IPv6 Operations	j. woodyatt, Ed.	
Internet-Draft	Apple	
Intended status: BCP	July 29, 2008	
Expires: January 30, 2009		

TOC

Recommended Simple Security Capabilities in Customer Premises Equipment for Providing Residential IPv6 Internet Service
draft-ietf-v6ops-cpe-simple-security-03

Status of this Memo

Abstract

This document makes specific recommendations to the makers of devices that provide "simple security" capabilities at the perimeter of local-area IPv6 networks in Internet-enabled homes and small offices.

Table of Contents

1. Introduction
    1.1. Special Language
2. Overview
    2.1. Basic Sanitation
    2.2. Internet Layer Protocols
    2.3. Transport Layer Protocols
3. Detailed Recommendations
    3.1. Stateless Filters

3.2. Connection-free Filters
3.2.1. Upper-layer Transport Protocols
3.2.2. UDP Filters
3.2.3. Teredo-specific Filters
3.2.4. IPsec and Internet Key Exchange (IKE)
3.2.5. Other Virtual Private Network Protocols
3.3. Connection-oriented Filters
3.3.1. TCP Filters
3.3.2. SCTP Filters
3.3.3. DCCP Filters
3.4. Passive Listeners
4. Summary of Recommendations
5. Contributors
6. IANA Considerations
7. Security Considerations
8. References
8.1. Normative References
8.2. Informative References
Appendix A. Change Log
A.1. draft-ietf-v6ops-cpe-simple-security-00 to draft-ietf-v6ops-cpe-simple-security-01
A.2. draft-ietf-v6ops-cpe-simple-security-01 to draft-ietf-v6ops-cpe-simple-security-02
A.3. draft-ietf-v6ops-cpe-simple-security-02 to draft-ietf-v6ops-cpe-simple-security-03
§ Author's Address
§ Intellectual Property and Copyright Statements

---

1. Introduction

TOC

In "Local Network Protection for IPv6" [RFC4864] (Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, “Local Network Protection for IPv6,” May 2007.), IETF recommends 'simple security' capabilities for gateway devices that enable delivery of Internet services in residential and small office settings. The principle goal of these capabilties is to improve security of the IPv6 Internet without increasing the perceived complexity for users who just want to accomplish useful work.

There is, at best, a constructive tension between the desires of users for transparent end-to-end connectivity on the one hand, and the need for local-area network administrators to detect and prevent intrusion by unauthorized public Internet users on the other. The specific recommendations in this document are intended to promote optimal local-area network security while retaining full end-to-end transparency for

users, and to highlight reasonable limitations on transparency where security considerations are deemed important.
Residential and small office network administrators are expected to have no expertise in Internet engineering whatsoever. Configuration interfaces for simple security in router/gateway appliances marketed toward them should be easy to understand and even easier to ignore. In particular, extra care should be taken in designing the baseline operating modes of unconfigured devices, since the security functions of most devices will never be changed from their factory set default.

1.1. Special Language

TOC

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 (Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," March 1997.) [RFC2119].
The key word "DEFAULT" in this document is to be interpreted as the configuration of a device, as applied by its vendor, prior to the operator changing it for the first time.

---

2. Overview

TOC

For the purposes of this document, residential Internet gateways are assumed to be fairly simple devices with a limited subset of the full range of possible features. They function as default routers [RFC4294] (Loughney, J., "IPv6 Node Requirements," April 2006.) for a single local-area network segment, e.g. an ethernet, a Wi-Fi network, a bridge between two or more such segments. They have a single interface by which they connect to the public Internet, and they can obtain service by any combination of sub-IP mechanisms, including tunnels and transition mechanisms. In referring to their security capabilities, it is reasonable to distinguish between the "interior" network, i.e. the local-area network, and the "exterior" network, i.e. the public Internet. This document is concerned with the behavior of packet filters that police the flow of traffic between the interior and exterior networks of residential Internet gateways.
The operational goals of security capabilities in Internet gateways are described with more detail in "Local Network Protection for IPv6" [RFC4864] (Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6," May 2007.), but they can be summarized as follows.

*Check all traffic to and from the public Internet for basic sanity, e.g. anti-spoofing and "martian" filters.

*Allow tracking of application usage by source and destination transport addresses.

*Provide a barrier against untrusted external influences on the interior network by requiring filter state to be activated by traffic originating at interior network nodes.

*Allow manually configured exceptions to the stateful filtering rules according to network administration policy.

*Isolate local network DHCP and DNS proxy resolver services from the public Internet.

Prior to the widespread availability of IPv6 Internet service, homes and small offices often used private IPv4 network address realms [RFC1918] (Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, “Address Allocation for Private Internets,” February 1996.) with Network Address Translation (NAT) functions deployed to present all the hosts on the interior network as a single host to the Internet service provider. The stateful packet filtering behavior of NAT set user expectations that persist today with residential IPv6 service. "Local Network Protection for IPv6" [RFC4864] (Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, “Local Network Protection for IPv6,” May 2007.) recommends applying stateful packet filtering at residential IPv6 gateways that conforms to the user expectations already in place.

It should be noted that NAT for IPv6 is both strictly forbidden by the standards documents and strongly deprecated by Internet operators. Only the perceived security benefits associated with stateful packet filtering, which NAT requires as a side effect, are thought relevant in the IPv6 residential usage scenario.

As the latest revision of this document is being drafted, conventional stateful packet filters are activated as a side effect of outbound flow initiations from interior network nodes. This requires applications to have advance knowledge of the addresses of exterior nodes with which they expect to communicate. Several proposals are currently under consideration for allowing applications to solicit inbound traffic from exterior nodes without advance knowledge of their addresses. While consensus within the Internet engineering community has emerged that such protocols are necessary to implement in residential IPv6 gateways, the best current practice has not yet been established.

TOC

2.1. Basic Sanitation

In addition to the functions required of all Internet routers [RFC4294] (Loughney, J., “IPv6 Node Requirements,” April 2006.), residential gateways are expected to have basic stateless filters for prohibiting certains kinds of traffic with invalid headers, e.g. martian packets, spoofs, routing header type code zero, etc.
Internet gateways that route multicast traffic are expected to implement appropriate filters for scoped multicast addresses.
Conversely, simple Internet gateways are not expected to prohibit the development of new applications. In particular, packets with end-to-end network security and routing extension headers for mobility are expected to pass Internet gateways freely.

2.2. Internet Layer Protocols

TOC

In managed, enterprise networks, virtual private networking tunnels are typically regarded as an additional attack surface. and they are often restricted or prohibited from traversing firewalls for that reason. However, it would be inappropriate to restrict virtual private networking tunnels by default in unmanaged, residential network usage scenarios. Therefore, this document recommends the DEFAULT operating mode for residential IPv6 simple security is to permit all virtual private networking tunnel protocols to pass through the stateful filtering function. These include IPsec transport and tunnel modes as well as other IP-in-IP protocols.
Where IPv6 simple security functions are integrated with an IPv4/NAT gateway of any of the types described in [RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.), it's important to keep IPv6 flows subject to a consistent policy. If the security functions of an IPv6 residential gateway can be bypassed through Teredo (Huitema, C., “Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs),” February 2006.) [RFC4380], then application developers will be encouraged to use it even at nodes where native IPv6 service is available. This will have the effect of impeding the completion of the transition to native IPv6.
Residential IPv6 gateways are expected to continue operating as IPv4/NAT gateways for the foreseeable future. To prevent Teredo from acquiring a utility that it was never meant to have on networks where both IPv4/NAT and native IPv6 services are available, gateways MUST impede Teredo tunnels by blocking clients from learning their mapped addresses and ports in the qualification procedure described in sections 5.2.1 and 5.2.2 of [RFC4380] (Huitema, C., “Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs),” February 2006.). (Note: this is a necessary addition to the "automatic

sunset" provision in section 5.5 of [RFC4380] (Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)," February 2006.) because it's all too common that nested IPv4/NAT gateways are deployed unintentionally in residential settings and without consideration for Internet architectural implications.)

