A Delay Bound for a Network with Aggregate Scheduling

Jean-Yves Le Boudec and Felix Farkas

Abstract

Given the increasing demand for services that can guarantee a maximum end-to-end delay in the Internet, it is important to understand the behaviour of different existing networks. We consider networks that implement scheduling per aggregate flow. We give a worst case bound that is valid for any network topology given the maximum hop count for every flow and the utilization at every link. Unfortunately, this bound is finite only for small utilization. Using the context of arrival and service curves we can improve the result for the case of a ring.

1 Introduction

We consider a network serving a high-priority service class in an aggregate way. All packets belonging to the high-priority class are seen at every node as a single, aggregate flow. This is a model for the expedited forwarding (EF) service class, a delay priority mechanism proposed in the context of differentiated services.

We assume that every flow is individually policed at the network boundary by one leaky bucket, so that it is made conformant to an arrival curve of the form $a(t) = rt + b$. This is the only place where flows are handled separately; after entering the network, flows are seen as aggregates.

The high priority aggregate is processed at every node according to some scheduling policy. The scheduler handles all priority traffic as one flow. We follow [1] and assume that every node can be abstracted by a service curve property, of the form $\sigma(t) = (St - \beta)^+$ [2, 3, 4]. A simplified way of thinking of the service curve property is to assume that during every period of duration t for which there is some backlog of high priority traffic at the node, the amount of service received is at least $\sigma(t)$. For example, if the node implements head of the line, non-preemptive priority queuing, then S is the rate of the outgoing link and $\beta = \text{MTU} \times S$.

We give in this paper a bound for the end-to-end queuing delay, which is valid for any network topology, provided that we know the maximum number of hops undergone by any flow, and the maximum intensity of high priority traffic. However, for large utilization factors, it becomes infinite. Explosion of bounds for aggregate scheduling is not new [5, 6]; it is also known that good bounds for aggregate scheduling may depend on complex, global conditions [7, 8]. In this working document, we do not have results about the tightness of our bound.

However, we show that in the specific example of a ring, better bounds do exist.

The paper is organized as follows. In Section 2 we give a delay bound for a general network topology. In Section 3 we study a more complex bound for the ring. Our conclusions are presented in Section 4.

2 A delay bound for a general network topology

Consider a network as defined in the Introduction, and call h the maximum hop count across all flows in the network. To simplify the mathematics here we assume that the service curve guaranteed to the aggregate high priority traffic is the same at all nodes; call it $\sigma(t) = (St - \beta)^+$.

We define b_{tot} as a bound on the sum of the burst tolerance of all individual flows using any particular link. More specifically, assume that for an individual flow i, the arrival curve enforced at the network access point is $a_i(t) = r_i t + b_i$. For a link l, call $\mathcal{S}(l)$ the set of flows that constitute the aggregate high priority traffic on link l. Then we must have:

$$\sum_{i \in \mathcal{S}(l)} b_i \leq b_{tot}$$

We also assume that the total average traffic intensity on any link is bounded by αS, namely we must have, for each link l

$$\sum_{i \in \mathcal{S}(l)} r_i \leq \alpha S$$

The parameter α is thus a bound on the high priority class utilization factor.

Also, we define

$$\Delta = \frac{\beta}{S}$$

In [1], the parameter b_{tot} is not used. Instead, the author assumes that all flows have the same arrival curve $a_i(t) = r_{min}t + b_{max}$, which is less general. We can map our result to the notation in [1] by letting

$$b_{tot} = \alpha S \frac{b_{max}}{r_{min}}$$

Finally, we assume that the peak rate of all incoming traffic at any link is bounded by some constant C. The model in [1] corresponds to $C = +\infty$. For a router with large internal speed and buffering only at the output, C is the sum of the bit rates of all incoming links.

Theorem 1 *If*

$$\alpha < \frac{C}{(C-S)(h-1)+S} \tag{1}$$

then a bound on the end-to-end delay for high priority traffic is

$$D = \frac{h}{1-u\alpha(h-1)}\left(\Delta + u\frac{b_{tot}}{S}\right) \tag{2}$$

with $u = \frac{C-S}{C-\alpha S}$.

