THE BROADBAND SERVICE OPTIMIZATION HANDBOOK

Allot communications

CHAPTER 4

Phase Two: Application Control

As discussed in previous chapters, deep packet inspection (DPI) allows you to gain visibility into Layers 4–7 and to complete the first phase of service optimization: the establishment of a performance baseline. But baseline monitoring of network behavior is merely the first step in controlling your network and distinguishing your service offerings from those of your competitors.

Knowing what is happening on your network without taking additional measures to control it is a lot like having a severe storm warning system in place but no emergency shelters or evacuation plans. Rather, once you have established a network performance baseline (see Chapter 3), you will need to deploy methods to control that traffic based on what you have learned. The data in your baseline statistics give you the performance information you need to set up appropriate application controls in the network.

Application control allows you to intelligently manage — some might say "dictate" — the way that resources are allocated to the different applications on your network for overall optimum performance. Application control, which relies on DPI for application recognition, involves classifying traffic and assigning actions to each traffic class, thereby creating network rules, or policies. The aim of application control is to meet subscribers' expectations for the quality of experience (QoE) associated with the different applications they use. The next chapter will discuss further controlling traffic, again using traffic classification and assigning policies, but drilling down to the subscriber level to enforce service-level agreements (SLAs) and to create tiered pricing and other service packages.

As noted, application control is made up of two basic steps: *traffic classification* and *assigning actions* for network policy creation and enforcement. Let's take a look at each.

Traffic Classification: Grouping Packets

The first step toward gaining control over network traffic's behavior is to identify the applications used over the network and to classify them into groups with similar QoE expectations. The types of traffic in each category should have enough in common to warrant similar treatment across the network. The size of your network, the number and types of subscribers, the types of applications, your specific service offerings, and service levels are all variables to consider when creating specific classifications.

For example, let's look at some common types of traffic: voice-over-IP (VoIP), gaming, and peer-to-peer (P2P) applications. Both VoIP and gaming are delay-sensitive applications. So you might choose to create a "latency-sensitive" traffic class and place VoIP and gaming in that group. You might then choose to assign the latency-sensitive traffic class a "high priority."

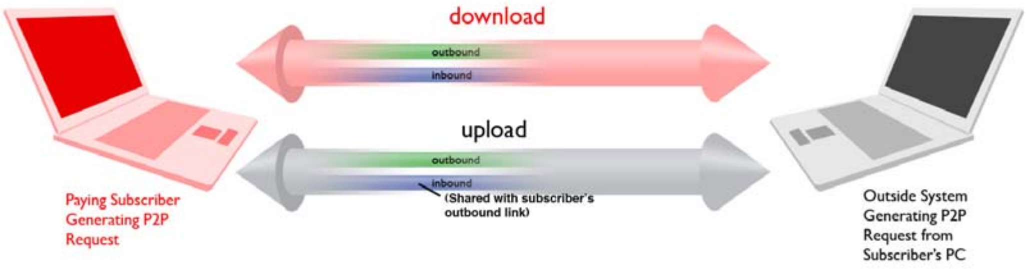
For the P2P traffic, you might want to create a class for P2P file-sharing applications and differentiate between P2P *download* and *upload* traffic. Both types of traffic generate inbound and outbound packets.

Download traffic is content accessed by your subscriber. It generates some traffic in the outbound direction in the form of initiation requests and packets necessary to maintain the content streaming session. The majority of the download traffic is in the inbound direction as the content flows from an Internet source to the subscriber's machine.

Upload traffic, on the other hand, is content requested by a non-subscriber directly from your subscriber's machine. It, too, generates inbound and outbound traffic. The inbound traffic toward the non-subscriber's computer shares an access connection with your subscriber's outbound traffic and can thus interfere with—and degrade—your subscriber's P2P streaming QoE.

