

Connection-Rate Filtering Based on Virus Throttle Technology



Technical Brief

Introduction	3
The Limitations of Current Responses.....	3
The Need for a New Solution.....	3
Virus Throttle: What it is & How it Works	4
What it is	4
Tests Show Quick Detection, Prevention	4
Benefits of Virus Throttle Technology	4
How it Works.....	5
Connection-Rate Filtering in ProCurve Switches	5
Filtering Options	5
Sensitivity and Application Options	6
ProCurve Manager Plus	7
Configuration Guidelines	7
For a network that is relatively attack-free.....	7
For a network that appears to be under significant risk of attack	7
ICMP Rate Limiting in ProCurve Switches	9
What is Rate Limiting.....	9
Effect of ICMP Rate Limiting.....	9
ICMP Rate Limiting Operation.....	9
Network Application.....	9
Summary.....	10

Introduction

As every information-technology manager knows, computer virus epidemics are only getting worse. In 2003, the SQL Slammer worm infected 75,000 computers in one minute, making it the fastest-moving virus yet seen, and caused major network disruptions worldwide. Nimbda, Blaster, Code Red, Sasser and Welchia are continual threats as well. Today, computer users are directly threatened by more than 97,000 viruses, worms and Trojan horses.¹ Increased usage of network applications such as Instant Messages also increases the risk of virus infection. In the 3rd quarter of 2005, the volume of IM threats was more than 3,000 percent higher than the previous year, according to IMlogic Threat Center.

To protect themselves from the onslaught of traffic generated by computer viruses, many corporations shut down portions of their network infrastructure; when they can't act fast enough, entire network subnets or even entire networks can be brought down by viruses. Either way, the viruses cost corporations incalculable sums in lost productivity. Beyond bringing normal operations in an office or enterprise to a halt, computer viruses can put attacker-defined code on a system to cause additional damage.

Network threats once were slow-moving and easy to defend against when information transfer was done largely by sharing floppies. Organizations had the time they needed to clean their networks and install defenses. However, as CPU speeds increase, bandwidth grows, networks become more business critical and clients become more mobile, network administrators increasingly lack the time to shut down operations or develop inoculations to cure the infections.

Nor is productivity the only victim of network viruses. The SQL Slammer virus took out a 911 emergency response center serving two police departments and 14 fire departments near Seattle. Protecting against computer viruses can ultimately be an effort to protect lives.

The Limitations of Current Responses

Current methods to stop the propagation of malicious agents rely on the use of signature recognition to prevent hosts from being infected. That is, they seek to prevent the virus or worm from entering the system. These methods concentrate on the physical characteristics of the virus—i.e., its program code—and use parts of this code to create a unique signature. Programs entering the system are compared against this signature and discarded if they match.

While this approach has been effective in protecting systems, it has several limitations which, as the number of viruses increase, decrease its effectiveness. It is fundamentally a reactive and case-by-case approach in that a new signature needs to be developed for each new virus or variant as it appears. Signature development is usually performed by skilled people who are able to produce only a certain number of signatures at a time. As the number of viruses increase, the time between initial detection and the release of a signature also increases, allowing a virus to spread further in the interim.

This latency between the introduction of a new virus or worm into a network and the implementation and distribution of a signature-based patch can be significant. Within this period, a network can be crippled by the abnormally high rate of traffic generated by infected hosts.

As long as attacks occur at “machine speed” and responses are implemented at “human speed,” computers will essentially be defenseless against new threats. As systems get bigger and more complex, so does the problem of addressing new threats.

The Need for a New Solution

A different solution is needed. A truly resilient infrastructure would include a solution that automatically hampers, contains and mitigates attacks by previously unknown threats, giving the people responsible for an infrastructure's security the time they need to implement a response.

Rather than replacing current, signature-and-patch-based protections, the new solution would complement them by allowing computers and humans to each do what they do best: computers can respond far more quickly than people, but are poor at gauging the nature of a previously unknown threat. Humans are good at making such decisions, but are slow—by machine standards—to act. A new solution would have computers acting quickly to stabilize a situation until humans could intervene.

