CISTER
Research Center in
Real-Time & Embedded
Computing Systems

Technical Report

Technical Report: Holistic Analysis for Fork-Join Parallel Distributed Real-Time Tasks using the FTT-SE Protocol

Ricardo Garibay-Martínez*

Geoffrey Nelissen*

Luis Lino Ferreira*

Paulo Pedreiras

Luis Miguel Pinho*

*CISTER Research Center

CISTER-TR-150102

2015

Technical Report: Holistic Analysis for Fork-Join Parallel Distributed Real-Time Tasks using the FTT-SE Protocol

Ricardo Garibay-Martínez*, Geoffrey Nelissen*, Luis Lino Ferreira*, Paulo Pedreiras, Luis Miguel Pinho*

*CISTER Research Center

Polytechnic Institute of Porto (ISEP-IPP)

Rua Dr. António Bernardino de Almeida, 431

4200-072 Porto

Portugal

Tel.: +351.22.8340509, Fax: +351.22.8321159

E-mail: rgmaz@isep.ipp.pt, grrpn@isep.ipp.pt, llf@isep.ipp.pt, lmp@isep.ipp.pt

http://www.cister.isep.ipp.pt

Abstract

Modern distributed real-time embedded applications have high processing requirements associated with strict deadlines. Such constraints cannot be fulfilled by existing single-core embedded platforms. A solution is to parallelise the execution of the applications, by allowing networked nodes to distribute their workload to remote nodes with spare capacity in the system. In this context, this paper presents a holistic timing analysis for fixed-priority fork-join parallel distributed tasks (P/D tasks). Furthermore, we extend the holistic approach to consider the interaction between parallel threads and messages interchanged through a Flexible Time Triggered - Switched Ethernet (FTTSE) network, and we show how the pessimism on the Worst-Case Response Time (WCRT) computation of such tasks can be improved by considering a pipeline effect in such distributed systems.

Technical Report: Holistic Analysis for Fork-Join Parallel Distributed Real-Time Tasks using the FTT-SE Protocol

Ricardo Garibay-Martínez*, Geoffrey Nelissen*, Luis Lino Ferreira*, Paulo Pedreiras†, Luís Miguel Pinho*
* CISTER/INESC-TEC Research Centre, ISEP/IPP
Polytechnic Institute of Porto, Porto, Portugal
{rgmaz, grrpn, llf, lmp}@isep.ipp.pt
†IEETA/University of Aveiro, Aveiro, Portugal
pbrp@ua.pt

In this technical report we present the proofs for **Theorem 2** and **Theorem 2**, related to the paper Holistic Analysis for Fork-Join Parallel Distributed Real-Time Tasks using the FTT-SE Protocol. The improvement is based on a pipeline effect that occurs when simultaneously transmitting P/D messages on an FTT-SE network and executing their respective P/D threads on remote nodes.

A. Overlap on the downlink

Let us consider the two following situations:

i) assume that a low priority thread θ^l is executing on a remote processor node π_i. If the execution of a thread of higher priority θ^h is triggered on π_i, θ^l is preempted by θ^h. However, because messages are non-preemptible, the message μ^l that triggered the execution of θ^l must have reached π_i before the message μ^h could start being transmitted, thereby implying that the transmission of μ^h occurred in parallel with the execution of θ^l.

ii) assume a thread of high priority θ^h executing on a remote node processor π_i. If the execution of a lower priority thread θ^l is triggered on π_i, θ^l is delayed until θ^h completes its execution. Similarly to the previous case, because only one message can be transmitted at a time, it implies that the transmission of μ^l occurred in parallel with the execution of θ^h.

Let $\mathsf{IntT}(\theta_{i,j,k})$ be the set of jobs that participated to the WCRT of a thread $\theta_{i,j,k}$ (including the job $\theta_{i,j,k}$ itself) on a remote processor node π_ℓ. And let $\mathsf{IntM}(\theta_{i,j,k})$ be the set of messages that participated to the WCRT of $\mu_{i,j-1,k}$ and triggered the execution of jobs in $\mathsf{IntT}(\theta_{i,j,k})$. Extrapolating the two situations discussed above, we can conclude that:

Property 1. *Only one message in $\mathsf{IntM}(\theta_{i,j,k})$ was not transmitted in parallel with the execution of the jobs in $\mathsf{IntT}(\theta_{i,j,k})$. This message is the message of the first job in $\mathsf{IntT}(\theta_{i,j,k})$ that started executing on π_ℓ.*

In the worst-case, the message that did not overlap with the response time of $\theta_{i,j,k}$ is the message with the largest WCML in $\mathsf{IntM}(\theta_{i,j,k})$. Let $M^{\max}_{i,j,k}$ be the WCML of that message. Then, the overlap $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$ is lower bounded by:

$$\mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k}) = \sum_{\mu_{p,q,r}\in \mathsf{IntM}(\theta_{i,j,k})} \mu_{p,q,r} - M^{\max}_{i,j,k} \quad (20)$$

Therefore, denoting $\mathsf{RT}(\mu_{i,j-1,k}+\theta_{i,j,k})$ the response time of a P/D message $\mu_{i,j-1,k}$, and its corresponding thread $\theta_{i,j,k}$ during a D-Fork operation, we prove the following theorem.

