No. 103
Technology
White Paper

HRL: Hardware Rate Limiting with Riverstone Networks

Abstract

To allow service providers to effectively manage bandwidth, Riverstone's RS Switch Routers offer an extremely powerful feature called Hardware Rate Limiting (HRL). Metro Service Providers can use this feature to create tiered levels of service, offering pricing schemes such as dollar-per-byte-per-second ($/bps). Content Hosting Providers and ASPs can offer tiered Service Level Agreements (SLA) based on users or applications. Bandwidth on demand is no longer a myth. By simply "turning on" a port, service providers can now increase bandwidth without getting involved with the hardware.

This white paper provides an overview on the HRL features supported in Riverstone's new rate-enabled line cards, and demonstrates how service providers can utilize these line cards in different environments to achieve their business goals by providing customers with value-added services.

Internet and Bandwidth Management

Since its inception in 1968, the Internet has undergone a tremendous transformation. It was initially only intended to be used by researchers and the government. But with the development of the World Wide Web in the 1990s, the Internet has become accessible to everyone, requiring higher speeds and more bandwidth which has driven changes for new technologies. As a consequence, networking infrastructures — especially in terms of hubs, bridges, and routers — have changed immensely. Today, high-density switches and ASIC-based switch routers have replaced most of the traditional equipment. This shift to hardware-based switch routers has been driven by the ever-increasing demand for faster speeds and higher bandwidth.

But even with the advanced technology that allows today's networks to put more bits and bytes on the cable, bandwidth is still not free. Blindly throwing bandwidth at a problem is not the solution. As with any resource in a company, bandwidth needs to be managed and controlled to ensure that it is being used for critical (rather than non-business-related) applications.

River STONE NETWORKS™

Through HRL, Riverstone Networks helps service providers obtain growth objectives with next-generation bandwidth management. Service providers offer a wide array of new services, ranging from network connectivity to Internet access and application maintenance. Just as a power company depends on electricity to generate revenue, bandwidth and applications are the sources of income for these service providers;

their requirements for tight bandwidth control and strict allocation are crucial to their business. The ability to control deliverable bandwidth lets service providers establish SLAs, or offer tiered levels of services, while also enabling them to guarantee traffic rate based on a per-user, per-application, or per-subnet basis. The RS Switch Router is ideal for service providers delivering applications, hosting Websites or wiring multi-tenant unit buildings. Whether local or over the WAN, hardware-based rate limiting will ensure that the service provider has complete control and awareness of how the network is being used.

Riverstone's eQoS Solution

The Riverstone RS Switch Router platform is one of the first hardware-based routers to offer a comprehensive set of electronic Quality of Service (eQoS) features in hardware. Riverstone's new rate-enabled LAN and WAN line cards provide the most comprehensive bandwidth management solution across LAN, MAN, or WAN networks. From Long Haul (70 km) Gigabit Ethernet for new Metro Service Providers to ATM or Packet over SONET, Riverstone's RS Switch Routers can control and manage bandwidth across the entire infrastructure.

This removes the traditional boundaries between LAN, MAN, and WAN networks. All bandwidth management features are implemented in hardware, so service providers can enable bandwidth control at any point in their network, without experiencing the degraded performance that usually comes with traditional software-based routers.

Riverstone's eQoS features include:

- *Hardware Rate Limiting (HRL),* which provides hardware-restricted bandwidth usage based on a predefined profile or per physical port. Traffic exceeding a preallocated bandwidth can either be dropped or reprioritized.
- *Queue Management* techniques that meet service provider or enterprise network requirements. Strict Priority Queuing (SPQ) ensures delivery of high-priority traffic at the expense of low-priority traffic, and Weighted Fair Queuing (WFQ) eliminates bandwidth starvation for low-priority traffic.
- *Traffic Classification and Prioritization* enables traffic to be classified and prioritized according to customer needs and application requirements.

Hardware Rate Limiting (HRL)

HRL enables a service provider to effectively utilize available bandwidth and resources in the network infrastructure. Unlike traditional routers, where rate limiting places a very heavy burden on the central processor — and thus reduces the overall system performance — Riverstone's RS Switch Routers implement HRL to ensure that other functions such as NAT, ACL, or WAN connections are not affected.