¹ <http://msnbc.msn.com/id/6679126/>; <http://msnbc.msn.com/id/4065701/?p1=0>

Virus Throttle: What it is & How it Works

Connection-rate filtering based on Virus Throttle technology is a new, HP-developed solution that overcomes the limitations of previous responses and meets the need for rapid containment and mitigation of attacks by malicious agents.

What it is

Traditional approaches to anti-viral protection are based on the actual code or signature of the virus. Virus Throttle, in contrast, is based on the behavior of malicious code and the ways in which that behavior differs from that of normal code. Virus Throttle is based on the observation that under normal activity, a computer will make fairly few outgoing connections to new computers, but instead is more likely to regularly connect to the same set of computers. This is in contrast to the fundamental behavior of a rapidly spreading worm, which will attempt many outgoing connections to new computers. For example, while computers normally make approximately one connection per second, the SQL Slammer virus tried to infect more than 800 computers per second.

The idea behind the Virus Throttle is to put a rate limit on connections to new computers, such that normal traffic remains unaffected but suspect traffic that attempts to spread faster than the allowed rate will be slowed. This creates large backlogs of connection requests that can be easily detected. Once the virus is slowed and detected, technicians and system administrators have the time they need to intervene in order to isolate and eradicate the threat by cleaning it from the system.

This approach differs from signature-and-patch approaches in three key ways:

1. It **focuses on the network behavior of the virus** and prevents certain types of behavior—in particular, the attempted creation of a large number of outgoing connections per second.
2. It is also unique in that, instead of stopping viruses from entering a system, it **restricts the code from leaving**.
3. Because connections exceeding the allowed rate can be blocked for configurable periods of time, **the system is tolerant to false positives** and is therefore robust.

Virus Throttle technology is not meant to replace signature-based solutions but, rather, to complement them. Virus Throttle fills a gap in anti-virus protection that previously allowed unknown threats to wreak significant damage before patches could be deployed. With Virus Throttle, previously unknown threats can be mitigated, giving administrators time to deploy signature updates and patches against further attack.

Tests Show Quick Detection, Prevention

Tests of Virus Throttle technology conducted at Hewlett-Packard Labs in Bristol, U.K.² show that Virus Throttle is able to very quickly detect and prevent worms spreading from an infected computer. For example, the throttle is able to stop the W32/Nimda-D worm in less than one second.

Since the throttle prevents subsequent infection, the effect on the global spread of a virus depends on how widely the throttle is deployed. HP Labs results show that when only 50 percent of computers are installed with the throttle, the global spread of both real and constructed worms is substantially reduced. Throttled machines do not contribute any network traffic in spite of being infected, significantly reducing the amount of network traffic produced by a virus.

Benefits of Virus Throttle Technology

The benefits of Virus Throttle technology include the following:

- **Works without knowing anything about the virus.** Because it is triggered by the behavior of a virus rather than by identifying the code of the virus, it can handle unknown threats without waiting for signature updates.
- **Protects network infrastructure** by slowing or stopping routed traffic from hosts exhibiting high connection rates. The infrastructure will stay up and running, even when it is under attack from a virus.
- **Provides event log and SNMP trap warnings** when worm-like behavior is detected.
- **Gives IT staff time to react** before the problem escalates to a crisis.

² "Implementing and testing a virus throttle" by Jamie Twycross and Matthew M. Williamson, HP Labs, March 3, 2003

- If deployed widely, **makes it difficult for viruses to spread** at all.

