The Schedulability Region of Two-Level Mixed-Criticality Systems based on EDF-VD

Dirk Müller and Alejandro Masrur
Department of Computer Science
TU Chemnitz, Germany

Abstract—The algorithm Earliest Deadline First with Virtual Deadlines (EDF-VD) was recently proposed to schedule mixed-criticality task sets consisting of high-criticality (HI) and low-criticality (LO) tasks. EDF-VD distinguishes between HI and LO mode. In HI mode, the HI tasks may require executing for longer than in LO mode. As a result, in LO mode, EDF-VD assigns virtual deadlines to HI tasks (i.e., it uniformly downscales deadlines of HI tasks) to account for an increase of workload in HI mode. Different schedulability conditions have been proposed in the literature; however, the schedulability region to fully characterize EDF-VD has not been investigated so far. In this paper, we review EDF-VD’s schedulability criteria and determine its schedulability region to better understand and design mixed-criticality systems. Based on this result, we show that EDF-VD has a schedulability region being around 85% larger than that of the Worst-Case Reservations (WCR) approach.

Keywords—real-time scheduling; mixed criticality; EDF-VD; resource efficiency

I. INTRODUCTION

In real-time systems, deadlines need to be guaranteed. To this end, a task model is used based on which a schedulability analysis is performed. Usually, this task model consists of a minimum inter-arrival time and a worst-case execution time (WCET) [1]. However, it is a challenge to determine the WCET of a task [2].

For the known two basic approaches, there is a trade-off between safety and resource utilization. Static methods yield safe results [2]. However, they are often pessimistic in terms of resource utilization.

On the other hand, measurement-based methods allow for a better resource utilization, but do not provide any guarantees on the safety of results since it is nearly impossible to cover all execution paths by measurement. Recently, techniques have been proposed combining static and measurement-based approaches [3]. Nevertheless, the problem of finding a proper trade-off between safety and resource utilization remains an open challenge.

Mixed-criticality (MC) systems are characterized by scheduling tasks of different importance (i.e., different criticality). In this paper, we consider two such levels: low-criticality (LO) level, and high-criticality (HI) level. All jobs of HI tasks must be certified to meet their deadlines under all possible circumstances, whereas those of LO tasks are not subject to such a certification [4].

Clearly, to certify a HI task, static methods should be used to determine the safe WCETs of HI tasks. However, as mentioned before, this leads to a conservative estimation and tends to overdesign. To allow for an increased resource efficiency, optimistic (measurement-based) WCET estimations can be used. In the latter case, additional provisions have to be made in order to take potential WCET overruns of jobs of HI tasks into account.

Towards this, the seminal algorithm Earliest Deadline First with Virtual Deadlines (EDF-VD) was recently proposed [5]. EDF-VD distinguishes between two modes: HI and LO mode. In LO mode, HI tasks are scheduled together with LO tasks, where optimistic WCETs are used for the HI tasks. In HI mode, only the HI tasks are allowed to run using their conservative WCETs. Here, the algorithm EDF-VD discards all LO tasks in order to accommodate an increase in the execution demands of HI tasks.

Different schedulability conditions have been published for EDF-VD [5] [6]. However, a full characterization of its schedulability region has not been proposed so far. In this paper, we review EDF-VD’s schedulability criteria and determine its schedulability region to fully characterize its behavior. Schedulability regions have been already used in the literature in order to better understand and compare schedulability criteria [7] [8].

Similarly, in this work, we compare EDF-VD with the well-known Worst-Case Reservations (WCR) approach both graphically and analytically based on schedulability regions.

A. Contributions

In this section, we state the contributions of this paper. These can be summarized in the following manner:

- We review the schedulability criteria of EDF-VD from the literature and derive a new, more intuitive, schedulability condition.
- We determine EDF-VD’s schedulability region which allows us to better understand MC systems.
- We compare schedulability regions of EDF-VD and WCR both in an analytical and graphical manner. We show that EDF-VD’s schedulability region is around 85% larger than that of WCR.
B. Structure of the Paper

The rest of this paper is structured as follows. Related work is discussed in Section II. Next, Section III explains the task and system model. We shortly revisit WCR in Section IV. Section V introduces EDF-VD and proposes a new upper bound on the design parameter deadline scaling factor. Next, EDF-VD’s necessary and sufficient schedulability conditions are discussed in Sections VI and VII. Based on this, we study the EDF-VD schedulability region in Section VIII and compare it to the one of WCR. Finally, some examples and concluding remarks will be given in Sections IX and X.

