Flow Control with (Min,+) Algebra

Euriell Le Corronc, Bertrand Cottenceau, and Laurent Hardouin

Laboratoire d'Ingénierie des Systèmes Automatisés, Université d'Angers,
62, Avenue Notre Dame du Lac, 49000 Angers, France,
{euriell.lecorronc,bertrand.cottenceau,laurent.hardouin}@univ-angers.fr,
WWW home page: http://www.istia.univ-angers.fr/LISA/

Abstract. According to the theory of Network Calculus based on the (min,+) algebra, analysis and measure of worst-case performance in communication networks can be made easily. In this context, this paper deals with traffic regulation and performance guarantee of a network *i.e.* with flow control. At first, assuming that a minimum service provided by a network is known, we aim at finding the constraint over the input flow in order to respect a maximal delay or backlog. Then, we deal with the window flow control problem in the following manner: The data stream (from the source to the destination) and the acknowledgments stream (from the destination to the source) are assumed to be different and the service provided by the network is assumed to be known in an uncertain way, more precisely it is assumed to be in an interval. The results are obtained by considering the Residuation theory which allows functions defined over idempotent semiring to be inverted.

1 Introduction

Theory of (min,+) algebra enables the study of Discrete Event Dynamic Systems (DEDS) characterized by delay and synchronization phenomena such as production systems, communication networks and transportation systems (see [2]). Such systems can be described by linear models, thanks to the particular algebraic structure called *idempotent semiring* or *dioid*. In particular, the theory of Network Calculus aimed at worst-case performance analysis in communication networks. For instance, end-to-end delay or backlog can be computed with curves representing constraints over traffic and service provided by a network. Furthermore, operations defined over idempotent semiring and residuation theory allows some traffic control elements to be computed. Indeed, some model matching problems are already solved by the way of control structures (open-loop or close-loop structures) as presented in [7, 9].

By leaning on Network Calculus as well as known control synthesis problem, the work introduced in this paper deals with control and performance guarantee of traffic in networks. On the one hand, the computation of the optimal constraint applied on the input flow in order to respect a maximum delay or backlog is given. By optimal we mean that it is the lower constraint such that the delay or backlog is satisfied: if the input flow is greater than this constraint, the resulting delay or backlog will be exceeded. This computation is made assuming that a minimum service provided by the network is known. On the other hand, optimal window size of a feedback configuration (window flow control) is studied. For that computation, a difference is made between the data stream (from the source to the destination) and the acknowledgments stream (from the destination to the source). Moreover, the service provided by the network is assumed to be included in an interval.

In order to introduce this work, the paper is organized as follows. Section 2 recalls the links between Network Calculus and (min,+) algebra. In particular some properties of the algebraic tools called *idempotent semiring* or *dioid* and classical operations of Network Calculus are presented. In the third section, the modelling of a communication network is given with cumulative functions, arrival and service curves and bounds on performances (delay and backlog) of a network. Finally, problems addressed previously are stated in the fourth section and an application is given in the last section.

2 An Algebraic Approach of Network Calculus

Network Calculus is a theory based on the (min,+) algebra and devoted to the analysis of performance guarantee in communication networks (see [5], [6] and [8]). This study lies on the particular algebraic structure called *idempotent semiring* whereas well-known operations as deconvolution and subadditive closure can be seen from the point of view of the residuation theory for the former and the solution of implicit equation $x = ax \oplus b$ for the latter. All these properties are recalled in this section.

2.1 (Min,+) Algebra

(Min,+) algebra is very closed to the lattice theory and the definition below of the *idempotent semiring* gives the basis of the particular algebraic structure used in this algebra (see [2]).

Definition 1. *An idempotent semiring $\mathcal{D}$ is a set endowed with two inner operations denoted $\oplus$ and $\otimes$. The sum $\oplus$ is associative, commutative, idempotent (i.e. $\forall a \in \mathcal{D}, a \oplus a = a$) and admits a neutral element denoted ε. The product $\otimes$ is associative, distributes over the sum and accepts e as neutral element.*

When $\otimes$ is commutative (*i.e.* $\forall a, b \in \mathcal{D}, a \otimes b = b \otimes a$), the idempotent semiring $\mathcal{D}$ is said to be commutative. Furthermore, an idempotent semiring is said to be complete if it is closed for infinite sums and if the product distributes over infinite sums. Then, the greatest element of $\mathcal{D}$ is denoted T (for *Top*) and represents the sum of all its elements ($T = \bigoplus_{x \in \mathcal{D}} x$).

Due to the idempotency of the addition, a canonical order relation can be associated with $\mathcal{D}$ by the following equivalences: $\forall a, b \in \mathcal{D}$, $a \succcurlyeq b \Leftrightarrow a = a \oplus b \Leftrightarrow b = a \wedge b$. Because of the lattice properties of a complete idempotent semiring, $a \oplus b$ is the least upper bound of $\mathcal{D}$ whereas $a \wedge b$ is its greatest lower bound.

An example of this structure is the idempotent semiring $\overline{\mathbb{R}}_{min}$ defined below.