All next-generation line cards are rate-enabled to ensure there is no performance impact on the system — hence providing the scalability required for service provider networks. There are three different types of HRL, discussed in the following sections:

- Port Rate Limiting
- Aggregate Rate Limiting
- Per-Flow Rate Limiting

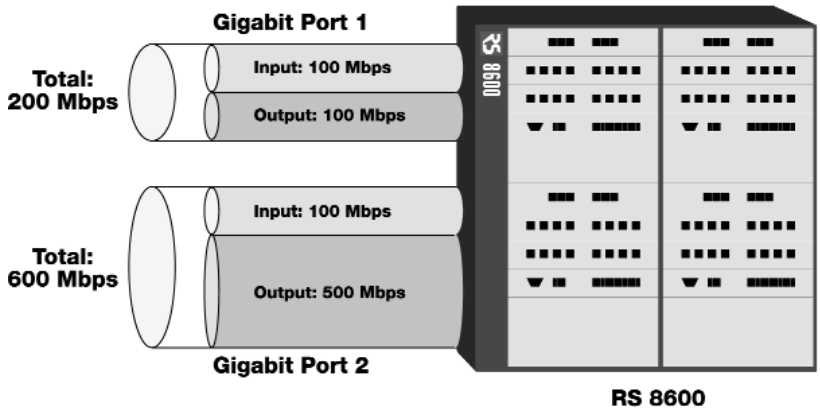
Port Rate Limiting

Port Rate Limiting (PRL) provides bidirectional rate limiting for service providers that want to restrict bandwidth on an inbound or outbound port, regardless of what is transmitted on the physical port.

PRL controls all available bandwidth on a physical port, turning a Fast Ethernet or Gigabit Ethernet connection into any other lower-speed connection, based on the customer's bandwidth requirements. Inbound and outbound traffic can be configured separately, offering fractional services on a physical port. This can be used to create asymmetric links for users who (for instance) might demand more download than upload bandwidth.

PRL is set separately for each physical port, enabling the service providers to configure different rate limits depending on utilization and traffic requirements. Thus, a service provider can determine how to handle excessive traffic. For instance, traffic can either be dropped or reprioritized.

In addition, a packet's "type of service" field can be rewritten. In PRL, inbound and outbound traffic can be rate-limited differently on both bridged and routed ports, allowing customers to fully utilize the switching and routing capability in the Riverstone RS Switch Router. PRL restricts traffic that enters or leaves a specific port, disregarding the traffic or protocol type; this is far less complex than configuring rate limits based on traffic or protocol type.



In the above example, a rate-enabled Gigabit Ethernet port is configured to a full duplex Fast Ethernet connection. The other Gigabit Ethernet port is configured to an asymmetric Fast Ethernet connection, with five times the output bandwidth.

Aggregate Rate Limiting

Aggregate Rate Limiting (ARL) provides a method to control bandwidth consumption on a specific traffic or protocol pattern, based on traffic policy. Traffic policy can be created to control the aggregate traffic to or from a subnet, and the aggregate traffic for a specific application within the subnet.

River STONE NETWORKS™

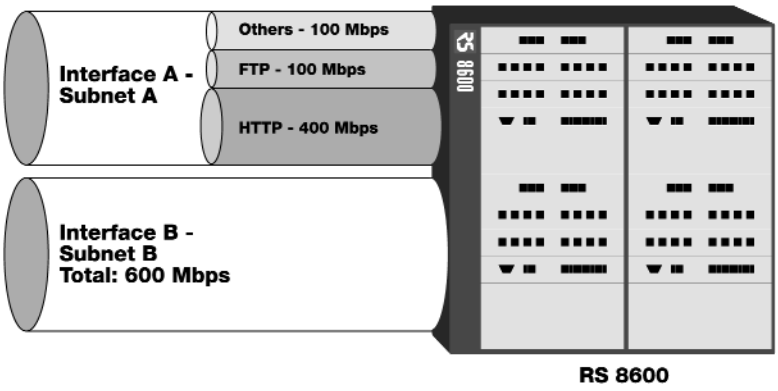
Each traffic policy can consist of multiple applications, allowing flexibility to specify the type of traffic to be controlled. Once the traffic type is identified, a rate limit is applied. Multiple traffic policies can be used to create different limits for different types of traffic utilizing the same interface, thereby enabling complete control of the amount of bandwidth allocated to particular applications.

