

Quality of Service in the Wireless Backhaul

Executive Summary

The growing demand for mobile data services is both a blessing and a challenge for wireless service providers. End-users, once content with simple voice and text messaging, are growing increasingly savvy. They want the convenience of accessing the Internet and IP-based applications from mobile devices, without sacrificing the Quality of Service (QoS) experienced on wired networks.

With the demand for IP-based services comes a dramatic increase in bandwidth requirements, best served by the transition to an Ethernet backhaul infrastructure. However, the need for the reliability, efficiency and timing functionality that TDM-based networks have historically provided remains. How can network engineers evolve the backhaul network while simultaneously supporting revenue-generating legacy services?

This white paper outlines a methodology to create a solid backhaul network foundation — that addresses both current and future needs — by examining the QoS parameters of the various services offered on wireless networks. By understanding how to prioritize network traffic based on Class of Service (CoS), it is possible to right-size bandwidth allocation and correctly configure network elements for superior end-to-end connectivity that best meets the needs of customers and service providers.

Evolving Wireless Backhaul Networks

Wireless networks revolutionized the way people use telecom services at work, at home and when traveling. As communication technology advances and demand for IP-based services grows, wireless backhaul (traffic from NodeB/BTS to the RNC) must also evolve to meet new demands on networks.

Traditional backhaul networks use TDM circuits, generally T1s. Ethernet has emerged as the technology to replace T1 circuits because it promises significant increases in bandwidth, reductions in costs and superior support for IP-based applications. However, in the near term, backhaul equipment must still support legacy ATM (UMTS networks) and/or TDM (UMTS/CDMA networks). To continue to provide legacy services over Ethernet backhaul, wireless providers employ MPLS Pseudowires to encapsulate ATM/TDM,

using either native-IP or MPLS encapsulation. The use of MPLS EXP bits, IP Diffserv, and/or Ethernet P-Bits enables the implementation of different QoS traffic classes to enforce service priorities.

Understanding QoS Requirements

Engineering a suitable QoS solution is a critical component of wireless backhaul. First, it is important to clearly understand how application protocols work and their associated QoS requirements.

Every application protocol is assigned to a specific Class of Service (CoS) queue. Individual queue configurations support the intended QoS requirements of the associated application. A suitable QoS solution allocates adequate resources at each hop in the packet backhaul network to ensure that each application receives required resources. Some applications have stringent delay guarantees. Others allow a higher range of delay, but still have strict requirements for packet loss.

To fine-tune a wireless network, the delay, jitter, packet loss and bandwidth of each protocol must be understood for every service offered. Constraints for services are similar to those defined in Table 1.

How many QoS classes are defined depends on the number of variations of application protocol QoS requirements the wireless service provider is willing to undertake. The main consideration is the complexity of the resulting design versus real benefits derived from additional QoS classes. A thorough understanding of application protocols usually leads to an optimal (minimal) solution.

The engineer must also understand how to classify each application protocol at various points in the network. Some applications may use fixed TCP/UDP port assignments providing for a simple means of identification. Others have dynamic assignments of port numbers. Some applications may set IP DiffServ or 802.1p values upon origination of the packet. In MPLS backhaul links, EXP bits likely indicate the application's service class.

IP QoS Class	Diffserv Class	EXP Bits	P Bits	Police	Queue Management	Shape	Schedule
Control	CS7	7	7	No	Tail drop, 34ms/hop	No	Strict Priority
Real time	EF	6	6	Green/Red	Tail drop, 2.5ms/hop	Yes	Strict Priority
Priority	AF41	5	5	Green	wred, 5ms/hop@10%p	Yes	wfq 95%
Priority-out-of-contract	AF42	4	4	Yellow	wred, @10%prob		
Best Effort	BE	0	0	No	Red, 50ms/hop@20%p	Yes	wfq 5%

Table 1. QoS requirements.

In any event, each application must be analyzed to identify which mechanisms are used for classification. Since MPLS EXP and Ethernet 802.1p bits allow for only 8 values, the wireless engineer uses other tools if more than 8 classes are defined. These tools might include mapping and creating a different PWE or RSVP-TE LSP per traffic type (RFC 3270).