You could assign, then, P2P download traffic flows a medium priority, because they directly serve your subscriber, but assign P2P traffic uploads a low priority, because they do not. (See Figure 1) This ability to differentiate, classify, and control traffic based on whether a paying subscriber or a non-paying source is generating the content request—a more granular capability than simply generically controlling "inbound" and "outbound" P2P packets, which could be generated from either source—is a strong tool in delivering the QoE expected by paying subscribers.

Figure 1. P2P Downloads and Uploads


download
outbound
inbound
upload
outbound
inbound
(Shared with subscriber's outbound link)
Paying Subscriber Generating P2P Request
Outside System Generating P2P Request from Subscriber's PC

Within both P2P downloads and uploads, there are inbound and outbound exchanges of information. DPI allows you to identify your subscribers' inbound/outbound flows (downloads) and give that traffic a greater priority than the inbound/outbound flows generated by an outside system (uploads) to prevent degradation in the subscriber's QoE.

The way you structure traffic classifications establishes the foundation of your IP service optimization framework, which ultimately dictates how your networked application traffic behaves. You can judge how detailed your classification system should be based on a number of variables: the complexity of your network, the size of your subscriber base, the range of your offerings and applications, your churn rate, your future growth, and long-range plans. Part of that decision-making process also involves subscriber management and policies you establish for user-definable traffic sessions, which will be detailed in the next chapter.

Assigning Actions to Build Network Policies

A network policy determines the actions that will be taken when certain network conditions exist for specific classes of traffic. For example, if P2P download traffic was considered a medium priority, the network manager might want to set a *rate limit* on the P2P download traffic so it did not “hog” bandwidth but still provided the subscriber the expected QoE. The actions you assign to these application flows using your service optimization system create the policy and can include any combination of the following:

Rate limiting ensures that certain traffic does not consume more than a reasonable amount of available bandwidth. This action sets a limit to the amount of bandwidth that can be consumed by a class of traffic or a given traffic flow within a class. By limiting bandwidth consumption, you improve your subscriber’s QoE while also conserving network resources.

Bandwidth guarantees assign a minimum amount of bandwidth to a specific traffic type or flow, ensuring that a particular application — such as low-bandwidth but latency-sensitive VoIP — always has the resources it needs to perform as expected.

Traffic blocking drops all packets of a certain traffic class, which can be useful in keeping malware and unwanted applications off the network and preventing access from specific IP source addresses.

Relative prioritization is a simple and logical way to ensure that, when congestion occurs, certain traffic flows are passed through the network ahead of others.

See Figure 2 for a summary of possible traffic classifications and corresponding actions. Additional actions can be assigned at the subscriber level to govern subscriber activities. These will be discussed in the next chapter on subscriber management.

Figure 2. Sample Traffic Classifications

Traffic Type	Group Members	Applications / Protocols	Priority or Action
Unwanted Traffic	Known worms/viruses	N/A	Drop/block
Latency-Sensitive Traffic	VoIP (own offering)	UDP	High priority
	Other VoIP Services	Google Talk Vonage Skype Net2Phone	Medium priority/ bandwidth guarantee
	Gaming	World of WarCraft Game Spy X Box Live Lineage	High priority/ bandwidth guarantee
Web Traffic	HTTP	Method (GET, POST, etc.) URL (e.g., file types) Host Name Mime Types	Medium to low priority
	Email	POP POP2 POP3 SMTP	Medium priority
	VPN	IPsec SSL/TLS	High to medium priority
P2P	P2P Upload	Ares Encrypted Ares KaZaA eDonkey Gnutella BitTorrent Encrypted BitTorrent Others	Low priority/rate limit
	P2P Download	Ares Encrypted Ares Kazaa eDonkey Gnutella BitTorrent Encrypted BitTorrent Others	Medium priority/rate limit

