

Katarína Bachratá *

EFFECTIVE BANDWIDTH FOR DETERMINISTIC NETWORKS

In present paper are given exact definitions of the studied notions. Some elementary properties of an effective bandwidth are deduced. Some terms and results are illustrated with examples for finite as well as for infinite packet flows in the internet network.

1. Introduction

Packets in packet switched communication networks create a random point process which can be described as it is known from the queueing theory. This randomness of packet inputs has an important impact to the quality of service that is offered by packet switched networks for real time applications. Therefore, it is necessary have a policy mechanism which controls a committed rate of incoming packets. The concept of the committed rate needs another description of the packet stream, which is known as an effective bandwidth. To use this concept in a proper way it is necessary to understand deeply stream behavior and consequently delay and packet loss introduced by an effective bandwidth as a link capacity. This paper explains a determinist case of the packet flow as a first step in understanding its random behavior.

These problems can be found in more abstract form in [1], [2].

2. Basic terms

The flow of packets will be modeled by the sequence $\{X(t_i)\}$ whose terms represent the number of packets arriving the communications system in discrete times

$$t_0, t_1, t_2, t_3, \dots$$

In this paper the following terms will be used:

The channel is that part of Internet network, which transmits the packets.

Analogical term in the queueing theory is *the serving link*.

The capacity of a channel is the maximum amount of packets transmitted in the unit of time.

Analogical term in the queueing theory is *the service rate*.

Buffer is the place where the packets are stored in the case that they have to wait for transmission.

Analogical term in the queueing theory is *the queue*.

The buffer size (backlog) is the maximum number of packets stored in buffer and waiting for transmission.

Analogical term in the queueing theory is *the queue length*.

The communication system is the system consisting of a buffer and a channel.

Analogical term in the queueing theory is *the single server system with queue*.

The virtual delay at time t is the gap between the arrival of a packet into the communication system at the given time t and the start of packet transmission through the channel. Because at given time t there may be no real arrival of packet, we talk about virtual delay of fictive packet.

Analogical term in the queueing theory is *the queueing time plus service time* of a customer arriving the system at a given time t .

The flow of packets is a sequence of packets transmitted through the communication system.

Analogical term in the queueing theory is *the arrival process of customers*.

The effective bandwidth (of a given input flow) is the minimum capacity of the channel (or transmission speed, packet rate) with which communication system is able to transmit the given flow under a given delay or queue characteristics (precise definition is given below).

The effective bandwidth is measured in the packets per second [p/s] units.

In the queueing theory for the given flow in the *stable single-server system with the fixed (finite or infinite) queue length* the number of served customers per unit of time is computed. In other words, this is the *minimum service rate*, which guarantees sustainable servicing without overloading the system.

* Katarína Bachratá

Department of InfoCom Networks, Faculty of Management Science and Informatics, University of Žilina,
E-mail: Katarina.Bachrata@fri.utc.sk

3. Calculation of an effective bandwidth

Next we will assume the flow of packets represented by the sequence $\{X(t_i)\}$.

The number of packets arriving by flow X at times $t_0, t_1, t_2, t_3, \dots, t_i, \dots$ will be

$$X(t_0), X(t_1), X(t_2), X(t_3), \dots, X(t_i), \dots$$

To simplify the situation we assume that for the arriving times, $t_0, t_1, t_2, t_3, \dots$, holds $t_i - t_{i-1} = 1$ second. At the beginning the flow is empty, so in time t_0 is $X(t_0) = 0$.

The arrival flow of packets may be *finite* (bounded in time), or *infinite* (unbounded in time).

We will present the computation method of the effective bandwidth for both cases separately.