How it Works

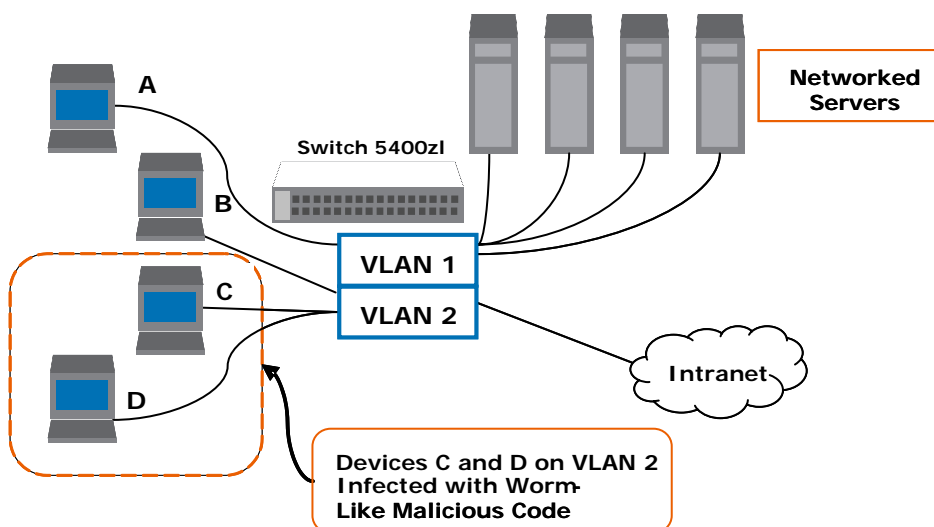
Virus Throttle works by intercepting all IP connection requests—that is, connections in which the source subnet and destination addresses are different. This applies to most common Layer 4-7 session and application protocols, including TCP connections, UDP packets, SMTP, IMAP, Web Proxy, HTTP, SSL and DNS—virtually any protocol where the normal traffic does not look like a virus spreading.

(Some protocols, such as NetBIOS and WINS, are not appropriate for Virus Throttle, because they initiate a broad burst of network traffic that could be misinterpreted by Virus Throttle technology as a threat. Similarly, applications that innocently generate suspicious-looking volumes of short traffic—such as network management scanners, notification services and some p2p file sharing—also are not suitable for Virus Throttle.)

The Virus Throttle tracks the number of recently made connections. If a new, intercepted request is to a destination to which a connection was recently made, the request is processed as normal. If the request is to a destination that has not had a recent connection, the request is processed only if the number of recent connections is below a pre-set threshold. The threshold specifies how many connections are to be allowed over a set amount of time, thereby enforcing a connection-rate limit. If the threshold is exceeded, because requests are coming in at an unusually high rate, it is taken as evidence of a virus. This causes the throttle to stop processing requests and, instead, to notify the system administrator.

In figure 1 below, devices C and D on VLAN 2 are infected and exhibiting a high connection-rate characteristic of virus attacks. Virus Throttle in the ProCurve Switch 5400zl series protects devices A, B and network servers, as well as the intranet, from infected C and D devices.

Figure 1: Throttling virus movement within and across VLANs



Connection-Rate Filtering in ProCurve Switches

In ProCurve's product portfolio, the Virus Throttle technology is implemented through connection-rate filtering feature in the ProCurve Switch 3500yl, 5300xl 5400zl, and 6200yl series. This feature can be enabled on per-port basis on all of these switches. However, for the Switch 5300xl series, Virus Throttle can function only when the switch is in routing mode. For the Switch 3500yl, 5400zl and 6200yl series, there is no such restriction and Virus Throttle can function when the switch is in routing or bridging mode.

Filtering Options

In the default configuration, connection-rate filtering is disabled. When enabled on a port, connection-rate filtering monitors inbound routed traffic for a high rate of connection requests from any given host on the port. If a host appears to exhibit the worm-like behavior of attempting to establish a large number of outbound IP connections (destination

addresses) in a short period of time, the switch responds in one of the following ways, depending on how connection-rate filtering is configured:

- **Notify only of potential attack:** While the apparent attack continues, the switch generates an Event Log notice identifying the offending host source address (SA) and (if a trap receiver is configured on the switch) a similar SNMP trap notice.
- **Notify and reduce spreading:** In this case, the switch temporarily blocks inbound routed traffic from the offending host SA for a “penalty” period and generates an Event Log notice of this action and (if a trap receiver is configured on the switch) a similar SNMP trap notice. When the penalty period expires, the switch re-evaluates the routed traffic from the host and continues to block this traffic if the apparent attack continues. (During the re-evaluation period, routed traffic from the host is allowed.)
- **Block spreading:** This option blocks forwarding of the host’s traffic on the switch. When a block occurs, the switch generates an Event Log notice and (if a trap receiver is configured on the switch) a similar SNMP trap notice. Note that system personnel must explicitly re-enable a host that has been previously blocked.

