IEEE 802.16 BASED LAST MILE BROADBAND WIRELESS MILITARY NETWORKS WITH QUALITY OF SERVICE SUPPORT

Kitti Wongthavarawat
Aura Ganz
Multimedia Networks Laboratory
Electrical and Computer Engineering Department
University of Massachusetts, Amherst, MA USA

ABSTRACT

Future Department of Defense warfighting concepts leverage information superiority and will require vast improvements in information transfer in terms of higher bandwidth, Quality of Service (QoS) support and connection to a high speed backbone. The new IEEE 802.16 broadband wireless access system is a viable alternative that can meet such requirements. In addition, this network can be swiftly deployed to interconnect the military theater, emergency response, and disaster relief operations to the backbone. Due to the diverse multimedia traffic with different priorities and QoS requirements, it is a well know fact that it is imperative to provide QoS support in military networks. However, the IEEE 802.16 provides only signaling mechanisms, but does not specify any scheduling or admission control algorithms that ultimately provide QoS support.

In this paper we introduce a new scheduling algorithm for IEEE 802.16 broadband wireless access standard. The proposed solution which is practical and compatible to the IEEE 802.16 standard, provides QoS support to different traffic classes. To the best of our knowledge this is the first such algorithm. The simulation studies show that the proposed solution includes QoS support for all types of traffic classes as defined by the standard. We have shown the relationship between traffic characteristics and its QoS requirements and the network performance. This study will help network architects to decide the system parameters as well as the kind of traffic characteristics for which the network can provide QoS support.

Acknowledgement: This project was supported in part by the following grants: grants: NSF-ANI-0230812, NSF-EIA-0080119, ARO-DAAD19-03-1-0195, DARPA-F33615-02-C-4031.

1. INTRODUCTION

Information superiority plays a crucial role in determining a decisive victory in the current warfare. Therefore, modernizing the communication infrastructure of the US military command and control is necessary for enhancing the war fighting capabilities. For example, it is pertinent to employ wireless networks that provide high bandwidth and Quality of Service (QoS) support. Figure 1 shows an example of the command and control communication infrastructure. Front-line forces that are constantly moving in the battlefield communicate with the base command through a satellite. The base command and branch units which are less mobile communicate with each other through a high speed wireless connection. All the information collected from the base command is sent to the headquarters through a high speed backbone.

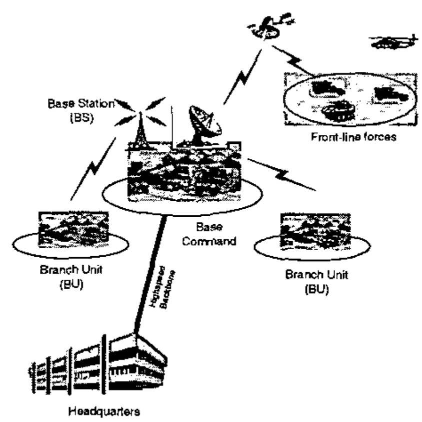


Figure 1: Military Networks

Broadband wireless systems, e.g., IEEE 802.16 standard that was recently approved [1], are viable solutions that enable high speed wireless connectivity between the base command and the branch units. Deployment of IEEE 802.16 provides the following benefits: 1) fast deployment 2) high speed wireless connectivity, and 3) low cost. It is important to mention that a variety of military applications such as just-in-time logistics, teleconferencing, telemedicine and multimedia applications, require QoS support in terms of delay and delay variation. Moreover, the network needs to be able to provide priority services to messages based on their criticality. In case such QoS support can not be provided, the quality of the application is severely degraded, and in fact it may compromise critical

0-7803-8140-0/03$17.00 (C) 2003 IEEE

war fighting decisions. Therefore, QoS support is a necessity in military deployments. However, the IEEE 802.16 provides only signaling mechanisms. It does not specify any scheduling or admission control algorithms that ultimately provide QoS support. In recent years, several packet scheduling algorithms for broadband wireless networks were published [3,4,5,6,7,8,9]. To the best of our knowledge, there is no proposed packet scheduling solution specifically designed for IEEE 802.16. In this paper, we propose a packet scheduling algorithm that provides QoS support for a wide range of real time applications as defined in IEEE 802.16. The proposed solution is practical and compatible with the IEEE 802.16 QoS signaling mechanisms. The simulation results we obtained show that the proposed solution can support diverse traffic classes of traffic with different QoS requirements in terms of bandwidth and maximum delay.