Example 1 ((Min,+) algebra). The set $\overline{\mathbb{R}}_{min} = (\mathbb{R} \cup \{-\infty, +\infty\})$ endowed with the min operator as sum $\oplus$ and the pointwise addition as product $\otimes$ is a complete idempotent semiring where $\varepsilon = +\infty$, $e = 0$ and $T = -\infty$. On $\overline{\mathbb{R}}_{min}$, the greatest lower bound $\wedge$ takes the sense of the max operator.

Remark 1. It is important to note that the order relation in $\overline{\mathbb{R}}_{min}$ corresponds to the reverse of the natural order:

$$5 \oplus 3 = 3 \quad \Leftrightarrow \quad 3 \succcurlyeq 5 \quad \Leftrightarrow \quad 3 \leq 5.$$

In the rest of this document, the order relation of $\overline{\mathbb{R}}_{min}$ is used ($\succcurlyeq$ and $\preccurlyeq$) but the natural order (respectively $\leq$ and $\geq$) will be written clearly too when it will be necessary.

2.2 Other Algebraic Preliminaries

Residuation is a general notion in lattice theory which allows "pseudo-inverse" of some isotone maps (see [3] and [2]) to be defined. In particular, the residuation theory provides optimal solutions to inequalities $f(x) \preceq b$, where f is an order-preserving mapping (*i.e.* an isotone mapping: $a \preccurlyeq b \Rightarrow f(a) \preccurlyeq f(b)$) defined over ordered sets.

Definition 2 (Residuation). *Let $f : \mathcal{D} \to \mathcal{C}$ be an isotone mapping where $\mathcal{D}$ and $\mathcal{C}$ are complete idempotent semirings. Mapping f is said residuated if $\forall b \in \mathcal{C}$, the greatest element denoted $f^\sharp(b)$ of subset $\{x \in \mathcal{D} | f(x) \preccurlyeq b\}$ exists and belongs to this subset. Mapping $f^\sharp$ is called the residual of f. Furthermore, when f is residuated, $f^\sharp$ is the unique isotone mapping such that $f \circ f^\sharp \preccurlyeq \mathsf{Id}_{\mathcal{C}}$ and $f^\sharp \circ f \succcurlyeq \mathsf{Id}_{\mathcal{D}}$, where $\mathsf{Id}_{\mathcal{C}}$ and $\mathsf{Id}_{\mathcal{D}}$ are respectively the identity mappings on $\mathcal{C}$ and $\mathcal{D}$.*

Fixed point theory allows one to find greatest finite solutions to equations $f(x) = x$, where f is an isotone mapping defined over complete idempotent semiring $\mathcal{D}$. In particular, thanks to the following theorem, the optimal solution of the implicit equation $x = ax \oplus b$ is provided.

Theorem 1. *[2, section 4.5.3] Implicit equation $x = ax \oplus b$ defined over a complete dioid $\mathcal{D}$ admits $x = a^\star b$ as lower solution where $\forall a \in \mathcal{D}, a^\star = \bigoplus_{i \geq 0} a^i$ and $a^0 = e$.*

These two theories will be necessary in the definition of operations linked to Network Calculus, as the next section shows it.

2.3 Operations of Network Calculus

Once (min,+) algebra and other tools are defined, first main operations used by Network Calculus as pointwise minimum and inf-convolution can be given. To this end, the set $\mathcal{F}$ brings together non-decreasing functions $f : \mathbb{R} \mapsto \overline{\mathbb{R}}_{min}$ where $f(t) = 0$ for $t < 0$. A restriction of this set is the set $\mathcal{F}_0$ where $f(0) = 0$. Let now f and g be two functions of $\mathcal{F}_0$, the following operations are defined:

- pointwise minimum

$$(f \oplus g)(t) = \min[f(t), g(t)],$$

- pointwise maximum

$$(f \wedge g)(t) = \max[f(t), g(t)],$$

- inf-convolution

$$(f * g)(t) \triangleq \bigoplus_{\tau \geq 0} \{f(\tau) \otimes g(t - \tau)\} = \min_{\tau \geq 0} \{f(\tau) + g(t - \tau)\}.$$

Thanks to these operations, another idempotent semiring can be defined.

Definition 3. *The set $\mathcal{F}_0$ endowed with the two inner operations $\oplus$ as pointwise minimum and $*$ as inf-convolution is a commutative idempotent semiring denoted $\{\mathcal{F}_0, \oplus, *\}$ where ε and e are defined by:*

$$\forall t, \varepsilon : t \mapsto +\infty \quad \textit{and} \quad e : t \mapsto \begin{cases} 0 & \textit{for } t < 0, \\ +\infty & \textit{for } t \geq 0. \end{cases}$$

Remark 2. As in usual algebra, operator $*$ can be omitted in order to save place:

$$ab = a * b.$$

Then, two another well-known operations of Network Calculus are the one of *deconvolution* denoted $\oslash$ and the one of *subadditive closure* denoted $\star$. Firstly, thanks to the residuation theory (see Definition 2), mapping $R_a : x \mapsto x * a$ defined over $\mathcal{F}_0$ is said to be residuated. Its residual is usually denoted $R_a^\sharp : x \mapsto x \oslash a$ and called *deconvolution*. Therefore, $b \oslash a$ is the greatest solution to inequality $x * a \preccurlyeq b$, *i.e.*:

$$b \oslash a = \hat{x} = \bigoplus \{x \mid x * a \preccurlyeq b\}.$$

Remark 3. This operation of deconvolution is also called *right quotient* and a similar mapping, called *left quotient* and denoted $L_a^\sharp : x \mapsto a \obslash x$ exists. This mapping is the residual of $L_a : x \mapsto a * x$ defined over $\mathcal{F}_0$ and $a \obslash b = \hat{x} = \bigoplus \{x \mid a * x \preccurlyeq b\}$. However, since $\mathcal{F}_0$ is commutative $b \oslash a = a \obslash b$.