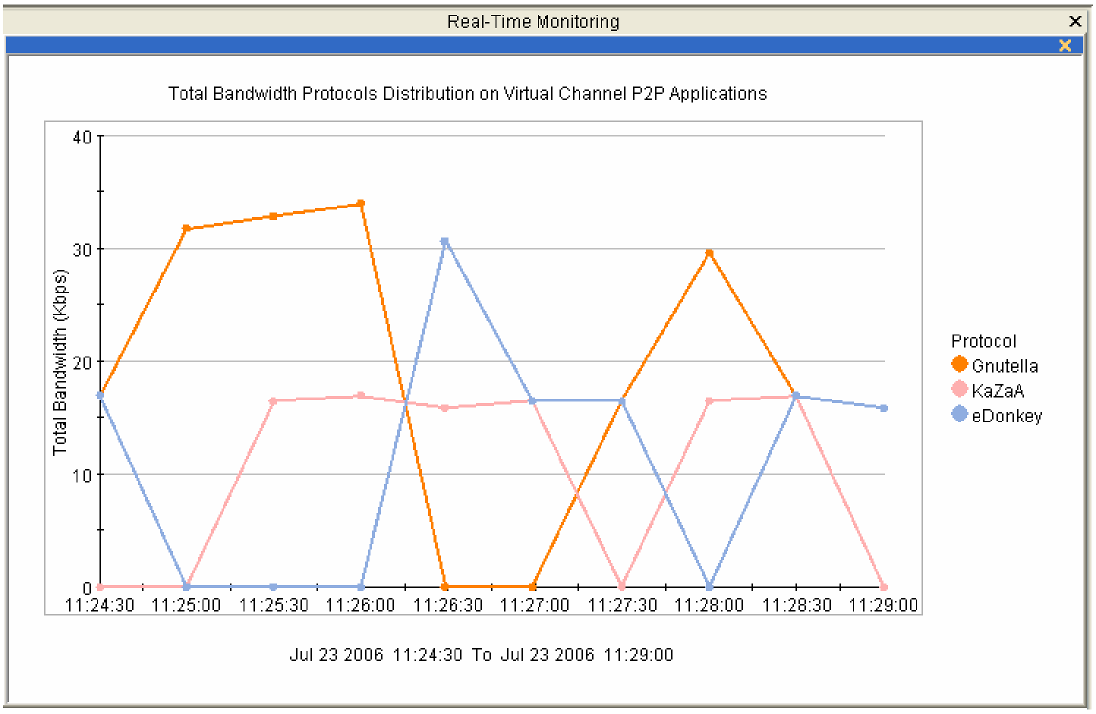
Curbing ”Bad” Traffic

Controlling traffic to enhance your subscribers’ experience includes isolating malicious traffic and security threats. Your network manager can gain visibility into the network through real-time graphical reports and will be able to identify when the network is being misused, by whom, and by what applications and protocols. Such real-time analysis speeds troubleshooting and allows a more rapid resolution to most problems. Figures 3 and 4 show two samples of graphical reports: “Protocol Distribution over a Four-Minute Period” and “Week-long Report Showing Bandwidth Consumption,” respectively.

A service optimization solution that offers active monitoring enables a service provider to make real-time adjustments to application and subscriber policies, to traffic rates and network policies, and to service-level prioritizations. Types of reports that enable service providers to troubleshoot and manage network performance in real-time include the following:

- Those with pie charts showing total policy and application distribution of traffic currently on the network
- Those that include details about specific network events
- Those that reveal trends by showing a detailed view of the fluctuations in network utilization, including peak and average usage for each circuit, even down to the subscriber level on a minute-by-minute basis
- Those that track response times and historical round-trip delay on all monitored circuits

