

Experimental Characterization of QoS in Commercial Ethernet Switches for Statistically Bounded Latency in Aircraft Networks

A. Jacobs, J. Wernicke, S. Oral, B. Gordon, A. D. George
{jacobs, wernicke, oral, gordon, george}@hcs.ufl.edu

*High-Performance Computing and Simulation (HCS) Research Lab, Department of Electrical and Computer Engineering,
University of Florida, P.O. Box 116200, Gainesville, Florida 32611, USA*

Abstract

Aircraft networks are used to service mission-critical avionics systems as well as cabin systems such as in-flight entertainment. These networks require that the switches used offer line-rate switching as well as bounded latency and jitter. Gigabit Ethernet offers an interesting replacement to traditional proprietary networks because of its high performance and low cost. We develop a framework for analyzing the abilities of Gigabit Ethernet switches to provide probabilistic guarantees for reliably low latency using Quality of Service (QoS) controls. From the perspective of control capabilities, management complexity, and implementation success we compare the strengths and weaknesses of three modern Ethernet switches. These switches exhibit a broad range of configuration options and are representative of the different levels of QoS implantation available today. We use the performance metrics of latency, jitter, and packet loss to characterize the success of a service policy. The results from a variety of network, load, and traffic scenarios are presented in terms of these metrics. Through the results, we are able to gain insight into the individual switch implementations of QoS.

1. Introduction

Key applications for future cockpit and cabin avionics systems promise to drive the performance and reliability requirements of their integrated networking infrastructure to ever-increasing levels. Avionics networks are used to connect mission-critical systems, and therefore have strict performance requirements [1]. The switches used in these networks must be able to operate at line rate while providing bounded latency and jitter, as well as offer a robust QoS scheme. Specifically, reliably low latencies are essential to handle the growing amounts of critical data. The need for higher bandwidth combined with the importance of cost makes Commercial Off-The-Shelf (COTS) networking technologies desirable. One of the most promising network protocols and technologies to provide these capabilities are those of Ethernet. The economy of scale advantages of Ethernet have led to low equipment costs and ever increasing line rates. While Gigabit Ethernet is currently under consideration, 10-Gigabit and 40-Gigabit Ethernet are also possible future choices. As the link rates increase, new switches and network interface cards will be required, but the network will remain interoperable with older equipment due to protocol backward compatibility.

The standard implementation of Ethernet makes no provision for treating different traffic in different ways. Using QoS, Ethernet switches can identify certain types of traffic and perform actions on the traffic if necessary. The development of Gigabit Ethernet has caused significantly increased levels of performance with much higher data rates, compared to the previous generation of Ethernet, and an increasing degree of support for quality of service. However, most QoS mechanisms have been developed to assure bandwidth and reduce packet loss instead of explicitly providing reliably low latency [2-5]. Further, the choice of implementation algorithms and management features changes the granularity and success of a quality of service implementation. Although committed bandwidth and low packet loss are important, avionics networks also require consistently low latency. Certain mission-critical systems such as navigation may require low latency, while other systems such as in-flight entertainment may require high bandwidth.

Our study analyzes three QoS-enabled Ethernet switches from the perspective of control capabilities, management complexity, and service implementation success. These switches were selected because they represent a range of prices and supported QoS abilities. The control capabilities and management complexity were examined qualitatively. Implementation success will be analyzed statistically in a variety of specific transmission scenarios. Combining these three perspectives, conclusions about the ability of modern COTS Ethernet switches to deliver desired performance guarantees for critical traffic, especially in an avionics setting, will be drawn.

The BayStack 5510-48T from Nortel Networks is an inexpensive 48-port Gigabit Ethernet switch. The switch supports several QoS features including 802.1p user priority and DiffServ. The BayStack can also be stacked to provide up to 384 Gigabit Ethernet ports with a total bandwidth of 640 Gbps.

The Catalyst 2970 from Cisco Systems also offers 24 ports of inexpensive Gigabit Ethernet. This switch represents the entry-level Gigabit Ethernet switch from Cisco. It has support for all of the QoS features that will be tested in this study. The Catalyst also implements a queuing method, shaped round-robin queuing, that is different from the other tested switches.

The E300 switch Force10 Networks is the most versatile switch that was tested. The switch has six removable line cards that can be replaced to upgrade the network hardware. The switch has a 400 Gbps of non-blocking backplane bandwidth with support for twelve Gigabit Ethernet ports per line card. Both 802.1p priority and DiffServ are supported. There are many security features as well as multiple redundant power supplies. This switch is designed to be used in high-bandwidth, mission-critical systems such as internet exchanges or campus backbones.

The organization of the paper is as follows. Section 2 is an overview of QoS terminology and standards for Ethernet. Section 3 describes related work in quality of service guarantees. Section 4 provides an overview of our experimental framework and procedure. Section 5 presents the results and analysis from our experiments. Finally, Section 6 contains the conclusions drawn from the study.

2. Background

Typically, Ethernet provides no performance guarantees and therefore operates on a best-effort basis. Although quality of service can have many definitions, our key interest is in the study of performance-centric network QoS [2]. That is, this paper is interested in the ability of a network to provide specific performance guarantees. In this study, the primary focus is to provide reliably low latency to critical data. Performance guarantees in a physical network are traditionally specified probabilistically [2]. In order to accomplish this task, the metrics of latency, jitter, and packet loss are used. Jitter is defined as the standard deviation of measured latencies.

In QoS, control mechanisms are specified and implemented to provide performance guarantees. In general, QoS mechanisms are categorized according to several key classifications: time scale, granularity, control carrier, and location of control [2]. QoS can work on multiple time scales. A switch can implement QoS on a per-packet basis, or QoS could be provided on the round-trip time scale by providing flow control. The granularity of QoS can be used to assign priorities based upon varying levels of information. QoS could be classified using a low-layer MAC address or using higher-layer information such as a destination port for a TCP packet. Finally, the control can be stored by a switch, or the information can be embedded inside of a packet header.

This paper focuses on packet-level QoS: mechanisms like classifiers, markers, and shapers that improve packet transfer performance. Classifiers are elements of a system that determine what level of service should be given to a specific packet. Markers can, if desired, use the results of the classifier to mark the packet header permanently to pass the classification to the next switch. Finally, shapers moderate the packet to provide the proper level of service inside the switch. Figure 1 shows the process that a switch uses to implement QoS.

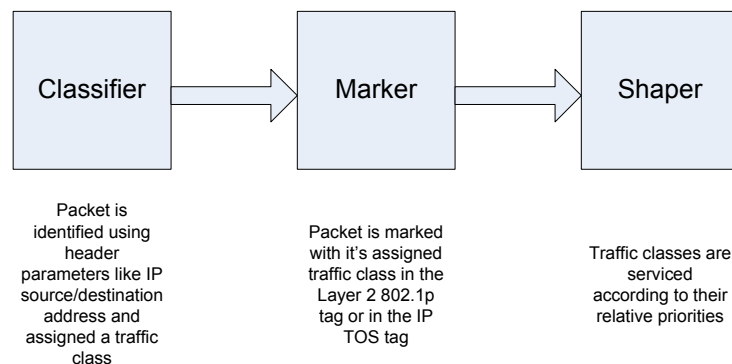


Figure 1: Classification, marking, and shaping.

There are two general levels of granularity of the control state of a QoS control mechanism: per-flow and aggregate. A per-flow control state provides different service for each *flow*, which is defined by an IP source and destination address, source and destination port, and protocol field. Aggregate control states combine several flows together and provide each flow in the group equal service to match the desired group profile. Thus, the level of flow aggregation greatly affects the fidelity and complexity of the QoS implementation [6-7].

The final category of QoS classification is the location of the control. Almost all current Ethernet QoS implementations use the switch to store control state information [4, 8]. For this report, the switch is the sole location of control.

A number of Ethernet QoS standards have been developed to provide sophisticated control over switching performance. First, a relatively simple system, the IEEE 802.1p standard, was created as part of the IEEE 802.1D [9]. 802.1p uses three bits from the Layer 2 tag to differentiate 8 levels of service. So, in a complex network many different types of traffic must be aggregated into a single 802.1p service class.

To provide greater fidelity of control, integrated services (IntServ) was proposed [3]. The IntServ standard required a packet to schedule a path of appropriate resources using the Resource ReSerVation Protocol (RSVP) before transmission so that guarantees or predictions of service can be made. Unfortunately, there are two drawbacks to using IntServ. First, it requires a large amount of data to be maintained in the switch about the current state of each flow. Second, each intermediate switch on the packet's path must implement IntServ. These drawbacks can be alleviated by decentralizing the QoS control mechanisms with a QoS implementation like DiffServ.

The differentiated services (DiffServ) standard extends the three-bit 802.1p marking to a full byte to provide 64 different classes of services [4]. DiffServ also specifies certain per-hop behaviors that can be implemented to assure service to each class. While the Class of Service field for an 802.1p packet is part of the MAC header, the DiffServ classification uses the Type of Service (ToS) field located in the IP header. In contrast to IntServ, DiffServ stores the QoS information inside of the packet header, allowing QoS packets to pass through networks without QoS.

An even more recent standard is Multi-Protocol Label Switching (MPLS). MPLS is a Layer 3 standard that tags each packet with an MPLS tag that holds a specific routing path [8]. Like integrated services, MPLS switches rely upon a full path of MPLS-enabled machines to provide performance guarantees. To date, relatively few COTS switches implement MPLS due to slow adoption of the specification by end users.

