Proceedings of QoFiS, Stockholm, October 2003

Deterministic end-to-end delay guarantees in a heterogeneous Route Interference environment

Florian-Daniel Oţel , Jean-Yves LeBoudec

11th August 2003

Abstract

Some of the known results for delivering deterministic bounds on end-to-end queuing delay in networks with constant packet sizes and constant link rates rely on the concept of Route Interference. Namely, it is required to know the number of flows joining on any output link in the whole network. In this paper we extend the existing results for the more generic cases of connection-oriented networks consisting of links with different capacities, carrying different traffic classes and packets of different sizes.

1 Introduction and related work

In the quest for delivering deterministic end-to-end delay guarantees in general networks it has been shown [2] that it is feasible to deliver deterministic bounds for queuing delay in networks using FIFO queuing but the bound is dependent on complex network conditions. Specifically, by strictly controlling the number of times flow paths join on output links and by performing ingress traffic shaping in accordance with these metrics, it is possible to compute tight bounds on queuing delay and required buffer capacities. However, the results presented in [2] are limited to very specific network setups i.e. to connection-oriented networks carrying packets of fixed size (ATM networks), with all links having same capacity and where time is considered to be divided in equal slots that are synchronized network-wide. Moreover, nodes were assumed to be globally FIFO and have zero internal propagation and processing delays.

Later results [3, 4] relaxed the requirement for synchronized time slots and improved the bounds on required buffer capacities and end-to-end queuing delay. However, they maintained the limiting requirement for equal packet sizes and equal capacity links.

In this paper we present an extension of the proofs in [1] that relaxes these requirements and generalizes the results to generic connection-oriented networks with links of different speeds and carrying different types of traffic with different packet sizes.

The outline of this paper is as follows: In the next subsection we state the assumed traffic, network, and time models. In Section 2 we introduce the concept of route interference, the source rate condition and we state the main result of this paper – the theorem for bounded buffer and delay. The proof of the theorem is given separately in Section 3 as it involves a detailed analysis of the mechanics of flow aggregation and queue busy periods along a flow path. Section 4 outlines how to compute the required buffer capacity (i.e. maximum amount of queue backlog) and the queuing delay. In addition to the algorithmic steps necessary to compute these values, the results are also given as closed form approximation formulas.

1.1 Assumed network, traffic and time models

Traffic model For the purpose of this paper we assume that transmission of delay-sensitive data is performed in a connection-oriented manner, with traffic organized in flows whose routes are pre-established before data transmission. Inside each flow data is transmitted in packets having a finite set of possible packet sizes.

Network model We consider that the network consists of nodes which offer service guarantees in the form of generic rate-latency service curves [1]. However, for the sake of clarity we will consider only the special case of non-preemptive schedulers performing strict priority FIFO queuing. Specifically, we will consider that nodes have a single FIFO queue per traffic class and that delay sensitive traffic has the highest priority in the network. Under these assumptions the node serves the delay-sensitive traffic with a rate-latency service curve $\beta_{r,\tau}$ with rate $r =$ the physical link rate and latency $\tau = \frac{MTU_L}{r}$, where MTU_L is the maximum packet size for lower priority traffic classes.

Also, we will assume that network links are unidirectional, with variable rates and propagation delays. Without the loss of generality we will consider that the (bounded) internal processing and transmission delays at network nodes are included in the upstream link propagation delays. As such for the rest of this paper we will assume that node internal delays are negligible.

Time model Time is assumed to be continuous and relevant network events have a time index sequentially numbered starting from time 0, when the network is in an idle state. In other words we assume that all packet receptions and transmissions time ordered, network-wide.

2 The Source Rate Condition and the theorem for bounded buffer and queuing delay

In order to present the main result of this paper – the theorem for bounded buffer and queuing delay – we first define the concept of flow joins and introduce the source rate condition as a requirement for ingress traffic shaping.

Definition 1 (Flow joins, Interference Event) *Two flows F and G are said to* join on link[1] j *if both flows share link* j *but do not share the link upstream from* j *in their respective paths.*

An interference event *is defined as a pair* (j,*{F,G})* *where* j *is a link and F and G are two flows joining at link* j. *As such the number of flows joining F on link* j *is given by the number of interference events that contain* j *and F.*

Definition 2 (Source Rate Condition) *We say that a flow F satisfies the* Source Rate Condition *if the inter-packet emission time T_F satisfies the inequality:*

$$T_F \geq \frac{MTU_H}{r_F^*} \sum_{j=src}^{dst} I_j + MTU_H \sum_{j=src}^{dst} S_j \left(\frac{1}{r_j} - \frac{1}{r_{prev_F(j)}} \right)^+ + \frac{MTU_H}{r_F^*} + \sum_{j=src}^{dst} \frac{MTU_L}{r_j}$$

where:

j - *node along the path of flow F, starting from source node and ending with destination node.*

r_j - *the rate of the link outgoing from node* j *along the flow path. If* j *is the last node then r_j is assumed to be infinite.*

$prev_F(j)$- *The link previous to* j *along the path of flow F i.e. the link upstream to* j *along the route of flow F. If node* j *is the first node along the flow path, then $r_{prev_F(j)}$ is infinite.*

r_F^*- *The minimum capacity link along the path of flow F i.e.* $r_F^* = \min_j r_j$.