Sensitivity and Application Options

All the ProCurve switches that support Virus Throttle include a global sensitivity setting that enables adjusting the ability of connection-rate filtering to detect relatively high instances of connection-rate attempts from a given source.

For the most part, normal network traffic is distinct from the traffic exhibited by malicious agents. However, when a legitimate network host generates multiple connections in a short period of time, connection-rate filtering may generate a “false positive” and treat the host as an infected client. Lowering the sensitivity or changing the filter mode may reduce the number of false positives. Conversely, relaxing filtering and sensitivity provisions lowers the switch’s ability to detect worm-generated traffic in the early stages of an attack, and should be carefully investigated and planned to ensure that a risky vulnerability is not created. As an alternative, the system administrator can use connection-rate ACLs (access control lists) or selective enabling to allow high-rate legitimate traffic.

Selective enabling. This option involves applying connection-rate filtering only to ports posing a significant risk of attack. For ports that are reasonably secure from attack, there may be little benefit in configuring them with connection-rate filtering.

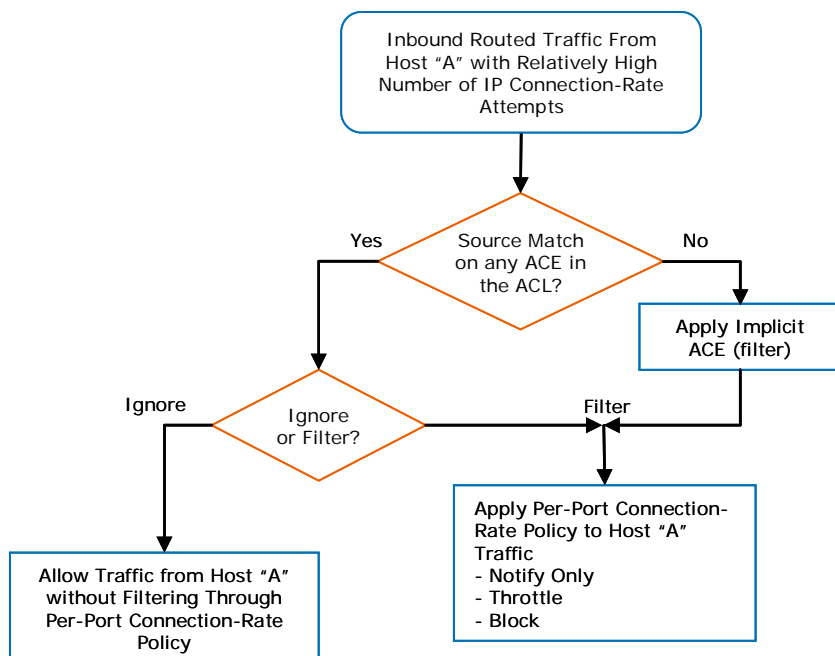
Connection-rate ACLs. As noted above, the basic connection-rate filtering policy is configured per-port as notify-only, throttle and block. A connection-rate ACL, consisting of a series of access control entries (ACEs), creates exceptions to these per-port policies by creating special rules for individual hosts, groups of hosts or entire subnets. Thus, the system administrator can adjust a connection-rate filtering policy to create and apply an exception to configured filters on the ports in a VLAN.

Connection-rate ACLs are useful if the system administrator needs to exclude legitimate high-rate inbound traffic from the connection-rate filtering policy. For example, a server responding to network demand might send a relatively high number of legitimate connection requests. This can generate a false positive by exhibiting the same elevated connection-rate behavior as a worm. Using a connection-rate ACL to apply an exception for this server allows the administrator to exclude the trusted server from connection-rate filtering and thereby keep the server running without interruption.

Figure 2 shows the logical process by which the ProCurve switches that supports Virus Throttle apply their list of ACEs to high-rate inbound traffic to determine whether to ignore it and allow it to pass into the VLAN, or whether to filter it and subject it to connection-rate filtering and notification, throttling or blocking (as pre-set by the administrator). Even traffic that matches an approved ACE is subject to implicit filtering unless specifically excluded, as an added security measure.



