). From this point forward, we thus refine the definition of the critical rem-job set as follows.

**Definition 14 (Critical rem-job set $J^{wc}_{\pi}$):** Assuming any transition from a specific mode $M^i$ to any other mode $M^j$, the critical rem-job set $J^{wc}_{\pi}$ is the set of jobs issued from the tasks of $\tau^i$ that leads to the largest idle-instants.

As it will be shown in Corollary 1 (page 14), the critical rem-job set $J^{wc}_{\pi}$ of every mode $M^j$ is the one that contains one job $J_{\tau^j}$ for each task $\tau^j$ and such that every job $J_{\tau^j}$ in $J^{wc}_{\pi}$ has a processing time equals to $C_{\tau^j}$, i.e., the WCET of $\tau^j$. Informally speaking, the worst-case scenario during any mode transition is the one in which (i) every old-mode task releases a job exactly when the mode change is requested and (ii) every released job executes for its WCET.

**Second**, we determine (for any given set $J$ of jobs) an upper-bound on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (for $k = 1, 2, \ldots, m$). As in the previous section (and for the same reason), we distinguish between FTP and FJP schedulers. That is, for FTP schedulers we focus on determining an upper-bound $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$ (for $k = 1, 2, \ldots, m$) assuming that the job priority assignment $\mathcal{P}$ is known beforehand, whereas for FJP schedulers, we determine an upper-bound $\overline{\text{idle}}_k(J, \pi)$ on each idle-instant $\text{idle}_k(J, \pi, \mathcal{P})$, with the interpretation that for every job priority assignment $\mathcal{X}$:

$$\overline{\text{idle}}_k(J, \pi) \geq \text{idle}_k(J, \pi, \mathcal{X})$$

Finally, we simulate Algorithm 9 at each of these upper-bounds. That is, we verify whether all the transition deadlines are met while enabling the new-mode tasks at each instant $\overline{\text{idle}}_k(J^{wc}_{\pi}, \pi, \mathcal{P})$ (or $\overline{\text{idle}}_k(J^{wc}_{\pi}, \pi)$ depending on the family of the old-mode scheduler). Obviously, if every transition deadline is met during this simulation then it will be met during the actual execution of the application.

It goes without saying that the prior knowledge of the jobs priority assignment allows for establishing tighter upper-bounds on the idle-instants, i.e., the upper-bounds $\overline{\text{idle}}_k(J, \pi, \mathcal{P})$ are tighter than $\overline{\text{idle}}_k(J, \pi)$. Notice that it results...
from these notations that $\overline{idle}_m^k(J, \pi)$ and $\overline{idle}_m^k(J, \pi, P)$ correspond to the upper-bounds $\overline{ms}(J_{s}^{wc}, \pi)$ and $\overline{ms}(J_{s}^{wc}, \pi, P')$ introduced in Validity Test 1, respectively.

Mathematical expressions of these upper-bounds $\overline{idle}_k^i(J_{s}^{wc}, \pi)$ and $\overline{idle}_k^i(J_{s}^{wc}, \pi, P')$ on the $k^{th}$ idle-instants are defined for both identical and uniform platforms in Sections VI–IX. Algorithm 10 provides details on the validity test for AM-MSO, where the upper-bounds $\overline{idle}_k^i(J_{s}^{wc}, \pi)$ must be replaced with $\overline{idle}_k^i(J_{s}^{wc}, \pi, P')$ at line 9 if the old-mode scheduler is FTP.

**Input: $\tau = \{\pi^1, \pi^2, \ldots, \pi^x\}$**

1: for (all $i, j \in [1, x]$ such that $i \neq j$) do
2: $\tau^{\text{disabled}} \leftarrow \tau^j$
3: $\tau^{\text{enabled}} \leftarrow \emptyset$
4: $\pi^{\text{avl}} \leftarrow \emptyset$
5: Sort $\tau^{\text{disabled}}$ by non-decreasing order of transition deadlines
6: for ($k = 1; k < m; k++)$ do
7: $\pi^{\text{avl}} \leftarrow \pi^{\text{avl}} \cup \pi_k$
8: for (all $\tau^j \in \tau^{\text{disabled}}$) do
9: if ($D^i_k(M') < \overline{idle}_k^i(J_{s}^{wc}, \pi)$) then
10: return $\text{false}$
11: end if
12: if (sched($\pi^{\text{avl}}$, $\tau^j$, $\tau^{\text{enabled}} \cup \{\tau^j\}$)) then
13: $\tau^{\text{enabled}} \leftarrow \tau^{\text{enabled}} \cup \{\tau^j\}$
14: $\tau^{\text{disabled}} \leftarrow \tau^{\text{disabled}} \setminus \{\tau^j\}$
15: end if
16: end for
17: end for
18: return $\text{true}$

![Fig. 10: Validity Test for AM-MSO](image)

Notice that Algorithm 10 enables new-mode tasks only at the instants $\overline{idle}_k^i(J_{s}^{wc}, \pi)$ (with $k = 1, 2, \ldots, m$). That is, it implicitly considers that every instant at which CPUs become available to the new-mode tasks are as late as possible. As a consequence, if all the transition deadlines are met while running Algorithm 10 then all these deadlines will be met during every transition phase at run-time\(^9\). Nevertheless, the fact that Algorithm 10 simulates every idle-instant of every mode transition by its corresponding upper-bound $\overline{idle}_k^i(J_{s}^{wc}, \pi)$ brings about the following situation: during the actual execution of the application, there could be some intervals of time (during any mode transition) during which the set of currently enabled new-mode tasks benefits from more (and faster) CPUs than during the execution of Algorithm 10. This kind of situation can occur upon identical and uniform platforms and for both FJP and FTP schedulers as shown in the following example.

**Example 5:** Let us consider a 5-processors uniform platform $\pi$ and a system which is transitioning from mode $M^1$ to mode $M^2$. Other details such as the CPU speeds, the characteristics of the jobs and the job priority assignment are not relevant in the scope of this example. Figures 11 and 12 illustrate a situation where during some intervals of time the set of currently enabled new-mode tasks benefits from more (and faster) CPUs than during the execution of Algorithm 10.

![Fig. 11: Illustration of the schedule assumed by the execution of Algorithm 10.](image)

For sake of clarity, Figure 12 uses the notations $\overline{idle}_k^i$ and $\overline{idle}_k^i$ instead of $\overline{idle}_k^i(J_{s}^{wc}, \pi)$ and $\overline{idle}_k^i(J_{s}^{wc}, \pi)$, respectively. In this latter schedule, there can be less rem-jobs and/or rem-jobs with lower processing times than in the schedule of Figure 11 since the schedule of Figure 11 is drawn while assuming the critical rem-job set of mode $M^1$. This is the reason why the schedule of Figure 12 seems less “loaded” than the one of Figure 11. Due to the fact that (i) the validity test provided by Algorithm 10 uses the same function sched($\pi$, $S$, $\tau$) as protocol AM-MSO at run-time and (ii) this function sched($\pi$, $S$, $\tau$) is independent of the current time, we know that the set of tasks enabled at each instant $\overline{idle}_k^i$ ($k = 1, 2, \ldots, m$) in Figure 12 is the same as the set of tasks enabled at each instant $\overline{idle}_k^i$ in Figure 11. Let us temporarily name this property the “equivalence property”. Let $\tau_{(k)}$ temporarily denote the set of tasks enabled at time $\overline{idle}_k^i$, $\forall k \in [1, m]$ and suppose that at time $\overline{idle}_k^i$ in Figure 11 some tasks are enabled (i.e., $\tau_{(3)} \neq \emptyset$) and at time $\overline{idle}_k^i$ no task
is enabled, i.e., \( \tau(4) = \phi \). Thanks to the equivalence property, we know that the tasks enabled at time \( \text{idle}_3 \) in Figure 12 are the tasks of \( \tau(3) \) and those enabled at time \( \text{idle}_4 \) are the tasks of \( \tau(4) \). Since we assumed in Figure 12 that \( \text{idle}_3 = \text{idle}_4 \), it holds that the tasks enabled at time \( \text{idle}_3 \) are the tasks of \( \tau(3) \cup \tau(4) = \tau(3) \) (since \( \tau(4) = \phi \)). It follows that in the time interval \([\text{idle}_3, \text{idle}_4] \), only 3 CPUs are available to the task set \( \tau(1) \cup \tau(2) \cup \tau(3) \) in Figure 11 while 4 CPUs are available to this task set in Figure 12. Moreover, during this time interval, the additional CPU \( \pi_4 \) in Figure 12 is faster (or of equal speed) than every CPU in the subset of CPUs \( \{\pi_1, \pi_2, \pi_3\} \) available to \( \tau(1) \cup \tau(2) \cup \tau(3) \) in Figure 11.

Lemma 3 proves that this kind of situation does not jeopardize the schedulability of the application during its execution.

**Lemma 3 (See [13]):** Any strongly work-conserving scheduler that is able to schedule a task set \( \tau \) upon a uniform platform \( \pi = [s_1, \ldots, s_m] \) is also able to schedule \( \tau \) upon any uniform platform \( \pi^* \) such that (i) \( \pi^* \supseteq \pi \) and (ii) \( \forall \pi_k \in \pi^* \) and \( \pi_k \notin \pi \) we have \( s_k \geq s_m \).

**Proof:** To obtain the proof, it is sufficient to show the lemma for \( \pi^* = [s_1, \ldots, s_m, s_{m+1}] \) where \( s_{m+1} \geq s_m \). The proof is made by contradiction. Suppose there exists a task set \( \tau \) that is schedulable by a strongly work-conserving scheduler \( S \) upon \( \pi \), but not upon \( \pi^* \supseteq \pi \). Consider the schedule upon \( \pi^* \) of a particular set \( J \) of jobs issued from \( \tau \) that leads to a deadline miss, and let \( J^* \) be another set of jobs derived from \( J \) by reducing the processing time of each job \( J_i \) by the amount of time \( J_i \) executes upon the sub-platform \( \pi^* \setminus \pi \), i.e., upon \( \pi_{m+1} \). Since the scheduler is strongly work-conserving, the schedule of \( J \) by \( S \) upon the CPUs in common with \( \pi \) is the same as the one that would be produced by \( S \) for \( J^* \) upon platform \( \pi \). Since a deadline is missed in the schedule of \( J \) upon \( \pi^* \), then a deadline is also missed in the schedule of \( J^* \) upon \( \pi \). But since the schedule is predictable from Lemma 2, a deadline will be missed on \( \pi \) even (a fortiori) with the more demanding jobs set \( J \), leading to a contradiction. The lemma follows.

Lemma 3 is proved while considering uniform platforms and strongly work-conserving schedulers but one can easily show that it also holds for identical platforms and weakly work-conserving schedulers.

**V. SOME BASIC RESULTS FOR DETERMINING VALIDITY TESTS**

**A. Introduction to the three required key results**

Three key results are required to establish a validity test for SM-MSO and AM-MSO.

**Key Result 1:** It must be proved that disabling the old-mode tasks upon any MCR does not jeopardize the schedulability of the rem-jobs when they continue to be scheduled by the old-mode scheduler. That is, it must be guaranteed that the absolute deadline \( d'_{a,b} \) of every rem-job \( \tau'_{a,b} \) is met during any mode transition from every mode \( M^i \).

**Key Result 2:** The critical rem-job set \( J'_{\text{wc}} \) for every mode \( M^i \) must be determined. Indeed, for every mode transition from mode \( M^i \) to any other mode \( M^j \), our validity test (see Algorithm 10) determines the upper-bounds on the idle-instants by basing the computations on the corresponding critical rem-job sets \( J'_{\text{wc}} \) (at line 10). In all cases (i.e., identical or uniform platforms and FJP or FTP schedulers), we will provide a proof that the critical rem-job set \( J'_{\text{wc}} \) of every mode \( M^i \) is the one that contains one job \( J_i \) for each task \( \tau_j \) and such that every job \( J_i \in J'_{\text{wc}} \) has a processing time equals to \( c'_{i,j} \), i.e., the WCET of the task \( \tau_j \).

