Darrel Lewis

James Gill
Verizon Business.
Darrel Lewis
Cisco Systems, Inc.
Paul Quinn
Cisco Systems, Inc.
Peter Schoenmaker
NTT America

October 2006

Service Provider Infrastructure Security
<draft-lewis-infrastructure-security-00>

Status of this Memo

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at

Copyright Notice

http://www.ietf.org/shadow.html.

This Internet-Draft will expire on October 28, 2006.

Abstract

This RFC defines best current practices for implementing Service Provider network infrastructure protection for network elements. This RFC complements and extends RFC 2267 and RFC 3704. RFC 2267 provides guidelines for filtering traffic on the ingress to service provider networks. RFC 3704 expands the recommendations described in RFC 2267 to address operational filtering guidelines for single and multi-homed environments. The focus of those RFCs is on filtering ingress packets ingress, regardless of destination, if those packets are have spoofed source address or fall within "reserved" address space. Deployment of RFCs 2267 and 3704 has limited the effects of denial of service attacks by dropping ingress packets with spoofed source addresses, which in turn offers other benefits by ensuring that packets coming into a network originate from validly allocated and consistent sources.

This document focuses solely on traffic destined to the network infrastructure itself to protect the network from denial of service and other attacks. This document presents techniques that, together with network edge ingress filtering and RFC 2267 and RFC 3704, create a layered approach for infrastructure protection.

This document does not present recommendations for protocol validation (i.e. "sanity checking") nor does it address guidelines for general security configuration.

Table of Contents

1. Introduction

This RFC defines best current practices for implementing Service Provider network infrastructure protection for network elements. RFC 2267 and RFC 3704 focuses on limiting the effects of denial of service attacks by filtering ingress packets with spoofed source addresses, which in turn offers other benefits by ensuring that packets coming into a network originate from validly allocated and consistent sources. RFC 3704 extends the recommendations described in RFC 2267 to address operational filtering guidelines for single and multi-homed environments. In both cases (RFC 2267 and RFC 3704), the focus is on dropping packets on ingress, regardless of destination, if those packets are have spoofed source address or fall within "reserved" address space. This document both refines and extends the filtering best practices outlined in RFC 2267 and RFC 3704 and focuses only on traffic destined to the network infrastructure itself to protect the service provider network from denial of service and other attacks. This document presents techniques that, together with network edge ingress filtering and RFC 2267 and RFC 3704, create a layered approach for infrastructure protection.

Denial of Service (DoS) attacks are common and the network infrastructure itself is a target. Attacks targeting the network infrastructure can take many forms, ranging from bandwidth saturation to crafted packets destined to a router. These attacks might use spoofed source address or they might use the true address of source of the traffic. Regardless of the nature of the attack, the network infrastructure must be protected from both accidental and intentional attacks.

The techniques outlined in this document and described in section 2 below, provide a layered approach for infrastructure protection: Edge policy (filtering and precedence), per device traffic policy enforcement for packets destined to a device and finally, routing/address advertising best practices to limit core network -- that is P and PE infrastructure -- exposure.

This document is aimed at network operators who would like to "harden" their infrastructure and make it more resilient to external attack. These techniques are designed to be used in addition to specific protocol or application security features implemented in network devices.

Infrastructure protection is a complex topic, improving protection is always beneficial.

2. Overview of Infrastructure Protection Techniques

This section provides an overview of the four recommended techniques that may be used to protect network infrastructure. The details of each area along with some deployment consideration are described in detail in subsequent sections.

- Edge Infrastructure Access Control List
- Edge Remarking
- Device and Element Protection
- Infrastructure Hiding

The above list is not exhaustive; other mechanisms can be used to provide a measure of protection. The techniques discussed in this document have been widely deployment and have proven operational security benefits in large networks.

2.1. Edge Infrastructure Access Control Lists

Edge infrastructure access control lists are ingress access control lists that filter traffic destined to the network only. They should permit all traffic through the network. Explicit filtering of traffic destined to network devices creates a first level of protection at the network edge: only traffic explicitly permitted into the network can reach a device beyond the PE router with the filter.

Although very effective, edge infrastructure access control lists are not perfect and, like any filter lists, must be maintained and updated. Furthermore, while widespread deployment on ingress interface provides the most protection (which in some cases will not be possible), some deployment is better than no deployment.

2.2. Edge Remarking

We define Edge Remarking as ensuring that ingress IP precedence or DSCP values match expected values within the context of security. This provides another layer of defense particularly for traffic permitted through any of the Edge Infrastructure Access Control

Lists. In this RFC we focus only on using Edge Remarking best practices to enforce security policies.

2.3. Device and Element Protection

Each device infrastructure device should enforce local rules for traffic destined to the device itself. These rules can take the form of filters (permit/deny) or rate limiting rules that allow ingress traffic at specified rates. These should complement any existing Edge Infrastructure Access Control Lists.

