The Duality of Maximum Delay and Maximum Backlog

Shahrokh Valaee

Abstract—**In this paper, we will establish a duality between maximum delay and maximum backlog in a single queue network. Min-plus algebra will be employed to devise the duality. In min-plus algebra, the maximum backlog is represented by the scalar projection of accumulated input traffic curve onto accumulated output traffic curve. We will define the adjoint operator of a nondecreasing function and show that the maximum delay can be represented as the scalar projection of the adjoint of accumulated output traffic onto the adjoint of accumulated input traffic. Based on this observation, we will draw a duality between the maximum backlog and the maximum delay.**

We will also define the left- and the right-seminorms of a nondecreasing function and will use it to introduce the concept of a "matched" queue. A matched queue is a queue in which the accumulated input and output curves are adjoint. In a matched queue, the maximum delay and the maximum backlog are identical and are equal to the left-norm of input process.

I. Introduction

Worst-case quality-of-service (QoS) provisioning has been the focus of recent research in high-speed networking [1] — [7]. QoS is usually indicated by performance measures such as delay, delay jitter, traffic loss, backlog, minimum bandwidth and so on. The QoS parameters can be studied in average-case and/or in worst-case paradigms. In the average-case analysis of QoS parameters, the input traffic is usually modelled by a canonical probability distribution function and the behaviour of network is measured in terms of the percentiles of the QoS parameters. In the worst-case approach, the QoS parameters are usually quantified in the extreme case in which the input traffic is greedy and resource allocation is parsimonious. In fact, the worst-case analysis quantifies the maximum deviation from the normal network operation. In this approach, the maximum value of QoS parameters are studied.

In this paper, we study maximum delay and maximum backlog as the two indices of network performance. For simplicity, the whole network is modelled as a single queue [1] [3]. The delay of each packet is the difference between the time that the packet arrives at the queue and the time that it leaves the queue. Backlog indicates the amount of traffic awaiting service inside the queue. Both parameters are directly related to the available bandwidth and the activity of other sources — usually modelled as a crossing traffic.

We will use the concept of *service curve* [3] [8] [9]. The service curve of each traffic flow represents the amount of service guaranteed for that traffic. The service curve provisioning will provide a certain degree of isolation between the crossing traffic and the flow under investigation. Using this approach, the delay and backlog for each traffic can be respectively calculated as the horizontal and the vertical separation between the accumulated input traffic and the corresponding service curve.

In the present paper, we will show that there exists a duality between the maximum delay and the maximum backlog. We will use *min-plus algebra* to establish the duality [10] [11]. In min-plus algebra, we use *minimization* and *addition* as the two main operators. In fact, the arithmetical addition of the conventional algebra is replaced with minimization and the arithmetical multiplication is replaced with the addition. It is possible to show that the set of increasing functions along with the minimization and addition can generate a *dioid* [4]. In addition to the "idempotency", a dioid in min-plus algebra has the same characteristics as a *ring* in the conventional algebra.

The maximum backlog of each traffic flow in a queue with a given service curve is represented in terms of the *scalar projection* of the accumulated input traffic curve onto the service curve [12]. We will define an *adjoint* operator in the set of non-decreasing functions and then will show that the maximum delay can be represented as the scalar projection of the adjoint of service curve onto the adjoint of the accumulated input traffic. Using this property, we will show that the maximum delay of a given queue is identical to the maximum backlog of a dual queue in which the input is driven by the adjoint of service curve and the output is equal to the adjoint of the accumulated input traffic of the first queue.

Furthermore, we will define the concept of a

The authors is with the Edward S. Rogers Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada, M5S 3G4, Email: valaee@comm.utoronto.ca.

TABLE I

CORRESPONDING OPERATORS IN THE CONVENTIONAL AND MIN-PLUS ALGEBRA.

Conventional Algebra	Min-plus Algebra
$\min\{a,b\}$	$a \oplus b$
$a+b$	$a \otimes b$
$a-b$	$\frac{a}{b}$ (with ⊖)
ab	$a^b = b^a$
$\frac{a}{b}$	$\sqrt[b]{a}$
$\max\{a,b\}$	$a \vee b$

"matched" queue. We will show that a queue is matched to the input if the service curve is the adjoint of accumulated input traffic. For a matched queue, the delay and the backlog are identical and are equal to a *seminorm* — to be defined in the course of the paper — of the accumulated input traffic curve.