1. Finite (time-bounded) flow of packets

The cumulative function $R_X(t)$ of the flow X counts the number of packets (seen on the flow in time interval $[0, t]$) arriving by the flow X from the time 0 until the time t :

$$R_X(t) = X(t_0) + X(t_1) + X(t_2) + \dots + X(t)$$

Computation $efb_B(X)$ of the effective bandwidth of a finite flow of packets

$$X = \{X(t_0), X(t_1), X(t_2), X(t_3), \dots, X(t_k)\}$$

and for a fixed buffer size B :

$$efb_B(X) = \sup_{0 \leq s \leq t \leq t_k} \left\{ \frac{R_X(t) - R_X(s) - B}{t - s} \right\} \quad (1)$$

Computation $efb_D(X)$ of the effective bandwidth of a finite flow of packets

$$X = \{X(t_0), X(t_1), X(t_2), X(t_3), \dots, X(t_k)\}$$

and for a given maximum virtual delay D :

$$efb_D(X) = \sup_{0 \leq s \leq t \leq t_k} \left\{ \frac{R_X(t) - R_X(s)}{t - s + D} \right\} \quad (2)$$

Example 1:

Consider a flow where in times t_i $X(t_i)$ packets arrive.

i	0	1	2	3
$X(t_i)$	0	1	1	1

We perform the computation of the effective bandwidth of a flow X for a fixed buffer size B according to formula (1).

B	0	1	2	3	4	5
$efb_B(X)$	0	2/3	1/3	0	0	0

Commentary:

The minimum transmission speed (minimum capacity of channel) which is able to transmit the flow X in a system without buffer is

$$efb_{B=0}(X) = 1[p/s]$$

The requirement of zero buffer size (the system has no buffer) can be replaced by the requirement of zero delay time (packets mustn't wait).

Then we have:

$$efb_{B=0}(X) = efb_{D=0}(X)$$

The minimum transmission speed (minimum capacity of channel) which is able to transmit the flow X for the buffer size equal to 1 is

$$efb_{B=1}(X) = \frac{2}{3} [p/s]$$

It means when 1 packet can wait in the buffer then with a speed $2/3[p/s]$ in 3 seconds we transmit 2 packets with 1 packet waiting.

The request of a buffer size 1 means that 1 packet can wait.

The delays (waiting times) of packets can be computed from the transmission speed. According to the table the transmission speed is $2/3[p/s]$. In the buffer there is at most 1 packet waiting. The waiting time (the time to empty the full buffer) is 1.5 second.

The packets are arriving by the flow X for 3 seconds and afterwards 1.5 seconds are necessary to transmit the rest of the flow.

We have:

$$efb_{B=1}(X) = efb_{D=1.5}(X)$$

If the buffer size is 3 or more, we can let all the 3 packets of a flow X wait. In the extreme case there is no need for transmission, and the "transmission" of flow X can be realized by buffering all the packets. For this case the transmission speed 0 $[p/s]$ is sufficient.

It is clear that if we can use a buffer with sufficient size (and the packets in buffer can wait for an arbitrary time) for transmission of the finite flow a very small transmission speed is suitable. In an extreme case all packets are waiting in the buffer and no transmission is realized.

So it will be useful to restrict the delay for a packet instead of restricting the buffer size.

We compute the effective bandwidth values for the same flow X , for a fixed maximum delay D . According to the formula (2).

D	0	1	1.5	2	3	4
$efb_D(X)$	1	3/4	2/3	3/5	3/6	3/7

Commentary:

The minimum transmission speed (minimum capacity of channel) which is able to transmit the flow X when there are no delays for packets allowed is

$$efb_{D=0}(X) = 1[p/s]$$

The minimum transmission speed which is able to transmit the flow X if 1 second delays of packets are allowed is

$$efb_{D=1}(X) = \frac{3}{4} [p/s]$$

In 3 seconds from 3 packets arriving with flow X , 3 times $3/4$ of packets are transmitted, it is $9/4$ of packets on the whole. The remaining $3/4$ of the packet will wait 1 second until it will be transmitted with speed $3/4 [p/s]$.

According to the table it is not necessary to use only integer values of delays. If the packets can wait 1.5 second, the minimum transmission speed can be $2/3 [p/s]$. Then 2 packets are transmitted in 3 seconds and then the last packet is transmitted 1.5 second (with speed $2/3 [p/s]$).