Unlike PRL, which provides bandwidth control on the physical port, ARL provides control of the total bandwidth consumed by the specific type of traffic that enters an IP interface. All traffic that does not match the traffic policy is not being limited. Similar to PRL, action can be taken relative to traffic that has exceeded the predefined threshold. This traffic can be dropped, reprioritized to a lower priority, or overwritten in the "type of service" field of the packet.

Traffic policy can be defined based on any of the IP headers listed right:

IP Protocol Header

Source IP Address
Destination IP Address
Source Port Number
Destination Port Number
Protocol Type
"Type of Service"



RS 8600

In the above example, multiple traffic policies can be configured on interface A to control the bandwidth allocation for different types of traffic. If you don't need granular bandwidth control, the configuration can be simplified by setting a limit on the total bandwidth allowed to enter an interface. On interface B, the total amount of bandwidth allowed to enter is configured to be 600 Mbps.

Hardware Credit Bucket

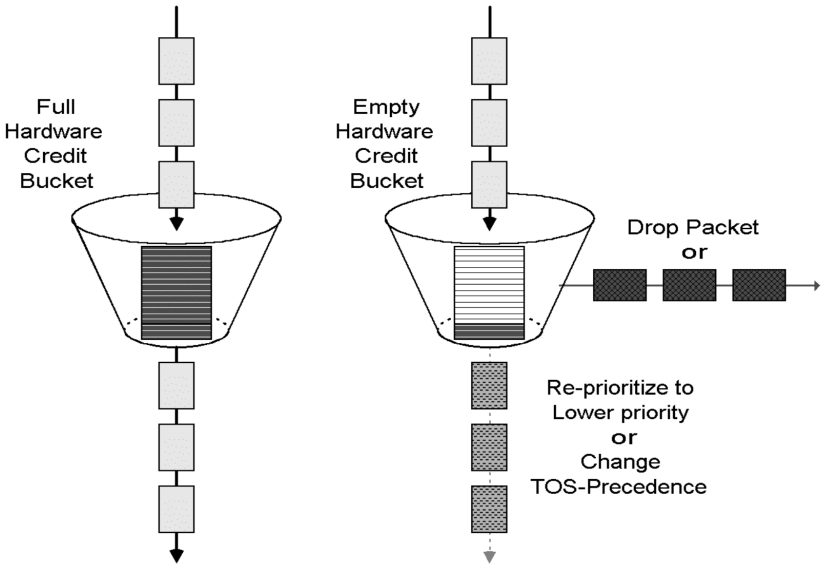
Both PRL and ARL use a hardware credit bucket scheme to provide bandwidth control on hardware. Hardware credit buckets are built into all rate-enabled line cards to ensure wire-speed performance. For PRL, the hardware credit bucket scheme limits how much traffic is allowed to enter or leave a physical port. As for ARL, the hardware credit bucket limits how much traffic is allowed to enter an interface.

The service provider controls the amount of credit in the bucket, depending on the amount of bandwidth allocated (the more bandwidth, the more credit for that type of traffic). When traffic passes through a physical port or matches the predefined profile passing through an interface, credit will be deducted from the bucket. When all credits have been used, and traffic has exhausted all the preallocated bandwidth, action will be taken on the excessive traffic by either dropping the packet, reprioritizing it to a lower priority, or rewriting the "type of service"-precedence value in the packet.

The credit bucket scheme consists of two major components — credit count and replenish rate.

Credit count is the amount of credits in each bucket, which define how much credit there is for a physical port or a specific type of traffic. The credit count determines when traffic is considered to be excessive (based on the availability of credit).

The replenish rate determines when or how often the credit buckets are refilled. If the replenish rate is too low, there will not be sufficient tokens when traffic arrives. If the replenish rate is too high, it will reduce the effectiveness of the rate-limiting policy. Both the credit count and replenish rate are calculated by the system according to the rate-limiting profile defined during the configuration process. Riverstone's RS Switch Routers use a special algorithm to determine the best credit count and replenish rate to achieve optimum bandwidth control.