II. Min-plus Algebra

Let the set of all non-decreasing functions over $I\!R^+$ be denoted by

$$\mathcal{J} \triangleq \{a(t) \quad | \quad 0 \le a(t_1) \le a(t_2), \text{ for } 0 \le t_1 \le t_2, \text{ and } a(0)=0\}. \tag{1}$$

We will use *min-plus algebra* to handle the algebraic manipulation of elements of $\mathcal{J}$. In min-plus algebra, the arithmetical addition of conventional algebra is replaced by the point-wise minimization — denoted here by the notation $\oplus$ — and the arithmetical multiplication is replaced by the point-wise addition, represented by $\otimes$. In fact, for $a(t), b(t) \in \mathcal{J}$, we have

$$a(t) \oplus b(t) \triangleq \min\{a(t), b(t)\}, \tag{2}$$
$$a(t) \otimes b(t) \triangleq a(t) + b(t). \tag{3}$$

It is straightforward to show that the set $\mathcal{J}$ is closed under the $\oplus$ and $\otimes$ operations. Subtraction, multiplication, and division in the conventional algebra also have corresponding counterparts in min-plus algebra as represented in Table I.

Definition 1: A set $\mathcal{A}$ supplied with two inner operations $\oplus$ and $\otimes$ is a *commutative dioid* if the following axioms hold:

- *Axiom 1*: (Associativity) $\forall\, a(t), b(t), c(t) \in \mathcal{A}$,

$$[a(t) \oplus b(t)] \oplus c(t) = a(t) \oplus [b(t) \oplus c(t)], \tag{4}$$
$$[a(t) \otimes b(t)] \otimes c(t) = a(t) \otimes [b(t) \otimes c(t)]. \tag{5}$$

- *Axiom 2*: (Commutativity) $\forall\, a(t), b(t) \in \mathcal{A}$,

$$a(t) \oplus b(t) = b(t) \oplus a(t), \tag{6}$$
$$a(t) \otimes b(t) = b(t) \otimes a(t). \tag{7}$$

- *Axiom 3*: (Distributivity) $\forall\, a(t), b(t), c(t) \in \mathcal{A}$,

$$[a(t) \oplus b(t)] \otimes c(t) = [a(t) \otimes c(t)] \oplus [b(t) \otimes c(t)]. \tag{8}$$

- *Axiom 4*: (Null and Identity Elements) $\forall\, a(t) \in \mathcal{A}, \exists \epsilon(t), e(t) \in \mathcal{A}:$

$$a(t) \oplus \epsilon(t) = a(t), \tag{9}$$
$$a(t) \otimes e(t) = a(t). \tag{10}$$

- *Axiom 5*: (Absorbing Null Element) $\forall\, a(t) \in \mathcal{A}$,

$$a(t) \otimes \epsilon(t) = \epsilon(t). \tag{11}$$

- *Axiom 6*: (Idempotency of Addition) $\forall\, a(t) \in \mathcal{A}$,

$$a(t) \oplus a(t) = a(t). \tag{12}$$

The null and the identity elements in Axiom 4 are defined as $\epsilon(t) = \infty$ for $t > 0$, and $\epsilon(0) = 0$, and $e(t) = 0$ for $t \ge 0$, respectively.

Definition 2: A dioid is *complete* iff it is closed for infinite sums, and Axiom 3 extends to infinite sums.

Remark 1: It is possible to show that $(\mathcal{J}, \oplus, \otimes)$ is a complete *dioid* [10].

In a dioid, a partial order is defined as

$$a(t) \le b(t) \Leftrightarrow a(t) = a(t) \oplus b(t). \tag{13}$$

With this ordering, the dioid is an *inf-semilattice.* Defining the inner operation $\vee$ as

$$a(t) \vee b(t) \triangleq \max\{a(t), b(t)\}, \tag{14}$$

the semilattice becomes a complete *lattice* [13] [11].