The deployment of these local device protection rules compliments the edge techniques by protecting the device from traffic that: i) was permitted but violates device policy, ii) could not be filtered at the edge, iii) entered the network on an interface that did not have ingress filtering enabled.

2.4. Infrastructure Hiding

Hiding the infrastructure of the network provides an elegant mechanism for protecting the network infrastructure. If the destination of an attack is to an infrastructure address that is unreachable, attacks become far more difficult. Infrastructure hiding can be achieved in several ways:

- MPLS techniques
- IGP configuration
- Route advertisement control

3. Edge Infrastructure Access Control Lists

Edge Infrastructure Access Control Lists (EIACLs) are a specific implementation of the more general Ingress Access List. As opposed to generic ingress filtering which denies data (sometimes referred to as user) plane traffic, edge infrastructure access control lists do

not attempt to deny traffic going through the devices, rather this form of access control limits traffic destined to infrastructure equipment while permitting -- if needed, explicitly -- traffic through the network.

3.1. Constructing the Access List

Edge Infrastructure Access Control Lists permit only required traffic to the network infrastructure, while allowing data plane traffic to flow through unaffected. The basic premise of EIACLs is that only a relatively limited subset of traffic, sourced from outside your AS, needs to be destined for a core router and that by explicitly permitting only that known and understood traffic, the core devices are not subjected to unnecessary traffic that might result in a denial of service attack.

Since edge infrastructure access control lists protect only the infrastructure, the development of the list differs somewhat from "traditional" access filter lists:

1. Review addressing scheme, and identify address block(s) that represent core devices.

2. Determine what traffic must be destined to the core devices from outside the AS.

3. Create a filter that allow the required traffic, denies all traffic destined to the core address block and then finally, permits all other traffic to all.

As with other ingress filtering techniques, EIACLs are applied on ingress into the network, and clearly comprehensive coverage (i.e. on as many interface as possible) yields the most protection.

3.2. Other Traffic

In addition to the explicitly permitted traffic, EIACLs can be combined with other common edge filters such as:

1. Source spoof prevention (as per RFC 3704) by denying internal AS addresses as external sources.

2. Filtering of reserved addresses (e.g. rfc1918 addresses) as traffic should not be sourced from reserved address.
3. Other unneeded or unnecessary traffic

Filtering this traffic can be part of the list explicitly or implicitly, however, explicit filters often provides log-able information that can be of use during a security event.

3.3. Edge Infrastructure Conclusion

Edge Infrastructure Access Control Lists provide a very effective first line of defense. To deploy them effectively, core address space must be identifiable and widespread deployment is necessary.

4. Edge Rewrite/Remarking

RFC 1812 section 5.3 defines the use of IP Preference in IPv4 packets for routing, management and control traffic. In addition it recommends devices use a mechanism for providing preferential forwarding for packets marked as routing, management or control traffic using IP Preference bits 6 or 7 (110 or 111 in binary.) RFC2474 defines DSCP and the compatibility of IP Preference bits when using DSCP.

All packets received from the Customer edge (CE,) and the Peer Edge by the Provider Edge (PE,) with IP Preference values of 6 or 7 or DSCP bits of 11xxxx, as specified in RFC2474 Differentiated Services Field Definition, should have the IP Preference bits rewritten. Routing traffic received from the CE and the Peer Edge can safely have the IP Preference bits rewritten, because only a limited number of protocols are transmitted beyond the first PE router. The bits may be rewritten to any value other than IP Preference values 6 or 7, or any DSCP value other than 11xxxx. The new value can be based on the network operators IP Preference or DSCP policy. If no policy exists the bits should be rewritten to 0.

Providers may not want to modify traffic that goes through their network in an effort offer a fully transparent service. If the provider relies on alternative means of classifying traffic for prioritized forwarding rewriting the IP Preference bits is not

required. Alternatives include encapsulating customer traffic into a second protocol, such as MPLS, GRE, and IP, or using an Access Control List (ACL) to classify legitimate routing, management, and control traffic. When encapsulating traffic into a second protocol, policy must ensure that IP Preference bits 6 and 7 are not transferred to the preference field of encapsulating protocol. In this example the EXP bits or IP Preference/DSCP bits. A longer tuple used for identifying routing, management and control traffic will provide a higher level of security than a shorter one. Other techniques may exist not covered in this document.

4.1. Edge Rewriting/Remarking Discussion

By default router vendors do not differentiate an interface on a PE router connected to a P router from an interface connected to a CE router. As a result any packet with the proper IP Preference or DSCP bits set may receive the same preferential forwarding behavior as legitimate routing, management, and control traffic. A malicious attack may be able to take advantage of the vulnerability to increase the effectiveness of the attack or to attack the routing, management, and/or control traffic directly.