Example 2:

Consider a flow where in times t_i $X(t_i)$ packets arrive:

i	0	1	2	3	4	5	6	7
$X(t_i)$	0	1	5	0	0	2	0	1

The values of the effective bandwidth of flow X , for a fixed buffer size B .

B	0	1	2	3	4	5
$efb_B(X)$	5	4	3	2	1	0

The values of the effective bandwidth of flow X , for a fixed maximum delay D :

D	0	0.5	1	1.5	2	2.5
$efb_D(X)$	5	3.33	2.5	2	1.667	1.4286

If the effective bandwidth of flow X is smaller than the capacity of the channel c ,

$$efb_D(X) \leq c$$

then it is possible to transmit the flow X through this channel.

For finite flows with zero sized buffer holds:

If we aggregate the flows X_1, X_2, \dots, X_k and compute the effective bandwidth of this sum, we obtain a value which is less than or equal to the sum of the effective bandwidths of these flows.

$$efb_B\left(\sum_{j=1}^k X_j\right) < \sum_{j=1}^k efb_B(X_j) \quad (4)$$

When computing the effective bandwidth using the maximum delay, then the formula holds without constraints:

$$efb_D\left(\sum_{j=1}^k X_j\right) < \sum_{j=1}^k efb_D(X_j) \quad (5)$$

This property can be used by transmission of more flows through the same channel. Although the sum of the effective bandwidths of some flows overreaches the capacity of some channel, still these flows can be transmitted through this channel. The effective bandwidth of the sum of these flows can be less than the capacity of the channel.

Example 3:

Consider the flows X and Y with the following number of packets in times t :

t	0	1	2	3	4	5	6	7	8	9	10
$X(t_i)$	0	2	0	2	0	2	0	2	0	2	0
$Y(t)$	0	1	3	0	2	0	4	0	0	1	3
$[X+Y](t)$	0	3	3	2	2	2	4	2	0	3	3

B	0	1	2	3	4	5
$efb_B(X)$	2	1	0.8889	0.7778	0.6667	0.5556
$efb_B(Y)$	4	3	2	1.2	1	0.9
$efb_B(X) + efb_B(Y)$	6	4	2.8889	1.9778	1.6667	1.4556
$efb_B(X + Y)$	4	3	2.3333	2.1667	2	1.9

D	0	1	2	3	4	5
$efb_D(X)$	2	1	0.9091	0.8333	0.7692	0.7143
$efb_D(Y)$	4	2	1.3333	1.125	1	0.9333
$efb_D(X) + efb_D(Y)$	6	3	2.2424	1.9583	1.7692	1.6476
$efb_D(X + Y)$	4	2.2857	2	1.8462	1.7143	1.6

Commentary:

In this example we can see that the relation (4) holds for zero value of B (buffer size). (That relation holds for such a small buffer that not too many packets will be stored and transmitted after arrival of the whole flow into the system.)

The relation (5) can be proved for an arbitrary finite delay D .

2. The computation of the effective bandwidth $efb(X)$ for an infinite flow of packets

The flow of packets represented by sequence $\{X(t_i)\}$ will be transmitted through some channel. This channel will be represented by sequence $\{\hat{A}(n)\}$, which we call the envelope of the flow. We will search for the smallest envelope $\{\hat{C}(n)\}$ with sufficient properties. (We try to find the (minimum) smallest channel, which is able to transmit the given flow.)

We call the envelope of a sequence $\{X(t_i)\}$ such sequence $\{\hat{A}(n)\}$, which is the upper bound of the sum of arbitrary n consecutive terms of sequence $\{X(t_i)\}$:

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + X(t_{i+2}) + \dots + X(t_{i+n-1})] \leq \hat{A}(n) \quad (6)$$

Because the condition (6) is satisfied, the flow $\{X(t_i)\}$ can be transmitted through the channel $\{\hat{A}(n)\}$ ("it fits"). The sequence $\{\hat{A}(n)\}$ can be replaced by a *nondecreasing and sub-additive* sequence $\{\hat{C}(n)\}$ which is also the envelope of the sequence $\{X(t_i)\}$.