Proof: The proof consists in two parts. Part 1 shows that if a finite bound exists, then the formula in the Theorem is true. Part 2 shows that a finite bound exists.

Part 1: We assume that a finite bound exists. Call D' the worst case delay across all nodes, and h the maximum hop count for any flow in the network. Consider a buffer at some link l. An arrival curve for the aggregate traffic is

$$a(t) = \min\left(Ct, \sum_{j\in\mathcal{S}(l)} a_j(t + (h-1)D')\right)$$

The former part in the formula is because the total incoming bit rate is limited by C; the latter is because any flow reaching that node has undergone a delay bounded by $(h-1)D'$. Thus

$$a(t) \leq a'(t) = \min\left(Ct, \alpha St + b'\right)$$

with $b' = b_{tot} + \alpha S(h-1)D'$.

A bound on the delay at our node is given by the horizontal deviation between the arrival curve $a'(t)$ and the service curve $\sigma(t) = (St-\beta)^+$ [2, 3]. The alert reader will enjoy doing the computation and will find that a bound on the delay is $\Delta + u\frac{b'}{S}$. Since D' is the worst case delay, we must have $D' \leq \Delta + u\frac{b'}{S}$ (to see why, simply consider a node l where the worst case delay D' is attained). Thus, we must have

$$D' \leq \Delta + \frac{ub_{tot}}{S} + u\alpha(h-1)D' \tag{3}$$

Define function g by $g(D')$ = the right hand-side in Equation (3), so that we can rewrite the equation as: $D' \leq g(D')$. The condition

$$\alpha < \frac{C}{(C-S)(h-1)+S}$$

means that $u\alpha(h-1) < 1$; a simple inspection of g shows that there is a unique fix-point D_1, namely, a value such that $g(D_1) = D_1$. Furthermore, if D' is finite, then $D' \leq g(D')$ implies that $D' \leq D_1$. Thus D_1 is a bound for the delay at any hop. The end-to-end delay is thus bounded by hD_1, which, after some algebra, provides the required formula.

Part 2: We now prove that a finite bound exists. We use the time-stopping method in [9]. For any time $\tau > 0$, consider the virtual system made of the original network, where all sources are stopped at time τ. This network satisfies the assumptions of part 1, since there is only a finite number of bits at the input. Call $D'(\tau)$ the worst case delay across all nodes for the virtual network indexed by τ. From the above derivation we see that $D'(\tau) \leq D_1$ for all τ. Letting τ tend to $+\infty$ shows that the worst case delay at any node remains bounded by D_1.

$\diamond$

Our bound explodes when α tends to $\frac{C}{(C-S)(h-1)+S}$, which is very unfortunate. It is not clear whether the delay does become unbounded for general topologies or whether it is an artifact of our bound. Explosion of bounds for aggregate scheduling is not new, and is encountered for example in [5, 6].

Note that if C is close to S, then $\frac{C}{(C-S)(h-1)+S}$ is close to 1. This is the case for a high speed add-drop ring network, where the bit rate is much higher for the ring than for the input links. In that case, u tends to 0 and

$$D = h(\Delta + u\frac{b_{tot}}{S}) + o(u)$$

with $u = \frac{C-S}{C-\alpha S}$; this confirms that the explosion does not take place when C is close to S.

It is also known that, in practice, worst case delays for FIFO networks are usually much less than predicted by simple bounding techniques. For example, in [7, 8], it is assumed that individual flows are shaped at network entry in such a way that the spacing between packets is at least equal to the route interference number (RIN) of the flow. The RIN is defined as the number of occurrences of a flow joining the path of some other flow. Then the end-to-end delay is bounded by the time to transmit a number of packets equal to the RIN. This usually results in much smaller bounds than the bounds here, albeit at the expense of more restrictions on the routes. If this is confirmed, this would indicate that a technology like MPLS is more suited for providing the guaranteed service than expedited forwarding.

