Comparison of different classes of service curves in Network Calculus

Anne Bouillard * **Laurent Jouhet** ** **Éric Thierry** **

* *ENS Cachan (Bretagne) / IRISA, Rennes, France (e-mail: anne.bouillard@bretagne.ens-cachan.fr).*
** *LIP, ENS Lyon, France, (e-mail: laurent.jouhet,eric.thierry@ens-lyon.fr)*

Abstract: In envelope-based models for worst-case performance evaluation like Network Calculus or Real-Time Calculus, several types of service curves have been introduced to quantify some deterministic service guarantees. This paper studies the expressiveness of those different definitions of service curves. We revisit the hierarchy ranging from the most restrictive definition linked to *variable capacity nodes* to the most general definition of *simple service curves*. We state the conditions when the different definitions overlap and discuss the existence of canonical descriptions for systems specified through those definitions.

Keywords: Network Calculus, (min,+)-algebra, service curves

1. INTRODUCTION

Network Calculus (NC) is a theory of deterministic queuing systems encountered in communications networks. It is based on $(\min,+)$ algebra and it can be seen as a $(\min,+)$ filtering theory by analogy with the $(+,\times)$ filtering theory used in traditional system theory. More than just a formalism, it enables to analyze complex systems and to prove deterministic bounds on delays, backlogs and other Quality-of-Service (QoS) parameters. The analysis usually focuses on worst-case performances. The information about the system is stored in functions such as arrival curves shaping the traffic or service curves quantifying the service guaranteed at the network nodes. These functions can be combined together thanks to special Network Calculus operations, in order to analyze the system and compute bounds on local performances (*i.e.* maximum buffer size at a node) or on end-to-end performances (*i.e.* maximum end-to-end delay). At the present time, the theory has developed and yielded accomplished results which are mainly recorded in two reference books: Chang (2000) and Le Boudec and Thiran (2001). A nice survey of NC including recent results can be found in Fidler (2010).

Nevertheless it remains difficult to draw the exact borders of Network Calculus at the time being. One of the main obstacles comes from the apparent variety of *service curve* definitions in the literature which might lead to different types of models. Readers are often warned to stick to the definition chosen in each paper in order to ensure the relevance of the model and the validity of its analysis. However it is rarely questioned whether another choice may lead to the same model or at least to the same performance evaluation.

Our general objective is to unveil the differences between models yielded by the different definitions of service curves. As a first step, we study their expressiveness. Such comparisons exist as folklore in the Network Calculus literature, but they are scattered, some proofs are missing and some comparisons are unset. Moreover the question of canonical service specifications has not been tackled. Given a system constrained by a family of service curves of a fixed type, is it possible to reduce or transform this family into a canonical one? Is it possible to translate it into a family of another type of service curves? We investigate those key issues in this paper. Those are necessary premises to present NC as an unified theory.

Section 2 introduces the NC framework and the service curve definitions we wish to classify: *variable capacity node service*, *strict service*, *weakly strict service*, *simple service*. In Section 3, we compare those different definitions which fit into each other but with some large gaps. We fully characterize the cases when they are equivalent and for instance we show that the equivalence between variable capacity node service and strict service, often quoted as a folklore result, is often true but not always. We also state several results about the translation of service curve families into other families of the same type or of a different type. As soon as simple services are involved, these are mainly impossibility results. Section 4 concludes the paper with first assessments about those comparisons and discusses future works required to complete the big picture.

Note that the study also concerns alternative theories like Real-Time Calculus (RTC - Thiele et al. (2000); Wandeler (2006) or Sensor Calculus (SC - Schmitt and Roedig (2005)) that use extremely close formalisms (envelope-based models, $(\min,+)$ algebra). In particular, RTC models are described as a combination of RTC Greedy Processing Components which have exactly the same behavior as NC Variable Capacity Nodes, as shown in Bouillard et al. (2009).

⋆ ONERA (CNRS 038361)/ ANR PEGASE SEGI 009 02

2. DEFINITIONS AND NOTATION

2.1 NC functions and operations

Network Calculus' primal objective is the performance analysis of communication networks. Flows and services in the network are modelled by non-decreasing functions $t \mapsto f(t)$ where t is *time* and $f(t)$ an amount of *data*. There are different models depending on whether t (resp. $f(t)$) takes discrete or continuous values, *e.g.* in $\mathbb{N}$ or $\mathbb{R}_+$. In this paper, we will present the results in a *fluid model* where time and data quantities belong to $\mathbb{R}_+$, but it can be easily checked that they directly apply to models where time or data are discrete. Note also that we do not exactly use the term *fluid* as in Le Boudec and Thiran (2001) where the *fluid model* adds the condition that the manipulated functions are continuous.

In Network Calculus, one must distinguish two kinds of objects: the real movements of data and the constraints that these movements satisfy. The real movements of data are mainly modeled by *cumulative functions*: a cumulative function $f(t)$ counts the total amount of data that has achieved some condition up to time t (*e.g.* the total amount of data which has gone through a given place in the network). In all the paper, *we make the usual assumption that cumulative functions are left-continuous.* This is not a huge restriction for the modeler. This assumption has nevertheless a technical importance in the Network Calculus edifice (*e.g.* when defining the start of backlogged periods). On the contrary, no assumption of (left- or right-)continuity is imposed to the constraint functions.

In the paper, Network Calculus functions will belong to $\mathcal{F}$ the set of functions from $\mathbb{R}_+$ into $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Cumulative functions usually belong to $\mathcal{F}_\uparrow = \{f \in \mathcal{F} \mid f$ non-decreasing, left-continuous, $f(0) = 0\}$.