In general, most switches implement both 802.1p and DiffServ and treat traffic of both types in a consistent manner. Most switches implement some variation of a priority queuing system to provide quality of service control. Ingress and egress queues are serviced according to internal priority mappings and queuing schemes. The priority mapping is used to map priority bits from any supported type of QoS to the proper destination queue for the incoming packet. All three of the switches analyzed in this study follow this pattern. MPLS is not implemented in most Gigabit Ethernet switches yet. Figure 2 is a graphical representation of a switch implementation of classifying, marking, and shaping using a packet-based QoS system.

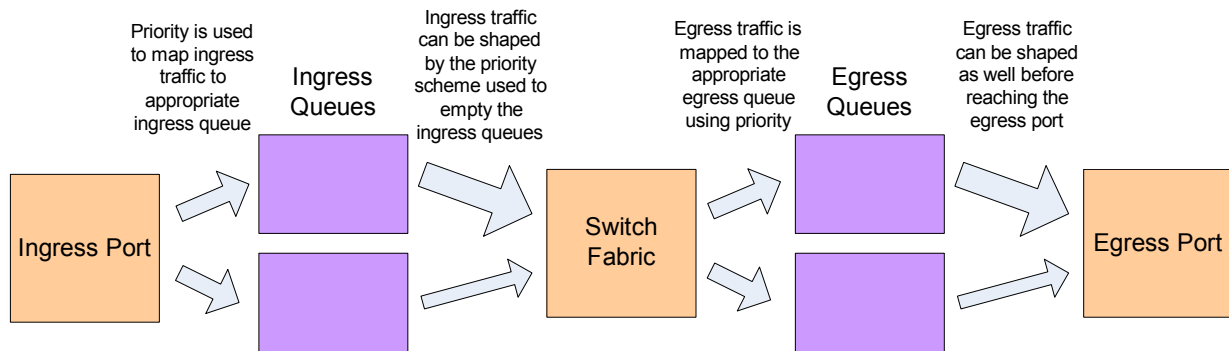


Figure 2: Generic packet-based QoS prioritization in a switch.

3. Related Work

A number of recent studies have explored QoS guarantees. Most of these papers concentrate on absolute guarantees in theoretical networks, but a few also discuss probabilistic guarantees in a real network. The following papers analyze performance guarantees in a theoretical or analytical method. In [5], T. Bauge, *et al.*, provide a general overview of several key areas of QoS and apply theoretical analysis to satisfy the performance demands of a specific application. It was determined that a guaranteed QoS service would be infeasible, but that IntServ would provide adequate service as long as the network was not over-provisioned. A. Jarayya, *et al.*, develop an integrated service protocol in [10] as well as evaluate the importance of resource allocation and scheduling. Using a reservation-based protocol, the effects of different scheduling protocols, such as weighted fair queuing or strict priority queuing, on different types of traffic are measured.

Several studies have also focused on using simulation or experimental analysis on performance guarantees. V. Laatu, *et al.*, experimentally analyze the effects of a specific DiffServ mechanism on flows of similar priorities using

latency and throughput as metrics [11]. Certain types of traffic were found to be more sensitive to the QoS policies than others. C. Bouras, *et al.*, use estimation to provide theoretical performance guarantees and then use simulative results to assess the accuracy of their predictions [12]. In [13], V. Firoiu, *et al.*, provide a framework for evaluating traffic engineering using modeling and then validate the model using simulation.

The use and development of performance metrics is a considerably large area of study. T. Chahed discusses in more detail performance metrics in [14]. Also, for a more complete study, refer to the Internet Protocol Performance Metrics (IPPM) RFCs. A good starting point is a “Framework for IP Performance Metrics” [15].

Previous work [1] compared several Gigabit Ethernet switches using latency and jitter as metrics. It was found that best-effort service between under-subscribed nodes exhibits low latency and line-rate switching. A very simple priority system was introduced, but latency was only measured on a single switch.

4. Experimental Framework

In order to evaluate the capability of a switch to provide probabilistic performance guarantees, both the management features and performance of each switch are evaluated. The fidelity and complexity of control mechanisms can vary considerably from switch to switch. For each switch, a short description of the management capabilities and the granularity of control is provided. In particular, we evaluate to what degree the switch can be configured to match our chosen scenarios and then generalize our conclusions to how well the switch configuration can solve any arbitrary QoS problem.

Section 5 of this paper will concentrate on the goal of evaluating the ability of COTS Ethernet switches to provide reliably low latency for high-priority streams in a congested network. Our experiments concentrate on developing a better understanding of which preconditions a switch can provide the requisite performance of statistically low latency and jitter. The metrics of mean latency and jitter will be critical to identifying expectations of performance. Furthermore, we use maximum and minimum latencies as well as the variation of measured mean latencies to analyze the reliability of the observed results.

Figure 3 shows the general setup for the experimental case studies in this paper. Several key many-to-one contention scenarios are used to study the ability of QoS controls in the switches to provide performance guarantees to specific flows. These scenarios are abstract models of more complex real QoS problems where many different flows compete for a limited number of resources.

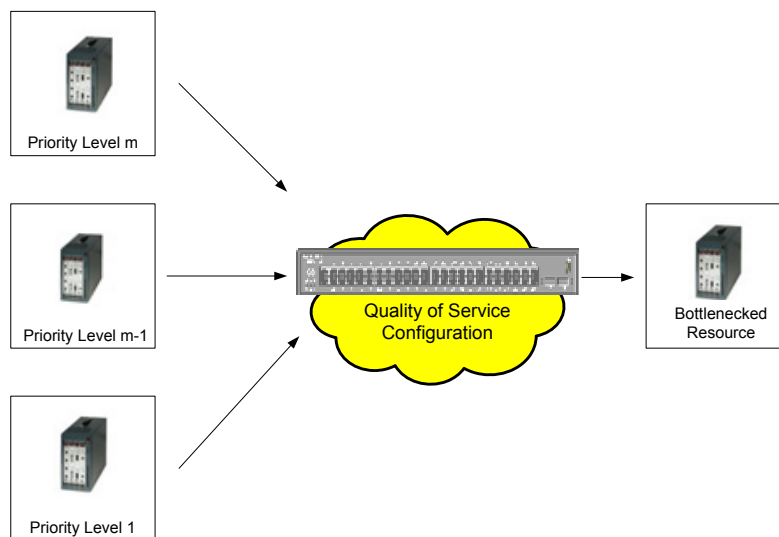


Figure 3: Conceptual diagram of the experimental case studies.

First, all flows were given best-effort service to provide a control set for our other measurements. This approach provides a baseline so that the results with QoS enabled can be put into a proper context. This test case will be referred to as the best-effort case. Then, a single *platinum* stream was given highest priority while leaving all other streams at lowest priority. The highest priority corresponds with an 802.1p value of 7, and the lowest priority corresponds with a value of 0. The 802.1p priority was set by the switch based on the source address of the incoming packets. The platinum stream models a single flow of critical data competing with less critical data for access to an egress port. This test case will be referred to as the single platinum stream case.

Additionally, multiple medium-priority streams (with an 802.1p value of 6) were generated with one platinum stream still holding the highest priority. These medium-priority streams may be referred to as *gold* streams for easier reference. This configuration examines the effect of priority assignments with weighted round-robin (WRR) or shaped round-robin (SRR) queuing on performance guarantees for both the platinum stream and the gold streams. Weighted round-robin assigns a weight to each queue, and then services the queues in a round-robin fashion based upon the weight. This test case will be referred to as the mixed priority case.

For consistency, streams with a priority of 7 will be referred to as platinum streams for the remainder of this paper. Streams with a priority of 6 will be called gold streams. Streams with a priority of zero will be referred to as best-effort streams.

To analyze latency and jitter at the packet level, an Ixia 400T Traffic generator with 8 Gigabit Ethernet ports was used [6]. The Ixia 400T is capable of measuring received latency to the nearest 20 nanoseconds. The Ixia chassis and ports were configured using a custom Tool Control Language (Tcl) script, which can be found in the appendix, designed to accurately measure real-time latency and packet loss. Tcl scripts were used because they provided more control over the types of latency measurements that could be made. This approach also permitted the automation of tests to facilitate data collection. Seven Ixia ports, used as transmission nodes, compete for the egress port of the switch to the eighth node, which is the eighth Ixia port. Thus, when each port is set at 15% of the full line rate, the reception port will be congested with 105% of its own line rate. The performance of each switch was measured both before and after the saturation point. The latency was measured from the time the first bit of the packet leaves the Ixia transmission port until the time that same bit reaches the Ixia receive port. Latencies were measured for successfully received packets only; dropped packets do not affect latency measurements. Packet sizes of 128, 512, and 1518 bytes were used and the line rate of the transmitting ports was varied from 0 to 30 percent. Trials were run three times for each data point. There were no significant variations between trials except where noted in Section 5.

5. Results and Analysis

The performance of three COTS Gigabit Ethernet switches was analyzed. In Section 5.1, we show the performance characteristics of the Nortel Networks BayStack 5510. The latency, jitter, and packet loss under the conditions listed in the previous section were determined. In Section 5.2, the results from the Cisco Systems Catalyst 2970 tests are analyzed. Section 5.3 contains the results of the final switch tested, the Force10 Networks E300. Section 5.4 analyzes the performance differences between the separate testbeds.