Figure 2: Connection-rate ACL applied to traffic through a given port



ProCurve Manager Plus

ProCurve Manager Plus is ProCurve's premier "command from the center" network management software, a key component of the ProCurve Networking Adaptive EDGE Architecture™. Administrators can use ProCurve Manager Plus software to receive alerts from the ProCurve switch's connection-rate filter, as well as to implement decisions to shut down ports in response to detected threats. The integration of Virus Throttle support in ProCurve Manager Plus enables the software to continue to be the all-in-one command center for network management even as the administrator adds Virus Throttle technology to the enterprise's security arsenal. ProCurve Manager Plus also provides greater Virus Throttle response than is possible from the ProCurve switches alone, including the capability to shut down the offending port.

Configuration Guidelines

For the ProCurve switches to apply connection-rate filters, IP routing and multiple VLANs with member ports must first be configured. System administrators can take one approach for networks that are relatively attack-free, and another for high-risk networks. The following summarizes the steps that an administrator takes to configure the switch for connection-rate filtering.

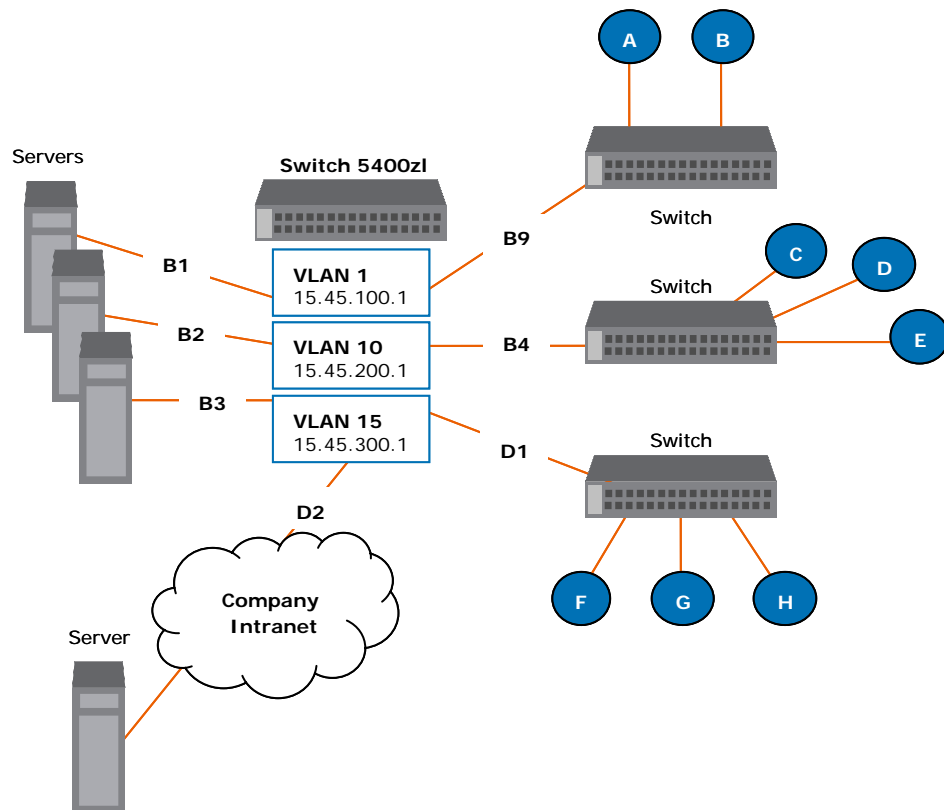
For a network that is relatively attack-free

In cases where a network is relatively attack-free, the network administrator can set global sensitivity on the ProCurve switch's connection filter to low. By monitoring event logs and SNMP trap receivers, if available, the administrator can identify hosts that show high connection rates. If the high rates are the result of legitimate activity—such as heavily used servers—then the administrator can configure connection-rate ACLs to create policy exceptions for trusted hosts. At this point, the sensitivity of the connection-rate filter can be raised to medium and the network monitored again.