Secondly, according to theorem 1, the subadditive closure operation $\star$ takes the sense of the optimal solution of a given implicit equation $a^\star = \bigwedge \{x \mid x = ax \oplus e\}$. Finally, $\forall f, g \in \mathcal{F}_0$ operations of deconvolution and subadditive closure are given:

- deconvolution

$$(f \oslash g)(t) \triangleq \bigwedge_{\tau \geq 0} \{f(\tau) - g(t - \tau)\} = \max_{\tau \geq 0} \{f(\tau) - g(\tau - t)\},$$

- subadditive closure

$$f^\star(t) \triangleq \bigoplus_{\tau \geq 0} f^\tau(t) = \min_{\tau \geq 0} f^\tau(t) \quad \text{with} \quad f^0(t) = e.$$

Numerous properties are associated with both deconvolution and subadditive closure. The following theorem brings together some of useful properties.

Theorem 2. *Firstly,* $\forall x, y, a \in \{\mathcal{F}_0, \oplus, *\}$*:*

$$x \preccurlyeq y \Rightarrow \begin{cases} a \mathbin{\diamond\!\!\!\backslash} x \preccurlyeq a \mathbin{\diamond\!\!\!\backslash} y & (x \mapsto a \mathbin{\diamond\!\!\!\backslash} x \textit{ is isotone}), \\ x \mathbin{\diamond\!\!\!\backslash} a \succcurlyeq y \mathbin{\diamond\!\!\!\backslash} a & (x \mapsto x \mathbin{\diamond\!\!\!\backslash} a \textit{ is antitone}). \end{cases} \tag{1}$$

Then:

$$(x * a) \oslash a \quad \succcurlyeq \quad x, \tag{2}$$

$$x \oslash (b * a) \quad = \quad (x \oslash a) \oslash b, \tag{3}$$

$$a \oslash b \succcurlyeq x \quad \Leftrightarrow \quad a \succcurlyeq xb, \tag{4}$$

$$(a^\star)^\star \quad = \quad a^\star. \tag{5}$$

And in particular about the subadditive closure:

$$a^\star = \bigoplus \{x \mid x^\star \preccurlyeq a^\star\}, \tag{6}$$

$$a^\star = \bigwedge \{x \mid x = x^\star, \ x \succcurlyeq a\}. \tag{7}$$

Proof. Proofs of these equations are found in literature. For equation (1), see [10] whereas for equations (2) until (5) see [2]. Finally, for equations (6) and (7) see [9] with the following precision for the latter: $a^\star$ is a solution of this equation because $a^\star \succcurlyeq a$ and it is also the smallest one because $(a^\star)^\star \succcurlyeq a^\star \Leftrightarrow a^\star \succcurlyeq a^\star \Leftrightarrow x^\star \succcurlyeq a^\star$. □

3 Network Calculus Modelling

3.1 Input and Output Flows, Arrival and Service Curves

Input and output flows. A communication network can be seen as a blackbox denoted S with an input flow and an output flow. These flows are respectively described by cumulative functions belonging to $\mathcal{F}_0$ and denoted u and y. Element $u(t)$ corresponds to the total amount of data introduced in the system until time t whereas $y(t)$ corresponds to the total amount of data that has left the system until this time. The main assumption made about input and output flows is a characteristic of causality:

$$u \preccurlyeq y,$$

which means that for all t, $u(t) \geq y(t)$. So, the amount of data leaving the network is always lower than the one getting in.

In order to guarantee performance in network, constraints are applied over these flows. For instance, an arrival curve is applied over the input flow whereas the service provided by S is constrained by a lower curve as well as an upper curve.

Arrival curve. One says that a given flow $u \in \mathcal{F}_0$ is constrained by an arrival curve $\alpha \in \mathcal{F}_0$ if it is such that $\forall s \leq t \in \mathbb{R}_+$, $u(t) - u(s) \leq \alpha(t-s)$ (u is said α-smooth). So, the amount of data arriving between time s and time t is at most $\alpha(t-s)$. Firstly, according to the definition of the inf-convolution:

$$u(t) - u(s) \leq \alpha(t-s) \quad \Leftrightarrow \quad u \leq \alpha u. \quad \Leftrightarrow \quad u \succcurlyeq \alpha u.$$

Secondly, thanks to the isotony of the inf-convolution, this inequality can be written as below:

$$u \succcurlyeq \alpha u \;\Rightarrow\; \alpha u \succcurlyeq (\alpha^2 u) \;\Rightarrow\; (\alpha^2 u) \succcurlyeq (\alpha^3 u) \;\Rightarrow\; \ldots$$

and therefore:

$$u = u \oplus (\alpha u) \oplus (\alpha^2 u) \oplus \ldots = \bigoplus_{n \geq 0} \alpha^n u.$$

Finally:

$$u \succcurlyeq \alpha u \quad \Leftrightarrow \quad u = \alpha^\star u. \tag{8}$$

So, α is an arrival curve[1] for u if and only if for the input flows considered we have: $u = \alpha^\star u$.