The example in Table 1 illustrates a wireless provider configuring three QoS traffic classes, or “general flows,” as a minimum.

QoS Classes

In Table 2, classes of QoS for different service types are outlined. Understanding these classes is an important part of creating a backhaul network design that is easy to deploy and maintain, yet encompasses all types of applications for wireless networks.

The *Control* class defines all the control signaling protocols needed to set up router connectivity (OSPF, LDP hellos, ARP, etc.). This class has the highest priority because if the control protocol fails, the routers cease to forward any of the other packets. In this category, wireless backhaul engineers might also include certain router management traffic such as *telnet* and *ssh*.

The *Real Time* class represents applications that allow groups of users to interact with each other by exchanging multimedia content. Examples include TDM Voice, VoIP, voice signaling and multimedia conferencing applications. These applications are intolerant to delay/jitter and impose strict performance requirements on the network. All similar applications that have similar delay and jitter parameters are included in this class. Real Time applications are generally UDP-based. So wireless backhaul engineers need to determine the UDP ports each application utilizes in order to classify the traffic. Or engineers might prefer to configure, and subsequently trust, the BTS/NodeB sites’ IP/DSCP or Ethernet 802.1p settings and map them accordingly. Refer to Table 3 for base QoS requirements.

The *Priority* traffic class represents non-real-time applications (i.e., no strict jitter/latency requirements) that require better service than Best Effort (see below). Examples include applications such as e-commerce, database access, web browsing or transactional services. Key aspects of Priority class applications are minimization of packet-loss and per-hop delay. Typically, this class of application utilizes the TCP protocol.

Finally, the *Best Effort* traffic class combines the *3GPP Streaming and Background* class. Streaming flows are capable of delivering multimedia sessions by satisfying time-dependent parameters through buffering. Examples are audio/video downloads and print server applications, which require guaranteed performance in terms of packet loss and data integrity, but do not have strict delay or jitter constraints. Applications in this class typically have relatively large buffers (10s or 100s of milli-seconds or more) in order to cope with any network-induced delay/jitter. Understanding the applications’ jitter buffer capabilities is critical to ensuring this class receives adequate servicing. Other “robust” applications (*Background*) may also live happily in this class since their flows are very tolerant to higher latency, jitter and packet loss. Most applications in this class

Traffic Class Name	Type of Flow
Control	Protocol signaling
Real time	Voice signaling
	Voice
	VoIP
	Multimedia conferencing
Priority	Transaction services
	Web browsing
	Telnet
Best effort	Print servers
	Video downloads
	High throughput data
	Low latency data
	Low priority data

Table 2. QoS classes for different service types.

IP QoS Class	Max Delay	Jitter	Max Data Rate	Max Packet Loss
Real time	150ms (25ms)	10ms	5Kbps	1%
Priority	500ms (50ms)	20ms	384Kbps	1%
Best Effort	n/a	n/a	384Kbps	3%

Table 3. Base QoS requirements.

are TCP-based and rely on TCP’s robustness to cope with the best-effort nature. Examples of Background include ftp and email. Typically, these applications are serviced only when bandwidth is available for them. A well-designed network assures adequate capacity for this QoS class which is likely to consume the greatest bandwidth.

Table 3 shows an example of the QoS characteristics that the applications belonging to each group should meet. Real Time traffic allows for maximum delay of 150ms end-to-end, of which only 25ms is allowed due to the backhaul network routers (the rest being propagation delay due to the transmission distances and other devices). The Real Time jitter indicates no more than 10ms is allowable. Priority traffic allows the routers and network to inject a maximum of 500ms and a jitter of 20ms, while Best Effort continues to operate with fewer than 10 seconds of maximum delay.