**Key Result 3:** A mathematical expression must be established that provides, for any given set \( J \) of jobs and platform \( \pi \):

1) an upper-bound \( \bar{\text{idle}}_{\text{c}}(J, \pi) \) (1 \( \leq k \leq m \)) on each idle-instant \( \text{idle}_k(J, \pi, \mathcal{X}) \), for every job priority assignment \( \mathcal{X} \). This concerns FJP schedulers.

2) an upper-bound \( \bar{\text{idle}}_{\text{c}}(J, \pi, \mathcal{P}) \) (1 \( \leq k \leq m \)) on each idle-instant \( \text{idle}_k(J, \pi, \mathcal{P}) \), for a specific job priority assignment \( \mathcal{P} \). This concerns FTP schedulers.

Note that the protocol SM-MSO requires only an upper-bound on the makespan, i.e., on the \( m^\text{th} \) idle-instant \( \bar{\text{idle}}_{\text{c}}(J, \pi) \) and \( \bar{\text{idle}}_{\text{c}}(J, \pi, \mathcal{P}) \).

**B. Proof of the first key result**

Lemma 4 proves the first key result introduced above for any uniform platform and strongly work-conserving scheduler, as well as any identical platform and weakly work-conserving scheduler. This result, which is essential to the validity tests of both protocols SM-MSO and AM-MSO, is based on the notion of predictability introduced on page 5. It has been drawn from [15] and extended to uniform platforms.

**Lemma 4:** Let \( M^i \) and \( M^j \) denote two distinct modes of the application. If the application is running in mode \( M^i \) and a MCR(\( j \)) occurs at time \( t_{\text{MCR}(j)} \) then every rem-job meets its deadline during the transition phase while being scheduled by the old-mode scheduler \( S^i \).

**Proof:** From our first assumption on page 5, the set of tasks \( \tau^i \) of the mode \( M^i \) is schedulable by \( S^i \) upon \( \pi \). When the MCR(\( j \)) is invoked at time \( t_{\text{MCR}(j)} \), the transition protocol disables every old-mode task, which is equivalent to set the processing time of all their future jobs to zero. Since \( S^i \) is predictable (from Lemma 1 or 2 depending on the scheduler family), the deadline of every rem-job is still met in the produced schedule. The lemma follows.

**C. Proof of the second key result**

Corollary 1 proves the second key result introduced above for any uniform platform and strongly work-conserving FTP (or FJP) scheduler, as well as any identical platform and weakly work-conserving FTP (or FJP) scheduler. It has been drawn from the following Lemma 5.

**Lemma 5:** Let \( \pi \) be any uniform multiprocessor platforms (including identical platforms) and let \( J \) and \( J' \) be any fixed set of \( n \) synchronous jobs such that \( J = \{J_1, J_2, \ldots, J_n\} \) of processing times \( c_1, c_2, \ldots, c_n \) and \( J' = \{J'_1, J'_2, \ldots, J'_n\} \) of processing times \( c'_1, c'_2, \ldots, c'_n \). For any job priority assignment \( \mathcal{P} \), if there exists a bijective function between \( J \) and \( J' \) such that every job \( J'_r \in J' \) is mapped to exactly one
In both Figures 13 and 14, we voluntarily omit the details about the CPU speeds, the jobs characteristics, etc. since they are useless in the scope of these examples.

Similarly, Figures 15 and 16 illustrate an example of schedules $S$, $S'$ on a 5-processors identical platform, respectively, where $idle_3(J, \pi, P) < idle_3(J', \pi, P)$. Since the platform is identical in these examples, the scheduler is assumed to be weakly work-conserving. Furthermore, note that in both examples no job is released after time 0.

By definition of the idle-instants, the schedule of any set $J$ of jobs upon any uniform or identical multiprocessor platform is such that $\forall k \in [1, m]$:

- the idle-instant $idle_k(J, \pi, P)$ corresponds to the completion time of a job,
- there is no waiting job at time $idle_k(J, \pi, P)$ and,
- there are at most $(m - k)$ running jobs at time $idle_k(J, \pi, P)$. “At most” since there can exist some $r > k$ such that $idle_r(J, \pi, P) = idle_k(J, \pi, P)$.

Since every idle-instant corresponds to the completion of a job, this implies that within the time interval $[idle_3(J, \pi, P), idle_3(J', \pi, P)]$ there are at most $(m - \ell)$ running jobs in $S$ while there are at least $(m - \ell + 1)$ running jobs in $S'$. Therefore, within $[idle_3(J, \pi, P), idle_3(J', \pi, P)]$, at least one job (say $J_r$) is already completed in $S$ while $J_r'$ is still running in $S'$. The fact that $J_r'$ completes later in $S'$ than $J_r$ in $S$ leads to a direct contradiction of Inequality 3.

As we can see in Figures 14 and 16, three jobs are running in $S'$ during the time interval $[idle_3(J, \pi, P), idle_3(J', \pi, P)]$ while only two jobs are running in $S$, meaning that there is
one job which is completed in \( S \) and still running in \( S' \). The lemma follows.

**Corollary 1:** For any uniform multiprocessor platforms \( \pi \) and for any transition of the system from mode \( M' \) to mode \( M \), let \( J^{\text{any}} \) denote any set of rem-jobs issued from the old-mode tasks and let \( J^{\text{wc}} \) be the set of rem-jobs that contains one job \( J_t \) for each task \( \tau^i_t \) and such that every job \( J_t \in J^{\text{wc}} \) has a processing time equals to \( C^i_t \). The \( k \)th idle-instants \( \text{idle}_{k}(J^{\text{wc}}, \pi) \) (\( \forall k \in [1, m] \)) in the schedule of \( J^{\text{wc}} \) is never lower than the \( k \)th idle-instance \( \text{idle}_{k}(J^{\text{any}}, \pi) \) in the schedule of \( J^{\text{any}} \), i.e., it holds \( \forall k \in [1, m] \) that
\[
\text{idle}_{k}(J^{\text{any}}, \pi) \leq \text{idle}_{k}(J^{\text{wc}}, \pi)
\]

**Proof:** The proof is a consequence of Lemma 5. Let \( c^{\text{any}}_{\tau^i_t} \) and \( c^{\text{wc}}_{\tau^i_t} \) denote the processing time of job \( J_t \) in \( J^{\text{wc}} \) and \( J^{\text{any}} \), respectively. By definition, \( J^{\text{wc}} \) contains one job \( J_t \) of processing time \( C^i_t \) for each task \( \tau^i_t \in \tau^i \), i.e., it holds \( \forall \tau^i_t \in \tau^i \) that
\[
c^{\text{wc}}_{\tau^i_t} = C^i_t
\]
and thus we know by definition of \( J^{\text{any}} \) that \( \forall J_t \in J^{\text{any}}, \)
\[
c^{\text{any}}_{\tau^i_t} \leq c^{\text{wc}}_{\tau^i_t}
\]
In addition, we know that there could be some jobs \( J_t \in J^{\text{wc}} \) such that \( J_t \notin J^{\text{any}} \) (since \( J^{\text{any}} \) does not necessarily contain one job for each old-mode task). For each such job \( J_t \) we add a fake job \( J'_t \) in \( J^{\text{any}} \) with \( c^{\text{any}}_{\tau^i_t} = 0 \). It results from this operation that the number of jobs in both \( J^{\text{wc}} \) and \( J^{\text{any}} \) are the same (we denote this number by \( n \)) and there is a bijective function between \( J^{\text{wc}} \) and \( J^{\text{any}} \) such that every job \( J_t \in J^{\text{wc}} \) is mapped to exactly one job \( J_t \in J^{\text{any}} \) and such that \( c^{\text{any}}_{\tau^i_t} \leq c^{\text{wc}}_{\tau^i_t} \). Thanks to this bijection, we know from Lemma 5 that \( \forall r \in [1, m] \) we have
\[
\text{idle}_{r}(J^{\text{any}}, \pi) \leq \text{idle}_{r}(J^{\text{wc}}, \pi)
\]
and the corollary follows.

By definition, for every mode transition from any mode \( M^i \) upon \( \pi \), each \( \text{idle}_{k}(J^{\text{wc}}, \pi) \) is an upper-bound on the \( k \)th idle-instant in the schedule of \( J^{\text{wc}} \) (this also holds for each upper-bound \( \text{idle}_{k}(J^{\text{wc}}, \pi, P) \) if the job priority assignment \( P \) is known beforehand). Thanks to Corollary 1, we are now aware that each upper-bound \( \text{idle}_{k}(J^{\text{wc}}, \pi, P) \) (and \( \text{idle}_{k}(J^{\text{wc}}, \pi, P) \)) is also an upper-bound on the \( k \)th idle-instance in the schedule of \( J^{\text{wc}} \) other set of rem-jobs issued from the old-mode tasks (i.e., the tasks of \( \tau^i \)). That is, for every mode transition from any mode \( M^i \) we have \( \forall k \in [1, m] \):
\[
\text{idle}_{k}(J^{\text{wc}}, \pi, P) \geq \text{idle}_{k}(J^{\text{wc}}, \pi) \geq \text{idle}_{k}(J^{\text{any}}, \pi, P) \geq \text{idle}_{k}(J^{\text{any}}, \pi)
\]
where \( J^{\text{any}} \) denotes any set of rem-jobs issued from the tasks of \( \tau^i \). As a result, the instants \( \text{idle}_{k}(J^{\text{wc}}, \pi, P) \) (and \( \text{idle}_{k}(J^{\text{wc}}, \pi, P, \pi) \)), with \( k = 1, 2, \ldots, m \), can be considered as the largest instants at which new-mode tasks are enabled during every transition from mode \( M^i \) and thus, these instants can be used in our validity test given by Algorithm 10.

**D. Organization for the third key result**

The third key result consists in determining a mathematical expression for each upper-bound \( \text{idle}_{k}(J, \pi) \) (or \( \text{idle}_{k}(J, \pi, P) \)) depending on the scheduler family, i.e., FJP or FTP, for all \( 1 \leq k \leq m \). Depending on the type of the platform (uniform or identical) and on the scheduler family (FJP or FTP), we distinguish between four different cases that are studied in turn in the following four sections. More precisely:

- Section VI addresses the identical and FJP case.
- Section VII addresses the identical and FTP case.
- Section VIII addresses the uniform and FJP case.
- Section IX addresses the uniform and FTP case.

Recall that the protocol SM-MSO requires only an upper-bound on the makespan, i.e., on the \( m \)th idle-instant. The organization for the third key result is as follows.

**VI. IDENTICAL PLATFORMS AND FJP SCHEDULERS**

This section is organized as follows. First, Section VI-A determines an upper-bound \( \text{idle}_{k}(J, \pi) \) on the earliest time-instant where at least \( k \) CPUs are idle and derives an upper-bound \( \text{ms}(J, \pi) \) on the maximum makespan. Then, Section VI-B shows that this upper-bound \( \text{ms}(J, \pi) \) is 2-competitive, with the interpretation that \( \text{ms}(J, \pi) \) is at most twice the exact value of the maximum makespan. Finally, Section VI-C establishes a sufficient validity test for protocols SM-MSO and AM-MSO.

**A. Upper-bounds \( \text{idle}_{k}(J, \pi) \) on the idle-instants**

Throughout this section, \( J \) refers to any set of \( n \) jobs. For sake of clarity, we will use the notation \( \text{idle}_{k} \) instead of \( \text{idle}_{k}(J, \pi) \) and similarly, we will use the notation \( \text{idle}_{k} \) to denote the exact value of the \( k \)th idle-instant. Before introducing the computation of these upper-bounds \( \text{idle}_{k} \), \( 1 \leq k \leq m \), let us introduce the following result taken from [15].

