Opsec WG
Internet-Draft
Intended status: Informational
Expires: December 14, 2006

P. Savola
CSC/FUNET

June 12, 2006

Experiences from Using Unicast RPF
draft-savola-bcp84-urpf-experiences-01.txt

Status of this Memo

Copyright Notice

Abstract

RFC 3704 (BCP 84) published in March 2004 provided an ingress filtering technique update to RFC 2827 (BCP 38). This memo tries to document operational experiences learned practising ingress filtering techniques, in particular ingress filtering for multihomed networks.

Table of Contents

1. Introduction

RFC 3704 [RFC3704] (BCP 84) published in March 2004 provided an ingress filtering technique update to RFC 2827 [RFC2827] (BCP 38). This memo tries to document operational experiences learned practising ingress filtering techniques, in particular ingress filtering for multihomed networks.

Specifically, this version describes the lessons learned in author's network where strict unicast RPF (uRPF) ingress filtering, using "feasible paths" variant [RFC3704] has been used for all the customer interfaces (whether single- or multihomed) for over two years. In feasible paths strict uRPF, only an accepted equal length prefix (even if not preferred) is considered feasible. While in some cases, a more specific or even a less specific might be acceptable, such condition would not necessarily be correct in general.

We use the typical "customer" and "ISP" terms to refer to the subject of strict uRPF filtering and the party doing filtering. The same considerations also apply for other business relationships (e.g., "internal customers" inside an ISP).

According to a study, there is substantial ingress filtering deployment, even 75% of addresses were not spoofable [SPOOFER].

We note explicitly that Loose mode RPF is NOT a sufficient solution in any way to ingress filtering as it creates a false sense of protection. Even its use as a "contract validation" [RFC3704] is tenuous at best.

NOTE IN DRAFT: comments should be directed to the author or the OPSEC mailing list (opsec@ops.ietf.org). However, it is not clear what should be the next steps wrt. these experiences. Update to the ingress filtering RFCs? Publish separately? Keep as a standing document for now? Integrate with OPSEC document work? In any case, feedback on other experiences is encouraged.

In the second section, we'll first look at the most common types of uRPF failures and their mitigation techniques. In the third section, we'll look at a few special cases observed on multihoming or multi-connecting scenarios. More special filtering failures are discussed in the fourth section.

2. Common uRPF Failures

We'll describe the most common uRPF failures which apply to both single- and multi-homed network, and respective fixes.

2.1. Unused Address Space Ping-Pong

By far, the most common cause for uRPF failures seems to be the case where a prefix P is routed to the customer (e.g., using a static route), but the customer doesn't use all of P, and an attacker A is port-scanning the unused address space.

In that case, typically packets destined to the unused part of "P" lack a more specific route, and are routed back to the ISP through a default route. The ISP's router sees these as sourced from attacker A (an IP address in the Internet), destined to the customer's prefix P. This fails uRPF check and is dropped.

Note: if uRPF is not employed, the scan may may cause ping-pong effect up to the remaining hop count/TTL of the packet, consuming even 250 times the bandwidth and packet processing. This has been briefly described in [I-D.ietf-ipngwg-p2p-pingpong].

The ping-pong effect has also been used in Internet Exchanges to game peer selection or traffic balance data.

Therefore, the customer should install static discard aggregate routes (or equivalent) for all of its address space upon assignment, so that if no better route exists, such probe packets are discarded. An alternative is applying a similar filtering in egress interface towards the ISP. There isn't much an ISP can do to prevent this unless it wants to create customer-specific uRPF access-lists.

2.2. Private Address Leak

Very often, packes from all kinds of private addresses also leak to the ISP, which are obviously dropped by uRPF. This is probably a result of misconfigured NATs or inadequate firewall rules. Even (constant) rates of hundreds of packets per second have been observed, which makes one wonder which kinds of users' communications must be failing or otherwise working in a non-optimal fashion due to this kind of misconfiguration...

This is actually one of the most convincing reasons from the users' perspective why (they or the ISP) using uRPF could give benefits: it allows them to notice and fix network misconfiguration and malfunction "at the source" and as a result, communication should work more reliably and new issues would be easier to notice.

The obvious fix is to ensure that the customer is filtering out (and logging) these packets, and based on that, figures out what is causing such address leaks and fixes the misconfiguration or other problem(s).

2.3. Wrong IP Address

It's also not atypical to see other kinds of wrong source addresses. These can be classified in three main categories: a) nomadic laptops trying their old IP from a previous network attachment point, b) spoofed/misconfigured/typoed public, routable IP address, or c) an unroutable ("bogon") IP address. (It should be noted that Loose uRPF would only spot the last category.)

Many spoofed attacks are usually a result of a worm or a botnet (DoS) attack. A recent case was using recursive DNS servers for reflection [I-D.ietf-dnsop-reflectors-are-evil], but a lot of different usages have been observed.

The same considerations as for leaking private addresses apply here, except that these wouldn't typically get this far if the customer had been using unicast RPF at its LAN interfaces (i.e., uRPF can and should be applied recursively [RFC3704]).

3. Multihoming uRPF Failures

We'll describe a few uRPF failure modes which only occur in scenarios with a multihomed/multi-connected network or host.

We note that a customer can multihome and even perform traffic engineering with feasible paths uRPF provided that the consistency requirement is fulfilled. In other words, AS-path prepending, setting communities to lower local-preference, etc. are all valid mechanisms to ensure the prefix is advertised to every provider, but actually may not ever end up being used.

