Advanced QoS White Paper

Executive Summary
An Ethernet service that is becoming more attractive to end-users is the Virtual Private LAN Service (VPLS) or Provider Provisioned Virtual Private Network (PPVPN). VPLS enables Service Providers (SPs) to deliver Virtual Private Network (VPN) services based on cost-effective Ethernet, with the same level of support and reliability as existing services that use Frame Relay and ATM.

The advanced QoS and shaping features on some newer Ethernet switches enable these devices to be positioned as access devices into Virtual Private LAN networks and for SPs to offer the Service Level Agreements (SLAs) that customers rely on for end-to-end data delivery. This advanced QoS white paper describes the four key new QoS features available in a range of Allied Telesis Ethernet switches and routers—Bandwidth Metering, RED Curves, Mixed Scheduling and Virtual Bandwidth. These features enable SPs to provide the same SLAs for their customers with Ethernet as they have been able to do in the past with Frame Relay and ATM services, but at much less cost.

This white paper contains the following sections:

A. Introduction
An introduction to Service Providers’ needs and what they want to achieve with QoS.

B. What are the new features available with advanced QoS?
Discusses the advanced QoS features that are now available.

C. How do Allied Telesis products support advanced QoS?
Outlines how Allied Telesis products support these advanced QoS features.
Contents

A. INTRODUCTION 3
   What do service providers want to achieve? 4
   What are the benefits of bandwidth profiles? 4
B. WHAT NEW FEATURES ARE AVAILABLE WITH ADVANCED QOS? 5
   Metering 5
      Packet Colour 6
      Two-Rate, Three-Colour Metering 6
      Bandwidth Profile Parameters 7
      Token Bucket Analogy 8
      Single Rate Three Colour Metering 9
      Colour Blind and Colour Aware Rate Metering 10
      How do downstream devices know which packets have which colour? 11
   Policing or Shaping 11
   Random Early Detection/Discard (RED) 12
   Mixed Scheduling 13
   Virtual Bandwidth 14
   Dynamic Application Recognition (DAR) 14
C. HOW DO ALLIED TELESIS PRODUCTS SUPPORT ADVANCED QOS? 15
   Allied Telesis x900 Series Switches 15
   Allied Telesis Routers 15
   The Metro Ethernet Forum 16
A. Introduction

Introduction

Customer demands for cheaper and more advanced technologies mean that Telecommunications Service Providers (SPs) need to constantly evaluate new and more efficient methods of delivering services to their customers. In order for SPs to provide services that generate the most revenue with the lowest operational costs their transport networks need an optimal design.

Today, more than 95% of all applications begin and end on Ethernet¹, so keeping packets as Ethernet and transporting them over Ethernet reduces costs by eliminating protocol conversions and other inter-networking challenges. In addition, because Ethernet services can be delivered over various transport networks, SPs can use their existing network infrastructure to offer more revenue generating services while they migrate to a packet-based infrastructure.

One Ethernet service that is becoming more attractive to end-users is the Virtual Private LAN Service (VPLS) or Provider Provisioned Virtual Private Network (PPVPN). VPLS enables SPs to deliver Virtual Private Network (VPN) services that are based on cost-effective Ethernet and have the same level of support and reliability as existing services that use Frame Relay and ATM. VPNs enable SPs to securely connect multiple customer sites together over a common shared network, ensuring that individual customer data remains separate.

A prime driver for the development of PPVPNs is the provision of seamless Ethernet WANs for business customers. VPLS networks also enable SPs to improve their services for residential users. As the revenue that can be made from telephone calls rapidly decreases, telecommunications companies need to be able to move into more data-intensive residential services, and seriously compete for customers against satellite and cable service providers.

For network managers, all Virtual Private LAN traffic over the WAN appears to be logically on the same LAN, which greatly simplifies management and reduces scaling issues. It is also possible to spread VLANs and IP subnets across several sites as though they were segments of a single LAN. In addition, because the VPLS customer/provider interface is Ethernet, not native traffic, it creates a “seamless” private/public interface.

One of the reasons why Ethernet has not been used to provide these services in the past is that it does not inherently have the traffic shaping and QoS features that Frame Relay and traditional ATM WAN services have. These QoS and traffic shaping features enable SPs to offer Service Level Agreements that customers rely on to ensure end-to-end data delivery.