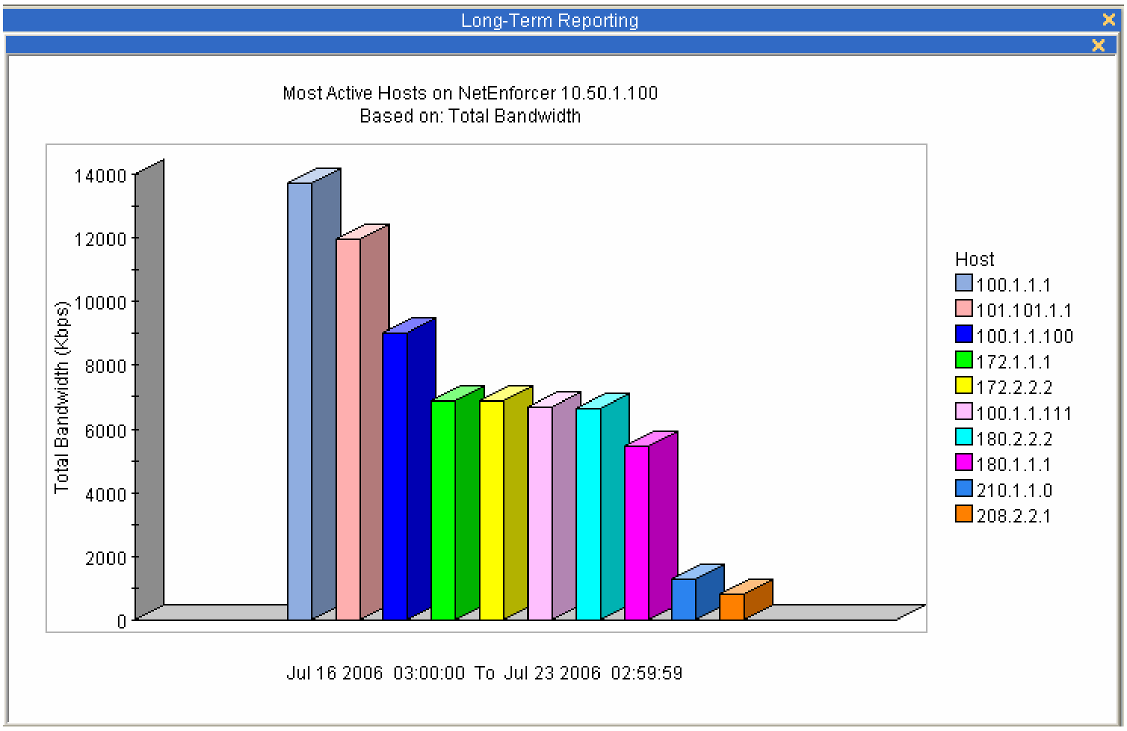
Figure 3. Protocol Distribution over a Four-Minute Period



In addition to real-time statistics, it is possible to have reports generated on a daily, weekly, or monthly basis to:

- keep your baseline up to date
- plan for capacity
- spot long-term trends
- verify that SLAs are being met
- monitor subscriber usage levels to spot opportunities for service upgrades that will increase average revenue per user (ARPU).

Figure 4. Week-long Report Showing Bandwidth Consumption



With the combination of real-time reports and statistics gathered over time, the network administrator has improved visibility into the network. Armed with this information he can set policies and create conditions that curb malicious traffic. For example, to thwart denial-of-service (DoS) attacks and distributed DoS (DDoS) attacks that flood the target system until it shuts down, a policy to control the number of connections per second can be set. Another policy might define a ceiling on the number of new connections per second that any application is allowed, such as 1000 connections per second. With that policy in place, a DoS or DDoS attack that tried to open tens of thousands of new connections would be prevented by the pre-set policy. At the same time, the real-time monitoring reports would show that someone or some application was exceeding the connections threshold, and the operator would be able to determine what is happening and who the culprit is.

Chapter Summary

Application control is the second phase of IP service optimization. Also reliant on the powerful capabilities of DPI, it involves classifying traffic and setting network rules to create automated policies that govern application behavior. The ability to accurately identify applications is essential to understanding and controlling your network, and provides the hierarchy upon which a range of actionable policies can be set. Service optimization also involves utilizing real-time graphical reports to help a traffic manager troubleshoot network issues as they occur to keep the network running smoothly.