

# Acceleration Systems Adaptive Queue Management & Quality of Service

May 2015, v1.4

## Table of Contents

- 3 Overview
- 3 TCP Traffic Management
- 4 Adaptive Queue Management
- 5 Traffic Prioritization

## Overview

The worldwide web had yet to be conceived when the TCP/IP protocols were defined in 1974. As TCP/IP standards were applied to larger and larger groups of users and extended from LANs contained in a single building to WANs extending across oceans, the volume of IP traffic exploded. During the development of the ubiquitous global web of interconnectivity, various limitations inherent to TCP/IP have surfaced. One of those issues is how to manage networks that send data faster than the network on the other side of a gateway has the capacity to handle. If not handled properly, excessive traffic at a gateway negatively affects overall network efficiency and impacts quality of service.

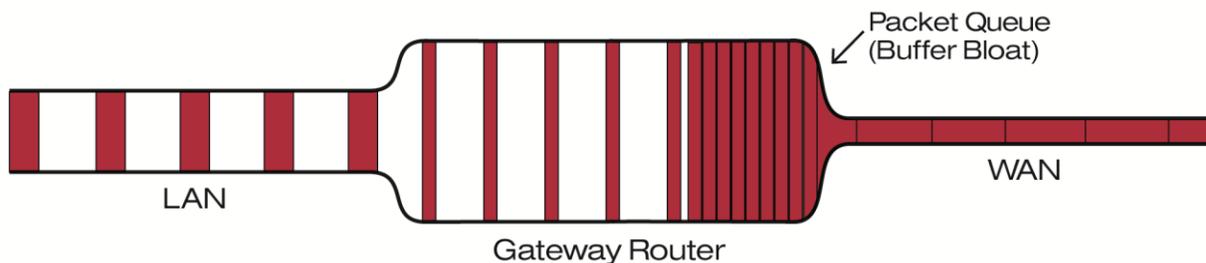
Acceleration Systems employs a first-to-market, advanced traffic management algorithm that senses changes in the *upstream* capacity of an Internet connection on a near real-time basis and regulates the rate that LAN-side applications send data to the Internet gateway device. This avoids long queues in the buffer at the gateway where QoS is applied. Avoiding backups improves overall throughput and assures that the queue manager prioritizes time sensitive data such as VoIP in a timely manner.

## TCP Traffic Management

When two networks running at different rates interface with one another (for example a 1 Gbps LAN connected to a 10 Mbps Internet connection), the flow of packets from the faster network must be managed. To deal with the difference in speeds, routers contain buffers that hold a specific number of data packets. These buffers serve much like shock absorbers on a car.

When bursts of data arrive at a router, the excess data queues up in the buffer. The buffer stores the sudden burst of excess data and feeds that information to the slower network at a rate that matches the capacity of the second link. This is the shock absorber affect.

Ideally, the queue quickly drops to a manageable level or dissipates altogether. But if the high speed side continues to send data at a rate above the capacity of the second leg of the connection, the queue will not dissipate. It will increase. This increase is referred to as buffer bloat.



The greater the disparity between LAN and WAN speeds, the more likely the gateway will develop bottlenecks that impede the transfer of data to the slower WAN and which impact the timely prioritization of packets by the Queue Manager.

The TCP mechanism used to manage congestion determines available bandwidth by sensing packet drops. It starts sending data slowly and increases data transfer rates until packets drop. Then it slows down the packet transfer. After some number of cycles of increasing and decreasing transfer rates, TCP theoretically finds the optimal data rate. The reliance upon dropped packets assumes that packets get dropped in a timely manner. TCP has no mechanism to detect if data packets are stored in a buffer. Thus, TCP does not slow until after the gateway buffer discards packets due to data overflow.

While intended to smooth the flow of data between networks, buffers reduce interactivity with the TCP congestion control algorithm and can exacerbate transmission delays on the network. This is especially important when managing two or more simultaneous transmissions over the same network. If all sources do not experience packet loss at the same time, some sources could continue transmitting too fast. Buffer bloat results when TCP fills the gateway buffer and continues to transmit to the gateway at rates which keep the buffer full or cause it to overflow, rather than slowing to a speed that matches the capacity of the next segment.

## Adaptive Queue Management

Adaptive Queue Management (AQM) does not control transmission rates based upon dropped data packets. AQM, instead, monitors the time packets spend in the gateway buffer. AQM keeps buffer induced delay at or below 5 milliseconds. If packets reside in the queue for more than 5 milliseconds, the algorithm slows the rate of packets arriving at the gateway.

Adaptive Queue Management offers a number of advantages.

- There are no parameters to configure.
- AQM responds directly to conditions in the buffer of the gateway router, not external conditions such as round-trip delay, connection rate, traffic load, and other factors that cannot be controlled or predicted at the local buffer.
- Local queue delay is determined when packets leave the buffer. This avoids further delay while waiting for acknowledgements from the next network segment.
- AQM adjusts to changing link rates with no negative impact on available network utilization.

If the gateway buffer is relatively empty or if the delay is below the 5 millisecond limit, Adaptive Queue Management takes no action. Tests have shown that AQM produces link utilization consistently near 100% of the available bandwidth.

## Traffic Prioritization

Configuring Quality of Service for a network has long been a challenge for network administrators. Establishing priorities and bandwidth allocations for various types of traffic can be a daunting task.

Conventional QoS assumes that connections run at a constant speed. But most connections are shared, so the capacity changes constantly. Even on dedicated links the amount of available bandwidth fluctuates as users contend for a finite resource. So, a key underlying assumption to conventional QoS is seldom, if ever, met.

As discussed previously, conventional TCP flow control (and subsequent refinements such as RED and explicit congestion notification) do not prevent buffer bloat. Once the gateway buffer fills or overflows, time sensitive packets such as VoIP may get delayed due their position in the queue. As a result, the queue manager may not prioritize VoIP packets quickly enough to maintain call quality – even when VoIP traffic is given a top priority in the QoS configuration.

Acceleration Systems developed a prioritization hierarchy for protocols and data types. Our priority stack has proven successful in over 95% of the environments in which it has been deployed. This virtually eliminates the need for custom configuration. Administrators do have the ability to adjust QoS settings to meet local requirements in the rare instances where doing so would be beneficial.

With Acceleration Systems' Adaptive Queue Management efficiently controlling the flow of data across the network, the queue manager is able to prioritize traffic in a timely manner based upon a predefined hierarchy of protocols and data types. This combination of flow control and prioritization produce a more efficient use of network capacity than has previously been achieved and does so without little or no effort on the part of network administrators.