Table 2 and Table 3 outline what traffic flows belong to what QoS class, what type of services the wireless providers offer to its customers, and finally the maximum delays, jitter, bandwidth and packet loss those services should expect via the backhaul network. A critical goal in end-to-end QoS in the wireless backhaul network is to be able to translate the service classes into specific parameters that the routers can understand. How many router-hops flows travel

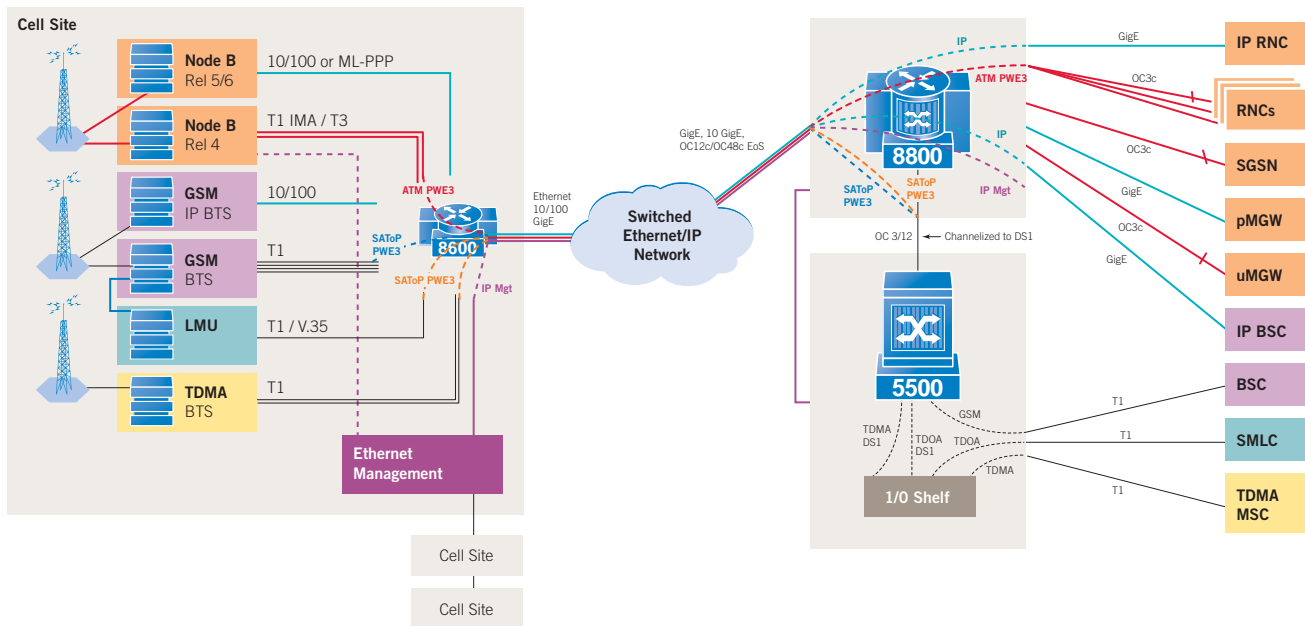


Figure 1. Example wireless backhaul network design.

through in the network must be analyzed in conjunction with how the router configurations meet the requirements set forth in Table 3. Configuring consistent QoS parameters end-to-end and being able to allocate adequate resources on each hop is an absolute requirement with “bursty” traffic over limited bandwidth.

Routing Legacy Services on Ethernet

Figure 1 illustrates how legacy service offerings can reside with emerging services. Utilizing IP/MPLS, ATM SAToP Pseudowires and/or Ethernet Pseudowires, the wireless provider supports both legacy and new services. Alternatively, if there is no need for MPLS, native IP packets can provide for new IP services. For fast fail-over with IP, the wireless backhaul engineers likely implement BFD, OSPF fast hellos or some other form of fast fail-over mechanism.

When deploying the wireless services corresponding to those in Table 2 via the wireless backhaul network in Figure 1, the QoS parameters defined in each router-hop provide the mechanism for classifying, marking, policing, shaping, scheduling and other queue management techniques. The actual QoS values for each QoS class depend on the network topology, the speed of the network links and the amount of traffic carried on each of these links.

