MIT Open Access Articles

Guidelines for sizing traffic queues in terminals of future protected satcom systems

Massachusetts Institute of Technology

GUIDELINES FOR SIZING TRAFFIC QUEUES IN TERMINALS OF FUTURE PROTECTED SATCOM SYSTEMS[1]

Mu-Cheng Wang, Jun Sun, and Jeffrey Wysocarski
MIT Lincoln Laboratory
Lexington, MA

ABSTRACT

Future Military Satellite Communication systems will feature Time-Division Multiple Access (TDMA) uplinks in which uplink resources will be granted on demand to each terminal by a centralized resource controller. Due to the time-shared nature of the uplink, a terminal will not be constantly transmitting. It will only transmit in its assigned timeslots so as not to cause interference to other terminal transmissions. Packets arriving at a terminal during idle transmission periods will have to be buffered or queued, potentially in a terminal router, else they will be dropped. At the next assigned timeslot these queues will be serviced via a queue scheduling policy that maintains Quality-of-Service (QoS) requirements to the different traffic classes. These queues must be sized large enough to ensure no packet loss when operating in an uncongested state; how large is a function of the distribution of timeslots assigned to the terminal. In this paper, we investigate the relationship between timeslot assignment distributions and queue requirements of a terminal router, providing insight of how to size router queues given an assigned timeslot distribution, or reciprocally, constraints placed on timeslot distribution given a set queue size, in order to avoid packet loss.

1. INTRODUCTION

Future protected MILSATCOM systems will carry IP traffic for voice, video, and data applications. Because many applications are sensitive to delay, jitter, and loss, the terminal must have the ability to process, queue and schedule IP traffic on the uplink in a manner that provides QoS and performance guarantees. One way to keep terminal cost and complexity down is to interface the terminal to a COTS router that provides the needed IP scheduling and QoS functionality. Uplink resources are assigned dynamically to terminals as they are needed using Dynamic Resource Allocation (DRA). DRA can potentially change the rate of the RF channel on *epoch* timescales, where an ep-

och is on the order of a second. The characteristics of the DRA algorithm used in this paper are as described in [4].

There may be a number of inputs into the DRA allocation algorithm, including terminal Committed Information Rates (the rate a terminal is allowed to transmit up to as per its Service Level Agreement), terminal priority, latency and jitter requirements, resource packing complexity, and overall system demand and available resources. The output the DRA algorithm is an assigned set of timeslots for an epoch in which a terminal is allowed to transmit. These inputs change over time and as they do, the number and distribution of timeslots assigned to a terminal will change. Thus, for a COTS router to be a viable option, the router must be able to interface with the terminal modem and control software, as well as with a dynamically varying RF link layer. Furthermore, a terminal's instantaneous data rate is zero during timeslots of an epoch not assigned to it. Packets arriving to the terminal during an idle timeslot must be queued until the next assigned timeslot, otherwise they will be dropped. Ideally, these packets are queued in the terminal router which allows the use of commercial Differentiated Services mechanisms to maintain QoS [1]. At the next transmission slot, these queues are serviced according to a configured scheduling policy, a number of which exist in commercial routers [2-3].

Assuming the desire is to operate with no packet loss due to queue overflow during uncongested states (e.g., when the average arrival rate at the terminal is less than the assigned RF rate (service rate) over an epoch), then the terminal router queue must be sized large enough to queue all packets that arrive during an idle period. Given a distribution of assigned timeslots, the necessary queue size to prevent packet loss can be determined. Alternatively, we can configure the queue sizes on a router to provide a certain queue latency which would then result in constraints on how timeslots assigned to the terminal must be distributed, in order to avoid packet loss. There is an inherent trade-

off between queue size and latency. In this work, we focus on preventing packet loss due to queue overflow and do not try to optimize the queue size to meet specific latency bounds.

While the focus of this paper is analyzing the aforementioned behaviors in a protected MILSATCOM system, the discussion and findings are appropriate for any packet-radio network employing TDMA channel access techniques.

This paper is organized as follows. In Section 2, a brief description of our terminal model is given. In, Section 3 the impact of timeslot assignment distributions on packet loss is investigated. This includes understanding the relationships between assigned RF rate, the distribution of slots that yield this rate, queue size, and router egress rate. In Section 4, we describe the experiment setup and present results.