3 Stability of a Ring Network with Service Curves Guarantees

We consider a unidirectional ring with n nodes each of them guaranteeing a service curve $\sigma(t) = (St - \beta)^+$. In this ring there are n (r, b) constrained flows going through n hops. Each flow enters the ring at a different node and uses the whole ring as presented in Fig. 1. For this ring, the utilization factor is $\alpha = \frac{nr}{S}$.

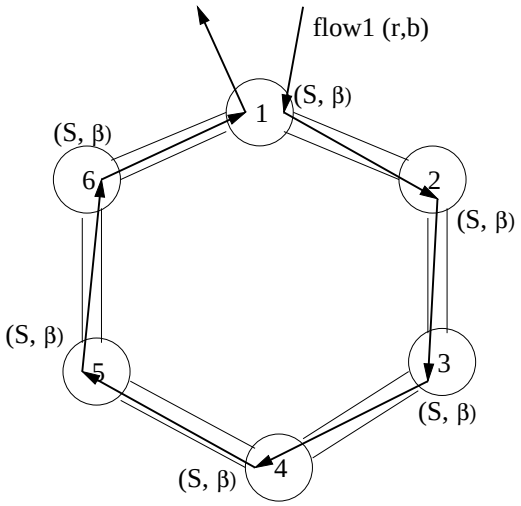

Figure 1: Ring with six nodes; on each node a flow enters the ring but only flow 1 is drawn

For this case we can improve the result obtained in the previous section. We can obtain stability in the network for a larger utilization value than the one obtained from Equation 1.

Tassiulas, in [10], proved that any work conserving policy stabilizes a ring with shapers, where all the flows are (r_i, b_i) constrained, given that the maximum load on every link is less than one. However the proof of that result does not hold for the case of a network with nodes which guarantee a minimum service curve like in the model studied in the previous section.

We consider the following result proven in [1]:

Statement 1 *If all flows at the entry of a node are constrained by leaky buckets with rate r_i and burst b_i and the node is guaranteeing a service curve $\sigma(t) = (St - \beta)^+$ then a bound on the burstiness of any flow at the*

output of the node is:

$$b'_i = b_i + \frac{r_i}{S}\left(\sum_i b_i + \beta\right), \tag{4}$$

given that the data in the queue are served in FIFO order and $\sum_i r_i \leq S$.

By bound on the burstiness, b'_i, we mean that flow i is constrained by the arrival curve $r_i t + b'_i$.

Proof: We will apply network calculus [3]. An arrival curve for the total traffic entering the node is given by:

$$a(t) = rt + b,$$

with $r = \sum_i r_i$ and $b = \sum_i b_i$.

Thus a bound on the delay across the node, given by the horizontal deviation between $a(t)$ and the service curve $\sigma(t)$ is equal to

$$D = \frac{b + \beta}{S} \tag{5}$$

Every individual flow receives a delay bounded by D. It follows immediately that an arrival curve on the output of flow i is

$$a'_i(t) = a_i(t + D) = r_i t + b_i + \frac{r_i}{S}(b + \beta) \tag{6}$$

◇

Statement 2 *If* $\alpha < \frac{2}{n+1}$ *then the bound on the burstiness on any flow after* i *hops is:*

$$b^{(i)} = \frac{b\left(1 + \frac{(i-n)(n+1)}{2}\frac{r}{S}\right) + i\frac{r}{S}\beta}{1 - \frac{n(n+1)}{2}\frac{r}{S}}. \tag{7}$$

Proof: The proof consists in two parts. Part 1 shows that if a finite bound exists, then the formula in the statement is true. Part 2 that a finite bound on the bursts exists.

Part 1: Given the symmetry of the ring model, we denote with $b^{(i)}$ a bound on the burstiness of any flow after i hops. Assuming that these bounds are finite from statement 1 we can write the following system of inequations:

$$\begin{cases} b^{(1)} \leq b^{(0)} + \frac{r}{S} * (\nu + \beta) \\ b^{(2)} \leq b^{(1)} + \frac{r}{S} * (\nu + \beta) \\ \dots \\ b^{(n)} \leq b^{(n-1)} + \frac{r}{S} * (\nu + \beta) \end{cases}$$

where $\nu = \sum_{i=0}^{n} b^{(i)}$ and $b^{(0)} = b$.

