Internet Engineering Task Force (IETF)
Request for Comments: 6092
Category: Informational
ISSN: 2070-1721

J. Woodyatt, Ed.
Apple

January 2011

Recommended Simple Security Capabilities in Customer Premises Equipment (CPE) for Providing Residential IPv6 Internet Service

Abstract

This document identifies a set of recommendations for the makers of devices and describes how to provide for "simple security" capabilities at the perimeter of local-area IPv6 networks in Internet-enabled homes and small offices.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 5741.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc6092.

Table of Contents

1. Introduction

Some IPv6 gateway devices that enable delivery of Internet services in residential and small-office settings may be augmented with "simple security" capabilities as described in "Local Network Protection for IPv6" [RFC4864]. In general, these capabilities cause packets to be discarded in an attempt to make local networks and the Internet more secure. However, it is worth noting that some packets sent by legitimate applications may also be discarded in this process, affecting reliability and ease of use for these applications.

There is 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. This document is intended to highlight reasonable limitations on end-to-end transparency where security considerations are deemed important to promote local and Internet security.

The reader is cautioned always to remember that the typical residential or small-office network administrator has no expertise whatsoever in Internet engineering. Configuration interfaces for router/gateway appliances marketed toward them should be easy to understand and even easier to ignore. In particular, extra care should be used in the design of baseline operating modes for unconfigured devices, since most devices will never be changed from their factory configurations.

1.1. Special Language

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 [RFC2119].

Additionally, the key word "DEFAULT" is to be interpreted in this document as pertaining to a configuration as applied by a vendor, prior to the administrator changing it for its initial activation.

1.2. Use of Normative Keywords

> NOTE WELL: This document is not a standard, and conformance with it is not required in order to claim conformance with IETF standards for IPv6. It uses the normative keywords defined in the previous section only for precision.

Particular attention is drawn to recommendation REC-49, which calls for an easy way to set a gateway to a transparent mode of operation.

2. Overview

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] for a single local-area network, e.g., an Ethernet network, a Wi-Fi network, or a bridge between two or more such segments. They have only one interface by which they can access the Internet service at any one time, using any of several possible sub-IP mechanisms, including tunnels and transition mechanisms.

In referring to the security capabilities of residential gateways, it is reasonable to distinguish between their "interior" network, i.e., the local-area network, and their "exterior" networks, e.g., the public Internet and the networks of Internet service providers. This document is concerned only with the behavior of IP packet filters that police the flow of traffic between the interior IPv6 network and the exterior IPv6 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], but they can be summarized as follows.

- Check all traffic to and from the public Internet for basic sanity, e.g., filter for spoofs and misdirected (sometimes called "Martian") packets [RFC4949].

- Allow tracking of application usage by source and destination network addresses and ports.

- 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 administrative policy.

- Isolate local network DHCPv6 and DNS 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] 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] recommends applying stateful packet filtering at residential IPv6 gateways that conforms to the user expectations already in place.

Conventional stateful packet filters activate new states as a side effect of forwarding 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.

2.1. Basic Sanitation

In addition to the functions required of all IPv6 routers [RFC4294], residential gateways are expected to have basic stateless filters for prohibiting certain kinds of traffic with invalid headers, e.g., "Martian" packets, spoofs, routing header type code zero, etc. (See Section 3.1 for more details.)

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.

Finally, Internet gateways that route multicast traffic are expected to implement appropriate filters for multicast traffic to limit the scope of multicast groups that span the demarcation between residential networks and service provider networks.

2.2. Internet Layer Protocols

As virtual private networking tunnels are regarded as an unacceptably wide attack surface, this document recommends that the DEFAULT operating mode for residential IPv6 simple security be to treat Generic Packet Tunneling [RFC2473] and similar protocols as opaque transport layers, i.e., inbound tunnel initiations are denied and outbound tunnel initiations are accepted.

IPsec transport and tunnel modes are explicitly secured by definition, so this document recommends that the DEFAULT operating mode permit IPsec. To facilitate the use of IPsec in support of IPv6

mobility, the Internet Key Exchange (IKE) protocol [RFC5996] and the Host Identity Protocol (HIP) [RFC5201] should also be permitted in the DEFAULT operating mode.

2.3. Transport Layer Protocols

IPv6 simple security functions are principally concerned with the stateful filtering of the Internet Control Message Protocol (ICMPv6) [RFC4443] and transport layers like the User Datagram Protocol (UDP) [RFC0768], the Lightweight User Datagram Protocol (UDP-Lite) [RFC3828], the Transmission Control Protocol (TCP) [RFC0793], the Stream Control Transmission Protocol (SCTP) [RFC4960], the Datagram Congestion Control Protocol (DCCP) [RFC4340], and potentially any standards-track transport protocols to be defined in the future.

