BCM963xx DSL Router Packet QoS Application Notes

Nov 4, 2005

Table of Contents

REVISION HISTORY

Version	*Date*	*Change Description*
1.0	Nov 4, 2005	V3.04L.01 formal release.

Introduction

The first time IP quality of service feature was introduced to the BCM963xx Linux DSL router platforms in DSL Router Linux release version 2.14L.01, which supports only IP layer QoS over PVCs in one of the router mode WAN protocols (PPPoE, PPPoA, MER, IPoA) but not in bridge mode. Version 2.14L.02 added QoS for PVCs in bridge mode and classification/queuing for 802.1p frame received from LAN ports. Since the scope has been expanded to cover both MAC layer and IP layer, the original name of "IP QoS" was changed to "Packet QoS" (versus ATM cell QoS). Version 3.02L.01 added several more improvements to packet QoS, including IP precedence marking, 802.1p priority marking, 802.1q VLAN frame over ATM PVCs and new classification rule based on physical LAN port including port in an external BCM5325E Ethernet switch.

The packet QoS feature provides three services, classification, marking and queuing, to upstream traffic over ATM PVCs that have packet QoS enabled. Each classification class can be assigned by a user defined name, a set of classification rules and a set of actions to a packet if it matches all of the classification rules. There are two sets of classification rules. Set-1 is based on various fields within TCP/UDP/IP headers plus physical LAN port; set-2 is based on MAC layer IEEE 802.1p priority field. The actions contain queuing action and a set of marking actions. The queuing action assigns a packet to one of the ATM TX queue priorities. The marking actions contain marking on IP precedence, IP TOS and priority field within IEEE 802.1q VLAN header.

The classification and marking were originally carried out by iptables in IP layer but later were replaced by ebtables within the bridge module. The 802.1p marking is executed within the rfc2684bridge module. The packet queuing is handled in the ATM driver and the ATM SAR.

There are three places in the WEB UI involving the Packet QoS setup:

1. **"Enable Quality of Service"** button in WAN Configuration WEB page after clicking "WAN" under "Advanced Setup". Selecting "Enable Quality of Service" button will assign three ATM TX queues to this PVC. Three different priorities are assigned to the three queues: high, medium and low. De-selecting this feature will cause the system to configure only one ATM TX queue with low priority. Packet QoS queuing will only be effective if the packet is forwarded to a PVC that has packet QoS enabled.

2. **"Quality of Service Setup"** WEB page after clicking "Quality of Service" under "Advanced Setup". A packet QoS summary table, as shown below, will be displayed and allows user to add or remove a QoS classification class. This is the main place to configure the classification, marking and queuing rules.

	MARK				TRAFFIC CLASSIFICATION RULES							
					SET-1						SET-2	
Class Name	Priority	IP Precedence	IP Type of Service	WAN 802.1P	Lan Port	Protocol	Source Addr./Mask	Source Port	Dest. Addr./Mask	Dest. Port	802.1P	Remove

3. **"Enable virtual ports"** in WEB UI "Port Mapping" page after clicking "Port Mapping" under "Advanced Setup". This is only necessary for BCM9634x platforms which contain the BCM5325E Ethernet switch and if classification is required to be based on specific BCM5325E port. More description can be found in the Classification section NOTES.

4. **"Enable 802.1q"** in the WAN configuration WEB page if IEEE 802.1q VLAN header is to be inserted to the rfc2684 bridged encapsulated MAC frame in upstream direction. In receiving downstream MAC frame, the 802.1q header will be stripped before it is

forwarded to the IP or the bridge module. If the 802.1p marking is configured in the packet Quality of Service, it will only be effective if the packet is forwarded to a PVC that has 802.1q VLAN enabled.

Classification

There are two sets of classification rules. Set-1 is based on different fields within TCP/UDP/IP layer plus physical LAN port; set-2 is based on MAC layer IEEE 802.1p priority field. Only one set can be used for each class. Each class can be assigned by a name to uniquely identify a class among multiple classes.

Set-1 Rules contain the following items:

- Physical LAN port: select one among USB port, Ethernet ports and wireless port
- Protocol: select one among TCP/UDP, TCP, UDP or ICMP protocols
- Source IP address
- Source subnet mask
- UPD/TCP source port or a range of ports
- Destination IP address
- Destination subnet mask
- UPD/TCP destination port or a range of ports

NOTES - Virtual interfaces corresponding to BCM5325E ports

The BCM9634xGW platforms contain a BCM5325E Ethernet switch. Selecting the “Enable virtual ports” button in WEB UI “Port Mapping” page will create three virtual interfaces within the Linux system. Each virtual interface represents a physical Ethernet port within the external BCM5325E Ethernet Switch. The WEB UI will display four Ethernet ports: ENET1, ENET2, ENET3, and ENET4. ENET1, ENET2, and ENET3 represent Ethernet port ID 0, 1 and 2 within the BCM5325E. ENET4 represents the BCM634x Ethernet MAC/PHY MDI port.