Applying QoS Classes

When packets arrive in the first backhaul router, the first QoS treatment on the traffic must be *classification*. Classification “marks” packets based on various fields in the packet or by which port the packet arrived. Classification also may “re-mark” packets based on existing markings. Queuing mechanisms use the markings to place the packet into the defined service classes.

A *Diffserv-like* approach is the recommended method to classify MPLS or IP traffic. Diffserv enables aggregation of the service classes for all services into a manageable set of groups, based on the similarity of the applications’ QoS requirements. Applications with similar requirements are placed within the same class.

In the earlier example, all types of applications flows are classified into one of three QoS classes: Real-time, Priority or Best Effort. The Priority class aggregate distinguishes itself by an additional drop probability and forwarding behavior. All Diffserv markings are kept constant as the traffic transverses the different devices in the wireless backhaul network. Traffic carried via MPLS will require mapping of the Diffserv code points into the MPLS EXP bits.

The wireless backhaul design in Figure 1 aggregates all 10/100Mbps Ethernet links coming from BTS/NodeB sites into a QoS-capable Ethernet switch. In order to enforce QoS at the Ethernet switch, mapping Diffserv/EXP code points into the Ethernet 802.1p bits is required. If the wireless provider opts to connect all BTS/NodeB sites directly into the RNC router (via TDM circuits), mapping to the Ethernet 802.1p bits is not required.

The second QoS treatment is to police the traffic. Real-time traffic is treated as Expedited Forwarding (EF) because the applications included in this QoS class are intolerant to delay and jitter. The Priority class uses a policer configured with a two-rate Three Color Marker (trTCM; RFC 2697), which marks the packets:

1. *Green* if the rate is below Committed Information Rate (CIR); these packets pass uninhibited
2. *Yellow* if the rate is below the Peak Information Rate (PIR) but greater than the CIR, remarking those packets with a higher drop probability within the same class
3. *Red* if the rate exceeds the PIR, in which case those packets are remarked with a lower QoS class, i.e., Best Effort.

The last QoS class, Best Effort, is not policed. The BTS/NodeB sites must be engineered to not send more traffic than what the links can handle. In case of aggregation of Best Effort traffic on the RNC GE link (refer to Figure 1), each BTS/NodeB router and “QoS-enabled” Ethernet switch must shape the traffic to the defined SLA.

The third QoS treatment is to queue the packets. Each QoS class is assigned its own queue. The queues have a maximum length configured to represent the maximum delay per hop for that class of traffic (i.e. Real Time has a maximum queue length of <2.5 ms as reflected in Table 1). Each queue can also be configured to tail-drop packets or use WRED/RED. In the case of Real Time, the simpler tail-drop mechanism discards packets if the queue is full. In the Priority traffic class, there are two drop-precedences for the traffic (AF41 and AF42). WRED is used to discard the green and yellow packets each with a maximum probability of 10% as the packets reaches the maximum length for the queue depicted in Table 1. Best Effort traffic is marked Red since only one drop-precedence exists within the queue.

The fourth QoS treatment is to shape each of the QoS classes in order to limit the output rate of each towards the egress port. Shaping can also occur to set limits on the aggregation of services in a subinterface or physical port/channel. Priority and Best Effort traffic are shaped according to the SLAs purchased by the customer. Shaping Real-Time traffic is not recommended because it injects delay. Instead, an egress policer is used if there is need to set a cap on egress on this type of traffic.