2. TERMINAL DESIGN

One way in which functionality of a terminal can be partitioned and implemented is presented in Figure 1. This terminal consists of terminal command software, the terminal modem or digital hardware core, and a COTS router. The terminal modem performs all the signaling functions of the waveform (e.g., link-layer framing, FEC coding, modulation, etc.). The terminal command software configures both the digital hardware core and also the COTS router, and also performs terminal command functions such as generating and processing the SATCOM control messages. The router performs the IPv6 functionality required of the terminal.
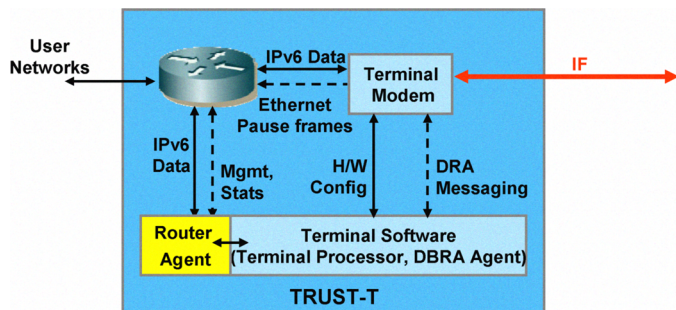

Figure 1: Lincoln Laboratory implementation of Future MILSATCOM Terminal.

Most commercial routers are designed to operate on fixed data rate links. However, to accommodate the time-varying rate of the RF link, a flow control mechanism between the COTS router and modem is required to prevent packet loss due to rate mismatch. In previous work [4], Ethernet pause frames are shown to be a viable flow control mechanism to moderate the router egress interface rate

to the assigned RF uplink rate. It is assumed that the router and modem are interconnected via Gigabit Ethernet and the modem transmits Ethernet pause frames to the router to enforce the RF uplink rate. The rate and pattern the pause frames generation is a function of the timeslot assignment distribution.

Leveraging on this work, we develop a way to analyze the effect the TDMA environment the SATCOM system has on a COTS router by emulating DRA timeslot assignments using pause frames. Since the flow control mechanism is driven by the time slot assignment, it is important to note that the work in this paper is also applicable if other flow control mechanisms are employed instead of Ethernet pause frames (e.g., PPPoE with extensions for credit flows [5]). Given the TDMA nature of the uplink, buffering during idle times will always be a necessity, regardless of the flow control mechanism employed, and the effects of the timeslot distribution will still be observed at scheduler mechanism (in the router in this terminal design).

An alternative to queueing packets in the router is to queue them in the modem. While this design removes the need for modem-to-router flow control, it increases the design complexity of the modem in other ways. More buffer space will be required in the modem in addition to the implementation of queueing, scheduling, and potentially other DiffServ mechanisms. In this design, the terminal router acts as a pass-through, in terms of queueing and scheduling, and the on/off effects of the TDMA environment occurs at the scheduler mechanism in the modem.

2.1 ETHERNET PAUSE FRAMES

The Ethernet standard includes an optional flow control operation known as pause frames [6]. Pause frames permit one end station to temporarily stop all traffic from the other end station (except MAC Control frames). For example, assume a full-duplex link that connects two devices called "station A" and "station B". Suppose station A transmits frames at a rate that causes station B to enter into a state of congestion (i.e., no buffer space remaining to receive additional frames). Station B may transmit a pause frame to station A requesting that station A stop transmitting packets for a specified period of time (called pause duration). Upon receiving the pause frame, station A will suspend further packet transmission until the specified pause duration has elapsed. This will allow station B time to recover from the congestion state. At the end of the pause duration, station A will resume normal transmission of frames.

In our model of a terminal, "station A" and "station B" are the terminal router and terminal modem, respectively. The

modem should only be buffering enough traffic to ensure an assigned timeslot is always used, given there is offered load in the terminal. Hence, when the terminal is in a timeslot not assigned to it, the modem should quickly enter a "congestion state" and send a pause frame to the router. Then, upon entering or just prior to an assigned timeslot, the modem should "unpause" the router, ensuring the assigned RF channel is always being used. This results in the majority of packets being queued in the router, where QoS mechanisms can be applied.

A pause frame contains a 16-bit value that specifies the duration of the pause event in units of 512-bit times (i.e., the time it takes to transmit 512 bits over the port media, which would be 0.512 microseconds when using Gigabit Ethernet). Valid values are 00-00 to FF-FF (hex). If a new pause frame arrives before the current pause duration has expired, its parameter replaces the current pause duration. Receiving a pause frame with pause duration of zero allows traffic to resume immediately.