**Lemma 6 (See [15]):** Suppose that \( J \) is sorted by non-decreasing job processing times, i.e., \( c_1 \leq c_2 \leq \cdots \leq c_n \). Then, whatever the job priority assignment we have \( \forall j, k \in [1, m] \) such that \( j < k \):
\[
\text{idle}_{j} \geq \text{idle}_{k} - c_{n-m+k}
\]
Based on this Lemma 6, the following result was proved in our previous work [15].

**Lemma 7 (See [15]):** Suppose that \( J \) is sorted by non-decreasing job processing times, i.e., \( c_1 \leq c_2 \leq \cdots \leq c_n \). Then, whatever the job priority assignment, an upper-bound \( \text{idle}_{k} \) on the idle-instant \( \text{idle}_{k} \), \( 1 \leq k \leq m \), is given by \( c_k \) if \( n = m \) or by
\[
\text{idle}_{k} = \max_{1 \leq i \leq m} \left\{ \sum_{j=1}^{n} c_{j} - \sum_{j=i+1}^{i+m-k+1} c_{j} + \sum_{j=m+1}^{m-k+1} c_{j} \right\} 
\]
otherwise \( n > m \).

Holding this result, we improve here this previous analysis by (i) successfully establishing another upper-bound \( \text{idle}_{k}(J, \pi) \) on each idle-instant \( \text{idle}_{k}(J, \pi) \) and (ii) proving that these alternative upper-bounds are always tighter than
those proposed in Lemma 7. In short, we complete our previous work [15] as follows.

- Lemma 8 shows that Expression 4 of $\text{idle}_k$ is always maximal for $i = n - m + k - 1$.
- Lemma 9 proposes another upper-bound $\text{idle}^\text{new}_k(J, \pi)$ on each idle-instant $\text{idle}_k$.
- Lemma 10 shows that these alternative upper-bounds $\text{idle}^\text{new}_k(J, \pi)$, $\forall k \in [1, m]$, are never larger than those provided by Expression 4.
- Finally, based on these alternative upper-bounds, Corollary 2 derives an upper-bound on the makespan.

**Lemma 8:** If $n > m$, Expression 4 is maximal for $i = n - m + k - 1$.

**Proof:** This result is presented in Lemma 2.10 in [38]. Due to the space limitation and because the proof is simply based on algebra, we do not repeat it here. □

Thanks to Lemma 8, Expression 4 can be rewritten as follows: $\text{idle}^\text{new}_k \overset{\text{def}}{=} c_k$ if $n = m$ or

$$\text{idle}^\text{old}_k \overset{\text{def}}{=} \sum_{j=1}^{n} c_j - \sum_{j=n-m+k}^{n} c_j + \sum_{j=n-m+k}^{n} c_j \cdot \frac{m-k+1}{m}$$  \hspace{1cm} (5)

otherwise ($n > m$).

**Lemma 9:** Suppose that $J$ is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \cdots \leq c_n$. Then, whatever the job priority assignment, an upper-bound $\text{idle}^\text{old}_k$ on the idle-instant $\text{idle}_k$, $1 \leq k \leq m$, is given by $\text{idle}^\text{old}_k \overset{\text{def}}{=} c_k$ if $n = m$ or by

$$\text{idle}^\text{old}_k \overset{\text{def}}{=} \sum_{i=1}^{n} c_i + (k-1) \cdot c_{n-m+k}$$ \hspace{1cm} (6)

otherwise ($n > m$).

**Proof:** The case where $n = m$ is obvious. Otherwise, the proof is made by contradiction. Suppose that there exists $k \in [1, m]$ such that $\text{idle}_k > \text{idle}^\text{old}_k$. The following properties hold:

- **Prop. (a):** $\forall j > k$: $\text{idle}_j \geq \text{idle}_k$ (by definition of the idle-instants).
- **Prop. (b):** $\forall j < k$: $\text{idle}_j \geq \text{idle}_k - c_{n-m+k}$ (from Lemma 6).

The proof starts with this obvious equality:

$$\sum_{j=1}^{m} \text{idle}_j = \sum_{j=1}^{k-1} \text{idle}_j + \sum_{j=k}^{m} \text{idle}_j$$

Then, applying properties (a) and (b) to the right-hand side yields

$$\sum_{j=1}^{m} \text{idle}_j \geq \sum_{j=1}^{k-1} (\text{idle}_k - c_{n-m+k}) + \sum_{j=k}^{m} \text{idle}_k \geq (k-1)(\text{idle}_k - c_{n-m+k}) + \sum_{j=k}^{m} \text{idle}_k \geq m \cdot \text{idle}_k - (k-1) \cdot c_{n-m+k}$$

Since by hypothesis $\text{idle}_k > \text{idle}^\text{old}_k$, replacing $\text{idle}_k$ with $\text{idle}^\text{old}_k$ in the above inequality leads to

$$\sum_{j=1}^{m} \text{idle}_j > m \cdot \text{idle}^\text{old}_k - (k-1) \cdot c_{n-m+k}$$

$$> m \left( \sum_{i=1}^{n} c_i + (k-1) \cdot c_{n-m+k} \right) \cdot \frac{m}{m} - (k-1) \cdot c_{n-m+k}$$

$$> \sum_{j=1}^{n} c_i$$

This leads to a contradiction since it obviously holds by definition of the idle-instants that $\sum_{j=1}^{m} \text{idle}_j = \sum_{i=1}^{n} c_i$. The lemma follows. □

**Lemma 10:** The upper-bounds $\text{idle}^\text{new}_k$ (with $k = 1, 2, \ldots, m$) provided by Expression 6 are never larger than those provided by Expression 5.

**Proof:** The proof is made by contradiction. Let $k$ be any integer in $[1, m]$. Let $\text{idle}^\text{old}_k$ and $\text{idle}^\text{new}_k$ denote the upper-bound provided by Expressions 5 and 6, respectively, and suppose that $\text{idle}^\text{new}_k > \text{idle}^\text{old}_k$. From Expressions 5 and 6 we get

$$\sum_{j=1}^{m} c_j + (k-1) \cdot c_{n-m+k} \geq \sum_{j=1}^{n} c_j - \sum_{j=n-m+k}^{n} c_j + \sum_{j=n-m+k}^{n} c_j \cdot \frac{m-k+1}{m}$$

By multiplying both sides by $m \cdot (m - k + 1)$ we get

$$(m-k+1) \cdot \left( \sum_{j=1}^{n} c_j + (k-1) \cdot c_{n-m+k} \right) \geq (m-k+1) \cdot \left( \sum_{j=1}^{n} c_j - \sum_{j=n-m+k}^{n} c_j \right) + m \sum_{j=n-m+k}^{n} c_j$$

Thus,

$$(m-k+1) \cdot (k-1) \cdot c_{n-m+k} > (k-1) \cdot \sum_{j=n-m+k}^{n} c_j$$

If $k = 1$ then we obviously get $0 > 0$ and the lemma follows. Otherwise, if $k > 1$ then dividing both sides by $(k-1)$ yields

$$(m-k+1) \cdot c_{n-m+k} > \sum_{j=n-m+k}^{n} c_j$$

In this case, in the right-hand side of the above inequality, there are $m - k + 1$ terms that are not larger than $c_{n-m+k}$ each. This therefore leads to a contradiction since $c_1 \leq c_2 \leq \cdots \leq c_n$. The lemma follows. □

The following corollary derives an upper-bound on the makespan from $\text{idle}^\text{old}_m$ provided by Expression 6.

**Corollary 2:** Suppose that $J$ is sorted by non-decreasing job processing times, i.e., $c_1 \leq c_2 \leq \cdots \leq c_n$. Then, whatever the job priority assignment, an upper-bound $\text{ms}^\text{ident}(J, \pi)$ on the makespan is given by $\text{ms}^\text{ident}(J, \pi) \overset{\text{def}}{=} c_n$ if $n = m$, or by

$$\text{ms}^\text{ident}(J, \pi) \overset{\text{def}}{=} \sum_{i=1}^{n-1} c_i + c_n$$ \hspace{1cm} (7)
otherwise.

Proof: Since the makespan corresponds to the $m^{th}$ idle-instant, an upper-bound on the makespan is given by $\text{idle}_m$. Therefore, the proof is obtained by simply replacing $k$ with $m$ in Expression 6.

\section*{B. Accuracy of the upper-bound $\text{ms}_{\text{id}}(J, \pi)$}

In this section, Lemma 11 proves that the upper-bound $\text{ms}_{\text{id}}(J, \pi)$ is 2-competitive, according to the following definition.

Definition 15 ($\alpha$-competitive): Any upper-bound is said to be $\alpha$-competitive if it provides at most $\alpha$ times the exact value of the approximated parameter.

This is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET. Without this assumption, the minimum makespan that could be produced is always 0 since it can always be the case that no old-mode task has an active job when the mode change is requested. For instance in Figure 5, the makespan would be zero if the MCR$(j)$ was released at time 110. However, in order to guarantee that our approach always provides an upper-bound on the makespan we have to consider the worst-case scenario in which every old-mode task releases a job exactly upon the mode change request and all these jobs executes for their WCET during the transition.

Lemma 11: For any set $J$ of jobs sorted by non-decreasing job processing time and for any identical multiprocessor platform $\pi$ composed of $m$ CPUs, the upper-bound $\text{ms}_{\text{id}}(J, \pi)$ is 2-competitive.

Proof: Recall from Expression 7 that,

$$\text{ms}_{\text{id}}(J, \pi) \equiv \begin{cases} c_n & \text{if } (n \leq m) \\ \frac{\sum_{i=1}^{n-1} c_i}{m} + c_n & \text{otherwise} \end{cases}$$

Let $\text{ms}(J, m)$ denote the exact makespan for the set $J$ of jobs and the $m$ identical CPUs. Since we do not have any mathematical expression for determining this exact makespan $\text{ms}(J, m)$, our analysis is performed while considering a lower-bound $\text{ms}_{\text{id}}(J, m)$ on the makespan rather than its exact value, i.e., $\alpha$ is determined in such a manner that

$$\frac{\text{ms}_{\text{id}}(J, \pi)}{\text{ms}(J, m)} \leq \alpha$$

where

$$\text{ms}_{\text{id}}(J, m) \equiv \begin{cases} c_n & \text{if } n \leq m \\ \max \left\{ \frac{c_n + \sum_{i=1}^{n-1} c_i}{m} \right\} & \text{if } n > m \end{cases}$$

The case where $n \leq m$ obviously leads to $\alpha = 1$ since both $\text{ms}_{\text{id}}(J, \pi)$ and $\text{ms}_{\text{id}}(J, \pi)$ return a makespan of $c_n$. Otherwise (if $n > m$) the “max” operator in the definition of $\text{ms}_{\text{id}}(J, m)$ leads to two different cases.

\begin{itemize}
  \item \textbf{Case 1}: If $c_n \geq \frac{\sum_{i=1}^{n-1} c_i}{m}$ then we get

$$\frac{\text{ms}_{\text{id}}(J, \pi)}{\text{ms}(J, m)} \leq \frac{\frac{\sum_{i=1}^{n-1} c_i}{m} + c_n}{\sum_{i=1}^{n-1} c_i + c_n} \leq \frac{c_n}{c_n + \frac{\sum_{i=1}^{n-1} c_i}{m}} \leq 2$$

and since in this case we have $c_n \geq \frac{\sum_{i=1}^{n-1} c_i}{m}$, it holds that

$$\frac{\text{ms}_{\text{id}}(J, \pi)}{\text{ms}(J, m)} \leq \frac{c_n + c_n}{c_n + \frac{\sum_{i=1}^{n-1} c_i}{m}} \leq 2$$

\item \textbf{Case 2}: If $c_n < \frac{\sum_{i=1}^{n-1} c_i}{m}$ then

$$\frac{\text{ms}_{\text{id}}(J, \pi)}{\text{ms}(J, m)} \leq \frac{\frac{\sum_{i=1}^{n-1} c_i}{m} + c_n}{\sum_{i=1}^{n-1} c_i} \leq \frac{\frac{\sum_{i=1}^{n-1} c_i}{m} + \frac{\sum_{i=1}^{n-1} c_i}{m}}{\sum_{i=1}^{n-1} c_i} \leq 2$$

The lemma follows.

