Technische Universiteit
Eindhoven
University of Technology

Dataflow Model for Credit-Controlled Static-Priority Arbitration

Firew Siyoum, Benny Akesson, Sander Stuijk, Kees Goossens,
and Henk Corporaal

Version 1.1

ES Reports
ISSN 1574-9517

ESR-2010-03
15 October 2010

Eindhoven University of Technology
Department of Electrical Engineering
Electronic Systems

Note:
This technical note is mainly intended to present the mathematical proof of the dataflow model of Credit-Controlled Static-Priority (CCSP) arbiter, which is omitted from the paper:

Resource-Efficient Real-time Scheduling Using Credit-Controlled Static-Priority Arbitration
Firew Siyoum, Benny Akesson, Sander Stuijk, Kees Goossens, *and Henk Corporaal*
Submitted to ECRTS-2011.

Therefore, it assumes that the reader is familiar with the CCSP arbiter as well as the formalism used for the analysis of the arbiter. If the reader does not have any prior acquittance with CCSP, it is recommended to read the following paper before going through this document.

Real-Time Scheduling Using Credit-Controlled Static-Priority Arbitration
Benny Åkesson, Liesbeth Steffens, Eelke Strooisma, and Kees Goossens
In: *Int'l Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA) 2008*

http://www.es.ele.tue.nl/esreports
esreports@es.ele.tue.nl

Eindhoven University of Technology
Department of Electrical Engineering
Electronic Systems
PO Box 513
NL-5600 MB Eindhoven
The Netherlands

Dataflow Model for Credit-Controlled Static-Priority Arbitration

Firew Siyoum, Benny Akesson, Sander Stuijk, Kees Goossens, Henk Corporaal
Eindhoven University of Technology

Abstract

A Credit-Controlled Static Priority (CCSP) arbiter has been proposed that circumvents two key downsides of alternative arbiters in real-time embedded systems. The first is that it enables a better resource utilization by decoupling essential properties, such as allocated rate, allocation granularity and latency. Secondly, it has a fast and small hardware implementation whose correctness is formally proved. However, the existing dataflow model of the arbiter, which is based on the general model for a class of latency-rate ($\mathcal{LR}$) servers, does not capture bursty provided service. As a result, system-level analysis of CCSP leads to unnecessarily large resource allocation to satisfy application requirements. In this document, we present a new dataflow model to address this drawback. The new dataflow model is based on a piecewise linear service guarantee that accurately models the provided service by CCSP. As a result of the new dataflow model, a given resource under CCSP arbitration can support more requestors, or a given set of requestors can be accommodated with less resource.

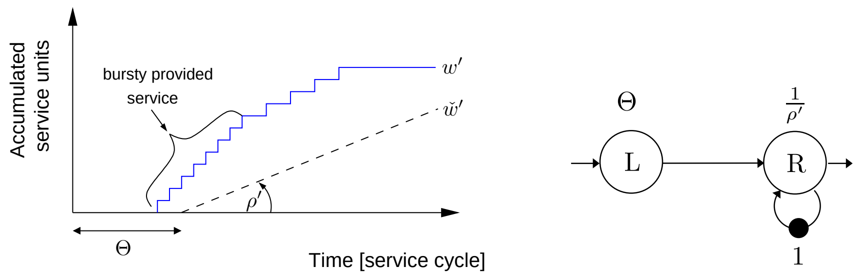
I. Introduction

The *service guarantee* of an arbiter is defined as the minimum service available to a *requestor*, irrespective of the situation of other requestors sharing the resource. This minimum service is normally computed by considering the worst-case scenario. For a requestor under a Credit-Controlled Static-Priority (CCSP) arbiter, this worst-case scenario happens when it experiences the maximum interference from higher priority requestors [1]. Given a set of requestors R that are sharing a common resource under CCSP arbitration, the set of requestors that have higher priority than requestor $r \in R$ is denoted R_r^+. Previous work [1] on the service guarantee of the CCSP arbiter shows that a requestor can be guaranteed a minimum service at its *allocated rate*, ρ', after a *maximum latency*, Θ. As illustrated in Figure 1(a), this service guarantee defines a linear lower bound, $\breve{w}'$, on the *provided service curve*, w', during a time interval, referred to as *active period*. We refer to this lower bound as the *latency-rate service guarantee*. Based on the latency-rate service guarantee, an active requestor $r \in R$ is guaranteed a minimum service during an active period $[\tau_1, \tau_2]$ according to the following relation: $\forall t \in [\tau_1, \tau_2] : \ \breve{w}'_r(\tau_1, t) = \max(0, \rho'_r \cdot (t - \tau_1 + 1 - \Theta_r))$, where

$$\Theta_r = \frac{\sum_{\forall s \in R_r^+} \sigma'_s}{1 - \sum_{\forall s \in R_r^+} \rho'_s} \tag{1}$$

The latency-rate service guarantee is used in [2] to show that CCSP arbiter belongs to the class of $\mathcal{LR}$ servers [3], which is a general framework for analyzing scheduling algorithms. In [4], a general dataflow model is presented for arbiters in the class of $\mathcal{LR}$ servers. This dataflow model, which we refer to as the *latency-rate dataflow model*, is illustrated in Figure 1(b). The latency-rate dataflow model is a two-actor homogeneous (single-rate) dataflow model that defines a linear lower bound on the provided service curve. Actor L_r, models the maximum blocking requests encounter and actor R models the time it takes to serve a request at the allocated rate. This means, a requestor $r \in R$ is guaranteed to receive service at its allocated rate, ρ'_r, after a maximum latency, Θ_r, during every active period.

However, a requestor under CCSP arbitration can receive service at a rate higher than its allocated rate, after the maximum latency, Θ_r. In other words, the provided service curve, w', can have bursty provided service intervals, as illustrated in Figure 1(a). This behavior is not captured in the linear



(a) Latency-rate service guarantee (b) Latency-rate dataflow model

Fig. 1. Latency-rate service guarantee and dataflow model for CCSP arbitration

latency-rate service guarantee and its corresponding dataflow model. This results in a pessimistic worst-case finishing time of requests, which ultimately leads to unnecessarily large resource allocation to satisfy application requirements. This drawback is addressed in [5] through a piecewise linear service guarantee for CCSP, which is referred to as *bi-rate service guarantee*.

Under the bi-rate service guarantee, a requestor $r \in R$ is guaranteed a minimum service at two different rates, ρ^*_r and ρ'_r, after a maximum latency Θ_r. This maximum latency, Θ_r, is the same as the one in the latency-rate service guarantee, which is given in Equation (1). The bi-rate guarantee is denoted $\breve{w}_r$ and is illustrated in Figure 2. It is defined based on two linear equations: a higher-rate guarantee $\breve{w}'^h_r$ and an allocated-rate guarantee $\breve{w}'^a_r$, which are given in Equation (2) and (3), respectively.

$$\breve{w}'^h_r(\tau_1, t) = \rho^*_r \cdot (t - \tau_1 + 1 - \Theta_r) \tag{2}$$

$$\breve{w}'^a_r(\tau_1, t) = \rho'_r \cdot (t - \tau_1 + 1 - \Gamma_r) \tag{3}$$

where

$$\Gamma_r = -\frac{\sigma'_r + \rho^*_r - 1}{\rho'_r} \tag{4}$$

The bi-rate service guarantee is, then, the lesser of the two at any time. It is shown in [5] that during an active period $[\tau_1, \tau_2]$, a requestor $r \in R$ is guaranteed a minimum service, $\check{\breve{w}}'_r(\tau_1, t)$, as given in Equation (5).

$$\check{\breve{w}}'_r(\tau_1, t) = max(0, min(\breve{w}'^h_r(\tau_1, t), \breve{w}'^a_r(\tau_1, t))) \tag{5}$$

The bi-rate service guarantee splits every active period into two intervals, depending on the rate at which the requestor receives service. The first interval spans from the beginning of the active period up to the time point referred to as the *boundary cycle*, τ^b_r, where a requestor receives service at the higher service rate (cf. Figure 2). The second interval spans from the boundary cycle up to the end of the active period, where the requestor receives service at its allocated rate, ρ'. For a requestor $r \in R$, the higher service rate, ρ^*_r, is given by Equation (6). From this equation, we can see that $\rho^*_r \geq \rho'_r$, since for a valid CCSP configuration it should hold that $\sum_{\forall r \in R} \rho' \leq 1$. The bi-rate guarantee applies to all requestors where $\rho^*_r > \rho'_r$. For $\rho^*_r = \rho'_r$, it converges to the latency-rate guarantee. Note that throughout this paper, we only consider requestors for which $\rho^*_r > \rho'_r$, since otherwise the higher service rate is equal to the allocated rate.