Beyond usual operations like the minimum or the addition of functions, Network Calculus makes use of several classical operations which are the translations of $(+, \times)$ filtering operations into the $(\min, +)$ setting, as well as a few other transformations. Here is a sample of operations that can be encountered: let $f, g \in \mathcal{F}$, $\forall t \in \mathbb{R}_+$,

- (Inf-)convolution: $(f * g)(t) = \inf_{0 \le s \le t}(f(s) + g(t-s))$.
- Sup-convolution: $(f \overline{*} g)(t) = \sup_{0 \le s \le t}(f(s) + g(t-s))$.
- (Inf-)deconvolution: $(f \oslash g)(t) = \sup_{u \ge 0}(f(t+u) - g(u))$.
- Sup-deconvolution: $(f \overline{\oslash} g)(t) = \inf_{u \ge 0}(f(t+u) - g(u))$.
- Positive rounding: $f_+(t) = \max(f(t), 0)$.
- Positive and non-decreasing upper closure: $f_\uparrow(t) = \max(\sup_{0 \le s \le t} f(s), 0)$.
- Sub-additive closure: $f^* = \inf_{n \in \mathbb{N}} f^{(n)}$ where $f^{(n)} = f * \cdots * f$, n times, and $f^{(0)} = 0$ at 0 and $= +\infty$ elsewhere.
- Super-additive closure: $f^{\overline{*}} = \sup_{n \in \mathbb{N}} f^{\overline{(n)}}$ where $f^{\overline{(n)}} = f \overline{*} \cdots \overline{*} f$, n times, and $f^{\overline{(0)}} = 0$ at 0 and $= -\infty$ elsewhere.

Such operations have interesting algebraic properties (*e.g.* see Baccelli et al. (1992)). Network Calculus formulas use such operations to combine the curves constraining the traffic and the services in the network, in order to output worst-case performance bounds.

2.2 NC input/output systems

An NC model for a communication network usually consists in a partition of the network into subsystems which may have different scales (from elementary hardware like a processor to large sub-networks), a description of data flows, where each flow follows a path through a specified sequence of subsystems and where each flow is shaped by some arrival curve just before entering the network, a description of the behavior of each subsystem, that is service curves bounding the performances of each subsystem, as well as service policies in case of multiplexing (several flows entering the same subsystem and thus sharing its service).

Systems or sub-systems are described as input/output systems (where the number of inputs is the same as the number of outputs). An *(acceptable) trajectory* for a system crossed by p flows is a set of cumulative functions $(A_k)_{1 \le k \le p}$ and $(B_k)_{1 \le k \le p}$ in $\mathcal{F}_\uparrow$ (where A_k and B_k respectively correspond to the cumulative functions of flow k at the input and the output of the system). For now, a *system* $\mathcal{S}$ over p flows will be simply defined as the set of all its acceptable trajectories, that is $\mathcal{S} \subseteq \mathcal{F}_\uparrow^p \times \mathcal{F}_\uparrow^p$. Such a black boxed view is usual in classical filtering theory and enables to deal with any scale of system. Note also that this definition allows to consider *deterministic dynamics* (one output for one input) and *non-deterministic dynamics* (several possible outputs for one input).

2.3 NC main performance measures: backlog & delay

Let (A, B) be an input/output trajectory for a flow in a system. Then the *global backlog* of the flow at time t is $b(t) = A(t) - B(t)$ and the delay (under the FIFO policy assumption) endured after z input bits is $d(z) = B^{(-1)}(z) - A^{(-1)}(z)$ where for all $f \in \mathcal{F}$, $f^{(-1)}(z) = \inf\{t \ge 0 \mid f(t) \ge z\}$ (*pseudo-inverse*). Now for a system $\mathcal{S}$, the *worst-case backlog over* $\mathcal{S}$ is $b_{\max} = \sup_{(A,B) \in \mathcal{S}} \sup_{t \ge 0} A(t) - B(t)$ and the *worst-case delay over* $\mathcal{S}$ is $d_{\max} = \sup_{(A,B) \in \mathcal{S}} \sup_{z \ge 0} B^{(-1)}(z) - A^{(-1)}(z)$.

Given a trajectory $(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow$, a *backlogged period* is an interval $I \subseteq \mathbb{R}_+$ of time during which the backlog is non-null, *i.e.* $\forall u \in I$, $A(u) > B(u)$. Let $t \in \mathbb{R}_+$, the *start of the backlogged period* of t is $start(t) = \sup\{u \le t | A(u) = B(u)\}$. Since the cumulative functions A and B are assumed left-continuous, we also have $A(start(t)) = B(start(t))$. If $A(t) = B(t)$, then $start(t) = t$. For any $t \in \mathbb{R}_+$, $]start(t), t[$ is a backlogged period ($]start(t), t]$ if $A(t) > B(t)$).

In the definition of backlogged period, the interval I can be closed, semi-closed or open. Such a flexible definition is convenient in some future definitions or proofs where the precise description of trajectories requires a particular type of intervals, *e.g.* semi-closed rather than open (see the consequences of such choices in Section 2.6). Note that in the literature, backlogged periods have been sometimes defined for open intervals only Bouillard et al. (2008) (page 885) or without worrying about this question Le Boudec and Thiran (2001) (Definition 1.3.2, page 21).

2.4 NC arrival curves: one definition

Given a data flow traversing a network, let $A \in \mathcal{F}_\uparrow$ be its *cumulative function* at some point in the network, *i.e.* $A(t)$ is the number of bits that have gone through this point until time t, with $A(0) = 0$. A function $\alpha \in \mathcal{F}$ is an *arrival curve* for A if $\forall\ s, t \in \mathbb{R}_+$, $0 \leq s \leq t$, we have $A(t) - A(s) \leq \alpha(t - s)$.

The set of all arrival curves for $A \in \mathcal{F}_\uparrow$ admits a minimum which remains an arrival curve: it is $\alpha = A \oslash A$ and it can be called the *canonical arrival curve* for A.

2.5 NC service curves: several definitions

In the literature, the definitions of service curves usually concern:

- *minimum service curves* which are lower bounds on the service provided in a system (useful for upper bounds on worst case performances).
- *single flow systems* $\mathcal{S}$, that is $\mathcal{S} \subseteq \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow$.

Note that in NC models with multiplexing, the aggregation of all the flows entering the system is often considered as a single flow to which the minimum service is applied.

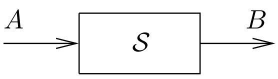

Fig. 1. A single flow input/output system.