II. RELATED WORK

Mixed-criticality scheduling was first proposed under this name by Vestal [4]. Barua et al. later analyzed per-task priority assignments and the resulting response times for sporadic task sets [9].

In [5], Barua et al. proposed the EDF-VD algorithm to schedule MC sporadic task sets with priority promotion by scaling deadlines of HI tasks. The speed-up factor of 1.619 (by rounding up the golden ratio \( \frac{\sqrt{5}+1}{2} \)) proposed in [5] was later improved to 4/3 in [6].

For multiprocessor MC scheduling, a partitioned and a global scheduling approach both based on EDF-VD are proposed in [10]. According to this work, partitioned scheduling behaves better than the global approach for MC task sets.

Further, in [11], Pathan studies global MC scheduling with task-level fixed priorities and gives a schedulability test based on response time analysis for > 2 criticality levels.

A more flexible approach with per-task deadline scaling factors was presented by Ekberg and Yi [12]. However, this approach makes it hard to derive utilization and speed-up bounds due to the big number of scaling factors, i.e., tuning parameters, involved.

Recently, other improvements compared to EDF-VD were proposed. In [13], Su and Zhu used an elastic task model [14] [15] to improve resource utilization in MC systems. Finally, in [16], Zhao et al. applied preemption thresholds [17] in MC scheduling in order to better utilize the processor.

III. TASK AND SYSTEM MODEL

We consider \( n \) independent sporadic tasks \( \tau_i \) with their instances or jobs separated by a minimum inter-arrival time \( T_i \) under preemptive, implicit-deadline (i.e., \( \forall i: D_i = T_i \) where \( D_i \) is a task’s relative deadline), uniprocessor scheduling. There is no self-suspension, and context-switch overheads are assumed to be zero.

In this paper, only dual-criticality systems with two levels of criticality are examined. In this MC setting, an additional task parameter is required: criticality denoted by \( \chi_i \in \{ \text{LO}, \text{HI} \} \). For LO tasks, there is one WCET \( C_i^{LO} \). Opposed to this, HI tasks are provided with an optimistic WCET \( C_i^{LO} \) and a conservative WCET \( C_i^{HI} \) being \( C_i^{LO} \leq C_i^{HI} \).

As mentioned before, the system operates in two modes: LO and HI mode. Initially, we assume the mode \( m \) of the system to be \( m = \text{LO} \) where all tasks are scheduled.

As soon as a job of a HI task executes for longer than its \( C_i^{LO} \), the system switches to HI mode where only the subset of HI tasks are subject to their deadlines. All active jobs of LO tasks are immediately discarded and no further such jobs will be allowed to run.

In a short notation, a LO task is then specified as \( (T_i, C_i^{LO}) \) while a HI task is expressed as \( (T_i, (C_i^{LO}, C_i^{HI})) \).

We define the utilization parameters with \( \chi, m \in \{ \text{LO}, \text{HI} \} \) where \( \chi \) denotes task criticality and \( m \) mode.

\[
u_m^\chi := \sum_{\chi_i = \chi} \frac{C_i^m}{T_i}
\]

Note that among the four potential criticality-to-mode combinations, only \( u_m^{LO}, u_m^{LO} \) and \( u_m^{HI} \) are defined. \( u_m^{LO} \) does not exist since there is no \( C_i^{HI} \) specified for LO tasks. Since \( C_i^{LO} \leq C_i^{HI} \) holds for all \( i \) for which \( \chi_i = \text{HI} \), i.e., the optimistic WCET is always less than or equal to the conservative WCET of all HI tasks, we obtain the following constraint on utilizations \( u_m^{LO} \) and \( u_m^{HI} \).

\[
 u_m^{LO} \leq u_m^{HI}
\]

IV. WORST-CASE RESERVATIONS

In the simple strategy of WCR, the optimistic \( C_i^{LO} \) of HI tasks in LO mode is replaced by the conservative \( C_i^{HI} \). It then applies the well-known EDF scheduling strategy. This way, we obtain (3) as an exact schedulability condition for the WCR approach using [1].

\[
 u_m^{LO} + u_m^{HI} \leq 1
\]

The WCR approach assumes that HI tasks always require their maximum WCETs. This is pessimistic (with respect to EDF-VD) since it ignores the fact that, in LO mode, the optimistic WCETs could be used.