Service curve. As regards to the service provided by S, it is framed by two service curves $\underline{\beta}$ and $\overline{\beta} \in \mathcal{F}_0$ such that $\underline{\beta} \preccurlyeq \overline{\beta}$ ($\underline{\beta} \geq \overline{\beta}$). These curves constitute interval $[\ \underline{\beta}\ ,\ \overline{\beta}\ ]$ where $\overline{\beta}$ corresponds to the minimum service provided by S for all input flows and $\underline{\beta}$ corresponds to its maximum service. Then, output flow y is included in an interval too:

$$\underline{\beta} u \;\preccurlyeq\; y \;\preccurlyeq\; \overline{\beta} u \quad \Leftrightarrow \quad y \in [\ \underline{\beta} u\ ,\ \overline{\beta} u\ ]. \tag{9}$$

All these Network Calculus elements (input and output flows, arrival and service curve) are illustrated in Figure 1.

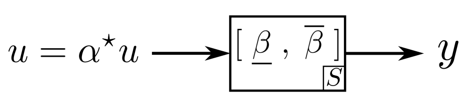


Fig. 1: Network Calculus diagram.

3.2 Performance Characteristics: Delay and Backlog

Two characteristics used as performance indicator in Network Calculus are the delay and the backlog (see [4]). The former denoted $d(t)$ corresponds to the

[1] And $\alpha^\star$ is also an arrival curve for u.

waiting time of a paquet in a FIFO order whereas the latter denoted $b(t)$ is the amount of paquets in a network at time t. Let u and y be the input and the output flow of a network:

$$d(t) \triangleq \inf_{\tau \geq 0} \{\tau \mid u(t) \leq y(t+\tau)\},$$
$$b(t) \triangleq u(t) - y(t).$$

These data are given for all time t in the network. However, according to the following theorem coming from the second order theory of (min,+) linear systems detailed in [10] and used in [11], upper bounds on their worst-case can be measured easily.

Theorem 3. *Let v_1 and v_2 be two functions of $\mathcal{F}$ where*[2] $v_1 \preccurlyeq v_2$*. Function $v_1 \not\phi v_2$ is called the correlation of v_1 over v_2 and contains the maximal distances, denoted $\tau_{\max}$ and $\nu_{\max}$, between v_1 and v_2 respectively in time and event domain. More precisely, $\tau_{\max}$ and $\nu_{\max}$ are such that:*

$$\tau_{\max} = \inf_{\tau \geq 0} \{\tau \mid (v_1 \not\phi v_2)(-\tau) \leq 0\},$$
$$\nu_{\max} = (v_1 \not\phi v_2)(0).$$

Remark 4. It is possible that $v_1 \not\phi v_2 = \varepsilon$. In such a case, maximal time and event distances $\tau_{\max}$ and $\nu_{\max}$ are infinite.

Then, thanks to theorem 3, we are able to provide two kinds of distances for a network S:

- if input and output flows u and y of S are assumed to be known then, its *maximal delay and backlog* can be computed,
- if arrival curve of u and minimum service curve of S are assumed to be known then, *upper bounds* on maximal delay and backlog can be computed.

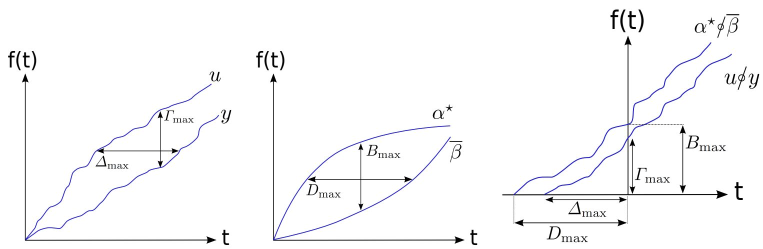
These measures as well as links between them are given in the following proposition. This is a different formulation of some well known results (see [8, §3.1.11]) but with different tools.

Proposition 1. *On the one hand, let $u, y \in \mathcal{F}_0$ be input and output flows of a network S such that $u \preccurlyeq y$. On the other hand, let α be the arrival curve of input u such that $u = \alpha^\star u$ and $\overline{\beta}$ be a minimum service curve of S such that $y \preccurlyeq \overline{\beta} u$. Then:*

$$\begin{array}{lllll} d(t) & \leq & \Delta_{\max} = \inf_{\Delta \geq 0} \{(u \not\phi y)(-\Delta) \leq 0\} & \leq & D_{\max} = \inf_{D \geq 0} \{(\alpha^\star \not\phi \overline{\beta})(-D) \leq 0\}, \\ b(t) & \leq & \Gamma_{\max} = (u \not\phi y)(0) & \leq & B_{\max} = (\alpha^\star \not\phi \overline{\beta})(0). \end{array}$$

Distances $\Delta_{\max}$ and $\Gamma_{\max}$ are the maximal delay and backlog of S whereas $D_{\max}$ and $B_{\max}$ are their upper bounds (see Figure 2).

[2] Recall that $v_1 \preccurlyeq v_2 \Leftrightarrow \forall t, v_1(t) \geq v_2(t)$.