The above figure depicts the operation of the hardware credit bucket. Hardware credit buckets are located on the rate-enabled line card, and all credit validation is locally performed by the hardware at wire speed. The control module is only involved when a new entry needs to be created in the hardware. It programs the entry to the advanced ASICs on the line card, and the subsequent packets are handled locally. The hardware credit buckets then measure credit based on the amount of traffic that matches the rate-limiting profile.

River STONE NETWORKS™

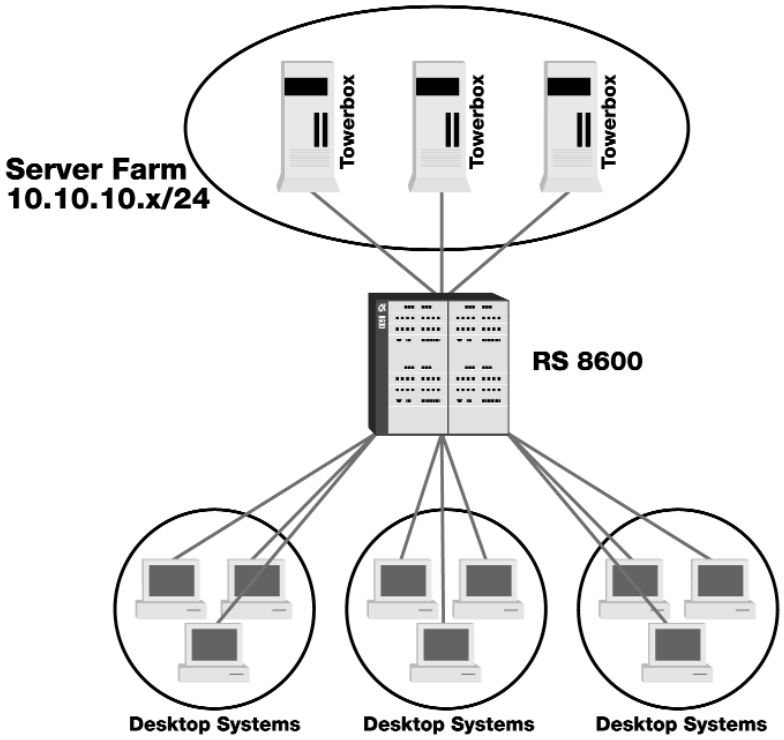
Per-Flow Rate Limiting

Per-Flow Rate Limiting restricts the bandwidth on a per-flow basis. To understand Per-Flow Rate Limiting, you must understand the concept of a flow, which is an entry that contains both the network addresses for all senders and receivers, the application port number, and the protocol type. The Riverstone RS Switch Router can identify each flow based on the IP header listed in Figure 3.

A flow is created when a user initiates a conversation. An IP flow will contain the source IP address, destination IP address, source IP port number, destination IP port number, protocol type, and "type of service" value. Based on the various hosts and applications, different flows will be created. By identifying the flow based on the above parameters, administrators can apply an upper rate limit to this flow to restrict bandwidth consumption.

If a customer doesn't adhere to an SLA, specific actions can be defined for packets that have exceeded this upper rate limit. Based on the policy, packets might either be dropped or reprioritized to different queues. Per-Flow Rate Limiting is configured on the inbound interface, and excessive traffic is dropped or reprioritized before being forwarded to any outbound port.

It is also important to understand that Per-Flow Rate Limiting restricts the bandwidth of each flow, which limits communications between two end stations for a specific application. This provides a bandwidth control on a per-connection basis.



In the example above, users can access multiple servers and multiple flows can be created between users and servers. Bandwidth used by each of the flows can be limited by using Per-Flow Rate Limiting.