V. THE EDF-VD ALGORITHM

The algorithm EDF-VD [5] is an adaption of the classic EDF [1] to MC scheduling. Its basic idea is to promote HI jobs in LO mode by shortening their relative deadlines in order to reserve processor time for the HI mode.

The switch from LO to HI mode is triggered by a HI job exceeding its \( C_i^{LO} \). The EDF-VD approach is to scale these relative deadlines homogeneously in LO mode. That is, \( D_i' = x \times T_i \) with \( x \in [0, 1] \) for all \( i \). \( D_i' \) is referred to as virtual deadline and is used to schedule HI tasks in LO mode. Opposed to this, there is no deadline scaling for LO tasks, they keep their original deadlines \( D_i \). In HI mode, again the original deadlines \( D_i \) are used to schedule tasks according to EDF\(^1\). As mentioned above, LO tasks are discarded in HI mode.

\(^1\)Note that EDF relies on absolute deadlines such that using virtual deadlines in HI mode might lead to inconsistencies.
Based on these considerations, Baruah et al. [6] found a utilization bound of $3/4$, see (4) below.

$$\max(u_{LO}^L + u_{HI}^L, u_{HI}^H) \leq 3/4 = 0.75 \quad (4)$$

A lower bound (5) of the scaling factor $x$ can be used as a sufficient condition for meeting all deadlines in LO mode.

$$\frac{u_{HI}^L}{1 - u_{LO}^L} \leq x \quad (5)$$

An upper bound (6) of $x$ can be found by considering the transition between LO and HI mode [6]. This serves as a sufficient schedulability condition for the HI mode.

$$x \leq \frac{1 - u_{HI}^H}{u_{LO}^L} \quad (6)$$

### A Straightforward Upper Bound on the Scaling Factor

Let us assume that the system switches from LO to HI mode at a time $t$ – recall this happens when one of the HI tasks executes for $C_{i}^{LO}$ time without signaling end of execution. As a result, no LO tasks will be executed after $t$. In addition, since (5) holds, $u_{LO}^L + \frac{u_{HI}^H}{x} \leq 1$ also holds and it can be guaranteed that all HI tasks execute for $C_{i}^{LO}$ time within their corresponding virtual deadlines (i.e., $D_{i}^{V} = x \times T_{i}$). Let us assume that only one task causes the change to HI mode, i.e., it has executed for $C_{i}^{LO}$ at $t$ without signaling its end. Now, since (5) holds guaranteeing the schedulability of all tasks in LO mode, $t$ has to be less than or equal to the virtual deadline of the job triggering the mode switch. Otherwise (5) would not hold. In worst case, $t$ is exactly equal to the deadline of that job. For this job to finish before its deadline in HI mode, it has to execute $C_{i}^{HI}$ time within $T_{i} - x \times T_{i}$, i.e., within an interval equal to the difference between its original and virtual deadline. As a result, the following inequalities must hold.

$$\frac{C_{i}^{HI}}{C_{i}^{HI} - C_{i}^{LO}} \leq \frac{T_{i} - xT_{i}}{T_{i}} \leq 1$$

The worst-case situation happens when all HI tasks switch to HI mode simultaneously\(^2\) at time $t$ and hence the following must hold for all of them to be schedulable, cf. [5].

$$\frac{u_{HI}^H}{1 - x} \leq 1 \quad (7)$$

$$x \leq 1 - u_{HI}^H$$

On the other hand, a HI job has already executed for $C_{i}^{LO}$ at $t$ and, hence, it only has to execute for additional $C_{i}^{HI} - C_{i}^{LO}$ time within $T_{i} - x \times T_{i}$ (instead of $C_{i}^{HI}$ as assumed in [5]). As a result, the following must hold.