Definition 3: The *scalar projection* of a function $a(t)$ onto $b(t)$ was defined in [12] as

$$\langle a(t), b(t)\rangle = \sup_{t\ge 0}[a(t) - b(t)] \vee 0. \tag{15}$$

Using the operator for subtraction, one can also represent (15) as

$$\langle a(t), b(t)\rangle = \sup_{t\ge 0} \frac{a(t)}{b(t)} \vee 0. \tag{16}$$

Definition 4: For $a(t) \in \mathcal{J}$, we define the *adjoint mapping* as

$$a^*(t) \triangleq t + \inf\{d \,|\, t \le a(t+d)\}. \tag{17}$$

The reverse mapping is defined as

$$a(t) \triangleq t - \inf\{d \,|\, a^*(t-d) \le t\}. \tag{18}$$

Proposition 1: For each $a(t) \in \mathcal{J}$, there exists a unique $a^*(t) \in \mathcal{J}$.

Lemma 1: For $a(t) \in \mathcal{J}$ with the adjoint mapping $a^*(t)$ we have:

(i) $a(t) \in \mathcal{J} \Longrightarrow a^*(t) \in \mathcal{J}$;

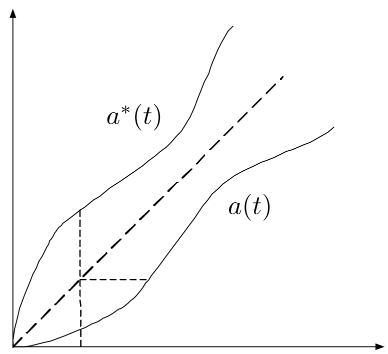

Fig. 1. The adjoint mapping for a function $a(t) \in \mathcal{J}$.

(ii) $a^{**}(t) = a(t)$;
(iii) If for all $t \in I\!R^+$, $a(t) \leq b(t)$, then $a^*(t) \geq b^*(t)$;
(iv) If $c(t) = a(t) \oplus b(t)$, then $c^*(t) = a^*(t) \vee b^*(t)$;
(v) If $c(t) = a(t) \otimes b(t)$, then $c^*(t) \leq a^*(t) \oplus b^*(t)$;
(vi) If $b(t) = \phi a(t)$, then $b^*(t) = a^*(\frac{t}{\phi})$ for any scalar $\phi \in I\!R^+$;
(vii) If $a(t)$ in convex (respectively concave), then $a^*(t)$ will be concave (respectively convex);
$(viii)$ If $a(t) \leq b(t)$, for all $t \geq 0$, then $a^*(t) - b^*(t) = \inf\{d \,|\, t \leq a(b^*(t) + d)\} = \inf\{d \,|\, b(a^*(t) - d) \leq t\}$;
(ix) $a^*(a(t)) = t$ and $a(a^*(t)) = t$;
(x) $\langle t, a(t) \rangle = \langle a^*(t), t \rangle$.

Using Lemma 1-(iv) and the fact that the adjoint mapping is *bijective* on $\mathcal{J}$, we can conclude that the sup-semilattice $(\mathcal{J}, \vee)$ is *conjugate* to the inf-semilattice $(\mathcal{J}, \oplus)$ and therefore that the structure $(\mathcal{J}, \oplus, \vee)$ is self-conjugate. Furthermore, if we define the operator $\otimes'$ as

$$a^*(t) \otimes' b^*(t) = \inf\{\tau \,|\, a(\tau) \otimes b(\tau) \geq t\}, \tag{19}$$

then

$$\Big(a(t) \otimes b(t)\Big)^* = a^*(t) \otimes' b^*(t) \tag{20}$$

and we can state that the dioid $(\mathcal{J}, \vee, \otimes')$ is conjugate to the dioid $(\mathcal{J}, \oplus, \otimes)$.

The adjoint mapping of a function is, in fact, the reflection of that function about the line $a(t) = t$. Fig. 1 illustrates an example of a function $a(t) \in \mathcal{J}$ along with its adjoint mapping. We call this operator "adjoint" mapping for the items (ix) and (x) in Lemma 1; similar properties hold between an operator and its adjoint in the conventional algebra. See also [6], for a similar definition in terms of pseudo-inverse.