$$\rho^*_r = 1 - \sum_{\forall s \in R^+_r} \rho'_r. \tag{6}$$

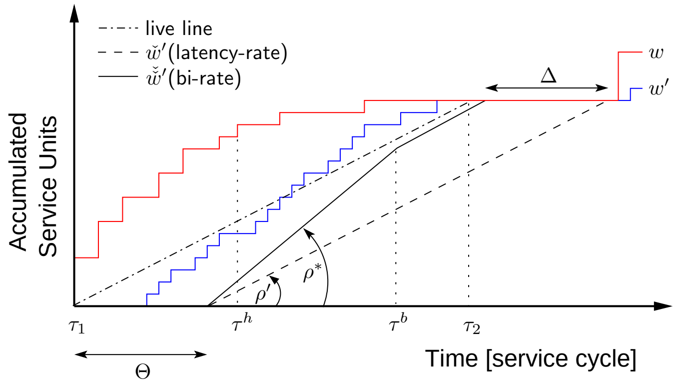


Fig. 2. Improved Worst Case Finishing Time

For a requestor $r \in R$ during the active period $[\tau_1, \tau_2]$ of the worst-case scenario, the boundary cycle, τ_r^b, is given by Equation (7). The case where the end of an active period is before this boundary cycle implies that the requestor receives service at the higher rate throughout the entire active period [5].

$$\tau_r^b = \left\lfloor \tau_1 + \frac{\sigma'_r - 1 + \rho'_r + \sum_{\forall s \in R_r^+} \sigma'_s}{\rho^*_r - \rho'_r} \right\rfloor \tag{7}$$

A bi-rate service guarantee that takes into account the higher service rate, ρ^*, improves the worst-case finishing time of requests, as illustrated in Figure 2. Δ in the figure illustrates the improvement as compared to the latency-rate service guarantee.

To guarantee a provided service at the higher service rate, ρ^*, the requestor has to ask for it. Thus, the new service guarantee applies only to requestors that have a minimum service request rate of ρ^* at least up to time point τ^h (cf. Figure 2), given in Equation (8). We refer to such requestors as *high-rate requestors*. A requestor $r \in R$ being served by a CCSP arbiter is said to be a high-rate requestor during an active period $[\tau_1, \tau_2]$ if Equation (9) holds for the requested service curve, w.

$$\tau_r^h = \tau_r^b - \Theta_r. \tag{8}$$

$$\forall t \in [\tau_1, \tau_2] : w_r(\tau_1, t) \geq \begin{cases} \rho^*_r \cdot (t - \tau_1 + 1) & t \leq \tau_r^h \\ w_r(\tau_1, \tau_r^h) + \rho'_r \cdot (t - \tau_r^h) & \text{Otherwise} \end{cases} \tag{9}$$

State-of-the-art dataflow-based system-level analysis techniques require the various aspects of the system, such as computation, storage and arbitration, to be modeled with dataflow components. We have said earlier in this section that the existing dataflow model of CCSP arbiter is too pessimistic, since it is based on the latency-rate service guarantee that does not capture the provided service at the higher service rate, ρ^*. A dataflow model of CCSP arbiter, which is based on the bi-rate service guarantee, can significantly improve utilization of resources under CCSP arbitration. Since, the the bi-rate service guarantee enables a given SoC resource under CCSP arbitration to support more requestors, or accommodate a given set of requestors with less resource. In the next section, we present a dataflow model of CCSP arbitration that is based on the bi-rate service guarantee.

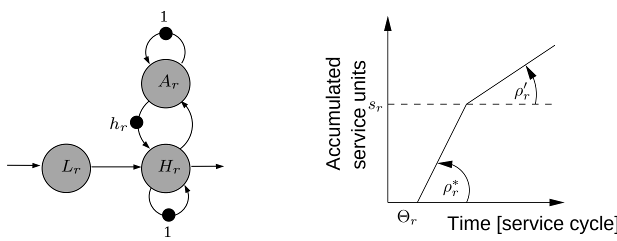
II. Dataflow Model

In this section, we present a dataflow model of CCSP arbitration, based on the bi-rate service guarantee given in Equation (5). Dataflow graphs (DFG) enable the modeling and analysis of real-time applications. These models have effective analysis techniques to compute throughput and storage requirements. Besides, they also capture cyclic data dependencies, which exist in many real-life applications. Because of these reasons, DFGs can play an indispensable role in today's MPSoC design flow that carry out application mapping and resource scheduling [6], [7].

A DFG consists of nodes, called *actors*, which communicate *tokens* through their *ports*. Tokens are sent from one actor to another over edges, called *channels*, which connect the ports of the sending and receiving actors. A variety of DFGs exist today, which vary with their level of analyzablity, expressiveness and implementation efficiency. Homogeneous Synchronous Dataflow (HSDF) graph is one of them. In a *self-timed* execution of HSDF graphs, when there exists at least one token on every input port of an actor, the actor *fires*. At the end of the firing, the actor produces one token on each of its output ports. The execution duration of a single firing of an actor X is denoted $\chi(X) \in \mathbb{R}^+$.

We present a HSDF model, which is shown in Figure 3, that models the service provision of CCSP arbitration, based on the bi-rate service guarantee. A number next to a black dot over a channel represents the number of *initial tokens* that should be available over the channel at the beginning of the execution of the graph to prevent deadlock. The number of initial tokens h_r for a requestor $r \in R$ is determined by its allocated service and the allocated services of higher-priority requestors, R_r^+.

The dataflow model consists of three actors: *latency actor (L_r)*, *higher-rate actor (H_r)* and *allocated-rate actor (A_r)*.



(a) CCSP dataflow model (b) Rate of provided service

Fig. 3. Dataflow model for CCSP arbitration

The execution durations of the three actors for a requestor $r \in R$ are given in Equations (10), (11) and (12). The case where $h_r = 1$ is unique from all other cases ($h_r > 1$), since actors H_r and A_r cannot be executed in parallel. This is because there can exist a maximum of only one token at a time over their cyclic dependency. As a result, the execution duration of actor A_r, $\chi(A_r)$, for $h_r = 1$ is different from the other cases.

$$\chi(L_r) = \Theta \tag{10}$$

$$\chi(H_r) = \frac{1}{\rho_r^*} \tag{11}$$

$$\chi(A_r) = \begin{cases} \frac{1}{\rho'_r} & \text{if } h_r > 1 \\ \frac{1}{\rho'_r} - \frac{1}{\rho_r^*} & \text{if } h_r = 1 \end{cases} \tag{12}$$

The rate of completion of service units by the dataflow model, which is shown in Figure 3(b), is equivalent to the bi-rate service guarantee, given in Equation (5). In this section, we formally prove this equivalence. Before that, we give an intuition to this equivalence by explaining the execution of the dataflow model.In this section we drop the subscript r in the notations from now on, since we will be referring to the same requestor throughout the section.

Arriving requests at the input of the CCSP arbiter can be multiple service units long, i.e. for the k^{th} request $s(\omega^k) \geq 1$. Each service unit is represented by a single token in our dataflow model. The arrival of a request at the input of the arbiter is analogous to the arrival of multiple tokens at the input of actor L. This means the arrival time of the j^{th} token marks the arrival time of the j^{th} service unit at the input of the arbiter. $\mathcal{E}(j)$ denotes the arrival time of the j^{th} token at actor L. Actor L produces a single token for every arriving token after an execution time of Θ service cycles, which models the maximum waiting time every service unit encounters. The production time of the j^{th} token by the latency actor equals $\mathcal{E}(j)+\Theta$.

In HSDF graphs, a *self-edge* refers to a special type of channel that has the same source and destination actor. The number of initial tokens on self-edges determines the *auto-concurrency* of the actor, which is the number of possible simultaneous firings of the actor. It is important to note that actor L does not have a self-edge. This implies that multiple tokens can wait in parallel, which accurately reflects what actually happens at the input of the arbiter. On the other hand, actors H and A have self-edges, each with a single initial token that implies only one service unit can be served at a time.

Output tokens produced by actor H represent completed service units. This means the production time of the j^{th} token by actor H marks the completion of the j^{th} service unit. $\mathcal{F}(j)$ denotes the finishing (production) time of the j^{th} token by actor H. The finishing time of the j^{th} token by actor H is determined by the arrival time of tokens from its three input channels. These are: (1) the arrival of the j^{th} token from the output of actor L, $\mathcal{E}(j)+\Theta$, (2) the finishing time of the previous firing, $\mathcal{F}(j-1)$, which models that only one service unit can be served at a time and (3) the production time of the $(j-h)^{th}$ token from actor A. The reasoning behind this is explained as follows. There are already h initial tokens on this channel. For every served service unit, one token is consumed from this channel by actor H. After all these h initial tokens are consumed, there may be still accumulated tokens available on this channel that are produced by actor A, during the consumption of the h initial tokens by actor H. As long as there are accumulated tokens available on this channel, service units can be served at the higher rate. This is because, in order to fire, actor H does not have to wait for the production of tokens by actor A. However, after all tokens are consumed, the rate of production of tokens by actor H is determined by the rate of production of tokens by the actor A, which is guaranteed to be lower. That means service units are served at the allocated rate of the requestor ρ'. When there are no more service units to be served, it marks the end of an active period. As we later show in Section II-D, the h initial tokens are completely restored before the beginning of a new active period, since this operation of the dataflow model is shown to be valid for sequences of active periods. Note that for $h=1$, the rate of token production by actor H is the sum of the execution times of both actor H and A. This is because the two actors cannot be fired simultaneously, since there can be a maximum of one token over the cyclic dependency at a time. Hence, $h=1$ is considered a special case and its proof is also given separately, as we will show later.