The advanced QoS and shaping features on some newer Ethernet switches enable these devices to be positioned as access devices into Virtual Private LAN networks to meet customer demands. This advanced QoS white paper describes the four key new QoS features that are available in a range of Allied Telesis Ethernet switches and routers—Bandwidth Metering, RED Curves, Mixed Scheduling and Virtual Bandwidth. These features enable SPs to provide the same SLAs for their customers with Ethernet as they have been able to do in the past with Frame Relay and ATM services, but at much less cost.

¹ Source: http://www.americasnetwork.com/americasnetwork/article/articleDetail.jsp?id=115749
What do service providers want to achieve?
The two main drivers for VPLS networks are to provide seamless Ethernet WANs for business customers and a multi-service delivery mechanism for residential customers. The prioritization and management of latency and jitter is important for both of these services but even more so for a multi-service residential offering. To be able to compete with current voice and television services, an Ethernet-based multi-service system must be able to achieve a very high quality management of video and voice traffic flows.

In order to provide these services, a SP’s Ethernet switching equipment must be capable of accurately shaping traffic to conform to set bandwidth limits, so they can then offer specific bandwidth profiles. A bandwidth profile outlines the service guarantees that the SP will provide by defining the traffic types and amounts of each traffic type that subscribers can send into the SP’s network. In order for the SP to meet the delivery guarantee, the Ethernet switching equipment must be able to give relative priorities to different traffic types, and manage the latency and jitter of particular traffic streams.

What are the benefits of bandwidth profiles?
Bandwidth profiles enable SPs to:

- Engineer their network resources to provide accurate performance assurances for in-profile packets.
- Offer bandwidth to subscribers in very fine increments, enabling subscribers to purchase the bandwidth they need and SPs to price services in a more fine-grained fashion than is possible with Time Division Multiplexing based services.
- Offer multiple services per interface with each service having its own bandwidth profile. This enables SPs to achieve higher profit margins with lower operational overheads while providing subscribers with more cost effective services that accurately meet their needs.
B. What New Features are Available with Advanced QoS?

Earlier implementations of QoS in Ethernet switches have been relatively simple\(^2\). However, as SPs move into new markets, they require more advanced QoS functionality. Figure 1 illustrates the basic functionality available in a simple QoS solution compared to the extra functionality available in an advanced QoS solution.

This section provides a brief overview of these advanced features and explains how they enable SPs to meet the needs of new market segments.

<table>
<thead>
<tr>
<th>Simple QoS Features</th>
<th>Advanced QoS Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification</td>
<td>Classification</td>
</tr>
<tr>
<td>Policing or Shaping</td>
<td>Metering</td>
</tr>
<tr>
<td>Remarking</td>
<td>Policing or Shaping</td>
</tr>
<tr>
<td>Queuing</td>
<td>Remarking</td>
</tr>
</tbody>
</table>

**Figure 1: Simple and Advanced QoS Features**

**Metering**

Bandwidth metering requires a bandwidth profile that specifies the average rate of ‘committed’ and ‘excess’ Ethernet packets allowed into the SP’s network at the switch port. Packets that are transmitted up to the ‘committed’ rate are allowed into the provider’s network and delivered per the service performance objectives specified in the Service Level Agreement (SLA) or

\(^2\) Allied Telesis’ QoS White Paper published in January 2007 introduces the basic QoS features and functionality.
Service Level Specification (SLS). These packets are ‘in-profile’ or ‘conformant’ with the bandwidth profile.

Packets sent above the ‘committed’ rate and below the ‘excess’ rate are allowed into the provider’s network but are delivered without any service performance objectives. These packets are ‘out-of-profile’ or ‘non-conformant’ to the bandwidth profile. Packets sent above the ‘excess’ rate are discarded. There are two main methods of measuring the bandwidth profile—twin-rate, three colour metering and single-rate, three colour metering. Both of these metering methods require packets to be coloured depending on their conformance with the SLA.

Packet Colour

The colours green, yellow and red have become the common terminology used to indicate a packet’s level of conformance with a bandwidth profile:

• Green packets
  If the packets conform to the committed rate of the bandwidth profile, they are marked green and delivered in accordance with the service performance objectives specified in the SLA.

• Yellow packets
  If the packets are over of the committed information rate and below the excess rate of the bandwidth profile, they are marked yellow.

• Red packets
  If the packets do not conform to either the committed or the excess rates of the bandwidth profile, they are marked red and are usually discarded immediately.