The paper is organized as follows. In Section 2 we introduce the IEEE 802.16 broadband wireless access systems. In Section 3, we describe the existing IEEE 802.16 QoS architecture as well as our proposed QoS architecture. In Section 4, we describe in details the proposed Uplink Packet Scheduling (UPS) algorithm. Section 5 provides simulation results of our proposed UPS algorithm and Section 6 concludes the paper.

2. IEEE 802.16 BROADBAND WIRELESS ACCESS SYSTEMS

The Physical layer operates at 10-66 GHz (IEEE 802.16) and 2-11 GHz (IEEE 802.16a) with data rates of 32 – 130 Mbps depending on the channel frequency width and modulation technique. IEEE 802.16 architecture consists of two kinds of stations: Branch Unit (BU) and a Base Station (BS). In practice, BS resides in the base command. The BS regulates all the communication in the network, i.e., there is no peer-to-peer communication directly between the BUs. The communication path between BU and BS has two directions: uplink (from BU to BS) and downlink (from BS to BU). When the system uses time-division multiplexing (TDM), for uplink and downlink transmissions, the frame is subdivided into an uplink subframe and a downlink subframe (see Figure 2). The duration of these subframes is dynamically determined by the BS. Each subframe consists of a number of time slots. BUs and BS have to be synchronized and transmit data into predetermined time slots.

IEEE 802.16 can support multiple communication services (data, voice, video) with different QoS requirements. The Media Access Control (MAC) layer defines QoS signaling mechanisms and functions that can control BS and BU data transmissions. On the downlink (from BS to BU), the transmission is relatively simple because the BS is the only one that transmits during the downlink subframe. The data packets are broadcasted to all BUs and a BU only picks up the packets destined to it. One of the modes of uplink arbitration (from BU to BS) uses a TDMA MAC. The BS determines the number of time slots that each BU will be allowed to transmit in an uplink subframe. This information is broadcasted by the BS through the Uplink Map Message (UL-MAP) at the beginning of each frame. UL-MAP contains Information Element (IE) which includes the transmission opportunities, i.e., the time slots in which the BU can transmit during the uplink subframe. After receiving the UL-MAP, each BU will transmit data in the predefined time slots as indicated in IE. The BS uplink scheduling module determines the IEs using Bandwidth Request PDU (BW-Request) sent from BUs to BS. IEEE 802.16 defines the required QoS signaling mechanisms such as BW-Request and UL-MAP, but it does not define the Uplink Scheduler, i.e., the mechanism that determines the IEs in the UL-MAP.

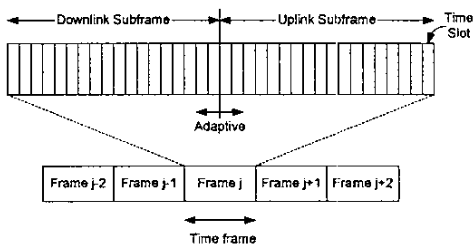


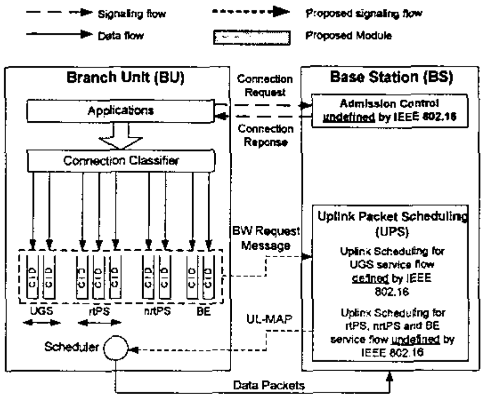
Figure 2: IEEE 802.16 TDM Frame Structure

3. QOS ARCHITECTURE

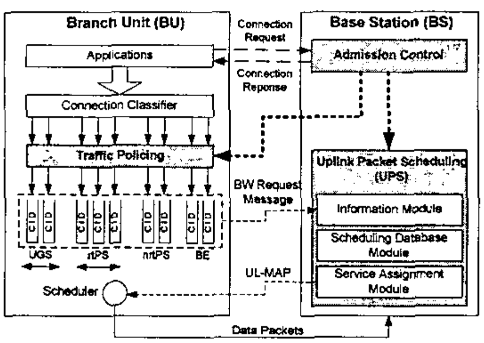
IEEE 802.16 defines four types of service flows, each with different QoS requirements and corresponding Uplink Scheduler policy:

1. Unsolicited Grant Service (UGS) – this service supports constant bit-rate (CBR) or CBR-like flows such as Voice over IP. These applications require constant bandwidth allocation.
2. Real-Time Polling Service (rtPS) -- this service is for real-time VBR-like flows such as MPEG video, just-in-time logistics, or teleconference. These applications have specific bandwidth requirements as well as a deadline (maximum delay). Late packets that miss the deadline will be useless.
3. Non-Real-Time Polling Service (nrtPS) – this service is for non-real-time flows which require better than best effort service, e.g., bandwidth intensive file transfer. These applications are time-insensitive and require minimum bandwidth allocation. .
4. Best Effort Service (BE) – this service is for best effort traffic such as HTTP. There is no QoS guarantee. The applications in this service flow receive the available bandwidth after the bandwidth is allocated to the previous three service flows.

For UGS, BW-Request is not required. For rtPS, nrtPS and BE, the current queue size is included in the BW-Request to represent the current bandwidth demand. In summary, IEEE 802.16 defines: 1) the signaling mechanism for information exchange between BS and BU such as the connection setup, BW-Request, and UL-MAP, 2) the Uplink Scheduling for UGS service flow. IEEE 802.16 does not define: 1) the Uplink Scheduling for rtPS, nrtPS, BE service flow, 2) the Admission Control and Traffic Policing process



(A) IEEE 802.16 QoS Architecture



(B) Proposed QoS Architecture

Figure 3: QoS architecture

Figure 3A shows the existing QoS architecture of IEEE 802.16. Uplink Packet Scheduling (UPS) resides in the BS to control all the uplink packet transmissions. Since IEEE 802.16 MAC protocol is connection oriented, the application first establishes the connection with the BS as well as the associated service flow (UGS, rtPS, nrtPS or BE). BS will assign the connection with a unique Connection ID (CID). The connection can represent either an individual application or a group of applications. IEEE 802.16 defines the connection signaling (Connection Request, Response) between BU and BS but it does not define the Admission Control process. All packets from the application layer in the BU are classified by the Connection Classifier based on CID and are forwarded to the appropriate queue. At the BU, the Scheduler will retrieve the packets from the queues and transmit them to the network in the appropriate time slots as defined by the UL-MAP sent by the BS. The UL-MAP is determined by the UPS module based on the BW-Request messages that report the current queue size of each connection in BU. Figure 3B shows the proposed QoS architecture that completes the missing parts in the IEEE 802.16 QoS architecture. At the BS we add a detailed description of the UPS module (scheduling algorithm that which supports all types of service flows), and Admission Control module. At the BU we add the Traffic Policing module. Here is a brief description of the connection establishment using the QoS architecture in Figure 3B:

- An application that originates at a BU establishes the connection with BS using connection signaling. The application includes in the connection request the traffic contract (bandwidth and delay requirement).
- The Admission Control module at the BS accepts or rejects the new connection.
- If the Admission Control module accepts the new connection, it will notify the UPS module at the BS and provide the token bucket parameters to the traffic policing module at the BU.

After the connection is established, the following steps are taken:

- Traffic policing enforces traffic based on the connection's traffic contract.
- At the beginning of each time frame, the UPS's Information Module collects the queue size information from the BW-Requests received during the previous time frame. The Information Module will process the queue size information and update the Scheduling Database Module.
- The Service Assignment Module retrieves the information from the Scheduling Database Module and generates the UL-MAP.
- BS broadcasts the UL-MAP to all BUs in the downlink subframe.
- BU's scheduler transmits packets according to the UL-MAP received from the BS.

4. PROPOSED UPLINK PACKET SCHEDULING

To support all types of service flows (UGS, rtPS, nrtPS and BE), the proposed Uplink Packet Scheduling uses a combination of Strict Priority service discipline, Earliest Deadline First (EDF) [10] and Weight Fair Queue (WFQ) [11]. The hierarchical structure of the bandwidth allocation is shown in Figure 4.