We will explain why we require these properties and how we can assure them:

Because the sequence $\{X(t_i)\}$ represents the flow of packets, its terms are nonnegative:

$$X(t_i) > 0$$

The *nondecreasing property* of the envelope

$$\hat{B}(n) < \hat{B}(n+1) \quad \forall n \in N \quad (7)$$

means that the number of packets transmitted by the channel during a longer period $(n+1)$ is equal or greater than the number of packets transmitted by the channel during a shorter period (n) .

We have:

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1})] \leq$$

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1}) + X(t_{i+n})] \leq \hat{A}(n+1)$$

and also

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1})] \leq \hat{A}(n+1)$$

It follows: there exists a number β , which holds

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1})] \leq \beta < \hat{A}(n+1)$$

Define now $\hat{B}(n) = \beta$.

$\{\hat{B}(n)\}$ is the envelope of the sequence $\{X(t_i)\}$ and in addition we assure that the envelope $\{\hat{B}(n)\}$ is nondecreasing and "smaller" than $\{\hat{A}(n)\}$:

$$\hat{B}(n) < \hat{A}(n) \quad \forall n \in N$$

The *sub-additivity* of envelope, it means the property

$$\hat{C}(n+m) + \hat{C}(n) + \hat{C}(m) \quad \forall n, m \in N \quad (8)$$

represents the fact that in the time period of length $m+n$, the channel can transmit at most such an amount of packets as in the time period m and n (these periods can start in an arbitrary moment).

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+m+n-1})] =$$

$$\sup_{\forall i} [X(t_i) + \dots + X(t_{i+n-1}) + X(t_{i+n-1+1}) + \dots +$$

$$+ X(t_{i+n-1+m})] \leq \hat{B}(n) + \hat{B}(m)$$

For $\hat{B}(n) + \hat{B}(m)$ is the upper boundary of the value sup

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+m+n-1})]$$

there exists a number γ , which holds

$$\sup_{\forall i} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+m+n-1})] \leq \gamma \leq \hat{B}(n) + \hat{B}(m)$$

Let $\hat{C}(n+m) = \gamma$.

The sequence $\{\hat{C}(n)\}$ is the envelope of the sequence $\{X(t_i)\}$ and moreover we assure that the envelope $\{\hat{C}(n)\}$ is non-decreasing, sub-additive and

$$\hat{C}(n) < \hat{A}(n) \quad \forall n \in N$$

If the sequence $\{X(t_i)\}$ is known, from all of nondecreasing and sub-additive envelopes $\{\hat{C}(n)\}$ we can choose the smallest one $\{C^*(n)\}$.

$$\{C^*(n)\} = \min_{\forall C} [\{\hat{C}(n)\}] \quad (9)$$

The terms of sequence $\{C^*(n)\}$ (the minimum envelope) can be also estimated as supreme of all the sums of n consecutive terms of the sequence

$$\{C^*(n)\} = \sup_{i \leq 0} [X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1})] \quad (10)$$

The value

$$c^* = \lim_{n \rightarrow \infty} \frac{C^*(n)}{n} \quad (11)$$

is the mathematical analogue of the *effective bandwidth*, (this is the *intensity* of the minimum envelope of flow X).

$$efb(X) = c^* \quad (12)$$

If the sequence $\{X(t_i)\}$ is not known, we can use the knowledge of any arbitrary envelope (any channel which was able to transmit the flow). We can minimize this envelope and transfer it into nondecreasing and sub-additive sequence, which still remains an envelope.

Example 4:

Consider the given envelope $\{\hat{A}(n)\}$ represented a channel which was able to transmit some flow.