It holds from Lemma 11 that, for any set $J$ of jobs and any identical platform composed of $m$ CPUs, the upper-bound on the maximum makespan provided by $\text{ms}_{\text{id}}(J, \pi)$ is at most twice the exact value of the maximum makespan. Additionally we can show that in some particular cases as the one provided in the following example, the upper-bounds $\text{idle}_{k}(J, \pi)$ (for $k \in [1, m]$) defined on page 15 are exact.

Example 6: Let us consider the set of 12 jobs with characteristics given in Table III to be scheduled on a 3-processors identical platform.

\begin{table}[ht]
\centering
\begin{tabular}{llllllll}
\hline
$c_1$ & $c_2$ & $c_3$ & $c_4$ & $c_5$ & $c_6$ \\
\hline
1 & 1 & 1 & 1 & 1 & 1 \\
\hline
$c_7$ & $c_8$ & $c_9$ & $c_{10}$ & $c_{11}$ & $c_{12}$ \\
\hline
3 & 3 & 6 & 6 & 9 & 12 \\
\hline
\end{tabular}
\caption{Processing times of the 12 jobs in $J$.}
\end{table}

For this set of jobs,

- the upper-bound $\text{idle}_1 = 15$ is reached with the job priority assignment $J_7 > J_9 > J_{10} > J_{12} > J_{11} > J_8 > J_1 > J_2 > J_3 > J_4 > J_5 > J_6 > J_{12}$.
- the upper-bound $\text{idle}_2 = 18$ is reached with the job priority assignment $J_{10} > J_9 > J_1 > J_2 > J_3 > J_4 > J_5 > J_6 > J_{12} > J_7 > J_8 > J_{11}$.
- the upper-bound $\text{idle}_3 = 23$ is reached with the job priority assignment $J_7 > J_{11} > J_{10} > J_3 > J_2 > J_9 > J_8 > J_3 > J_5 > J_4 > J_6 > J_{12}$.

Due to the space limitation, we did not draw the schedules corresponding to these priority assignments.
C. Validity tests for SM-MSO and AM-MSO

From Corollaries 1 and 2, the sufficient validity test given by Test 1 on page 7 can be rewritten as follows.

**Validity Test 2 (SM-MSO, Identical and FJP):** For any multi-mode real-time application $\tau$ and any identical platform $\pi$ composed of $m$ CPUs, the protocol SM-MSO is valid provided that, for every mode $M^i$,

$$\text{ms}_i^{\text{ident}}(J^{\text{wc}}_k, \pi) \leq \min \left\{ \min_{j \in J_k} \left\{ D_j^i(M^i) \right\} \right\}$$

where $\text{ms}_i^{\text{ident}}(J^{\text{wc}}_k, \pi)$ is defined as in Expression 7 and $J^{\text{wc}}_k$ is defined as follows:

- $J^{\text{wc}}_k$ is defined as in Expression 7 and $J^{\text{wc}}_k$ is defined as follows:
  - each job $J_k \in J^{\text{wc}}_k$ has a processing time equal to the WCET $C^i_k$ of task $\tau^i_k$.
  - $J^{\text{wc}}_k$ is sorted by non-decreasing processing time.

Concerning the protocol AM-MSO, the upper-bounds $idl^i_k(J^{\text{wc}}_k, \pi)$ (for all $1 \leq k \leq m$) defined as in Lemma 9 can be used at line 10 of the validity algorithm given by Algorithm 10 (on page 11).

VII. IDENTICAL PLATFORMS AND FTP SCHEDULERS

This section is organized as follows. First, Section VII-A determines an upper-bound $idl^i_k(J, \pi, P)$ on each idle-instant $idl^i_k(J, \pi, P)$ for any given job priority assignment $P$ and derives an upper-bound $\text{ms}_i^{\text{ident}}(J, \pi, P)$ on the maximum makespan. Then, Section VII-B shows that this upper-bound $\text{ms}_i^{\text{ident}}(J, \pi, P)$ is 1-competitive, with the interpretation that $\text{ms}_i^{\text{ident}}(J, \pi, P)$ corresponds to the exact value of the maximum makespan. Finally, Section VII-C establishes a sufficient validity test for the protocols SM-MSO and AM-MSO.

A. Upper-bounds $idl^i_k(J, \pi, P)$ on the idle-instants

As introduced earlier, this section focuses on determining a mathematical expression for the upper-bounds $idl^i_k(J, \pi, P)$, where $J$ refers to any set of $n$ jobs, $\pi$ denotes any identical multiprocessor platform composed of $m$ CPUs and $P$ is a specific given job priority assignment. Indeed, for a given FTP scheduler the priority of every task (and thus of every job) is known beforehand. This prior knowledge allows us to determine tighter upper-bounds than those proposed in the previous section. Once again, for sake of clarity, we will use the notations $idl^i_k$ and $idl^i_k$ instead of $idl^i_k(J, \pi, P)$ and $idl^i_k(J, \pi, P)$, respectively.

For any transition from a given mode $M^i$ to any other mode $M^j$, the knowledge of the critical rem-job set $J^{\text{wc}}_k$ and the fact that the job priority assignment is known beforehand allow us to compute the exact maximum idle-instants $idl^i_k$—exact in the sense that they are actually reached if every job executes for its WCET—simply by drawing the schedule of $J^{\text{wc}}_k$ and by measuring the idle-instants $idl^i_k$ in that schedule. Indeed, from Corollary 1 (on page 14), each idle-instant $idl^i_k(J^{\text{wc}}_k, \pi, P)$ is an upper-bound on the idle-instant $idl^i_k(J, \pi, P)$ derived from the schedule of any other set $J$ of rem-jobs. Before expressing these exact maximum idle-instants, let us introduce the following definition.

**Definition 16 (Processed work $W_{\pi}^k$):** Let $\pi$ denote any identical multiprocessor platform and let $S$ be any global, weakly work-conserving and FTP scheduler. Let $J = \{J_1, J_2, \ldots, J_n\}$ denote any set of $n$ jobs sorted by decreasing $S$-priority, i.e., $J_1 >_S J_2 >_S \cdots >_S J_n$ and let $S^i$ denote the schedule by $S$ of the $i$ highest priority jobs of $J$ upon $\pi$. The processed work $W_{\pi}^k(1 \leq k \leq m$ and $0 \leq i \leq n)$ denotes the amount of processing time executed on CPU $\pi_k$ in $S^i$.

In order to familiarize the reader with this notation $W_{\pi}^k$, we provide the following example.

**Example 7:** Let us consider the set $J$ of 7 jobs with characteristics given in Table IV to be scheduled on a 4-processors identical platform, following the priority assignment: $J_1 > J_2 > \cdots > J_7$.

<table>
<thead>
<tr>
<th>Job</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>$c_5$</th>
<th>$c_6$</th>
<th>$c_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_1$</td>
<td>7</td>
<td>2</td>
<td>5</td>
<td>16</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

**TABLE IV:** Processing times of the 7 jobs in $J$.

---

Figure 17 illustrates the schedule of $J$ upon the 4 CPUs. In this schedule, we have $W_{\pi}^5 = 8$ because, in the schedule $S^5$ of the 5 highest priority jobs $J_1, J_2, J_3, J_4, J_5$, the amount of processing time units executed on $\pi_3$ is $c_2 + c_5 = 8$. Similarly, $W_{\pi}^3 = 7, W_{\pi}^4 = 2, W_{\pi}^2 = 5$ and $W_{\pi}^4 = 0$ because, in the schedule $S^3$ of jobs $J_1, J_2, J_3$, we can see that 7 processing time units are executed on $\pi_4$ (i.e., job $J_4$), 2 processing time units are executed on $\pi_3$ (i.e., job $J_2$), 5 processing time units are executed on $\pi_2$ (i.e., job $J_3$) and no processing time unit is executed on $\pi_1$. Notice that $W_{\pi}^k = 0 \forall k = 1, 2, \ldots, m$.

Lemma 12 provides the exact values of $W_{\pi}^k$ ($\forall i = 1, m$ and $\forall k \in [1, m]$) when each job executes for its WCET. Then, Corollary 3 derives the exact maximum idle-instants $idl^i_k$ ($1 \leq k \leq m$, for the scheduling of any set $J$ of $n$ jobs upon any $m$-processors identical platform.

**Lemma 12:** Let $\pi$ denote any identical multiprocessors platform composed of $m$ CPUs. Let $S$ be any global, weakly work-conserving and FTP scheduler and let $J$ be any set of
n jobs sorted by decreasing $S$-priority, i.e., $J_1 \succ_S J_2 \succ_S \cdots \succ_S J_n$. It holds $\forall k \in [1, m]$ and $\forall i \in [1, n]$ that

$$\text{Work}^k = \begin{cases} \text{Work}^k_{i-1} + c_i & \text{if } k = \max \left\{ \arg\min_{\ell \in [1, m]} \{ \text{Work}^\ell_{i-1} \} \right\} \\ \text{Work}^k_{i-1} & \text{otherwise} \end{cases}$$

(8)

where $\text{Work}^k_0 = 0 \forall k$ by definition of the processed work.

Proof: The proof directly follows from the definition of $\text{Work}^k \forall i, k$ and from the second condition of our definition of a weakly work-conserving scheduler (see Definition 8, page 4). Indeed, whenever a subset $P$ of several CPUs idle (or complete a job) at the same time, $S$ dispatches the waiting job (if any) with the highest priority to the CPU of $P$ with the highest index (this is the reason for the condition “if $k$ is the highest value of $\ell$ that minimizes $\text{Work}^\ell_{i-1}$”).

Corollary 3: An upper-bound $\text{idle}_k$, $1 \leq k \leq m$, is given by the $k^{th}$ element of the vector $\{\text{Work}^1, \text{Work}^2, \ldots, \text{Work}^m\}$ sorted by non-decreasing order.

Proof: The proof directly follows from the definition of the processed work $\text{Work}^i \forall k \in [1, m]$.

Corollary 4: The maximum makespan $\text{ms}^\text{id}etr (J, \pi, \mathcal{P})$ is given by $\text{idle}_m$, where $\text{idle}_m$ is determined as in Corollary 3.

B. Accuracy of the upper-bound $\text{ms}^\text{id}etr (J, \pi, \mathcal{P})$

In this section we prove that the upper-bound $\text{ms}^\text{id}etr (J, \pi, \mathcal{P})$ is $1$-competitive, i.e., exact—exact in the sense that it can actually be reached if every job executes for their WCET. Again, this is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET as we have to consider the worst-case scenario in which every old-mode task releases a job exactly upon the mode change request and all these jobs execute for their WCET during the transition.

For any transition from a given mode $M^j$ to any other mode $M^l$, the knowledge of the critical rem-job set $J^\text{wc}_i$ and the fact that we proceed by simulation allow us to compute the exact maximum idle-instants simply by drawing the schedule of $J^\text{wc}_i$ following $\mathcal{P}$ and by measuring the idle-instants in this schedule. Using this approach, the measured upper-bound $\text{ms}^\text{id}etr (J, \pi, \mathcal{P})$ is nothing else but $1$-competitive.

C. Validity tests for SM-MSO and AM-MSO

From Corollary 4, the sufficient validity test given by Test 1 (on page 7) can be rewritten as follows.