5.1 Nortel Networks BayStack 5510

Intended as an edge switch rather than a core router, the BayStack 5510 offers fairly straightforward quality of service control [7]. The 5510 supports both 802.1p and DiffServ classification and marking. The switch enforces policies using 8 egress queues which can be configured to have a strict priority or WRR queue emptying scheme. Strict priority will always favor packets with the highest 802.1p value, while the WRR approach will give some access to the lower-priority queues. Each queue can be assigned a specific amount of bandwidth to be given in order to provide more fidelity to the QoS implementation. Depending on the packet's 802.1p/DiffServ priority, the packet is mapped to one of the queues by a configurable mapping.

The switch by default disables quality of service but has a simple priority mapping in memory to make setting up a WRR priority scheme simple. A WRR scheme is usually preferred so that lower priority streams do not starve under heavy loading conditions. Since the hypothetical platinum stream represents critical data in an avionics network, the highest priority should be allocated to the stream. At the same time, the low-priority traffic should not starve whenever the platinum stream is transmitting. Based on the requirements of the platinum stream, network engineers could adjust the WRR queuing. Since the 5510 only has 8 queues, all traffic must be grouped into 8 classes of service for shaping.

5.1.1 BayStack 5510 Best-Effort Performance Results

The following results were measured when the BayStack 5510 was configured to give best-effort service to all of the seven flows. Figures 4 and 5 show the minimum, mean, and maximum latency of the transmission packets for 128-byte and 1518-byte packets, respectively. Below saturation of the receiving link, less than 15% line rate per node, the switch produces average latencies of 9 and 59 microseconds for small and large packets, respectively.

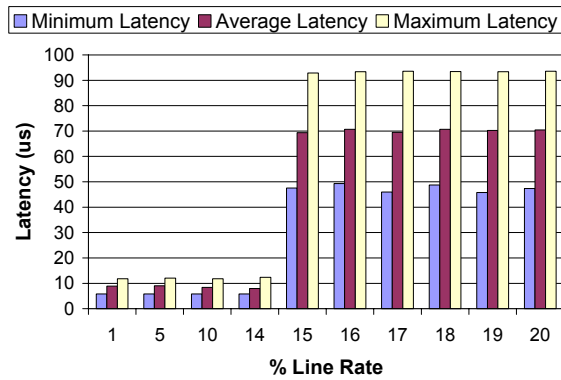


Figure 4: Best-effort 128-byte packet latency.

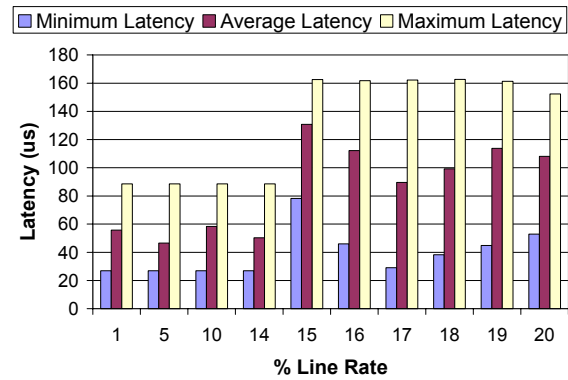


Figure 5: Best-effort 1518-byte packet latency.

Below saturation, average latencies are around 8-9 microseconds for 128-byte packets and the maximum latency measured is 12 microseconds. For 1518-byte packets, the average latency is around 50 microseconds and the maximum latency observed was 89 microseconds. Above saturation, the latencies are much higher. However, these latencies are only those of successful packets since dropped packets cannot be measured. Without using QoS controls, the latency of important data can increase ten times if the channel becomes momentarily saturated.

The packet loss from best-effort is 0% for all line rates below saturation as shown in Figure 6. The packet loss increases steadily after saturation however. This outcome provides a baseline to be able to quantify improvements from using QoS controls.

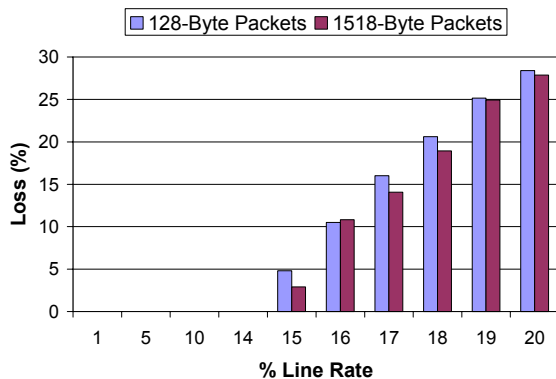


Figure 6: Best-effort packet loss.

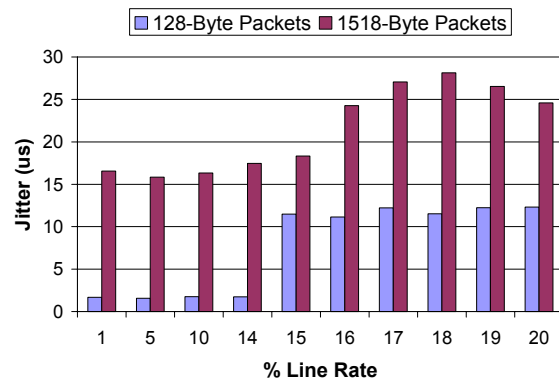


Figure 7: Best-effort jitter

Figure 7 is the measured jitter of the packets in the best-effort case. For small packets the jitter is extremely low (less than 2 microseconds) below saturation but afterwards increases more than five times. Large packet jitter is much higher than small packets below saturation (around 17 microseconds). After saturation, large packet jitter increases by around 10 microseconds.

Best-effort is clearly not suitable for delivering performance guarantees to critical data. High latencies, dropped packets, and high jitter would cause the breakdown of key applications under high traffic conditions. QoS controls are needed in order to ensure reliably low latency independent of line rate.

5.1.2 BayStack 5510 Single Platinum Stream Performance Results

Once the BayStack 5510 was configured to give the platinum stream the highest priority, and the other six streams all with simply best-effort service, the after-saturation latencies for the platinum stream were found to decrease dramatically versus the previous best-effort case. For small packets below saturation, the mean latency remained around 8 microseconds as seen in Figure 8. However, after saturation, the mean latency decreased from about 70 microseconds (in the best-effort case) to around 11 microseconds. The highest observed latency was under 15 microseconds. The QoS controls of the BayStack 5510 are successful in lowering the latency of a single platinum stream in an environment of congestion with lower-priority streams.

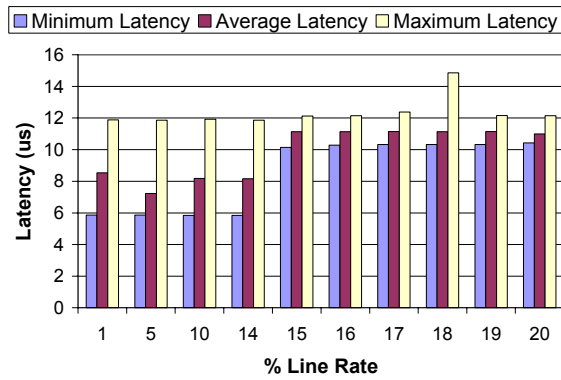


Figure 8: Single platinum stream 128-byte packet latency.

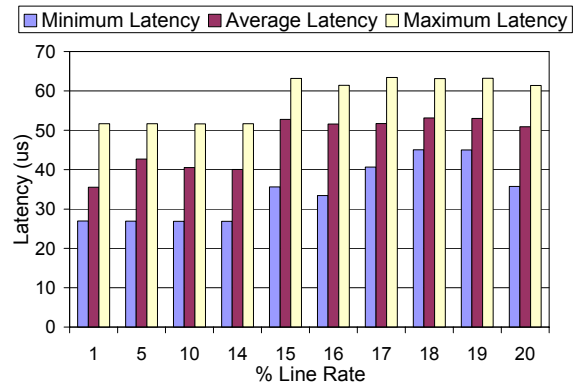


Figure 9: Single platinum stream 1518-byte packet latency.

For larger packets, Figure 9 shows that mean latencies have been reduced from over 100 microseconds to slightly above 50 microseconds. Thus, for both large and small packets, giving the platinum stream high priority allows the critical data to be assured low latencies. Also, neither packet size of the priority stream experienced any packet loss even after saturation, ensuring reliable delivery.

Figure 10 shows the jitter for the platinum stream. The jitter of the platinum stream before saturation is lower than in the best-effort case, and it is reduced further after saturation. Each packet from the platinum stream is given equal treatment and experiences approximately the same background load in the switch. By successfully reducing the latency and by radically decreasing the jitter, the critical application would be able to continue to operate successfully even after heavy loading on this switch.

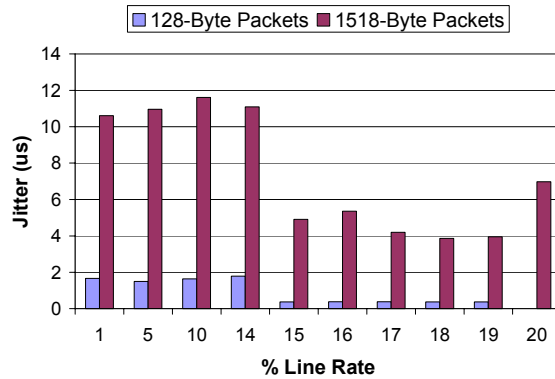


Figure 10: Single platinum stream jitter.