(a) Flows u and y with distances $\Delta_{\max}$ and $\Gamma_{\max}$

(b) Curves $\alpha^\star$ and $\overline{\beta}$ with distances $D_{\max}$ and $B_{\max}$

(c) Correlations $u \not\circ y$ and $\alpha^\star \not\circ \beta$

Fig. 2: Maximal delay and backlog and their upper bounds.

Proof. Since $u = \alpha^\star u$ and $y \preccurlyeq \overline{\beta} u$ (see equations (8) and (9)) are the relation between real flows of network S and constraints over these flows, the following inequality shows that from the correlation $u \not\circ y$ another correlation with $\alpha^\star$ and $\overline{\beta}$ is found:

$$\begin{array}{lcllr} y \preccurlyeq \overline{\beta} u & \Rightarrow & u \not\circ y & \succcurlyeq (\alpha^\star u) \not\circ (\overline{\beta} u) & \text{since } u = \alpha^\star u \text{ and see (1)}, \\ & & & \succcurlyeq ((\alpha^\star u) \not\circ u) \not\circ (\overline{\beta}) & \text{see (3)}, \\ & & & \succcurlyeq \alpha^\star \not\circ \overline{\beta} & \text{see (2)}. \end{array}$$

So, according to the $\overline{\mathbb{R}}_{min}$ order relation:

$$u \not\circ y \succcurlyeq \alpha^\star \not\circ \overline{\beta} \quad \Leftrightarrow \quad u \not\circ y \leq \alpha^\star \not\circ \overline{\beta}.$$

□

Remark 5. In the Network Calculus literature, maximal distances $D_{\max}$ and $B_{\max}$ are obtained by horizontal and vertical deviations between elements of correlation $\alpha^\star \not\circ \overline{\beta}$ as shown in Figure 2b.

3.3 Functions associated to Delay and Backlog

In the next sections dealing with some control problems, we will need to handle given delay and backlog as fixed value of pure delay and amount of data. To this end, particular functions are defined below.

Definition 4. *Let τ and ν be respectively a pure delay and an amount of data. Then, function denoted δ_τ is defined by:*

$$\delta_\tau(t) = \begin{cases} 0 & \text{for } t < \tau, \\ +\infty & \text{for } t \geq \tau, \end{cases}$$

and the one denoted γ_ν by:

$$\gamma_\nu(t) = \begin{cases} \nu & \text{for } t < 0, \\ +\infty & \text{for } t \geq 0. \end{cases}$$

Remark 6. Some properties can be associated to these functions:

$$\delta_\tau * \delta_{-\tau} = e,$$
$$\gamma_\nu * \gamma_{-\nu} = e.$$

Moreover, in relation to input flow u, these functions are such that:

$$\forall t,\ (\delta_\tau * u)(t) = u(t-\tau),$$
$$\forall t,\ (\gamma_\nu * u)(t) = \nu + u(t).$$

4 Flow Control

In this section, we consider the traffic regulation in order to get a guaranteed performance of a network, this is known as the flow control.

4.1 Arrival Curve Computation

The first problem addressed in this paper is the next one. Assuming that a minimum service provided by a network is known, we aim at finding the arrival curve, *i.e.* a constraint applied over input flow, in order to respect a maximal delay or backlog. By definition of an arrival curve (see equation (8) with $u \succcurlyeq \alpha u$ and so $u = \alpha^\star u$), this optimal curve is a subadditive closure. Moreover, an optimal curve represents the minimal constraint applied over the input in order to eventually reach but not exceed the given delay or backlog. The problem from the point of view of time performance is given in the following proposition.

Proposition 2. *Let $\overline{\beta}$ be a minimal service curve of a network S and τ be a fixed worst end-to-end delay. The optimal arrival curve $\hat{\alpha}^\star$ which guarantees the respect of τ is given by:*

$$\hat{\alpha}^\star = \bigwedge\{\alpha^\star \mid \alpha^\star \succcurlyeq \delta_{-\tau}\overline{\beta}\} = (\delta_{-\tau}\overline{\beta})^\star$$

where $\delta_{-\tau}$ is the function associated with τ.

Proof. First of all, according to proposition 1 and definition 4, upper bound $D_{\max}$ of worst end-to-end delay is given by correlation $\alpha^\star \oint \overline{\beta}$ and can be represented by function $\delta_{-D_{\max}}$. So, the following relation is given:

$$D_{\max} = -\sup\{D \mid (\alpha^\star \oint \overline{\beta})(D) \leq 0\} \quad \Rightarrow \quad \alpha^\star \oint \overline{\beta} \succcurlyeq \delta_{-D_{\max}}.$$

Then, if the worst end-to-end delay τ is chosen $D_{\max} = \tau \Leftrightarrow \delta_{-D_{\max}} = \delta_{-\tau}$ and thanks to equation (4), arrival curve $\alpha^\star$ has to follow these following inequalities:

$$\alpha^\star \oint \overline{\beta} \succcurlyeq \delta_{-\tau} \quad \Leftrightarrow \quad \alpha^\star \succcurlyeq \delta_{-\tau}\overline{\beta}.$$

Finally, thanks to equation (7), the minimal $\alpha^\star$ which respects the inequality is $(\delta_{-\tau}\overline{\beta})^\star$. $\square$

The next proposition states the problem from the point of view of data performance.

