The ‘Virtual Wire’ Per Domain Behaviour

Analysis and Extensions

Guven Mercankosk and John Siliquini

TELECOMMUNICATIONS, ELECTRONICS AND NETWORKING RESEARCH GROUP

Department of Electrical and Electronic Engineering

The University of Western Australia
35 Stirling Highway, CRAWLEY WA 6009, Australia

Abstract: This paper provides an analysis and extensions of the Virtual Wire Per Domain Behaviour as defined by the Internet Engineering Task Force. A formalised model of the Virtual Wire Per Domain Behaviour is developed by explicitly identifying key timing and decision variables and associated design parameters. The necessary and sufficient conditions for creating and establishing a Virtual Wire flow is derived.

Key words: Diffserv, QoS, Virtual Wire

1. INTRODUCTION

With the tremendous growth of the Internet in the past few years, and the wide variety of new applications that have appeared, the convergence of other networks -- telephone, radio, and television -- to the Internet is underway. Moreover, network traffic has increased as the number of users and applications has increased. In accordance with Moore's Law computer systems are capable of transferring more data than ever. The question is whether increasing bandwidth -- the data carrying capacity of the network -- is sufficient to accommodate these increased demands. The answer is no, it is not. Internet traffic has not only increased, but it has changed in character. New applications have new service requirements, and as a result the Internet needs to change as well.

Some of the new breed of Internet applications are multimedia and require significant bandwidth as well as having strict timing requirements on

---

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35673-0_28

C. McDonald (ed.), *Converged Networking*

information delivery. These require network services beyond the simple "best-effort" service that current IP networks deliver. IP Telephony is today's "killer application." More than any other, the desire to provide telephone service over the Internet is driving the convergence of the telephone and Internet industries.

In an attempt to enable the Internet to support real time and multimedia applications, the Internet Engineering Task Force (IETF), as part of its Differentiated Services (diffserv or DS) architecture, describes a Per Hop Behavior (PHB) called Expedited Forwarding (EF) intended for use in building a scalable edge-to-edge service that appears to the end points like an un-shared, point-to-point connection or `Virtual Wire' [1]. The document [2] provides a set of specifications necessary on the aggregate traffic (in DS terminology, a Per Domain Behavior or PDB) in order to meet these requirements and thus defines a new PDB, the `Virtual Wire' PDB of some fixed capacity. For scalability, it does not require `per-flow' state to exist anywhere other than the ingress and egress routers. Despite the lack of per-flow state, if the aggregate input rates are appropriately policed and the EF service rates on interior links are appropriately configured, the edge-to-edge service supplied by the DS-domain will be indistinguishable from that supplied by a dedicated wire between the end points. This paper provides an analysis of `Virtual Wire' PDB with limited extensions and formalizes a model for the `Virtual Wire' PDB by explicitly identifying key timing and decision variables and associated design parameters. It derives the necessary and sufficient conditions for creating and establishing a `Virtual Wire' flow. It also provides methods for quantifying design parameters and setting decision parameters.

2. DESCRIPTION OF ‘VIRTUAL WIRE’ PDB WITH EXTENSIONS

An illustration of the model of a Virtual Wire transfer over a DS- domain is given in Figure 1. At the ingress end of the micro flow, the model contains an ingress link (S to I) and a DS-domain ingress border router (at I). Over the DS-domain (I to E), the packets from the micro flow are subject to random delays, according to EF-PHB service. At the egress end of the flow, the remaining components of the model are a DS-domain egress border router (at E) and an egress link (E to D). The ingress link and the egress link are assumed to run at the same constant rate as the virtual wire rate R with negligible line clock jitter.

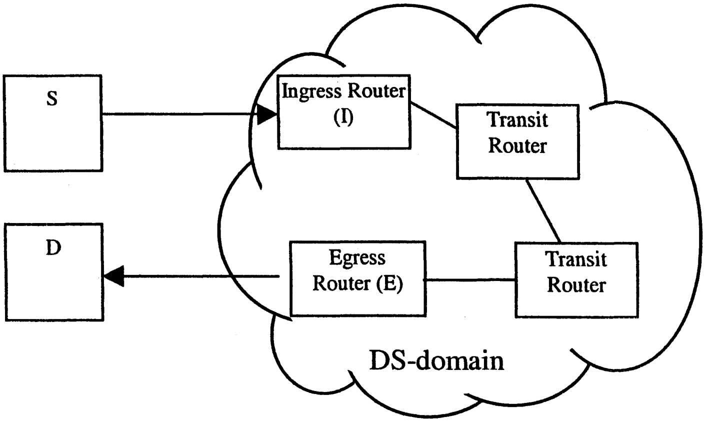

Figure 1. An illustration of virtual wire transfer over a DS-domain

2.1 Ingress edge of DS-domain

A periodic flow of packets of size S is sourced at a rate R on the ingress link. Accordingly, one packet is presented at the DS-domain ingress border router in every $T=S/R$ time units. The number of complete packets that are in transit in the DS-domain at time t is denoted by $N_{DS}(t)$. Without loss of generality, we choose the time origin such that the last bit of the first packet presented at the ingress border router is at time T. Let t_k represent the time at which the last bit of the k^{th} packet is presented to the DS-domain at the ingress border router, then