Bringing the above three limiting factors together, the finishing time of service units can be expressed with a max-plus equation [8] that bounds the completion time of service units as shown in Equation (13). $\mathcal{G}(j)$ denotes the production time of the j^{th} token at the output of actor A.

$$\mathcal{F}(j)=\begin{cases} max(\mathcal{E}(j)+\Theta,\mathcal{F}(j-1),\mathcal{G}(j-h))+\frac{1}{\rho^*} & j>0 \\ 0 & \text{Otherwise} \end{cases} \tag{13}$$

In the remaining part of this section, the formal proof of the equivalence between the bi-rate service guarantee and the dataflow model is given. The strategy of the proof is to first derive a relation that bounds the finishing time of the j^{th} service unit based on the bi-rate service guarantee. Then, we prove the equivalence of this bound on finishing time and Equation (13) for both execution of service units within an active period as well as for sequences of active periods. However, to bound the production of tokens according to Equation (13), first we need to know the number of service units that can be served at the higher rate to derive h as well as a bound on the arrival time of service units ($\mathcal{E}(j)$). Hence, the proof is organized as follows. In Section II-A, the number of service units that can be served at the higher rate is derived. Relations that bound the arrival time of the j^{th} service unit and the finishing time of the j^{th} service unit are derived in Section II-B and Section II-C, respectively. Finally, Section II-D presents the proof that shows the equivalence between the bound on the finishing time of service units and Equation (13). Table I presents the notation, for a requestor $r\in R$, that is going to be used throughout the proof.

TABLE I
THE NOTATION USED IN THE FORMAL PROOF

Notation	Description
$[\tau_1, \tau_2]$	Duration of an active period
k	The first service unit arrived during the active period
j	Any service unit arrived during the active period, where $j \geq k$
n	The last service unit that is completed before the boundary cycle, τ^b
$W(\tau_1, t)$	The number of service units arrived in the time interval $[\tau_1, t]$
$\breve{W}(\tau_1, t)$	A minimum bound on $W(\tau_1, t)$
$W'(\tau_1, t)$	The number of service units completed in the time interval $[\tau_1, t]$
$\breve{W}'(\tau_1, t)$	A minimum bound on $W'(\tau_1, t)$

A. Number of Service Units Served at the Higher Rate

In this section, we derive the number of service units that can be served at the higher service rate during the active period of the worst-case scenario. In terms of k and n, the number of service units that can be served at the higher rate, which is denoted as s_h, is given by Equation (14).

$$s_h = n - k + 1 \tag{14}$$

For a requestor $r \in R$, s_h is determined by the allocated service of r and the allocated services of higher-priority requestors, R_r^+. The derivation is given as follows.

The number of service units that can be served in an active period $[\tau_1, \tau_2]$ is bounded by the service guarantee given in Equation (5). Since the number of completed service units is an integer, we take the largest integer value, as given in Equation (15).

$$\breve{W}^{'}(\tau_1 - t) = max(0, \lfloor min(\rho^* \cdot (t - \tau_1 + 1 - \Theta), \rho' \cdot (t - \tau_1 + 1 - \Gamma)) \rfloor) \tag{15}$$

Since we have two different rates in our service guarantee, part of the service units in the active period are served at a higher rate and the rest at the allocated rate. From Equation (15), we can compute how many of these service units can be served at the higher rate, ρ^*. This is shown in Lemma 1.

Lemma 1: The number of service units that can be served at the higher service rate is given as:

$$s_h = n - k + 1 = \left\lfloor \frac{\Theta - \Gamma}{\frac{1}{\rho'} - \frac{1}{\rho^*}} \right\rfloor. \tag{16}$$

Proof: The number of finishes in the interval $[\tau_1, \tau^b]$ is equal to $\lfloor \rho^* \cdot (\tau^b - \tau_1 + 1 - \Theta) \rfloor$, according to Equation (15). By equating this with Equation (14), we get the value of s_h, as shown below.

$$
\begin{aligned}
s_h &= n-k+1 = \lfloor \rho^* \cdot (\tau^b - \tau_1 + 1 - \Theta) \rfloor \\
&= \left\lfloor \rho^* \cdot \left(\frac{\sigma' - 1 + \rho' + \sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^* - \rho'} + 1 - \Theta \right) \right\rfloor \text{(substitution by Equation (7))} \\
&= \left\lfloor \rho^* \cdot \left(\frac{\sigma' - 1 + \rho^* + \sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^* - \rho'} - \frac{\sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^*} \right) \right\rfloor \text{(substitution by Equation (1) and algebraic reduction)} \\
&= \left\lfloor \frac{\rho^* \cdot \rho^* + \rho' \cdot \sum_{\forall s \in R_r^+} \sigma_s^{'} + \rho^* \cdot (\sigma' - 1)}{\rho^* - \rho'} \right\rfloor \text{(by algebraic operations)} \\
&= \left\lfloor \frac{\Theta - \Gamma}{\frac{1}{\rho'} - \frac{1}{\rho^*}} \right\rfloor \text{(by algebraic reduction - divide both the numberator and denominator by } \rho_r^* \cdot \rho_r'\text{)}
\end{aligned}
$$

This shows that in the worst-case active period, the number of service unit a requestor can be served at the higher rate is determined entirely by its allocated service (allocated rate and burstiness) and the allocated services of its higher priority requestors. □

In the previous lemma, we have derived the number of service units that can be served at the higher rate, s_h. Another key information we need for the proofs of subsequent sections is a bound on the arrival time of these s_h service units. The arrival time of the first service unit, $\mathcal{E}(k)$, is τ_1, the start of the active period. Lemma 2 shows that the n^{th} service unit arrives before $= \tau^b - \Theta$. That means the first s_h service units in the active period, which are served at the higher service rate, arrive at the latest at $\tau^h = \tau^b - \Theta$ (cf. Figure 2).

Lemma 2: The arrival time of the n^{th} service unit, $\mathcal{E}(n)$, is less than or equal to $\tau^b - \Theta$.

$$\mathcal{E}(n) \leq \tau^b - \Theta. \tag{17}$$

Proof: For the arrival time of the n^{th} service unit to be in the interval $[\tau_1, (\tau^b - \Theta)]$, we need to show that the total number of arrivals in the interval $[\tau_1, \tau^b - \Theta]$ i.e. $w(\tau_1, \tau^b - \Theta)$ is at least equal to $n - k + 1$. We know from Equation (9) that for a high-rate requestor, it holds that $\forall t \in [\tau_1, (\tau^b - \Theta)], w(\tau_1, t) \geq \rho^* \cdot (t - \tau_1 + 1)$. Hence,

$$
\begin{aligned}
w(\tau_1, \tau^b - \Theta) &\geq \rho^* \cdot (\tau^b - \Theta - \tau_1 + 1) \\
&= \left(\frac{\sigma' - 1 + \rho' + \sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^* - \rho'} + 1 - \Theta \right) \cdot \rho^* \text{ (by substituting Equation (7))} \\
&= \left(\frac{\sigma' - 1 + \rho^* + \sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^* - \rho'} - \frac{\sum_{\forall s \in R_r^+} \sigma_s^{'}}{\rho^*} \right) \cdot \rho^* \text{ (by substituting Equation (1))} \\
&= \left(\frac{\Theta - \Gamma}{\frac{1}{\rho'} - \frac{1}{\rho^*}} \right) \\
&\geq n - k + 1 \text{ (from Lemma (1))}
\end{aligned}
$$

This shows us that at least $n-k+1$ service units arrive at $\tau^h = \tau^b - \Theta$, which proves Lemma 2. □

Since we now know how many service units are served at the higher rate, we can also derive the number of initial tokens h for our dataflow model, illustrated in Figure 3(a). The derivation for h is given in Lemma 3.