2.3. Transport Layer Protocols

TOC

IPv6 simple security functions are principally concerned with the stateful filtering of transport layers like User Datagram Protocol (UDP) (Postel, J., "User Datagram Protocol," August 1980.) [RFC0768] (and Lightweight User Datagram Protocol (UDP-Lite) (Larzon, L-A., Degermark, M., Pink, S., Jonsson, L-E., and G. Fairhurst, "The Lightweight User Datagram Protocol (UDP-Lite)," July 2004.) [RFC3828]), Transport Control Protocol (TCP) (Postel, J., "Transmission Control Protocol," September 1981.) [RFC0793], the Stream Control Transmission Protocol (SCTP) (Stewart, R., "Stream Control Transmission Protocol," September 2007.) [RFC4960], the Datagram Congestion Control Protocol (DCCP) (Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)," March 2006.) [RFC4340], and potentially any standards-track transport protocols to be defined in the future.
The general operating principle is that transport layer traffic is only permitted into the interior network of a residential IPv6 gateway when it has been solicited explicitly by interior nodes. All other traffic is expected to be discarded or rejected with an ICMPv6 error message to indicate the traffic is administratively prohibited.

---

3. Detailed Recommendations

TOC

This section describes the specific recommendations made by this document in full detail. They are summarized into a convenient list in Section 4 (Summary of Recommendations).
Some recommended filters are to be applied to all traffic that passes through residential Internet gateways regardless of the direction they are to be forwarded. However, most filters are expected to be sensitive to the direction that traffic is flowing. Packets are said to be "outbound" if they originate from interior nodes to be forwarded to the Internet, and "inbound" if they originate from exterior nodes to be forwarded to any node or nodes on the interior prefix. Flows, as opposed to packets, are said to be "outbound" if the initiator is an interior node and one or more of the participants are at exterior addresses. Flows are said to be "inbound" if the initiator is an exterior node and one or more of the participants are nodes on the interior network. The initiator of a flow is the first node to send a

packet in the context of a given transport association, e.g. a TCP connection, et cetera.

3.1. Stateless Filters

TOC

Certain kinds of IPv6 packets MUST NOT be forwarded in either direction by residential Internet gateways regardless of network state. These include packets with multicast source addresses, packets to destinations with certain non-routable and/or reserved prefixes, and packets with deprecated extension headers.
Other stateless filters are recommended to guard against spoofing, to enforce multicast scope boundaries, and to isolate certain local network services from the public Internet.
R1: Packets bearing in their outer IPv6 headers multicast source addresses MUST NOT be forwarded or transmitted on any interface.
R2: Packets bearing in their outer IPv6 headers multicast destination addresses of equal or narrower scope that the configured scope boundary level of the gateway MUST NOT be forwarded in any direction. The DEFAULT scope boundary level SHOULD be organization-local scope.
R3: Packets bearing deprecated extension headers prior to their first upper-layer-protocol header MUST NOT be forwarded or transmitted on any interface. In particular, all packets with routing extension header type 0 [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.) preceding the first upper-layer-protocol header MUST NOT be forwarded.
R4: Outbound packets MUST NOT be forwarded if the source address in their outer IPv6 header does not have a unicast prefix assigned for use by globally reachable nodes on the interior network.
R4: Inbound packets MUST NOT be forwarded if the source address in their outer IPv6 header has a global unicast prefix assigned for use by globally reachable nodes on the interior network.
R5: Packets MAY be discarded if the source and/or destination address in the outer IPv6 header is a unique local address. By DEFAULT, gateways SHOULD NOT forward packets across unique local address scope boundaries.
R6: By DEFAULT, inbound non-recursive DNS queries received on exterior interfaces MUST NOT be processed by any integrated DNS proxy resolving server.
R7: Inbound DHCP discovery packets received on exterior interfaces MUST NOT be processed by any integrated DHCP server.

TOC

3.2. Connection-free Filters

Some Internet applications use connection-free transport protocols with no release semantics, e.g. UDP. These protocols pose a special difficulty for stateful packet filters because most of the application state is not carried at the transport level. State records are created when communication is initiated and abandoned when no further communication is detected after some period of time.

3.2.1. Upper-layer Transport Protocols

TOC

Residential IPv6 gateways are not expected to prohibit the use of applications to be developed using future upper-layer transport protocols. In particular, transport protocols not otherwise discussed in subsequent sections of this document are expected to be treated consistently, i.e. as having connection-free semantics and no special requirements to inspect the transport headers.
In general, upper-layer transport filter state records are expected to be created when an interior endpoint sends a packet to an exterior address. The filter allocates (or reuses) a record for the duration of communications, with an idle timer to delete the state record when no further communications are detected.
R9: Filter state records for generic upper-layer transport protocols MUST BE indexable by a 3-tuple comprising the interior node address, the exterior node address and the upper-layer transport protocol identifier.
R10: Filter state records for generic upper-layer transport protocols MUST NOT be deleted or recycled until an idle timer not less than two minutes has expired without having forwarded a packet matching the state in some configurable amount of time. By DEFAULT, the idle timer for such state records is five minutes.

3.2.2. UDP Filters

TOC

"NAT Behaviorial Requirements for UDP" (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.) [RFC4787] defines the terminology and best current practice for stateful filtering of UDP applications in IPv4 with NAT, which serves as the model for behaviorial requirements for simple UDP security in IPv6 gateways, notwithstanding the requirements related specifically to network address translation.
An interior endpoint initiates a UDP exchange through a stateful packet filter by sending a packet to an exterior address. The filter allocates (or reuses) a filter state record for the duration of the exchange. The

state record defines the interior and exterior IP addresses and ports used between all packets in the exchange.
State records for UDP exchanges remain active while they are in use and only abandoned after an idle period of some time.
R11: A state record for a UDP exchange where both interior and exterior ports are outside the well-known port range (ports 0-1023) MUST NOT expire in less than two minutes of idle time. The value of the UDP state record idle timer MAY be configurable. The DEFAULT is five minutes.
R12: A state record for a UDP exchange where one or both of the interior and exterior ports are in the well-known port range (ports 0-1023) MAY expire after a period of idle time shorter than two minutes to facilitate the operation of the IANA-registered service assigned to the port in question.
As [RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.) notes, outbound refresh is necessary for allowing the interior endpoint to keep the state record alive. Inbound refresh may be useful for applications with no outbound UDP traffic. However, allowing inbound refresh can allow an attacker in the exterior or a misbehaving application to keep a state record alive indefinitely. This could be a security risk. Also, if the process is repeated with different ports, over time, it could use up all the state record memory and resources in the filter.
R13: A state record for a UDP exchange MUST be refreshed when a packet is forwarded from the interior to the exterior, and it MAY be refreshed when a packet is forwarded in the reverse direction.
As described in section 5.5 of [RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.), the connection-free semantics of UDP pose a difficulty for packet filters in trying to recognize which packets comprise an application flow and which are unsolicited. Various strategies have been used in IPv4/NAT gateways with differing effects.
R14: If application transparency is most important, then a stateful packet filter SHOULD have "Endpoint independent filter" behavior for UDP. If a more stringent filtering behavior is most important, then a filter SHOULD have "Address dependent filtering" behavior. The filtering behavior MAY be an option configurable by the network administrator, and it MAY be independent of the filtering behavior for TCP and other protocols.
Applications mechanisms may depend on the reception of ICMP error messages triggered by the transmission of UDP messages. One such mechanism is path MTU discovery.
R15: If a gateway forwards a UDP exchange, it MUST also forward ICMP Destination Unreachable messages containing UDP headers that match the exchange state record.
R16: Receipt of any sort of ICMP message MUST NOT terminate the state record for a UDP exchange.