5.1.3 BayStack 5510 Mixed Priority Performance Results

In this case of one platinum stream and six gold streams, the platinum stream receives approximately the same behavior under light loading conditions as in the single platinum stream case. Figures 11 and 12 show that average latencies were below 12 microseconds for 128-byte packets and slightly above 50 microseconds for 1518-byte packets. When more traffic uses high priority levels, the highest priority will maintain low latency and jitter.

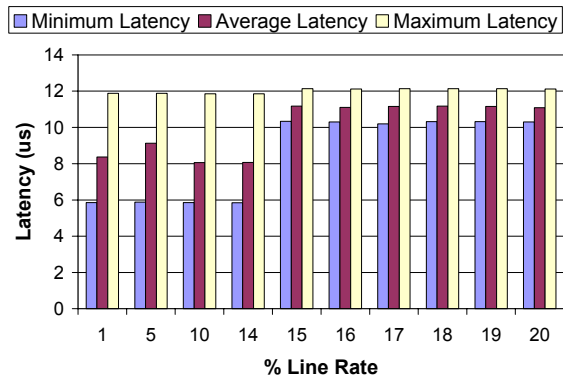


Figure 11: Platinum stream 128-byte latency.

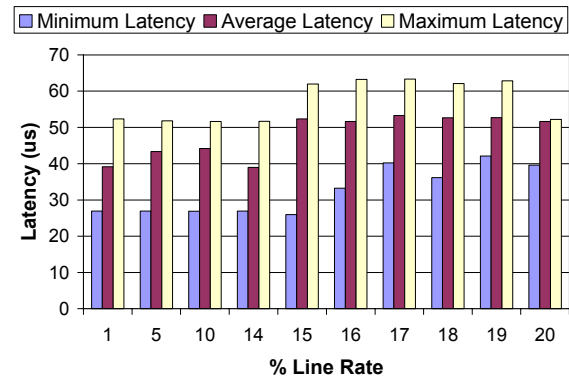


Figure 12: Platinum stream 1518-byte packet latency.

The platinum stream also has approximately the same jitter as in the single platinum stream case as demonstrated by a comparison of Figures 13 and 10. After saturation, the jitter of the platinum stream is reduced. With QoS enabled the amount of jitter after saturation remains below 5 microseconds, worst case. There was no packet loss at any of the tested line rates.

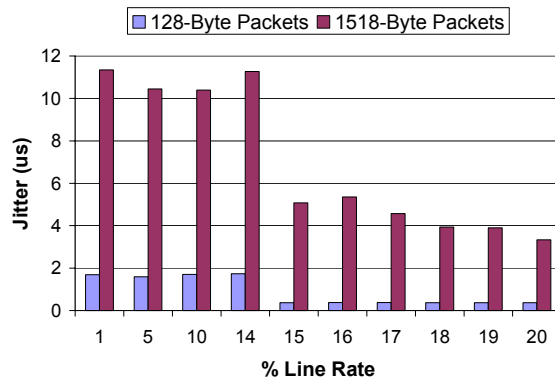


Figure 13: Platinum stream packet jitter.

5.2 Cisco Catalyst 2970

The Catalyst 2970 features a slightly more complicated quality of service control than the BayStack 5510 in terms of possible configurations. Like the 5510, the Catalyst 2970 features 802.1p and DiffServ classification and marking. The switch offers the ability to queue the packets at ingress according to priority as well as at the egress port. Two ingress ports and four egress queues are available at each port. Using the packet's 802.1p/DiffServ priority, the packet is mapped to one of the queues based on the current priority mapping.

The Catalyst 2970 has a configurable SRR priority scheme to empty each buffer at both the ingress and egress queues. The SRR algorithm specifies a maximum amount of bandwidth that a queue may use. However, if other queues are empty, the bandwidth may be shared. The control over the queues is more complicated than the configuration of the BayStack 5510, but also has greater fidelity since you can specify both the SRR scheme and also the size of each queue. Multiple ingress queues would be useful for distinguishing between different flows that arrive on the same port, but with only four egress queues, multiple flows must be aggregated into a single class of service for shaping.

5.2.1 Cisco Catalyst 2970 Best-Effort Performance Results

The performance of the Catalyst 2970 was relatively equal to the BayStack 5510 below saturation; both switches produce approximately 9 microseconds of average latency. However, as Figures 14 and 15 demonstrate, the latencies after saturation for the Catalyst 2970 are more than two orders of magnitude greater than the latencies before saturation. The Catalyst 2970 is able to queue more packets and since, only successful packets are measured, the higher latencies reflect a larger queuing delay while waiting to reach the egress port. Below saturation, small packets have around 9 microsecond latency and large packets have around 45 microsecond latency. The maximum

latency for small packets below saturation was around 15 microseconds and for large packets was slightly below 100 microseconds.

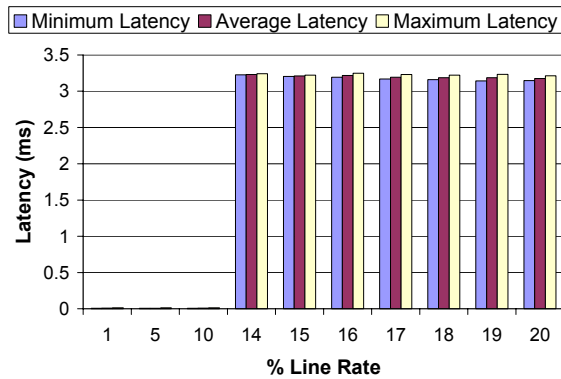


Figure 14: Best-effort 128-byte packet latency.

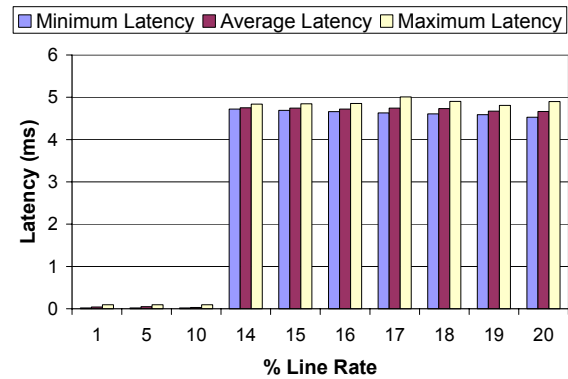


Figure 15: Best-effort 1518-byte packet latency.

The packet loss from best-effort is 0% for all line rates below saturation as shown in Figure 16. However, after saturation, the small packet loss is typically larger than the large packet loss. While the packet loss increases with the line rate, it does not follow in a strictly linear fashion. The reason for this behavior is not evident; it is dependent on the switch’s queuing strategy and implementation. This behavior was not demonstrated in any of the other tested switches.

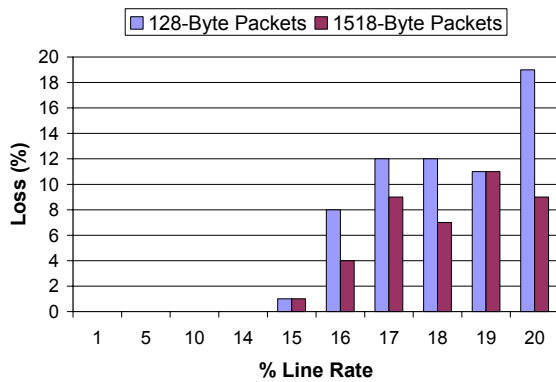


Figure 16: Best-effort packet loss.

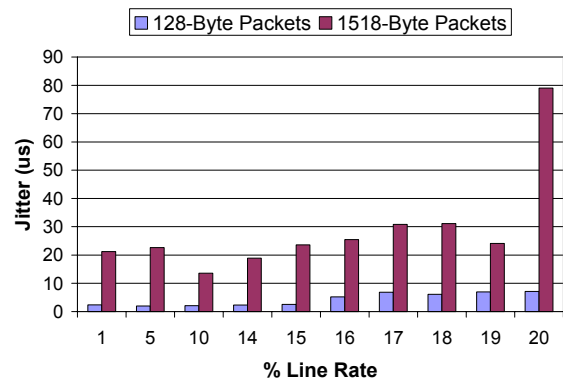


Figure 17: Best-effort jitter.

Figure 17 shows the jitter of the packets for the best-effort case. For small packets the jitter is less than 3 microseconds below saturation and reaches over 7 microseconds at 20% line rate per flow. Large packet jitter is lower than 25 microseconds below saturation but after saturation increases slightly. At 20% line rate, the large packet jitter is very large but since only successful packets are measured and the latencies in question are very high the jitter this far above saturation is not a very meaningful metric. Without QoS controls enabled, the Catalyst 2970 produces highly variable jitter after saturation.

5.2.2 Cisco Catalyst 2970 Single Platinum Stream Performance Results

Using QoS to give highest priority to the platinum stream again resulted in greatly reducing the latencies of both small and large packets after saturation. After saturation, the mean latency was below 13 microseconds for small packets and below 45 microseconds for large packets as seen in Figures 18 and 19. The highest observed latency was under 15 microseconds for small packets and under 54 microseconds for large packets. Once again, neither packet size showed any packet loss after saturation.

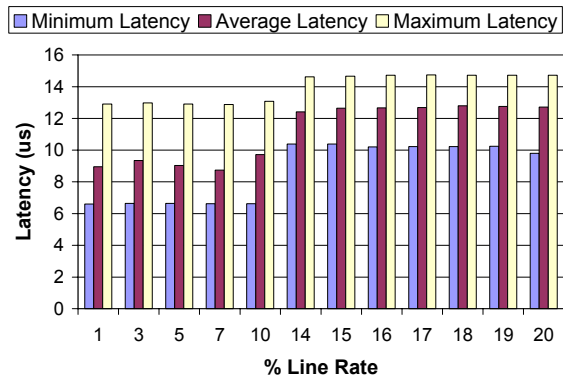


Figure 18: Single platinum stream 128-byte packet latency.