I_j - *number of flows that join flow F at link* j *i.e. the number of interference events that contain both* j *and F.*

S_j - *number of all flows except F that share both link j and $prev_F(j)$.*

MTU_H - *the maximum packet size for the delay-sensitive, high priority traffic.*

MTU_L - *the maximum packet size of lower priority traffic classes.*

In the above formula – as well as for the rest of this paper – the $expression^+$ notation is a shorthand for $\max(expression, 0)$.

Theorem 1 (Theorem for bounded buffer and delay) *Provided that the source rate condition holds for all flows, then:*

- *The network is stable i.e. the maximum amount of backlog at any queue and the corresponding required buffer capacity are bounded.*

[1] We will alternatively use the expression "node" or "link" as meaning the same thing i.e. the corresponding network event occurring at the named/implied outgoing link of the named/implied node immediately upstream of that link.

Also, unless explicitly noted otherwise, when referring to "flows", "traffic", "packets" or "interfering segments" we implicitly refer to the high-priority traffic for which delay guarantees must be delivered.

- *The queuing delay at any node is bounded.*

The proof of the theorem – presented in the next section – involves a complex analysis of the queue busy periods and the relations between queue backlogs and interference events. The net result of the theorem are the bounds for the maximum backlog and maximum queuing delay. The bounds are given both as a description of the algorithmic steps necessary to compute them and as closed form approximation formulas.

3 Proof for bounded buffer and delay theorem

Before delving into the theorem proof proper we introduce some technical definitions that express the concept of chained busy periods i.e. queue busy periods at successive nodes along a flow path.

Definition 3 (Delay operation) *For two packets* p *and* q *and for some link* j *we say that* $p \prec_j q$ *if* p *and* q *are in the same busy period of the queue for high-priority traffic at* j *and* p *is transmitted on* j *before* q*. Also by* $p \preceq_j q$ *we say that* p *leaves on* j *no later than* q *(or, alternatively,* q *leaves on* j *no earlier than* p*).*

It must be noted that, since we refer to packets belonging to highest priority traffic class and since nodes perform priority queuing, the delay relationship between two packets implicitly states that there are no low-priority packets being transmitted between them on the output link.

Definition 4 (Super-Chain, Super-chain path) *Consider a sequence of packets* $\underline{p} = (p_0......p_i....p_k)$ *and a sequence of nodes* $\underline{f} = (f_1.......f_k)$*. We say that* $(\underline{p},\underline{f})$ *is a* super-chain *if:*

- $f_1, ..., f_k$ *are all on* P *- the path of packet*[2] p_0*, not necessarily consecutive but distinct.*
- $p_{i-1} \prec_{f_i} p_i$ *for* $i = 1$ *to* k*.*
- *The path of packet* p_i *from* f_i *to* f_{i+1} *is a sub-path of* P*.*

The path of the super-chain *is defined as the sub-path of* p_0 *that spans from* f_1 *to* f_k*.*

Definition 5 (Relevant network events, arrival and departure time) For the purpose of this paper we define a *relevant network event* as the enqueuing or dequeuing of a packet at a link/node. Also, we will denote the time index of these events as a_j^k for the arrival of packet number k at link j (defined as the time index when the last bit of packet k is received) and, respectively, d_j^k for

[2] *For simplicity we refer to "the path of a packet" as meaning the network route of the flow the said packet belongs to.*

the corresponding departure time (defined as the time index when the last bit of packet k is transmitted).

It is to be noted that since the set of possible packet sizes was assumed to be finite we cannot have an infinite number of network events occurring in a finite time interval.

Definition 6 (Segment interfering with a super-chain) *For a given super-chain we call* segment *an ordered pair (*s,P*) where* P *is a sub-path of the path of the super-chain,* s *is a packet whose path has* P *as a sub-path and* P *is maximal (namely we cannot extend* P *to be a common sub-path of both* s *and the super-chain).*

We say that the segment *(*s,*P) is* interfering with the super-chain *(*$\underline{p}$,$\underline{f}$*) if there is some node* f_i *on* P *such that* $s \prec_{f_i} p_i$.