Applications

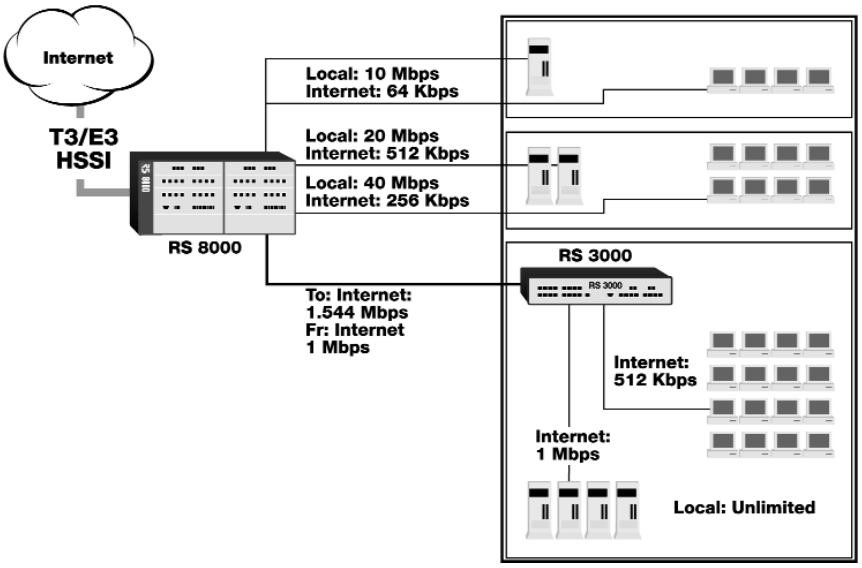
The following sections describe eQoS applications for different types of service providers.

Multi-Tenant-Building Service Providers

Multi-Tenant-Building (MTB) Service Providers represent a new breed of service providers that offer connectivity and Internet services. Property owners and managers are aware of the importance and the value of connectivity and Internet services to their tenants. As such, they are beginning to provide these services in the same way they provide water, gas, and electricity. The idea is to offer network connections to tenants who want network connectivity and Internet access in their offices, but do not have the expertise, personnel, or resources to set up a small office network with Internet access. The MTB Service Providers provide direct connections to building tenants, and allow them to have connectivity and Internet access through the "building" network.

With PRL, MTB Service Providers can control the bandwidth for each network connection, and therefore invoice tenants by their access rate. Together with ARL, customized plans can be created to offer differentiated network and Internet access rates. This allows the service provider to have separate rates for local and Internet traffic. Users who want to have higher-speed Internet access can be billed according to the extra bandwidth, while users who only need limited access rate can opt for the minimum package.

Due to the nature of applications, upload and download traffic patterns — especially with Internet traffic — are often very different. For Internet access, users often require more download bandwidth, whereas extranet applications typically require more upload bandwidth. In response to these needs, MTB Service Providers can implement an asymmetric service plan to sell download and upload bandwidth to their tenants at different prices, based on specific requirements.

