Internet Engineering Task Force (IETF)
Request for Comments: 6752
Category: Informational
ISSN: 2070-1721

A. Kirkham
Palo Alto Networks

September 2012

Issues with Private IP Addressing in the Internet

Abstract

The purpose of this document is to provide a discussion of the potential problems of using private, RFC 1918, or non-globally routable addressing within the core of a Service Provider (SP) network. The discussion focuses on link addresses and, to a small extent, loopback addresses. While many of the issues are well recognised within the ISP community, there appears to be no document that collectively describes the issues.

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/rfc6752.

Table of Contents

1. Introduction

In the mid to late 1990s, some Internet Service Providers (ISPs) adopted the practice of utilising private (or non-globally unique) [RFC1918] IP addresses for the infrastructure links and in some cases the loopback interfaces within their networks. The reasons for this approach centered on conservation of address space (i.e., scarcity of public IPv4 address space) and security of the core network (also known as core hiding).

However, a number of technical and operational issues occurred as a result of using private (or non-globally unique) IP addresses, and virtually all these ISPs moved away from the practice. Tier 1 ISPs are considered the benchmark of the industry and as of the time of writing, there is no known tier 1 ISP that utilises the practice of private addressing within their core network.

The following sections will discuss the various issues associated with deploying private [RFC1918] IP addresses within ISP core networks.

The intent of this document is not to suggest that private IP addresses can not be used with the core of an SP network, as some providers use this practice and operate successfully. The intent is to outline the potential issues or effects of such a practice.

Note: The practice of ISPs using "squat" address space (also known as "stolen" space) has many of the same, plus some additional, issues (or effects) as that of using private IP address space within core networks. The term "squat IP address space" refers to the practice

of an ISP using address space for its own infrastructure/core network addressing that has been officially allocated by an RIR (Regional Internet Registry) to another provider, but that provider is not currently using or advertising within the Internet. Squat addressing is not discussed further in this document. It is simply noted as an associated issue.

2. Conservation of Address Space

One of the original intents for the use of private IP addressing within an ISP core was the conservation of IP address space. When an ISP is allocated a block of public IP addresses (from an RIR), this address block was traditionally split in order to dedicate some portion for infrastructure use (i.e., for the core network) and the other portion for customer (subscriber) or other address pool use. Typically, the number of infrastructure addresses needed is relatively small in comparison to the total address count. So unless the ISP was only granted a small public block, dedicating some portion to infrastructure links and loopback addresses (/32) is rarely a large enough issue to outweigh the problems that are potentially caused when private address space is used.

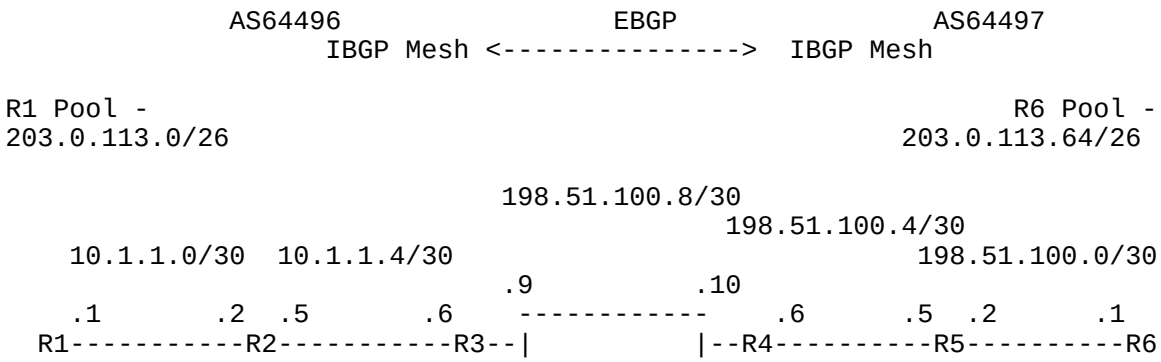
Additionally, specifications and equipment capability improvements now allow for the use of /31 subnets [RFC3021] for link addresses in place of the original /30 subnets -- further minimising the impact of dedicating public addresses to infrastructure links by only using two (2) IP addresses per point-to-point link versus four (4), respectively.

The use of private addressing as a conservation technique within an Internet Service Provider (ISP) core can cause a number of technical and operational issues or effects. The main effects are described below.