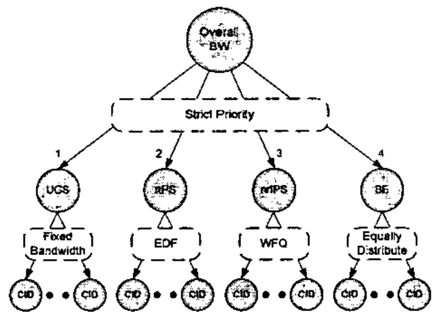


Figure 4: Hierarchical structure of bandwidth allocation

The proposed UPS principles:

1. Overall bandwidth allocation: bandwidth allocation per flow follows strict priority, from highest to lowest: UGS, rtPS, nrtPS and BE. One disadvantage of the strict priority service discipline is that higher priority connections can starve the bandwidth of lower priority connections. To overcome this problem, we include the Traffic Policing Module in each BU which forces the connection's bandwidth demand to stay within its traffic contract. This will prevent the higher priority connections from using bandwidth more than their allocation.
2. Bandwidth allocation within UGS connections: The UPS allocates fixed bandwidth (fixed time duration) to UGS connections based on their fixed bandwidth requirement. This policy is determined by the IEEE 802.16 standard.
3. Bandwidth allocation within rtPS connections: We apply Earliest Deadline First (EDF) service discipline to this service flow. Packets with earliest deadline will be scheduled first. The Information Module determines the packets' deadline.
4. Bandwidth allocation within nrtPS connections: We apply Weight Fair Queue (WFQ) service discipline to this service flow. We schedule nrtPS packets based on the weight of the connection (ratio between the connection's nrtPS average data rate and total nrtPS average data rates)
5. Bandwidth allocation within BE connections: The remaining bandwidth is equally allocated to each BE connection.

The proposed UPS consists of three modules: Information Module, Scheduling Database Module and Service Assignment Module.

A. Information Module

The Information Module performs the following tasks:

1. Retrieves the queue size information of each connection from the BW-Request messages.
2. Determines the number of packets (in bits) that arrived from rtPS connection in the previous time frame using the Arrival-Service curve concept [12].
3. Determines rtPS packets' arrival time and deadline and updates this information in the Scheduling Database Module.
4. Queuing information from nrtPS and BE BW-Requests is passed directly to Scheduling Database Module.

Note: Since UGS requires only fixed bandwidth allocation and does not need BW-Requests, there is no need for processing UGS connections

Information Module for rtPS connections

The Information Module needs to find the rtPS deadline information (see Figure 5). Based on this deadline information, the UPS will know exactly when to schedule packets such that packets' delay requirements are met. We apply the Arrival-Service curve concept to determine the packets' arrival and deadline. The packets' deadline is their arrival time plus the connection's maximum delay requirement.

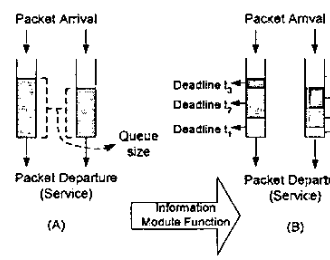


Figure 5: Concepts of Information Module operation for rtPS messages

In summary, the output of the Information Module which updates the Scheduling Database Module is given by:

1. rtPS connections – $Nd_i[a,b]$ the number of bits waiting in the queue of rtPS connection i with deadline in interval [a,b]
2. nrtPS connections – $q_j(t)$, the current queue size of nrtPS connection j
3. BE connections – $q_k(t)$, the current queue size of BE connection k

B. Scheduling Database Module

The Scheduling Database Module serves as the information database of all connections in the network. Figure 6 and Figure 7 show the database structure of the Scheduling Database Module. The database module (at time t) includes four types of databases based on each service flow as follows:

1. UGS database (Figure 6A) – this is a per connection database. Each item i in the database contains the number of bits ($N_{UGS,i}$) of connection i that need to be

serviced. This number is fixed and determined by the UGS connections' bandwidth requirement.

2. nrtPS database (Figure 6B) – this is a per connection database. Each item m in the database contains $q_m(t)$ (as found by the Information Module), i.e., the number of bits (current queue size) of connection m.
3. BE database (Figure 6C) – this is a per connection database. Each item w in the database contains $q_w(t)$ (as found by the Information Module), i.e., the number of bits (current queue size) of connection w.
4. rtPS database (Figure 7) – this is a two dimensional database, per connection and deadline (frame). Item (i, [t, t + f]) includes $Nd_i[t, t + f]$ (received from the Information Module) which is the number of bits to be transmitted in frame [t, t + f]. Figure 7 also shows the number of bits in the database table that correspond to the actual packets waiting in queue.