$$t_k = kT \tag{1}$$

2.2 Over the DS-domain

Once a packet that belongs to a micro flow is presented at the ingress border router, it hops from one EF router to another along its path over the DS-domain before reaching the egress border router. There are two main components for the delay experienced by a packet over the DS-domain, namely the propagation delay and the queueing delay. The propagation delay refers to the time it would take for a packet to traverse through the DS-domain if the packet experienced no queueing delay along its path. Accordingly, the propagation delay, d_p, is fixed and includes the time necessary to process a packet at routers. Let d_k denote the total queueing delay experienced by the k^{th} packet along its path. It is assumed that the queueing delay experienced by each packet of a micro flow over the DS-domain is statistically bounded by a known value, *D*, that is

$$0 \leq d_k < D \tag{2}$$

The value, *D*, can be obtained by some (1-α) quantile of the total queueing delay. Note that the value *D* refers to the difference between the best and the worst case expectation of the packet transfer delay. The best case is equal to d_p and the worst case is equal to $d_p + D$, a value likely to be exceeded with a probability less than α. Therefore, the value *D* is a measure of variation of the delay distribution and may be referred to as expected maximum jitter.

Also let τ_k denote the time at which the last bit of the k^{th} packet reaches the egress border router. So we have

$$\tau_k = t_k + d_k + d_p \tag{3}$$

2.3 Egress edge of DS-domain

The packets that arrive at the egress border router are placed in a micro flow specific buffer of size B_E. The number of complete packets in the egress buffer at time *t* is denoted by $N_E(t)$. Similar to the ingress link, one packet is transmitted onto the egress link in every *T* time units. It is clear that the timing structure of the sequence $\{\tau_k\}$ is not necessarily the same as that of the sequence $\{t_k\}$. A commonly used strategy for recovering the initial timing structure is to delay the transmission of the first packet onto the egress link by a fixed amount of time, say D_ξ, and set the size of the egress buffer B_E such that the continuity of the data flow is maintained when packets are transmitted onto the egress link using the very same timing structure $\{t_k\}$[3][4]. Let ξ_k denote the time at which the first bit of the k^{th} packet is transmitted onto the egress link. Accordingly, we have

$$\xi_k = T + d_1 + d_p + D_\xi + (k-1)T \tag{4}$$

3. MICRO FLOW RELATIONS

In this section, we express key system quantities in terms of the virtual wire rate and the equalization delay D_ξ. We assume that buffers over the

DS-domain and at the egress border are sufficiently large that they do not overflow. We also assume that the initial delay is made sufficiently large so that when the transmission of a packet onto the egress link is completed, there is always another packet in the egress buffer ready for transmission. Subsequently in section 4, we determine the bounds on the buffer sizes and the initial delay.

3.1 Packets presented and delivered

The number of packets presented to the *DS*-domain at the ingress border in the interval $(0,t]$ is given by the expression

$$\lfloor t/T \rfloor \tag{5}$$

and illustrated in Figure 2 (for real x, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x).

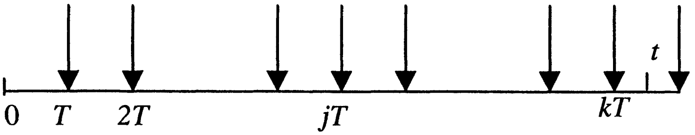

Figure 2: Packet presentation instants at the ingress

Similarly, the number of packets transmitted onto the egress link in the interval $(d_1 + d_p + D_\xi, t]$ is given by the expression

$$\lfloor (t - (d_1 + d_p + D_\xi))/T \rfloor \tag{6}$$

and illustrated in Figure 3.