De-selecting the ““Enable virtual ports”” button will disable the system capability to recognize individual Ethernet port within the BCM5325E switch. The WEB UI will display two Ethernet ports, ENET(1-3) and ENET4. The ENET(1-3) represents the BCM634x Ethernet MAC MII port. The ENET4 represents the BCM634x Ethernet MAC/PHY MDI port.

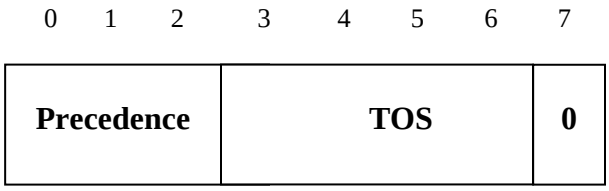
The mapping is shown in below table:

Table 1 BCM96345/BCM96348GW board Ethernet Ports

Ethernet ports	BCM5325E Ethernet switch (connected to BCM634x Ethernet MAC MII port)			BCM634x Ethernet MAC/PHY MID port	BCM634x Ethernet MAC MII port
	BCM5325E port 0	BCM5325E port 1	BCM5325E port 2		
WEB Display	ENET1	ENET2	ENET3	ENET4	n/a
CLI Command “ifconfig –a”	eth1.2	eth1.3	eth1.4	eth0	eth1

Marking

The Marking service will modify the original packet IP header or 802.1q header if all the rules defined within the classification class are matched. The IP marking service can overwrite the IP precedence field and/or TOS field as defined in RFC1394 “Type of Service in the Internet Protocol Suite”.



Mark IP Precedence: select blank if don’t overwrite, or select a priority level from 0 to 7 to overwrite the original Precedence bits.

Mark IP Type of Service: select blank if don’t overwrite, or select one of the following TOS types to overwrite the original TOS bits.

1000 -- minimize delay
0100 -- maximize throughput
0010 -- maximize reliability
0001 -- minimize cost
0000 -- normal service

Mark 802.1p if 802.1q is enabled on WAN: select blank if don’t overwrite the default priority (which is zero) of the 802.1q VLAN header if the outbound PVC has 802.1q VLAN enabled; or select a number from 0 to 7 to overwrite the default priority value in the 802.1q VLAN header if the outbound PVC has 802.1q enabled.

Queuing

Three-priority-level queuing and scheduling in packet QoS is executed in the second stage of the multiple-hierarchy scheduler after the first stage ATM cell level scheduling is done. The ATM SAR scheduler transmits the data among multiple PVCs according to the following order:

1. ATM service category from the highest to lowest order: CBR, rt-VBR, rrt-VBR, UBR
2. For the same ATM service category of different PVCs, or for multiple TX queues from the same PVC: three priority level from high, medium, to low. However for the same PVC, it will not switch to another TX queue until the entire packet is transmitted completely.
3. For the same priority level in multiple PVCs: round robin.

A PVC without packet QoS enabled will be assigned with one TX queue and set to the same priority level as the **low** priority of a packet QoS enabled PVC.

The queuing in packet QoS will become effective only when the packet is forwarded to a QoS-enabled PVC. Packet forwarding is determined by IP routing or bridging, not under control of the packet QoS.

Both packet and cell level scheduling are handled by hardware SAR within the BCM634x chip. For BCM6338, both schedulers are handled by software.

Within WAN configuration, packet QoS can only be configured for a PVC if it meets all of the following conditions:

1. The ATM service category is set to either one of following:
   - UBR-with-PCR
   - UBR-without-PCR
   - non-realtime-VBR

2. The system has at least three un-used ATM TX queues.

Each packet QoS enabled PVC will consume three ATM TX queues. The ATM SAR within the BCM634x chip supports 8 ATM TX queues total. Therefore at maximum, only two PVCs can be enabled with packet QoS. The following shows three examples of maximum number of PVC can be configured with a mixture of PVCs with and without packet QoS enabled:

Consume all 8 ATM TX queues	Number of PVCs with packet QoS	Number of PVCs without packet QoS
Example Configuration #1	2	2
Example Configuration #2	1	5
Example Configuration #3	0	8