Proposition 3. *Let $\overline{\beta}$ be the known minimal service curve of a network S and ν be a fixed worst backlog. The optimal arrival curve $\hat{\alpha}^\star$ which guarantees the respect of ν is given by:*

$$\hat{\alpha}^\star = \bigwedge\{\alpha^\star \mid \alpha^\star \succcurlyeq \gamma_\nu \overline{\beta}\} = (\gamma_\nu \overline{\beta})^\star$$

where γ_ν is the function associated with ν.

Proof. The proof is the same as in proposition 2. □

4.2 Window Flow Control

The second problem of traffic regulation and performance guarantee is the one of the window flow control where its optimal window size is computed.

First of all, let us recall this control context. A window flow controller aims at bounding the amount of data admitted in a network in such a way that its total amount in transit is always less than some positive number, *i.e.* the window size. This problem has already been treated in literature but not in the same manner. The window flow control introduced in [8] do not have the same model than in [6]. In this paper, we adopt the Chang's modelling which is homogeneous with the one introduced in [7].

Moreover, this problem is studied here with two associated configurations. On the one hand, the service provided by the network is assumed to be included in interval. Indeed, assuming that minimum and maximum service curves are known, the size of the window can be computed as well as for the worst case than for the best case of traffic without damaging the service provided. On the other hand, a difference is made between the data stream (from the source to the destination) and the acknowledgments stream (from the destination to the source) since these acknowledgments requires considerably less bandwidth than the data itself (see [1]). So, the computation of the window size will have benefit of this profit of bandwidth. This configuration is described in the following proposition and illustrated in Figure 3.

Proposition 4. *Let S_1 be the system representing the data stream where $[\ \underline{\beta}_1\ ,\ \overline{\beta}_1\ ]$ ($\underline{\beta}_1 \preccurlyeq \overline{\beta}_1$) is the interval containing its provided service. In the same way, let S_2 be the system representing the acknowledgments stream where $[\ \underline{\beta}_2\ ,\ \overline{\beta}_2\ ]$ ($\underline{\beta}_2 \preccurlyeq \overline{\beta}_2$) is the interval containing its provided service. Then, let γ_w be the representative function of the window size w. The service curve of the whole system is included in the interval:*

$$[\ \underline{\beta}_1(\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star\ ,\ \overline{\beta}_1(\gamma_w \overline{\beta}_2 \overline{\beta}_1)^\star\ ].$$

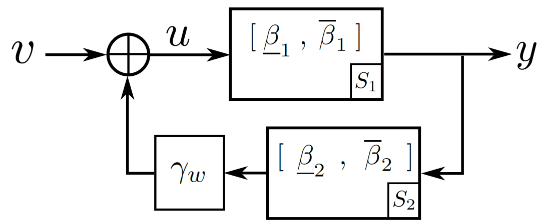


Fig. 3: Chosen configuration of the window flow control system.

Proof. The output flow y is described by the following equation:

$$\underline{\beta}_1 u \preccurlyeq \ y \ \preccurlyeq \overline{\beta}_1 u,$$

whereas intermediate flow u is included in:

$$\begin{aligned}
\min(v, \gamma_w \underline{\beta}_2 \underline{\beta}_1 u) \preccurlyeq \ & u \ \preccurlyeq \min(v, \gamma_w \overline{\beta}_2 \overline{\beta}_1 u), & \\
v \oplus \gamma_w \underline{\beta}_2 \underline{\beta}_1 u \preccurlyeq \ & u \ \preccurlyeq v \oplus \gamma_w \overline{\beta}_2 \overline{\beta}_1 u & \text{see operator } \oplus, \\
(\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star v \preccurlyeq \ & u \ \preccurlyeq (\gamma_w \overline{\beta}_2 \overline{\beta}_1)^\star v & \text{see theorem 1.}
\end{aligned}$$

Therefore:

$$\underline{\beta}_1 (\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star v \preccurlyeq \ y \ \preccurlyeq \overline{\beta}_1 (\gamma_w \overline{\beta}_2 \overline{\beta}_1)^\star v.$$

□

By considering this configuration, the computation of the optimal window size $\hat{w}$ can be studied. The chosen point of view is to compute a minimal window size such that the global network behavior, *i.e.* the controlled one, is the same as the open-loop network behavior, *i.e.* the one of S_1 only. The behavior of S_1 is described by the interval of service curve $[\ \underline{\beta}_1\ ,\ \overline{\beta}_1\ ]$, this objective can be stated as follows:

$$\widehat{\gamma_w} \ = \ \bigoplus \{\gamma_w \mid \underline{\beta}_1 (\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star = \underline{\beta}_1 \quad \text{and} \quad \overline{\beta}_1 (\gamma_w \overline{\beta}_2 \overline{\beta}_1)^\star = \overline{\beta}_1\}. \qquad (10)$$

The following proposition puts forward the computation of such a window size.