We assume, that the flow represented by the sequence $\{X(t_i)\}$ was transmitted through channel $\{\hat{A}(n)\}$, whose capacity is:

n	1	2	3	4	5	6	7
$\hat{A}(n)$	8	6	7	21	14	25	19

Our goal is to minimize the envelope $\{\hat{A}(n)\}$ and preserve the non-decreasing property of the envelope. We change the values of the envelope setting:

$$\hat{B}(n) = \inf_{m \geq n} \hat{A}(m)$$

(Thus if for transmission of the flow X in $n = 2$ time units the channel with capacity $\hat{A}(2) = 6$ is sufficient, then for transmission of the flow X in $n = 1$ time unit the channel of capacity $\hat{B}(1) = 6$ is sufficient.)

n	1	2	3	4	5	6	7
$\hat{B}(n)$	6	6	7	21	14	25	19

In the same way we change the values $\hat{A}(4)$ and $\hat{A}(6)$.

n	1	2	3	4	5	6	7
$\hat{B}(n)$	6	6	7	14	14	19	19

$\{\hat{B}(n)\}$ is the envelope of the sequence $\{X(t_i)\}$ and we ensure that the envelope $\{\hat{B}(n)\}$ is non-decreasing and

$$\hat{B}(n) < \hat{A}(n) \quad \forall n \in N$$

Next we ensure the sub-additivity by setting:

$$\hat{C}(n+m) = \min\{\hat{B}(n+m); \hat{B}(n) + \hat{B}(m)\}$$

(So if for transmission in $n = 2$ time units the channel with capacity $\hat{B}(2) = 6$ is sufficient, then for transmission in $n = 2 + 2 = 4$ time units the channel with intensity $\hat{C}(4) = 6 + 6 = 12$ packets is sufficient.)

n	1	2	3	4	5	6	7
$\hat{C}(n)$	6	6	7	12	14	19	19

Using the same advance we recalculate the values $\hat{C}(5)$ and $\hat{C}(6)$.

n	1	2	3	4	5	6	7
$\hat{C}(n)$	6	6	7	12	13	14	19

The envelope $\{\hat{C}(n)\}$, which we obtain using this method is wide enough to transmit the flow, is non-decreasing and sub-additive.

If we know all of the possible nondecreasing and sub-additive envelopes of the given (and unknown) flow, we can write:

$$\begin{aligned} C^*(n) &= \min_{\forall C} \{\hat{C}(n)\} = \\ &= \sup_{i \leq 0} \{X(t_i) + X(t_{i+1}) + \dots + X(t_{i+n-1})\} \end{aligned}$$

For sub-additive functions $f(t)$ holds

$$\lim_{n \rightarrow \infty} \frac{f(t)}{t} = \inf_{t \geq 1} \frac{f(t)}{t}$$

For the computation of c^* we can use formula:

$$c^* = \lim_{n \rightarrow \infty} \frac{C^*(n)}{n} = \inf_{t \geq 1} \frac{C^*(n)}{n}$$

For the infinite flow X can be proved:

If $efb(X) < \infty$, then there exists such $D \in R_0^+$, that the packets of the flow are waiting in the buffer for maximum delay D .

Moreover, the inequality (5) holds in the formSummary:

$$efb\left(\sum_{j=1}^k X_j\right) \leq \sum_{j=1}^k efb_B(X_j)$$

We have shown some examples of the effective bandwidths in the deterministic case.

The given examples illustrated some important properties of the characteristic of the flow of packets called effective bandwidth. The given formulas can be used for computing minimal envelopes.

References

- [1] LE BOUDEC, J.-Y., THIRAN, P.: *Network Calculus, A Theory of Deterministic Queueing Systems for the Internet*, Springer-Verlag Berlin, Heidelberg 2001
- [2] CHANG, CH.-SH: *Stability, Queue Length and Delay of Deterministic and Stochastic Queueing Networks*, IEEE Trans. Automatic Control, Vol. 39, pp. 913-931 (1994).