$$\frac{C_{i}^{HI} - C_{i}^{LO}}{1 - xT_{i}} \leq 1$$

$$\frac{C_{i}^{HI} - C_{i}^{LO}}{(1 - x)T_{i}} \leq 1$$

Again, the worst-case situation happens when all HI tasks switch to HI mode simultaneously at time $t$. As a result, we obtain the following expression.

$$\frac{u_{HI}^H - u_{HI}^L}{1 - x} \leq 1 \quad (8)$$

### VI. Necessary Schedulability Criteria for EDF-VD

Since EDF-VD applies traditional EDF in both LO and HI mode, two necessary EDF-VD schedulability conditions (9) and (10) can be directly derived using the exact EDF uniprocessor utilization bound [1].

$$u_{LO}^L + u_{HI}^L \leq 1 \quad (9)$$

$$u_{HI}^H \leq 1 \quad (10)$$

### VII. Sufficient Schedulability Criteria for EDF-VD

Unfortunately, conditions (9) and (10) are not sufficient for EDF-VD to be schedulable since these disregard the schedulability of mode transitions.

The existence of a lower and an upper bound on the scaling factor $x$ allows us to derive a third condition to guarantee schedulability of mode transitions. Note that this is possible since the upper bound on $x$ – see (8) – is obtained taking mode transitions into account.

If a valid $x$ exists, (11) must hold from (5) and (8).

$$\frac{u_{HI}^H}{1 - u_{LO}^L} \leq 1 - u_{HI}^H + u_{HI}^L \quad (11)$$

Condition (11) can be easily shown to be equivalent (i.e., if one of them holds, then all of them hold) to (12) as given in [6] and to (13) as proposed in [10] – see Appendix A.

$$\frac{u_{HI}^H}{1 - u_{LO}^L} \leq \frac{1 - u_{HI}^H}{u_{LO}^L} \quad (12)$$

$$u_{LO}^L \leq \frac{1 - u_{HI}^H}{1 - u_{HI}^H} \quad (13)$$

Note that (12) is based on (5) and (6), while (13) results from solving (12) for $u_{LO}^L$.

Another sufficient schedulability condition (14) results from combining the more conservative upper bound of (7) with (5) as proposed in [5].

$$\frac{u_{HI}^H}{1 - u_{LO}^L} \leq 1 - u_{HI}^H \quad (14)$$
VIII. THE SCHEDULABILITY REGION

A. EDF-VD

Fig. 1 illustrates EDF-VD’s schedulability region. The abscissas represent values of $u_{LO}^O$ whereas the ordinates represent values of $u_{HI}^O$. Clearly, since (9) must hold for EDF-VD to be schedulable, the schedulability region of EDF-VD must be below the line crossing by \((u_{HI}^O = 1, u_{LO}^O = 0)\) and \((u_{HI}^O = 0, u_{LO}^O = 1)\).

As discussed previously, the utilization bound of EDF-VD is $3/4$ as shown in (4). This is represented by the lower line crossing by \((u_{HI}^O = 3/4, u_{LO}^O = 0)\) and \((u_{HI}^O = 0, u_{LO}^O = 3/4)\) in Fig. 1. That is, utilization values that are below this lower line are schedulable without needing further tests. Clearly, from (4), nothing can be concluded if the utilization is between $3/4$ and 1, i.e., for the section between the lower and the upper line in Fig. 1.

To characterize the schedulability region between these two lines, let us set $u_{HI}^O = 3/4$. Again, this is the maximum value of $u_{HI}^O$ by which schedulability is guaranteed by (4). Now, replacing this value into (11) and solving for $u_{LO}^O$, we obtain (15) – note that this can also be obtained by replacing $u_{HI}^O = 3/4$ and solving for $u_{LO}^O$ in (12) or (13).

$$u_{LO}^O \leq \frac{1 - u_{LO}^O}{4u_{LO}^O} \quad (15)$$

Inequality (15) expresses how $u_{LO}^O$ varies with respect to $u_{LO}^O$ under a fixed $u_{HI}^O = 3/4$. This is a hyperbolic function of $u_{LO}^O$ as shown in Fig. 1. All pairs of $u_{LO}^O$ and $u_{HI}^O$ below this curve are schedulable under EDF-VD, whereas values of $u_{LO}^O$ and $u_{HI}^O$ above it cannot be guaranteed to be schedulable. Since condition (2) must hold, the hyperbolic curve is truncated at $u_{HI}^O = u_{HI}^O = 3/4$. EDF-VD’s schedulability region is given by the area below the truncated hyperbolic curve. It should be noticed that the area below the line given by (4) is fully contained below the truncated hyperbolic curve. Now, considering $u_{HI}^O = 3/4$ and solving for $u_{LO}^O$ in (14) we obtain (16) illustrated in Fig. 1.