Hence, the rate at which pause frames are sent and the duration of the pause are two parameters that can be adjusted to yield different service profiles. In this work, we emulate different DRA assigned distributions of timeslots via different series of pause frame patterns. Two different ways of generating pause frame streams using a commercial traffic generator are now described: constant rate pause streams and variable rate pause streams. These represent a uniform and non-uniform distribution of timeslot assignments, respectively. They are described below.

2.1.1 CONSTANT RATE PAUSE STREAM

We refer to the state of an interface as being **ON** when it is able to forward packets at the full service rate. The service rate is determined by the speed negotiation between the communication devices over a communication link (e.g., 1 Gbps for Gigabit Ethernet). The interface is **OFF** when the interface is in a suspended transmission state, resulting from receiving a pause frame on that interface; no packets can be forwarded out an interface in an OFF state.

A pause stream consists of a sequence of pause frames. Let β_i be the length of an OFF period, and α_i be the time interval of an ON period be between two consecutive OFF periods, as showed in Figure 2. A cycle is defined to be an instance of consecutive ON and OFF periods, and the ith cycle is displayed in Figure 2. A pause stream is a constant rate stream if α and β are fixed for every cycle within a pause stream. Otherwise, it is a variable rate pause stream. Constant pause streams can be used to emulate DRA uplink assignments consisting of uniformly distributed groups of contiguous timeslots
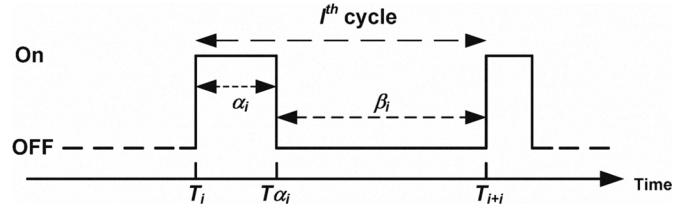

Figure 2: The ON and OFF periods of a pause frame.

2.1.2 VARIABLE RATE PAUSE STREAM

A pause stream is variable rate stream if either α or β is not a constant. Due to dynamic nature of the uplink quality in addition to traffic demand from terminals sharing the same uplink beam, it is likely that time slot assignments will change over a number of epochs, and depending on the above mentioned inputs, in addition to uplink resource packing constraints, there is no guarantee the groups of timeslots assigned to a terminal will be uniformly distributed. While it is trivial to create constant rate pause streams on a commercial traffic generator, no commercial traffic generator currently provides the option of generating a variable rate pause stream. However, a variable rate pause stream can be obtained by combining multiple constant rate pause streams.

As stated earlier, if an additional pause frame arrives before the current pause time has expired, its parameter replaces the current pause time. This feature can be utilized to generate a variable rate pause stream by multiplexing multiple constant pause streams, which may have unique α and β. Depending on the degree of overlap among pause frames from different streams, α and β of the resulting stream may be different from cycle to cycle, yielding a way to emulate a pseudo-random variable rate pause stream. By tuning α and β values of individual streams we can create streams in which we know the boundary values for α and β of the variable rate stream. For example, given two constant rate pause streams P1 and P2, the resulting variable rate pause stream is showed in Table 1 and Figure 3.

Pause Stream	α (msec)	β (msec)	# of Pause Frames per Second	Pause Duration per Second (msec)
P1	72.76	18.15	11	200
P2	40	10	20	200
P1 + P2	0 − 40	10 − 28.15	24	< 400