Proposition 5. *In order to obtain a behavior of the closed-loop system unchanged in comparison to the one of the open-loop (see equation (10)), the optimal window size $\hat{w}$ represented by function $\widehat{\gamma_w}$ is given below:*

$$\widehat{\gamma_w} = (\underline{\beta}_1 \circ\!\!\!\!\backslash\, \underline{\beta}_1 \circ\!\!\!\!/\, (\underline{\beta}_2 \underline{\beta}_1)) \wedge (\overline{\beta}_1 \circ\!\!\!\!\backslash\, \overline{\beta}_1 \circ\!\!\!\!/\, (\overline{\beta}_2 \overline{\beta}_1)).$$

Proof. Firstly, by considering the minimal bound, let $\underline{G}_c$ be the minimal behavior of the controlled network and $\underline{G}_{ref}$ be the reference behavior we want to reach, so the one of $\underline{\beta}_1$:

$$\underline{G}_c = \underline{\beta}_1 (\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star \quad \text{and} \quad \underline{G}_{ref} = \underline{\beta}_1.$$

Equation (10) can be written as follow:

$$\hat{\gamma_w} \;=\; \bigoplus\{\gamma_w \mid \underline{G}_c \preccurlyeq \underline{G}_{ref}\}.$$

Then:

$$\begin{array}{rcll}
\underline{G}_c \preccurlyeq \underline{G}_{ref} & \Leftrightarrow & \underline{\beta}_1(\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star \preccurlyeq \underline{\beta}_1, & \\
& \Leftrightarrow & (\gamma_w \underline{\beta}_2 \underline{\beta}_1)^\star \preccurlyeq \underline{\beta}_1 \circ\!\!\!\!\diagdown \underline{\beta}_1 & \text{see (4)}, \\
& \Leftrightarrow & \gamma_w \underline{\beta}_2 \underline{\beta}_1 \preccurlyeq \underline{\beta}_1 \circ\!\!\!\!\diagdown \underline{\beta}_1 & \text{since } a \preccurlyeq a^\star, \\
& \Leftrightarrow & \gamma_w \preccurlyeq \underline{\beta}_1 \circ\!\!\!\!\diagdown \underline{\beta}_1 \circ\!\!\!\!\diagup (\underline{\beta}_2 \underline{\beta}_1) & \text{see (4)}.
\end{array} \tag{11}$$

Secondly, this proof is the same for the maximal bound and thus we obtain:

$$\gamma_w \preccurlyeq \overline{\beta}_1 \circ\!\!\!\!\diagdown \overline{\beta}_1 \circ\!\!\!\!\diagup (\overline{\beta}_2 \overline{\beta}_1). \tag{12}$$

Then, in order to satisfy both equations (11) and (12), function γ_w is given by:

$$\hat{\gamma_w} = (\underline{\beta}_1 \circ\!\!\!\!\diagdown \underline{\beta}_1 \circ\!\!\!\!\diagup (\underline{\beta}_2 \underline{\beta}_1)) \wedge (\overline{\beta}_1 \circ\!\!\!\!\diagdown \overline{\beta}_1 \circ\!\!\!\!\diagup (\overline{\beta}_2 \overline{\beta}_1)).$$

□

Remark 7. Thanks to this optimal window size, the interval including the service curve of the controlled system is the same as the one of the open-loop system:

$$[\ \underline{\beta}_1(\hat{\gamma_w} \underline{\beta}_2 \underline{\beta}_1)^\star\ ,\ \overline{\beta}_1(\hat{\gamma_w} \overline{\beta}_2 \overline{\beta}_1)^\star\ ] \quad = \quad [\ \underline{\beta}_1\ ,\ \overline{\beta}_1\ ].$$

5 Application: Window Flow Control with a Given Delay

Let us see an example of a window flow control with a given delay to respect. This application takes the main propositions of this paper into account, namely proposition 2 about the computation of an arrival curve according to a given delay (its backlog version given in proposition 3 is not treated in this example) and proposition 5 about optimal size of a window flow controller.

5.1 Configuration

For this application, the scheme of the network is the same as described in proposition 4 and illustrated in Figure 3. All the service provided by element S_1 is included in interval $[\ \underline{\beta}_1\ \ \overline{\beta}_1\ ]$. Services curves $\underline{\beta}_1$ and $\overline{\beta}_1$ are rate-latency functions with a latency of $16ms$ for the former, $20ms$ for the latter and a rate of $100Mb/s$ for both of them:

$$\underline{\beta}_1(t) = 16ms + 100Mb/s \cdot t \quad \text{and} \quad \overline{\beta}_1(t) = 20ms + 100Mb/s \cdot t.$$

By considering the service provided by element S_2 included in interval $[\ \underline{\beta}_2\ \ \overline{\beta}_2\ ]$ with rate-latency functions as services curves $\underline{\beta}_2$ and $\overline{\beta}_2$:

$$\underline{\beta}_2(t) = 12ms + 100Mb/s \cdot t \quad \text{and} \quad \overline{\beta}_2(t) = 14ms + 100Mb/s \cdot t.$$

For this network, the upper bound $D_{\max}$ of the worst end-to-end delay from v to y is fixed to $90ms$. This delay is represented by function $\delta_{-D_{\max}}$.