The general operating principle is that transport layer traffic is not forwarded into the interior network of a residential IPv6 gateway unless it has been solicited explicitly by interior transport endpoints, e.g., by matching the reverse path for previously forwarded outbound traffic, or by matching configured exceptions set by the network administrator. 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

This section describes the specific recommendations made by this document in full detail. Section 4 is a summary.

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. Other recommended filters are intended to be sensitive to the "direction" of traffic flows. Applied to bidirectional transport flows, "direction" has a specific meaning in this document.

Packets are said to be "outbound" if they originate at nodes located in the interior network for exterior destinations, and "inbound" if they arrive from exterior sources with interior destinations.

Flows are said to be "outbound" if the originator of the initial packet in any given transport association is an interior node and one or more of the participants are located in the exterior. Flows are said to be "inbound" if the originator of the initial packet is an exterior node and one or more of the participants are nodes on the interior network.

3.1. Stateless Filters

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 implement ingress filtering (see [RFC2827] and [RFC3704]), to enforce multicast scope boundaries, and to isolate certain local network services from the public Internet.

REC-1: Packets bearing multicast source addresses in their outer IPv6 headers MUST NOT be forwarded or transmitted on any interface.

REC-2: Packets bearing multicast destination addresses in their outer IPv6 headers of equal or narrower scope (see "IPv6 Scoped Address Architecture" [RFC4007]) than 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, and it SHOULD be configurable by the network administrator.

REC-3: Packets bearing source and/or destination addresses forbidden to appear in the outer headers of packets transmitted over the public Internet MUST NOT be forwarded. In particular, site-local addresses are deprecated by [RFC3879], and [RFC5156] explicitly forbids the use of address blocks of types IPv4-Mapped Addresses, IPv4-Compatible Addresses, Documentation Prefix, and Overlay Routable Cryptographic Hash IDentifiers (ORCHID).

REC-4: Packets bearing deprecated extension headers prior to their first upper-layer-protocol header SHOULD NOT be forwarded or transmitted on any interface. In particular, all packets with routing extension header type 0 [RFC2460] preceding the first upper-layer-protocol header MUST NOT be forwarded. See [RFC5095] for additional background.

REC-5: Outbound packets MUST NOT be forwarded if the source address in their outer IPv6 header does not have a unicast prefix configured for use by globally reachable nodes on the interior network.

REC-6: 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.

REC-7: By DEFAULT, packets with unique local source and/or destination addresses [RFC4193] SHOULD NOT be forwarded to or from the exterior network.

REC-8: By DEFAULT, inbound DNS queries received on exterior interfaces MUST NOT be processed by any integrated DNS resolving server.

REC-9: Inbound DHCPv6 discovery packets [RFC3315] received on exterior interfaces MUST NOT be processed by any integrated DHCPv6 server or relay agent.

NOTE WELL: Nothing in this document relieves residential Internet gateways, when processing headers to identify valid sequences of upper-layer transport packets, from any of the requirements of the "Internet Protocol, Version 6 (IPv6) Specification" [RFC2460], including any and all future updates and revisions.

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 are abandoned when no further communication is detected after some period of time.

3.2.1. Internet Control and Management

Recommendations for filtering ICMPv6 messages in firewall devices are described separately in [RFC4890] and apply to residential gateways, with the additional recommendation that incoming "Destination Unreachable" and "Packet Too Big" error messages that don’t match any filtering state should be dropped.

REC-10: IPv6 gateways SHOULD NOT forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing IP headers that do not match generic upper-layer transport state records.

3.2.2. Upper-Layer Transport Protocols

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.

One key aspect of how a packet filter behaves is the way it evaluates the exterior address of an endpoint when applying a filtering rule. A gateway is said to have "endpoint-independent filtering" behavior when the exterior address is not evaluated when matching a packet with a flow. A gateway is said to have "address-dependent filtering" behavior when the exterior address of a packet is required to match the exterior address for its flow.

REC-11: If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" behavior for generic upper-layer transport protocols. 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 other protocols. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

REC-12: 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.

The Internet security community is never completely at rest. New attack surfaces, and vulnerabilities in them, are typically discovered faster than they can be patched by normal equipment upgrade cycles. It’s therefore important for vendors of residential gateway equipment to provide automatic software updates to patch vulnerabilities as they are discovered.

REC-13: Residential IPv6 gateways SHOULD provide a convenient means to update their firmware securely, for the installation of security patches and other manufacturer-recommended changes.

Vendors can expect users and operators to have differing viewpoints on the maintenance of patches, with some preferring automatic update and some preferring manual procedures. Those preferring automatic update may also prefer either to download from a vendor site or from one managed by their network provider. To handle the disparity, vendors are advised to provide both manual and automatic options. In

the automatic case, they would do well to facilitate pre-configuration of the download URL and a means of validating the software image, such as a certificate.