The proof of the theorem is organized as follows: First, based on the number of segments interfering with a super-chain, we will derive an expression for the delay experienced along a super-chain path. Second, we will prove the non intra-flow property i.e. show that, if the source rate condition is imposed for all flows, then there cannot be two packets from the same flow in a super-chain. Finally, using the formula for the delay along a super-chain path and the non intra-flow interference property, we show that buffer requirements and queuing delay at any node in the network are bounded, thus concluding the proof.

3.1 Delay along a super-chain path

Let ($\underline{p}$,$\underline{f}$) be a super-chain and consider the f_j node on the super-chain path (see Fig. 1). Let v_j be the beginning of the busy period (for the queue for high-priority traffic) that a_j^{j-1} is in, i.e. $v_j = a_j^n$ for some packet number n with $n \leq j-1$. Assume without the loss of generality that packet p_{j-1} arrives on input link i (by necessity belonging to the super-chain path) and packet p_j arrives on input link k, not necessarily distinct from i. Define $\mathcal{B}_j$ as the set interference segments (s,P) such that s is arriving at the node no earlier than time v_j on a link other than input link i (i.e. on a link incident to the super-chain path), $s \preceq_{f_j} p_j$ and P is the maximal common sub-path for s and the path of the super-chain. Define $\mathcal{B}_j^0$ in the same manner but with packet s arriving on the same input link as packet p_{j-1} i.e. on the path of the super-chain. Also define $\mathcal{A}_j^0$ as the subset of $\mathcal{B}_j^0$ that contains only the packets that depart no earlier than p_{j-1} i.e. $p_{j-1} \preceq_{f_j} s$. Let B_j (resp. B_j^0, A_j^0) be the number of elements in $\mathcal{B}_j$ (resp. $\mathcal{B}_j^0$, $\mathcal{A}_j^0$). Please note that by definition $p_j \notin \mathcal{B}_j \bigcup \mathcal{B}_j^0$ and[3] $p_{j-1} \in \mathcal{A}_j^0$.

With r_i^j and r_j being the rates of input link i and, respectively, link j – the

[3]By an abuse of notation we will write packet $p_j \notin \mathcal{B}_j$ as meaning segment $(p_j, P) \notin \mathcal{B}_j$ for any path P and, respectively, packet $p_j \in \mathcal{B}_j$ as meaning segment $(p_j, P) \in \mathcal{B}_j$ for some path P.

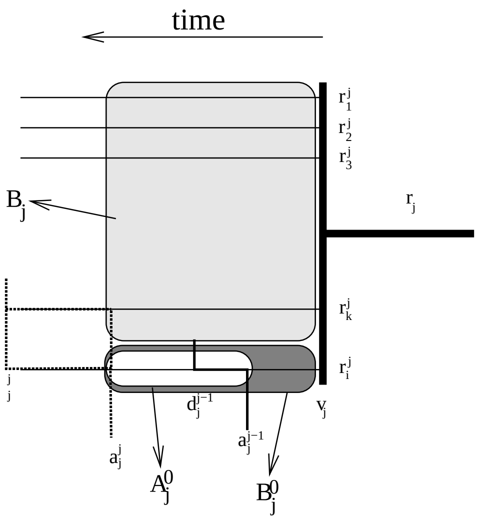

Figure 1: Node f_j on super-chain path and notation used in section 2.

link on the super-chain path outgoing from f_j, we have:

$$a_j^{j-1} - v_j \geq \frac{1}{r_i^j} \sum_{n \in \mathcal{B}_j^0 \setminus \mathcal{A}_j^0} l^n$$

$$d_j^j - v_j \leq \frac{1}{r_j} \sum_{n \in \mathcal{B}_j \cup \mathcal{B}_j^0} l^n + \frac{l^j}{r_j} + \frac{MTU_L}{r_j}$$

where l^n is the length of packet n, l^j is the length of packet p_j, l^j/r_j is the transmission time for packet p_j and MTU_L/r_j is the maximum node latency due to non-priority cross-traffic. Subtracting the two we obtain:

$$d_j^j - a_j^{j-1} \leq \frac{1}{r_j} \sum_{n \in \mathcal{B}_j \cup \mathcal{B}_j^0} l^n - \frac{1}{r_i^j} \sum_{n \in \mathcal{B}_j^0 \setminus \mathcal{A}_j^0} l^n + \frac{l^j}{r_j} + \frac{MTU_L}{r_j}$$