5.2 Computation of the arrival curve $\hat{\alpha}^\star$

Firstly, proposition 2 is applied in order to find the minimal arrival curve $\hat{\alpha}^\star$ which allows $D_{\max}$ to be respected in the open-loop context. Thus:

$$\hat{\alpha}^\star = (\delta_{-D_{\max}}\overline{\beta}_1)^\star \quad \Rightarrow \quad \begin{aligned} \hat{\alpha}^\star(t) &= (-90ms) * (20ms + 100Mb/s \cdot t), \\ &= 9Mb + 100Mb/s \cdot t. \end{aligned}$$

This arrival curve is the lowest one, so the less restrictive one enabling eventually to reach but not to exceed the given delay. If an arrival curve still less restrictive is chosen, the network will be subjected to congestions and the maximum end-to-end delay of the network will increase.

5.3 Computation of the window size $\hat{w}$

Secondly, we can compute the optimal window size $\hat{w}$ for this configuration. However, the maximum end-to-end delay $D_{\max}$ has to be respected again and thus the optimal arrival curve $\hat{\alpha}^\star$ previously computed is used as follows: input v of the global system is constrained by $\hat{\alpha}^\star$ such that $v = \hat{\alpha}^\star v$. So, the open-loop behavior is the following interval:

$$\hat{\alpha}^\star[\ \underline{\beta}_1\ ,\ \overline{\beta}_1\ ]$$

whereas the closed-loop one is:

$$\hat{\alpha}^\star[\ \underline{\beta}_1(\hat{\gamma_w}\underline{\beta}_2\underline{\beta}_1)^\star\ ,\ \overline{\beta}_1(\hat{\gamma_w}\overline{\beta}_2\overline{\beta}_1)^\star\ ].$$

Then, proposition 5 is applied in order to find the minimal window size $\hat{w}$ which allows the same behavior in closed-loop context than in open-loop context to be obtained and function $\hat{\gamma_w}$ is given by:

$$\hat{\gamma_w} = ((\underline{\beta}_1\hat{\alpha}^\star) \mathbin{\circ\backslash} (\underline{\beta}_1\hat{\alpha}^\star) \mathbin{\circ/} (\underline{\beta}_2\underline{\beta}_1)) \wedge ((\overline{\beta}_1\hat{\alpha}^\star) \mathbin{\circ\backslash} (\overline{\beta}_1\hat{\alpha}^\star) \mathbin{\circ/} (\overline{\beta}_2\overline{\beta}_1)).$$

The proof of this result is left to reader by following the one of proposition 5. Finally, the optimal window size is obtained:

$$\hat{w} = 2,8Mb.$$

This window size is the minimal one for the largest bandwidth of the network without congestion it. Moreover, this window size respect the maximum end-to-end delay given in the assumption.

6 Conclusion

In this paper, traffic regulation and performance guarantee of a network have been treated. First of all, we recalled the algebraic linked to Network Calculus operations thanks to the (min,+) algebra and the residuation theory. Once these useful properties defined, we used them in the context of flow control in order to solve two problems enabling to avoid congestion in the network.

The first case shows the computation of an optimal arrival curve in order to respect a maximal delay or backlog, assuming that the minimum service provided by a network is known. This arrival curve is said to be optimal because it is the less restrictive one where the given delay is not exceeded.

The second case brings forward the computation of a window size in a closed-loop structure. Assuming that the data stream and the acknowledgments are different, this window size is said to be optimal. Moreover, the service provided by network elements are included in an interval so the window flow controller is computed as well as for the worst case than for the best case of traffic.

Finally, an example applies propositions made in order to solve these two problems and optimal arrival curve and window size are found.

References

1. Agrawal R, Cruz RL, Okino C, Rajan R (1999) Performance bounds for flow control protocols. IEEE/ACM Transactions on Networking (TON), IEEE Press 7(3):310–323.
2. Baccelli F, Cohen G, Olsder GJ, Quadrat J-P (1992) Synchronisation and linearity: an algebra for discrete event systems. Wiley and sons.
3. Blyth TS, Janowitz MF (1972) Residuation theory. Pergamon.
4. Bouillard A, Jouhet L, Thierry E (2008) Computation of a (min,+) multi-dimensional convolution for end-to-end performance analysis. Proceedings of the 3rd International Conference on Performance Evaluation Methodologies and Tools. ValueTools'08.
5. Chang CS (1998) On deterministic traffic regulation and service guarantees: asystematic approach by filtering. IEEE Transactions on Information Theory, 44(3):1097–1110.
6. Chang CS (2000) Performance guarantees in communication networks. Springer Verlag.
7. Cottenceau B, Hardouin L, Boimond J-L, Ferrier, J-L (2001) Model reference control for timed event graphs in dioids. Automatica, Elsevier 37(9):1451–1458.
8. Le Boudec J-Y, Thiran P (2001) Network calculus: a theory of deterministic queuing systems for the internet. Springer.
9. Maia CA, Hardouin L, Santos-Mendes R, Cottenceau B (2003) Optimal closed-loop control of timed event graphs in dioids. IEEE Transactions on Automatic Control 48(12):2284–2287.
10. Max Plus (1991) Second order theory of min-linear systems and its application to discrete event systems. Proceedings of the 30th IEEE Conference on Decision and Control, Brighton. CDC'91.
11. Santos-Mendes R, Cottenceau B, Hardouin L (2005) Adaptive feedback control for (max,+)-linear systems. Proceedings of the 10th IEEE Conference on Emerging Technologies and Factory Automation. ETFA'05.