Lemma 3: To guarantee that s_h number of service units are served at the higher service rate by the dataflow model, illustrated in Figure 3(a), the number of initial tokens h that should be available on the channel from actor A to actor H is given by Equation (18).

$$h = s_h - \frac{(n-k-1)\cdot\rho'}{\rho^*} \tag{18}$$

Proof: Let t_{s_h} denotes the total time interval the dataflow model takes to serve those s_h service units at the higher rate. This gives us Equation (19).

$$t_{s_h} = s_h \cdot \frac{1}{\rho^*} \tag{19}$$

h' denotes tokens produced by actor A during the time interval t_{s_h}. This means h' equals the number of firings of actor A during time interval t_{s_h}. Taking into account the number of initial tokens h, this gives us Equation (20).

$$s_h = n-k+1 = h + h' \tag{20}$$

Let t_{str} denotes the starting time of the first firing of actor A and t_{end} denotes the time the last firing of actor A should end to the latest. Actor A can start firing only after the first service unit is served i.e. $\frac{1}{\rho^*}$ service units after the start of t_{s_h}. In addition, the last firing of actor A should be completed at least $\frac{1}{\rho^*}$ service units before the end of t_{s_h} so that the n^{th} service unit can be served before the end of t_{s_h}. Then, h' can be computed as follows.

$$\begin{aligned} h' &= \frac{(t_{end}-t_{str})}{\frac{1}{\rho'}} \\ &= ((t_{s_h} - \frac{1}{\rho^*}) - \frac{1}{\rho^*})\cdot\rho' \\ &= (\frac{s_h}{\rho^*} - \frac{2}{\rho^*})\cdot\rho' \text{ (according to Equation (19))} \\ &= \frac{(s_h-2)\cdot\rho'}{\rho^*} \\ &= \frac{(n-k-1)\cdot\rho'}{\rho^*} \text{ (according to Equation (20))} \end{aligned} \tag{21}$$

Putting back h' into Equation (20), we get h, the number of initial tokens, as given in Equation (18). □

From this section, we have now obtained relations for the number of service units that can be served at the higher rate, s_h, the number of initial tokens, h, as well as a bound for arrival time of the n^{th} service unit. In the next section, we derive a bound on the arrival time of each service unit in the active period.

B. Bound on Arrival Time of Service Units

In this section, we aim to derive a bound on the arrival time of each service unit. The basis of this derivation is Equation (9), which defines a lower bound on the requested service curve, w. We use it to bound the number of service units arrived since the start of the active period, as given in Equation (22). We use Equation (22) in Lemma 4 to derive the bound on the arrival time of each service unit.

$$\forall t \in [\tau_1, \tau_2],\, W(\tau_1, t) \geq \breve{W}(\tau_1, t):$$
$$\breve{W}(\tau_1, t) = \begin{cases} \rho^* \cdot (t - \tau_1 + 1) & t \leq \tau^h \\ \breve{W}(\tau_1, \tau^h) + \rho' \cdot (t - \tau^h + 1) & \text{Otherwise} \end{cases} \tag{22}$$

Lemma 4: Let $\mathcal{E}(k) = \tau_1$ and $\mathcal{E}(j) \leq \tau_2$. The bound ϕ_j on the arrival time of the j^{th} service unit $\mathcal{E}(j)$, given in Equation (23) is equivalent to bounding the number of arriving service units as in Equation (22) during every active period.

$$\mathcal{E}(j) \leq \phi_j = \begin{cases} \tau_1 + \frac{j-k}{\rho^*} & j \leq n \\ \phi_n + \frac{j-n}{\rho'} & j > n \end{cases} \tag{23}$$

Proof: Let us split the proof into two cases where $j \leq n$ and $j > n$.

- **Case 1**: $j \leq n$. From Lemma 2, we can tell that $j \leq n$ implies $\mathcal{E}(j) \leq \tau^b - \Theta$ and, hence, the arrival rate is determined by the relation $\breve{W}(\tau_1, t) = \rho^* \cdot (t - \tau_1 + 1)$, according to Equation (22). Since ϕ_j is a bound on the arrival time of the j^{th} service unit, it should hold that starting from ϕ_j the number of arrived service units is at least $j - k$: i.e. $\forall t' \geq \phi_j : \breve{W}(\tau_1, t' - 1) \geq j - k$. Expanding $\breve{W}(\tau_1, t' - 1)$ in this relation gives $\rho^* \cdot (t' - \tau_1) \geq j - k$. Rearranging gives:

$$t' \geq \tau_1 + \frac{j-k}{\rho^*} \tag{24}$$

From Equation (24), we get the bound ϕ_j on the arrival time of the j^{th} service unit as:

$$\phi_j = \tau_1 + \frac{j-k}{\rho^*}. \tag{25}$$

- **Case 2:** $j > n$. We use the same proof strategy as in the first case. From Lemma 2, since $j > n$, the arrival time is determined by the relation $\breve{W}(\tau_1, t) = \breve{W}(\tau_1, \tau^b - \Theta) + \rho' \cdot (t - \tau^b + \Theta)$, according to Equation (23). Since ϕ_j is a bound on the arrival time of the j^{th} service unit, it should hold that starting from ϕ_j the number of arrived service units is at least $j - k$: i.e. $\forall t' \geq \phi_j : \breve{W}(\tau_1, t' - 1) \geq j - k$. After expanding $\breve{W}(\tau_1, t' - 1)$ in this relation, we get the relation $\rho^* \cdot (\tau^b - \Theta - \tau_1 + 1) + \rho' \cdot (t' - \tau^b + \Theta) \geq j - k$, which after rearranging gives:

$$t' \geq \tau_1 + \frac{n-k}{\rho^*} + \frac{j-n}{\rho'}. \tag{26}$$

From Equation (26), we get the bound ϕ_j on the arrival time of the j^{th} service unit as:

$$\phi_j = \tau_1 + \frac{n-k}{\rho^*} + \frac{j-n}{\rho'}. \tag{27}$$

□

Now, we have a relation that bounds the arrival time of individual service units. In the next section, we follow a similar strategy to bound finishing time of individual service units. The bounds on the arrival and finishing time of service units are later used in Section II-D to prove the equivalence between Equation (13) and the bi-rate service guarantee.

C. Bound on Finishing Time of Service Units

In this section, we bound the finishing time of individual service units, as shown in Lemma 5. The basis of this lemma is Equation (15), which bounds the minimum number of service units completed since the beginning of the active period.

Lemma 5: Let $\mathcal{E}(k) = \tau_1$ and $\mathcal{E}(j) \leq \tau_2$. The bound ψ_j on the finishing time of the j^{th} service unit as provided by Equation (28) is conservative with respect to Equation (15).

$$\mathcal{F}(j) \le \psi_j = \begin{cases} \mathcal{E}(k) + \Theta + \frac{j-k+1}{\rho^*} & j \le n \\ \psi_n + \frac{j-n}{\rho'} & j > n \end{cases} \tag{28}$$

Proof: Let us split the proof into two cases where $j \le n$ and $j > n$.

- **Case 1**: $j \le n$. Since $j \le n$, $\psi_j \le \tau^b$. Hence, the minimum service guarantee is determined by the relation $\breve{W}'(\tau_1, t) = \lfloor \rho^* \cdot (t - \tau_1 + 1 - \Theta) \rfloor$. Since ψ_j is a bound on the finishing time of the j^{th} service unit, at least $j - k + 1$ service units should be completed from ψ_j on: i.e. $\forall t' \ge \psi_j : \breve{W}'(\tau_1, t' - 1) \ge j - k + 1$. After expanding $\breve{W}'(\tau_1, t' - 1)$ in this relation, we get $\rho^* \cdot (t' - \tau_1 - \Theta) \ge j - k + 1$. Rearranging gives:

$$t' \ge \tau_1 + \Theta + \frac{j-k+1}{\rho^*}. \tag{29}$$

From Equation (29), we get ψ_j, which is a bound on the finishing time of the j^{th} service unit as:

$$\psi_j = \tau_1 + \Theta + \frac{j-k+1}{\rho^*}. \tag{30}$$

- **Case 2**: $j > n$. We use the same proof strategy as in the first case. $j-k+1$ service units should be completed from ψ_j on: i.e. $\forall t' \ge \psi_j : \breve{W}'(\tau_1, t' - 1) \ge j - k + 1$. Expanding the above step, we get $\rho^* \cdot (\tau^b - \Theta - \tau_1 + 1) + \rho' \cdot (t' - \tau^b - \Gamma) \ge j - k + 1$. Substitution and algebraic reduction gives:

$$t' \ge \tau_1 + \Theta + \frac{n-k+1}{\rho^*} + \frac{j-n}{\rho'} \tag{31}$$

From Equation (31), we get ψ_j, which is a bound on the finishing time of the j^{th} service unit as:

$$\psi_j = \tau_1 + \Theta + \frac{n-k+1}{\rho^*} + \frac{j-n}{\rho'}. \tag{32}$$

□

We now have relations from Section II-B and II-C that bound the arrival and finishing time of each service unit in the active period, respectively. We use these bounds in the next section to prove the equivalence between the bi-rate guarantee and the dataflow model, illustrated in Figure 3(a).