R17: UDP-Lite exchanges [RFC3828] (Larzon, L-A., Degermark, M., Pink, S., Jonsson, L-E., and G. Fairhurst, “The Lightweight User Datagram Protocol (UDP-Lite),” July 2004.) SHOULD be handled in the same way as UDP exchanges, except that the upper-layer transport protocol identifier for UDP-Lite is not the same as UDP, and therefore UDP packets MUST NOT match UDP-Lite state records, and vice versa.

3.2.3. Teredo-specific Filters

TOC

Transitional residential IPv6 gateways that also feature integrated IPv4/NAT gateways require special filtering for Teredo tunnels.
R18: Where an IPv6 prefix is advertised on an interior interface alongside an IPv4 private address [RFC1918] (Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, “Address Allocation for Private Internets,” February 1996.) and IPv4 Internet service is provided with NAT [RFC4787] (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.), the Teredo qualification procedure (see section 5.2.1 and 5.2.2 of [RFC4380] (Huitema, C., “Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs),” February 2006.)) for clients in the interior MUST be prohibited by the IPv4/NAT stateful filter. This SHOULD be done by blocking outbound UDP initiations to port 3544, the port reserved by IANA for Teredo servers. This MAY be done by discarding Teredo packets identified by the heuristic defined in "Teredo Security Concerns Beyond What Is In RFC 4380" (Hoagland, J. and S. Krishnan, “Teredo Security Concerns Beyond What Is In RFC 4380,” July 2007.) [HOAGLAND].
[EDITOR: In the event [HOAGLAND] (Hoagland, J. and S. Krishnan, “Teredo Security Concerns Beyond What Is In RFC 4380,” July 2007.) does not advance to publication as an RFC, then that heuristic will be reproduced here.]

3.2.4. IPsec and Internet Key Exchange (IKE)

TOC

Internet protocol security (IPsec) offers greater flexibility and better overall security than the simple security of stateful packet filtering at network perimeters. Therefore, residential IPv6 gateways need not prohibit IPsec traffic flows.
R19: In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of packets, to and from legitimate node addresses, with destination extension headers of type "Authenticated Header (AH)" (Kent, S., “IP Authentication Header,” December 2005.) [RFC4302] in their outer IP extension header chain.

R20: In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of packets, to and from legitimate node addresses, with an upper layer protocol of type "Encapsulating Security Payload (ESP)" (Kent, S., "IP Encapsulating Security Payload (ESP)," December 2005.) [RFC4303] in their outer IP extension header chain.
R21: In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of any UDP packets, to and from legitimate node addresses, with a destination port of 500, i.e. the port reserved by IANA for the Internet Key Exchange Protocol (Kaufman, C., "Internet Key Exchange (IKEv2) Protocol," December 2005.) [RFC4306].
R22: In all operating modes, IPv6 gateways SHOULD use filter state records for Encapsulating Security Payload (ESP) (Kent, S., "IP Encapsulating Security Payload (ESP)," December 2005.) [RFC4303] that are indexable by a 3-tuple comprising the interior node address, the exterior node address and the ESP protocol identifier. In particular, the IPv4/NAT method of indexing state records also by security parameters index (SPI) SHOULD NOT be used. Likewise, any mechanism that depends on detection of Internet Key Exchange (IKE) (Kaufman, C., "Internet Key Exchange (IKEv2) Protocol," December 2005.) [RFC4306] initiations SHOULD NOT be used.

3.2.5. Other Virtual Private Network Protocols

TOC

Residential IPv6 gateways are not expected to prohibit the use of virtual private networks in residential usage scenarios.
R23: In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding, to and from legitimate node addresses, with upper layer protocol of type IP version 6, and SHOULD NOT prohibit the forwarding of other tunneled networking protocols commonly used for virtual private networking, e.g. IP version 4, Generic Routing Encapsulation, etcetera.

3.3. Connection-oriented Filters

TOC

Most Internet applications use connection-oriented transport protocols with orderly release semantics. These protocols include the Transport Control Protocol (TCP) [RFC0793] (Postel, J., "Transmission Control Protocol," September 1981.), the Stream Control Transmission Protocol (SCTP) [RFC4960] (Stewart, R., "Stream Control Transmission Protocol," September 2007.), the Datagram Congestion Control Protocol (DCCP) [RFC4340] (Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)," March 2006.), and potentially any future IETF standards-track transport protocols that use such semantics. Stateful packet filters track the state of individual transport connections and

prohibit the forwarding of packets that do not match the state of an active connection and do not conform to a rule for the automatic creation of such state.

3.3.1. TCP Filters

TOC

An interior endpoint initiates a TCP connection through a stateful packet filter by sending a SYN packet. The filter allocates (or reuses) a filter state record for the connection. The state record defines the interior and exterior IP addresses and ports used for forwarding all packets for that connection.

Peer-to-peer applications use an alternate method of connection initiation termed simultaneous-open (Fig. 8, [RFC0793] (Postel, J., "Transmission Control Protocol," September 1981.)) to traverse stateful filters. In the simultaneous-open mode of operation, both peers send SYN packets for the same TCP connection. The SYN packets cross in the network. Upon receiving the other end's SYN packet, each end responds with a SYN-ACK packet, which also cross in the network. The connection is established at each endpoint once the SYN-ACK packets are received.

To provide stateful packet filtering service for TCP, it is necessary for a filter to receive, process and forward all packets for a connection that conform to valid transitions of the TCP state machine (Fig. 6, [RFC0793] (Postel, J., "Transmission Control Protocol," September 1981.)).

R24: All valid sequences of TCP packets (defined in [RFC0793] (Postel, J., "Transmission Control Protocol," September 1981.)) MUST be forwarded for outbound connections and explicitly permitted inbound connections. In particular, both the normal TCP 3-way handshake mode of operation and the simultaneous-open modes of operation MUST be supported.

It is possible to reconstruct enough of the state of a TCP connection to allow forwarding between an interior and exterior node even when the filter starts operating after TCP enters the established state. In this case, because the filter has not seen the TCP window-scale option, it is not possible for the filter to enforce the TCP window invariant by dropping out-of-window segments.

R25: The TCP window invariant MUST NOT be enforced on connections for which the filter did not detect whether the window-scale option (see [RFC1323] (Jacobson, V., Braden, B., and D. Borman, "TCP Extensions for High Performance," May 1992.)) was sent in the 3-way handshake or simultaneous open.

A stateful filter can allow an existing state record to be reused by an externally initiated connection if its security policy permits. Several different policies are possible as described in "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP (Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral

Requirements for Unicast UDP," January 2007.) [RFC4787] and extended in "NAT Behaviorial Requirements for TCP" (Guha, S., Biswas, K., Sivakumar, S., Ford, B., and P. Srisuresh, "NAT Behavioral Requirements for TCP," April 2007.) [BEHAVE-TCP].