3.2.3. UDP Filters

"Network Address Translation (NAT) Behavioral Requirements for Unicast UDP" [RFC4787] defines the terminology and best current practice for stateful filtering of UDP applications in IPv4 with NAT, which serves as the model for behavioral requirements for simple UDP security in IPv6 gateways, notwithstanding the requirements related specifically to network address translation.

An interior endpoint initiates a UDP flow 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 flow. The state record defines the interior and exterior IP addresses and ports used between all packets in the flow.

State records for UDP flows remain active while they are in use and are only abandoned after an idle period of some time.

REC-14: A state record for a UDP flow where both source and destination 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.

REC-15: A state record for a UDP flow where one or both of the source and destination 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] 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.

REC-16: A state record for a UDP flow 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 of [RFC4787], 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.

REC-17: If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" 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. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

Application mechanisms may depend on the reception of ICMPv6 error messages triggered by the transmission of UDP messages. One such mechanism is path MTU discovery [RFC1981].

REC-18: If a gateway forwards a UDP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing UDP headers that match the flow state record.

REC-19: Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a UDP flow.

REC-20: UDP-Lite flows [RFC3828] SHOULD be handled in the same way as UDP flows, except that the upper-layer transport protocol identifier for UDP-Lite is not the same as UDP; therefore, UDP packets MUST NOT match UDP-Lite state records, and vice versa.

3.2.4. IPsec and Internet Key Exchange (IKE)

The Internet Protocol security (IPsec) suite 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.

REC-21: 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 "Authentication Header (AH)" [RFC4302] in their outer IP extension header chain.

REC-22: 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)" [RFC4303] in their outer IP extension header chain.

REC-23: If a gateway forwards an ESP flow, it MUST also forward (in the reverse direction) ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing ESP headers that match the flow state record.

Internet Key Exchange (IKE) is a secure mechanism for performing mutual authentication, exchanging cryptographic material, and establishing IPsec Security Associations between peers. Residential IPv6 gateways are expected to facilitate the use of IPsec security policies by allowing inbound IKE flows.

REC-24: 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 (IKE) protocol [RFC5996].

REC-25: In all operating modes, IPv6 gateways SHOULD use filter state records for Encapsulating Security Payload (ESP) [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 the security parameters index (SPI) SHOULD NOT be used. Likewise, any mechanism that depends on detection of Internet Key Exchange (IKE) [RFC5996] initiations SHOULD NOT be used.

The Host Identity Protocol (HIP) is a secure mechanism for establishing host identity and secure communications between authenticated hosts. Residential IPv6 gateways need not prohibit inbound HIP flows.

REC-26: 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 "Host Identity Protocol (HIP)" [RFC5201] in their outer IP extension header chain.

3.2.5. Mobility Support in IPv6

Mobility support in IPv6 [RFC3775] relies on the use of an encapsulation mechanism in flows between mobile nodes and their correspondent nodes, involving the use of the Type 2 IPv6 Routing Header, the Home Address destination header option, and the Mobility

extension header. In contrast to mobility support in IPv4, mobility is a standard feature of IPv6, and no security benefit is generally to be gained by denying communications with either interior or exterior mobile nodes.

Not all usage scenarios of mobility support in IPv6 are expected to be compatible with IPv6 simple security. In particular, exterior mobile nodes are expected to be prohibited from establishing bindings with interior correspondent nodes by the filtering of unsolicited inbound Mobility Header messages, unless they are the subject of an IPsec security policy.

REC-27: The state records for flows initiated by outbound packets that bear a Home Address destination option [RFC3775] are distinguished by the addition of the home address of the flow as well as the interior care-of address. IPv6 gateways MUST NOT prohibit the forwarding of any inbound packets bearing type 2 routing headers, which otherwise match a flow state record, and where A) the address in the destination field of the IPv6 header matches the interior care-of address of the flow, and B) the Home Address field in the Type 2 Routing Header matches the home address of the flow.

REC-28: Valid sequences of Mobility Header [RFC3775] packets MUST be forwarded for all outbound and explicitly permitted inbound Mobility Header flows.

REC-29: If a gateway forwards a Mobility Header [RFC3775] flow, then it MUST also forward, in both directions, the IPv4 and IPv6 packets that are encapsulated in IPv6 associated with the tunnel between the home agent and the correspondent node.

REC-30: If a gateway forwards a Mobility Header [RFC3775] flow, then it MUST also forward (in the reverse direction) ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing any headers that match the associated flow state records.