Theorem 1. *The response time $\mathsf{RT}(\mu_{i,j-1,k}+\theta_{i,j,k})$ is upper bounded by:*

$$\mathsf{RT}(\mu_{i,j-1,k}+\theta_{i,j,k}) \le \mathsf{WR}(\mu_{i,j-1,k}) + \mathsf{WR}(\theta_{i,j,k}) - \mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k})$$

where $\mathsf{WR}(\mu_{i,j-1,k})$ and $\mathsf{WR}(\theta_{i,j,k})$ are computed with Eq. (7) increased by 1 or 3 ECs (see Section VI-A) and Eq. (13).

Proof. The proof is done by contradiction. Assume that there exists a scenario such that:

$$\mathsf{RT}(\mu_{i,j-1,k}+\theta_{i,j,k}) > \mathsf{WR}(\mu_{i,j-1,k}) + \mathsf{WR}(\theta_{i,j,k}) - \mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k}) \quad (21)$$

We know that there is an overlap $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$ between the transmission of the messages and the execution of the threads participating to the response time of $\mu_{i,j-1,k}$ and $\theta_{i,j,k}$. Therefore, we have at the least:

$$\mathsf{RT}(\mu_{i,j-1,k}+\theta_{i,j,k}) \le \mathsf{WR}(\mu_{i,j-1,k}) + \mathsf{WR}(\theta_{i,j,k}) - \mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$$

This implies that Eq. (21) is true iff $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k}) < \mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k})$. The only possible reason for such a situation to happen, is that at least one transmission of a message $\mu^h_{i,j-1,k} \in \mathsf{IntM}(\theta_{i,j,k})$ accounted in $\mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k})$ does not contribute to $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$. Assume that there is only one such instance[1]. Then,

$$\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k}) = \mathsf{OvF}^{*}_{\pi_i,l_{SW_x}}(\theta_{i,j,k}) - M^h_{i,j-1,k} \quad (22)$$

[1]If multiple message instances in $\mathsf{IntM}(\theta_{i,j,k})$ do not contribute to $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$, then the reasoning developed in the following of this proof can be applied iteratively by considering one more instance at each iteration.

Two cases must be considered:

1) the thread $\theta^h_{i,j,k}$ triggered by $\mu^h_{i,j-1,k}$ does not interfere with the execution of $\theta_{i,j,k}$. This implies that $\mathsf{RT}(\theta_{i,j,k}) \leq \mathsf{WR}(\theta_{i,j,k}) - C^h_{i,j,k}$, and because by assumption $C^h_{i,j,k} \geq M^h_{i,j-1,k}$, we get:

$$\begin{aligned}&\mathsf{RT}(\theta_{i,j,k} + \mu_{i,j-1,k})\\ \leq&\mathsf{WR}(\theta_{i,j,k}) - C^h_{i,j,k} + \mathsf{WR}(\mu_{i,j-1,k}) - \mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})\\ \leq&\mathsf{WR}(\theta_{i,j,k}) - M^h_{i,j,k} + \mathsf{WR}(\mu_{i,j-1,k}) - \mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})\\ \leq&\mathsf{WR}(\theta_{i,j,k}) + \mathsf{WR}(\mu_{i,j-1,k}) - \mathsf{OvF}^*_{\pi_i,l_{SW_x}}(\theta_{i,j,k})\end{aligned}$$

thereby contradicting Eq. (21).

2) the thread $\theta^h_{i,j,k}$ triggered by $\mu^h_{i,j-1,k}$ interferes with the execution of $\theta_{i,j,k}$. Because by Property 1, only one message in $\mathsf{IntM}(\theta_{i,j,k})$ does not contribute to $\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k})$, we get:

$$\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k}) = \sum_{\mu_{p,q,r}\in\mathsf{IntM}(\theta_{i,j,k})} \mu_{p,q,r} - M^h_{i,j,k}$$

And using Eq. (20)

$$\begin{aligned}\mathsf{OvF}^{\pi_i}_{l^d_{SW_x}}(\theta_{i,j,k}) &= \mathsf{OvF}^*_{\pi_i,l_{SW_x}}(\theta_{i,j,k}) + M^{\max}_{i,j-1,k} - M^h_{i,j,k}\\ &\geq \mathsf{OvF}^*_{\pi_i,l_{SW_x}}(\theta_{i,j,k})\end{aligned}$$

which contradicts Eq. (22) and therefore Eq. (21).