D. Proof of the Dataflow Model

In this section, we prove the equivalence between Equation (28) that bounds the finishing time of each service unit and Equation (13) that determines the production time of tokens by actor H in the dataflow model, illustrated in Figure 3(a). For the sake of simplicity, the main proof focuses on the case where $h > 1$. This is because $h = 1$ is a trivial case that can be proved with simpler steps as follows. $h = 1$ implies the requestor receives service at the higher rate only for the first service unit in the active period. Since we have one initial token on the edge from actor A to actor H, the first token is produced by actor H at $\mathcal{E}(k) + \Theta + \frac{1}{\rho^*}$. However, for all subsequent tokens, actor H has to wait the production of tokens by actor A before it can fire. That means, the rate of token production by actor H is the sum of the rate of the two actors, $\chi(A) + \chi(H)$, which equals $\frac{1}{\rho'}$. Therefore, all service units in the active period, except the first one, are served at the allocated rate. This proves that the dataflow model, illustrated in Figure 3(a), is an accurate model for the bi-rate service guarantee for $h = 1$.

In the remaining part of this section, we carry out the proof for $h > 1$ in two parts. First, in Section II-D1, we derive a relation for the production time of tokens by actor A. Then, we use this relation in Section II-D2 to finally prove the equivalence of Equation (28) and Equation (13).

1) Bounding Token Production by the Allocated-rate Actor: The production of tokens by actor A is determined by the production of tokens by actor H and the finishing time of its previous firing, as shown in Figure 3(a). $\mathcal{G}(j)$ denotes the production time of the j^{th} token by the allocated-rate actor. Equation (33) defines it based on the finishing time of its previous firing, $\mathcal{G}(j-1)$, and the production time of the j^{th} token by actor H, $\mathcal{F}(j)$.

$$\mathcal{G}(j) = \begin{cases} max(\mathcal{F}(j), \mathcal{G}(j-1)) + \frac{1}{\rho'} & j > 0 \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

We need a conservative bound on the production time of tokens by actor A to bound the production time of tokens by actor H, which corresponds to the bound on the completion time of service units. Lemma 6 presents such a conservative bound on the production time of tokens by actor A.

Lemma 6: Function $g(j)$, shown in Equation (34), is a bound to the production time of tokens by actor A, $\mathcal{G}(j)$, shown in Equation (33), during any active period.

$$g(j) = \begin{cases} \psi_k + \frac{j-k+1}{\rho'} & j > 0 \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

Proof: We prove this lemma by induction in two parts. In the first part, we use induction to prove for service units within the first active period. In the second part, we again use induction to prove for sequences of active periods, but using the result of the first part as a base step.

Since $g(j)$ is a bound on $\mathcal{G}(j)$, we should be able to get an equivalent relation to $g(j)$ if we bound $\mathcal{F}(j)$ in Equation (33). In this proof, we show this equivalence. To bound $\mathcal{F}(j)$, we use ψ_j, given in Equation (28).

Part 1: For service units within the first active period

Base step: $k = j = 1$.

$g(1) = \psi_1 + \frac{1}{\rho'}$, according to Equation (34). $\mathcal{G}(1) = max(\psi_1, \mathcal{G}(0)) + \frac{1}{\rho'} = \psi_1 + \frac{1}{\rho'}$, according to Equation (33), proving the equivalence for the base step.

Inductive step: Assume $g(j) = \mathcal{G}(j) = \psi_k + \frac{j-k+1}{\rho'}$. We now need to show that $g(j+1) = \mathcal{G}(j+1)$. $g(j+1)$ is given by Equation (35), according to Equation (34).

$$g(j+1) = \psi_k + \frac{j+1-k+1}{\rho'} \tag{35}$$

$\mathcal{G}(j+1)$ is given as shown in Equation (36), according to Equation (33), after bounding $\mathcal{F}(j+1)$.

$$\mathcal{G}(j+1) = max(\psi_{j+1}, \mathcal{G}(j)) + \frac{1}{\rho'} \tag{36}$$

To evaluate the max term in Equation (36), we need to compare ψ_{j+1} and $\mathcal{G}(j)$. We know $\mathcal{G}(j)$ from the assumption in the inductive step. But for ψ_{j+1}, we need to consider three cases: for $j < n$, $j > n$ and also for $j = n$, since it involves equations from both the case $j \leq n$ and $j > n$.

Case 1: $j < n$. This case implies $j + 1 \leq n$. Then, according to Equation (28), ψ_{j+1} is as given in Equation (37).

$$\psi_{j+1} = \mathcal{E}(k) + \Theta + \frac{j+1-k+1}{\rho^*} \tag{37}$$

Now comparing, Equation (37) and $\mathcal{G}(j)$ shows that $\mathcal{G}(j)$ is the maximum of the two. This comparison is shown below.

$$\begin{aligned}
\frac{j-k+1}{\rho^*} &\leq \frac{j-k+1}{\rho'} \text{ (since } \rho' < \rho^*) \\
\mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{j-k+1}{\rho^*} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{j-k+1}{\rho'} \\
\mathcal{E}(k)+\Theta+\frac{j+1-k+1}{\rho^*} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{j-k+1}{\rho'} \\
\psi_{j+1} &\leq \mathcal{G}(j)
\end{aligned}$$

Case 2: $j = n$. This case implies $j+1 > n$. Then, according to Equation (28), ψ_{j+1} is as given in Equation (38).

$$\psi_{j+1} = \mathcal{E}(k)+\Theta+\frac{n-k+1}{\rho^*}+\frac{1}{\rho'} \tag{38}$$

Now comparing, Equation (38) and $\mathcal{G}(j)$ shows that $\mathcal{G}(j)$ is the maximum of the two. This comparison is shown below.

$$\begin{aligned}
\frac{n-k}{\rho^*} &\leq \frac{n-k}{\rho'} \text{ (since } \rho' < \rho^*) \\
\mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{n-k}{\rho^*}+\frac{1}{\rho'} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{n-k}{\rho'}+\frac{1}{\rho'} \\
\mathcal{E}(k)+\Theta+\frac{n-k+1}{\rho^*}+\frac{1}{\rho'} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{j-k+1}{\rho'} \\
\psi_{j+1} &\leq \mathcal{G}(j)
\end{aligned}$$

Case 3: $j > n$. This case implies $j+1 > n$. Then, according to Equation (28), ψ_{j+1} is as given in Equation (39).

$$\psi_{j+1} = \mathcal{E}(k)+\Theta+\frac{n-k+1}{\rho^*}+\frac{j+1-n}{\rho'} \tag{39}$$

Now comparing, Equation (39) and $\mathcal{G}(j)$ shows that $\mathcal{G}(j)$ is the maximum of the two. This comparison is shown below.

$$\begin{aligned}
\frac{n-k}{\rho^*} &\leq \frac{n-k}{\rho'} \text{ (since } \rho' < \rho^*) \\
\frac{n-k}{\rho^*}+\frac{1}{\rho'} &\leq \frac{n-k+1}{\rho'} \\
\frac{n-k}{\rho^*}+\frac{j-n}{\rho'}+\frac{1}{\rho'} &\leq \frac{n-k+1}{\rho'}+\frac{j-n}{\rho'} \text{ (since } j > n) \\
\mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{n-k}{\rho^*}+\frac{j-n+1}{\rho'} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{n-k+1}{\rho'}+\frac{j-n}{\rho'} \\
\mathcal{E}(k)+\Theta+\frac{n-k+1}{\rho^*}+\frac{j+1-n}{\rho'} &\leq \mathcal{E}(k)+\Theta+\frac{1}{\rho^*}+\frac{j-k+1}{\rho'} \\
\psi_{j+1} &\leq \mathcal{G}(j)
\end{aligned}$$

For all the three cases, we have showed that $\psi_{j+1} \leq \mathcal{G}(j)$. Hence, Equation (36) can be simplified as shown below.

$$\begin{aligned}\mathcal{G}(j+1) &= max(\psi_{j+1}, \mathcal{G}(j)) + \frac{1}{\rho'} \\ &= \mathcal{G}(j) + \frac{1}{\rho'} \\ &= \mathcal{E}(k) + \Theta + \frac{1}{\rho^*} + \frac{j-k+1}{\rho'} + \frac{1}{\rho'} \\ &= \psi_k + \frac{j+1-k+1}{\rho'} \\ &= g(j+1)\end{aligned}$$

This proves that $g(j)$ is a bound to $\mathcal{G}(j)$ for the first active period. Now, we need to prove if the bound holds for sequence of active periods. This is given in the second part of the proof.