3.3. Connection-Oriented Filters

Most Internet applications use connection-oriented transport protocols with orderly release semantics. These protocols include TCP, SCTP, DCCP, and potentially any future IETF Standards-Track transport protocols that use such semantics. Stateful packet filters track the state of individual transport flows and prohibit the forwarding of packets that do not match the state of an active flow and do not conform to a rule for the automatic creation of such state.

3.3.1. TCP Filters

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

Some peer-to-peer applications use an alternate method of connection initiation termed "simultaneous-open" ([RFC0793], Figure 8) to traverse stateful filters. In the simultaneous-open mode of operation, both peers send SYN packets for the same TCP flow. 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 flow that conform to valid transitions of the TCP state machine ([RFC0793], Figure 6).

REC-31: All valid sequences of TCP packets (defined in [RFC0793]) MUST be forwarded for outbound flows and explicitly permitted inbound flows. In particular, both the normal TCP 3-way handshake mode of operation and the simultaneous-open mode of operation MUST be supported.

It is possible to reconstruct enough of the state of a TCP flow 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.

REC-32: The TCP window invariant MUST NOT be enforced on flows for which the filter did not detect whether the window-scale option (see [RFC1323]) 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 flow if its security policy permits. Several different policies are possible, as described in [RFC4787] and extended in [RFC5382].

REC-33: If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" 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. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

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 ICMPv6 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 [RFC5382], but the basic outcome of it is that filters need to wait on signaling errors until simultaneous-open will not be impaired.

REC-34: By DEFAULT, a gateway MUST respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited) to any unsolicited inbound SYN packet after waiting at least 6 seconds without first forwarding the associated outbound SYN or SYN/ACK from the interior peer.

A TCP filter maintains state associated with in-progress connections and established flows. 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 flow state records for crashed endpoints, followed by closed flows and partially open flows. 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 cannot determine whether the endpoint is active, then the associated state record needs to be retained until the TCP flow has been idle for some time.

> Note: An established TCP flow 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] and [RFC4294] can reduce the chances of abandoning a live flow.

TCP flows can stay in the established phase indefinitely without exchanging packets. Some end-hosts can be configured to send keep-alive packets on such idle flows; by default, such packets are sent every two hours, if enabled [RFC1122]. Consequently, a filter that waits for slightly over two hours can detect idle flows with keep-alive packets being sent at the default rate. TCP flows 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].

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

REC-35: If a gateway cannot determine whether the endpoints of a TCP flow 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 flow idle-timeout" MUST NOT be less than two hours four minutes, as discussed in [RFC5382]. The value of the "transitory flow 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 RST packets or TCP flows in the TIME-WAIT state is left unspecified. A gateway MAY hold state for a flow in the 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 flow, holding state for a closed flow can limit the throughput of flows through a gateway with limited resources. [RFC1337] 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 flow, or abandons a flow in the TIME-WAIT state before the four-minute TIME-WAIT period expires. The decision either to discard or to respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited) is left up to the implementation.

Behavior for notifying endpoints when abandoning live flows is left unspecified. When a gateway abandons a live flow, 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 ICMPv6 error messages triggered by the transmission of TCP segments. One such mechanism is path MTU discovery, which is required for correct operation of TCP.

REC-36: If a gateway forwards a TCP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing TCP headers that match the flow state record.

REC-37: Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a TCP flow.

3.3.2. SCTP Filters

Because Stream Control Transmission Protocol (SCTP) [RFC4960] flows 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 connection-oriented filtering behaviors similar to those 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.

Some peer-to-peer SCTP 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 the 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 chunk without an ERROR or ABORT chunk is received at one end.

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 ([RFC4960], Figure 3).

REC-38: All valid sequences of SCTP packets (defined in [RFC4960]) MUST be forwarded for outbound associations and explicitly permitted inbound associations. In particular, both the normal SCTP association establishment and the simultaneous-open mode 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 ICMPv6 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 the simultaneous-open mode of operation.

REC-39: By DEFAULT, a gateway MUST respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited), to any unsolicited inbound INIT packet after waiting at least 6 seconds without first forwarding the associated outbound INIT from the interior peer.

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 record 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 recommendations in [RFC4294] 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] 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]).

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.

REC-40: 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] regarding the hazards of TIME-WAIT assassination is 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 flow, 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 ICMPv6 "Destination Unreachable" error code 1 (Communication with destination 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 ICMPv6 error messages triggered by the transmission of SCTP packets.

REC-41: If a gateway forwards an SCTP association, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing SCTP headers that match the association state record.

REC-42: Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for an SCTP association.

3.3.3. DCCP Filters

The connection semantics described in the "Datagram Congestion Control Protocol (DCCP)" [RFC4340] are very similar to those of TCP. An interior endpoint initiates a DCCP flow 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 flow that conform to valid transitions of the DCCP state machine ([RFC4340], Section 8).