Validity Test 3 (SM-MSO, identical and FTP): For any multi-mode real-time application $\tau$ and any identical platform $\pi$ composed of $m$ CPUs, the protocol SM-MSO is valid provided that, for every mode $M^j$,

$$\text{ms}^\text{id}etr (J^\text{wc}, \pi, \mathcal{P}) \leq \min_{j \neq i} \left\{ \min_{k=1} \left\{ D_k^i (M^j) \right\} \right\}$$

where $\text{ms}^\text{id}etr (J^\text{wc}, \pi, \mathcal{P})$ is defined as in Corollary 4, $\mathcal{P}^i$ is obtained from the old-mode scheduler $S^i$ and $J^\text{wc}_i$ is defined as follows:

$\triangleright J^\text{wc}_i \triangleq \{ J_1, J_2, \ldots, J_n \}$

$\triangleright$ each job $J_k \in J^\text{wc}_i$ has a processing time equal to the WCET $C_k^i$ of task $\tau_k^i$

$\triangleright J^\text{wc}_i$ is sorted by decreasing $S^i$-priority.

Concerning the protocol AM-MSO, the upper-bounds $\text{idle}_k (J^\text{wc}, \pi, \mathcal{P}^i)$ (for all $1 \leq k \leq m$) determined in Corollary 3 can be used at line 10 of the validity algorithm given by Algorithm 10 on page 11).

VIII. UNIFORM PLATFORMS AND FJP SCHEDULERS

A. Some useful observations

In this section, we show that the maximum makespan determination problem is highly counter-intuitive upon uniform platforms and the methods for solving this problem cannot be straightforwardly extended from those proposed for identical multiprocessor platforms. First, recall that the schedulers are assumed to be strongly work-conserving here since we focus on uniform platforms.

Observation 2: For a given set of jobs, an intuitive idea for maximizing the makespan upon any $m$-processor uniform platform is to execute, at any time, the longest job upon the slowest CPU, i.e., the shorter the computation requirement of a job, the higher its priority. We name this priority assignment “Shortest Job First” (SJF). However, we can show by using the following example that this intuitive idea is erroneous, as SJF does not lead to the maximum makespan.

Example 8: Let us consider the set $J$ of 4 jobs $J_1, J_2, J_3, J_4$ of respective processing times 4, 4, 16 and 22, and suppose that they are scheduled on the 2-processors uniform platform $\pi = [1, 2]$. The priority assignment SJF (i.e., $J_1 > J_2 > J_3 > J_4$) provides a makespan of 17.75 whereas the priority assignment $J_3 > J_1 > J_2 > J_4$ leads to a makespan of 19. Notice that the problem of determining in a polynomial time (i.e., without trying every priority assignment) a priority assignment leading to the maximum makespan remains an open question and is out of the scope of this study.

Observation 3: Another intuitive idea is to naively extend to uniform platforms the result (replicated below) of Corollary 2 on page 15, i.e., for any identical platform $\pi$ composed of $m$ CPUs, an upper-bound on the makespan is given by

$$\text{ms}^\text{id}etr (J, \pi) \triangleq \begin{cases} c_m & \text{if } (n = m) \\ \sum_{i=1}^{n-1} c_i + c_n & \text{otherwise} \end{cases}$$

(9)

where $c_i$ is assumed to be such that $c_i \geq c_{i-1} \forall i \in [2, n]$.

Upon identical platforms there is a sense in distinguishing the case $n = m$ from the case $n > m$, because the rem-jobs never migrate between CPUs during mode transitions. Therefore, in the particular case where $n = m$, the maximum makespan does not depend on the job priority assignment and can be determined exactly by $\text{ms}^\text{id}etr (J, \pi) = c_n$. In contrast, we can easily show that this property does not hold upon uniform platforms. That is, the maximum makespan in the case $n = m$ is not independent from the job priority assignment upon uniform platforms. This is shown through the following example.
Example 9: Consider the uniform platform $\pi = [1, 2]$ and the two jobs $J_1, J_2$ of processing time 4 and 6, respectively. If $J_1 > J_2$ then $J_1$ completes on $\pi_2$ at time 2—time during which $J_2$ executes 2 execution units on $\pi_1$—and $J_2$ completes on $\pi_2$ at time 4, thus leading to a makespan of 4. On the other hand, if $J_2 > J_1$ then $J_2$ completes on $\pi_2$ at time 3—time during which $J_1$ executes 3 execution units on $\pi_1$—and $J_1$ completes on $\pi_2$ at time 3.5, thus leading to a makespan of 3.5. As a result, the maximum makespan in the case $n = m$ depends on the job priority assignment on uniform platforms and the case $n = m$ can no longer be distinguished from the case $m < n$.

From the previous example, naively extending Expression 9 to uniform platforms yields the following “1-piece” expression:\(^{10}\)

\[
\overline{m}_{\text{unif}}(J, \pi) = \sum_{i=1}^{n-1} \frac{c_i}{s(1)} + \frac{c_n}{s_m} \tag{10}
\]

Unfortunately, we show in the following example that this extension does not provide an upper-bound on the maximum makespan.

Example 10: Let us consider the set $J$ of 3 jobs $J_1, J_2, J_3$ of respective processing times 50, 80 and 99, and suppose that they are scheduled on the 3-processors uniform platform $\pi = [1, 2, 10]$. The maximum makespan is 20, reached using the job priority assignment $J_1 > J_2 > J_3$ (see Figure 18). On the other hand, Expression 10 yields $\overline{m}_{\text{unif}}(J, \pi) = \frac{50 + 80}{13} + \frac{99}{19} = 19.9$. This approximation made by Expression 10 is illustrated in Figure 19. This simple example is much more important than what it seems to be at first blush and we will deeply examine its impacts in Section VIII-D (page 21).

B. Upper-bounds $\overline{m}_{\text{idle}}(J, \pi)$ on the idle-instants

Once more but this time for any uniform platform $\pi$, we focus on determining a mathematical expression that provides an upper-bound $\overline{m}_{\text{idle}}(J, \pi)$ on the $k$th idle-instant, $\forall k \in [1, m]$. For sake of clarity, the following two lemmas use the notations $\overline{m}_{\text{idle}}$ instead of $\overline{m}_{\text{idle}}(J, \pi)$ and similarly, the notation $\overline{m}_{\text{idle}}$ will be used to denote the exact value of the $k$th idle-instant. First, Lemma 13 determines a lower-bound $\overline{m}_{\text{idle}}$ on each idle-instant $\overline{m}_{\text{idle}}$, $1 \leq k \leq m$. Then, Lemma 14 determines an upper-bound $\overline{m}_{\text{idle}}$ on each idle-instant $\overline{m}_{\text{idle}}$.

Lemma 13 (See [13]): Let $\pi = [s_1, s_2, \ldots, s_m]$ be any $m$-processors uniform platform such that $s_i \geq s_{i-1} \forall i, 2 \leq i \leq m$. Let $J = \{J_1, J_2, \ldots, J_n\}$ be any set of $n$ jobs of respective processing times $c_1, c_2, \ldots, c_n$ such that $c_1 \leq c_2 \leq \cdots \leq c_n$. Let $S$ be the schedule of $J$ upon $\pi$ following any global, strongly work-conserving and FJP scheduler. A lower bound $\overline{m}_{\text{idle}}$ on each idle-instant $\overline{m}_{\text{idle}}$ ($1 \leq k \leq m$) in $S$ is given by

\[
\overline{m}_{\text{idle}} = \frac{\sum_{i=1}^{n-m+k} c_i}{s(1)} \tag{11}
\]

Proof: According to the definition of the idle-instants, at most $(m-k)$ jobs are not completed at time $\overline{m}_{\text{idle}}$, meaning that at least $(n-m+k)$ jobs are already completed. Let $J_{\text{any}}$ be any subset of $J$ composed of $r$ jobs, where $(n-m+k) \leq r \leq n$. Obviously, a lower bound $t$ on the instant at which the $r$ jobs of $J_{\text{any}}$ are completed is given by

\[
t = \frac{\sum_{i \in J_{\text{any}}} c_i}{s(1)}
\]

and since $c_1 \leq c_2 \leq \cdots \leq c_n$, $t$ is minimal if (i) the number of jobs in $J_{\text{any}}$ is low as possible, i.e., $r = n - m + k$, and (ii) the processing time of each job of $J_{\text{any}}$ is low as possible. As a result, $t$ is minimum for $J_{\text{any}} = \{J_1, J_2, \ldots, J_{n-m+k}\}$ and then yields a lower-bound for $\overline{m}_{\text{idle}}$.

Lemma 14 (See [13]): Using the same notations as in the previous lemma, an upper-bound $\overline{m}_{\text{idle}}$ on each idle-instant $\overline{m}_{\text{idle}}$ ($1 \leq k \leq m$) in $S$ is given by

\[
\overline{m}_{\text{idle}} = \frac{\sum_{i=1}^{n} c_i - \sum_{i=1}^{k-1} \overline{m}_{\text{idle}} \cdot \Delta s_i}{s(k)} \tag{12}
\]

where $s(k) \equiv \sum_{i=k}^{m} \Delta s_i$ (as defined in Expression 1, page 3).

Proof: From the “staircase” property derived from the definition of a strongly work-conserving scheduler on uniform
platform (see page 4 for details) and from the fact that all the jobs are assumed to be synchronous at time 0, we know that CPU $\pi_j$ becomes idle at time $idle_j$, $\forall j = 1, \ldots, m$. Let $\omega_j$ ($1 \leq j \leq m$) denotes the amount of work executed on CPU $\pi_j$, within $[0, idle_j]$, i.e., $\omega_j \leq idle_j \cdot s_j$. The proof is made by contradiction. Let $\ell$ be any integer in $[1, m]$ and suppose that $idle_{\ell} > idle_{\ell}$. By definition of $\omega_j$, we know that
\[
\sum_{j=1}^{m} \omega_j = \sum_{i=1}^{n} c_i
\]  
(13)
and from the definition of $\omega_j$ we know that
\[
\sum_{j=1}^{m} \omega_j = \sum_{j=1}^{m} idle_j \cdot s_j
\]
By definition of the idle-instants, it holds $\forall j \geq \ell$ that $idle_j \geq idle_{\ell}$. Therefore, replacing “idle$_j$” with “idle$_{\ell}$” in the second term of the right-hand side of the above equality yields
\[
\sum_{j=1}^{m} \omega_j \geq \sum_{j=1}^{\ell-1} (idle_j \cdot s_j) + \sum_{j=\ell}^{m} (idle_{\ell} \cdot s_j)
\]
By hypothesis we have $idle_{\ell} > idle_{\ell}$. Therefore, replacing $idle_{\ell}$ with $idle_{\ell}$ in the right-hand side of the above inequality yields
\[
\sum_{j=1}^{m} \omega_j > \sum_{j=1}^{\ell-1} (idle_j \cdot s_j) + idle_{\ell} \sum_{j=\ell}^{m} s_j
\]
\[
> \sum_{j=1}^{\ell-1} (idle_j \cdot s_j) + \sum_{i=1}^{n} c_i - n \sum_{i=1}^{n} idle_i \cdot s_i
\]
\[
> \sum_{i=1}^{n} c_i + \sum_{j=1}^{\ell-1} (idle_j - idle_{\ell}) \cdot s_j
\]
Since from Lemma 13 it holds that $idle_{\ell} \leq idle_i \forall i = 1, 2, \ldots, m$, it holds that
\[
\sum_{j=1}^{\ell-1} (idle_j - idle_{\ell}) \cdot s_j \geq 0
\]
and thus
\[
\sum_{j=1}^{m} \omega_j > \sum_{i=1}^{n} c_i
\]
leading to a contradiction with Equality 13. The lemma follows. ■

Corollary 5 (See [13]): Whatever the job priority assignment, an upper-bound $\text{ms}_1^{\text{unif}}(J, \pi)$ on the makespan is given by
\[
\text{ms}_1^{\text{unif}}(J, \pi) = \frac{1}{sm} \left( \sum_{i=1}^{n} c_i - \sum_{i=1}^{m-1} idle_i \cdot s_i \right)
\]  
(14)

**Proof:** Since the makespan corresponds to the idle-instant $idle_m$, an upper-bound on the makespan is given by $\text{ms}_1^{\text{unif}}(J, \pi)$. Therefore, the proof is obtained by simply replacing $k$ with $m$ in Expression 12. ■

C. Accuracy of the upper-bound $\text{ms}_1^{\text{unif}}(J, \pi)$

In this section we prove that the upper-bound $\text{ms}_1^{\text{unif}}(J, \pi)$ is $s(1)$-competitive, with the interpretation that the value returned by $\text{ms}_1^{\text{unif}}(J, \pi)$ is at most $\frac{s(1)}{sm}$ times the exact value of the maximum makespan for any given set $J$ of jobs and uniform platform $\pi$. Once again, this is achieved under the assumption that during any mode transition all the rem-jobs execute for their WCET as we have to consider the worst-case scenario in which every old-mode task releases a job exactly upon the mode change request and all these jobs executes for their WCET during the transition.