or, since $\mathcal{B}_j$ and $\mathcal{B}_j^0$ are disjoint and $\mathcal{A}_j^0 \subseteq \mathcal{B}_j^0$:

$$d_j^j - a_j^{j-1} \leq \frac{1}{r_j} \left(\sum_{n \in \mathcal{B}_j} l^n + \sum_{n \in \mathcal{B}_j^0 \setminus \mathcal{A}_j^0} l^n + \sum_{n \in \mathcal{A}_j^0} l^n \right) - \frac{1}{r_i^j} \sum_{n \in \mathcal{B}_j^0 \setminus \mathcal{A}_j^0} l^n + \frac{l^j}{r_j} + \frac{MTU_L}{r_j}$$

$$d_j^j - a_j^{j-1} \leq \frac{1}{r_j} \left(\sum_{n \in \mathcal{B}_j} l^n + \sum_{n \in \mathcal{A}_j^0} l^n \right) + \sum_{n \in \mathcal{B}_j^0 \setminus \mathcal{A}_j^0} l^n \left(\frac{1}{r_j} - \frac{1}{r_i^j} \right) + \frac{1}{r_j}(l^j + MTU_L)$$

Since $d_j^j - a_j^{j-1} \geq \frac{l^j + l^{j-1}}{r_j}$, the right hand side of the inequality must be equal or greater than this quantity for any combination of rates r_j and r_i^j. As $p_{j-1} \in \mathcal{A}_j^0$, the inequality above becomes:

$$d_j^j - a_j^{j-1} \leq \frac{1}{r_j}\left(\sum_{n\in\mathcal{B}_j} l^n + \sum_{n\in\mathcal{A}_j^0} l^n\right) + \sum_{n\in\mathcal{B}_j^0\setminus\mathcal{A}_j^0} l^n \left(\frac{1}{r_j} - \frac{1}{r_i^j}\right)^+ + \frac{1}{r_j}(l^j + MTU_L)$$

As $l^n \leq MTU_H$ for any packet n (including p_j) and $\sum_{n\in\mathcal{B}_j} l^n \leq MTU_H\, B_j$ (with the corresponding inequality also holding for, respectively, $\mathcal{B}_j^0$, $\mathcal{A}_j^0$ and $\mathcal{B}_j^0\setminus\mathcal{A}_j^0$) we obtain:

$$d_j^j - a_j^{j-1} \leq MTU_H \frac{B_j + A_j^0}{r_j} + MTU_H (B_j^0 - A_j^0)\left(\frac{1}{r_j} - \frac{1}{r_i^j}\right)^+ + \frac{1}{r_j}(MTU_H + MTU_L) \tag{1}$$

By iterative use of relation 1 along the super-chain path (i.e. along the subscripts) and packet numbers (superscripts) we obtain:

$$\begin{aligned} d_k^k - a_1^0 \quad \leq \quad & MTU_H \sum_{j=f_1}^{f_k} \frac{B_j + A_j^0}{r_j} + MTU_H \sum_{j=f_1}^{f_k} (B_j^0 - A_j^0)\left(\frac{1}{r_j} - \frac{1}{r_i^j}\right)^+ + \\ & + (MTU_H + MTU_L) \sum_{j=f_1}^{f_k} \frac{1}{r_j} + \tau_{1,k} \end{aligned}$$

$$\begin{aligned} d_k^k - a_1^0 \quad \leq \quad & \frac{MTU_H}{r_f^*} \sum_{j=f_1}^{f_k} (B_j + A_j^0) + MTU_H \sum_{j=f_1}^{f_k} (B_j^0 - A_j^0)\left(\frac{1}{r_j} - \frac{1}{r_i^j}\right)^+ + \\ & + (MTU_H + MTU_L) \sum_{j=f_1}^{f_k} \frac{1}{r_j} + \tau_{1,k} \end{aligned} \tag{2}$$

where $r_f^* = \min_j r_j$ corresponds to the smallest capacity link along the super-chain path, $\tau_{1,k}$ is the propagation time along the links in the super-chain path and the penultimate term denotes the transmission times and node latencies, cumulated along the path.