Two-Rate, Three-Colour Metering

The two-rate, three colour metering process, also known as twin-rate metering, is commonly implemented through a token bucket algorithm. Four parameters are used to determine the bandwidth profiles for twin-rate metering: two bandwidth limits—the Committed Information Rate (CIR) and the Excess Information Rate (EIR)—and two burst sizes—the Committed Burst Size (CBS) and the Excess Burst Size (EBS).
Bandwidth Profile Parameters

The Committed Information Rate (CIR) is the average rate up to which packets are marked green. These packets are referred to as CIR-conformant.

The Excess Information Rate (EIR) specifies the average rate up to which packets are admitted to the SP’s network. The EIR is greater than or equal to the CIR. Packets that exceed the CIR, but are below the EIR are marked yellow. Because these packets do not conform to the CIR, the SP does not provide any guarantees with regard to their delivery. Packets that exceed the EIR, do not conform and are marked red, and are typically discarded.

Because traffic levels can fluctuate, the twin rate, three colour metering process enables the traffic to burst above the CIR and EIR a certain amount before marking the packets yellow and red, respectively.

The Committed Burst Size (CBS) is the maximum number of bytes allowed for incoming packets to burst above the CIR, but still be marked green.

The Excess Burst Size (EBS) is the maximum number of bytes allowed for incoming packets to burst above the EIR and still be marked yellow. When the burst size has been exceeded, packets above the EIR are marked red.

Figure 2: Two-Rate Three Colour Metering Graph
Token Bucket Analogy

A ‘token bucket’ analogy is used to describe the algorithm that performs the metering. The algorithm decides which particular packets are within the bandwidth limits, and which are in excess of the limit.

The token bucket analogy uses two buckets, one with a volume equal to the CBS—the C bucket—and one with volume equal to the EBS—the E bucket. Tokens are dropped into the buckets at rates equal to the CIR and EIR, respectively. Simultaneously, every time a packet comes past, a set of tokens equal to the size of the packet are taken out of the buckets. As long as the C bucket is not empty, packets are marked green. When the C bucket is empty, but the E bucket is not, packets are marked yellow. At moments when both buckets are empty, packets are marked red.

![Figure 3: Token Bucket Analogy - Two-Rate, Three-Colour Metering](image-url)
Single Rate Three Colour Metering

Another commonly used metering method is single-rate three-colour metering\(^5\), also known as single-rate metering. In this process, there is only one bandwidth limit, but two burst sizes.

![Figure 4: Single-Rate Three-Colour Metering Graph](image)

The algorithm used in this process can also be described using a token bucket analogy. Single-rate, three-colour metering also uses two buckets with volumes equal to the CBS and EBS. The first bucket is filled with tokens at a rate equal to the CIR, and any overflow pours into the second bucket.

Tokens are removed from the buckets when packets come along. Packets are marked green if the first bucket is not empty, yellow if the first bucket is empty, but the second isn’t, and red when both buckets are empty.

---

\(^5\) Defined in RFC 2697.
Colour Blind and Colour Aware Rate Metering

The metering processes described above are referred to as ‘colour blind’ metering. That is, packets are considered to have had no conformance judgement made on them before they arrive at the meter. Effectively, all packets are considered to be green when they enter the metering process and are only marked yellow or red if the traffic class exceeds the bandwidth limits on that device.

It is also possible to have a metering process work in colour aware mode where the packets
have already been allocated a colour by an upstream device before entering the metering process. So when a red packet enters the meter; it remains red. If a yellow one enters the meter; it can be marked red if the max bandwidth and max burst size have been exceeded, otherwise it remains yellow. In colour-aware mode the device should not remark non-conformant or semi-conformant packets as conformant.

How do downstream devices know which packets have which colour?

As mentioned above, SPs guarantee to deliver green packets, but make no promises about the delivery of yellow packets. In order for the SPs to meet the SLA, their devices need to know which packets are marked with which colour. The re-marking process achieves this. After the metering has been performed, the 802.1p and/or DSCP fields within the packets are set to specific values to indicate the result of the metering. The service provider defines the particular 802.1p/DSCP values used to indicate different packet colours6.

Policing or Shaping

This discussion of the metering process has indicated that packets marked red are typically discarded. However, the immediate discarding of red-marked packets is a choice known as policing. An alternative method of dealing with non-conformant packets is known as shaping.