This document is aimed at protecting network infrastructure from traffic to the device rather than traffic through the device. Even though the edge rewrite/remarking deals primarily with traffic through a device it is included because the traffic has a direct impact on traffic to a device. The forwarding prioritization given to routing, management, and control traffic by default leaves devices vulnerable to indirect attacks to the core infrastructure.

5. Device/Element Protection

Even with the widest possible deployment of the techniques described above in the section Infrastructure Edge Access Control, the individual devices of the network must implement access control mechanisms. This is because in addition to the case of incomplete or imperfect deployment of edge infrastructure control, threats may occur from trusted sources within the perimeter of the network.

5.1. Service Specific Access Control

Typically these mechanisms are not concerned with protecting the system as a whole, but the service from exploitation. The goal is not overall system availability, but maximizing the security of the particular service.

5.1.1. Common Services

While each service implemented by network equipment manufacturers differs in its available security features there are some common services and security features for those services that have been widely deployed.

The most important first step for the operator is to disable any unneeded/unused services.

Second, the operator should utilize the services access control mechanisms to limit the access to the devices service to only required sources. Examples are using virtual terminal access control lists, or SNMP Community access control lists.

5.2. Aggregate Device Access Control

The device must be protected from denial of service threats, in addition, aggregating the security policy allows for a simplified view of the access policies traffic going to the device.

A key requirement of these mechanisms is that it must not impact transit data plane traffic. In addition, these mechanisms should not make the device more vulnerable to malicious traffic than not using them.

5.2.1. IP Fragments

Traffic destined to a router is not typically fragmented. Fragment keywords or other mechanisms to deny fragments to the device are recommended.

5.2.2. Performance Considerations

Care should be taken to understand a vendors implementation of this functionality and to make sure that device operation is not impaired during DoS attacks against the device.

5.2.3. Access Control Implementation Guide

Implementing a complex set of access controls for all traffic going to and from a router is non trivial. The following is a recommended set of steps that has been used successfully by many carriers.

-Develop list of required protocols
-Develop source address requirements
-Determine destination interface on router
Does the protocol access a single interface?
Does the protocol access many interfaces?
Does the protocol access a virtual or physical interfaces?
-Deployment should be an iterative process
-Start with relatively open lists then tighten as needed

5.3. Device Access Authorization and Accounting

Operators should use per command authorization and accounting wherever possible. Aside from their utility in mitigating other security threats, they provide an invaluable tool in the post event forensics.

6. Infrastructure Hiding

Hiding the infrastructure of the network provides an elegant mechanism for protecting the network infrastructure. If the destination of an attack is to an infrastructure address that is unreachable, attacks become far more difficult. Infrastructure hiding can be achieved in several ways:

6.1. Use less IP

One way to reduce exposure of network infrastructure is to use unnumbered links wherever possible. This is particularly useful for customers in the simple case of a single provider with a default path to the Internet.

6.2. MPLS techniques

While it may not be feasible to hide the entire infrastructure of large networks from edge to edge using MPLS, it is certainly possible to reduce exposure of critical core infrastructure beyond the first hop by creating an MPLS mesh where TTL is not decremented as packets pass through it. In this manner the number, addresses, and even existence of intermediary devices can be hidden from traffic as it passes through the core.

6.3. IGP configuration

Using a non-IP control plane for the core routing protocol can substantially reduce the number of IP addresses that [comprise/expose] the core. This simplifies the task of maintaining edge ACLs or route announcement filters. IS-IS is an elegant and mature protocol that may be suitable for this task.

6.4. Route advertisement control

6.4.1. Route Announcement filtering

Inasmuch as it is unavoidable that some network elements must be configured with IP addresses, it may be possible to assign these address out of netblocks for which the routing advertisement can be filtered, thereby limiting possible sources of traffic to core netblocks down to customers for which you provide a default route, or direct peers who would make the effort to create a static route for your core netblock into your AS.

Further, it may be possible in those situations where customer point-to-point links must be numbered, to address such links out of another range of addresses for which announcements could be similarly filtered. While this has implications for a customer's ability to remote-monitor their circuit, this can often be overcome with application of an address from the customer's routed space to the CPE loopback.

6.4.2. Address core out of rfc1918 space

In addition to filtering the visibility of core addresses to the wider Internet, it may be possible to use rfc1918 netblocks for numbering infrastructure when IP addresses are required (eg, loopbacks). This added level of obscurity takes prevention of wide distribution of your infrastructure address space one step further. Many networks filter out packets with rfc1918 address at ingress/egress points as a matter of course. In this circumstance, tools such as traceroute can work through your core, but reverse-resolution of descriptive names should be restricted to queries from internal/support groups.