Part 2: For sequence of active periods

Base step: We have showed in the first part of the proof that the bound holds for the first active period.

Inductive step: Assuming the bound holds for the l^{th} active period, A^l, we need to show if it holds for the $(l+1)^{th}$ active period, A^{l+1}. Let k^l and k^{l+1} denote the first service units in active period A^l and A^{l+1}, respectively.

The production time of the first token produced by actor A in the $(l+1)^{th}$ active period is as given in Equation (40).

$$\mathcal{G}(k^{l+1}) = max(\mathcal{F}(k^{l+1}), \mathcal{G}(k^{l+1}-1)) + \frac{1}{\rho'} \tag{40}$$

Let us now evaluate the max term in Equation (40). $\mathcal{F}(k^{l+1})$ is as given in Equation (41), according to Equation (28).

$$\mathcal{F}(k^{l+1}) = \mathcal{E}(k^{l+1}) + \Theta + \frac{1}{\rho^*} \tag{41}$$

$\mathcal{G}(k^{l+1}-1)$ is as given in Equation (42), according to Equation (34). This is because $k^{l+1}-1$ belongs to active period A^l and based on our assumption in the inductive step, the bound holds for active period A^l.

$$\begin{aligned}\mathcal{G}(k^{l+1}-1) = g(k^{l+1}-1) &= \psi_{k^l} + \frac{k^{l+1}-1-k^l+1}{\rho'} \\ &= \mathcal{E}(k^l) + \Theta + \frac{1}{\rho^*} + \frac{k^{l+1}-k^l}{\rho'}\end{aligned} \tag{42}$$

Comparing Equation (41) and (42) shows that $\mathcal{F}(k^{l+1})$ is the maximum of the two. This comparison is given below. For Equation (43), note that $k^{l+1}-k^l$ equals the total number of service units requested (also served) in active period A^l. The relation is based on the fact that the total service a requestor would get if it were continuously served at its allocated rate, i.e $\rho' \cdot (\tau_2 - \tau_1 + 1)$, is always greater than or equal to the total service it actually receives.

$$
\begin{aligned}
\rho' \cdot (\tau_2^l - \tau_1^l + 1) &\geq k^{l+1} - k^l \text{ (from concept of active periods)} \quad (43)\\
\rho' \cdot (\tau_1^{l+1} - \tau_2^l - 1) + \rho' \cdot (\tau_2^l - \tau_1^l + 1) &\geq k^{l+1} - k^l \text{ (since } \tau_1^{l+1} - \tau_2^l \geq 1)\\
\rho' \cdot (\tau_1^{l+1} - \tau_1^l) &\geq k^{l+1} - k^l\\
\rho' \cdot (\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l)) &\geq k^{l+1} - k^l\\
\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l)) &\geq \frac{k^{l+1} - k^l}{\rho'}\\
\mathcal{E}(k^{l+1}) &\geq \mathcal{E}(k^l) + \frac{k^{l+1} - k^l}{\rho'}\\
\mathcal{E}(k^{l+1}) + \Theta + \frac{1}{\rho^*} &\geq \mathcal{E}(k^l) + \Theta + \frac{1}{\rho^*} + \frac{k^{l+1} - k^l}{\rho'}\\
\mathcal{F}(k^{l+1}) &\geq \mathcal{G}(k^{l+1} - 1) \text{ (by substituting Equations (41) and (42))}
\end{aligned}
$$

Now, Equation (40) can be simplified as shown below.

$$
\begin{aligned}
\mathcal{G}(k^{l+1}) &= max(\mathcal{F}(k^{l+1}), \mathcal{G}(k^{l+1} - 1)) + \frac{1}{\rho'}\\
&= \mathcal{F}(k^{l+1}) + \frac{1}{\rho'}\\
&= \mathcal{E}(k^{l+1}) + \Theta + \frac{1}{\rho^*} + \frac{1}{\rho'}\\
&= \psi_{k^{l+1}} + \frac{j - k^{l+1} + 1}{\rho'}\\
&= g(k^{l+1})
\end{aligned}
$$

This shows that the bound holds for the first service unit of active period A^{l+1}. In addition, in the inductive step of the first part of this proof, we showed that if $\mathcal{G}(j) = g(j)$, then $\mathcal{G}(j+1) = g(j+1)$. In that proof, we did not assume that service units j and $j+1$ belong to the first active period. That means, the proof also holds for any active period. Hence, with this we show that the bound holds for active period A^{l+1}.

This proves that the bound holds for all active periods, since we showed that the bound holds for the $(l+1)^{th}$ active period, assuming it holds for the l^{th} active period. □

2) Equivalence between Dataflow Model and Bi-rate Guarantee: In this section, we give the proof for the equivalence between the dataflow model, illustrated in Figure 3(a), and the bi-rate service guarantee of CCSP. We do this by showing the equivalence between Equation (28) and Equation (13). Equation (28) bounds the finishing time of each service unit according to the bi-rate service guarantee. Equation (13) is a relation that gives the production time of tokens by actor H of the dataflow model. Since the production of each token by actor H corresponds to the completion of one service unit, the equivalence of these two equations shows the equivalence of the bi-rate service guarantee and the dataflow model.

The proof is split into two theorems for the sake of readability. The theorems are proved by induction. Theorem 1 first proves that the equivalence holds for the first active period. Then, Theorem 2 proves that the equivalence also holds for sequences of active periods, making use of the proof from Theorem 1 as a base step.

Theorem 1 (Equivalence for the first active period): For the first active period, the bound on the finishing time of service unit j, ψ_j, as defined in Equation (28) is equivalent to the bound $\mathcal{F}(j)$, as defined by Equation (13).

Proof: We prove the theorem by induction.

Base step: For the first service unit, we have that

$$\begin{aligned}\mathcal{F}(1) &= max(\mathcal{E}(1)+\Theta, \mathcal{F}(0), \mathcal{G}(j-h)) + \frac{1}{\rho^*}\\ &= \mathcal{E}(1)+\Theta+\frac{1}{\rho^*} \text{ (since } \mathcal{F}(0)=\mathcal{G}(1-h)=0 \text{ according to Equations (13) and (33))}\\ &= \tau_1+\Theta+\frac{1}{\rho^*}\\ &= \psi_1\end{aligned}$$

Inductive step: Assume for any service unit j in the active period, it holds that $\mathcal{F}(j)=\psi_j$. Then, we need to prove that it also holds for the $j+1^{th}$ active period: i.e. $\mathcal{F}(j+1)=\psi_{j+1}$. According to Equation (13), $\mathcal{F}(j+1)$ is given as given in Equation (44).

$$\mathcal{F}(j+1) = max(\mathcal{E}(j+1)+\Theta, \mathcal{F}(j), \mathcal{G}(j+1-h)) + \frac{1}{\rho^*} \tag{44}$$

To evaluate the max term in Equation (44), we need to consider three cases. The cases are for $j<n$, $j=n$ and $j>n$.

Case 1: $j<n$. This case implies that $j+1\leq n$. Let us expand all the three terms in Equation (44). The arrival time of the $(j+1)^{th}$ service unit is as given in Equation (45), according to Equation (23).

$$\mathcal{E}(j+1)+\Theta = \tau_1+\Theta+\frac{j+1-k}{\rho^*} \tag{45}$$

We know from the assumption of the inductive step that $\mathcal{F}(j)$ holds. It is as given in Equation (46), according to Equation (28).

$$\mathcal{F}(j) = \tau_1+\Theta+\frac{j-k+1}{\rho^*} \tag{46}$$

The third term $\mathcal{G}(j+1-h)$ is as given in Equation (47), according to Equation (33).

$$\mathcal{G}(j+1-h) = g(j+1-h) = \psi_k+\frac{j+1-h-k+1}{\rho'} \tag{47}$$

Now let us compare the above three equations to compute the max term in Equation (44). It can be seen that Equation (45) and Equation (46) are equal. Comparing them with Equation (47) reveals that Equation (47) is the lesser. The comparison is shown below here.

$$\begin{aligned}\frac{j-k}{\rho^*} &= \frac{j-n+1}{\rho^*}+\frac{n-k-1}{\rho^*}\\ \frac{j-k}{\rho^*} &\geq \frac{j-n+1}{\rho'}+\frac{n-k-1}{\rho^*} \text{ (since } j+1-n\leq 0 \text{ and } \rho'<\rho^*)\\ \frac{j-k}{\rho^*} &\geq \frac{j-n+1}{\rho'}+\frac{h'}{\rho'} \text{ (according to Equation (21))}\\ \frac{j-k}{\rho^*} &\geq \frac{j-n+h'+1}{\rho'}\\ \frac{j-k}{\rho^*} &\geq \frac{j+1-h-k+1}{\rho'} \text{ (according to Equation (20))}\\ \tau_1+\Theta+\frac{1}{\rho^*}+\frac{j-k}{\rho^*} &\geq \tau_1+\Theta+\frac{1}{\rho^*}+\frac{j+1-h-k+1}{\rho'}\\ \tau_1+\Theta+\frac{1}{\rho^*}+\frac{j-k}{\rho^*} &\geq \psi_k+\frac{j+1-h-k+1}{\rho'}\\ \mathcal{F}(j) &\geq \mathcal{G}(j+1-h) \text{ (according to Equations (46) and (47))}\end{aligned}$$