For a network that appears to be under significant risk of attack

For a network under likely threat of attack, connection-rate filtering steps include policies for managing the hosts that exhibit high connection rates. This allows better network performance for unaffected hosts and helps to identify hosts that may require updates or patches to eliminate malicious code. For example, the administrator is advised to set connection-rate filtering to "throttle" on all ports, with global sensitivity set to medium. Event log and SNMP trap monitoring should be conducted as above to identify hosts with high connection rates. To immediately halt an attack from a specific host, group of hosts or a subnet, the administrator should use the per-port block mode on the appropriate port or ports. After gaining control of the situation, the administrator can use connection-rate ACLs to manage traffic more selectively to allow receipt of normal routed traffic from reliable hosts.

Figure 3: Basic network configuration



For example, in Figure 3, above, the administrator could choose to

- Throttle potentially malicious high-rate traffic from ports B1-3.
- Notify-only in response to high-rate traffic from the more secure sources C, D and E, connected to B4.
- Immediately block high-rate traffic from potentially high-risk locations such as the company intranet, entering the VLAN via port D2.
- Use an ACL to allow known, legitimate high-rate traffic originating at sources F, G and H to pass into the VLAN at port D1.

ICMP Rate Limiting in ProCurve Switches

What is Rate Limiting

Network threats do not always come in the form of viruses or worms. For example, another dangerous breed of threat—called “Denial-of-Service (DoS)” —is not a virus but rather a method used by attackers to prevent (deny) legitimate users access to network or hosts.

One of the effective means of executing a DoS attack is by misusing ICMP traffic. In IP networks, ICMP messages are generated in response to either inquiries or requests from routing and diagnostic functions. These messages are directed to the applications originating the inquiries. In unusual situations, if the messages are generated rapidly with the intent of overloading network circuits, they can threaten network availability, causing a denial of service.

This problem is visible in DoS attacks like the Smurf Attack, where an attacker sends a large number of ping packets (ICMP echo requests) to IP broadcast addresses, all of them having a spoofed source address of a victim. When all hosts on the broadcast domain respond to the ping with ICMP echo replies directed to the victim host, a Denial-of-Service might result.

ICMP traffic can also be misused by viruses and worms as the initial step in reconnaissance attempts to discover the live hosts in a target network. W32.Welchia.Worm³ is a classic example of a virus using ICMP echo requests to find an active host to infect.

In the ProCurve Switch 3500yl, 5300xl, 5400zl and 6200yl series, the amount of bandwidth that can be utilized for inbound ICMP traffic can be controlled with the ICMP rate limiting feature. This feature allows users to restrict ICMP traffic to levels that permit necessary functions, while throttling excessive ICMP traffic that might be due to ICMP-based DoS attacks or worms or viruses—thus reducing their effectiveness and the speed with which they spread. In addition, this preserves inbound port bandwidth for non-ICMP traffic.

Effect of ICMP Rate Limiting

ICMP rate limiting allows only a specified percentage of a port’s inbound bandwidth to be used for ICMP traffic. As a result, inbound bandwidth is preserved for non-ICMP traffic, and the port or trunk throttles any sudden flood of inbound ICMP traffic generated by a worm or virus attack (or any other cause). Notice that ICMP rate limiting does not throttle non-ICMP traffic. In cases where you want to throttle both ICMP traffic and all other inbound traffic on a given port, you can configure both ICMP rate limiting and all-traffic rate limiting.