R26: If application transparency is most important, then a stateful packet filter SHOULD have "Endpoint independent filter" behavior for TCP. If a more stringent filtering behavior is most important, then a filter SHOULD have "Address dependent filtering" behavior. The filtering behavior MAY be an option configurable by the network administrator, and it MAY be independent of the filtering behavior for UDP and other protocols.

If an inbound SYN packet is filtered, either because a corresponding state record does not exist or because of the filter's normal behavior, a filter has two basic choices: to discard the packet silently, or to signal an error to the sender. Signaling an error through ICMP messages allows the sender to detect that the SYN did not reach the intended destination. Discarding the packet, on the other hand, allows applications to perform simultaneous-open more reliably. A more detailed discussion of this issue can be found in [BEHAVE-TCP] (Guha, S., Biswas, K., Sivakumar, S., Ford, B., and P. Srisuresh, "NAT Behavioral Requirements for TCP," April 2007.), but the basic outcome of it is that filters need to wait on signaling errors until simultaneous-open will not be impaired.

R27: A gateway MUST NOT signal an error for an unsolicited inbound SYN packet for at least 6 seconds after the packet is received. If during this interval the gateway receives and forwards an outbound SYN for the connection, then the gateway MUST discard the original unsolicited inbound SYN packet without signaling an error. Otherwise, the gateway SHOULD send an ICMP Destination Unreachable error, code 1 (administratively prohibited) for the original SYN-- unless sending any response violates the security policy of the network administrator.

A TCP filter maintains state associated with in-progress and established connections. Because of this, a filter is susceptible to a resource-exhaustion attack whereby an attacker (or virus) on the interior attempts to cause the filter to exhaust its capacity for creating state records. To defend against such attacks, a filter needs to abandon unused state records after a sufficiently long period of idleness.

A common method used for TCP filters in IPv4/NAT gateways is to abandon preferentially sessions for crashed endpoints, followed by closed TCP connections and partially-open connections. A gateway can check if an endpoint for a session has crashed by sending a TCP keep-alive packet on behalf of the other endpoint and receiving a TCP RST packet in response. If the gateway connot determine whether the endpoint is active, then the associated state record needs to be retained until the TCP connection has been idle for some time. Note: an established TCP connection can stay idle (but live) indefinitely; hence, there is no fixed value for an idle-timeout that accommodates all applications. However, a large idle-timeout motivated by recommendations in [RFC1122]

(Braden, R., “Requirements for Internet Hosts - Communication Layers,” October 1989.) and [RFC4294] (Loughney, J., “IPv6 Node Requirements,” April 2006.) can reduce the chances of abandoning a live connection. TCP connections can stay in the established phase indefinitely without exchanging packets. Some end-hosts can be configured to send keep-alive packets on such idle connections; by default, such packets are sent every two hours, if enabled [RFC1122] (Braden, R., “Requirements for Internet Hosts - Communication Layers,” October 1989.). Consequently, a filter that waits for slightly over two hours can detect idle connections with keep-alive packets being sent at the default rate. TCP connections in the partially-open or closing phases, on the other hand, can stay idle for at most four minutes while waiting for in-flight packets to be delivered [RFC1122] (Braden, R., “Requirements for Internet Hosts - Communication Layers,” October 1989.).

The "established connection idle-timeout" for a stateful packet filter is defined as the minimum time a TCP connection in the established phase must remain idle before the filter considers the associated state record a candidate for collection. The "transitory connection idle-timeout" for a filter is defined as the minimum time a TCP connection in the partially-open or closing phases must remain idle before the filter considers the associated state record a candidate for collection. TCP connections in the TIME_WAIT state are not affected by the "transitory connection idle-timeout" parameter.

R28: If a gateway cannot determine whether the endpoints of a TCP connection are active, then it MAY abandon the state record if it has been idle for some time. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than four minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

Behavior for handing RST packets, or connections in the TIME_WAIT state is left unspecified. A gateway MAY hold state for a connection in TIME_WAIT state to accommodate retransmissions of the last ACK. However, since the TIME_WAIT state is commonly encountered by interior endpoints properly closing the TCP connection, holding state for a closed connection can limit the throughput of connections through a gateway with limited resources. [RFC1337] (Braden, B., “TIME-WAIT Assassination Hazards in TCP,” May 1992.) discusses hazards associated with TIME_WAIT assassination.

The handling of non-SYN packets for which there is no active state record is left unspecified. Such packets can be received if the gateway abandons a live connection, or abandons a connection in the TIME_WAIT state before the four minute TIME_WAIT period expires. The decision either to discard or to respond with an ICMP Destination Unreachable error, code 1 (administratively prohibited) is left up to the implementation.

Behavior for notifying endpoints when abandoning live connections is left unspecified. When a gateway abandons a live connection, for example due to a timeout expiring, the filter MAY send a TCP RST packet

to each endpoint on behalf of the other. Sending a RST notification allows endpoint applications to recover more quickly, however, notifying endpoints might not always be possible if, for example, state records are lost due to power interruption.
Several TCP mechanisms depend on the reception of ICMP error messages triggered by the transmission of TCP segments. One such mechanism is path MTU discovery, which is required for correct operation of TCP.
R29: If a gateway forwards a TCP connection, it MUST also forward ICMP Destination Unreachable messages containing TCP headers that match the connection state record.
R30: Receipt of any sort of ICMP message MUST NOT terminate the state record for a TCP connection.

3.3.2. SCTP Filters

TOC

Because Stream Control Transmission Protocol (SCTP) (Stewart, R., “Stream Control Transmission Protocol,” September 2007.) [RFC4960] connections can be terminated at multiple network addresses, IPv6 simple security functions cannot achieve full transparency for SCTP applications. In multipath traversal scenarios, full transparency requires coordination between all the packet filter processes in the various paths between the endpoint network addresses. Such coordination is not "simple" and it is, therefore, beyond the scope of this recommendation.
However, some SCTP applications are capable of tolerating the inherent unipath restriction of IPv6 simple security, even in multipath traversal scenarios. They expect similar connection-oriented filtering behaviors as for TCP, but at the level of SCTP associations, not stream connections. This section describes specific recommendations for SCTP filtering for such traversal scenarios.
An interior endpoint initiates SCTP associations through a stateful packet filter by sending a packet comprising a single INIT chunk. The filter allocates (or reuses) a filter state record for the association. The state record defines the interior and exterior IP addresses and the observed verification tag used for forwarding packets in that association.
Peer-to-peer applications use an alternate method of association initiation termed simultaneous-open to traverse stateful filters. In the simultaneous-open mode of operation, both peers send INIT chunks at the same time to establish an association. Upon receiving the other end's INIT chunk, each end responds with an INIT-ACK packet, which is expected to traverse the same path in reverse. Because only one SCTP association may exist between any two network addresses, one of the peers in simultaneous-open mode of operation will send an ERROR or ABORT chunk along with the INIT-ACK chunk. The association is