For each type $\mathcal{T}$ of service curve, we define for any $\beta \in \mathcal{F}$ and for any input/output trajectory $(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow$ the conditions so that *β is a $\mathcal{T}$-service curve for (A, B)* (we also say that *(A, B) admits β as a $\mathcal{T}$-service curve*). We then define for all $\beta \in \mathcal{F}$, $\mathcal{S}_\mathcal{T}(\beta)$ the set of all trajectories admitting β as a $\mathcal{T}$-service curve. We say that a system $\mathcal{S}$ admits β as a $\mathcal{T}$-service curve if it is true for all its trajectories, *i.e.* $\mathcal{S} \subseteq \mathcal{S}_\mathcal{T}(\beta)$.

- ***Simple service curve:*** $\mathcal{S}_{simple}(\beta) = \{(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow \mid A \geq B \geq A * \beta\}$.
- ***Strict service curve (weak sense):*** $\mathcal{S}_{wstrict}(\beta) = \{(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow \mid A \geq B,$ and $\forall t \geq 0,\ B(t) \geq B(start(t)) + \beta(t - start(t))\}$.
- ***Strict service curve:*** $\mathcal{S}_{strict}(\beta) = \{(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow \mid A \geq B,$ and $\forall$ backlogged period $]s, t], B(t) \geq B(s) + \beta(t - s)\}$.
- ***Variable capacity node:*** $\mathcal{S}_{vcn}(\beta) = \{(A, B) \in \mathcal{F}_\uparrow \times \mathcal{F}_\uparrow \mid \exists C \in \mathcal{F}_\uparrow, \forall t \geq 0,\ B(t) = \inf_{0 \leq s \leq t}\left[A(s) + C(t) - C(s)\right]$ and $\forall 0 \leq s \leq t,\ C(t) - C(s) \geq \beta(t - s)\}$.

Two classical functions used as service curves are:

- *Pure delay* $T \in \mathbb{R}_+ \cup \{+\infty\}$: $\delta_T(t) = 0$ if $t \leq T$, $= +\infty$ otherwise.
- *Constant rate* $r \in \mathbb{R}_+ \cup \{+\infty\}$: $\lambda_r(t) = rt$ (if $r = +\infty$, we set $\lambda_r = \delta_0$).

Nothing prevents us from using some service curves which are not in $\mathcal{F}_\uparrow$, *e.g.* with negative values, decreasing parts or left-discontinuities. Nevertheless note that it is usually required that at least $\beta(0) \leq 0$, otherwise $\beta(0) > 0$ implies that $B(0) > A(0)$ and thus $\mathcal{S}_{simple}(\beta) = \mathcal{S}_{wstrict}(\beta) = \mathcal{S}_{strict}(\beta) = \mathcal{S}_{vcn}(\beta) = \emptyset$.

2.6 Remark on strict service curves

The definition of strict service curves presented here is the one used by Schmitt and Zdarsky (2006); Schmitt et al. (2006). Some papers do not choose exactly the same definition for strict service curves Bouillard et al. (2007, 2008): they replace the backlogged interval $]s, t]$ in the definition by $]s, t[$ (both definitions allow $B(s) = A(s)$, but this variant also allows $B(t) = A(t)$). For $\beta \in \mathcal{F}$, let us denote $\mathcal{S}'_{strict}(\beta)$ the set of trajectories satisfying this variant of our definition. How do those slightly different definitions compare? It is clear that $\forall \beta \in \mathcal{F}$, $\mathcal{S}'_{strict}(\beta) \subseteq \mathcal{S}_{strict}(\beta)$ and the equality holds if β is left-continuous. If β is not left-continuous we may have a strict inclusion as illustrated by Figure 2 where $A(t) = 1/2$ if $t > 0$ and $= 0$ if $t = 0$, $B(t) = \min(t/2, 1/2)$, $\beta(t) = \lfloor t \rfloor$, and $(A, B) \in \mathcal{S}_{strict}(\beta)$ but $(A, B) \notin \mathcal{S}'_{strict}(\beta)$ (see Bouillard et al. (2009) for details).

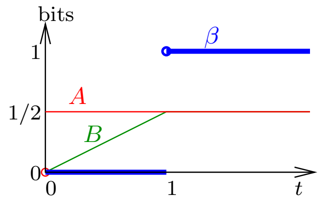

Fig. 2. Beware of the definition of strict service curves.

2.7 Remark on variable capacity nodes

Lemma 1. Let $A, C \in \mathcal{F}_\uparrow$ and $B \in \mathcal{F}$ such that $\forall t \geq 0$, $B(t) = \inf_{0 \leq s \leq t}\left[A(s) + C(t) - C(s)\right]$. Then $B \in \mathcal{F}_\uparrow$ and $\forall t \geq 0$, $B(t) = A(start(t)) + C(t) - C(start(t))$.

The proof of that result is a bit technical and can be found in Bouillard et al. (2009).

3. COMPARISON OF NC SERVICE CURVES

3.1 Monotony

All the definitions from the literature share the same natural monotonic behavior about trajectories.

Proposition 1. (Monotony). For any type $\mathcal{T}$ of service curve in the literature, for all $\beta, \beta' \in \mathcal{F}$ (not necessarily in $\mathcal{F}_\uparrow$), if $\beta \leq \beta'$ then $\mathcal{S}_\mathcal{T}(\beta) \supseteq \mathcal{S}_\mathcal{T}(\beta')$.

Moreover, for variable capacity node, strict and weakly strict service curves, one can replace service curves by some of their closures.

Proposition 2. Let $\beta \in \mathcal{F}$, then $\mathcal{S}_{wstrict}(\beta) = \mathcal{S}_{wstrict}(\beta_\uparrow)$, $\mathcal{S}_{strict}(\beta) = \mathcal{S}_{strict}(\beta_\uparrow)$, $\mathcal{S}_{strict}(\beta) = \mathcal{S}_{strict}(\overline{\beta^*})$. $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{vcn}(\beta_\uparrow)$ and $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{vcn}(\overline{\beta^*})$.