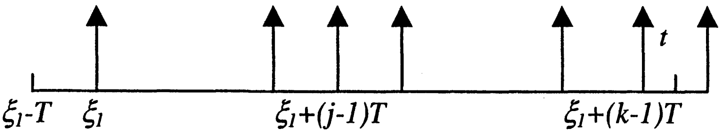

Figure 3: Packet departure instants at egress

3.2 Egress router buffer fill

The arrival instants at the egress border buffer partition the time axis into contiguous intervals of variable length. That is, for a given time t, there exists an integer such that

$$\tau_n \le t < \tau_{n+1} \tag{7}$$

where τ_n is defined as in (3) and illustrated in Figure 4.

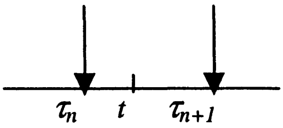

Figure 4: Arrivals at egress border buffer

The fill level of the egress border buffer at time t is equal to a difference between the total number of packets that have arrived at the egress buffer and the number of packets transmitted onto the egress link. Equivalently,

$$N_E(t) = n - \left\lfloor (t - (d_1 + d_p + D_\xi))/T \right\rfloor \tag{8}$$

which follows from (6) and (7). The fill level $N_E(t)$ is a non-increasing function of t between the two successive arrivals at the egress border buffer or over the interval $[\tau_n, \tau_{n+1})$.

3.3 Packets in transit

Packets in transit are the packets that have been presented to the *DS*-domain at the ingress router but have not yet reached the egress border buffer. So, the number of packets in transit at a given time t is given by

$$N_{DS}(t) = \left\lfloor t/T \right\rfloor - n \tag{9}$$

which follows from (5) and (7).

3.4 End-to-end delay

The initial fill for a particular flow is the number of packets presented to the *DS*-domain just ahead of the transmission of the first packet onto the egress link minus the first packet. Noting that $N_{DS}(\xi_1)$ is the number of packets in transit at time ξ_1 and $N_E(\xi_1)$ is the egress border buffer fill at the same time ξ_1, we have

$$N_{DS}(\xi_1) + N_E(\xi_1) = \left\lfloor \xi_1/T \right\rfloor - 1 = \left\lfloor \left(d_1 + d_p + D_\xi\right)/T \right\rfloor \tag{10}$$

Next, we derive an expression for the end-to-end delay over the *VW*. The end-to-end delay experienced by the k^{th} packet is the time from when the last bit of the k^{th} packet is presented to the *DS*-domain to when the first bit is transmitted onto the egress link. It can be found by rearranging (4) and substituting t_k for kT as

$$\xi_k - t_k = \xi_1 - T = d_1 + d_p + D_\xi \tag{11}$$

which shows that the end-to-end delay is constant as it would be over a dedicated wire. Adding and subtracting $\lfloor \xi_1 / T \rfloor T$ and rearranging terms lead to

$$\xi_k - t_k = \left(\xi_1 - \lfloor \xi_1 / T \rfloor T\right) + \left(N_{DS}(\xi_1) + N_{DS}(\xi_1)\right)T \tag{12}$$

where we also make use of (4) for *k=1* and (10). Since the packet presentation to the *DS*-domain and packet transmission onto the egress link are periodic with period *T*, the first term on the right hand side of the (12) can be considered as the phase difference of the egress link relative to the ingress link which is

$$0 \le \phi_R = \xi_1 - \lfloor \xi_1 / T \rfloor T < T \tag{13}$$

4. CONTINUITY OF FLOW

In section 3, we assume sufficiently large buffers and initial total fill so that the data flow continuity could be maintained. In this section, we derive the necessary and sufficient conditions for data flow continuity.

Continuity of data flow requires that every packet presented to the *DS*-domain at the ingress router at fixed intervals has to be transmitted onto the egress link after a fixed delay but again at the same fixed intervals, without being subject to packet loss or interruption during the life time of the flow. Packet loss occurs if a packet has to be discarded due to lack of storing capacity upon arrival at the egress border router. On the other hand, the flow is interrupted if a transmission cannot be started due to unavailability of a packet at the egress border buffer. In the latter case, the buffer is said to be underflowing.

In what follows, we show that having a bounded queueing delay over the *DS*-domain is sufficient to have limits on the maximum fill level that the egress buffer could reach and the minimum fill level that it could fall. These limits can then be used to derive the necessary and sufficient conditions to avoid both overflows and underflows of the egress buffer, i.e. to maintain the continuity of flow.