$$u_{HI}^O \leq 1/4 - \frac{u_{LO}^O}{4} \quad (16)$$

B. WCR

Let us now draw the schedulability region for WCR. To this end, we set $u_{HI}^O = 3/4$ and solving for $u_{LO}^O$ in (3) we obtain (17), which is a constant as illustrated in Fig. 1.

$$u_{LO}^O \leq 1 - 3/4 \quad (17)$$

This means that all pairs \((u_{HI}^O, u_{LO}^O)\) are schedulable under WCR if $u_{LO}^O$ is not greater than 1/4. This results in the area to the left of $u_{LO}^O = 1/4$ as illustrated in Fig. 1. Again, since condition (2) must hold, WCR’s schedulability region is limited from above by $u_{HI}^O = u_{HI}^O = 3/4$.

C. Comparing EDF-VD and WCR

Clearly, EDF-VD outperforms WCR. That is, if a set of MC tasks is schedulable under WCR, it will also be schedulable under EDF-VD; cf. Fig. 1, where WCR’s schedulability region is fully contained by EDF-VD’s one.

In this section, we further make a quantitative comparison between these two algorithms. To this end, let us assume that $u_{HI}^O = u_{LO}^O$ holds. Replacing this in (15) and reshaping we obtain a quadratic expression in $u_{LO}^O$.

$$4(u_{LO}^O)^2 + u_{LO}^O - 1 \leq 0 \quad (18)$$

Equalizing (18) to zero, we can compute its roots and take the positive one as given by (19).

$$u_{LO}^O = \frac{-1 + \sqrt{17}}{8} \approx 0.39 \quad (19)$$

Hence, for $u_{HI}^O = u_{LO}^O$ and $u_{HI}^O = 3/4$, a set of MC tasks is schedulable under EDF-VD if $u_{LO}^O$ is at most 0.39. On the other hand, for WCR, we obtain from (3) that $u_{LO}^O \leq 0.25$.

Note that this is more than 56% more allowable utilization for EDF-VD compared to WCR.

The above comparison is valid for the particular case where $u_{HI}^O = 3/4$. To obtain a more general comparison of EDF-VD with WCR, let us reshape (12) as follows – note that (11) or (13) can also be reshaped to this form.

$$u_{HI}^O \leq \frac{(1 - u_{HI}^O)(1 - u_{LO}^O)}{u_{LO}^O} \quad (20)$$

Inequality (20) describes the general case for $u_{HI}^O$. There, we have a scaling by $1 - u_{HI}^O$ instead of 1/4 as in (15).

By decreasing $u_{HI}^O$ from 3/4 towards 0 – note that $u_{HI}^O \geq u_{LO}^O$ must still hold, the boundary of EDF-VD’s
schedulability region (20) approaches the necessary condition (9). The corner cases at \( (u_{HI}^{LO} = 0, u_{LO}^{LO} = 1) \) and \( (u_{HI}^{HI} = u_{HI}^{LO}, u_{LO}^{LO} = 1 - u_{HI}^{HI}) \) remain. The advantage of EDF-VD compared to WCR worsens with a decreasing \( u_{HI}^{HI} \) since, in this case, \( u_{HI}^{HI} \) approaches \( u_{HI}^{LO} \) and, hence, makes reservations less wasteful. Clearly, increasing \( u_{HI}^{HI} \) from 3/4 towards 1 improves the advantage of EDF-VD over WCR.

In general, the area of a schedulability region is a measure of performance for a given scheduling algorithm\(^3\). Thus, the area ratio of two schedulability regions allows quantifying the advantage or efficiency of one algorithm with respect to another, see also [18] and [8].