Definition 5: For $a(t) \in \mathcal{J}$, we define the *left-seminorm* as

$$\|a(t)\|_\ell \triangleq \Big[\langle a(t), a^*(t) \rangle\Big]^{\frac{1}{2}} \tag{21}$$

and the *right-seminorm* as

$$\|a(t)\|_r \triangleq \Big[\langle a^*(t), a(t) \rangle\Big]^{\frac{1}{2}} \tag{22}$$

Note that the notations in the definition of the left and right seminorms are those of min-plus algebra. In fact, the left-seminorm (respectively right-seminorm) of a function $a(t)$ is half the maximum vertical distance between $a(t)$ (respectively $a^*(t)$) and its adjoint $a^*(t)$ (respectively $a(t)$).

Proposition 2: The left-seminorm satisfies the following properties:
(i) $\|a(t)\|_\ell \geq 0$;
(ii) $\|a(t) \oplus b(t)\|_\ell \leq \|a(t)\|_\ell \oplus \|b(t)\|_\ell$
Similar properties hold for the right-seminorm.

Proposition 3: The left and the right seminorm of a function $a(t) \in \mathcal{J}$ will satisfy:

$$\begin{aligned} \|a(t)\|_\ell &\neq \|a(t)\|_r, && (23)\\ \|a(t)\|_\ell &= \|a^*(t)\|_r, && (24)\\ \|a(t)\|_r &= \|a^*(t)\|_\ell. && (25) \end{aligned}$$

Lemma 2: If $\lim_{t\to\infty} \frac{a(t)}{t} > 1$, then the left-seminorm is infinite. If $\lim_{t\to\infty} \frac{a(t)}{t} < 1$, then the right-seminorm is infinite. If $\lim_{t\to\infty} \frac{a(t)}{t} = 1$, then both left-seminorm and right-seminorm are finite.

Remark 2: If $a(t) \leq t$ for all $t \geq 0$, then $\|a(t)\|_\ell = 0$ and if $a(t) \geq t$ for all $t \geq 0$, then $\|a(t)\|_r = 0$.

III. Delay-backlog Duality

The accumulated arrival process in a queue over the interval $[0, t]$ is represented by $A(t) = \int_0^t r(s)\, ds$ where the nonnegative function $r(t)$ is the normalized instantaneous rate of the input traffic and we have $r(t) = 0$ for $t < 0$. With this assumption, we have $A(t) \in \mathcal{J}$.

In [4], it has been shown that the maximum delay and the maximum backlog can be obtained using the *upper envelope* of the accumulated input traffic. The upper envelope is a subadditive closure of input traffic and reflects the worst-case behaviour of input flow. In the sequel, we will assume that $A(t)$ is, in fact, the upper envelope of accumulated input traffic.

Assume that the input $A(t)$ is served by a node with the service curve $S(t)$. The cumulative output traffic $B(t)$ will satisfy

$$B(t) \geq \min_{0 \leq s \leq t} \{A(s) + S(t - s)\} \tag{26}$$

In terms of min-plus algebra, (26) can be represented by

$$B(t) \geq A(t) * S(t) \tag{27}$$

where $*$ denotes the convolution operator defined as

$$a(t) * b(t) = \bigoplus_{0 \leq s \leq t} \{a(s) \otimes b(t - s)\}. \tag{28}$$

Using this definition, a single FIFO queue is contemplated as a linear filter in min-plus algebra.

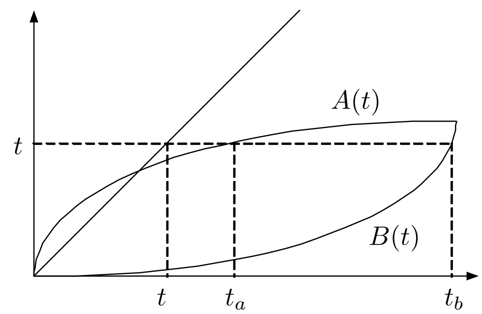

Fig. 2. The delay is defined as the horizontal distance between the input traffic curve and the output traffic curve.

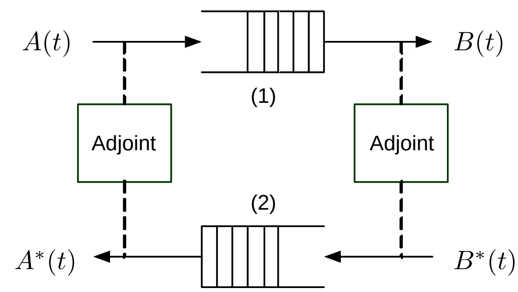

Fig. 3. Two queues illustrating the duality of maximum delay and maximum backlog.