Table 1: The configuration of constant rate pause streams P1 and P2 and the resulting combined stream
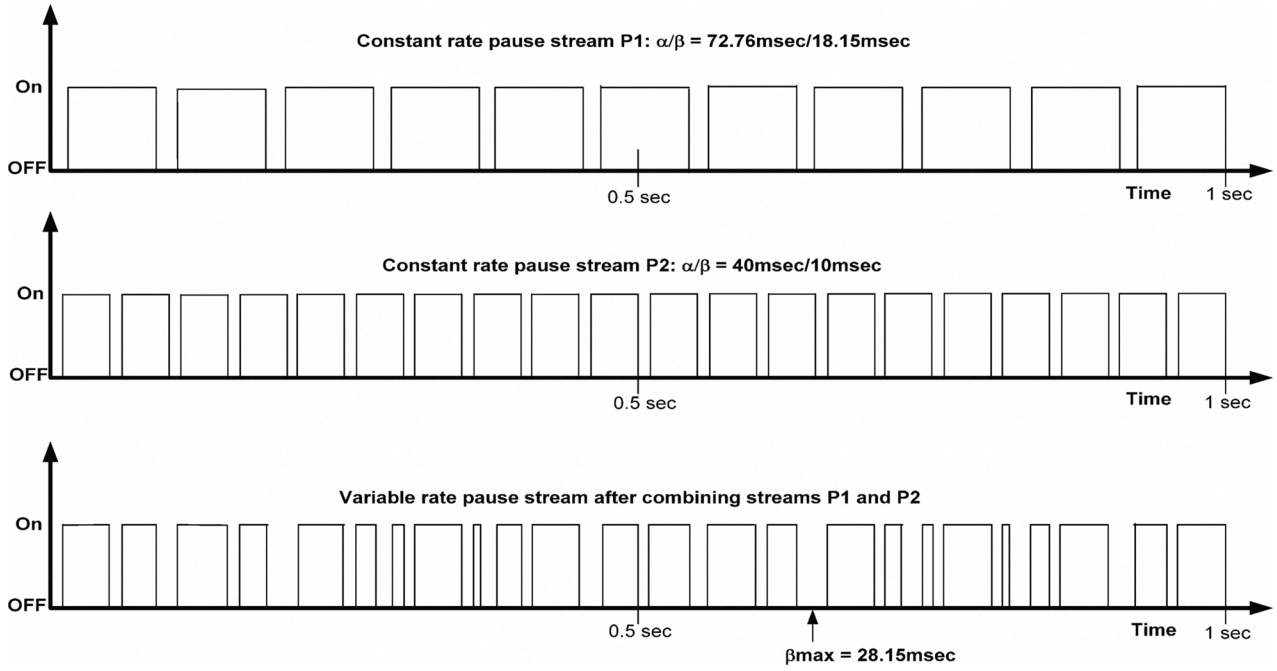
Figure 3: A variable rate pause stream resulting from two constant rate pause streams.

3. IMPACT OF TIMESLOT DISTRIBUTION ON PACKET LOSS

To avoid packet drops, most COTS routers provide a limited buffer space to hold packets. A buffer is used to hold packets due to the momentary imbalance of arrival and departure rates. Instead of functioning as a single FIFO queue, depending on the router's architecture, the buffer can be divided into multiple queues. To provide differentiated service, each queue can hold a specific class or classes of traffic and the rate at which the queues are serviced can be configured by the user by selecting a specific type of scheduling.

At an interface, during periods of time when the arrival rate is constantly larger than the transmission rate, the traffic queue will fill and any additional packet arriving will also be queued. The maximum number of packets a queue can hold is determined by the allocated buffer space and the packet size. Obviously if the average arrival rate is larger than the service rate long enough, packet drop will occur regardless of the buffer size being allocated. In this study, it is assumed that we are in an uncongested state, in that average arrival rate is less than or equal to the service rate over one epoch.

As stated above, packets will be queued when an interface is in an OFF state. To simplify the following discussion, it is assumed that only one queue is associated with an egress interface; a multiple queue configuration is described later.

Whether packet drops occur during a given OFF period is determined based on many factors, such as the queue length at the beginning of the OFF period, the number of packets received during the OFF period, the queue size, and the duration of OFF period. To derive the queue length at a given time, the following terms are used:

Term	Definition
AR	Average data rate of a traffic stream
TR	Transmission (service) rate of an egress interface
PS_{avg}	Average packet size
Q_i	Queue length at the beginning of ith cycle in terms of number of packets
Q_{max}	Queue size in terms of max number of packets it can hold
α_i	Duration during which the interface is allowed to transmit packets in the ith cycle (**ON** period)
β_i	Duration during which the interface remains off in the ith cycle (**OFF** period)

Referring back to Figure 2, during the ith cycle, the queue length at $T_{\alpha i}$ is equal to the queue length at T_i, plus the number of packets received during α_i, minus the number of packets allowed to be transmitted during this period, i.e.,

$$Q_{\alpha i} = Q_i + (AR* \alpha_i) / PS_{avg} - (TR* \alpha_i) / PS_{avg}$$
$$= Q_i - ((TR - AR)* \alpha_i) / PS_{avg} \qquad (1)$$