established at each endpoint once an INIT-ACK chunks is receivedÂ at one end without an ERROR or ABORT chunk.
To provide stateful packet filtering service for SCTP, it is necessary for a filter to receive, process and forward all packets for an association that conform to valid transitions of the SCTP state machine (Fig. 3, [RFC4960] (Stewart, R., “Stream Control Transmission Protocol,” September 2007.)).
R31: All valid sequences of SCTP packets (defined in [RFC4960] (Stewart, R., “Stream Control Transmission Protocol,” September 2007.)) MUST be forwarded for outbound associations and explicitly permitted inbound associations. In particular, both the normal SCTP association establishment and simultaneous-open modes of operation MUST be supported.
If an inbound INIT packet is filtered, either because a corresponding state record does not exist or because of the filter's normal behavior, a filter has two basic choices: to discard the packet silently, or to signal an error to the sender. Signaling an error through ICMP messages allows the sender to detect that the INIT packet did not reach the intended destination. Discarding the packet, on the other hand, allows applications to perform simultaneous-open more reliably. Delays in signaling errors can prevent the impairment of simultaneous-open mode of operation.
R32: A gateway MUST NOT signal an error for an unsolicited inbound INIT packet for at least 6 seconds after the packet is received. If during this interval the gateway receives and forwards an outbound INIT packet for the association, the the gateway MUST discard the original unsolicited inbound INIT packet without signaling an error. Otherwise, the gateway SHOULD send an ICMP Destination Unreachable error, code 1 (administratively prohibited) for the original INIT-- unless sending any response violates the security policy of the network administrator.
An SCTP filter maintains state associated with in-progress and established associations. Because of this, a filter is susceptible to a resource-exhaustion attack whereby an attacker (or virus) on the interior attempts to cause the filter to exhaust its capacity for creating state records. To defend against such attacks, a filter needs to abandon unused state records after a sufficiently long period of idleness.
A common method used for TCP filters in IPv4/NAT gateways is to abandon preferentially sessions for crashed endpoints, followed by closed associations and partially opened associations. A similar method is an option for SCTP filters in IPv6 gateways. A gateway can check if an endpoint for an association has crashed by sending HEARTBEAT chunks and looking for the HEARTBEAT ACK response. If the gateway cannot determine whether the endpoint is active, then the associated state records needs to be retained until the SCTP association has been idle for some time.
Note: an established SCTP association can stay idle (but live) indefinitely, hence there is no fixed value of an idle-timeout that accommodates all applications. However, a large idle-timeout motivated

by [RFC4294] (Loughney, J., “IPv6 Node Requirements,” April 2006.) can reduce the chances of abandoning a live association.

SCTP associations can stay in the ESTABLISHED state indefinitely without exchanging packets. Some end-hosts can be configured to send HEARTBEAT chunks on such idle associations, but [RFC4960] (Stewart, R., “Stream Control Transmission Protocol,” September 2007.) does not specify (or even suggest) a default time interval. A filter that waits for slightly over two hours can detect idle associations with HEARTBEAT packets being sent at the same rate as most hosts use for TCP keep-alive, which is a reasonably similar system for this purpose. SCTP associations in the partially-open or closing states, on the other hand, can stay idle for at most four minutes while waiting for in-flight packets to be delivered (assuming the suggested SCTP protocol parameter values in Section 15 of [RFC4960] (Stewart, R., “Stream Control Transmission Protocol,” September 2007.)).

The "established association idle-timeout" for a stateful packet filter is defined as the minimum time an SCTP association in the established phase must remain idle before the filter considers the corresponding state record a candidate for collection. The "transitory association idle-timeout" for a filter is defined as the minimum time an SCTP association in the partially-open or closing phases must remain idle before the filter considers the corresponding state record a candidate for collection.

R33: If a gateway cannot determine whether the endpoints of an SCTP association are active, then it MAY abandon the state record if it has been idle for some time. In such cases, the value of the "established association idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory association idle-timeout" MUST NOT be less than four minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

Behavior for handling ERROR and ABORT packets is left unspecified. A gateway MAY hold state for an association after its closing phases have completed to accommodate retransmissions of its final SHUTDOWN ACK packets. However, holding state for a closed association can limit the throughput of associations traversing a gateway with limited resources. The discussion in [RFC1337] (Braden, B., “TIME-WAIT Assassination Hazards in TCP,” May 1992.) regarding the hazards of TIME_WAIT assassination are relevant.

The handling of inbound non-INIT packets for which there is no active state record is left unspecified. Such packets can be received if the gateway abandons a live connection, or abandons an association in the closing states before the transitory association idle-timeout expires. The decision either to discard or to respond with an ICMP Destination Unreachable error, code 1 (administratively prohibited) is left to the implementation.

Behavior for notifying endpoints when abandoning live associations is left unspecified. When a gateway abandons a live association, for example due to a timeout expiring, the filter MAY send an ABORT packet to each endpoint on behalf of the other. Sending an ABORT notification

allows endpoint applications to recover more quickly, however, notifying endpoints might not always be possible if, for example, state records are lost due to power interruption.
Several SCTP mechanisms depend on the reception of ICMP error messages triggered by the transmission of SCTP packets.
R34: If a gateway forwards an SCTP association, it MUST also forward ICMP Destination Unreachable messages containing SCTP headers that match the association state record.
R35: Receipt of any sort of ICMP message MUST NOT terminate the state record for an SCTP association.

3.3.3. DCCP Filters

TOC

The connection semantics described in Datagram Congestion Control Protocol (DCCP) (Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” March 2006.) [RFC4340] are very similar to those of TCP. An interior endpoint initiates a DCCP connection through a stateful packet filter by sending a DCCP-Request packet. Simultaneous open is not defined for DCCP.
In order to provide stateful packet filtering service for DCCP, it is necessary for a filter to receive, process and forward all packets for a connection that conform to valid transitions of the DCCP state machine (Section 8, [RFC4340] (Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” March 2006.)).
R36: All valid sequences of DCCP packets (defined in [RFC4340] (Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” March 2006.)) MUST be forwarded for all connections to exterior servers and those connections to interior servers with explicitly permitted service codes.
It is possible to reconstruct enough of the state of a DCCP connection to allow forwarding between an interior and exterior node even when the filter starts operating after DCCP enters the OPEN state. Also, a filter can allow an existing state record to be reused by an externally initiated connection if its security policy permits. As with TCP, several different policies are possible, with a good discussion of the issue involved presented in Network Address Translation (NAT) Behavioral Requirements for Unicast UDP (Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” January 2007.) [RFC4787] and extended in NAT Behaviorial Requirements for TCP (Guha, S., Biswas, K., Sivakumar, S., Ford, B., and P. Srisuresh, “NAT Behavioral Requirements for TCP,” April 2007.) [BEHAVE-TCP].
If an inbound DCCP-Request packet is filtered, either because a corresponding state record does not already exist for it or because of the filter's normal behavior of refusing connections not explicitly permitted, then a filter has two basic choices: to discard the packet

silently, or to signal an error to the sender. Signaling an error through ICMP messages allows the sender to detect that the DCCP-Request did not reach the intended destination. Discarding the packet, on the other hand, only delays the failure to connect and provides no measurable security.

A DCCP filter maintains state associated with in-progress and established connections. Because of this, a filter is susceptible to a resource-exhaustion attack whereby an attacker (or virus) on the interior attempts to cause the filter to exhaust its capacity for creating state records. To prevent such an attack, a filter needs to abandon unused state records after a sufficiently long period of idleness.

A common method used for TCP filters in IPv4/NAT gateways is to abandon preferentially sessions for crashed endpoints, followed by closed TCP connections and partially-open connections. No such method exists for DCCP, and connections can stay in the OPEN phase indefinitely without exchanging packets. Hence, there is no fixed value for an idle-timeout that accommodates all applications. However, a large idle-timeout motivated by [RFC4294] (Loughney, J., “IPv6 Node Requirements,” April 2006.) can reduce the chances of abandoning a live connection. DCCP connections in the partially-open or closing phases can stay idle for at most eight minutes while waiting for in-flight packets to be delivered.