Now, evaluating Equation (44) shows that $\mathcal{F}(j+1) = \psi_{j+1}$, as shown below here.

$$\begin{aligned}\mathcal{F}(j+1) &= max(\mathcal{E}(j+1)+\Theta, \mathcal{F}(j), \mathcal{G}(j+1-h)) + \frac{1}{\rho^*} \\ &= \mathcal{F}(j) + \frac{1}{\rho^*} \\ &= \tau_1 + \Theta + \frac{j+1-k+1}{\rho^*} \\ &= \psi_{j+1}\end{aligned}$$

Case 2: $j = n$. This case implies that $j+1 > n$. Let us expand all the three terms in Equation (44). The arrival time of the $(j+1)^{th}$ service unit is as given in Equation (48), according to Equation (23).

$$\mathcal{E}(j+1) + \Theta = \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{1}{\rho'} \tag{48}$$

We know from the assumption of the inductive step that $\mathcal{F}(j)$ holds. It is as given in Equation (46). The third term $\mathcal{G}(j+1-h)$ is as given in Equation (47). Now let us compare Equations (48), (46) and (47) to compute the max term in Equation (44). Equation (48) is greater than Equation (46), as shown below here.

$$\begin{aligned}\frac{1}{\rho'} &\geq \frac{1}{\rho^*} \\ \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{1}{\rho'} &\geq \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{1}{\rho^*} \\ \mathcal{E}(j+1) + \Theta &\geq \mathcal{F}(j)\end{aligned}$$

Next, comparing Equation (48) and Equation (47) reveals that they are equal. The comparison is shown below here.

$$\begin{aligned}\frac{n-k}{\rho^*} + \frac{1}{\rho'} &= \frac{1}{\rho^*} + \frac{n-k-1}{\rho^*} + \frac{j-n+1}{\rho'} \text{ (since } j = n\text{)} \\ \frac{n-k}{\rho^*} + \frac{1}{\rho'} &= \frac{1}{\rho^*} + \frac{h'}{\rho'} + \frac{j-n+1}{\rho'} \text{ (according to Equation (21))} \\ \frac{n-k}{\rho^*} + \frac{1}{\rho'} &= \frac{1}{\rho^*} + \frac{j-n+h'+1}{\rho'} \\ \frac{n-k}{\rho^*} + \frac{1}{\rho'} &= \frac{1}{\rho^*} + \frac{j+1-h-k+1}{\rho'} \text{ (according to Equation (20))} \\ \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{1}{\rho'} &= \tau_1 + \Theta + \frac{1}{\rho^*} + \frac{j+1-h-k+1}{\rho'} \text{ (according to Equation (20))} \\ \mathcal{E}(j+1) + \Theta &= \mathcal{G}(j+1-h) \text{ (according to Equations (48) and (47))}\end{aligned}$$

Now, evaluating Equation (44) shows that $\mathcal{F}(j+1) = \psi_{j+1}$, as shown below here.

$$\begin{aligned}
\mathcal{F}(j+1) &= max(\mathcal{E}(j+1)+\Theta, \mathcal{F}(j), \mathcal{G}(j+1-h)) + \frac{1}{\rho^*} \\
&= \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{1}{\rho'} + \frac{1}{\rho^*} \\
&= \tau_1 + \Theta + \frac{n-k+1}{\rho^*} + \frac{1}{\rho'} \\
&= \tau_1 + \Theta + \frac{n-k+1}{\rho^*} + \frac{j+1-n}{\rho'} \text{ (since } j=n) \\
&= \psi_{j+1}
\end{aligned}$$

Case 3: $j > n$. This case implies that $j+1 > n$. Let us expand all the three terms in Equation (44). The arrival time of the $(j+1)^{th}$ service unit is as given in Equation (49), according to Equation (23).

$$\mathcal{E}(j+1) + \Theta = \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{j+1-n}{\rho'} \tag{49}$$

We know from the assumption of the inductive step that $\mathcal{F}(j)$ holds. It is as given in Equation (50), according to Equation (28).

$$\mathcal{F}(j) = \tau_1 + \Theta + \frac{n-k+1}{\rho^*} + \frac{j-n}{\rho'} \tag{50}$$

The third term $\mathcal{G}(j+1-h)$ is as given in Equation (47). Now let us compare Equations (49), (50) and (47) to compute the max term in Equation (44). Equation (49) is greater than Equation (50), as shown below here.

$$\begin{aligned}
\frac{1}{\rho'} &\geq \frac{1}{\rho^*} \\
\tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{j-n}{\rho'} + \frac{1}{\rho'} &\geq \tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{j-n}{\rho'} + \frac{1}{\rho^*} \\
\mathcal{E}(j+1) + \Theta &\geq \mathcal{F}(j)
\end{aligned}$$

Next, comparing Equation (49) with Equation (47) reveals that they are equal. The comparison is shown below here.

$$\begin{aligned}
\frac{n-k}{\rho^*} + \frac{j-n+1}{\rho'} &= \frac{1}{\rho^*} + \frac{n-k-1}{\rho^*} + \frac{j-n+1}{\rho'} \\
\frac{n-k}{\rho^*} + \frac{j-n+1}{\rho'} &= \frac{1}{\rho^*} + \frac{h'}{\rho'} + \frac{j-n+1}{\rho'} \text{ (according to Equation (21))} \\
\frac{n-k}{\rho^*} + \frac{j-n+1}{\rho'} &= \frac{1}{\rho^*} + \frac{j-n+h'+1}{\rho'} \\
\frac{n-k}{\rho^*} + \frac{j-n+1}{\rho'} &= \frac{1}{\rho^*} + \frac{j+1-h-k+1}{\rho'} \text{ (according to Equation (20))} \\
\tau_1 + \Theta + \frac{n-k}{\rho^*} + \frac{j-n+1}{\rho'} &= \tau_1 + \Theta + \frac{1}{\rho^*} + \frac{j+1-h-k+1}{\rho'} \text{ (according to Equation (20))} \\
\mathcal{E}(j+1) + \Theta &= \mathcal{G}(j+1-h) \text{ (according to Equations (49) and (47))}
\end{aligned}$$

Now, evaluating Equation (44) shows that $\mathcal{F}(j+1) = \psi_{j+1}$, as shown below here.

$$\begin{aligned}\mathcal{F}(j+1) &= max(\mathcal{E}(j+1)+\Theta, \mathcal{F}(j), \mathcal{G}(j+1-h)) + \frac{1}{\rho^*} \\ &= \mathcal{E}(j+1)+\Theta+\frac{1}{\rho^*} \\ &= \tau_1+\Theta+\frac{n-k+1}{\rho^*}+\frac{j-n+1}{\rho'} \\ &= \psi_{j+1}\end{aligned}$$

We are now able to show for all the three cases that $\mathcal{F}(j+1) = \psi_{j+1}$ assuming $\mathcal{F}(j) = \psi_j$. Hence, by induction, this proves that the equivalence between $\mathcal{F}(j)$ and ψ_j holds for the first active period. □

Theorem 1 now forms the base step for the proof by induction of Theorem 2 that shows that this equivalence also holds for all active periods.

Theorem 2 (Equivalence for sequences of active periods): For any active period, the bound on the finishing time of service unit j, ψ_j, as defined in Equation (28) is equivalent to the bound $\mathcal{F}(j)$, as defined by Equation 13.

Proof: We prove this theorem by induction.

Base step: We have proved the equivalence holds for the first active period in Theorem 1.

Inductive step Let us assume the equivalence holds for the l^{th} active period, A_l. Let k^{l+1} is the first service unit in the $(l+1)^{th}$ active period, A^{l+1}, and k^l the first service unit in active period A^l. In addition, τ_1^l denotes the beginning of the l^{th} active period, and n^l and n^{l+1} denote the last service units served at the higher rate in the l^{th} and $(l+1)^{th}$ active periods, respectively.

Now, we show that the equivalence holds for active period A^{l+1}, by induction.