We now show that all the sets in the collection $\{\mathcal{B}_j \bigcup \mathcal{A}_j^0\}_{j=1\,to\,k}$ are two-by-two disjoint: First if $(s,P) \in \mathcal{B}_j$ then f_j is the first node of P and thus (s,P) cannot be in some $\mathcal{B}_{j'}$ with $j' \neq j$. Thus $\mathcal{B}_j$ are two-by-two disjoint. Second, if $(s,P) \in \mathcal{B}_j$ and $(s,P) \in \mathcal{A}_{j'}^0$ it is obvious from their definition that for a fixed j, $\mathcal{B}_j$ and $\mathcal{A}_j^0$ are disjoint, so we must have that $j' \neq j$. Since f_j is the first node on P and j' is on P then by necessity $j < j'$. From the definition of $\mathcal{B}_j$ and $\mathcal{A}_{j'}^0$ we have that $s \prec_{f_j} p_j$ and $p_{j'-1} \preceq_{f_{j'}} s$, which contradicts the FIFO assumption. Thus $\mathcal{B}_j$ and $\mathcal{A}^0_{j'}$ are two-by-two disjoint. The same type of reasoning leads us to show that $\mathcal{A}^0_j$ and $\mathcal{A}^0_{j'}$ are two-by-two disjoint, thus proving the proposition.

By definition every element in $\{\mathcal{B}_j \bigcup \mathcal{A}^0_j\}$ is an interfering segment, so given that $\{\mathcal{B}_j \bigcup \mathcal{A}^0_j\}_{j=1\,to\,k}$ are two-by-two disjoint we have that $\sum_{j=f_1}^{f_k}(B_j + A^0_j) \leq I_{1,k}$ where $I_{1,k}$ is the number of interfering segments along the super-chain path. Thus relation 2 becomes:

$$\begin{aligned} d^k_k - a^0_1 \quad \leq \quad & \frac{MTU_H}{r^*_f} \sum_{j=f_1}^{f_k} I_{1,k} + MTU_H \sum_{j=f_1}^{f_k} (B^0_j - A^0_j) \left(\frac{1}{r_j} - \frac{1}{r^j_i}\right)^+ + \\ & + (MTU_H + MTU_L) \sum_{j=f_1}^{f_k} \frac{1}{r_j} + \tau_{1,k} \end{aligned} \tag{3}$$

3.2 The non intra-flow interference property

Assume that the source rate condition holds. Let (p,f) be a super-chain.

1. For every interference event of packet p_0 there is at most one segment interfering with the super-chain.
2. B^0_j is upper bounded by the number of flows that share the same input link as packet p_{j-1} and same output link as packet p_j.
3. p_k does not belong to the same flow as packet p_0.

Proof: Define the *time of the super-chain* as the time index for the exit of packet p_k from the last node f_k.We use a recursion[4] on time t.

At time index $t = 1$ the proposition is true because any flow has transmitted at most one packet. Assume now that the proposition holds for *any* super-chain with time index $\leq t-1$ and consider a super-chain with time index t.

First, we associate an interference event to any segment (s,P) interfering with the super-chain as follows: The paths of s and p_0 may share several non contiguous sub-paths and P is one of them. Call f the first node of P. To s we associate the interference event (f,$\{j_0$,$j\}$), where j_0 (resp. j) is the flow of packet p_0 (resp. s).

We now show that this mapping is injective i.e. different s packets correspond to different interference events. Assume that another segment $(s', P') \neq (s, P)$ is associated with the same interference event (f,$\{j_0$,$j\}$). Without the loss of generality we can assume that s was emitted before s'. Since s and s' both belong to flow j and since P and P' are maximal, we must have $P = P'$. By hypothesis we have an interference with the super-chain at some node on P. Let

[4] Please note that this is permissible in this case since we have no accumulation point along time indexes.

f_l be a node on the super-chain and on P such that $s \prec_{f_l} p_l$. Assume that s' leaves node f_l before p_l. Since s was emitted before s' we have $s \prec_{f_l} s'$ and thus $((s,s'),(f_l))$ is a super-chain. Since s' is an interfering packet it necessarily must leave f_l before t, thus the proposition is true for the super-chain $((s,s'),(f_l))$, which contradicts item 3. Thus s' must leave f_l after p_l. But there is some index $m \leq k$ such that $s' \prec_{f_m} p_m$, thus packet s' leaves node f_m before p_m. Define l' the smallest index with $l < l' \leq m$ such that s' leaves node $f_{l'}$ after packet $p_{l'-1}$ and before $p_{l'}$. Then $((s, p_l, ...p_{l'-1}, s'), (f_l, ..., f_{l'}))$ is a super-chain with time index $\leq t-1$ which would again contradict item 3. Thus in all cases we have a contradiction, the mapping is injective and item 1 is proven for the super-chain.

For item 2, consider an interfering packet $s \in \mathcal{B}_j^0$. Assume there exists another interfering packet $s' \in \mathcal{B}_j^0$, with s' belonging to the same flow as s. Consider without the loss of generality that s was emitted before s'. Since by definition $s \prec_{f_j} p_j$ and $s' \prec_{f_j} p_j$, then we must have that $s \prec_{f_j} s'$ and $((s,s'),(f_j))$ is a super-chain with exit time $\leq t-1$, which contradicts item 3. As such we cannot have two packets in the same flow in $\mathcal{B}_j^0$, which proves item 2.