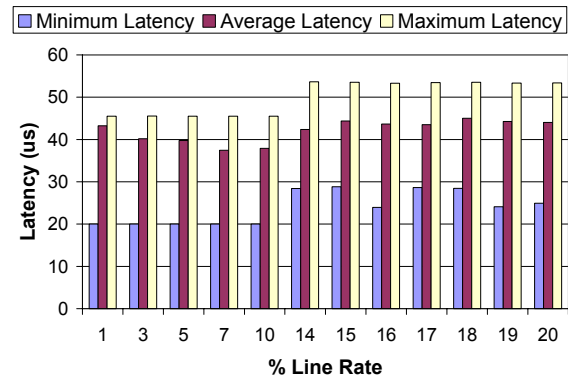


Figure 19: Single platinum stream 1518-byte packet latency.

Figure 20 shows that the jitter remains reduced even after saturation. Once again the jitter is lower after saturation for both small and large packets at less than a microsecond. Thus, the Catalyst also was able to deliver low latency and low jitter to the critical data. For 128-byte packets, the jitter was less than 2 microseconds at all measured data points. The 1518-byte packets experienced a maximum jitter of 10 microseconds. After the saturation point, the maximum jitter was reduced to less than 5 microseconds.

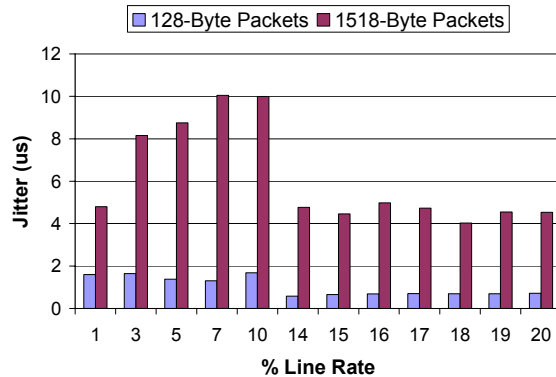


Figure 20: Single platinum stream jitter.

5.2.3 Cisco Catalyst 2970 Mixed Priority Performance Results

The platinum stream still receives good performance when competing with the six gold streams. Figures 21 and 22 show the latency results for the mixed priority experiments. Below saturation, the 128-byte platinum stream has a latency of 10 microseconds, which is 1 microsecond higher than the single platinum stream case. This performance result is evidence that latency will increase when the amount of prioritized traffic increases. After saturation, the latency decreases to 8.5 microseconds. One possible explanation for the decrease in latency of the platinum stream is that packets from the gold streams are dropped at a much higher rate than the best-effort case.

Figure 23 shows that the jitter for the platinum stream is also very low in the mixed priority case. For small packets, the jitter is below 2 microseconds and for large packets below 5 microseconds. The Catalyst 2970 was able to deliver low latency and jitter to the platinum stream more efficiently in this case of mixed priorities than it was in the single platinum stream case where the competition for bandwidth was from best-effort streams.

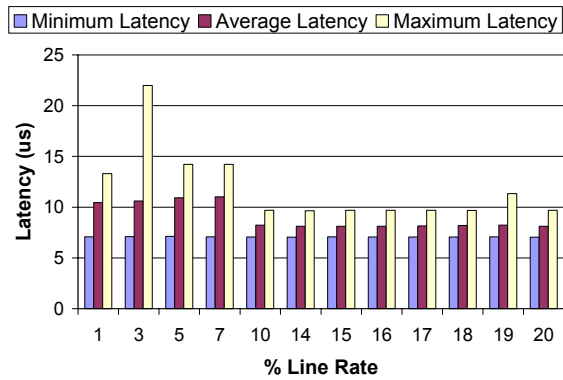


Figure 21: Platinum stream 128-byte packet latency.

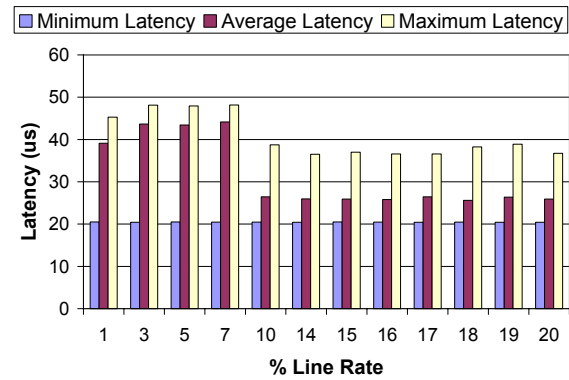


Figure 22: Platinum stream 1518-byte packet latency.

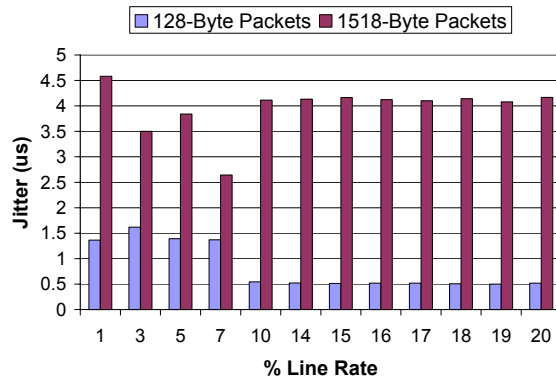


Figure 23: Platinum stream packet jitter.

5.3 Force10 E300

The Force10 E300 is quite different from the other two switches in our study. Intended as a core switch with multiple 10-Gigabit Ethernet ports, the E300 has large buffers capable of storing up to 200 milliseconds worth of data [16]. As in the Catalyst 2970, after classification, the switch chooses which packets to send from the ingress queues based on priority and congestion avoidance. Congestion avoidance consists of Random Early Drop (RED) or Weighted RED (WRED). These methods eliminate packets before a port reaches saturation in order to protect high priority streams from packet loss. Then, the packet is placed in one of 8 egress queues where egress traffic is also shaped. All of the mappings are configurable with a default that satisfies simple QoS problems.

The E300 allows the bandwidth percentages assigned to each egress queue to be set in the configuration of the switch. The granularity is 1%, so using the bandwidth percentage command does not give quite as much control to the network engineer. Also, one can specify the committed, peak, and burst rates allowed for each ingress and egress class so that a more complicated scheme can be created.

5.3.1 Force10 E300 Best-Effort Performance Results

The Force10 performed quite differently from the other switches in our study. Figures 24 and 25 show the small and large packet latencies, respectively. Below saturation, the mean latency was around 45 microseconds for small packets and around 95 microseconds for larger packets. Also, below saturation the maximum latencies observed for small packets was slightly below 66 microseconds and for larger packets was slightly above 134 microseconds. After saturation the E300 had very high latencies above 10 milliseconds starting at 16% line rate per flow.

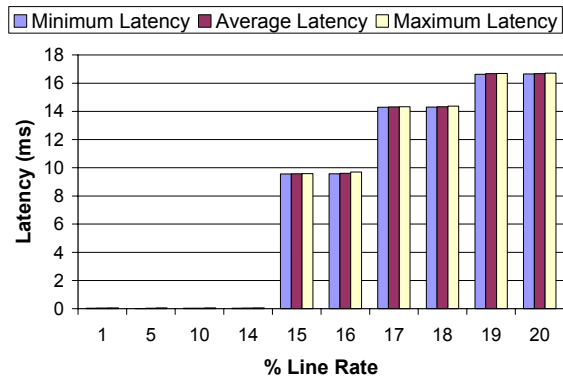


Figure 24: Best-effort 128-byte packet latency.

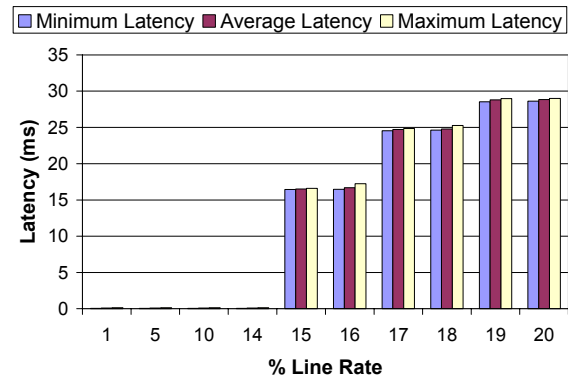


Figure 25: Best-effort 1518-byte packet latency.

The packet loss from the best-effort stream is 0% for all line rates below saturation as shown in Figure 26. For the E300, the small and large packet loss is almost identical for each line rate. The jitter is quite high both below and above saturation as seen in Figure 27. Below saturation, jitter was observed as high as 8 microseconds for small packets and around 27 microseconds for large packets. At 16% line rate, the jitter has a very large value in the 1518-byte case, and the variation of measured jitter from test to test was at worst 45%. The minimum jitter recorded at this line rate was 78 microseconds. It is evident that the 16% line rate case exhibits different jitter characteristics than the other cases measured. Since this value is above saturation with no quality of service, the reasons for this variation were not investigated further.