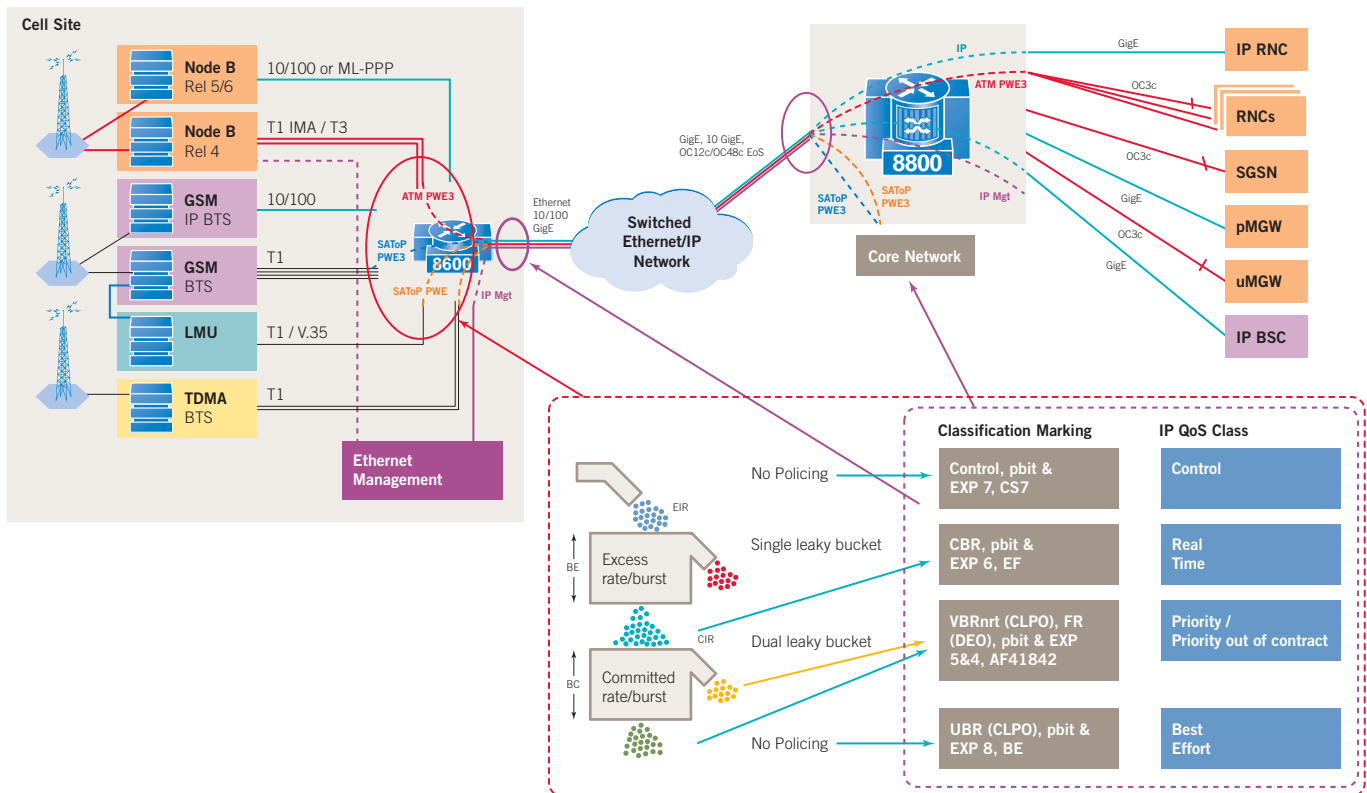


Figure 2. Ingress policing, classification and marking controls.