Before explaining how shaping is carried out, it is necessary to review the mechanism where packets are queued at an egress port. Each egress port has a set of egress queues, which are allocated different priorities or weights7. The QoS mechanisms place packets into the appropriate egress queue, but the queues are of a limited length, so packets cannot be placed into them indefinitely. If the switch is congested, the queues may fill up and no more packets can be added, so even high priority packets can be dropped from the end of queues.

The shaping process uses rules to decide which packets are allowed to enter the egress queues instead of simply dropping all the red packets. The rules are different for each packet colour:

- **Green packets** are subject to quite lenient rules, and most of these packets will be accepted into the egress queue.
- **Yellow packets** are subject to more stringent rules, so less of them will get into the egress queues.
- **Red packets** are subject to very stringent rules, especially when the interface becomes congested.

In this way, if there are multiple traffic classes passing through the device, each with different bandwidth limits, it is possible for an over-limit traffic class to make use of bandwidth made available by another traffic flow that is well below its bandwidth limit. But, if all traffic flows are at or above their limit, then the shaping process will make sure the flows do not encroach on each other’s allocated bandwidth.

The most common method used to achieve this selective admission of packets into the egress queues is called Random Early Detection/Discard (RED).

---


7 For more information, see the ‘Mixed Scheduling’ section of this white paper.
Random Early Detection/Discard (RED)

When congestion occurs, RED curves enable packets to be dropped before the egress queue exceeds its allocated maximum length. RED curves start dropping the less conformant packets, and then progressively more conformant packets until the congestion is eased.

A single RED curve consists of a set of ‘start’ ‘stop’ ‘drop’ values. These three values define a “curve” as illustrated in Figure 6.

![Figure 6: A Single RED Curve](image)

**Start** defines the length that the queue must reach before the packets start being dropped.

**Stop** defines the length that the queue must reach before the shaper stops dropping randomly, and just drops all further packets.

**Drop** defines the percentage of packets that are being dropped at the point when the length of the queue reaches the stop value. So, effectively, drop defines how quickly the rate of dropping packet must increase as the queue length grows from the start value to the stop value.

These fundamental curves are collected into RED curve groups. A group is a collection of three curves, one for each of the three possible packet colours. In this way, red packets start being dropped when only a small amount of data has been backed up in the egress queues, yellow packets start getting dropped when the queues are backed up a bit more, and the green packets start to be dropped when the congestion is quite severe.
Mixed Scheduling

The net effect of all the marking, metering and RED curve activity is to decide which packets make it into which queue. From this point, the scheduling process determines the order that the packets are picked out of different queues and transmitted. There are a number of different scheduling algorithms and the choice of algorithm determines how the packets are transmitted.

The scheduling algorithms fall into two main categories - Priority Scheduling and Round Robin Scheduling.

With priority scheduling the queues are assigned a set of priorities and packets are always sent from the highest-priority queue first with very little delay. Whenever there are packets in the highest-priority queue, they are transmitted; they do not have to wait for lower priority queues to be processed. Consequently, if there is so much traffic coming into the higher priority queue that it always has packets to send, then the queues below it will never get a chance to send any packets.

With round robin scheduling the queues are given turns at sending packets. There are a lot of different algorithms within the round robin scheduling category, including:

- simple round robin
- weighted round robin
- deficit weighted round robin
- self-clocked fair queuing

Round robin scheduling gives every queue a chance to forward packets, so no queue is ever totally starved. However, round robin scheduling can be problematic because the packets belonging to delay-sensitive flows can be held up as they wait for packets in other queues to have their turn at being transmitted.

The best solution is to use a mixture of priority and round robin scheduling, which is referred
to as mixed scheduling. With mixed scheduling, the traffic queues are divided into two sets—one set of queues operates according to priority scheduling, and one set operates according to round robin scheduling. The priority queues take precedence over the round robin queues. So the round robin queues can only send packets when all of the priority queues are empty.

The way to make use of a mixed scheduling system is to assign low-bandwidth delay-sensitive traffic flows to the priority queues and higher bandwidth delay-tolerant flows to the round-robin queues. This ensures that the delay-sensitive traffic will avoid delays, as the priority queues get serviced in preference to the round robin queues. However, provided that the traffic flows assigned to the priority queues are low bandwidth, there will not be any problem with traffic flows being starved. Even if a traffic flow assigned to one of the round-robin queues is of high-bandwidth, it cannot starve the other round-robin queues.

Mixed scheduling is an effective compromise that combines the advantages of the two scheduling types and avoids their pitfalls, provided the assignment of traffic flows to queues is done properly.