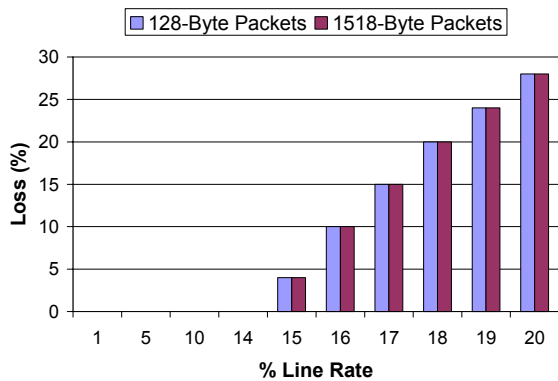


Figure 26: Best-effort packet loss.

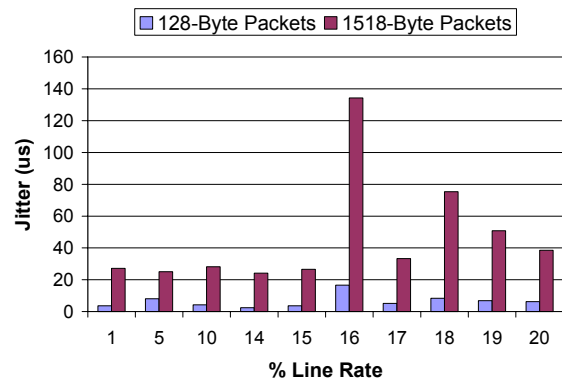


Figure 27: Best-effort jitter.

5.3.2 Force10 E300 Single Platinum Stream Performance Results

The platinum stream experienced latencies that were much lower than in best-effort, but still ten times higher than the Catalyst or the BayStack results. Interestingly, in both large and small packets, adding quality of service reduced the latencies below saturation as well as above saturation. This trait shows that the QoS is always active, not only after saturation. Once again, no packet loss occurred for the platinum stream. The latency results are shown below in Figures 28 and 29.

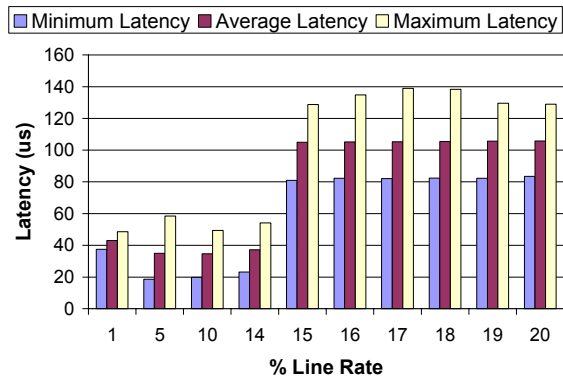


Figure 28: Single platinum stream 128-byte packet latency.

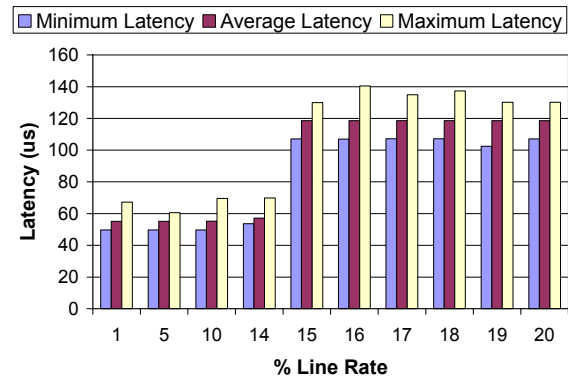


Figure 29: Single platinum stream 1518-byte packet latency.

The single platinum stream configuration bounded the observed jitter to below 12 microseconds for small packets and below 5 microseconds for large packet sizes, as seen in Figure 30. Interestingly, the packet jitter above saturation was increased by about 2 microseconds. The 128-byte jitter is not significantly changed from the best-effort case, although the 1518-byte packets experience between 5 and 10 times less jitter. This difference indicates that the E300 is optimized to support larger packet sizes.

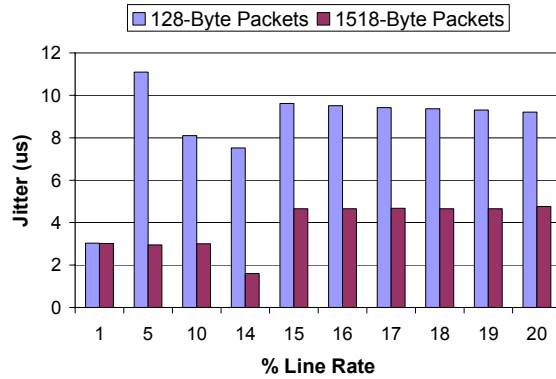


Figure 30: Single platinum stream jitter.

5.3.3 Force10 E300 Mixed Priority Performance Results

The platinum stream does not receive quite as good performance as the stream in the single platinum stream case. Comparing Figures 31 and 32 with Figures 28 and 29, small packet latency above saturation increased by about 50 microseconds and large packet latency above saturation increases by about 30 microseconds. The maximum latencies below saturation for 1518-byte packets are also much higher than in the single platinum stream case.

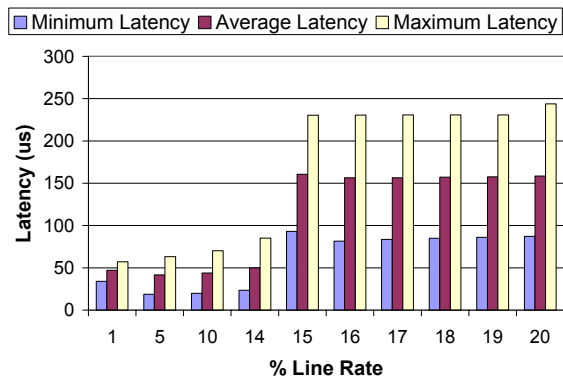


Figure 31: Platinum stream 128-byte packet latency.

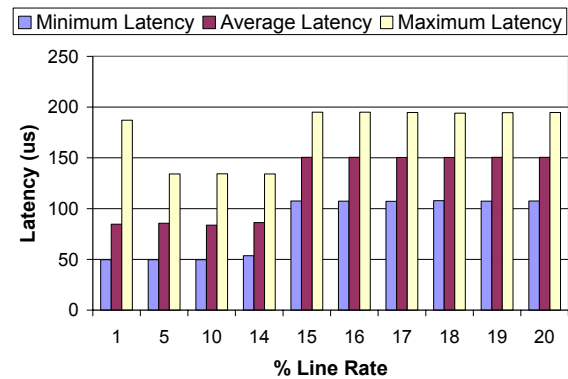


Figure 32: Platinum stream 1518-byte packet latency

Figure 33 demonstrates that the jitter is also relatively high. For the small packets, the jitter is higher after saturation, but for large packets the jitter is better regulated after saturation. The after-saturation case is similar to the single platinum stream case, where the 128-byte packets exhibit higher latency than the 1518-byte packets. Further study is needed in order to explain the below-saturation performance.

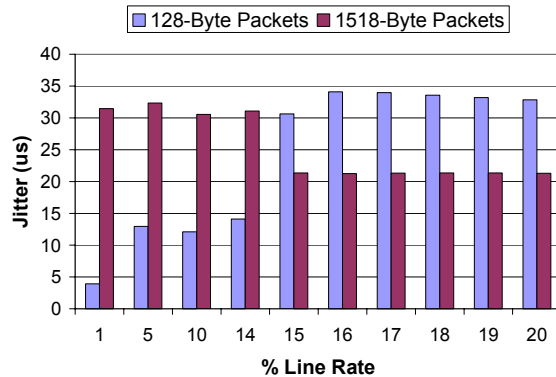


Figure 33: Platinum stream packet jitter.

5.4 Switch Comparisons

This section will present the QoS performance comparisons between the Nortel Networks BayStack 5510, the Cisco Systems Catalyst 2970, and the Force10 Networks E300. The switches are compared on their performance in terms of average latency, jitter, and packet loss. Also, comparisons of performance in two different situations, below saturation and above saturation, are analyzed. Results are shown from the best-effort, single platinum stream, and the mixed priority streams cases. By comparing the performance of the switches, insight into the algorithms that control QoS can be gained.

5.4.1 Best-Effort Below Saturation

Switches will operate below saturation given normal traffic conditions. Therefore, the ability of a switch to provide reliably low latencies and low jitter below saturation are of particular interest. Figures 34 and 35 show the best-effort latencies for small and large packets for all three switches in this study. Notice that the BayStack 5510 and the Catalyst 2970 have similar latencies but the E300 has consistently higher latencies. Although the E300 has the highest latencies measured, the difference between the other switches is less when larger packet sizes are used. This difference indicates that the Force10 switch favors larger packet sizes.

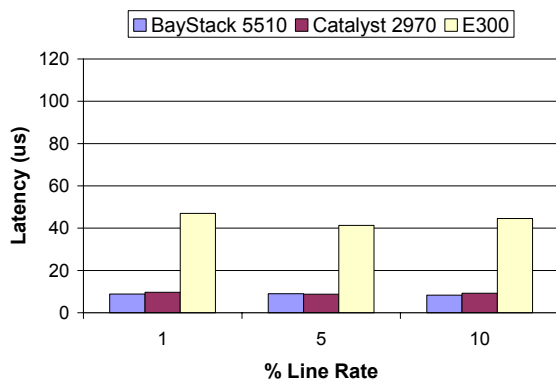


Figure 34: Best-effort 128-byte packet latency.