The "open connection idle-timeout" for a stateful packet filter is defined as the minimum time a DCCP connection in the open state must remain idle before the filter considers the associated state record a candidate for collection. The "transitory connection idle-timeout" for a filter is defined as the minimum time a DCCP connection in the partially-open or closing phases must remain idle before the filter considers the associated state record a candidate for collection. DCCP connections in the TIMEWAIT state are not affected by the "transitory connection idle-timeout" parameter.

R37: A gateway MAY abandon a DCCP state record if it has been idle for some time. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than eight minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

Behavior for handing DCCP-Reset packets, or connections in the TIMEWAIT state is left unspecified. A gateway MAY hold state for a connection in TIMEWAIT state to accommodate retransmissions of the last DCCP-Reset. However, since the TIMEWAIT state is commonly encountered by interior endpoints properly closing the DCCP connection, holding state for a closed connection can limit the throughput of connections through a gateway with limited resources. [RFC1337] discusses hazards associated with TIME_WAIT assassination in TCP, and similar hazards exists for DCCP.

The handling of non-SYN packets for which there is no active state record is left unspecified. Such packets can be received if the gateway

abandons a live connection, or abandons a connection in the TIMEWAIT state before the four minute 2MSL period expires. The decision either to discard or to respond with an ICMP Destination Unreachable error, code 1 (administratively prohibited) is left up to the implementation. Behavior for notifying endpoints when abandoning live connections is left unspecified. When a gateway abandons a live connection, for example due to a timeout expiring, the filter MAY send a DCCP-Reset packet to each endpoint on behalf of the other. Sending a DCCP-Reset notification allows endpoint applications to recover more quickly, however, notifying endpoints might not always be possible if, for example, state records are lost due to power interruption.
Several DCCP mechanisms depend on the reception of ICMP error messages triggered by the transmission of DCCP packets. One such mechanism is path MTU discovery, which is required for correct operation.
R38: If a gateway forwards a DCCP connection, it MUST also forward ICMP Destination Unreachable messages containing DCCP headers that match the connection state record.
R39: Receipt of any sort of ICMP message MUST NOT terminate the state record for a DCCP connection.

3.4. Passive Listeners

TOC

Some applications expect to solicit traffic from exterior nodes without any advance knowledge of the exterior address. This requirement is met by IPv4/NAT gateways typically by the use of either [NAT-PMP] (Cheshire, S., Krochmal, M., and K. Sekar, “NAT Port Mapping Protocol (NAT-PMP),” November 2001.) or [UPnP-IGD] (UPnP Forum, “Universal Plug and Play Internet Gateway Device Standardized Gateway Device Protocol,” September 2006.).
One proposal that has been offered as an Internet Draft is the Application Listener Discovery Protocol (Woodyatt, j., “Application Listener Discovery (ALD) for IPv6,” May 2007.) [IPv6-ALD]. It remains to be seen whether the Internet Gateway Device profile of the Universal Plug And Play protocol will be extended for IPv6. Other proposals of note include the Middlebox Communication Protocol (Stiemerling, M., Quittek, J., and T. Taylor, “Middlebox Communications (MIDCOM) Protocol Semantics,” February 2005.) [RFC3989] and the Next Steps in Signaling framework (Hancock, R., Karagiannis, G., Loughney, J., and S. Van den Bosch, “Next Steps in Signaling (NSIS): Framework,” June 2005.) [RFC4080]. No consensus has yet emerged in the Internet engineering community as to which proposal is most appropriate for residential IPv6 usage scenarios.
R31: Gateways MUST implement a protocol to permit applications to solicit inbound traffic without advance knowledge of the addresses of exterior nodes with which they expect to communicate. This protocol MUST have a specification that meets the requirements of [RFC3978]

(Bradner, S., “IETF Rights in Contributions,” March 2005.), [RFC3979] (Bradner, S., “Intellectual Property Rights in IETF Technology,” March 2005.) and [RFC4748] (Bradner, S., “RFC 3978 Update to Recognize the IETF Trust,” October 2006.).

---

4. Summary of Recommendations

TOC

This section collects all of the recommendations made in this document into a convenient list.

R1 Packets bearing in their outer IPv6 headers multicast source addresses MUST NOT be forwarded or transmitted on any interface.

R2 Packets bearing in their outer IPv6 headers multicast destination addresses of equal or narrower scope that the configured scope boundary level of the gateway MUST NOT be forwarded in any direction. The DEFAULT scope boundary level SHOULD be organization-local scope.

R3 Packets bearing deprecated extension headers prior to their first upper-layer-protocol header MUST NOT be forwarded or transmitted on any interface. In particular, all packets with routing extension header type 0 [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.) preceding the first upper-layer-protocol header MUST NOT be forwarded.

R4 Outbound packets MUST NOT be forwarded if the source address in their outer IPv6 header does not have a unicast prefix assigned for use by globally reachable nodes on the interior network.

R4 Inbound packets MUST NOT be forwarded if the source address in their outer IPv6 header has a global unicast prefix assigned for use by globally reachable nodes on the interior network.

R5 Packets MAY be discarded if the source and/or destination address in the outer IPv6 header is a unique local address. By DEFAULT, gateways SHOULD NOT forward packets across unique local address scope boundaries.

R6 By DEFAULT, inbound non-recursive DNS queries received on exterior interfaces MUST NOT be processed by any integrated DNS proxy resolving server.

R7 Inbound DHCP discovery packets received on exterior interfaces MUST NOT be processed by any integrated DHCP server.

R8
Inbound packets not matching any existing filter state record for a permitted transport flow MUST NOT be forwarded to the interior network, and an ICMP Error message of type Administratively Prohibited MUST be sent to the source address.

R9 Filter state records for generic upper-layer transport protocols MUST BE indexable by a 3-tuple comprising the interior node address, the exterior node address and the upper-layer transport protocol identifier.

R10 Filter state records for generic upper-layer transport protocols MUST NOT be deleted or recycled until an idle timer not less than two minutes has expired without having forwarded a packet matching the state in some configurable amount of time. By DEFAULT, the idle timer for such state records is five minutes.

R11 A state record for a UDP exchange where both interior and exterior ports are outside the well-known port range (ports 0-1023) MUST NOT expire in less than two minutes of idle time. The value of the UDP state record idle timer MAY be configurable. The DEFAULT is five minutes.

R12 A state record for a UDP exchange where one or both of the interior and exterior ports are in the well-known port range (ports 0-1023) MAY expire after a period of idle time shorter than two minutes to facilitate the operation of the IANA-registered service assigned to the port in question.

R13 A state record for a UDP exchange MUST be refreshed when a packet is forwarded from the interior to the exterior, and it MAY be refreshed when a packet is forwarded in the reverse direction.

R14 If application transparency is most important, then a stateful packet filter SHOULD have "Endpoint independent filter" behavior for UDP. If a more stringent filtering behavior is most important, then a filter SHOULD have "Address dependent filtering" behavior. The filtering behavior MAY be an option configurable by the network administrator, and it MAY be

independent of the filtering behavior for TCP and other protocols.

R15 If a gateway forwards a UDP exchange, it MUST also forward ICMP Destination Unreachable messages containing UDP headers that match the exchange state record.

R16 Receipt of any sort of ICMP message MUST NOT terminate the state record for a UDP exchange.

R17 UDP-Lite exchanges [RFC3828] (Larzon, L-A., Degermark, M., Pink, S., Jonsson, L-E., and G. Fairhurst, "The Lightweight User Datagram Protocol (UDP-Lite)," July 2004.) SHOULD be handled in the same way as UDP exchanges, except that the upper-layer transport protocol identifier for UDP-Lite is not the same as UDP, and therefore UDP packets MUST NOT match UDP-Lite state records, and vice versa.