3.1. Incorrect Source Address Selection

Hosts attaching to multiple LANs with different IP address need to be careful with their source address selection. The same applies to networks with multiple prefixes as explored in [I-D.huitema-shim6-ingress-filtering].

For example, assume the host has a default route through interface 1 with address A1 from prefix P1, and only a more specific route through interface 2 with address A2 from prefix P2. When a host in P1 sends a packet to A2, the response may go out through interface 1; similarly, when a host in P2 sends a packet to A1, the response may go out through interface 2.

This problem can be fixed by the customers by setting up source-based routing so packets go through the right route, or by making an

exclusion in the uRPF filter list to allow sourcing from the other prefix. The latter is typically not a good solution, especially if the ISP doesn't control both the prefixes, because an ISP originating these excluded packets would be indistinguishable from IP address spoofing.

3.2. Point-to-Point Interface Routes

Feasible path strict uRPF works well, but assumes that the routes in all the directions are consistent (i.e., exist). This principle is often violated with the interface routes between the ISP and the customer (ie., point-to-point links).

In some cases, the point-to-point link may be unnumbered but this has other issues (e.g., eBGP is more complicated). If the links have

addresses, the address blocks usually need to be separate. The addresses might be more specifics of the customer's aggregate(s) or from the ISP's address space. In either case, the similar source address selection issue as described in the previous section applies for communication (e.g., pinging the CPE's p2p address) to the customer's point-to-point addresses.

The easiest fix is to add dummy static routes with a higher preference/distance on all the border routers, so that every router facing the customer knows all the point-to-point address blocks used on other routers; using a higher preference implies that the route is actually never used, but is still valid from uRPF perspective. Another possibility, if the addresses come from the customer's aggregate, is to not propagate the point-to-point addresses in iBGP or IGP at all so that there are no more specifics to mess up the uRPF feasible path consistency, but this may have manageability concerns if the aggregate goes down (i.e., can't ping the point-to-point address except on the router connecting the customer). As already mentioned, using unnumbered interfaces is also possible in some cases but may have manageability or configuration concerns.

3.3. Multiple Routers on a LAN use LAN for Transit

When multiple routers attach to the same network subnet (typically when e.g., VRRP is used), packets destined to router 2 (R2)'s interface addresses towards the LAN transiting router 1 use the LAN interface to reach R2. (In most cases, the primary path between routers should go via dedicated link(s), not via a LAN.) These packets fail uRPF check at R2 (and vice versa at R1).

There are two obvious fixes: have R2 advertise such LAN addresses in iBGP or IGP (or set up static routes), resulting a more specific so the LAN interface is not used, or make an exception to uRPF

configuration to allow such "transit LAN" usage. However, the latter allows an attacker in the LAN to spoof an address to the LAN router's interface address(es) (for example, circumventing remote login access lists), which usually makes it a suboptimal solution.

4. Special uRPF Failures Cases

4.1. PMTUD and Private/Non-routed Addresses

A disturbing issue is that some large operators seem to think it's perfectly legitimate to send private-source addressed packets ICMP messages (e.g., from PMTUD) across AS boundaries [PRIVIP]. While the reasoning is different, the result is similar for non-routed, but uniquely assigned address space.

Private IP addresses for infrastructure are a bad idea. But even worse than that is deploying links in such infrastructure which have lower MTU than the egress link, i.e., are guaranteed to send ICMP fragmentation needed messages under certain circumstances. Deploying such networks that require PMTUD to work while happily originating RFC1918 traffic (and translating it at the edge) seems like very bad design from network hygiene perspective.

5. IANA Considerations

This memo makes no request to IANA.

6. Acknowledgements

Danny McPherson and Matsuzaki Yoshinobu provided comments on the first revision of this document.

7. Security Considerations

This document describes uRPF experiences. The most important security impact comes from applying particular fixes to uRPF issues noted, i.e., what kind of spoofing window or other unintended usage that would allow.

As already stated, in invalid source address selection scenario, making an exception to allow prefixes which you don't control is typically a big mistake, as then you become indistinguishable from someone spoofing that address. Also as already stated, in the case of transit LAN, making an exception might allow one to spoof an

address destined to the LAN router's interface address(es) which

usually has a security impact.

8. References

8.1. Normative References

[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.

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

8.2. Informative References

[I-D.huitema-shim6-ingress-filtering] Huitema, C., "Ingress filtering compatibility for IPv6 multihomed sites", draft-huitema-shim6-ingress-filtering-00 (work in progress), September 2005.

[I-D.ietf-dnsop-reflectors-are-evil] Damas, J. and F. Neves, "Preventing Use of Nameservers in Reflector Attacks", draft-ietf-dnsop-reflectors-are-evil-00 (work in progress), May 2006.

[I-D.ietf-ipngwg-p2p-pingpong] Hagino, J., JINMEI, T., and B. Zill, "Avoiding ping-pong packets on point-to-point links", draft-ietf-ipngwg-p2p-pingpong-00 (work in progress), July 2001.

[PRIVIP] NANOG mailing-list thread, "private IP addresses from ISP", May 2006, <http://www.merit.edu/mail.archives/nanog/msg00279.html>.

[SPOOFER] MIT ANA, "Spoofer Project", <http://spoofer.csail.mit.edu>.

Author's Address

Pekka Savola
CSC/FUNET
Espoo
Finland

Email: psavola@funet.fi

Full Copyright Statement

Intellectual Property

ietf-ipr@ietf.org.

Acknowledgment

Funding for the RFC Editor function is provided by the IETF Administrative Support Activity (IASA).