Because **TR** is assumed to be larger than or equal to **AR**, the result of Eq. (1) can be a negative value, i.e., the number of packets delivered is greater than the number of packets in the queue initially and received during α_i. Obviously, that is incorrect. Thus, Eq. (1) is further revised as following:

$$Q_{\alpha i} = Q_i - ((TR - AR)* \alpha_i) / PS_{avg},$$
$$\text{if } Q_i \geq ((TR - AR)* \alpha_i) / PS_{avg},$$
$$Q_{\alpha i} = 0, \quad\quad \text{otherwise.} \quad\quad (2)$$

Next, the number of packets received during the off period, β_i, is $(AR* \beta_i) / PS_{avg}$. Thus, the queue length at T_{i+1} $Q_{i+1} = Q_{\alpha i} + (AR* \beta_i) / PS_{avg}$. That is:

$$Q_{i+1} = Q_i - ((TR - AR)*\alpha_i) / PS_{avg} + (AR* \beta_i) / PS_{avg},$$
$$\text{if } Q_i \geq ((TR - AR)*\alpha_i) / PS_{avg},$$
$$Q_{i+1} = (AR* \beta_i) / PS_{avg}, \quad \text{otherwise.} \quad\quad (3)$$

At an egress queue, no packet drop will occur if the following two conditions are met:
1. An egress interface can transmit all the packets which are held in the queue at the beginning of a ON period and received during that period, i.e., $0 \leq ((TR - AR)*\alpha) / PS_{avg} - Q_{max}$. That is no packet will remain in the queue at the end of ON period.

2. The number of packets received during the OFF period is less than or equal to the queue size, i.e., $Q_{max} \geq (AR* \beta) / PS_{avg}$

These two conditions can be restated as following:
$$\alpha \geq (Q_{max}* PS_{avg}) / (TR - AR) \quad \text{and}$$
$$\beta \leq (Q_{max}* PS_{avg}) / AR \quad\quad\quad (4)$$

To verify Eq. (4), several tests have been conducted and the results are described in the following section. While Eq. (4) provides a region of operation that guarantees no packet loss, it is possible to operate with values of α less than this without loss of packets, as is demonstrated in section 4.2.1. When operating in this region, whether packet loss occurs is constrained by whether the queue is unstable at steady state (arrival rate is greater than service rate).

This single queue model can be easily extended to a multi-queue model. For each queue in a multi-queue configuration, the ON period (α_i) in Eq. (3) should be replaced by the portion of ON time being allocated to this queue and the OFF period (β_i) should be extended to include the time when the egress port is on but this queue is not being served.

4. EXPERIMENT SETUP

For the experiments discussed in this paper, two Gigabit Ethernet ports on a router are connected to an Ixia traffic generator as shown in Figure 4. One router port receives the data streams and the second router port sends the traffic back to the tester. This second port also receives pause frames generated from port 2 on the Ixia.
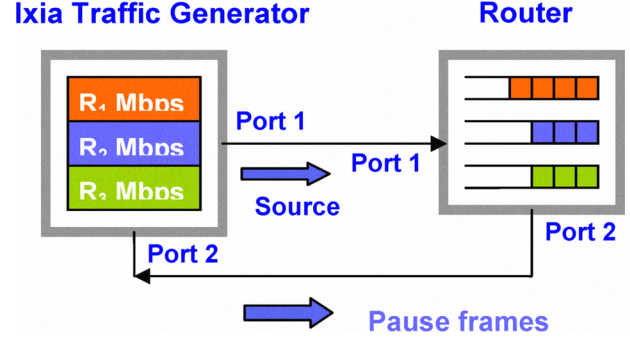

Figure 4: Test setup

Both constant rate and variable rate (bursty) traffic streams are considered in turn, in both the constant and variable rate pause stream environment. The traffic generated is IPv6. A Juniper M120 is used as the router with an 8-Port Type 3 Gigabit Ethernet IQ2 PIC.

In practice, queue size is configurable and may be defined in terms of maximum number of packets it can hold during congestion (for Cisco routers) or the percentage of total available buffer space (for both Cisco and Juniper routers). The queue size is fixed for these experiments to be 10% of the total buffer space of a M120 Gigabit interface, which equates to approximately ~8,200 packets. The Gigabit Ethernet interface is able to deliver packets up to 1 Gbps.

4.1 RELATIONSHIP OF PAUSE STREAMS WITH PACKET LOSS

Only one constant rate data stream (400 Mbps) is considered in this subsection and the packet size is fixed to 100 bytes. Due to the preamble and inter-packet gap, the actual rate generated from Ixia is 333 Mbps. For packets of 100 bytes, the queue on the egress interface can hold up to ~8200 packets. Thus, $Q_{max} = 8200$ packets, $PS_{avg} = 100*8$ bits, $TR = 10^9$ bps, and $AR = 3.33*10^8$ bps. Then, according to Eq. (4), packet drop will not occur as long as the duration of the ON period (α) ≥ 9.8 msec and the duration of OFF period (β) ≤ 19.7 msec.