The previous result is considered as folklore and the proof is rather simple.The next theorem is new.

Theorem 1. (Monotony refined). Let $\beta, \beta' \in \mathcal{F}$,

(1) $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \Leftarrow \beta \leq \beta'$.
(2) $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \not\Rightarrow \beta \leq \beta'$.
(3) $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \Rightarrow \beta_\uparrow \leq \beta'_\uparrow$.
(4) $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \not\Leftarrow \beta_\uparrow \leq \beta'_\uparrow$.

(5) $\mathcal{S}_{simple}(\beta_\uparrow) \supseteq \mathcal{S}_{simple}(\beta'_\uparrow)$ if and only if $\beta_\uparrow \le \beta'_\uparrow$.
(6) $\mathcal{S}_{wstrict}(\beta) \supseteq \mathcal{S}_{wstrict}(\beta')$ if and only if $\beta_\uparrow \le \beta'_\uparrow$.
(7) $\mathcal{S}_{strict}(\beta) \supseteq \mathcal{S}_{strict}(\beta')$ if and only if $(\beta_\uparrow)^{\overline{*}} \le (\beta'_\uparrow)^{\overline{*}}$.
(8) $\mathcal{S}_{vcn}(\beta) \supseteq \mathcal{S}_{vcn}(\beta')$ if and only if $(\beta_\uparrow)^{\overline{*}} \le (\beta'_\uparrow)^{\overline{*}}$.

Proof. (1) Proposition 1.

(2) Take $\beta'(t) = 0$ if $t = 0$ or $t \in]1,2]$ and $\beta'(t) = +\infty$ otherwise and $\beta(t) = 0$ if $t \in [0,1]$ or $t \in]2,+\infty[$ and $\beta(t) = +\infty$ otherwise. Then, $\beta \not\le \beta'$ but, $\forall A \in \mathcal{F}_\uparrow$, $(A * \beta')_\uparrow(t) = A(t)$ if $t \in [0,1]$, $= A(1)$ if $t \in]1,2]$ and $= \max(A(1), A(t-2))$ otherwise, whereas $(A * \beta)_\uparrow(t) = A(0) = 0$ if $t \in [0,1]$, $= A(t-1)$ if $t \in]1,2]$ and $= A(1)$ otherwise. Then, $(A * \beta')_\uparrow \ge (A * \beta)_\uparrow$ and $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta')$. Figure 3 illustrates this construction.

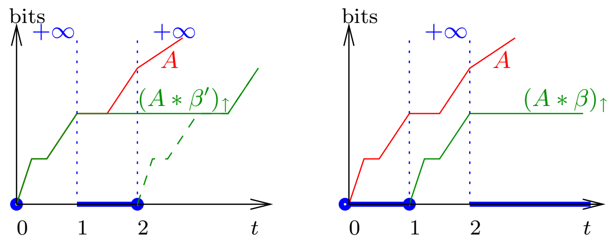

Fig. 3. $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \not\Rightarrow \beta \le \beta'$.

(3) $(\delta_0, \beta'_\uparrow) \in \mathcal{S}_{simple}(\beta')$, so $(\delta_0, \beta'_\uparrow) \in \mathcal{S}_{simple}(\beta)$ and then $\beta'_\uparrow \ge (\delta_0 * \beta)_\uparrow = \beta_\uparrow$.

(4) Take $\beta'(t) = 0$ if $t = 0$ or $t \in]1,+\infty[$ and $\beta'(t) = +\infty$ otherwise, and $\beta = \beta'_\uparrow = \delta_0$. We have $\beta_\uparrow \le \beta'_\uparrow$ and $\mathcal{S}_{simple}(\beta) = \{(A, A) \mid A \in \mathcal{F}_\uparrow\}$, but $\forall A \in \mathcal{F}_\uparrow$, $(A, A(\min(.,1)) \in \mathcal{S}_{simple}(\beta')$. Figure 4 illustrates this construction.

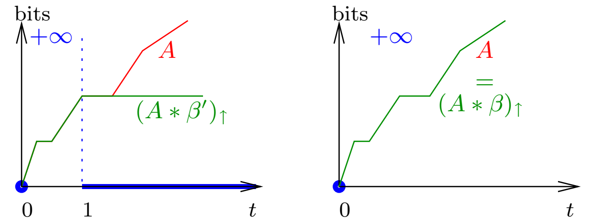

Fig. 4. $\mathcal{S}_{simple}(\beta) \supseteq \mathcal{S}_{simple}(\beta') \not\Leftarrow \beta_\uparrow \le \beta'_\uparrow$.

(5) $\Rightarrow$: idem (3); $\Leftarrow$: Proposition 1.

(6-8) Proposition 2 and $\Rightarrow$: idem (3); $\Leftarrow$: Proposition 1. $\square$

3.2 Families of service curves

We now discuss the fact that a system may admit several service curves of one or several types. Let $(\beta_i)_{i \in I}$ be a (possibly infinite) family of functions from $\mathcal{F}$, let $\mathcal{T}$ be a particular type of service curve. The fact that a system $\mathcal{S}$ admits all the functions β_i as $\mathcal{T}$-service curve can be also written $\mathcal{S} \subseteq \bigcap_{i \in I} \mathcal{S}_{\mathcal{T}}(\beta_i)$.

Theorem 2. (Families of curves). (1) Let I and J be finite sets and $(\beta_i)_{i \in I}$ and $(\beta'_j)_{j \in J}$ be two families in $\mathcal{F}_\uparrow$. Then, $\bigcap_{i \in I} \mathcal{S}_{simple}(\beta_i) = \bigcap_{j \in J} \mathcal{S}_{simple}(\beta'_j)$ iff $\{\beta \in \mathcal{F} \mid \exists i \in I,\ \beta \le \beta_i\} = \{\beta \in \mathcal{F} \mid \exists j \in J,\ \beta \le \beta'_j\}$.
(2) $\bigcap_{i \in I} \mathcal{S}_{wstrict}(\beta_i) = \mathcal{S}_{wstrict}((\sup_{i \in I} \beta_i)_\uparrow)$.
(3) $\bigcap_{i \in I} \mathcal{S}_{strict}(\beta_i) = \mathcal{S}_{strict}((\sup_{i \in I} \beta_i)^{\overline{*}})$.
(4) $\bigcap_{i \in I} \mathcal{S}_{vcn}(\beta_i) = \mathcal{S}_{vcn}((\sup_{i \in I} \beta_i)^{\overline{*}})$.