REC-43: All valid sequences of DCCP packets (defined in [RFC4340]) MUST be forwarded for all flows to exterior servers, and for any flows to interior servers that have explicitly permitted service codes.

It is possible to reconstruct enough of the state of a DCCP flow 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 flow if its security policy permits. As with TCP, several different policies are possible, with a good discussion of the issue involved presented in [RFC4787] and extended in [RFC5382].

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 flows 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 ICMPv6 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 connections and established flows. 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 flows and partially open flows. No such method exists for DCCP, and flows 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 recommendations in [RFC4294] can reduce the chances of abandoning a live flow.

DCCP flows 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 flow idle-timeout" for a stateful packet filter is defined as the minimum time a DCCP flow in the open state must remain idle before the filter considers the associated state record a candidate

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

REC-44: A gateway MAY abandon a DCCP state record if it has been idle for some time. In such cases, the value of the "open flow idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory flow idle-timeout" MUST NOT be less than eight minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

Behavior for handling DCCP-Reset packets or flows in the TIMEWAIT state is left unspecified. A gateway MAY hold state for a flow in the 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 flow, holding state for a closed flow can limit the throughput of flows through a gateway with limited resources. [RFC1337] discusses hazards associated with TIME-WAIT assassination in TCP, and similar hazards exist 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 flow, or abandons a flow in the TIMEWAIT state before the four-minute 2MSL period (two times the maximum segment lifetime [RFC4340]) expires. The decision either to discard or to respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited) is left up to the implementation.

Behavior for notifying endpoints when abandoning live flows is left unspecified. When a gateway abandons a live flow, 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 ICMPv6 error messages triggered by the transmission of DCCP packets. One such mechanism is path MTU discovery, which is required for correct operation.

REC-45: If an Internet gateway forwards a DCCP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing DCCP headers that match the flow state record.

REC-46: Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a DCCP flow.

3.3.4. Level 3 Multihoming Shim Protocol for IPv6 (Shim6)

While IPv6 simple security is applicable to residential networks with only one Internet service provider at a time, the use of the Level 3 Multihoming Shim Protocol for IPv6 (Shim6) [RFC5533] is necessary for communications with some multihomed exterior destinations. No special recommendations are made in this document for processing the Shim6 message format (protocol 140) beyond the recommendations in Section 3.2.2. The content of the Shim6 payload extension header may be ignored.

REC-47: Valid sequences of packets bearing Shim6 payload extension headers in their outer IP extension header chains MUST be forwarded for all outbound and explicitly permitted flows. The content of the Shim6 payload extension header MAY be ignored for the purpose of state tracking.

3.4. Passive Listeners

Some applications expect to solicit traffic from exterior nodes without advance knowledge of the exterior addresses of their peers. This requirement is met by IPv4/NAT gateways, typically by the use of either the NAT Port Mapping Protocol [NAT-PMP] or the Universal Plug and Play Internet Gateway Device [UPnP-IGD] standardized device control protocol. On IPv4/NAT networks connected by gateways without such services, applications must use techniques like Session Traversal Utilities for NAT (STUN) [RFC5389] to obtain and maintain connectivity, despite the translation and filtering effects of NAT.

While NAT for IPv6 is unlikely to be used in most residential gateways, the simple security functions recommended by this document, and their filtering effects, are derived from comparable functions already in widespread use on the IPv4 Internet. A similar barrier to communication at passive listeners is a natural outcome of the deployment of NAT for IPv6. To avoid the need for IPv6 applications to use techniques like STUN for opening and maintaining dynamic filter state, something similar to NAT-PMP and UPnP-IGD, but without actually supporting NAT, could be deployed. Alas, no consensus has yet emerged in the Internet engineering community as to what is most appropriate for residential IPv6 usage scenarios.

One proposal that has been offered is the Application Listener Discovery Protocol [WOODYATT-ALD] document. 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 [RFC5189] and the Next Steps in Signaling framework [RFC4080]. Until a consensus emerges around a specific method, the following recommendations are the best guidance available.

REC-48: Internet gateways with IPv6 simple security capabilities SHOULD 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.

REC-49: Internet gateways with IPv6 simple security capabilities MUST provide an easily selected configuration option that permits a "transparent mode" of operation that forwards all unsolicited flows regardless of forwarding direction, i.e., not to use the IPv6 simple security capabilities of the gateway. The transparent mode of operation MAY be the default configuration.

In general, "transparent mode" will enable more flexibility and reliability for applications that require devices to be contacted inside the home directly, particularly in the absence of a protocol as described in REC-48. Operating in transparent mode may come at the expense of security if there are IPv6 nodes in the home that do not have their own host-based firewall capability and require a firewall in the gateway in order not to be compromised.