For item 3, let us compute a bound on maximum queuing delay for packet p_0. Consider u_0 its emission time, P_0 the sub-path of p_0 from its source up to, but excluding, node f_1, and T the total propagation *and transmission* time for p_0 along P_0 and the super-chain path. Consider that the component of T along the super-chain path is $T_{1,k}$. Applying relation (3) along P_0 and separating the summation terms for packet transmission times from node latencies, we have:

$$\begin{aligned} a_1^0 \quad &\leq \quad d_1^0 \leq u_0 + (T - T_{1,k}) + \frac{MTU_H}{r_F^*} \sum_{j=src\,node}^{prev_F(f_1)} I_{0,1} + \\ &+ MTU_H \sum_{j=src\,node}^{prev_F(f_1)} (B_j^0 - A_j^0) \left(\frac{1}{r_j} - \frac{1}{r_i^j} \right)^+ + \sum_{j=src\,node}^{prev_F(f_1)} \frac{MTU_L}{r_j} \end{aligned}$$

where F is the flow the packet p_0 belongs to, r_i^j is infinite for j=source node and $I_{0,1}$ is the number of interference events for F along P_0. From item 2 above $B_j^0 \leq$ number of flows sharing both link j and $prev_F(j)$. Since $p_{j-1} \in \mathcal{A}_j^0$ we have that $A_j^0 \geq 1$ and thus $B_j^0 - A_j^0 \leq S_j$, where S_j is the number of flows sharing link j, minus 1 (i.e. the number of flows sharing links $prev_F(j)$ and j, except the flow p_{j-1}belongs to). As such we can re-write the above expression as:

$$\begin{aligned} a_1^0 \quad &\leq \quad u_0 + (T - T_{1,k}) + \frac{MTU_H}{r_F^*} \sum_{j=src\,node}^{prev_F(f_1)} I_{0,1} + \\ &+ MTU_H \sum_{j=src\,node}^{prev_F(f_1)} S_j \left(\frac{1}{r_j} - \frac{1}{r_{prev_F(j)}} \right)^+ + \sum_{j=src\,node}^{prev_F(f_1)} \frac{MTU_L}{r_j} \end{aligned}$$

Using the same reasoning, along the super-chain path we have that:

$$\begin{aligned} d_k^k &\le a_1^0 + T_{1,k} + \frac{MTU_H}{r_F^*}\sum_{j=f_1}^{f_k} I_{1,k} + \\ &+ MTU_H \sum_{j=f_1}^{f_k} S_j\left(\frac{1}{r_j} - \frac{1}{r_{prev_F\,(j)}}\right)^+ + \sum_{j=f_1}^{f_k}\frac{MTU_L}{r_j} \end{aligned}$$

Combining the last two inequalities we obtain:

$$\begin{aligned} d_k^k &\le u_0 + T + \frac{MTU_H}{r_F^*}\sum_{j=src\,node}^{f_k} I_j + MTU_H \sum_{j=src\,node}^{f_k} S_j\left(\frac{1}{r_j} - \frac{1}{r_{prev_F\,(j)}}\right)^+ + \\ &+ \sum_{j=src\,node}^{f_k}\frac{MTU_L}{r_j} \end{aligned}$$

Assuming that p_k and p_0 belong to the same flow and u_k is the emission time of packet p_k , since the source rate condition holds (including for the sub-path of F from the source node to node f_k), by applying the Source Rate Condition we have that (with $k > 0$):

$$\begin{aligned} u_k &\ge u_0 + \frac{MTU_H}{r_F^*}\sum_{j=src\,node}^{f_k} I_j + MTU_H \sum_{j=src\,node}^{f_k} S_j\left(\frac{1}{r_j} - \frac{1}{r_{prev_F\,(j)}}\right)^+ + \\ &+ k\left(\frac{MTU_H}{r_F^*} + \sum_{j=src\,node}^{f_k}\frac{MTU_L}{r_j}\right) \end{aligned}$$

which – by adding on both sides T, the transmission and propagation times for packet p_k from its source to node f_k – translates at node f_k into (since $d_k^k \ge u_k + T$):

$$\begin{aligned} d_k^k &\ge u_0 + T + \frac{MTU_H}{r_F^*}\sum_{j=src\,node}^{f_k} I_j + MTU_H \sum_{j=src\,node}^{f_k} S_j\left(\frac{1}{r_j} - \frac{1}{r_{prev_F\,(j)}}\right)^+ + \\ &+ \sum_{j=src\,node}^{f_k}\frac{MTU_L}{r_j} + k\left(\frac{MTU_H}{r_F^*} + \sum_{j=src\,node}^{f_k}\frac{MTU_L}{r_j}\right) \end{aligned}$$

which contradicts the relation above. As such p_k and p_0 cannot belong to the same flow, which proves item 3 of the non intra-flow interference property.