ICMP Rate Limiting Operation

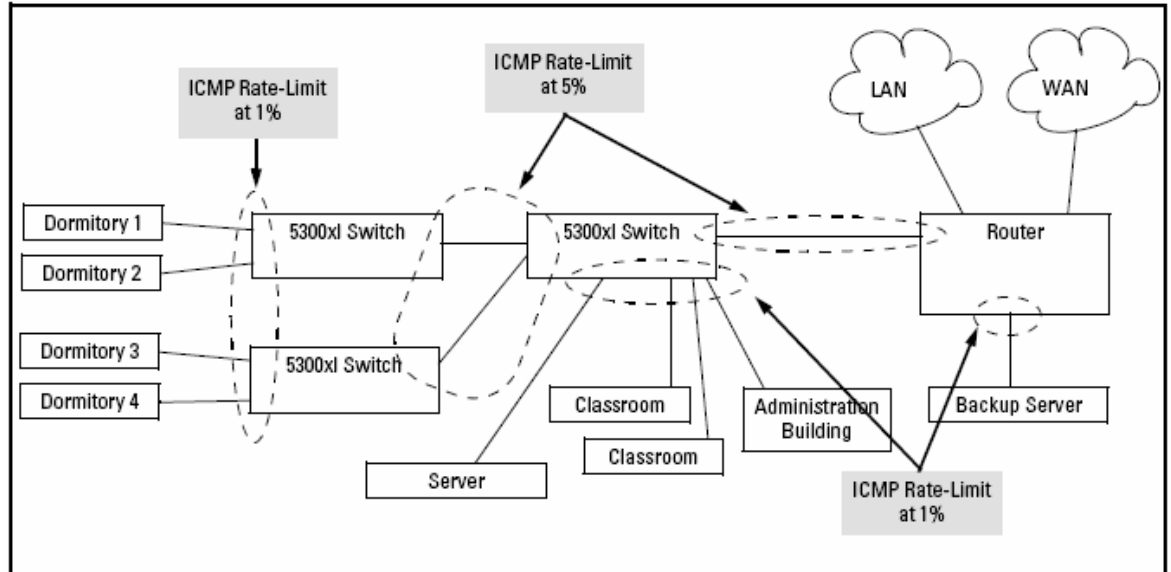
ICMP rate limiting operates on an interface (per-port or per-trunk) basis, and it should be configured to allow the highest expected amount of legitimate inbound ICMP traffic. If an interface experiences an inbound flow of ICMP traffic in excess of its configured limit, the switch throttles that traffic and generates a log messages and an SNMP trap (if an SNMP trap receiver is configured). For example, if a 100 Mbps port negotiates a link to a switch at 100 Mbps and is ICMP rate-limit configured at 5%, then the inbound ICMP traffic flow through that port is limited to 5 Mbps and any excess ICMP traffic is throttled.

Network Application

Apply ICMP rate limiting on all connected ports on the switch to effectively throttle excessive ICMP messaging from any source. On edge ports, where ICMP traffic should be minimal, a threshold of 1% of available bandwidth should be sufficient for most applications. On core ports, such as switch-to-switch and switch-to-router, a maximum threshold of 5% should be sufficient for normal ICMP traffic. (“Normal” ICMP traffic levels should be the maximums that occur when the network is rebooting.)

³ <http://securityresponse.symantec.com/avcenter/venc/data/w32.welchia.worm.html>

Figure 4: ICMP rate limiting



Summary

Traditional methods of addressing viruses, worms and other malicious code depend on signatures and patches. That leaves systems vulnerable to previously unknown threats until protective code can be written and deployed. At a time when viruses spread more quickly than ever before, often generating paralyzing amounts of network traffic, this is a significant lapse.

In contrast, Virus Throttle technology developed by HP Labs focuses on the behavior of malicious code, rather than on its content, so it can identify and respond even to previously unknown threats. By focusing on blocking such code from spreading across VLANs, it addresses the immense network traffic that is one of the most destructive aspects of such code. Virus Throttle is highly effective, stopping the spread of the W32/Nimda-D virus in tests, for example, in less than one second.

Connection-rate filtering based on Virus Throttle technology is now implemented in the ProCurve Switch 3500yl, 5300xl, 5400zl and 6200yl series. System administrators can configure this technology to throttle (slow) or completely block suspect traffic, or merely to notify administrators of potential threats. They can selectively apply ACLs that allow known, legitimate traffic that shares the high-rate profile of malicious code to pass through the infrastructure without problem. In all cases, the administrator can configure the technology to leave human agents in control of decisions about which traffic to block and which to pass. By slowing or blocking suspect traffic until administrators have time to act, connection-rate filtering adds a crucial tool to the defenses of today's networks.

To find out more about
ProCurve Networking
products and solutions,
visit our web site at

www.hp.com/go/procurve



© 2006 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

4AA0-0662ENW, 12/2006