Here, we determine the area ratio of WCR over EDF-VD. From Fig. 1, WCR's region (a rectangle) has an area of (21).

\[
A_{\text{WCR}} = u_{HI}^{HI} \left( 1 - u_{HI}^{HI} \right) \quad (21)
\]

The EDF-VD schedulability region comprises the one of WCR plus the area under a hyperbolic curve given by (20). According to Fig. 1 and (20), the difference in the area can be calculated as the definite integral (22).

\[
A_{\text{EDF-VD}} - A_{\text{WCR}} = \int_{1-u_{HI}^{HI}}^{1} \frac{(1-u_{LO}^{LO})(1-u_{HI}^{HI})}{u_{LO}^{LO}} \, du_{LO}^{LO} = (u_{HI}^{HI} - 1) \ln(1-u_{HI}^{HI}) + u_{HI}^{HI} - u_{HI}^{LO} \quad (22)
\]

Finally, (21) and (22) give the area ratio of WCR over EDF-VD as a function of \( u_{HI}^{HI} \) (23), cf. Fig. 2, see Appendix B.

\[
\text{Eff}_{\text{WCR-to-EDF-VD}} = \frac{A_{\text{WCR}}}{A_{\text{EDF-VD}}} = -\frac{u_{HI}^{HI}}{\ln(1-u_{HI}^{HI})} \quad (23)
\]

For \( u_{HI}^{HI} = 0 \), WCR is as efficient as EDF-VD. With increasing \( u_{HI}^{HI} \), WCR's efficiency decreases. Up to \( u_{HI}^{HI} \approx 0.9 \), this decrease is almost linear with a slope of \(-2/3\). From 0.9 onwards, WCR's relative performance rapidly falls to zero at \( u_{HI}^{HI} = 1 \). While WCR might be acceptable for small \( u_{HI}^{HI} \) values, high \( u_{HI}^{HI} \) values call for EDF-VD.

For \( u_{HI}^{HI} = 3/4 \), EDF-VD's advantage to WCR is obtained from (23) by taking the inverse, giving ca. 1.85. EDF-VD has a schedulability area ca. 85\% greater than WCR's.

IX. EXAMPLES

First, let us consider, from [13], the task set \( \{(8, 2), (30, 3), (10, (2, 4)), (25, (4, 10))\} \). Using (1), we obtain \( u_{LO}^{LO} = 0.35, u_{HI}^{HI} = 0.36 \) and \( u_{HI}^{HI} = 0.8 \). The necessary conditions (9) and (10) are met with \( 0.71 \leq 1 \) and \( 0.8 \leq 1 \). In this example, \( u_{LO}^{LO} + u_{HI}^{HI} = 1.51 > 1 \) holds which calls for EDF-VD. (5) and (6) give an interval of [36/65, 4/7] or ca. [0.5538, 0.5714] for EDF-VD's scaling factor \( x \). Our approach (8) easily gives the sound result of 0.56 which is in the middle of the interval and, thus, increases tolerance to arrival time jitter.

We take \( \{(6, 2), (10, (1, 2)), (20, (2, 10))\} \) as a 2nd example. Based on (1), \( u_{LO}^{LO} = 1/3, u_{HI}^{HI} = 0.2 \) and \( u_{HI}^{HI} = 0.7 \) are obtained. The necessary conditions (9) and (10) are met with \( 8/15 \approx 0.533 \leq 1 \) and \( 0.7 \leq 1 \). But, \( u_{LO}^{LO} + u_{HI}^{HI} = 31/30 \approx 1.033 > 1 \) holds, again calling for EDF-VD. We obtain a min. scaling factor (5) of \( x = 0.3 \). The max. scaling factor (6) is \( x = 0.9 \). Our new approach (8) returns 0.5 for \( x \) (in the middle of the interval), improving jitter tolerance.

X. CONCLUDING REMARKS

In this paper, we obtained the schedulability region of EDF-VD. Based on this, we showed how different schedulability conditions for EDF-VD relate to each other and which among them are the most accurate. In addition, we compared the schedulability regions of EDF-VD and WCR in a graphical and analytical manner. We have shown analytically that the area of the schedulability region of EDF-VD is ca. 85\% larger than that of WCR. Our technique to compare schedulability regions is general enough and can be used to analyze the performance of scheduling algorithms and schedulability criteria in other domains.

REFERENCES


First we show by equivalent transformations that (11) implies (12) and vice versa.