3.3 Bounded buffer requirements and queuing delay

Given the non intra-flow interference property it follows immediately that, if the all flows rates satisfy the source rate condition, at any output queue for delay-sensitive traffic there can be at most one packet from each flow during any busy period. As such the total amount of traffic that transiently shares the queue during any period is upper bounded by the number of flows sharing that link multiplied the maximum packet size. Consequently, at any node the amount of queue backlog at any instant is bounded and the network is stable. Since the nodes perform FIFO queuing and offer service guarantees in the form of rate-latency curves to the delay-sensitive traffic, the amount of queuing delay at any node is bounded. This completes the proof for the theorem of bounded buffer and delay.

4 Buffer requirements and queuing delay computation

Without the loss of generality, consider a single-node with I fan-in links to the same output link i.e. for a given output link consider the input links carrying flows which join on that output link. Let r_i be the rate of fan-in link i and N_i be the number of flows that share both input link i and the output link. Let $\theta_i = \frac{N_i-1}{r_i} MTU_H \geq 0$ (since $N_i \geq 1$) and assume, without the loss of generality, that input links are numbered in the increasing order of θ_i, from 1 to I.

Since during any busy period there can be at most one packet from each flow in the queue, due to packetization effects [1] the envelope for the input link i is $\alpha_i(t) = \min(N_i MTU_H,\, r_i t + MTU_H)$. As a result the envelope for the traffic aggregate at the output link queue is:

$$\begin{aligned}
\alpha(t) &= \sum_{i=1}^{I} \alpha_i(t) = \sum_{i=1}^{I} \min_t(N_i MTU_H,\, r_i t + MTU_H) = \\
&= \sum_{i=1}^{I} \Big(MTU_H + \min_t((N_i - 1)MTU_H,\, r_i t) \Big) = \\
&= I\, MTU_H + \sum_{i=1}^{I} r_i \min_t(\theta_i, t) = I\, MTU_H + \alpha'(t)
\end{aligned}$$

where

$$\alpha'(t) = \sum_{i=1}^{I} r_i \min(\theta_i, t)$$

The $\alpha'(t)$ function – illustrated in Fig. 2 – is piece-wise linear, with the linear segment with $t \in [\theta_{k-1}, \theta_k]$ having a slope $R_k = r_k + r_{k+1} + ... + r_I$, for $k = 1$ to I, with $\theta_0 = 0$, $R_I = r_I$ and $R_{I+1} = 0$. The ordinates at discontinuity points θ_k are $MTU_H \sum_{i=1}^{k}(N_i - 1) + R_{k+1}\theta_k$, with a maximum value of $MTU_H \sum_{i=1}^{I}(N_i - 1)$ at θ_I.

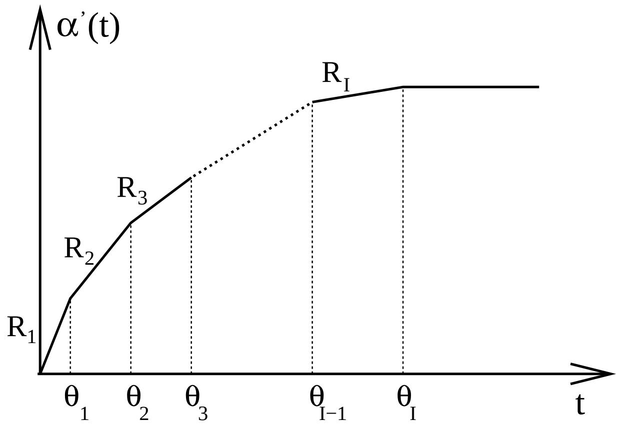

Figure 2: The $\alpha'(t)$ function

Since the node offers service guarantees in the shape of a rate-latency curve $\beta_{r\,,\,\tau}(t)$ (with r being the physical rate of the output link and $\tau = \frac{MTU_L}{r}$) the maximum backlog is:

$$\begin{aligned} maxbacklog &= \max_t \left(\alpha(t) - \beta_{r\,,\,\tau}(t)\right) = \\ &= I\, MTU_H + \max_t \left(\alpha'(t) - \beta_{r\,,\,\tau}(t)\right) \end{aligned}$$

The value $\max_t \left(\alpha'(t) - \beta_{r\,,\,\tau}(t)\right)$ depends on two factors: On one hand, on the relative order between server latency τ and the inflection points θ_k, and on the other hand on the relation between rate r and rates R_k. Since the values for the rates R_k are decreasing for increasing values of k, we can distinguish three cases:

1. $\tau < \theta_k$ and $R_k \geq r > R_{k+1}$. In this situation the maximum will occur at discontinuity point θ_k for the largest value of k for which the three inequalities are satisfied. Since the server rate is non-zero at this point, the maximum is $MTU_H \sum_{i=1}^{k}(N_i - 1) + R_{k+1}\theta_k - r\,(\theta_k - \tau)$.