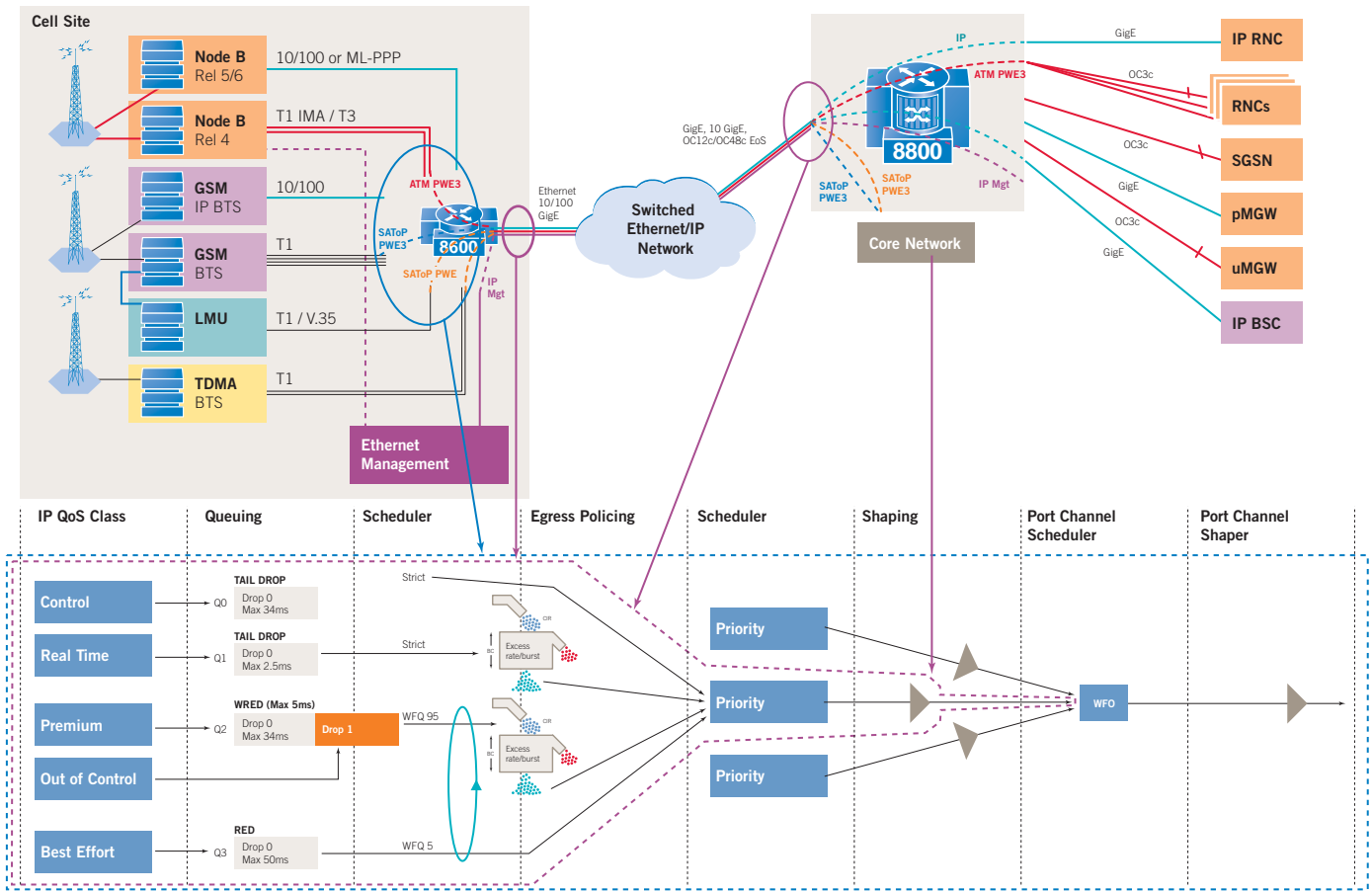


Figure 3. Sample configuration to achieve specific QoS class objectives.

The final QoS treatment is to schedule the traffic. Traffic is scheduled as strict priority for Real-Time or Weighted Fair Queuing (WFQ) for Priority and Best Effort traffic. Under congestion conditions, all Real Time traffic has first priority and is always serviced as long as there are packets waiting in the queue. Priority and Best Effort only receive service if there is not any Real Time traffic and are serviced in proportion to the weight given to the WFQ algorithm. In the example shown in Table 1, 9.5 packets in the Priority class are serviced before one packet of Best Effort is served. Furthermore, WFQ can be applied to prioritize aggregation of traffic from different customers sharing the same physical port/channel. A more desirable weight is given to higher priority customers.

When classifying traffic, it is important to note that Priority traffic should be TCP-based. A UDP-based packet, if remarked due to policing, could potentially be reordered within the stream, which may cause issues with the application.

Applying QoS Classifications

Various ingress and egress QoS treatments are described in this white paper. Figures 2 and 3 depict the different points of entry and exit through the wireless backhaul network.