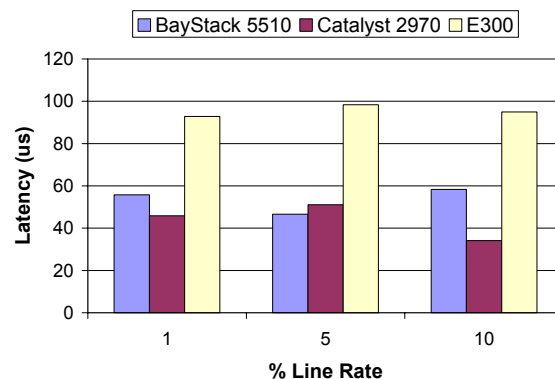


Figure 35: Best-effort 1518-byte packet latency.

Comparing the jitter below saturation under best-effort conditions shows that all three switches provide relatively consistent latencies. Figures 36 and 37 show that the E300 jitter is slightly higher than the other two switches but not by the same relative differences as found in latency. The BayStack 5510 typically has the lowest jitter. Without QoS controls enabled, the BayStack 5510 is able to provide the lowest latency and jitter compared to the other tested switches.

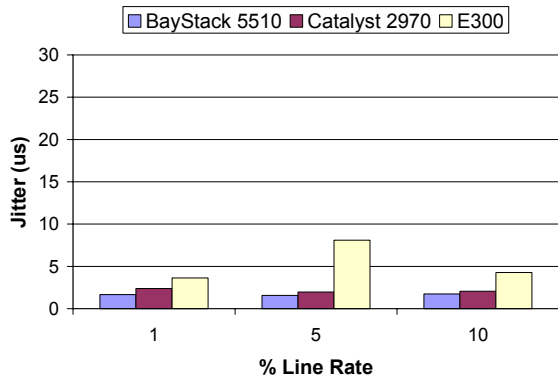


Figure 36: Best-effort 128-byte jitter.

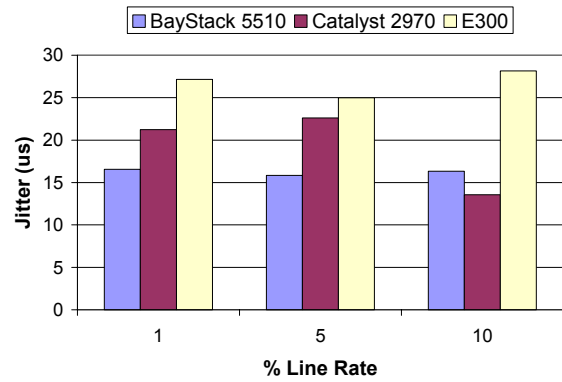


Figure 37: Best-effort 1518-byte jitter.

The packet loss observed from each switch for the best effort had several interesting results. Figures 38 and 39 show the packet loss for each switch. The packet loss for the BayStack 5510 and E300 were roughly equivalent. The Catalyst 2970 had lower packet loss than both. While the packet loss is expected to rise linearly with increased line rate, the Catalyst 2970 does not exhibit this behavior. This behavior may be caused by the way the switch buffers data. However, if this was the case, the E300 would be expected to offer similar results. The E300 may be using congestion avoidance techniques such as RED, but further investigation is needed.

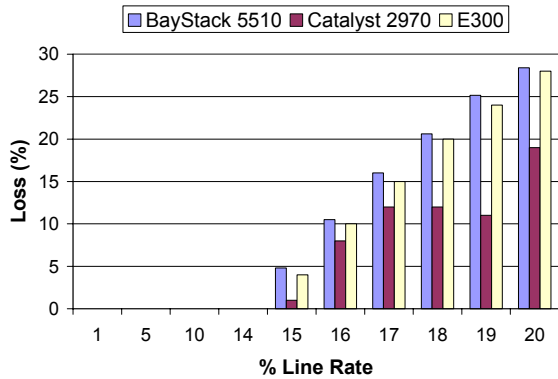


Figure 38: Best-effort 128-byte packet loss.

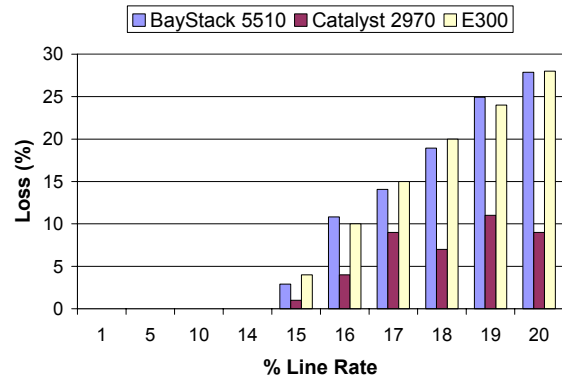


Figure 39: Best-effort 1518-byte packet loss.

5.4.2 Single Platinum Stream Below Saturation

How the switch performs below saturation after QoS is enabled is quite important because this is the normal mode in ordinary operation. The platinum stream latencies are shown for all three switches in Figures 40 and 41. After QoS is enabled, the BayStack and Catalyst latencies are one microsecond less than the best-effort case. However, the latency of the E300 is approximately 10 microseconds lower. While the Force10 switch has higher latencies, it also benefits more from the use of QoS.

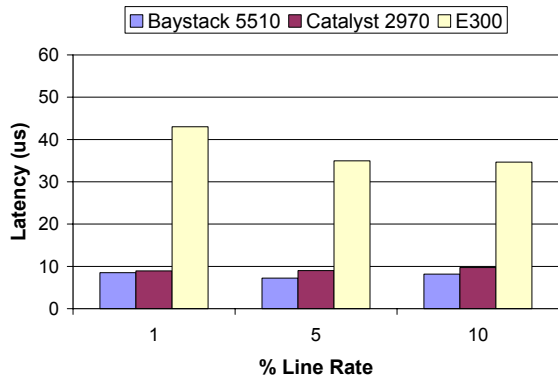


Figure 40: Single platinum stream 128-byte packet latency.

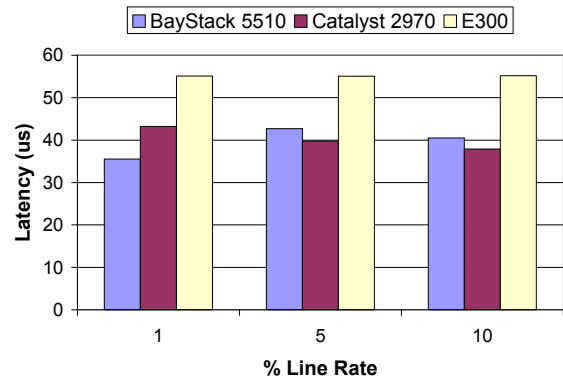


Figure 41: Single platinum stream 1518-byte packet latency.

The single platinum stream jitter results are shown in Figures 42 and 43. For 1518-byte packets, the jitter is reduced significantly after applying QoS for the Catalyst and E300. The BayStack jitter is also reduced by five microseconds, but exhibits the highest jitter observed. The packet loss for all three switches was 0% for the platinum stream. Thus, all three are capable of properly preventing packet drop with QoS for critical data.

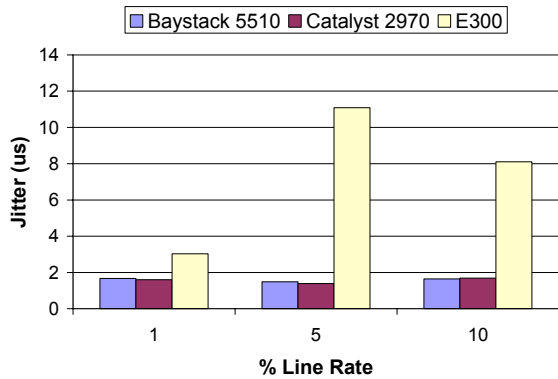


Figure 42: Single platinum stream 128-byte jitter.

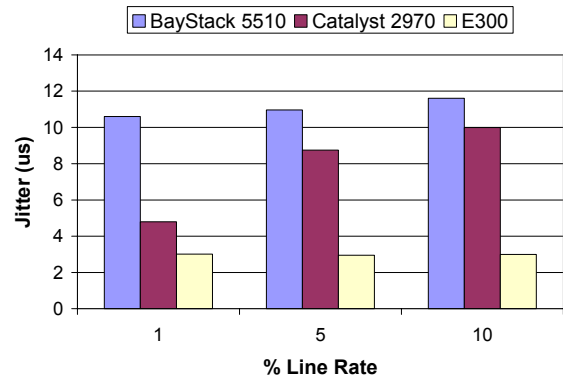


Figure 43: Single platinum stream 1518-byte jitter.

5.4.3 Single Platinum Stream Above Saturation

The latency and jitter performance of all three switches, both below and above saturation, was examined. The data measured was that of the platinum stream in this mix. Figure 44 shows that the BayStack and Catalyst vastly outperform the E300 in latency above saturation for small packets. Figure 45 shows that the difference is not quite as serious in larger packets, but small packets could be important to an avionics network.

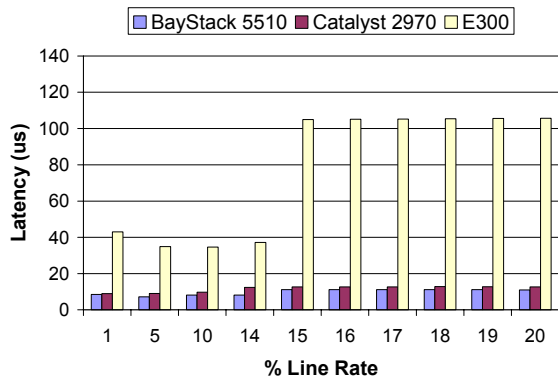


Figure 44: Single platinum stream 128-byte packet latency.