2. $\tau > \theta_k$ and $r \geq R_{k+1}$. In this situation the maximum occurs at instant τ and – since the server rate is zero at that point – it is equal to $\alpha'(\tau) = MTU_H \sum_{i=1}^{j}(N_i - 1) + R_{j+1}\theta_j$, where j is the index that satisfies $\theta_j \leq \tau < \theta_{j+1}$ (with the convention $\theta_{I+1} = \infty$).

3. If $\tau > \theta_k$ and $R_{k+1} > r$, then the maximum occurs at an instant later than θ_k . Since $R_{I+1} = 0$ and $r > 0$ one of the be above cases must apply i.e. it maximum will occur either at instant θ_i with $i > k$ or at instant τ.

Similarly to backlog, the maximum jitter can be computed as the maximum *horizontal* distance between envelope $\alpha(t)$ and the service curve $\beta_{r\,,\,\tau}(t)$ (see [1]).

These numerical computations can be implemented in an algorithm that at each link along a flow path computes the maximum required buffer capacity and maximum jitter encountered at that link. The computation at a node simply loops through the values of θ_k and τ sorted in increasing order and, at each such envelope/service curve inflection point, computes the horizontal and vertical distance from the envelope to the service curve. The maximum backlog and respectively maximum jitter at the node are these values accumulated throughout the loop.

Instead of performing the numerical computations described above, we can use a closed form approximation formula by assuming that the node has as strict service curve the rate-latency curve $\beta_{r^*,\tau}$, with $r^* = \min_i (r, r_i)$. In this case the maximum backlog occurs at θ_I and has a value of $MTU_H \, N - r\,(\theta_I - \tau)^+$. In this case the maximum jitter is upper bounded by:

$$MTU_H \left(\frac{N}{r} - \frac{N_I - 1}{r_I} \right) + \frac{MTU_L}{r}$$

Along a flow path an upper bound for the end-to-end queuing delay is the sum of the per-node queuing delay bounds along the path. For example in the case of the closed form approximation the end-to-end queuing delay is upper bounded by:

$$MTU_H \sum_{j=src}^{dst} \left(\frac{N^j}{r_j^*} - \frac{N_i^j - 1}{r_i^j} \right) + \sum_{j=src}^{dst} \frac{MTU_L}{r_j}$$

where N^j is the number of flows sharing link j, N_i^j is the number of flows sharing both output link j and fan-in link i, and $r_j^* = \min_i (r_j, r_i^j)$ is the smallest capacity among output link j and all its fan-in links.

5 Conclusions

In this paper we presented a method for guaranteeing deterministic worst case queuing delays bounds. By extending previous results to networks with different link sizes and carrying different types of traffic consisting of packets of different sizes, we showed how – by corespondingly shaping flows at network ingress – it is possible to guarantee worst case queuing delays. As both the queuing delay bounds and the flow shaping parameters depend only on routing topology metrics (the so called Route Interference metrics) we illustrated the algorithmic steps necessary for computing these quantities for each flow using network calculus concepts of input traffic envelope, service curve and known network calculus bounds for maximum jitter and maximum required buffer capacity.

References

[1] Jean-Yves LeBoudec, Patrick Thiran, "Network calculus: A theory of deterministic queuing systems for the Internet". Online version, July 2002.

[2] I. Chlamtac, A. Farago, H. Zhang, A. Fumagalli, "A deterministic approach to the end-to-end analysis of packet flows in Connection-oriented networks" in *IEEE/ACM Transactions on Networking, Vol. 6, No. 4, Aug 1998.*

[3] J.Y. LeBoudec, G. Hebutrene, "Comments on A Deterministic Approach to the End-to-End Analysis of Packet Flows in Connection Oriented Networks" in *IEEE/ACM Transactions on Networking, Vol. 8, No. 1, Feb 2000.*

[4] H. Zhang, "A note on deterministic end-to-end delay analysis in connection oriented networks" in *Proc. of IEEE ICC'99, Vancouver, p. 1223-1227, 1999.*