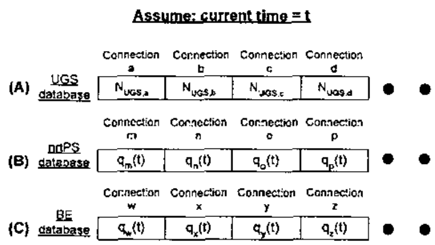


Figure 6: Database structure of UGS, nrtPS and BE in Scheduling Database Module

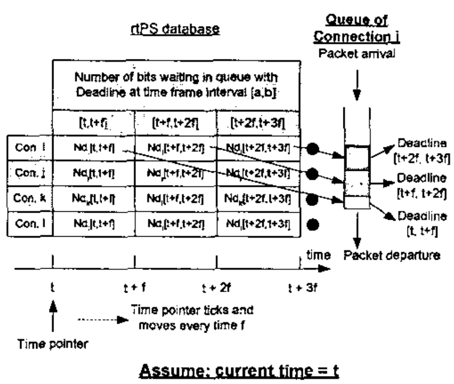


Figure 7: rtPS database structure in Scheduling Database Module

C. *Service Assignment Module*

The Service Assignment Module determines the UL-MAP, using the database tables created in the Scheduling Database Module. As depicted in Figure 4, we employ the following service disciplines: fixed bandwidth allocation for UGS, EDF for rtPS, WFQ for nrtPS and equal bandwidth allocation for BE. The Service Assignment Module determines the uplink subframe allocation in terms of the number of bits per connection. The number of bits will eventually be converted to the number of time slots which are the units used in the information elements (IE) of the UL-MAP. The number of bits per time slot is determined by the physical layer of the wireless network.

5. SIMULATION RESULTS

We have developed a simulation model in C++ that demonstrates that our proposed uplink packet scheduling (UPS) provides QoS support to real time applications. Packet arrivals occur at the beginning of each frame. The packet arrival process for each connection follows the token bucket envelope with parameters: token bucket rate (r_i), token bucket size (b_i) and max. burst size. Each connection has specific QoS parameters in terms of 1) average bandwidth requirement which is equal to the token bucket rate, and 2) maximum delay requirement. Simulation output: 1) the arrival curve which depicts the arrival pattern of the input traffic, 2) the service curve which shows the service pattern provided by UPS. The goal of this experiment is to show that the proposed UPS can provide QoS support in terms of bandwidth and delay for rtPS traffic. Assumptions: 1) There are only two types of traffic (rtPS and BE), 2) All traffic is already admitted to the network, 3) BE traffic requires uplink bandwidth at all times, and 4) C_{total} = 10 Mbps, C_{uplink} = 5 Mbps, $C_{downlink}$ = 5 Mbps The goal of the experiment is to show that the proposed UPS can provide QoS support in terms of bandwidth and delay for rtPS traffic.

- Frame size (f) = 10 msec.
- Three rtPS sessions with average total bandwidth (C_{rtPS}) of 3 Mbps.
- rtPS traffic characteristics are shown in Table 1 We specify token bucket rate (r_i), token bucket size (b_i), max. burst size and max. delay requirement for each session. The corresponding peak rate and burstiness (peak rate/average rate) are calculated from r_i, b_i, and max. burst size.

Session	1	2	3
Token bucket rate, r_i (kbps)	500	1000	1500
Token bucket size, b_i (bits)	10000	20000	30000
Maximum Burst Size (msec.)	10	10	10
Peak rate (kbps)	1500	3000	4500
Burstiness (peak rate/average rate)	3	3	3
Max. Delay Requirement (msec.)	20	40	60

Table 1: Input Traffic

Figure 8 shows the bandwidth allocation for rtPS and BE connections. Since our UPS is work conserving and there is always BE traffic available, rtPS and BE bandwidth allocation complement each other, i.e., in each frame the total rtPS and BE bandwidth equal to 5 Mbps. In this experiment there are no packets that miss their deadline. Figure 9 shows the arrival and service curves of all three rtPS connections. The graphs clearly show that the service curve adapts and follows the arrival curve. Our UPS dynamically allocates bandwidth based on the bandwidth demand of each session. The delay of each session is also guaranteed since there are no packets that miss their deadline. We observe that the horizontal distance between these two curves (arrival curve and service curve) of each session is bound by the maximum delay of each session.