Base step: for active period A^{l+1} - For the first service unit of active period A^{l+1}, Equation (51) holds, according to Equation (13).

$$\mathcal{F}(k^{l+1}) = max(\mathcal{E}(k^{l+1})+\Theta, \mathcal{F}(k^{l+1}-1), \mathcal{G}(k^{l+1}-h)) + \frac{1}{\rho^*} \tag{51}$$

Now, let us compare the three terms in Equation (51) to evaluate the max term. Comparing $\mathcal{E}(k^{l+1})+\Theta$ and $\mathcal{F}(k^{l+1}-1)$ shows that $\mathcal{E}(k^{l+1})+\Theta > \mathcal{F}(k^{l+1}-1)$. This comparison is presented as follows.

Since k^{l+1} is the first service unit of active period A^{l+1}, $k^{l+1}-1$ belongs to active period A^l and Equation (52) holds, according to Equation (23) .

$$\mathcal{E}(k^{l+1}-1) \le \phi_{k^{l+1}-1} = \begin{cases} \tau_1^l + \frac{(k^{l+1}-1)-k^l}{\rho^*} & k^{l+1}-1 < n^l \\ \tau_1^l + \frac{n^l-k^l}{\rho^*} & k^{l+1}-1 = n^l \\ \tau_1^l + \frac{n^l-k^l}{\rho^*} + \frac{(k^{l+1}-1)-n^l}{\rho'} & k^{l+1}-1 > n^l \end{cases} \tag{52}$$

Since service unit k^{l+1} does not belong to the active period A^l, we have Equation (53) where the right side of the the equation would be the bound on the arrival time of k^{l+1} if it were in active period A^l.

$$\mathcal{E}(k^{l+1}) > \begin{cases} \tau_1^l + \frac{k^{l+1}-k^l}{\rho^*} & k^{l+1}-1 < n^l \\ \tau_1^l + \frac{n^l-k^l}{\rho^*} + \frac{1}{\rho'} & k^{l+1}-1 = n^l \\ \tau_1^l + \frac{n^l-k^l}{\rho^*} + \frac{k^{l+1}-n^l}{\rho'} & k^{l+1}-1 > n^l \end{cases} \tag{53}$$

From the inductive step, we have assumed the equivalence holds for active period A_l. Hence, the bound on the finishing time of $k^{l+1}-1$ is given as shown in Equation (54), according to Equation (28).

$$\psi_{k^{l+1}-1} = \begin{cases} \tau_1^l + \Theta + \frac{k^{l+1}-k^l}{\rho^*} & k^{l+1}-1 \leq n^l \\ \tau_1^l + \Theta + \frac{n^l-k^l+1}{\rho^*} + \frac{(k^{l+1}-1)-n^l}{\rho'} & k^{l+1}-1 > n^l \end{cases} \tag{54}$$

Now evaluating $\mathcal{E}(k^{l+1}) + \Theta$ and $\psi_{k^{l+1}-1}$, we can see that $\mathcal{E}(k^{l+1}) + \Theta > \psi_{k^{l+1}-1}$.

1) If $k^{l+1} - 1 < n^l$,

$$\begin{aligned} \mathcal{E}(k^{l+1}) + \Theta &> \tau_1^l + \frac{k^{l+1} - k^l}{\rho^*} + \Theta \\ &= \psi_{k^{l+1}-1} \end{aligned}$$

2) If $k^{l+1} - 1 = n^l$,

$$\begin{aligned} \mathcal{E}(k^{l+1}) + \Theta &> \tau_1^l + \Theta + \frac{n^l - k^l}{\rho^*} + \frac{1}{\rho'} \\ &> \tau_1^l + \Theta + \frac{k^{l+1} - 1 - k^l}{\rho^*} + \frac{1}{\rho^*} \\ &= \tau_1^l + \Theta + \frac{(k^{l+1} - 1) - k^l + 1}{\rho^*} \\ &= \psi_{k^{l+1}-1} \end{aligned}$$

3) If $k^{l+1} - 1 > n^l$,

$$\begin{aligned} \mathcal{E}(k^{l+1}) + \Theta &> \tau_1^l + \Theta + \frac{n^l - k^l}{\rho^*} + \frac{k^{l+1} - n^l}{\rho'} \\ &> \tau_1^l + \Theta + \frac{n^l - k^l}{\rho^*} + \frac{k^{l+1} - n^l}{\rho'} + \frac{1}{\rho^*} - \frac{1}{\rho'} \\ &= \tau_1^l + \Theta + \frac{n^l - k^l + 1}{\rho^*} + \frac{k^{l+1} - 1 - n^l}{\rho'} \\ &= \psi_{k^{l+1}-1} \end{aligned}$$

Therefore, from the above steps we have

$$\mathcal{E}(k^{l+1}) + \Theta > \psi_{k^{l+1}-1} \tag{55}$$

Since we know the equivalence holds for active period A^l from the inductive hypothesis, we have $\mathcal{F}(k^{l+1} - 1) = \psi_{k^{l+1}-1}$. Hence, from Equation (55), we get $\mathcal{E}(k^{l+1}) + \Theta > \mathcal{F}(k^{l+1} - 1)$. Now to compute $\mathcal{F}(k^{l+1})$ of Equation (51), we need to compare $\mathcal{E}(k^{l+1}) + \Theta$ and $\mathcal{G}(k^{l+1} - h)$. This comparison is given below.

$$
\begin{aligned}
\rho' \cdot (\tau_2^l - \tau_1^l + 1) &\geq k^{l+1} - k^l \text{ (similar to Equation (43))} \\
\rho' \cdot (\tau_1^{l+1} - \tau_2^l - 1) + \rho' \cdot (\tau_2^l - \tau_1^l + 1) &\geq k^{l+1} - k^l \text{ (since } \tau_1^{l+1} - \tau_2^l \geq 1) \\
\rho' \cdot (\tau_1^{l+1} - \tau_1^l) &\geq k^{l+1} - k^l \\
\rho' \cdot (\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l)) &\geq k^{l+1} - k^l \\
\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l) &\geq \frac{k^{l+1} - k^l}{\rho'} \\
\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l) &\geq \frac{k^{l+1} - k^l}{\rho'} \\
\mathcal{E}(k^{l+1}) - \mathcal{E}(k^l) &\geq \frac{1}{\rho^*} + \frac{-h+1}{\rho'} + \frac{k^{l+1} - k^l}{\rho'} \text{ (since } \frac{1}{\rho^*} + \frac{-h+1}{\rho'} \leq 0) \\
\mathcal{E}(k^{l+1}) &\geq \mathcal{E}(k^l) + \frac{1}{\rho^*} + \frac{-h+1}{\rho'} + \frac{k^{l+1} - k^l}{\rho'} \\
\mathcal{E}(k^{l+1}) &\geq \mathcal{E}(k^l) + \frac{1}{\rho^*} + \frac{k^{l+1} - h - k^l + 1}{\rho'} \\
\mathcal{E}(k^{l+1}) + \Theta &\geq \mathcal{E}(k^l) + \Theta + \frac{1}{\rho^*} + \frac{k^{l+1} - h - k^l + 1}{\rho'} \\
\mathcal{E}(k^{l+1}) + \Theta &\geq \mathcal{G}(k^{l+1} - h) \text{ (according to Equation (34)}
\end{aligned}
$$

Now, let us evaluate the max term in Equation (51).

$$
\begin{aligned}
\mathcal{F}(k^{l+1}) &= max(\mathcal{E}(k^{l+1}) + \Theta, \mathcal{F}(k^{l+1} - 1), \mathcal{G}(k^{l+1} - h)) + \frac{1}{\rho^*} \\
&= \mathcal{E}(k^{l+1}) + \Theta + \frac{1}{\rho^*} \\
&= \psi_{k+1}
\end{aligned}
$$

This shows that the equivalence between $\mathcal{F}$ and ψ holds for the first service unit of active period A^{l+1}.

Inductive step: *for active period* A^{l+1} - In the inductive step of Theorem 1, we have showed for the first active period that if $\mathcal{F}(j) = \psi_j$ then $\mathcal{F}(j+1) = \psi_{j+1}$. That proof did not assume that service units j and $j+1$ belong to the first active period, and therefore it also holds for any active period. Hence, with this we show the equivalence between $\mathcal{F}$ and ψ holds for active period A^{l+1}.

With this, we prove by induction that the equivalence between $\mathcal{F}$ and ψ holds for all active periods.

□

III. Conclusion

The Credit-Controlled Static-Priority (CCSP) arbiter fits very well for arbitrating accesses to heavily loaded shared System-on-Chip resources with latency-critical requestors. However, the existing dataflow model is based on a service guarantee that does not capture bursty provided service. As a result, the system-level analysis gives pessimistic worst-case finishing time of requests. This ultimately leads to over-allocation of resources to satisfy real-time requirements of applications. In this document, we present a novel dataflow model of the arbiter and formally prove its equivalence to the bi-rate service guarantee that captures bursty service provision. The new dataflow model significantly improves the worst-case finishing time of requests, which leads to efficient resource utilization of resources under CCSP arbitration.