Lemma 15: For any set $J$ of jobs sorted by non-decreasing job processing time and any uniform platform $\pi = [s_1, s_2, \ldots, s_m]$ with $s_i \geq s_{i-1} \forall i$, $\text{ms}_1^{\text{unif}}(J, \pi)$ is $\alpha_1(\pi)$-competitive, where $\alpha_1(\pi) = \frac{1}{sm}$.

**Proof:** Recall from Expression 14 that
\[
\text{ms}_1^{\text{unif}}(J, \pi) = \frac{\sum_{i=1}^{n} c_i}{sm} - \frac{\sum_{k=1}^{m-1} (\sum_{i=1}^{n} c_i \cdot s_k)}{sm \cdot s(1)}
\]
Let $\text{ms}(J, \pi)$ denote the exact makespan for any given set $J$ of jobs and any uniform platform $\pi$. Since we do not have any mathematical expression for determining this exact makespan $\text{ms}(J, \pi)$, our analysis of $\alpha_1(\pi)$ is performed while considering a lower-bound $\text{ms}_1(J, \pi)$ on the makespan rather than its exact value, i.e., $\alpha_1(\pi)$ is determined in such a manner that
\[
\frac{\text{ms}_1^{\text{unif}}(J, \pi)}{\text{ms}(J, \pi)} \leq \alpha_1(\pi)
\]
Obviously, we know that $\text{ms}(J, \pi) \geq \frac{\sum_{i=1}^{n} c_i}{s(1)}$ and this implies that $\text{ms}_1(J, \pi) \equiv \frac{\sum_{i=1}^{n} c_i}{s(1)}$ is a lower-bound on the makespan. This yields
\[
\frac{\text{ms}_1^{\text{unif}}(J, \pi)}{\text{ms}(J, \pi)} \leq \frac{\text{ms}_1^{\text{unif}}(J, \pi)}{\text{ms}_1(J, \pi)}
\]
and thus,
\[
\frac{\text{ms}_1^{\text{unif}}(J, \pi)}{\text{ms}(J, \pi)} \leq \frac{\sum_{i=1}^{n} c_i}{sm \cdot s(1)} - \frac{\sum_{k=1}^{m-1} (\sum_{i=1}^{n} c_i \cdot s_k)}{sm \cdot s(1)}
\]
D. Another analysis of the maximum makespan

In Example 10 on page 19, we have showed that the naive extension of $m_{\text{id}}(J, \pi)$ (given by $m_{\text{id}}^{\text{naive}}(J, \pi)$ in Expression 10, page 19) does not provide an upper-bound on the maximum makespan considering uniform platforms. Essentially, in addition to refute the fact that $m_{\text{id}}^{\text{naive}}(J, \pi)$ provides an upper-bound on the maximum makespan, this example also refutes the main concept behind the expression of $m_{\text{id}}(J, \pi)$. Indeed, in the expression of $m_{\text{id}}(J, \pi)$, it can be easily shown that the term $\sum_{n-1}^{n} c_i$ is a bound on the time at which $J_n$ starts its execution, i.e., its dispatching time. Therefore, the whole expression can be interpreted as follows: upper-bound on the makespan = upper-bound on the dispatching time of $J_n + c_n$, where $J_n$ is the (any) job with the largest processing time. That is, this expression of $m_{\text{id}}(J, \pi)$ is based on the intuition that the maximum makespan is reached when the longest job is dispatched as late as possible and executes for its WCET. This intuition has revealed to be true for the case of identical platforms, but not for the uniform case (as shown by Example 10)\footnote{Indeed, we can also easily show that the term $(\sum_{n-1}^{n} c_i) / s(1)$ in Expression 10 is an upper-bound on the dispatching time of $J_n$ and at that time $J_n$ is dispatched to the fastest CPU $\pi_m$, leading to a WCET of $\frac{c_n}{s_m}$}. The whole concept is not extendable to uniform platforms and in order to figure out the underlying cause, let us focus on Example 10.

Let $S^{\text{naive}}$ and $S^{\text{ms}}$ denote the two schedules depicted in Figure 20, issued from the approximation $m_{\text{id}}^{\text{naive}}(J, \pi)$ and from the priority assignment $\pi_3 > \pi_2 > \pi_1$ which leads to the maximum makespan, respectively. The reason why $m_{\text{id}}^{\text{naive}}(J, \pi)$ under-approximates the maximum makespan comes from the following fact: if $t$ denotes the instant at which job $J_3$ is dispatched to CPU $\pi_3$ in $S^{\text{ms}}$ (here, $t = 12$), then during the time interval $[0, t]$, $J_3$ has executed a lower amount of execution units in the stairs of $S^{\text{ms}}$ than upon $\pi_3$ in $S^{\text{naive}}$. In other words the cumulated green areas in Figure 20 represent a lower amount of execution units than the red area. Indeed, $J_3$ executes $5 + 14 = 19$ execution units within $[0, t]$ in $S^{\text{ms}}$ whereas it executes $20$ execution units on $\pi_3$ in $S^{\text{naive}}$. As a result, the remaining processing time of $J_3$ at time $t$ is higher in $S^{\text{ms}}$ (here, 80) than in $S^{\text{naive}}$ (here, 79), implying that $J_3$ completes later in $S^{\text{ms}}$ than in $S^{\text{naive}}$. This is the reason why the expression $m_{\text{id}}^{\text{naive}}(J, \pi)$ does not provide the maximum makespan in the example above: on uniform platforms, the schedule in which any job $J_i$ reaches its maximum completion time is not necessarily the schedule in which $J_i$ is dispatched as late as possible.

Based on this fundamental observation, we propose and prove correct in [38] (pages 138–163 and 351–367) two additional upper-bounds $m_{\text{id}}^{\text{naive}}(J, \pi)$ and $m_{\text{id}}^{\text{naive}}(J, \pi)$ on the maximum makespan, considering uniform platforms and FJP schedulers. These upper-bounds are replicated below.

$$m_{\text{id}}^{\text{naive}}(J, \pi) \equiv \frac{1}{s_m} \sum_{i=1}^{n} c_i + \frac{s(1)}{s(1)} \cdot K_n - 1$$

where $K_j$ is such that $\forall j$, $K_j \equiv \begin{cases} 1 & \text{if } s_1 = s_m \text{ and } j = 0 \\ \left(1 - \frac{s_j}{s_m}\right)^j & \text{otherwise} \end{cases}$ and

$$m_{\text{id}}^{\text{naive}}(J, \pi) \equiv \frac{1}{s_m} \sum_{i=1}^{n} c_i + \frac{s(1)}{s(1)} \cdot H_{n-\ell}$$

where

$$x = \arg\min_{i \in [1, m]} \left\{ \frac{s_i}{\sum_{j=1}^{x} s_j} \right\}$$

and $H_j$ is such that $\forall j$, $H_j \equiv \begin{cases} 1 & \text{if } s_1 = s_m \text{ and } j = 0 \\ \left(1 - \frac{s_j}{s_m}\right)^j & \text{otherwise} \end{cases}$

Each of these two upper-bounds is based on a distinct upper-bound on the amount of execution units that can be executed in the green areas (see Figure 20), and then derives an upper-bound on the completion time of every job, and finally on the makespan.
E. Validity tests for SM-MSO and AM-MSO

From Expressions 14, 17, and Corollary 1, a sufficient validity test for the protocol SM-MSO can therefore be formalized as follows.

Validity Test 4 (SM-MSO, uniform and FIP): For any multi-mode real-time application \(\tau\) and any uniform platform \(\pi = [s_1, s_2, \ldots, s_m]\) composed of \(m\) CPUs, the protocol SM-MSO is valid provided that, for every mode \(M^i\),

\[
\min_{\pi} (J_{\text{wc}}^{\text{unif}}, \pi) \leq \min \left\{ \min_{j \neq i} \left\{ \min_{k=1}^{n_j} \left\{ D_k^{\text{def}}(M^i) \right\} \right\} \right\}
\]

where \(\min_{\pi} (J_{\text{wc}}^{\text{unif}}, \pi)\) is defined as \(\min_{\pi} (J_{\text{wc}}^{\text{unif}}, \pi) \defeq \min \left\{ \min_{1} (J_{\text{wc}}^{\text{unif}}, \pi), \min_{2} (J_{\text{wc}}^{\text{unif}}, \pi), \min_{3} (J_{\text{wc}}^{\text{unif}}, \pi) \right\}\) (19)

and \(\min_{1} (J_{\text{wc}}^{\text{unif}}, \pi), \min_{2} (J_{\text{wc}}^{\text{unif}}, \pi)\) and \(\min_{3} (J_{\text{wc}}^{\text{unif}}, \pi)\) are defined as in Expressions 14, 17 and 18, respectively.

This is performed considering the set \(J_{\text{wc}}^{\text{unif}}\) composed of \(n_j\) jobs of processing time \(C_j^i, C_j^i, \ldots, C_n^i\) such that \(C_j^i \geq C_j^{i-1} \forall j = 2, 3, \ldots, n_j\).

Concerning the protocol AM-MSO, the upper-bounds \(\min_{\pi} (J_{\text{wc}}^{\text{unif}}, \pi)\) (for all \(1 \leq k \leq m\)) defined as in Lemma 14 can be used at line 10 of the validity algorithm given by Algorithm 10 (on page 11).

F. Simulation results

Because our analysis of the competitive factor did not lead to a constant \(\alpha\) for the upper-bound \(\min_{\pi} (J_{\text{wc}}^{\text{unif}}, \pi)\) (as well as for the upper-bounds \(\min_{1} (J_{\text{wc}}^{\text{unif}}, \pi), \min_{2} (J_{\text{wc}}^{\text{unif}}, \pi)\) and \(\min_{3} (J_{\text{wc}}^{\text{unif}}, \pi)\) as shown in [38]), this section reports on the results of simulations in order to quantify the precision of the three upper-bounds \(\min_{1} (J_{\text{wc}}^{\text{unif}}, \pi), \min_{2} (J_{\text{wc}}^{\text{unif}}, \pi)\) and \(\min_{3} (J_{\text{wc}}^{\text{unif}}, \pi)\). These simulations are performed considering a single job \(J\) of jobs scheduled and multiple uniform platforms. We consider only a single set \(J\) of jobs for which the exact processing times are given in Table V. We will explain below where these parameters are drawn from and why we consider only a single set of jobs rather than generating numerous job sets.

<table>
<thead>
<tr>
<th></th>
<th>(c_1)</th>
<th>(c_2)</th>
<th>(c_3)</th>
<th>(c_4)</th>
<th>(c_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_1)</td>
<td>3896</td>
<td>3964</td>
<td>878</td>
<td>1378</td>
<td>2228</td>
</tr>
<tr>
<td>(c_2)</td>
<td>3612</td>
<td>1230</td>
<td>1232</td>
<td>1668</td>
<td>4672</td>
</tr>
</tbody>
</table>

TABLE V: Processing times of the 10 jobs in \(J\).