4.1.1 CONSTANT RATE PAUSE STREAM

Assume α and β are the same among all the cycles in a pause stream. By varying the value of α and β, four different constant rate pause streams have been tested and the results are showed in Table 2.

Test #	1	2	3	4
α (msec)	15	15	11.4	9.7
β (msec)	20	18	18	18
Packet Drop	Yes	No	No	Yes

Table 2: Packet drop experiment by varying the ON and OFF durations.

In this experiment, the configuration of test 1 fails to meet the second condition described in Eq. (4), and packet drop occurs after a pause frame is received. On the other hand, because tests 2 and 3 satisfy both requirements, no packet drop is observed during the test. The configuration of test 4 fails to meet the first condition described in Eq. (4). Thus, the egress interface is unable to transmit all the packets which are held in the queue at the beginning of α and received during α. Thus, some residual packets are held in the queue before the next pause frame arrives. Consequently the amount of buffer space available to hold the packets received during β is reduced, eventually resulting in dropping packets.

4.1.2 VARIABLE RATE PAUSE STREAM

The same test configuration is used in this subsection. As derived earlier, to avoid packet drop, the duration of ON period α should be larger than 9.8 msec and the duration of OFF period β should be smaller than 19.7 msec. Instead of using constant rate pause streams, variable rate pause streams are considered here, i.e., either α or β is not a constant. As described in Subsection 2.2, a variable pause stream must be constructed by using multiple constant rate pause streams.

The combined pause stream of P1+P2 characterized in Table 1 is now evaluated. Given the actual data arrival rate 333 Mbps, neither P1 nor P2 should cause the data stream to lose packets according to Eq. (4). However, packet drops are expected when the resulting variable rate pause stream (P1 + P2) is deployed. That is because the destined queue is not big enough to hold all the packets received during the OFF period which is ranging from 10 msec to 28.15 msec. This has been confirmed experimentally and the results are showed in Table 3.

Variable rate pause streams provide the greatest flexibility to perform flow control because they can generate a pause frame whenever there is a need. This test demonstrates the challenge of constructing a variable rate pause stream that reduces the service rate of an egress port without causing packet drops. Given the targeted service rate, there are many ways to construct a variable rate pause stream by varying pause duration and rate. However, to reduce the possibility of packet loss, the preference should be given to the stream with short pause duration, i.e., the maximum of β is less than $(Q_{max} * PS_{avg}) / AR$.

Actual Arrival Rate (Mbps)	333	333	333	333
Pause Stream	–	P1	P2	P1+P2
α (msec)	–	72.76	40	0 – 40
β (msec)	–	18.15	10	10 – 28.15
Reduced Line Rate (Mbps)	1000	800	800	> 600
Throughput (Mbps)	333	333	333	330
Packet Drops	No	No	No	Yes

Table 3: Packet drop experiment by deploying a variable rate pause stream.

4.2 RELATIONSHIP OF PAUSE STREAMS AND TRAFFIC LOAD ON LATENCY

In the prior section, how pause durations impact packet loss was modeled and discussed. The impacts of the pause duration in conjunction with offered load on packet latency is presented here. The measurements being considered include the average and maximum cut-through latencies. The *cut-through latency* of a packet is the time interval between the first data bit out of the traffic generator transmit port and the first data bit received by the traffic generator receive port is measured [7]. The average (or maximum) cut-through latency is the average (or maximum) delay seen on packets associated with a group, e.g., packets with the same DSCP value.

Because the constant rate pause stream is a special case of the variable rate pause stream, i.e., having constant α and β for all pause frames, only the variable rate pause stream is considered in this subsection. We consider the variable rate pause stream, P3+P4, characterized in Table 4, that is obtained by the multiplexing of constant rate pause streams P3 and P4, which are arbitrarily selected. For the resulting stream P3+P4, α=0.02 msec and β takes the value of either 0.2 msec or 0.42 msec.

For the traffic load used in this experiment, both constant rate and variable rate (bursty) traffic streams are considered. Network traffic generally consists of both constant rate streams, e.g., voice and video, and bursty streams,

e.g., ftp and web services. It is important to know how the pause duration can affect the packet latency and how to construct a pause stream such that it can have the minimum impact on packet latency. That is particularly important when the bounded delay is required for certain applications, such as voice.