**Virtual Bandwidth**

If more than one traffic class is sending packets to one egress queue and the total bandwidth allowed from all of these traffic classes needs to be limited, a bandwidth limit can be assigned to the common egress queue. This bandwidth limit is known as applying a virtual bandwidth to the egress queue.

Virtual bandwidth limits the rate at which the egress queue can emit packets, which can also cause packets to accumulate in the queue. When packets accumulate in the queue, the RED Curves come into play and make decisions about which packets to drop.

Virtual bandwidth provides flexibility in the way that traffic flows use their available bandwidth. Virtual bandwidth enables users to prevent some traffic flows from starving others, and if some of the traffic flows are quiet, then others are able to use a bigger slice of the virtual bandwidth and send more of their non-conformant packets.

The three processes of metering, RED Curves and virtual bandwidth can combine to provide much more flexibility than a simple QoS solution that would only configure a traffic class to drop red packets.

**Dynamic Application Recognition (DAR)**

Because it is difficult to predict which UDP ports will be used by voice or video sessions and require advanced QoS, Dynamic Application Recognition (DAR) is used to snoop for session setup exchanges and dynamically create classifiers that match the voice and video packets in the session.

In the DAR process, the device examines the voice or video session initiation messages arriving at the interface and compares them against a preconfigured DAR object. The DAR object tells the device what kind of session to match on that interface. The device then creates a dynamic classifier to match the session and uses the dynamic classifier to sort voice and video packets into traffic classes and apply the configured QoS processing.

---

8 DAR is available on Allied Telesis products that support advanced QoS in their software.
C. How do Allied Telesis Products Support Advanced QoS?

A range of Allied Telesis switches and routers support the advanced QoS functionality described in this paper. Our advanced routers and Layer 3+ switches reach new and unmatched heights in performance, flexibility, and reliability. Packaged in a one-rack unit standard rack mount chassis, the following switches incorporate these advanced QoS features in their leading edge silicon:

**Allied Telesis x900 Series Switches**

- **x900-48FE**
  4 x 1000BASE-X SFP uplinks, 48 x 10/100BASE-T (Rj-45) copper ports

- **AT-9924T**
  24 x 10/100/1000BASE-T (Rj-45) copper ports and 4 x 1000BASE-X SFP combo ports

- **AT-9924SP**
  24 x 1000BASE-X SFP ports

- **x900-24XT & x900-24XT-N**
  2 x 30 Gbps expansion bays + 24 x 10/100/1000BASE-T (Rj-45) copper ports

- **x900-24XS**
  2 x 30 Gbps expansion bays + 24 x 1000BASE-X SFP ports

**Allied Telesis Routers**

These advanced QoS features are also implemented through our superior AlliedWare® Software on the following routers:

- **AR415S/AR410S**
  4 x 10/100 BASE-T ports, 1 x asynchronous port, 1 x 10/100 Ethernet port, 1 x Port Interface Card (PIC) bay, and an integrated security engine

- **AR440S/AT-AR441S9**
  1 x ADSL port, 5 x 10/100BASE-T ports, 1 x PIC slot (Annex A ADSL router)

- **AR450S**
  1 x 10/100BASE-TX WAN port, 1 x 10/100B-TX DMZ port, 5 x 10/100BASE-TX LAN ports, 2 x Asynchronous RS232 ports

- **AR725 & AR745**
  Modular Enterprise Routers with 2 x 10/100BASE-T ports

- **AR750S**
  5 x LAN 10/100BASE-T ports, 2x WAN 10/100BASE-T ports, 2x PIC slots, 1 x Asynchronous console/modem port

- **AR770S**
  2 x WAN combo ports (SFP or 10/100/1000TX), 4 x LAN 10/100/1000TX ports, 2 x PIC slots, 1 x Asynchronous console / Modem port

For more information about our products, contact your local Allied Telesis representative or visit our website: www.alliedtelesis.com

---

9 AR441S is an Annex B ADSL router
Advanced QoS White Paper

The Metro Ethernet Forum

The Metro Ethernet Forum is a non-profit organisation established to accelerate the adoption of optical Ethernet as the technology of choice in metro networks worldwide. The Metro Ethernet Forum is made up of more than 60 members, including service providers, local exchange carriers, network equipment and test equipment vendors, and other networking companies with an interest in metro Ethernet.

Allied Telesis’ QoS implementation follows the recommendations outlined and recommended by the Metro Ethernet Forum. For more information visit the MEF website: www.metroethernetforum.org