Proof. (1) First set $I = \{0, \dots, k\}$ and consider $\cap_{i \in I} \mathcal{S}_{simple}(\beta_i)$ such that functions β_i are pairwise non-comparable. Then, there exists $t_1, \dots, t_k$ such that $\forall i \in I \setminus \{0\}$, $\beta_0(t_i) > \beta_i(t_i)$ and one can assume without loss of generality that $0 < t_1 \le \dots \le t_k$. Set

$$A(t) = \begin{cases} 0 & \text{if } t = 0, \\ \beta_0(t_k) - \beta_0(t_i) & \text{if } t_i \le t_k - t < t_{i+1}, \\ +\infty & \text{if } t > t_{k-1}. \end{cases}$$

Now, for all $i \in I$, $A * \beta_i(t_k) = \inf_j \beta_i(t_j) + A(t_k - t_j) = \inf_j \beta_i(t_j) + \beta_0(t_k) - \beta_0(t_j)$. Then, $A * \beta_0(t_k) = \beta_0(t_k)$ and $\forall i \in \{1, \dots, k\}$, $A * \beta_i(t_k) \le \beta_i(t_i) + \beta_0(t_k) - \beta_0(t_i) < \beta_0(t_k)$. Then, $\cap_{i \in I} \mathcal{S}_{simple}(\beta_i) \subsetneq \cap_{i \in I/\{0\}} \mathcal{S}_{simple}(\beta_i)$. Figure 5 illustrates this.

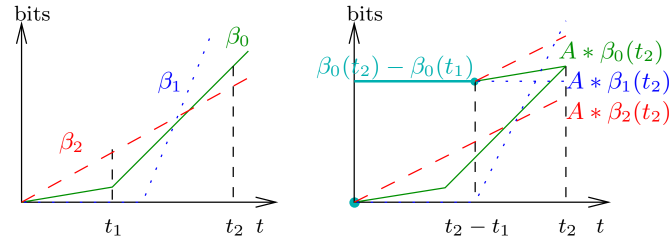

Fig. 5. Non-inclusion of families of curves.

Come back to $\cap_{i \in I} \mathcal{S}_{simple}(\beta_i) = \cap_{j \in J} \mathcal{S}_{simple}(\beta'_j)$ where the β_i's are pairwise non comparable and so are the β'_j's. We also have for all $j \in J$, $\cap_{i \in I} \mathcal{S}_{simple}(\beta_i) \bigcap \mathcal{S}_{simple}(\beta'_j) = \cap_{i \in I} \mathcal{S}_{simple}(\beta_i)$. Then, by the contraposition of the previous paragraph, $\exists i \in I$, such that $\beta'_j \le \beta_i$. By symmetry, $\forall i \in I$, $\exists j$ such that $\beta_i \le \beta'_j$. As functions in I and J are not two-by-two comparable, then this means that $\forall i \in I$, $\exists j \in J$ such that $\beta_i = \beta'_j$ and the symmetric.

(2) $\forall i \in I$, $\mathcal{S}_{wstrict}(\beta_i) \supseteq \mathcal{S}_{wstrict}(\sup_{i \in I} \beta_i)$, thus $\bigcap_{i \in I} \mathcal{S}_{wstrict}(\beta_i) \supseteq \mathcal{S}_{wstrict}(\sup_{i \in I} \beta_i)$. Let $(A, B) \in \bigcap_{i \in I} \mathcal{S}_{wstrict}(\beta_i)$. Then, $\forall t \in \mathbb{R}_+$, $\forall i \in I$, $B(t) \ge A(start(t)) + \beta_i(t - start(t))$, so $B(t) \ge A(start(t)) + \sup_{i \in I} \beta_i(t - start(t))$ and $(A, B) \in \mathcal{S}_{wstrict}(\sup_{i \in I} \beta_i) = \mathcal{S}_{wstrict}((\sup_{i \in I} \beta_i)_\uparrow)$.

(3) Idem 2., except: Let $(A, B) \in \bigcap_{i \in I} \mathcal{S}_{strict}(\beta_i)$. Then, $\forall s < t \in \mathbb{R}_+$, $\forall i \in I$, $B(t) \ge B(s) + \beta_i(t - s)$, so $B(t) \ge B(s) + \sup_{i \in I} \beta_i(t - s)$ and $(A, B) \in \mathcal{S}_{wstrict}(\sup_{i \in I} \beta_i) = \mathcal{S}_{wstrict}((\sup_{i \in I} \beta_i)^{\overline{*}})$.

(4) Idem 3., replacing B by C. $\square$

3.3 Hierarchy

The next hierarchy between the different notions of service curves is often considered as folklore, but the cases of equality have never been investigated (*e.g.* $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{strict}(\beta)$ is claimed with no assumption on β in Le Boudec and Thiran (2001) and Wandeler (2006)).

Theorem 3. (Hierarchy). For all $\beta \in \mathcal{F}$, we have the following inclusions:

$$\mathcal{S}_{vcn}(\beta) \subseteq \mathcal{S}_{strict}(\beta) \subseteq \mathcal{S}_{wstrict}(\beta) \subseteq \mathcal{S}_{simple}(\beta).$$

The equalities require:

- $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{strict}(\beta)$ iff $\beta \oslash \beta$ has only finite values.
- $\mathcal{S}_{strict}(\beta) = \mathcal{S}_{wstrict}(\beta)$ iff $\beta_\uparrow = \delta_T$, $T \in \mathbb{R}_+ \cup \{+\infty\}$.
- $\mathcal{S}_{wstrict}(\beta) = \mathcal{S}_{simple}(\beta)$ iff $\beta_\uparrow = \delta_0$ or 0.