For experimental purposes, let us introduce the parameter \(\lambda_{\pi}\) defined in [25] for any \(m\)-processor uniform platform \(\pi = [s_1, s_2, \ldots, s_m]\)

\[
\lambda_{\pi} \defeq \max_{j=1}^{m} \left\{ \sum_{k=1}^{j-1} s_k / s_j \right\}
\]

Informally speaking, this parameter \(\lambda_{\pi}\) measures the “degree” by which \(\pi\) differs from an identical multiprocessor platform, i.e., its “degree of heterogeneity”. For any identical platform composed of \(m\) CPUs, it holds that \(s_1 = s_2 = \cdots = s_m\) and thus, \(\lambda_{\pi} \defeq \max_{j=1}^{m} \left\{ \sum_{k=1}^{j-1} s_k / s_j \right\} \) is maximum for \(j = m\), leading to \(\lambda_{\pi} \defeq \sum_{k=1}^{m} s_k / s_j = m - 1\). The more homogeneous the platform \(\pi\) is, the closer to \((m - 1)\) is its corresponding \(\lambda_{\pi}\). For instance, the uniform platform \(\pi = [1, 500, 1000]\) has a corresponding \(\lambda_{\pi} = 0.835\) whereas \(\lambda_{\pi} = 0.835\) for the uniform platform \(\pi = [1, 500, 600]\) and \(\lambda_{\pi} \approx 1.67\) for the platform \(\pi = [500, 500, 600]\). In short, \(\lambda_{\pi} \approx (m - 1)\) if \(\pi\) is comprised of \(m\) identical CPUs and becomes progressively smaller as the speeds of the CPUs differ from each other by greater amounts.

The platform \(\pi\) considered in our simulations is composed of \(m = 4\) CPUs for which we make their computing speed varying within \([1, 101]\) with an increment of 10. More precisely, we consider all possible combinations of the CPU speeds in the range \([1, 101]\) with an increment of 10, i.e., the first simulation is performed considering \(\pi = [1, 1, 1, 1]\), the second simulation considers \(\pi = [1, 1, 1, 1]\), the third one considers \(\pi = [1, 1, 1, 21]\), and so on until reaching the speed assignment \(\pi = [101, 101, 101, 101]\). For every speed assignment, we determine the corresponding parameter \(\lambda_{\pi}\) as well as the exact value \(\text{ms}(J, \pi)\) of the maximum makespan. This exact maximum makespan \(\text{ms}(J, \pi)\) is determined by building the schedule of \(J\) upon \(\pi\) for every job priority assignment and by retaining only the maximum generated makespan. This is a highly computational-intensive operation that requires the exhaustive enumeration of every possible job priority assignment. This is the reason why we consider only a single set \(J\) of jobs in our simulations.

Indeed, according to this approach, our simulation process considers 11 different speeds for each CPU, leading to a total of \(11^m = 11^4 = 14,641\) different platforms \(\pi\). For each platform \(\pi\), the computation of the exact makespan requires to generate the schedules derived from every job priority assignment. Since there are 10 jobs, the number of considered priority assignments is \(10! = 3,628,800\). Multiplied by the number of platforms, this leads to 53, 129, 260, 800 operations. Our simulations were performed on HYDRA, the Scientific Computer Configuration at the VUB/ULB Computing Centre, where we fully distributed the computations among 15 processors AMD Opteron dual-core @ 2.8GHz. Distributing the computations allowed us to complete the simulation in about 2 hours but unfortunately, the computation time grows exponentially with the number of CPUs and in a factorial manner with the number of jobs. For instance, considering 13 jobs would result in 91, 169, 811, 532, 800 operations, 14 jobs to approximately \(20 \cdot 10^{15}\) operations, resulting in a computation time of about 82 years. The processing times of the jobs have been drawn from [39] where the authors present realistic parameters that concern the avionic domain. But since the number of operations of our algorithm is strongly restricted by the number of jobs, we arbitrarily selected 10 WCETs from these parameters.

For each speed assignment of the platform we computed the error \(E_{\text{unif}}(J, \pi)\) corresponding to the difference (in percent)
between $\overline{m}^\text{unif}(J, \pi)$ and $\text{ms}(J, \pi)$. Formally,

$$E^\text{unif}_1(J, \pi) = \frac{|\overline{m}^\text{unif}(J, \pi) - \text{ms}(J, \pi)|}{\text{ms}(J, \pi)} \cdot 100$$

and in a similar way we also computed the errors $E^\text{unif}_2(J, \pi)$ and $E^\text{unif}_3(J, \pi)$.

The errors $E^\text{unif}_1(J, \pi)$, $E^\text{unif}_2(J, \pi)$ and $E^\text{unif}_3(J, \pi)$ are displayed in Figure 21 relative to the corresponding $\lambda_\pi$. The horizontal black line is the error “$\text{EXACT MAKESPAN}$” of $\text{ms}(J, \pi)$ over the exact value of the maximum makespan. Obviously, this error is always 0. Also, for every speed assignment of $\pi$, we define the estimator $\overline{m}^\text{unif}(J, \pi)$ as in Expression 19 and its associated error $E^\text{unif}_\text{min}(J, \pi)$. This error is displayed in Figure 22 relative to the corresponding $\lambda_\pi$. Finally, Table VI provides the reader with some statistics issued from the simulation.

For obvious reason, the most accurate estimator (i.e., the most accurate upper-bound on the maximum makespan) is $\overline{m}^\text{unif}(J, \pi)$. As presented in Table VI, the most important error that we obtained for $\overline{m}^\text{unif}(J, \pi)$ is 22.89% and the minimal one is 1.57%. The average error is 10.44% with a squared distance of 5.78%. Hence, we believe that this is a promising path to go for more competitive bounds and for practical use. An open question remains however. For $\lambda_\pi \in [0, 2]$, we can see in Figure 21 that $\overline{m}^\text{unif}(J, \pi)$ is clearly lower than both $\overline{m}^\text{unif}_2(J, \pi)$ and $\overline{m}^\text{unif}_3(J, \pi)$, i.e., $\overline{m}^\text{unif}_\text{min}(J, \pi) = \overline{m}^\text{unif}(J, \pi)$ for $\lambda_\pi \in [0, 2]$. Within this interval $[0, 2]$, when the parameter $\lambda_\pi$ reaches an integer value (here, 1 and 2), something happens that considerably improves the accuracy of $\overline{m}^\text{unif}(J, \pi)$. But up to now, we did not find any interpretation to that phenomenon.

![Fig. 21: The three estimation errors $E^\text{unif}_1(J, \pi)$, $E^\text{unif}_2(J, \pi)$ and $E^\text{unif}_3(J, \pi)$ displayed relative to the corresponding $\lambda_\pi$.](image1)

![Fig. 22: The estimation error $E^\text{unif}_\text{min}(J, \pi)$ displayed relative to the corresponding $\lambda_\pi$.](image2)

**TABLE VI: Statistics issued from the simulation**

<table>
<thead>
<tr>
<th></th>
<th>$E^\text{unif}_1(J, \pi)$</th>
<th>$E^\text{unif}_2(J, \pi)$</th>
<th>$E^\text{unif}_3(J, \pi)$</th>
<th>$E^\text{unif}_\text{min}(J, \pi)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>1.57%</td>
<td>1.89%</td>
<td>2.7%</td>
<td>1.57%</td>
</tr>
<tr>
<td>1st Qu.</td>
<td>6%</td>
<td>21.74%</td>
<td>13.28%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Median</td>
<td>12.72%</td>
<td>41.07%</td>
<td>27.11%</td>
<td>9.92%</td>
</tr>
<tr>
<td>Mean</td>
<td>13.68%</td>
<td>37.91%</td>
<td>29.25%</td>
<td>10.44%</td>
</tr>
<tr>
<td>3rd Qu.</td>
<td>20.72%</td>
<td>55.5%</td>
<td>43.99%</td>
<td>15.08%</td>
</tr>
<tr>
<td>Max.</td>
<td>32.96%</td>
<td>88.78%</td>
<td>68.01%</td>
<td>22.89%</td>
</tr>
<tr>
<td>Variance</td>
<td>69.76%</td>
<td>359.37%</td>
<td>320.47%</td>
<td>33.36%</td>
</tr>
<tr>
<td>SD$^{12}$</td>
<td>8.35%</td>
<td>18.96%</td>
<td>17.9%</td>
<td>5.78%</td>
</tr>
<tr>
<td>Bias</td>
<td>42.35%</td>
<td>133.69%</td>
<td>94.05%</td>
<td>26.93%</td>
</tr>
<tr>
<td>MSE$^{13}$</td>
<td>1863.25%</td>
<td>18233.15%</td>
<td>9165.97%</td>
<td>758.43%</td>
</tr>
</tbody>
</table>

**IX. UNIFORM PLATFORMS AND FTP SCHEDULERS**

This section follows the same reasoning as the one for identical platforms and FTP schedulers. For any transition from a specific mode $M^i$ to any other mode $M^j$, the knowledge of the critical rem-job set $J^\text{wc}$ and the fact that the priorities are known beforehand enable us to compute the exact maximum idle-instants idl$e_k$, $1 \leq k \leq m$, simply by simulating the scheduling of the critical rem-job set and by measuring the idle-instants idl$e_k$, $1 \leq k \leq m$, in that schedule (from Corollary 1 presented on page 14). Thus, each idle-instant idl$e_k$ measured in the schedule of the critical rem-job set is an upper-bound on the idle-instants idl$e_k$ in the schedule derived from any other set of rem-jobs. In conclusion, FTP schedulers enable us to determine the exact$^{14}$ maximum idle-instants idl$e_k$, $1 \leq k \leq m$, rather than over-approximating them (as done for the FJP schedulers).

**A. Upper-bounds $\overline{\text{idle}}_j(J, \pi, P)$ on the idle-instants**

Lemma 16 provides the exact values of idl$e_j(J, \pi, P)$ $\forall j \in [1, m], i \in [1, n]$ and $\forall P$, assuming that every job $J_i$ executes for its WCET. However, in this particular case of FTP scheduler, we redefine the idle-instants idl$e_j(J, \pi, P)$ as follows.

**Definition 17 (Idle-instant idl$e_j(J, \pi, P)$):** If $S^i$ denotes the schedule upon $\pi$ of only the jobs with a higher (or equal) priority than $J_i$ according to $P$, then idl$e_j(J_i, \pi, P)$ is the earliest instant in $S^i$ at which at least $j$ CPUs idle.

The only difference w.r.t. the previous one resides in the “higher (or equal) priority than (..)””. The reason for this redefinition is that, with the previous one, it was not possible to express the idle-instants idl$e_j(J, \pi, P)$ (for $j = 1, \ldots, m$) as in Definition 13 (page 10). Indeed, these idle-instants idl$e_j(J, \pi, P)$ consider that every job of $J$ are scheduled, while the previous definition of the idle-instants idl$e_j(J, \pi, P)$ requires a job index $i$ and considers that only the jobs with a

$^{14}$Exact in the sense that this value is actually reached if every job executes for its WCET.
higher priority than \(J_i\) are scheduled. Thereby, this previous definition always excludes the job \(J_i\) in the computation of the idle-instants. Now, thanks to this new definition, the idle-instants idle\(_i^j\)(\(J_i, \pi, P\)) (for \(j = 1, \ldots, m\)) can be expressed by idle\(_i^j\)(\(J_{\text{low}}, \pi, P\)), where \(J_{\text{low}}\) is the lowest priority job according to \(P\). Once again, we use in Corollary 6 the notations idle\(_i^j\) to refer to the idle-instants idle\(_i^j\)(\(J_i, \pi, P\)) defined as in Definition 17.