3.5. Management Applications

Subscriber-managed residential gateways are unlikely ever to be completely zero-configuration, but their administrators will very often possess no particular expertise in Internet engineering. In general, the specification of management interfaces for residential gateways is out of scope for this document, but the security of subscriber-managed gateways merits special attention here.

REC-50: By DEFAULT, subscriber-managed residential gateways MUST NOT offer management application services to the exterior network.

4. Summary of Recommendations

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

REC-1 Packets bearing multicast source addresses in their outer IPv6 headers MUST NOT be forwarded or transmitted on any interface.

REC-2 Packets bearing multicast destination addresses in their outer IPv6 headers of equal or narrower scope (see "IPv6 Scoped Address Architecture" [RFC4007]) than 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, and it SHOULD be configurable by the network administrator.

REC-3 Packets bearing source and/or destination addresses forbidden to appear in the outer headers of packets transmitted over the public Internet MUST NOT be forwarded. In particular, site-local addresses are deprecated by [RFC3879], and [RFC5156] explicitly forbids the use of address blocks of types IPv4-Mapped Addresses, IPv4-Compatible Addresses, Documentation Prefix, and Overlay Routable Cryptographic Hash IDentifiers (ORCHID).

REC-4 Packets bearing deprecated extension headers prior to their first upper-layer-protocol header SHOULD NOT be forwarded or transmitted on any interface. In particular, all packets with routing extension header type 0 [RFC2460] preceding the first upper-layer-protocol header MUST NOT be forwarded. See [RFC5095] for additional background.

REC-5 Outbound packets MUST NOT be forwarded if the source address in their outer IPv6 header does not have a unicast prefix configured for use by globally reachable nodes on the interior network.

REC-6 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.

REC-7 By DEFAULT, packets with unique local source and/or destination addresses [RFC4193] SHOULD NOT be forwarded to or from the exterior network.

REC-8 By DEFAULT, inbound DNS queries received on exterior interfaces MUST NOT be processed by any integrated DNS resolving server.

REC-9 Inbound DHCPv6 discovery packets [RFC3315] received on exterior interfaces MUST NOT be processed by any integrated DHCPv6 server or relay agent.

REC-10 IPv6 gateways SHOULD NOT forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing IP headers that do not match generic upper-layer transport state records.

REC-11 If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" behavior for generic upper-layer transport protocols. 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 other protocols. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

REC-12 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.

REC-13 Residential IPv6 gateways SHOULD provide a convenient means to update their firmware securely, for the installation of security patches and other manufacturer-recommended changes.

REC-14 A state record for a UDP flow where both source and destination 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.

REC-15 A state record for a UDP flow where one or both of the source and destination 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.

REC-16 A state record for a UDP flow 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.

REC-17 If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" 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. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

REC-18 If a gateway forwards a UDP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing UDP headers that match the flow state record.

REC-19 Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a UDP flow.

REC-20 UDP-Lite flows [RFC3828] SHOULD be handled in the same way as UDP flows, except that the upper-layer transport protocol identifier for UDP-Lite is not the same as UDP; therefore, UDP packets MUST NOT match UDP-Lite state records, and vice versa.

REC-21 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 "Authentication Header (AH)" [RFC4302] in their outer IP extension header chain.

REC-22 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)" [RFC4303] in their outer IP extension header chain.

REC-23 If a gateway forwards an ESP flow, it MUST also forward (in the reverse direction) ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing ESP headers that match the flow state record.

REC-24 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 (IKE) Protocol [RFC5996].

REC-25 In all operating modes, IPv6 gateways SHOULD use filter state records for Encapsulating Security Payload (ESP) [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) [RFC5996] initiations SHOULD NOT be used.

REC-26 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 "Host Identity Protocol (HIP)" [RFC5201] in their outer IP extension header chain.

REC-27 The state records for flows initiated by outbound packets that bear a Home Address destination option [RFC3775] are distinguished by the addition of the home address of the flow as well as the interior care-of address. IPv6 gateways MUST NOT prohibit the forwarding of any inbound packets bearing type 2 routing headers, which otherwise match a flow state record, and where A) the address in the destination field of the IPv6 header matches the interior care-of address of the flow, and B) the Home Address field in the Type 2 Routing Header matches the home address of the flow.

REC-28 Valid sequences of Mobility Header [RFC3775] packets MUST be forwarded for all outbound and explicitly permitted inbound Mobility Header flows.

REC-29 If a gateway forwards a Mobility Header [RFC3775] flow, then it MUST also forward, in both directions, the IPv4 and IPv6 packets that are encapsulated in IPv6 associated with the tunnel between the home agent and the correspondent node.

REC-30 If a gateway forwards a Mobility Header [RFC3775] flow, then it MUST also forward (in the reverse direction) ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing any headers that match the associated flow state records.