Solving the system, if $\alpha \leq \frac{2}{n+1}$ we obtain:

$$\nu \leq \frac{(n+1)b + \frac{r}{S}\frac{n(n+1)}{2}\beta}{1 - \frac{n(n+1)}{2}\frac{r}{S}}$$

and thus

$$b^{(i)} \leq \frac{b\left(1 + \frac{(i-n)(n+1)}{2}\frac{r}{S}\right) + i\frac{r}{S}\beta}{1 - \frac{n(n+1)}{2}\frac{r}{S}}. \tag{8}$$

Part 2: We prove that a finite bound on the burstiness exists. We use the time-stopping method in [9]. For any time $\tau > 0$, consider the virtual system made of the original ring, where all sources are stopped at time τ. This network satisfies the assumptions of part 1, since there is only a finite number of bits at the input. Call $B(\tau)$ the worst case bound which occurs in the ring for the virtual network indexed by τ. From the above derivation we see that $B(\tau) \leq b^{(n)}$ for all τ. Letting τ tend to $+\infty$ shows that the worst case bound on the burstiness at any node remains bounded by

$$b^{(n)} = b\frac{1 + \frac{nr}{S}\beta}{1 - \frac{n(n+1)}{2}\frac{r}{S}}.$$

◇

It can be easily seen that this result is better than the one which is obtained from equation (1):

$$\alpha < \frac{1}{n-1} \tag{9}$$

since $h = n$.

4 Conclusion

We have shown that a ring with n nodes, and n (r, b) constrained flows going through the n hops is stable if the utilization of the a node satisfies the condition:

$$\alpha < \frac{2}{n+1} \tag{10}$$

Using the same methods, the result obtained in this paper should be generalized for a network with a general topology.

References

[1] Anna Charny and Jean-Yves Le Boudec. Delay bounds in networks with aggregate scheduling. available at ftp://ftpeng.cisco.com/acharny/qofis2000.ps.

[2] R.L.Cruz. Quality of service guarantees in virtual switched networks. *IEEE Journal on Selected Areas of Communications*, pages 1048–1056, August 1995.

[3] J.-Y. Le Boudec. Application of network calculus to guaranteed service networks. *IEEE Trans. on Information Theory*, 44:1087–1096, May 1998.

[4] R. Agrawal, R.L. Cruz, C. Okino, and R. Rajan. Performance bounds for flow control protocols. *IEEE/ACM Trans. on Networking*, 7(3):310–323, June 1999.

[5] A.K. Parekh and R.G. Gallager. A generalized processor sharing approach to flow control in integrated services:the multiple node case. *IEEE/ACM Trans. on Networking*, 2(2):137–150, April 1994.

[6] M. Andrews, B. Awerbuch, A. Fernandez, J. Kleinberg, T. Leighton, and Z. Liu. Universal stability results for greedy contention resolution protocols. *37th Annual Symposium on Foundations of Computer Science (FOCS'96), Burlington VT,*, October 1996.

[7] J.-Y. Le Boudec and G. Hebuterne. Comment on deterministic approach to the end-to-end analysis of packet flows in connection oriented networks. *IEEE/ACM Trans. on Networking*, February 2000.

[8] I. Chlamtac, A. Faragó, H.Zhang, and A. Fumagalli. A deterministic approach to the end-to-end analysis of packet flows in connection oriented networks. *IEEE/ACM Trans. on Networking*, 6(4):422–431, August 1998.

[9] C.S. Chang. *A filtering theory for Communication Networks*. Sprigne Verlag, 1999.

[10] Leandros Tassiulas. Any work-conserving policy stabilizes the ring with spatial re-use. *IEEE/ACM Trans. on Networking*, 4(2):205–208, 1996. Seminare Maurey-Schwartz (1975-1976).