Backlog $Q(t)$ at time t is defined as the vertical distance between the input traffic $A(t)$ and the output traffic $B(t)$, that is

$$Q(t) \triangleq A(t) - B(t). \tag{29}$$

Using the subtraction operator of min-plus algebra, backlog can also be represented as

$$Q(t) = \frac{A(t)}{B(t)} \tag{30}$$

Similarly, delay is defined as the horizontal distance between these two curves. We formulate the delay as

$$D(t) \triangleq t_b - t_a \tag{31}$$

where

$$\begin{aligned} t_b &= \inf\{\tau | B(\tau) \geq t\}, && (32)\\ t_a &= \sup\{\tau | A(\tau) \leq t\}. && (33) \end{aligned}$$

Refer to Fig. 2 for an illustration. Note that t_a and t_b are respectively equal to the adjoint of $A(t)$ and $B(t)$. Using this equality and the subtraction operator of min-plus algebra, we have

$$D(t) = \frac{B^*(t)}{A^*(t)} \tag{34}$$

Using (15) and Definition 4, it is possible to prove that the maximum backlog and the maximum delay are given by [4] [14]

$$\begin{aligned} Q_{_M} &= \langle\, A(t)\,,\, S(t)\rangle = \sup_{t\geq 0} \frac{A(t)}{S(t)} && (35)\\ D_{_M} &= \langle S^*(t)\,,\, A^*(t)\rangle = \sup_{t\geq 0} \frac{S^*(t)}{A^*(t)} && (36) \end{aligned}$$

Note that (35) and (36) illustrate a duality between the maximum backlog and the maximum delay. The maximum backlog is the scalar projection of the input traffic onto the output traffic. Similarly, the maximum delay can be formulated as the projection of the adjoint of accumulated output traffic onto the adjoint of accumulated input traffic. The duality can be employed to prescribe the application of similar approaches to the management of these QoS parameters. Fig. 3 illustrates two dual queues. The maximum delay of queue (1) is identical to the maximum backlog of queue (2), and the maximum delay of queue (2) is the same as the maximum backlog of queue (1).

Now consider a queue in which the service curve is the adjoint of input traffic, that is $S(t) = A^*(t) \oplus A(t)$. For this queue,

$$D_{_M} = Q_{_M} = \|A(t)\|_\ell^2. \tag{37}$$

In fact, the queue is "matched" to the input traffic. In a matched queue, for all $t \geq 0$, we have

$$Q(t) = D(t). \tag{38}$$

Using the notation for division in min-plus algebra, we can write (38) as

$$\frac{A(t)}{S(t)} = \frac{S^*(t)}{A^*(t)} \tag{39}$$

Therefore, for a matched queue we have

$$A(t) \otimes A^*(t) = S(t) \otimes S^*(t) \tag{40}$$

It is possible to show that the matched queue is the only queue that satisfies (38).

Similarly, for the input traffic $A^*(t)$, the queue will be matched to the input if the same service curve, $S(t) = A^*(t) \oplus A(t)$, is used. In this case, we have

$$D'_{_M} = Q'_{_M} = \|A^*(t)\|_\ell^2 = \|A(t)\|_r^2. \tag{41}$$

See Fig. 4 for an illustration. In Fig. 4, we have assumed $\lim_{t\to\infty} A(t)/t = 1$ in order to have finite left and right seminorms.

Example: A leaky bucket regulator is represented by a pair of parameters (σ, ρ) where σ is the size of token pool and ρ is token replenishment rate. A greedy traffic

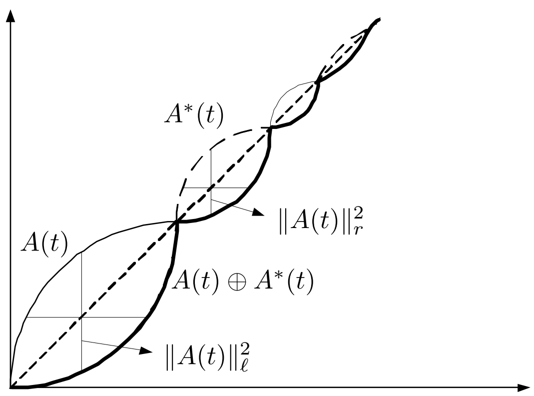

Fig. 4. An illustration for a matched queue.