REC-31 All valid sequences of TCP packets (defined in [RFC0793]) MUST be forwarded for outbound flows and explicitly permitted inbound flows. In particular, both the normal TCP 3-way handshake mode of operation and the simultaneous-open mode of operation MUST be supported.

REC-32 The TCP window invariant MUST NOT be enforced on flows for which the filter did not detect whether the window-scale option (see [RFC1323]) was sent in the 3-way handshake or simultaneous-open.

REC-33 If application transparency is most important, then a stateful packet filter SHOULD have "endpoint-independent filtering" 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. Filtering behavior SHOULD be endpoint independent by DEFAULT in gateways intended for provisioning without service-provider management.

REC-34 By DEFAULT, a gateway MUST respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited), to any unsolicited inbound SYN packet after waiting at least 6 seconds without first forwarding the associated outbound SYN or SYN/ACK from the interior peer.

REC-35 If a gateway cannot determine whether the endpoints of a TCP flow 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 flow idle-timeout" MUST NOT be less than two hours four minutes, as discussed in [RFC5382]. The value of the "transitory flow idle-timeout" MUST NOT be less than four minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

REC-36 If a gateway forwards a TCP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing TCP headers that match the flow state record.

REC-37 Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a TCP flow.

REC-38 All valid sequences of SCTP packets (defined in [RFC4960]) MUST be forwarded for outbound associations and explicitly permitted inbound associations. In particular, both the normal SCTP association establishment and the simultaneous-open mode of operation MUST be supported.

REC-39 By DEFAULT, a gateway MUST respond with an ICMPv6 "Destination Unreachable" error code 1 (Communication with destination administratively prohibited) to any unsolicited inbound INIT packet after waiting at least 6 seconds without first forwarding the associated outbound INIT from the interior peer.

REC-40 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.

REC-41 If a gateway forwards an SCTP association, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing SCTP headers that match the association state record.

REC-42 Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for an SCTP association.

REC-43 All valid sequences of DCCP packets (defined in [RFC4340]) MUST be forwarded for all flows to exterior servers, and for any flows to interior servers with explicitly permitted service codes.

REC-44 A gateway MAY abandon a DCCP state record if it has been idle for some time. In such cases, the value of the "open flow idle-timeout" MUST NOT be less than two hours four minutes. The value of the "transitory flow idle-timeout" MUST NOT be less than eight minutes. The value of the idle-timeouts MAY be configurable by the network administrator.

REC-45 If an Internet gateway forwards a DCCP flow, it MUST also forward ICMPv6 "Destination Unreachable" and "Packet Too Big" messages containing DCCP headers that match the flow state record.

REC-46 Receipt of any sort of ICMPv6 message MUST NOT terminate the state record for a DCCP flow.

REC-47 Valid sequences of packets bearing Shim6 payload extension headers in their outer IP extension header chains MUST be forwarded for all outbound and explicitly permitted flows. The content of the Shim6 payload extension header MAY be ignored for the purpose of state tracking.

REC-48 Internet gateways with IPv6 simple security capabilities SHOULD 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.

REC-49 Internet gateways with IPv6 simple security capabilities MUST provide an easily selected configuration option that permits a "transparent mode" of operation that forwards all unsolicited flows regardless of forwarding direction, i.e., not to use the IPv6 simple security capabilities of the gateway. The transparent mode of operation MAY be the default configuration.

REC-50 By DEFAULT, subscriber-managed residential gateways MUST NOT offer management application services to the exterior network.

5. Contributors

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

Jari Arkko	Ran Atkinson
Fred Baker	Norbert Bollow
Cameron Byrne	Brian Carpenter
Remi Despres	Arnaud Ebalard
Fabrice Fontaine	Jun-ichiro "itojun" Hagino
Thomas Herbst	Christian Huitema
Joel Jaeggli	Cullen Jennings
Suresh Krishnan	Erik Kline
Julien Laganier	Kurt Erik Lindqvist
Mohamed Boucadair	Keith Moore
Robert Moskowitz	Teemu Savolainen
Hemant Singh	Yaron Sheffer
Mark Townsley	Iljitsch van Beijnum
Magnus Westerlund	Dan Wing

The editor thanks them all for their contributions.

It must be noted that a substantial portion of the text describing the detailed requirements for TCP and UDP filtering is derived or transposed from [RFC4787] and [RFC5382]. The editors of those documents, Francois Audet and Saikat Guha, also deserve substantial credit for the form of the present document.

6. Security Considerations

The IPv6 stateful filtering behavior described in this document is intended to be similar in function to the filtering behavior of commonly used IPv4/NAT gateways, which have been widely sold as a security tool for residential and small-office/home-office networks. As noted in the Security Considerations section of [RFC2993], the true impact of these tools may be a reduction in security. It may be generally assumed that the impacts discussed in that document related to filtering (and not translation) are to be expected with the simple IPv6 security mechanisms described here.