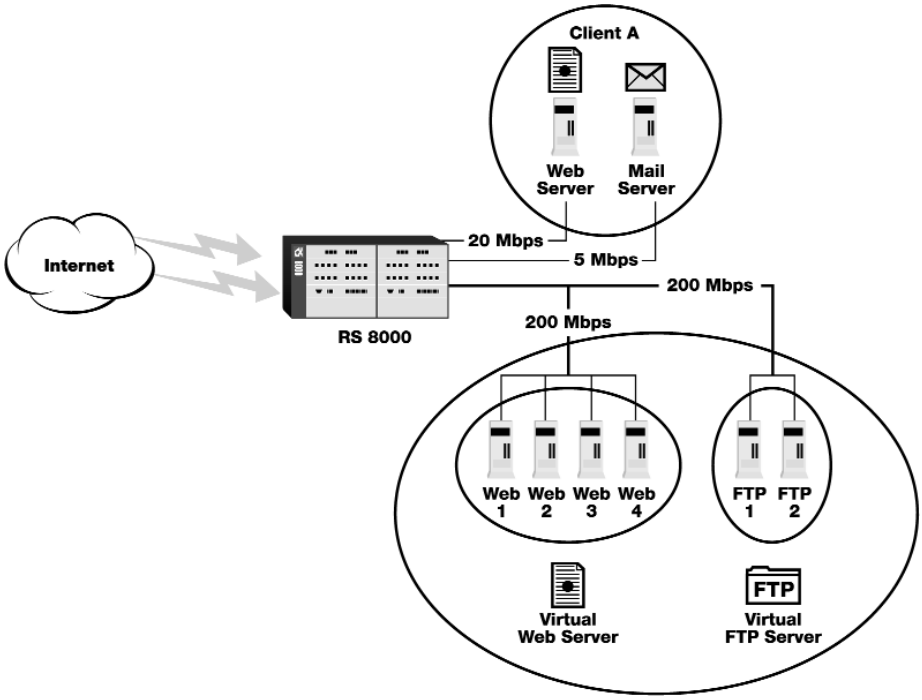


River STONE NETWORKS™

Application Service Providers

An Application Service Provider (ASP) offers consultation and application outsourcing to clients who need certain applications but don't have the resources and knowledge to implement or maintain them on their own. Clients often want an ASP to design, run, and maintain the application on their behalf. This drastically reduces the learning curve and deployment time for the customer. Today, hundreds of ASPs are offering or selling applications to their clients, ranging from e-mail and Web hosting to Enterprise Resource Planning (ERP) systems.

The primary role of an Application Service Provider is to live up to the customer's SLA. It is critical that ASPs offer tiered levels of services with predefined rates that are defined in the SLA. By combining Port Rate Limiting, Aggregate Rate Limiting, and other features such as Smart Trunking and Server Load Balancing, Application Service Providers can create a highly-reliable and fault-tolerant infrastructure — and still have full control over the bandwidth allocation for each individual application. In



this example, the ASP is hosting different applications for two clients. Client A, a small startup company, requires only minimal bandwidth and does not want to pay for any redundancy on its servers. Client B, meanwhile, requires highly reliable and maximum bandwidth for its Web-hosting application, and moderate bandwidth for its FTP server. By offering different connection speeds to the servers, the ASP can provide multiple access rates for each application — and therefore demand different premiums from customers.

River STONE NETWORKS™

Conclusion

eQoS features from Riverstone's award-winning RS Switch Routers enable service providers to offer differentiated services by utilizing rate-enabled line cards that support per-port, per-flow, and aggregate hardware rate limiting. These powerful features provide service providers the edge to succeed and win in this highly competitive market.

Riverstone Networks provides service providers with an infrastructure that yields increased network reliability, improved performance and enhanced services.

Acronyms

ACL	*Access Control List*
ANSI	*American National Standards Institute*
ASIC	*Application-Specific Integrated Circuit*
ASP	*Application Service Provider*
ATM	*Asynchronous Transfer Mode*
CBR	*Constant Bit Rate*
DS1/DS3	*Digital Signal, Level 1 (1.54 Mbps) or 3 (44.7 Mbps)*
DSL	*Digital Subscriber Line*
E1/E2	*European Trunk 1/2 (2 Mbps/34.3 Mbps)*
ERP	*Enterprise Resource Planning*
HSSI	*High Speed Serial Interface*
ISP	*Internet Service Provider*
ITU	*International Telecommunications Union*
LAN	*Local Area Network*
LEC	*Local Exchange Carrier*
MAC	*Media Access Control*
MAN	*Metropolitan Area Network*
MDU	*Multiple Dwelling Unit*
MLPPP	*Multi Layer Point-to-Point Protocol*
MTU	*Multiple Tenant Unit*
OC-3/OC-12	*Optical Carrier 3/12 (155 Mbps/622 Mbps)*
POS	*Packet over SONET*
PPP	*Point-to-Point Protocol*
PVC	*Private Virtual Circuit*
QoS	*Quality of Service*
RED	*Random Early Discard*
SLA	*Service Level Agreement*
T1	*Trunk 1 (1.544 Mbps)*
TCP/IP	*Transport Control Protocol/Internet Protocol*
TDM	*Time Division Multiplexing*
UBR	*Undefined Bit Rate*
VBR	*Variable Bit Rate*
VLAN	*Virtual LAN*
VoD	*Video on Demand*
WAN	*Wide Area Network*
WDM	*Wave Division Multiplexing*
WRED	*Weighted Random Early Discard*

River STONE NETWORKS™

Riverstone Networks, Inc.
5200 Great America Parkway, Santa Clara, CA 95054 USA

408 / 878-6500 *or* ***www.riverstonenet.com***

© 2000 Riverstone Networks, Inc. All rights reserved. RS, IA, Intrinsic Persistence Checking, Sticky Ports, and Comprehensive Server Checking are trademarks and service marks of Riverstone Networks, Inc. All other product names mentioned herein may be trademarks or registered trademarks of their respective owners. All specifications are subject to change without notice.

Printed in the USA *v 1.4 9/00*