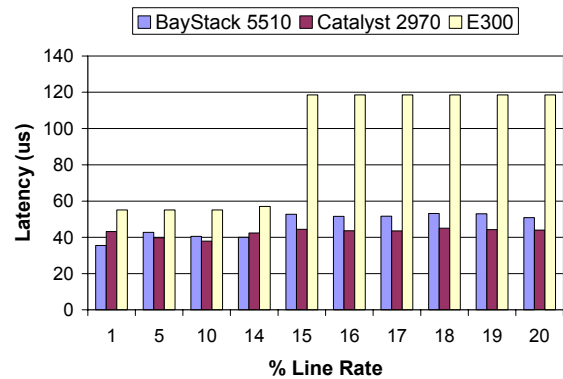


Figure 45: Single platinum stream 1518-byte packet latency.

Figures 46 and 47 show the single platinum stream jitter results above saturation. The small-packet jitter is very low for the BayStack and Catalyst, while more considerable but manageable for the E300. However, the large-packet jitter has all three switches having about the same jitter after saturation and the E300 having more consistent jitter below saturation. The E300 was designed with 10GigE operation in mind. Therefore, it may be optimized for larger packet sizes to get the best use from faster operation because of the effects of the inter-packet gap.

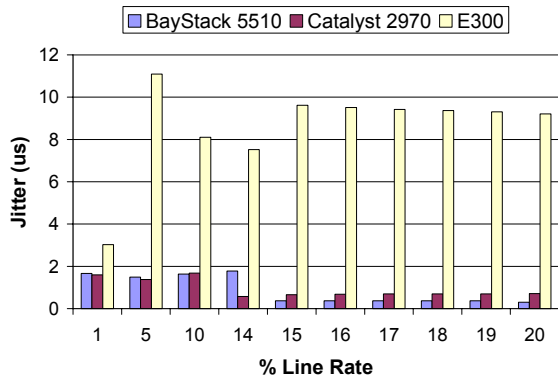


Figure 46: Single platinum stream 128-byte jitter.

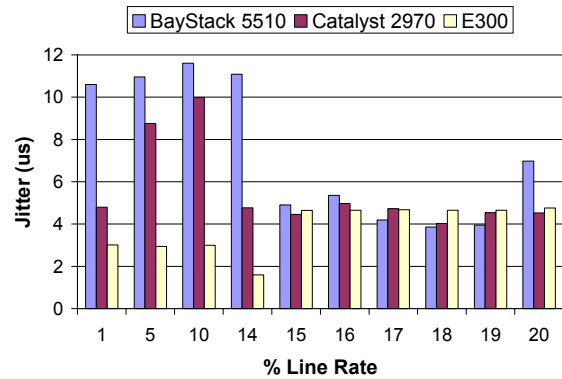


Figure 47: Single platinum stream 1518-byte jitter.

6. Conclusions

Aircraft networks demand both high bandwidth and low latency within bounds. For this reason, switches built for these networks were usually custom designed for each generation of network. Current Gigabit Ethernet switches offer high throughput and low latency in under-provisioned cases. With the use of QoS-enabled switches, these benefits can be extended to cases where the network is slightly over-provisioned. Current Gigabit Ethernet switches which implement QoS offer a low-cost solution for avionics networks. This paper discusses the performance of three such switches: the Nortel Networks BayStack 5510, the Cisco Systems Catalyst 2970, and the Force10 Networks E300.

Each switch has a good amount of QoS control with options for creating a full solution to match desired performance capabilities. Latencies and jitter for all switches decreased dramatically for critical data as long as it is given sufficient priority. However, these switches have not yet reached their potential in terms of matching the versatility of quality of service standards. Although each switch implements DiffServ classification and tagging, none of the three switches examined had the ability to differentiate between 64 separate traffic classes for purposes of queue shaping. Thus, engineers are given fewer options to implement complicated priority schemes.

The BayStack 5510 from Nortel Networks featured very impressive latency and jitter performance, especially for smaller packets. The switch was capable of assuring jitter for critical data below 2 microseconds for small packet sizes. Further, the BayStack configuration utility was easy to use while providing a good amount of QoS control to create more powerful QoS solutions. The Nortel Networks had eight egress queues which allows for a variety of traffic profiles.

The Catalyst 2970 from Cisco Systems also featured low latency and jitter for small packets. For large packet sizes, QoS controls decrease large-packet jitter by at least 10 microseconds compared to the best-effort case. The QoS configuration of this switch allows the user to have control over all aspects of the QoS policies. The ability to shape at the ingress queue is a significant difference, which would possibly help packets that have already been classified by a previous switch. This switch also implements WRED and other congestion avoidance techniques that are useful for protecting critical data.

The E300 from Force10 Networks also featured quite low jitter, but its latency was much higher than the other two switches. For low traffic loads, the E300 has jitter of approximately 3 microseconds for large packet sizes. The E300 is intended as a core switch for high-end applications. The large amount of QoS control however means that the E300 can implement the most diverse set of policies of any of the switches analyzed.

In the future, simulation models will be used to extend the gathered data. The simulations will investigate the use of new QoS services that are beyond the capabilities of the current experimental testbed. Various network, traffic, and load scenarios will be analyzed for the QoS mechanisms under study. The simulations will provide probability distributions of arrival latencies for different QoS algorithms. The models will be verified against the experimental data that has already been gathered. This data will provide a comprehensive analysis of the relation between current and emerging QoS services and switch technologies in terms of statistically bounded latencies.

7. Acknowledgements

This research was sponsored by Rockwell Collins, Inc. in Cedar Rapids, Iowa. Further support was received in terms of equipment and technical support from Cisco Systems, Force10 Networks, Ixia, and Nortel Networks. We would also like to thank all of the members of the HCS research lab at the University of Florida for providing support and insight during our study.

8. References

- [1] J. Meier, S. Kim, A. George, and S. Oral, "Gigabit COTS Ethernet Switch Evaluation for Avionics," *Proc. of the 27th Annual IEEE Conference on Local Computer Networks*, Nov. 2002, pp. 739-740.
- [2] V. Firoiu, J. Y. Le Boudec, D. Towsley, and Z. L. Zhang, "Theories and Models for Internet Quality of Service," *Proc. of the IEEE*, Vol. 90, No. 9, Sept. 2002, pp. 1565-1591.
- [3] R. Braden, D. Clark, and S. Shenker, "Integrated Services in the Internet Architecture: an Overview," *Internet Engineering Task Force (IETF) specification, RFC 1633*, Jun. 1994.
- [4] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," *Internet Engineering Task Force (IETF) specification, RFC 2475*, Dec. 1998.
- [5] T. Bauge, R. Egan, P. Flegkas, P. Trimintzios, I. Andrikopoulos, and G. Pavlou, "QoS Provisioning in an IP-based Studio Production Network." *Proc. of PGNet 2000, 1st Annual Post-Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting*, Liverpool, UK, Jun. 2000, pp. 81-88.
- [6] Ixia, Ixia Hardware Test Platforms Specifications, 2004, http://www.ixiacom.com/datasheets/pdfs/ch_1600t_400t_100.pdf.
- [7] Nortel Networks, BayStack 5510 Product Brief, 2004, http://www.nortelnetworks.com/products/02/bstk/switches/baystack_5510/collateral/nn104742-012804.pdf.
- [8] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," *Internet Engineering Task Force (IETF) specification, RFC 3031*, Jan. 2001.
- [9] IEEE Standard 802.1D, "IEEE Standard for Information Technology—Telecommunications and Information Exchange Between Systems—Local and Metropolitan Area Networks—Common Specifications. Part 3: Media Access Control (MAC) Bridges," *IEEE*, 1998.
- [10] A. Koubaa, A. Jarraya, and Y. Q. Song, "SBM Protocol for Providing Real-time QoS in Ethernet LANs," *Proc. of the 1st Intl. Workshop on Real-Time LANs in the Internet Age RTLLA'02*, Austria, Jun. 2002, pp. 45-49.
- [11] V. Laatu, J. Harju, and P. Loula, "Evaluating Performance among Different TCP flows in a Differentiated Services Enabled Network." *Proc. of the ICT'2003 International Conference on Telecommunications*, Papeete, Tahiti, French Polynesia, Feb. 2003, pp. 701-715.
- [12] C. Bouras and A. Sevasti, "Analytical Approach and Verification of a DiffServ-based Priority Service," *Proc. of High Speed Networks and Multimedia Communications HSNMC '03*, Portugal, Jul. 2003, pp. 11-20.
- [13] V. Firoiu, I. Yeom, and X. Zhang, "A Framework for Practical Performance Evaluation and Traffic Engineering in IP Networks," *Proc. of IEEE International Conf. on Telecommunications*, Jun. 2001.
- [14] T. Chahed, "IP QoS Parameters." TF-NGN, Nov. 2000, http://axgarr.dir.garr.it/~cmp/tf-ngn/QoS_parameters.ps.
- [15] V. Paxson, G. Almes, J. Mahdavi and M. Mathis, "Framework for IP Performance Metrics," *Internet Engineering Task Force (IETF) specification, RFC 2330*, May 1998.
- [16] Force10 Networks, Gigabit Ethernet Performance Evaluation, 2002, <http://www.force10networks.com/products/pdf/336GE-ports-v3.1.pdf>.