R18 Where an IPv6 prefix is advertised on an interior interface alongside an IPv4 private address [RFC1918] (Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets," February 1996.) and IPv4 Internet service is provided with NAT [RFC4787] (Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," January 2007.), the Teredo qualification procedure (see section 5.2.1 and 5.2.2 of [RFC4380] (Huitema, C., "Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs)," February 2006.)) for clients in the interior MUST be prohibited by the IPv4/NAT stateful filter. This SHOULD be done by blocking outbound UDP initiations to port 3544, the port reserved by IANA for Teredo servers. This MAY be done by discarding Teredo packets identified by the heuristic defined in "Teredo Security Concerns Beyond What Is In RFC 4380" (Hoagland, J. and S. Krishnan, "Teredo Security Concerns Beyond What Is In RFC 4380," July 2007.) [HOAGLAND].

R19 In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of packets, to and from legitimate node addresses, with destination extension headers of type "Authenticated Header (AH)" (Kent, S., "IP Authentication Header," December 2005.) [RFC4302] in their outer IP extension header chain.

R20 In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of packets, to and from legitimate node addresses, with an upper layer protocol of type "Encapsulating Security Payload (ESP)" (Kent, S., "IP Encapsulating Security Payload (ESP)," December 2005.) [RFC4303] in their outer IP extension header chain.

R21

In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding of any UDP packets, to and from legitimate node addresses, with a destination port of 500, i.e. the port reserved by IANA for the Internet Key Exchange Protocol (Kaufman, C., “Internet Key Exchange (IKEv2) Protocol,” December 2005.) [RFC4306].

R22 In their DEFAULT operating mode, IPv6 gateways MUST NOT prohibit the forwarding, to and from legitimate node addresses, with upper layer protocol of type IP version 6, and SHOULD NOT prohibit the forwarding of other tunneled networking protocols commonly used for virtual private networking, e.g. IP version 4, Generic Routing Encapsulation, etcetera.

R23 In all operating modes, IPv6 gateways SHOULD use filter state records for Encapsulating Security Payload (ESP) (Kent, S., “IP Encapsulating Security Payload (ESP),” December 2005.) [RFC4303] that are indexable by a 3-tuple comprising the interior node address, the exterior node address and the ESP protocol identifier. In particular, the IPv4/NAT method of indexing state records also by security parameters index (SPI) SHOULD NOT be used. Likewise, any mechanism that depends on detection of Internet Key Exchange (IKE) (Kaufman, C., “Internet Key Exchange (IKEv2) Protocol,” December 2005.) [RFC4306] initiations SHOULD NOT be used.

R24 All valid sequences of TCP packets (defined in [RFC0793] (Postel, J., “Transmission Control Protocol,” September 1981.)) MUST be forwarded for outbound connections and explicitly permitted inbound connections. In particular, both the normal TCP 3-way handshake mode of operation and the simultaneous-open modes of operation MUST be supported.

R25 The TCP window invariant MUST NOT be enforced on connections for which the filter did not detect whether the window-scale option (see [RFC1323] (Jacobson, V., Braden, B., and D. Borman, “TCP Extensions for High Performance,” May 1992.)) was sent in the 3-way handshake or simultaneous open.

R26 If application transparency is most important, then a stateful packet filter SHOULD have "Endpoint independent filter" behavior for TCP. If a more stringent filtering behavior is most important, then a filter SHOULD have "Address dependent filtering" behavior. The filtering behavior MAY be an option configurable by the network administrator, and it MAY be independent of the filtering behavior for UDP and other protocols.

R27

A gateway MUST NOT signal an error for an unsolicited inbound SYN packet for at least 6 seconds after the packet is received. If during this interval the gateway receives and forwards an outbound SYN for the connection, then the gateway MUST discard the original unsolicited inbound SYN packet without signaling an error. Otherwise, the gateway SHOULD send an ICMP Destination Unreachable error, code 1 (administratively prohibited) for the original SYN-- unless sending any response violates the security policy of network administrator.

R28 If a gateway cannot determine whether the endpoints of a TCP connection are active, then it MAY abandon the state record if it has been idle for some time. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than four minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

R29 If a gateway forwards a TCP connection, it MUST also forward ICMP Destination Unreachable messages containing TCP headers that match the connection state record.

R30 Receipt of any sort of ICMP message MUST NOT terminate the state record for a TCP connection.

R31 Gateways MUST implement a protocol to permit applications to solicit inbound traffic without advance knowledge of the addresses of exterior nodes with which they expect to communicate. This protocol MUST have a specification that meets the requirements of [RFC3978] (Bradner, S., “IETF Rights in Contributions,” March 2005.), [RFC3979] (Bradner, S., “Intellectual Property Rights in IETF Technology,” March 2005.) and [RFC4748] (Bradner, S., “RFC 3978 Update to Recognize the IETF Trust,” October 2006.).

R33 All valid sequences of SCTP packets (defined in [RFC4960] (Stewart, R., “Stream Control Transmission Protocol,” September 2007.)) MUST be forwarded for outbound associations and explicitly permitted inbound associations. In particular, both the normal SCTP association establishment and simultaneous-open modes of operation MUST be supported.

R34 A gateway MUST NOT signal an error for an unsolicited inbound INIT packet for at least 6 seconds after the packet is received. If during this interval the gateway receives and forwards an outbound INIT packet for the association, the the gateway MUST discard the original unsolicited inbound INIT packet without signaling an error. Otherwise, the gateway SHOULD send an ICMP

Destination Unreachable error, code 1 (administratively prohibited) for the original INIT-- unless sending any response violates the security policy of the network administrator.

R33 If a gateway cannot determine whether the endpoints of an SCTP association are active, then it MAY abandon the state record if it has been idle for some time. In such cases, the value of the "established association idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory association idle-timeout" MUST NOT be less than four minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

R34 If a gateway forwards an SCTP association, it MUST also forward ICMP Destination Unreachable messages containing SCTP headers that match the association state record.

R35 Receipt of any sort of ICMP message MUST NOT terminate the state record for an SCTP association.

R36 All valid sequences of DCCP packets (defined in [RFC4340] (Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)," March 2006.)) MUST be forwarded for all connections to exterior servers and those connections to interior servers with explicitly permitted service codes.

R37 A gateway MAY abandon a DCCP state record if it has been idle for some time. In such cases, the value of the "established connection idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory connection idle-timeout" MUST NOT be less than eight minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

R38 If a gateway forwards a DCCP connection, it MUST also forward ICMP Destination Unreachable messages containing DCCP headers that match the connection state record.

R39 Receipt of any sort of ICMP message MUST NOT terminate the state record for a DCCP connection.

---

5. Contributors

TOC

Comments and criticisms during the development of this document were received from the following IETF participants:

Fred Baker

Norbert Bollow

Brian Carpenter

Jun-ichiro itojun Hagino

Thomas Herbst

Christian Huitema

Cullen Jennings

Suresh Krishnan

Erik Kline

Kurt Erik Lindqvist

Iljitsch van Beijnum

Yaron Sheffer

Dan Wing

Much of the text describing the detailed requirements for TCP and UDP filtering is derived or transposed from [BEHAVE-TCP] (Guha, S., Biswas, K., Sivakumar, S., Ford, B., and P. Srisuresh, "NAT Behavioral Requirements for TCP," April 2007.) and [RFC4787] (Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP," January 2007.), and some form of attribution here may therefore be appopriate.