Pause Stream	α (msec)	β (msec)	# of Pause Frames per Second	Pause Duration per Second (msec)
P3	0.02	0.2	4500	900
P4	1.22	1	450	450
P3 + P4	0.02	0.2 / 0.42	4050	> 900

Table 4: Configuration of two constant rate pause streams and its resulting variable rate stream.

4.2.1 CONSTANT RATE TRAFFIC STREAM

Assume the data rate is fixed and the packet size is uniformly distributed between 100 bytes and 1518 bytes. The traffic loads being evaluated include 50Mbps, 75Mbps, 100Mbps, and 125Mbps. Due to the preamble and interpacket gap overheads, the actual rates generated from Ixia are 48.8Mbps, 73.2Mbps, 97.6Mbps, and 122Mbps, respectively. Given the traffic loads and a variable rate pause stream, P3 + P4, the results of this experiment are showed in Table 5.

Actual Arrival Rate (Mbps)	48.8	73.2	97.6	122
α (msec)	0.02	0.02	0.02	0.02
β (msec)	0.2/0.42	0.2/0.42	0.2/0.42	0.2/0.42
Reduced Line Rate (Mbps)	< 100	< 100	< 100	< 100
Throughput (Mbps)	48.4	73.2	97.6	103.5
Packet Drop	No	No	No	Yes
Avg/Max Latency (msec)	0.15/0.47	0.17/0.58	0.29/1.70	1500/1500

Table 5: Latencies of constant rate data streams.

When the arrival rate is far below the reduced line rate (e.g., 48.8 and 73.2 Mbps), most packets received during the ON period can be forwarded to the destined egress port without being held in the queue. Only the packets received during the OFF period will be held in the queue. This delay incurred by a queued packet waiting for the

next ON period is the major contributor to the average latency and the maximum latency is determined by the maximum pause duration (β). When the arrival rate is approaching or larger than the reduced line rate, packet drops occur and the egress queue remains almost full throughout the test. The latency when operating in these scenarios is mainly dominated by the queueing delay associated with a full queue. Hence, the larger the queue, the longer the latency experienced.

Note that in Table 5, α is less than that specified by (4) but packet loss does not occur for arrival rates of 48.8 to 97.6 Mbps. For these scenarios, the packet arrivals within an OFF cycle can all be serviced during an ON state and the buffer remains very shallow on average (less than approximately 3 packets). An arrival rate of 97.6 Mbps represents a special case in which the arrival rate is slightly greater than the service rate (3.7 versus 3.1 packets per pause cycle) but packet loss still does not occur due to a property of pause frames. Namely, when a pause frame is received by an Ethernet interface during transmission of a packet, the interface will finish transmitting the packet before pausing. Hence, the service rate is actually 4 packets per ON state, resulting in no packet loss.

4.2.2 VARIABLE RATE TRAFFIC STREAM

In this section a variable rate data stream is considered. Similar to the construction of variable rate pause streams, to generate a variable rate data stream, multiple constant rate data streams are used. Given two constant rate streams with data rates *R1* and *R2*, the switch pattern *X − Y* describes an Ixia-generated stream with rate *R1* for *X* seconds and then with rate *R2* for *Y* seconds, alternatively. Figure 5 shows a bursty stream with the switch pattern *X − Y*.
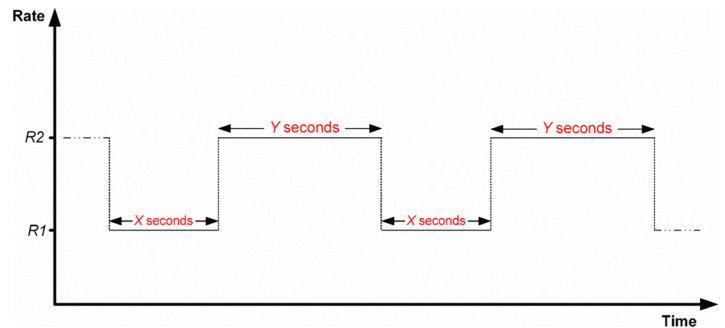

Figure 5: Constructing a bursty stream.