**Lemma 16 (See [13]):** Let \(\pi = [s_1, s_2, \ldots, s_m]\) denote any uniform multiprocessor platform composed of \(m\) CPUs and assume that \(s_i \geq s_{i-1}, \forall i = 2, 3, \ldots, m\). Let \(J = \{J_1, J_2, \ldots, J_n\}\) be any set of \(n\) jobs, all released at time \(t = 0\), with respective computation time \(c_1, c_2, \ldots, c_n\). Let \(S\) denote any global, FTP and strongly work-conserving scheduler and suppose that \(J\) is sorted by decreasing \(S\)-priority, i.e., \(J_i \geq J_{i+1}\). If these jobs are scheduled by \(S\) upon \(\pi\), then idle\(_i^j\) is inductively defined as follows:

**Initialization:**

\[
\forall 1 \leq j \leq m : \quad \text{idle}^0_i \triangleq 0
\]

\[
\forall 1 \leq i \leq n : \quad \text{idle}^{n+1}_i \triangleq \infty
\]

**Iteration:**

for \((i = 1 \text{ to } n)\)

for \((j = m \text{ to } 1)\)

\[
\text{idle}^j_i \triangleq \begin{cases} 
\text{idle}^{j-1}_i \text{ if } \text{idle}^{j-1}_i = \text{idle}^{j-1}_{\pi_j} \\
\text{idle}^{j+1}_i \text{ else if } c_i \geq \sum_{k=1}^{j}(\text{idle}^{k-1}_i - \text{idle}^{k-1}_{\pi_k}) \cdot s_k \\
\text{idle}^{j+1}_i + \frac{(c_i - \sum_{k=1}^{j+1}(\text{idle}^{k-1}_i - \text{idle}^{k-1}_{\pi_k})) \cdot s_j}{s_j} \quad \text{otherwise}
\end{cases}
\]

**Proof:** Initially, the \(m\) CPUs idle and thus, idle\(_i^0\) = 0, \(\forall j, 1 \leq j \leq m\). We find convenient to define idle\(_m^{n+1}\) = \(\infty\), \(\forall i\), which means that we have at most \(m\) CPUs available. In the following, we prove the correctness of the value of idle\(_i^j\) (\(\forall j, 1 \leq j \leq m\)) assuming that idle\(_i^1\) are defined (\(\forall j \leq m + 1\)). The idle-instants idle\(_i^{j-1}\) define a staircase as illustrated in Figure 23 for the scheduling of jobs \(J_1, \ldots, J_{i-1}\). Thus, job \(J_i\) can only progress into the blue areas and two cases have to be distinguished:

\(a\) **Case 1:** idle\(_i^{j-1}\) = idle\(_{\pi_j}^{j-1}\), meaning that at least one CPU faster than \(\pi_j\) becomes available at time idle\(_i^{j-1}\) (the blue area on CPU \(\pi_j\) is idle in that case). This situation is depicted in Figure 23 where idle\(_i^{j-1}\) = idle\(_{\pi_2}^{j-1}\) = idle\(_{\pi_3}^{j-1}\) and the blue area is void on CPUs \(\pi_2\) and \(\pi_3\). In this kind of situation, the job \(J_i\) is executed (if not completed) upon a faster CPU and the first instant at which at least \(j\) CPUs idle remains unchanged after having scheduled the job \(J_i\), i.e., idle\(_i^j\) = idle\(_i^{j-1}\).

\(b\) **Case 2:** Otherwise, \(J_i\) is dispatched to CPU \(\pi_j\) at instant idle\(_i^{j-1}\) and keeps executing on \(\pi_j\) as long as (i) no faster CPUs become idle or (ii) \(J_i\) completes. In the first case, \(J_i\) executes on \(\pi_j\) until the next idle-instant idle\(_{\pi_j}^{j+1}\), leading to the first sub-case idle\(_i^j\) = idle\(_{\pi_j}^{j+1}\). In the second case, \(J_i\) executes on CPU \(\pi_j\) but completes before time idle\(_{\pi_j}^{j+1}\). Thus, the idle-instant idle\(_i^j\) is the instant idle\(_i^{j-1}\) at which \(J_i\) was dispatched to \(\pi_j\) plus its remaining processing time on CPU \(\pi_j\) at time idle\(_i^{j-1}\). Since \(\sum_{k=1}^{j}(\text{idle}^{k-1}_i - \text{idle}^{k-1}_{\pi_k}) \cdot s_k\) corresponds to the amount of work that \(J_i\) has executed in the interval of \([0, \text{idle}^{j-1}_i]\), its remaining processing time on CPU \(\pi_j\) at time idle\(_i^{j-1}\) is given by 

\[
\text{idle}^{j-1}_i \cdot s_j \quad \text{leading to the second sub-case.}
\]

**Corollary 6 (See [13]):** The maximum idle-instant idle\(_k\)(\(J, \pi, P\)) (\(\forall k \in [1, m]\)) is given by idle\(_m\) computed as in Lemma 16.

**Corollary 7:** The maximum makespan \(\text{m}n^\text{eff}_i(j, \pi)\) is given by idle\(_m\) computed as in Lemma 16.

**B. Validity tests for SM-MSO and AM-MSO**

From Corollary 7, a sufficient validity test for the protocol SM-MSO can therefore be formalized as follows.

**Validity Test 5 (SM-MSO, uniform and FTP):** For any multi-mode real-time application \(\pi\) and any identical platform \(\pi\) composed of \(m\) CPUs, the protocol SM-MSO is valid provided that, for every mode \(M_i\),

\[
\text{idle}^{m^\text{eff}_i(j, \pi, \pi^*)} \leq \min_{j \neq i} \left\{ \frac{n_j}{k=1} \left\{ D_k(M_i) \right\} \right\}
\]

where idle\(_m^i(j, \pi, \pi^*)\) is computed as idle\(_m^i\) in Lemma 16, considering the critical rem-job set \(J_{\pi^*}^I\) composed of \(n_i\) jobs \(J_1, J_2, \ldots, J_{n_i}\) of respective processing time \(C_{1, i}, C_{2, i}, \ldots, C_{n_i, i}\) and such that \(J_{\pi^*}^I\) is sorted by decreasing \(\pi^*\)-priority.

Similarly, the upper-bounds idle\(_k\)(\(J, \pi, P\)) (\(\leq k \leq m\)) corresponds to the job priority assignment of the old-mode scheduler \(\pi^*\) determined in Lemma 16 can be used at line 10 of the validity algorithm of AM-MSO (see Algorithm 10 page 11), as long as these upper-bounds are computed while assuming the critical rem-job set \(J_{\pi^*}^I\) for the transitions from every mode \(M_i\).

**X. CONCLUSION AND OPEN PROBLEMS**

In this paper, we addressed the scheduling problem of multi-mode real-time applications upon identical and uniform multiprocessor platforms. We assumed that every mode of the
application was scheduled by following a global and Fixed-Task-Priority or Fixed-Job-Priority scheduler. Under these assumptions, we proposed two protocols for managing every transition between every pair of modes of the system, namely SM-MSO and AM-MSO. For both protocols, we established validity tests that allow the system designer to predict whether the given application can meet all the expected timing requirements upon the given platform. We prove the correctness of our schedulability analyses by extending the theory about the makespan determination problem.

In our future work, we aim at taking into account mode-independent tasks, i.e., tasks whose the periodic (or sporadic) activation pattern is not affected by the mode changes. Moreover, instead of scheduling the rem-jobs by using the scheduler of the old-mode during the transitions, it could be better, in term of the enablement delays applied to the new-mode tasks, to propose a dedicated priority assignment which meets the deadline of every rem-job, while minimizing the makespan. To the best of our knowledge, the problem of minimizing the makespan while meeting job deadlines is not yet addressed in the literature and remains open. Table VII outlines a brief overview of all different problems, considering the task and platform model introduced in this paper. For each problem, we indicated either the reference(s) where solutions have been proposed or T.W. (This Work) or F.W. (Future Work) or O.P. (Open Problem).

<table>
<thead>
<tr>
<th>Protocols without periodicity</th>
<th>Protocol</th>
<th>Platform</th>
<th>Scheduler</th>
<th>Existing results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchronous</td>
<td>identical</td>
<td>FJP</td>
<td>[15], [16], [38], T.W.</td>
<td></td>
</tr>
<tr>
<td>Synchronous</td>
<td>identical</td>
<td>FTP</td>
<td>[15], [16], [38], T.W.</td>
<td></td>
</tr>
<tr>
<td>Synchronous</td>
<td>uniform</td>
<td>FJP</td>
<td>[13], [38], T.W.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>identical</td>
<td>FTP</td>
<td>[15], [38], T.W.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>uniform</td>
<td>FJP</td>
<td>[13], [38], T.W.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>uniform</td>
<td>FTP</td>
<td>[13], [38], T.W.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protocols with periodicity</th>
<th>Protocol</th>
<th>Platform</th>
<th>Scheduler</th>
<th>Existing results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Synchronous</td>
<td>identical</td>
<td>FJP</td>
<td>[14], F.W.</td>
<td></td>
</tr>
<tr>
<td>Synchronous</td>
<td>identical</td>
<td>FTP</td>
<td>[14], F.W.</td>
<td></td>
</tr>
<tr>
<td>Synchronous</td>
<td>uniform</td>
<td>FJP</td>
<td>F.W.</td>
<td></td>
</tr>
<tr>
<td>Synchronous</td>
<td>uniform</td>
<td>FTP</td>
<td>F.W.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>identical</td>
<td>FJP</td>
<td>O.P.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>identical</td>
<td>FTP</td>
<td>O.P.</td>
<td></td>
</tr>
<tr>
<td>Asynchronous</td>
<td>uniform</td>
<td>FTP</td>
<td>O.P.</td>
<td></td>
</tr>
</tbody>
</table>

*TABLE VII: State-of-the-art at a glance.*

REFERENCES


Vincent Nély was born in Brussels, Belgium, in 1984. In 2006 he received his master degree in Computer Science at U.L.B. (Université Libre de Bruxelles), Belgium. During the same year, his Master Thesis was awarded by a prestigious Belgian prize: Solvay Awards 2006. From 2006 to 2010, Vincent worked in two successful departments of U.L.B.: the "Scheduling Group" and the BEAMS (Bio-, Electro- And Mechanical Systems). In 2010, he earned his Ph.D. degree in Computer Science at U.L.B under the supervision of Prof. Joël Goossens. He is currently working at CISTER/IPP-HURRAY as a research scientist in the area of real-time scheduling theory and real-time operating systems.

Patrick Meumeu Yomsi received his Ph.D. degree in 2009 at the Université Paris Sud, Orsay in France. He is currently a Belgian National Fund for Scientific Research (F.N.R.S.) Postdoctoral Researcher in the Scheduling group of the Computer Science Department at the Université Libre de Bruxelles (U.L.B.). Before this position, he was a member of the Project-Team Models and Methods of Analysis and Optimization for Systems with Real-Time and Embedding constraints (AOSTE) at the French National Institute for Research in Computer Science and Control (INRIA). His research interests include real-time scheduling theory, real-time scheduling algorithms and real-time operating systems.

Björn Andersson received his master of science degree in electrical engineering at Chalmers University of Technology in Sweden in 1999 and his Ph.D. degree at Chalmers University of Technology. He is Research Scientist at CISTER Research unit at ISEP/IPP and is the author of 60 conference papers, 9 journal articles, 2 book chapters. Research interests: sensor networks, cyber-physical systems, real-time scheduling, real-time communication, data aggregation.

Joël Goossens is Associate Professor at the Université Libre de Bruxelles (U.L.B.), since October 2006. He founded and chairs the “Parallel Architectures for Real-Time Systems” research group. U.L.B. Joël Goossens received his M.Sc. degree in computer science in 1992, his M.Sc. degree in networks and management in 1993 and his Ph.D. degree in computer science in 1999, all from the Université Libre de Bruxelles, Belgium. He teaches algorithms and programming, operating systems and real-time scheduling. His main research interests are presently in real-time scheduling systems.