---

6. IANA Considerations

TOC

This memo includes no request to IANA.

---

7. Security Considerations

TOC

The IPv6 stateful filtering behavior described in this document is intended to be similar in function to the filtering behavior of commonly use IPv4/NAT gateways, which have been widely sold as a security tool for residential and small-offce/home-office networks. As noted in the security considerations section of [RFC2993] (Hain, T., "Architectural Implications of NAT," November 2000.), the true impact of these tools may be a reduction in security. It may be generally

assumed that the impacts discussed there related to filtering (and not translation) are to be expected with the simple IPv6 security mechanisms described here.

In particular, it's worth noting that stateful filters create the illusion of a security barrier, but without the managed intent of a firewall. Appropriate security mechanisms implemented in the end nodes, in conjunction with the [RFC4864] (Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6," May 2007.) local network protection methods, function without reliance on network layer hacks and transport filters that may change over time. Also, defined security barriers assume that threats originate in the exterior, which may lead to practices that result in applications being fully exposed to interior attack and which therefore make breaches much easier.

Finally, residential gateways that implement simple security functions are a bastion between the interior and the exterior, and therefore are a target of denial of service attacks against the interior network itself by processes designed to consume the resources of the gateway, e.g. a ping or SYN flood. Gateways should employ the same sorts of protection techniques as application servers on the Internet.

8. References

TOC

8.1. Normative References

TOC

[RFC0768]	Postel, J., "User Datagram Protocol," STD 6, RFC 768, August 1980 (TXT).
[RFC0793]	Postel, J., "Transmission Control Protocol," STD 7, RFC 793, September 1981 (TXT).
[RFC1323]	Jacobson, V., Braden, B., and D. Borman, "TCP Extensions for High Performance," RFC 1323, May 1992 (TXT).
[RFC2119]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC2460]	Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," RFC 2460, December 1998 (TXT, HTML, XML).
[RFC3828]	Larzon, L-A., Degermark, M., Pink, S., Jonsson, L-E., and G. Fairhurst, "The Lightweight User Datagram Protocol (UDP-Lite)," RFC 3828, July 2004 (TXT).
[RFC3978]	Bradner, S., "IETF Rights in Contributions," RFC 3978, March 2005 (TXT).
[RFC3979]	

	Bradner, S., “Intellectual Property Rights in IETF Technology,” BCP 79, RFC 3979, March 2005 (TXT).
[RFC4302]	Kent, S., “IP Authentication Header,” RFC 4302, December 2005 (TXT).
[RFC4303]	Kent, S., “IP Encapsulating Security Payload (ESP),” RFC 4303, December 2005 (TXT).
[RFC4306]	Kaufman, C., “Internet Key Exchange (IKEv2) Protocol,” RFC 4306, December 2005 (TXT).
[RFC4340]	Kohler, E., Handley, M., and S. Floyd, “Datagram Congestion Control Protocol (DCCP),” RFC 4340, March 2006 (TXT).
[RFC4380]	Huitema, C., “Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs),” RFC 4380, February 2006 (TXT).
[RFC4748]	Bradner, S., “RFC 3978 Update to Recognize the IETF Trust,” RFC 4748, October 2006 (TXT).
[RFC4787]	Audet, F. and C. Jennings, “Network Address Translation (NAT) Behavioral Requirements for Unicast UDP,” BCP 127, RFC 4787, January 2007 (TXT).
[RFC4960]	Stewart, R., “Stream Control Transmission Protocol,” RFC 4960, September 2007 (TXT).

8.2. Informative References

TOC

[BEHAVE-TCP]	Guha, S., Biswas, K., Sivakumar, S., Ford, B., and P. Srisuresh, “NAT Behavioral Requirements for TCP,” April 2007.
[HOAGLAND]	Hoagland, J. and S. Krishnan, “Teredo Security Concerns Beyond What Is In RFC 4380,” July 2007.
[IPv6-ALD]	Woodyatt, j., “Application Listener Discovery (ALD) for IPv6,” May 2007.
[NAT-PMP]	Cheshire, S., Krochmal, M., and K. Sekar, “NAT Port Mapping Protocol (NAT-PMP),” November 2001.
[RFC1122]	Braden, R., “Requirements for Internet Hosts - Communication Layers,” STD 3, RFC 1122, October 1989 (TXT).
[RFC1337]	Braden, B., “TIME-WAIT Assassination Hazards in TCP,” RFC 1337, May 1992 (TXT).
[RFC1918]	Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, “Address Allocation for Private Internets,” BCP 5, RFC 1918, February 1996 (TXT).
[RFC2993]	Hain, T., “Architectural Implications of NAT,” RFC 2993, November 2000 (TXT).
[RFC3989]	Stiemerling, M., Quittek, J., and T. Taylor, “Middlebox Communications (MIDCOM) Protocol Semantics,” RFC 3989, February 2005 (TXT).

[RFC4080]	Hancock, R., Karagiannis, G., Loughney, J., and S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework," RFC 4080, June 2005 (TXT).
[RFC4294]	Loughney, J., "IPv6 Node Requirements," RFC 4294, April 2006 (TXT).
[RFC4864]	Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6," RFC 4864, May 2007 (TXT).
[UPnP-IGD]	UPnP Forum, "Universal Plug and Play Internet Gateway Device Standardized Gateway Device Protocol," September 2006.

Appendix A. Change Log

TOC

A.1. draft-ietf-v6ops-cpe-simple-security-00 to draft-ietf-v6ops-cpe-simple-security-01

TOC

*Added requirements for sequestering DHCP and DNS proxy resolver services to the local network.

*Fixed numbering of recommendations.

*Local Network Protection is now [RFC4864] (Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6," May 2007.).

*SCTP is now [RFC4960] (Stewart, R., "Stream Control Transmission Protocol," September 2007.).

*Moved some references to informative.

*Corrected the reference for [HOAGLAND] (Hoagland, J. and S. Krishnan, "Teredo Security Concerns Beyond What Is In RFC 4380," July 2007.).

TOC

A.2. draft-ietf-v6ops-cpe-simple-security-01 to draft-ietf-v6ops-cpe-simple-security-02

*Inserted R20, i.e. do not enforce TCP window invariant unless the TCP window-scale is known for the state.

*Filled out Section 4 (Summary of Recommendations).

*Added reference to [RFC1323] (Jacobson, V., Braden, B., and D. Borman, “TCP Extensions for High Performance,” May 1992.).

*Updated the reference for [HOAGLAND] (Hoagland, J. and S. Krishnan, “Teredo Security Concerns Beyond What Is In RFC 4380,” July 2007.).

*Expanded list of contributers and commenters.

*Mention that UDP-Lite should be handled just like UDP.

*Added section for generic upper layer transport protocols.

*Expanded on recommendations for IPsec ESP filtering.

*Expanded overview of recommendations with discussion about IP mobility and IPsec interactions.

*Added a security considerations section.

A.3. draft-ietf-v6ops-cpe-simple-security-02 to draft-ietf-v6ops-cpe-simple-security-03

TOC

*Fixed some spelling errors.

*Replace "prevent such attacks" with "defend against such attacks" everywhere.

*Replace "mapping" with "state record" in the TCP filters section.

*Added recommendations for SCTP and DCCP.

Author's Address

TOC

	james woodyatt (editor)
	Apple Inc.

	1 Infinite Loop
	Cupertino, CA 95014
	US
Email:	jhw@apple.com

Full Copyright Statement

TOC

Intellectual Property