4.1 Egress buffer overflow

As the fill level $N_E(t)$ is a non-increasing function of over the interval $[\tau_n, \tau_{n+1})$, it has its peak level at the start of the interval, when the n^{th} packet arrives. Substituting τ_n for t in (8), we have

$$N_E(t) \le n - \left\lfloor \left(t_n + d_n + d_p - \left(d_1 + d_p + D_\xi\right)\right)/T \right\rfloor \tag{14}$$

where we make use of (3). Again from the identity given in (1), we see that t_n/T is an integer. This leads to

$$\begin{aligned} N_E(t) &\le -\left\lfloor \left(d_n - d_1 - D_\xi\right)/T \right\rfloor \\ &\le \left\lceil \left(d_1 + D_\xi - d_n\right)/T \right\rceil \\ &< \left(d_1 + D_\xi - d_n\right)/T + 1 \end{aligned} \tag{15}$$

where for real x, $\lceil x \rceil$ denotes the largest integer greater than or equal to x. An upper bound for $N_E(t)$ can be found by considering a packet which experiences no queueing delay over the DS-domain and the case where the first packet experiences a queueing delay arbitrarily close to the maximum. That is,

$$\begin{aligned} N_E(t) &< \left(D + D_\xi\right)/T + 1 \\ &\le \left\lceil \left(D + D_\xi\right)/T \right\rceil \end{aligned} \tag{16}$$

Therefore, the maximum value $N_{E,\max}$ that the fill level $N_E(t)$ of the egress buffer can reach satisfies the relation

$$N_E(t) \le N_{E,\max} = \left\lceil \left(D + D_\xi\right)/T \right\rceil \tag{17}$$

So if we choose the size of the egress buffer B_E greater than or equal to $N_{E,\max}$, then the egress buffer never overflows. Otherwise whenever the fill level $N_E(t)$ reaches $N_{E,\max}$, that is

$$N_E(t_{\max}) = B_{E,\max} > B_E \tag{18}$$

the egress buffer overflows. This leads us to our first dimensioning rule

$$\left\lceil \left(D + D_\xi\right)/T \right\rceil \le B_E \tag{19}$$

which is necessary and sufficient to avoid egress buffer overflow.

4.2 Egress link starvation

Since the fill level $N_E(t)$ is a non-increasing function of t over the interval $[\tau_n, \tau_{n+1})$, it falls to its minimum level after the last, say the k^{th}, packet is transmitted onto the egress link but before the $(n+1)^{st}$ packet arrives. Substituting τ_{n+1} for t in (8), we obtain

$$N_E(t) \geq n - \left\lfloor \left(t_{n+1} + d_{n+1} + d_p - \left(d_1 + d_p + D_\xi\right)\right)/T \right\rfloor \tag{20}$$

where we again make use of (3). Using the fact that is an integer and the identity (1), we are led to

$$\begin{aligned} N_E(t) &\geq -1 - \left\lfloor \left(d_{n+1} - d_1 - D_\xi\right)/T \right\rfloor \\ &\geq \left\lceil \left(d_1 + D_\xi - d_{n+1}\right)/T \right\rceil - 1 \\ &\geq \left(d_1 + D_\xi - d_{n+1}\right)/T - 1 \end{aligned} \tag{21}$$

Over all packet transfers, a lower bound for $N_E(t)$ can be found this time by considering a packet which experiences a queueing delay arbitrarily close to the maximum and the case where the first packet experiences no queueing delay over the *DS*-domain. That is,

$$\begin{aligned} N_E(t) &> \left(D_\xi - D\right)/T - 1 \\ &\geq \left\lfloor \left(D_\xi - D\right)/T \right\rfloor \end{aligned} \tag{22}$$

Therefore, the minimum value $N_{E,\min}$ that the fill level $N_E(t)$ of the egress buffer can fall satisfies the relation

$$N_E(t) \geq N_{E,\min} = \left\lfloor \left(D_\xi - D\right)/T \right\rfloor \tag{23}$$

Note that the minimum value $N_{E,\min}$ is determined at a time instant such that there will be at least one packet arrival ahead of the next scheduled transmission onto the egress link. So if we delay the transmission of the first packet onto the egress link longer than the maximum queueing delay possible over the *DS*-domain, then the fill level of the egress buffer never falls below zero. Otherwise, whenever circumstances occur for the fill level $N_E(t)$ to fall down to $N_{E,\min}$, that is

$$N_E(t_{\min}) = N_{E,\min} < 0 \tag{24}$$

the transmission is interrupted. This leads us to our second dimensioning rule

$$\left\lfloor \left(D_\xi - D\right)/T \right\rfloor \geq 0 \tag{25}$$

which is necessary and sufficient to avoid egress link starvation.