Proof. We suppose that $\beta(0) \leq 0$, otherwise all the sets are empty.

The inclusion $\mathcal{S}_{strict}(\beta) \subseteq \mathcal{S}_{wstrict}(\beta)$ is clear, since for all $t \in \mathbb{R}_+$, either $]start(t), t]$ is a backlogged period and thus $B(t) - B(start(t)) \geq \beta(t - start(t))$, or $]start(t), t[$ is a backlogged period and $B(t) = A(t)$ but then $start(t) = t$ and $B(t) - B(start(t)) = 0 \geq \beta(t - start(t)) = 0$.

The inclusion $\mathcal{S}_{wstrict}(\beta) \subseteq \mathcal{S}_{simple}(\beta)$ comes from the remark that if $(A, B) \in \mathcal{S}_{wstrict}(\beta)$, then $B(t) \geq B(start(t)) + \beta(t - start(t)) = A(start(t)) + \beta(t - start(t)) \geq \inf_{0\leq s\leq t}(A(s) + \beta(t - s))$.

Now let us show that $\mathcal{S}_{vcn}(\beta) \subseteq \mathcal{S}_{strict}(\beta)$. Let $(A, B) \in \mathcal{S}_{vcn}(\beta)$, there exists $C \in \mathcal{F}_\uparrow$ such that $\forall t \geq 0$, $B(t) = \inf_{0\leq s\leq t} \left[A(s)+C(t)-C(s)\right]$ and $\forall 0 \leq s \leq t, C(t)-C(s) \geq \beta(t-s)$. Consider a backlogged period $]s, t]$ for (A, B), then by definition $start(s) = start(t) = p$. From Lemma 1, we have $B(t) = B(p)+C(t)-C(p)$ and $B(s) = B(p)+C(s)-C(p)$. Thus $B(t) - B(s) = C(t) - C(s) \geq \beta(t - s)$ and we have proved that β is a strict service curve for (A, B).

Conditions of equality:

- $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{strict}(\beta)$: As $\mathcal{S}_{vcn}(\beta) = \mathcal{S}_{vcn}(\overline{\beta^*})$ and $\mathcal{S}_{strict}(\beta) = \mathcal{S}_{strict}(\overline{\beta^*})$, one can consider that β is super-additive. Then, one has $\beta(t - s) \leq \beta \oslash \beta(t) - \beta \oslash \beta(s)$ and $\sum_{i=1}^n \beta \oslash \beta(t_i) + \beta(t_0) \geq \beta(\sum_{i=0}^n t_i)$. Suppose that $\beta\oslash\beta < \infty$. Let $(A, B) \in \mathcal{S}_{strict}(\beta)$. This trajectory can be seen as a succession of "idle" periods (intervals of length ≥ 0 during which the backlog is 0) and backlogged period of positive length. Let $(t_i)_{i\geq0}$ be the increasing sequence of the start of those periods (t_{2i} are the start of idle periods and t_{21+1} are the start of backlogged periods). Set $\ell_i = t_{i+1} - t_i$ and

$$C(t) = \begin{cases} A(t_i) + \beta \oslash \beta(t_i) + \sum_{j=0}^{i} \beta \oslash \beta(\ell_j) & \text{if } t = t_i \\ A(t) + \beta \oslash \beta(t) + \sum_{j=0}^{2i} \beta \oslash \beta(\ell_j) & \text{if } t_{2i} < t < t_{2i+1} \\ B(t) - B(t_{2i+1}) - C(t_{2i+1}) & \text{if } t_{2i-1} < t < t_{2i} \end{cases}$$

Then, $\forall s < t$,

- if $t_{2i-1} \leq s < t < t_{2i}$, then $C(t) - C(s) = B(t) - B(s) \geq \beta(t - s)$;
- if $t_{2i} \leq s < t < t_{2i+1}$ then $C(t) - C(s) = \beta \oslash \beta(t) - \beta \oslash \beta(s) + A(t) - A(s) \geq \beta(t - s)$;
- if $t_j \leq s < t_{j+1} \leq t_{2i+1} \leq t \leq t_{2i+2}$, then $C(t) - C(s) = C(t) - C(t_{2i+1}) + C(t_{2i+1}) - C(s) \geq \beta(t - t_{2i+1}) + \sum_{k=j+1}^{2i+1} \beta \oslash \beta(\ell_j) \geq \beta(t - t_{2i+1} + t_{2i+1} - t_j) \geq \beta(t - s)$;
- if $t_j \leq s < t_{j+1} \leq t_{2i} \leq t \leq t_{2i+1}$, then $C(t) - C(s) \geq C(t) - C(t_{2i}) + C(t_{2i}) - C(s) \geq \beta(t - t_{2i}) + \sum_{k=j+1}^{2i} \beta \oslash \beta(\ell_j) \geq \beta(t - t_{2i} + t_{2i} - t_j) \geq \beta(t - s)$;

Moreover, $B(t) = A(start(t)) + C(t) - C(start(t))$, $\forall s \in [start(t), t]$, $B(t) \leq A(s) + C(t) - C(s)$, and $\forall s < start(t)$, $C(s) - A(s) \geq C(start(t)) - A(start(t))$ so $B(t) \leq A(s) + C(t) - C(s)$. Then, $(A, B) \in \mathcal{S}_{vcn}(\beta)$.

On the other hand, if $\beta \oslash \beta(t_0) = \infty$, take $(\delta_{t_0}, \beta * \delta_{t_0}) \in \mathcal{S}_{strict}(\beta)$. In order to make the computations possible, $C(t_0)$ must be finite. But, one must have $C(t + t_0) - C(t_0) = \beta(t)$ and $C(t + t_0) - C(0) \geq \beta(t + t_0)$. Then, $\forall t \in \mathbb{R}_+$, $C(t_0) = C(t_0) - C(0) \geq \beta(t + t_0) - \beta(t)$ and $C(t_0) \geq \beta \oslash \beta(t_0) = +\infty$.