at the output of a leaky bucket regulator can be shown as $A(t) = \sigma + \rho t$. Assume that this traffic is served by a matched queue. The output of the queue will be $B(t) = \min\left\{ \max\{\frac{t-\sigma}{\rho}, 0\}, \sigma + \rho t \right\}$. In fact, the output corresponds to a constant rate server with a constant initial delay. The delay is σ units and the service rate is $1/\rho$. For this queue, if $\rho \leq 1$, then $Q_{_M} = D_{_M} = \|\sigma + \rho t\|_\ell^2 = \sigma(1+\rho)$. If $\rho > 1$, the queue is unstable and $Q_{_M} = D_{_M} = \infty$.

IV. Conclusion

In this paper, we have used min-plus algebra to devise a duality between maximum delay and maximum backlog in a singe-queue network. The two main operators of min-plus algebra are pointwise minimization and pointwise addition. We have shown that under these operators the set of nondecreasing positive functions is a "dioid" — a dioid has the same properties as a "ring" in the conventional algebra. Using min-plus algebra, maximum backlog is represented as the scalar projection of accumulated input traffic curve onto accumulated output traffic curve. We have defined an "adjoint" operator in the dioid of nondecreasing functions. We have shown that the adjoint operator in min-plus algebra has the properties similar to the adjoint operator of conventional algebra. Using this definition, we have shown that maximum delay can be represented as the scalar projection of the adjoint of accumulated output traffic onto the adjoint of accumulated input traffic. This observation establishes a duality between the maximum delay and the maximum backlog.

We have also defined the left and the right seminorms. Using these definitions, we have shown that if the accumulated output traffic is the adjoint of input traffic, a "matched" queue will be obtained. In a matched queue, the delay and the backlog are identical for all time instants. Furthermore, the maximum delay is equal to the maximum backlog and can be represented as the square of the left norm of input traffic.

References

[1] R. L. Cruz, "A calculus for network delay, part I: network elements in isolation," *IEEE Trans. Inform. Theory*, vol. 37, pp. 114–131, Jan. 1991.

[2] R. L. Cruz, "A calculus for network delay, part II: network analysis," *IEEE Trans. Inform. Theory*, vol. 37, pp. 132–141, Jan. 1991.

[3] A. K. Parekh and R. G. Gallager, "A generalized processor sharing approach to flow control in integrated services network: the single node case," *IEEE/ACM Trans. Networking*, vol. 1, pp. 334–357, June 1993.

[4] C. S. Chang, "On deterministic traffic regulation and service guarantees: a systematic approach by filtering," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1097–1110, May 1998.

[5] J. Y. Le Boudec, "Application of network calculus to guaranteed service networks," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1087–1096, May 1998.

[6] J. Y. Le Boudec and P. Thiran, *Network Calculus*. Springer Verlag, 2001. [Online]. Available: http://ica1www.epfl.ch/PS_files/NetCal.htm.

[7] C. S. Chang, *Performance Guarantees in Communication Networks*. Springer-Verlag, 2000.

[8] R. L. Cruz, "Quality of service gurantees in virtual circuit switched networks," *IEEE J., Select., Areas Commun.*, vol. 13, pp. 1048–1057, August 1995.

[9] H. Sariowan, *A service curve approach to performance guarantees in integrated-services networks*. PhD thesis, University of California, San Diego, 1996.

[10] G. Cohen, P. Moller, J. P. Quadrat, and M. Viot, "Algebraic tools for the performance evaluation of discrete event systems," *Proceedings of IEEE*, vol. 77, pp. 39–58, Jan. 1989.

[11] F. Baccelli, G. Cohen, G. J. Olster, and J. P. Quadrat, *Synchronization and Linearity, An Algebra for Discrete Event Systems*. New York: Wiley, 1992.

[12] C. S. Chang, "Deterministic traffic specification via projections under the min-plus algebra," *Proc. IEEE INFOCOMM*, pp. 43–50, 1999.

[13] R. Cuninghame-Green, *Minimax Algebra*. Lecture Notes in Economics and Mathematical Systems, Springer Verlag, 1979.

[14] S. Valaee, "A methodology for virtual network partitioning: The deterministic approach," *preprint*, 2001.