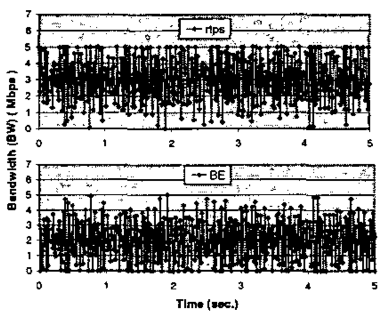


Figure 8: Bandwidth allocation for rtPS and BE connections (C_{uplink} = 5 Mbps, C_{rtPS} = 3 Mbps, C_{BE} = 2 Mbps)

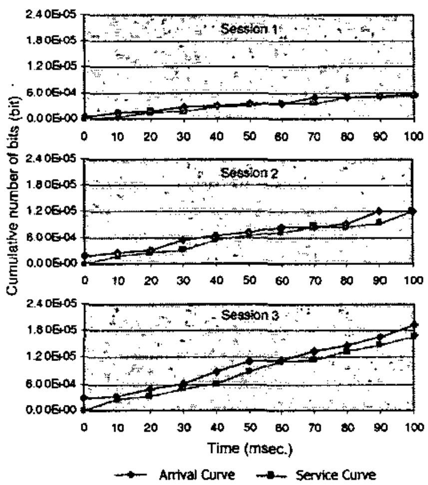


Figure 9: Arrival and service curves for Session 1,2,3

6. CONCLUSION

In this paper we have presented a scheduling algorithm and admission control policy for IEEE 802.16 standard. To the best of our knowledge this is the first such algorithm. The proposed solution is practical and compatible to the existing IEEE 802.16 standard. The simulation studies show that the proposed solution includes QoS support for all types of traffic classes as defined by the standard. We have shown the relationship between traffic characteristics and its QoS requirements and the network performance.

REFERENCES

1. IEEE 802.16 Standard – Local and Metropolitan Area Networks – Part 16, *IEEE Draft P802.16/D3-2001*
2. IEEE 802.16 Working Group on Broadband Wireless Access, http://wirelessman.org/
3. P. Bhagwat, A. Krisna, and S. Tripathi, "Enhancing throughput over wireless LAN's using channel state dependent packet scheduling," in *IEEE INFOCOM 96*, March 1996, pp 1133-1140
4. C. Fragouli, V. Sivaraman, and M. Srivastava, "Controlled multimedia wireless link sharing via enhanced class-based queuing with channel-state dependent packet scheduling," in *IEEE INFOCOM 98*, vol. 2, March 1998, pp 572-580
5. S. Lu and V. Bharghvan, "Fair scheduling in wireless packet networks," *IEEE/ACM Trans. Networking*, vol. 7, no. 4, pp 473-489
6. T.S. Eugene Ng, I. Stoica, and H. Zhang, "Packet fair queuing algorithms for wireless networks with location-dependent errors," *IEEE INFOCOM 98*, March 1998, pp. 1103-1111
7. P. Ramanathan and P. Agrawal, "Adapting packet fair queuing algorithms to wireless networks," *ACM/IEEE MOBICOM 98*, Dallas, TX pp. 1-9
8. J. Gome, A.T. Campbell, and H. Morikawa, "The Havana framework for supporting application and channel dependant QoS in wireless networks," *Proc. ICNP'99*, Nov. 1999, pp. 235-244
9. Y. Cao, and V.O.K. Li, "Scheduling Algorithms in Broad-Band Wireless Networks," *Proceeding of the IEEE*, Vol 89, pp 76-86
10. L. Georgiadis, R. Guerin, and A. Parekh, "Optimal Multiplexing on a Single Link: Delay and Buffer Requirements", *Proc. IEEE INFOCOM' 94* vol. 2, 1994, pp. 524 - 532
11. A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queuing Algorithm," *SIGCOMM CCR 19*, no. 4 (1989).
12. R. L. Cruz, "A Calculus for network delay, Part I: Network elements in isolation," *IEEE Transaction of Information Theory*, 37(1):114-121, 1991