In particular, it is 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] 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.

The security functions described in this document may be considered redundant in the event that all IPv6 hosts using a particular gateway have their own IPv6 host firewall capabilities enabled. At the time of this writing, the vast majority of commercially available operating systems with support for IPv6 include IPv6 host firewall capability.

Also worth noting explicitly, a practical side-effect of the recommendations in Section 3.2.4, to allow inbound IPsec and IKE flows from exterior to interior, is to facilitate more transparent communication by the use of an unauthenticated mode of IPsec, as described in "Better-Than-Nothing-Security: An Unauthenticated Mode of IPsec" [RFC5386], and this may be a departure from expectations of transparency set by traditional IPv4/NAT residential gateways.

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.

The IETF makes no statement, expressed or implied, as to whether using the capabilities described in this document ultimately improves security for any individual users or for the Internet community as a whole.

7. References

7.1. Normative References

[RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.

[RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC1323] Jacobson, V., Braden, B., and D. Borman, "TCP Extensions for High Performance", RFC 1323, May 1992.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[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.

[RFC3879] Huitema, C. and B. Carpenter, "Deprecating Site Local Addresses", RFC 3879, September 2004.

[RFC4007] Deering, S., Haberman, B., Jinmei, T., Nordmark, E., and B. Zill, "IPv6 Scoped Address Architecture", RFC 4007, March 2005.

[RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.

[RFC4302] Kent, S., "IP Authentication Header", RFC 4302, December 2005.

[RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.

[RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[RFC4443] Conta, A., Deering, S., and M. Gupta, "Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification", RFC 4443, March 2006.

[RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.

[RFC4890] Davies, E. and J. Mohacsi, "Recommendations for Filtering ICMPv6 Messages in Firewalls", RFC 4890, May 2007.

[RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC5095] Abley, J., Savola, P., and G. Neville-Neil, "Deprecation of Type 0 Routing Headers in IPv6", RFC 5095, December 2007.

[RFC5156] Blanchet, M., "Special-Use IPv6 Addresses", RFC 5156, April 2008.

[RFC5201] Moskowitz, R., Nikander, P., Jokela, P., and T. Henderson, "Host Identity Protocol", RFC 5201, April 2008.

[RFC5996] Kaufman, C., Hoffman, P., Nir, Y., and P. Eronen, "Internet Key Exchange Protocol Version 2 (IKEv2)", RFC 5996, September 2010.

7.2. Informative References

[NAT-PMP] Cheshire, S., Krochmal, M., and K. Sekar, "NAT Port Mapping Protocol (NAT-PMP)", Work in Progress, April 2008.

[RFC1122] Braden, R., "Requirements for Internet Hosts - Communication Layers", STD 3, RFC 1122, October 1989.

[RFC1337] Braden, B., "TIME-WAIT Assassination Hazards in TCP", RFC 1337, May 1992.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[RFC1981] McCann, J., Deering, S., and J. Mogul, "Path MTU Discovery for IP version 6", RFC 1981, August 1996.

[RFC2473] Conta, A. and S. Deering, "Generic Packet Tunneling in IPv6 Specification", RFC 2473, December 1998.

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 2000.

[RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993, November 2000.

[RFC3704] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", BCP 84, RFC 3704, March 2004.

[RFC4080] Hancock, R., Karagiannis, G., Loughney, J., and S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework", RFC 4080, June 2005.

[RFC4294] Loughney, J., "IPv6 Node Requirements", RFC 4294, April 2006.

[RFC4864] Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6", RFC 4864, May 2007.

[RFC4949] Shirey, R., "Internet Security Glossary, Version 2", RFC 4949, August 2007.

[RFC5189] Stiemerling, M., Quittek, J., and T. Taylor, "Middlebox Communication (MIDCOM) Protocol Semantics", RFC 5189, March 2008.

[RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", BCP 142, RFC 5382, October 2008.

[RFC5386] Williams, N. and M. Richardson, "Better-Than-Nothing Security: An Unauthenticated Mode of IPsec", RFC 5386, November 2008.

[RFC5389] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for NAT (STUN)", RFC 5389, October 2008.

[RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, June 2009.

[UPnP-IGD] UPnP Forum, "Universal Plug and Play Internet Gateway Device Standardized Device Control Protocol", September 2010, <http://upnp.org/specs/gw/igd2/>.

[WOODYATT-ALD] Woodyatt, J., "Application Listener Discovery (ALD) for IPv6", Work in Progress, July 2008.

Author's Address

James Woodyatt (editor)
Apple Inc.
1 Infinite Loop
Cupertino, CA 95014
US

EMail: jhw@apple.com