- $\mathcal{S}_{strict}(\beta) = \mathcal{S}_{wstrict}(\beta)$: if $\beta_\uparrow = \delta_T$, $T \in \mathbb{R}_+ \cup \{+\infty\}$, then let $(A, B) \in \mathcal{S}_{wstrict}(\beta)$. Let $s < t$ in the same backlogged period. Then, $t - s < T$ and $B(t) \geq B(s) = B(s) + \beta_\uparrow(t - s)$. Then $(A, B) \in \mathcal{S}_{strict}(\beta)$.

Otherwise $\beta_\uparrow$ is not a delay: let $A = \delta_0$. There exists $s > 0$ such that $0 < \beta_\uparrow(s) < \infty$. Let $t_0 = \sup\{t \geq 0 \mid \beta_\uparrow(t) = 0\} \leq s$. Define $B(t) = \beta_\uparrow(t)$ in $t < (s - t_0)/2$ or $t > s$, $= \beta_\uparrow(s)$ if $t \in [(s-t_0)/2, s]$. We have $(A, B) \in \mathcal{S}_{wstrict}(\beta)\setminus \mathcal{S}_{strict}(\beta)$. Indeed, $B(s) - B((s - t_0)/2) = 0 < \beta_\uparrow((s + t_0)/2)$.

- $\mathcal{S}_{wstrict}(\beta) = \mathcal{S}_{simple}(\beta)$: $\mathcal{S}_{simple}(0) = \mathcal{S}_{wstrict}(0) = \{(A, B) \in \mathcal{F}_\uparrow^2 \mid A \geq B\}$; $\mathcal{S}_{simple}(\delta_0) = \mathcal{S}_{wstrict}(\delta_0) = \{(A, A) \mid A \in \mathcal{F}_\uparrow^2\}$. If $\beta \notin \{0, \delta_0\}$, we consider three cases: let $t_0 = \inf\{t \geq 0 \mid \beta(t) > 0\}$. If $t_0 > 0$, take s such that $\beta(s) > 0$ and $A(t) = (\beta(s)/s)\,t$. As $\forall t > 0$, $A*\beta(t) \leq A(t-t_0) < A(t)$, the trajectory $(A, A * \beta) \in \mathcal{S}_{simple}(\beta)$ has only one infinite backlogged period, but $A(0) + \beta(s) = A(s)$, then the first backlogged period of a server offering a weakly strict service curve must be of length at most s: $(A, A * \beta) \notin \mathcal{S}_{wstrict}(\beta)$.

If $t_0 = 0$, if there exists ρ such that $A(t) = \rho t$ and there exists $s > 0$ with $\forall t \in [0, s]$, $A(t) > \beta(t)$ and $\beta^{-1}(A(s)) > s$. Let B defined by $B(t) = \beta(t)$ if $t < s$, $= A(s)$ if $t = s$, $\max(\beta(t), A(s))$ otherwise. We have $A \geq B \geq \beta \geq A * \beta$, then $(A, B) \in \mathcal{S}_{simple}(\beta)$, but there is a backlogged period beginning at s, and nothing is served between s and $\beta^{-1}(A(s))$. Then $(A, B) \notin \mathcal{S}_{wstrict}(\beta)$.

Otherwise, $t_0 = 0$ and there exist $\rho, u \in \mathbb{R}_+$, $A(t) = \rho t$ such that $A \not\geq \beta$ but $\forall t \in]0, u[$, $A(t) > \beta(t)$. Then, $A*\beta \leq \min(A, \beta) < A$, but $(A, B) \in \mathcal{S}_{wstrict}(\beta) \Rightarrow B = A$. Then $(A, A * \beta) \notin \mathcal{S}_{wstrict}(\beta)$. □

The equality cases between *strict* and *vcn* curves is not very restrictive, as most of the time, service curves have a finite asymptotic growing rate, and thus have finite auto-deconvolution. An important family of service curves where the equality does not hold is the pure delays $\delta_T, T \in \mathbb{R}_+$.

The other equality cases seem very restrictive. Nevertheless, if one considers the equality of the output process when the service is exact (that is when the last inequality are replaced by an equality in the definition of $\mathcal{S}_\mathcal{T}(\beta)$, $\mathcal{T} \in \{simple, wstrict, strict\}$) for any arrival process, equality cases are more frequent: for simple and weakly strict service curves, the equality holds for any $\lambda_r, r \in \mathbb{R}_+\cup \{+\infty\}$ and for weakly strict and strict service curves, the equality holds for any super-additive function.

Finally, the following theorem states that in the case where the equality does not hold, there is no chance to express one type of service curve as a combination of service curves of another type.

Theorem 4. (No translation with families). Let $\mathcal{T}$ and $\mathcal{T}'$ be two different types of service curves among *vcn*, *simple*, *strict (weak)*, *strict*. Then

$$\nexists(\beta_i)_{i\in I} \in \mathcal{F}^I,\ (\beta'_j)_{j\in J} \in \mathcal{F}^J,\ \bigcap_{i\in I} \mathcal{S}_{\mathcal{T}}(\beta_i) = \bigcap_{j\in J} \mathcal{S}_{\mathcal{T}'}(\beta'_j),$$

except for the equality cases defined in Theorem 3.

Proof. Translation between variable capacity node and strict, and between weakly strict and strict service curves: one only need to consider the case where I and J are singletons, due to Theorem 2.With $(\mathcal{T},\mathcal{T}') = \{(strict, vcn),(wstrict, strict)\}$, Suppose that $\mathcal{S}_{\mathcal{T}}(\beta) = \mathcal{S}_{\mathcal{T}}(\beta')$. Then, from Theorem 3, $\mathcal{S}_{\mathcal{T}'}(\beta) \subseteq \mathcal{S}_{\mathcal{T}'}(\beta')$ and from Theorem 1, $\beta' \leq \beta$. Now, $(\delta_0, \beta') \in \mathcal{S}_{\mathcal{T}'}(\beta') = \mathcal{S}_{\mathcal{T}}(\beta)$, then $\beta' \geq \beta$. In conclusion, one must have $\beta = \beta'$.