Consequently, Eq. (21) can never be true. □

B. Non-interference on the uplink

If all the P/D threads $\theta_{i,j,k}$ of a same parallel segment $\sigma_{i,j}$ share the same priority, then they do not preempt each other when executing on the same remote node π_ℓ. Consequently, the messages $\mu_{i,j,k}$ sent by those threads from π_ℓ to their master processor, start their transmissions at least $C_{i,j,k}$ time units apart. Because by assumption the WCML $M_{i,j,k}$ is smaller than or equal to the WCET $C_{i,j,k}$ of the threads triggering their execution, a non-interference effect $\mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,k})$ between the messages of the same P/D segment sent from the same remote node occurs during the D-join operation. This effect is given by:

$$\mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,k}) = \sum_{\substack{\forall \mu_{i,j,p}\in\pi_i\\ p\neq k}} M_{i,j,p}$$

where $\mu_{i,j,p} \in \pi_i$ means that the message $\mu_{i,j,p}$ has π_i as service node. This gives the following theorem:

Theorem 2. *The response time $\mathsf{RT}(\mu_{i,j,k})$ of a P/D message $\mu_{i,j,k}$ during a D-Join operation is upper bounded by:*

$$\mathsf{RT}(\mu_{i,j,k}) \leq \mathsf{WR}(\mu_{i,j,k}) - \mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,k})$$

Proof. Let us assume that only two messages[2] of the same segment $\sigma_{i,j}$ are sent from the remote processor π_ℓ. Let us denote them by $\mu_{i,j,1}$ and $\mu_{i,j,2}$ and assume that $\mu_{i,j,1}$ is the first to be triggered. Therefore, $\mu_{i,j,2}$ does not participate to the response time of $\mu_{i,j,1}$, yet Eq. (7) assumes that $\mu_{i,j,2}$ interferes with $\mu_{i,j,1}$. Therefore, we have:

$$\mathsf{RT}(\mu_{i,j,1}) \leq \mathsf{WR}(\mu_{i,j,1}) - M_{i,j,1} = \mathsf{WR}(\mu_{i,j,1}) - \mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,1})$$

Two cases must be considered for $\mu_{i,j,2}$:

1) $\mu_{i,j,1}$ and $\mu_{i,j,2}$ are triggered in the same EC. Because $\mu_{i,j,2}$ was triggered at least $M_{i,j,1}$ time units after $\mu_{i,j,1}$, the node queuing delay cannot be longer than $|EC| - M_{i,j,1}$. Since $\mathsf{WR}(\mu_{i,j,2})$ always considers a node queuing delay of $1 \times |EC|$ (see Section VI-A), there is:

$$\mathsf{RT}(\mu_{i,j,2}) \leq \mathsf{WR}(\mu_{i,j,2}) - M_{i,j,2} = \mathsf{WR}(\mu_{i,j,2}) - \mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,2})$$

2) $\mu_{i,j,1}$ and $\mu_{i,j,2}$ are triggered in different ECs. If $\mu_{i,j,1}$ already completed its transmission, then it does not interfere with $\mu_{i,j,2}$ and the theorem obviously holds. Otherwise, if $\mu_{i,j,1}$ is still waiting to be transmitted when $\mu_{i,j,2}$ is triggered, then it means that $\mu_{i,j,1}$ was delayed by higher priority messages for a time at least equal to the length LW of its transmission window. Those higher priority messages cannot interfere with $\mu_{i,j,2}$ anymore and because $LW \geq M_{i,j,1}$, the theorem holds for $\mu_{i,j,2}$.

□

In conclusion, combining the results of Theorem 1 and 2, we can improve the WCRT of a distributed execution path (Eq. (12)) as follows:

$$\begin{aligned}\mathsf{WR}(DP_{i,2j,k}) &= \mathsf{WR}(\mu_{i,2j-1,k}) + \mathsf{WR}(\theta_{i,2j,k})\\ &+ \mathsf{WR}(\mu_{i,2j,k}) - \mathsf{OvF}^*_{\pi_i,l_{SW_x}}(\theta_{i,j,k}) - \mathsf{OvJ}^{l^d_{SW_x}}_{\pi_i}(\mu_{i,j,k}).\end{aligned}$$

[2] If more than two messages of the same segment should be sent from the same remote processor, the proof still holds by applying the argumentation iteratively, adding one more message at each iteration.