Figure 2 depicts the configuration ingress policing, classification and marking controls as traffic enters the network.

Scheduling and Buffer Management, with the help of Egress Policing and Shaping, are the tools that provide the desired forwarding behavior for packets of the same class.

Figure 3 is an example of the configuration of Scheduling and Buffer Management and Egress Policing and Shaping to achieve the specific QoS objectives of each class.

The scheduler has two levels of hierarchy: physical port and logical port. A WFQ scheduler is implemented on the logical interface and on the physical port. On the logical interface each QoS class is weighted according to its priority. On the physical port, each logical

customer gets a piece of the link capacity, according to its scheduler weight. This is important in Ethernet networks as one customer of a wireless service provider may have contracted higher SLA treatments than another customer.

An egress policer is also defined to limit the amount of real-time and priority traffic. Since IP services are based on a one-to-many topology, a customer could potentially void an SLA if traffic is not policed at egress.

Shaping can also be deployed per customer interface or per whole port. If between-customers priority scheduling is offered versus WFQ, shaping would limit one particular customer from overtaking the whole Ethernet link. Furthermore, to avoid congestion on the physical link a shaper could be implemented.

Summary

Emerging wireless IP applications can be classified into four classes: Control, Real Time, Priority and Best Effort. Both Real Time and Priority applications require low loss rates to achieve assured throughputs. In addition, Real Time applications also require numerical bounds on delay and delay variation. The Best Effort class encompasses applications tolerant to higher latency, jitter and packet loss.

By following the Diffserv convention, packets of real-time applications are marked with DSCP corresponding to EF PHB and are policed at ingress and egress. Packets in Priority applications are marked with DSCPs corresponding to AF PHBs and also policed at ingress and egress.

For this type of traffic, the ingress policer marks the yellow packets as out-of-contract premium traffic and discards them if congestion is present at egress, using WRED. Best Effort traffic is marked with DSCP corresponding to BE and requires a policer. Premium traffic and Best Effort are weighted as well as the aggregation all of the QoS classes competing against others for the same bandwidth.

Both the Tellabs® 8600 Managed Edge System and the Tellabs® 8800 Multiservice Router Series are enabled with the QoS techniques described in this paper. These platforms offer a scalable solution from 100s Mbps to over 300 Gbps, meeting Ethernet backhaul needs from the base stations (NodeB/BTS) to hub sites and large MSCs.

The ultimate goal of this white paper is to provide a methodology for deploying QoS service across the wireless backhaul network to support both legacy TDM and new IP-based applications.

References

3GPP TS 23.207 V5.8.0, "Universal Mobile Telecommunication System (UMTS); End-to-End QoS Concept and Architecture," December 2000, <http://www.3gpp.org>

3GPP TS 23.107 V5.5.0, "Universal Mobile Telecommunication System (UMTS); QoS Concept and Architecture", March 2001, <http://www.3gpp.org>

F. Le Faucheur, et al, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services." RFC 3270, May 2002

J. Heinanen, et al, "A Single Rate Three Color Marker." RFC 2697, Sept 1999

"Quality of Service in the Tellabs® 8600 Managed Edge System," Tellabs White Paper

"Tellabs 8800 MSR Quality of Service: Architecture and Services," Tellabs Technical Note

North America

Tellabs
One Tellabs Center
1415 West Diehl Road
Naperville, IL 60563
U.S.A.
+1 630 798 8800
Fax: +1 630 798 2000

Asia Pacific

Tellabs
3 Anson Road
#14-01 Springleaf Tower
Singapore 079909
Republic of Singapore
+65 6215 6411
Fax: +65 6215 6422

Europe, Middle East & Africa

Tellabs
Abbey Place
24-28 Easton Street
High Wycombe, Bucks
HP11 1NT
United Kingdom
+44 871 574 7000
Fax: +44 871 574 7151

Latin America & Caribbean

Tellabs
Rua James Joule No. 92
EDIFÍCIO PLAZA I
São Paulo – SP
04576-080
Brasil
+55 11 3572 6200
Fax: +55 11 3572 6225