Between simple and weakly strict service curves: consider $(\beta_i)_{i\in I}$ and β', and suppose that $\cap_{i\in I}\mathcal{S}_{simple}(\beta_i) = \mathcal{S}_{wstrict}(\beta')$. From Proposition 1 and Theorem 3, one has: $\mathcal{S}_{wstrict}(\sup_{i\in I}\beta_i) \subseteq \mathcal{S}_{simple}(\sup_{i\in I}\beta_i) \subseteq \cap_{i\in I}\mathcal{S}_{simple}(\beta_i) = \mathcal{S}_{wstrict}(\beta')$. Then, $\beta' \leq (\sup_{i\in I}\beta_i)_\uparrow$. On the other hand, $(\delta_0, \beta') \in \mathcal{S}_{wstrict}(\beta') = \cap_{i\in I}\mathcal{S}_{simple}(\beta_i)$, then $\beta' \geq \sup_{i\in I}(\delta_0 * \beta_i)_\uparrow = (\sup_{i\in I}\beta_i)_\uparrow$ and $\beta' = (\sup_{i\in I}\beta_i)_\uparrow$. Moreover, $\mathcal{S}_{wstrict}((\sup_{i\in I}\beta_i)_\uparrow) = \mathcal{S}_{wstrict}(\sup_{i\in I}\beta_i) \subseteq \mathcal{S}_{simple}(\sup_{i\in I}\beta_i) \subseteq \cap_{i\in I}\mathcal{S}_{simple}(\beta_i) = \mathcal{S}_{strict}(\beta')$. Then, all inequalities are in fact equalities and then one must have $\mathcal{S}_{simple}(\sup_{i\in I}\beta_i) = \cap_{i\in I}\mathcal{S}_{simple}(\beta_i)$ and $\beta' = \sup_{i\in I}\beta_i = \delta_0$ or 0. □

4. CONCLUSION

We hope that this study casts new light on the main service curve definitions used in envelope-based models. We have shown that there exist strong gaps in the hierarchy, like the impossibility to express a strict service as a family of simple services, even if we allow infinite families. As a consequence, providing an unified framework to achieve tight performance analyses for all the corresponding models, may turn out to be very difficult. As a matter of fact, while some tight analyses of simple curve models exclusively rely on $(\min, +)$ algebra (window flow control or multimedia traffic smoothing in Chang (2000); Le Boudec and Thiran (2001)), some tight analyses of strict curve models completely avoid any reference to $(\min, +)$ (worst case performances in some acyclic networks in Bouillard et al. (2010)).

Nevertheless further research is required about the notion of service curves in order to fill those gaps or take into account additional features.

One can investigate the possibility to mix different definitions, as suggested with *adaptative service curves* (mixing strict service and simple service) in Le Boudec and Thiran (2001), or try to introduce a full range of intermediate definitions between the classical definitions exposed in the article.

One should also include maximum service curves in the study, like *maximum simple service curves* defined in Network Calculus (Le Boudec and Thiran, 2001, Chapter 1, pages 42-48) or *upper service curves* intensively used in Real-Time Calculus (Wandeler (2006)). Their use introduces new connections between the different notions. For instance, a constant rate server can be seen as a server with a minimum and a maximum simple service curves both equal to $\lambda_r(t) = rt$. In this case, it is easy to check that this server admits λ_r as a strict service curve. In contrast a system defined exclusively by minimum service curves can only ensure either λ_0 or δ_0 as strict services.

Beyond the expressiveness issue, other important aspects include the robustness of the definitions with regard to flow multiplexing. For instance, a server with fixed priorities offering a simple minimum service, does not guarantee any non null simple service to the flow with lowest priority.

Ultimately such studies will be useful to adjust both the modeling power (to be relevant) and the computational complexity of analyses (to be effective), in Network Calculus and its extensions.

REFERENCES

Baccelli, F., Cohen, G., Olsder, G., and Quadrat, J. (1992). *Synchronization and linearity*. Wiley.

Bouillard, A., Gaujal, B., Lagrange, S., and Thierry, E. (2007). Optimal routing for end-to-end guarantees: the price of multiplexing. In *Proceedings of Valuetools'07*.

Bouillard, A., Gaujal, B., Lagrange, S., and Thierry, E. (2008). Optimal routing for end-to-end guarantees using network calculus. *Performance Evaluation*, 65(11-12), 883–906.

Bouillard, A., Jouhet, L., and Thierry, E. (2010). Tight performance bounds in the worst-case analysis of feed-forward networks. In *Proceedings of Infocom'2010*.

Bouillard, A., Jouhet, L., and Thierry, E. (2009). Service curves in Network Calculus: dos and don'ts. Research Report RR-7094, INRIA.

Chang, C.S. (2000). *Performance Guarantees in Communication Networks*. TNCS, Springer-Verlag.

Fidler, M. (2010). A survey of deterministic and stochastic service curve models in the network calculus. *IEEE Communications Surveys & Tutorials*, 12(1).

Le Boudec, J.Y. and Thiran, P. (2001). *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, volume LNCS 2050. Springer-Verlag, revised version 4, may 10, 2004 edition.

Schmitt, J.B. and Roedig, U. (2005). Sensor network calculus: A framework for worst case analysis. In *Proceedings of 1st International Conference on Distributed Computing in Sensor Systems*.

Schmitt, J.B. and Zdarsky, F.A. (2006). The disco network calculator: a toolbox for worst case analysis. In *Proceedings of Valuetools'2006*.

Schmitt, J.B., Zdarsky, F.A., and Martinovic, I. (2006). Performance bounds in feed-forward networks under blind multiplexing. Technical Report 349/06, University of Kaiserslautern, Germany.

Thiele, L., Chakraborty, S., and Naedele, M. (2000). Real-time calculus for scheduling hard real-time systems. In *Proceedings of ISCAS'2000*.

Wandeler, E. (2006). *Modular Performance Analysis and Interface-Based Design for Embedded Real-Time Systems*. Ph.D. thesis, ETH Zurich.