In this experiment, *R1* and *R2* are arbitrarily chosen to be 75Mbps and 125Mbps, respectively. Because of the preamble and inter-packet gap, the actual arrival rates are 73.2Mbps and 122Mbps. By varying *X* and *Y* values, five different bursty streams are considered in this experiment. The results are summarized in Table 6. For comparison

purpose, the result of a constant rate data stream (100Mbps) showed in Table 5 is also included.

Similar to the constant rate experiment, when the average arrival rate is much smaller than the reduced line rate, the latency is mainly determined by the packet processing time and pause duration. When the average rate is approaching or larger than the reduced line rate, the queueing delay becomes the dominating factor in latency. The burstiness of arriving traffic causes more packets to be queued and thus, increases the queue depth and prolongs the delay time. For example, consider the cases where the throughput is 97.8Mbps. The average / maximum latency of the constant rate stream is 0.29 msec / 1.70 msec. For the bursty streams, with the switch patterns $1 - 1$, $2 - 2$, and $4 - 4$, the average/maximum latency becomes 76.93 msec / 189.38 msec, 154.72 msec / 370.41 msec, and 305.25 msec / 738.49 msec, respectively. Doubling the burst duration effectively doubles the latency even though the average data rate remains the same.

5. CONCLUSION

In future protected MILSATCOM systems, terminal uplink assignments may change per epoch, implying the average uplink data rate has the ability to change on epoch time scales. Furthermore, due to the TDMA nature of the uplink, a terminal's instantaneous data rate is zero during timeslots of an epoch not assigned to it.

In this paper we investigated the impact timeslot assignment distribution has on both queue overflow and latency. A formula was derived which can be used to determine the occurrence of queue overflow given the arrival rate, egress rate of an interface, queue size, and assigned uplink timeslot distribution. For all tests described in this paper, only one queue is considered at a router's egress port. However, it was discussed how this model can be easily extended to include a multi-queue configuration.

An increase in length of the idle period between assigned transmission slots results in an increase in the number of packets queued and latency experienced by packets in the terminal router. If bounds on the length of the idle and transmission periods can be obtained from DRA or mission planning, then router queues can be sized to ensure no packet loss when in the terminal is in an uncongested state. Additionally, information about available buffer size and buffer size impact on queue latency may also be one input in the configuration of the DRA timeslot assignment algorithm. As the time between timeslot allocations decreases, queue latency and required queue size also decrease. However, minimizing the length of the idle period between slot assignments may not be straight-forward, as timeslot allocation is a function of many things, including terminal committed information rate, terminal priority, resource packing complexity, and overall demand and availability of system resources.

REFERENCES

1. RFC 2475, "An Architecture for Differentiated Services", 1998.

2. Vijay. Bollapragada, Russ White, and Curtis Murphy, "Inside Cisco IOS Software Architecture (CCIE Professional Development)", *Cisco Press*, 2000.

3. Thomas M. Thomas, Doris Pavlichek, Lawrence H. Dwyer, III, Rajah Chowbay, Wayne W. Downing, and James Sonderegger, "Juniper Networks reference guide: JUNOS routing, configuration, and architecture", *Addison-Wesley*, 2002.

4. Aradhana Narula-Tam, Jeff Wysocarski, Huan Yao, Mu-Cheng Wang, Thomas Macdonald, Orton Huang, and Julee Pandya, "QoS considerations for future military sitcom networks with link layer dynamic resource allocation" *International Journal of Satellite Communications and Networking*. Sep. 2006, Volume 24, Issue 5, pp. 343 – 365.

5. RFC 4938, "PPP over Ethernet (PPPoE) Extensions for Credit Flow and Link Metrics", 2007.

6. IEEE 802.3x Standard.

7. Dhiman Deb Chowdhury, "High speed LAN technology handbook", *Springer*, 2000.

Switch Pattern $X - Y$	Const. Rate	$1 - 1$	$2 - 2$	$4 - 4$	$10 - 3$	$3 - 10$
α (msec)	0.02	0.02	0.02	0.02	0.02	0.02
β (msec)	0.2/0.42	0.2/0.42	0.2/0.42	0.2/0.42	0.2/0.42	0.2/0.42
Reduced Line Rate (Mbps)	< 100	< 100	< 100	< 100	< 100	< 100
Throughput (Mbps)	97.8	97.8	97.9	97.7	84.5	103.5
Packet Drop	No	No	No	No	No	Yes
Avg. Latency (msec)	0.29	76.93	154.72	305.25	120.06	1221
Max Latency (msec)	1.70	189.38	370.41	738.49	551.38	1500

Table 6: Latency measurement of variable rate data streams with pause stream.