5. ESTABLISHING A VIRTUAL WIRE

In section 4, we list a number of micro flow properties over a virtual wire and express key micro flow quantities in terms of the virtual wire rate R and the equalization delay D_ξ. We observe that the end-to-end delay experienced by each packet of the micro flow is constant as it would be over a dedicated wire and given by

$$\xi_k - t_k = d_1 + d_p + D_\xi \tag{26}$$

We note that the end-to-end delay depends on the queueing delay d_l experienced by the first packet, the propagation delay d_p over the virtual wire, and the equalization delay D_ξ at the egress border router. We will later observe that, in establishing a virtual wire, no explicit knowledge of d_l and d_p is required.

In expressing the key micro flow quantities, we assume that buffers over the DS-domain and at the egress border are sufficiently large that they do not overflow. We also assume that the equalization delay is made sufficiently large so that when the transmission of a packet onto the egress link is completed, there is always another packet in the egress buffer ready for transmission.

Continuity of micro flow requires that every packet presented to the DS-domain at the ingress router at fixed intervals has to be transmitted onto the egress link after a fixed delay but again at the same fixed intervals, without being subject to packet loss or interruption during the life time of the flow. Packet loss occurs if a packet has to be discarded due to lack of storing capacity upon arrival at the egress border router. On the other hand, the flow is interrupted if a transmission cannot be started due to unavailability of a packet at the egress border buffer.

In section 5, we showed that having a bounded queueing delay over the DS-domain, as indicated by (2), is sufficient to have maximum and minimum limits on the egress border buffer fill $N_E(t)$. Accordingly, the egress buffer fill $N_E(t)$ is bounded above by ceiling $\left\lceil (D + D_\xi)/T \right\rceil$ and below by $\left\lfloor (D_\xi - D)/T \right\rfloor$. We then use these limits to derive the necessary and sufficient conditions to maintain the continuity of flow. In relation to avoiding egress link starvation, we consider the situation where the first packet of the flow experiences no queueing delay and some other packet experiences a queueing delay arbitrarily close to the maximum D to obtain

$$\left\lfloor (D_\xi - D)/T \right\rfloor \geq 0 \tag{27}$$

In relation to avoiding egress border buffer overflow however, we consider the situation where the first packet experiences a queueing delay arbitrarily close to the maximum D and some other packet experiences no queueing delay to obtain

$$\left\lceil \left(D + D_{\xi}\right)/T \right\rceil \le B_E \tag{28}$$

As indicated earlier, in determining the maximum queueing delay D, the rarity of events has been taken into account. Accordingly, some level of loss is allowed in case events considered rare do occur. The conditions given in (27) and (28) ensure that there would be no further losses. The smallest value for D_{ξ} that satisfies (27) is D. Setting the initial delay as

$$D_{\xi} = D \tag{29}$$

also minimizes the end-to-end delay over the DS-domain. The resulting end-to-end delay is

$$\xi_k - t_k = \xi_1 - T = d_1 + d_p + D \tag{30}$$

Now substituting (29) into (28), we have

$$\left\lceil 2D/T \right\rceil \le B_E \tag{31}$$

Consequently, the left hand side of the inequality in (10) determines the minimum buffer size for egress border uffer as

$$B_E = \left\lceil 2D/T \right\rceil \tag{32}$$

6. CONCLUSIONS

The Virtual Wire as presented in this paper provides a synthesis of various methods in delivering a particular but versatile solution for the provision of real time services over the Internet. The methods described are scalable, they provide well defined edge-to-edge behavior characteristics independent of other traffic, they offers simple management of a DS-domain and provide a basic building block for expanding the services that can be offered over the Internet

7. ACKNOWLEDGEMENTS

This document is based on and inspired by an earlier joint work with Professor Cantoni. Numerous discussions on related topics with Professor Budrikis have found their way into the arguments presented in this document. We would also like to acknowledge the context provided by Van Jacobson and his team.

This work was supported by the Australian Telecommunications-Cooperative Research Centre (AT-CRC).

8. REFERENCES

[1] V. Jacobson, K. Nichols, K. Poduri, "The 'Virtual Wire' Behavior Aggregate", *IETF draft*, draft-ietf-diffserv-ba-vw-00.txt, March 2000.

[2] V. Jacobson, K. Nichols, K. Poduri, "An Expedited forwarding PHB" *RFC 2598*, June 1999.

[3] Z.L. Budrikis, J.L. Hullet, D.Q. Phiet, "Transient-Mode Buffer Stores for Non Uniform Code TV", *IEEE Transactions on Communication Technology*, Volume COM-19, No6, pp. 913-922, December 1971.

[4] G. Mercankosk, Z. Budrikis, "Establishing a Real-time VBR Connection over an ATM Network," *IEEE Globecom'96*, December 1996.