7. IPv6

IPv6 Networks contain the same infrastructure security risks as IPv4.

All techniques described in this document for IPv4 should be directly applicable to IPv6 networks. Limitations exist where devices do not have feature parity between IPv4 and IPv6. Different techniques maybe required where IPv4 and IPv6 networks deviate in implementation. Multi-vendor networks create greater difficulties when each vendor does not have feature parity with each other.

Hardware differences in devices that support both IPv4 and IPv6 must also be taken into consideration. Because IPv6 uses a longer address space the scaling, and performance characteristics of ACLs maybe lower for IPv6 vs IPv4. The fields or number of fields that an ACL can match on may also differ.

The fact that all PE devices do not support all the recommended ipv6 security features should not preclude the implementation of the recommendations in this document on the devices that do support the security features.

With the number of Network Operators deploying IPv6 growing, along with the continued availability of IPv6 Tunnel services, connecting to the IPv6 internet is less difficult. Dual stack IPv6 networks run on 10Gbps and greater backbones with edge speeds equal to IPv4. Neither the edge nor the core limit potential IPv6 attacks.

7.1. IPv6 Edge Infrastructure Access Control List

The same process should be used for constructing the IPv6 eiacl as the IPv4 eiacl.

7.2. IPv6 Edge Remarking

IPv6 DSCP bits should be rewritten in the same manner that IPv4 DSCP bits.

7.3. IPv6 Device and Element Protection

IPv6 device and element protection should be implemented using the same policy as IPv4.

7.4. IPv6 Infrastructure Hiding

Network operators may deploy IPv4 differently from IPv6 in their network. Providers may use native forwarding for IPv6 while using MPLS for IPv4, other combinations. IPv6 infrastructure hiding should have parity with IPv4 infrastructure hiding even if the technique used is different.

Implementation of IPv6 route advertisement control for infrastructure hiding is difficult when using global address space. It is difficult to get non-continuous network blocks from the address registries, and de-aggregation of IPv6 address space is not an acceptable alternative. It is still possible to use private address space as a way of restricting IPv6 advertisements.

8. IP Multicast

IP Multicast behaves differently from IP unicast therefore must be secured in a different manner. Some of the protocols used with Multicast rely on IP unicast to transport the routing, and control information. Unicast based protocols should be secured using the technique described in much of this document. Because this document is focused on hardening a service providers infrastructure rather than validating routing announcements, much of IP Multicast filtering will be better covered in other documents.

In much the same way a host must listen on a certain IP address and port for an IP unicast connection, Multicast must join a group in order to receive any information via Multicast. The major difference is that multicast groups are global and not assigned to a specific customer or end user. Administrative boundaries and scope are created to isolate Multicast groups within one network or desired area.

8.1. Multicast Group Protection

Certain Multicast groups should never been joined from outside an operators network or administrative boundary. Filters should be placed on the protocols used to communicate with external hosts and networks. IGMP should have a join filter to prevent hosts from joining internal groups. MSDP should be configured with a Source Address (SA) filter to prevent other networks from joining internal groups.

EIACLs should include administratively bounded multicast groups, along with any groups used for protocols internal to a providers network.

When constructing router Access Control as described in section 5.2.4, multicast protocols must be taken into consideration.

8.2. Performance Considerations

Multicast protocols and implementation have different performance and scaling limitation than IP unicast. Multicast users create state on the router every time the user joins a group. Router resources can be exhausted if the amount of state created exceeds the resources available on the router. Placing limits on the resources used by the Multicast protocols can prevent collateral damage to services other than Multicast on a router. MSDP should have a limit placed on the number of SA announcements received. A fixed limit should be placed on the number of entries the router stores in the IP Multicast routing table. The number of SAP entries should have a limit placed on them.

8.3. IPv6 and Multicast

IPv6 Multicast policy should be consistent with the IP Multicast policy. 9.0 Security Considerations

9. Acknowledgments

10. References

10.1. Normative References

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

[REF] Reference....

10.2. Informative References

[RFC3667] Bradner, S., "IETF Rights in Contributions", BCP 78, RFC 3667, February, 2004.

[RFC3668] Bradner, S., "Intellectual Property Rights in IETF Technology", BCP 79, RFC 3668, February, 2004.

[RFC2434] Narten, T., and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

11. Authors' Addresses

James Gill
TBD

Darrel Lewis
Cisco Systems Inc.
170 West Tasman Drive
San Jose, CA 95134
Phone: +1 408 853 3653
EMail: darlewis@cisco.com

Paul Quinn
170 West Tasman Drive
San Jose, CA 95134
Phone: +1 408 527 3560
Email: paulq@cisco.com

Peter Schoenmaker
NTT America
101 Park Ave., FL 41
New York, NY 10178
+1-212-808-2298
pds@ntt.net

Intellectual Property Statement

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.