3. Effects on Traceroute

The single biggest effect caused by the use of private addressing [RFC1918] within an Internet core is the fact that it can disrupt the operation of traceroute in some situations. This section provides some examples of the issues that can occur.

A first example illustrates the situation where the traceroute crosses an Autonomous System (AS) boundary, and one of the networks has utilised private addressing. The following simple network is used to show the effects.



R1 Loopback: 10.1.1.101
R2 Loopback: 10.1.1.102
R3 Loopback: 10.1.1.103
R4 Loopback: 198.51.100.103
R5 Loopback: 198.51.100.102
R6 Loopback: 198.51.100.101

Using this example, performing the traceroute from AS64497 to AS64496, we can see the private addresses of the infrastructure links in AS64496 are returned.

```
R6#traceroute 203.0.113.1
Type escape sequence to abort.
Tracing the route to 203.0.113.1

  1 198.51.100.2 40 msec 20 msec 32 msec
  2 198.51.100.6 16 msec 20 msec 20 msec
  3 198.51.100.9 20 msec 20 msec 32 msec
  4 10.1.1.5 20 msec 20 msec 20 msec
  5 10.1.1.1 20 msec 20 msec 20 msec
R6#
```

This effect in itself is often not a problem. However, if anti-spoofing controls are applied at network perimeters, then responses returned from hops with private IP addresses will be dropped. Anti-spoofing refers to a security control where traffic with an invalid source address is discarded. Anti-spoofing is further described in [BCP38] and [BCP84]. Additionally, any [RFC1918] filtering mechanism, such as those employed in most firewalls and many other network devices can cause the same effect.

The effects are illustrated in a second example below. The same network as in example 1 is used, but with the addition of anti-spoofing deployed at the ingress of R4 on the R3-R4 interface (IP Address 198.51.100.10).

```
R6#traceroute 203.0.113.1

Type escape sequence to abort.
Tracing the route to 203.0.113.1

  1 198.51.100.2 24 msec 20 msec 20 msec
  2 198.51.100.6 20 msec 52 msec 44 msec
  3 198.51.100.9 44 msec 20 msec 32 msec
  4  *  *  *
  5  *  *  *
  6  *  *  *
  7  *  *  *
  8  *  *  *
  9  *  *  *
 10  *  *  *
 11  *  *  *
 12  *  *  *
```

In a third example, a similar effect is caused. If a traceroute is initiated from a router with a private (source) IP address, located in AS64496 and the destination is outside of the ISP's AS (AS64497), then in this situation, the traceroute will fail completely beyond the AS boundary.

```
R1# traceroute 203.0.113.65
Type escape sequence to abort.
Tracing the route to 203.0.113.65

  1 10.1.1.2 20 msec 20 msec 20 msec
  2 10.1.1.6 52 msec 24 msec 40 msec
  3  *  *  *
  4  *  *  *
  5  *  *  *
  6  *  *  *
R1#
```

While it is completely unreasonable to expect a packet with a private source address to be successfully returned in a typical SP environment, the case is included to show the effect as it can have implications for troubleshooting. This case will be referenced in a later section.

In a complex topology, with multiple paths and exit points, the provider will lose its ability to trace paths originating within its own AS, through its network, to destinations within other ASes. Such a situation could be a severe troubleshooting impediment.

For completeness, a fourth example is included to show that a successful traceroute can be achieved by specifying a public source address as the source address of the traceroute. Such an approach can be used in many operational situations if the router initiating the traceroute has at least one public address configured. However, the approach is more cumbersome.

```
R1#traceroute
Protocol [ip]:
Target IP address: 203.0.113.65
Source address: 203.0.113.1
Numeric display [n]:
Timeout in seconds [3]:
Probe count [3]:
Minimum Time to Live [1]:
Maximum Time to Live [30]: 10
Port Number [33434]:
Loose, Strict, Record, Timestamp, Verbose[none]:
Type escape sequence to abort.
Tracing the route to 203.0.113.65

  1 10.1.1.2 0 msec 4 msec 0 msec
  2 10.1.1.6 0 msec 4 msec 0 msec
  3 198.51.100.10 [AS 64497] 0 msec 4 msec 0 msec
  4 198.51.100.5 [AS 64497] 0 msec 0 msec 4 msec
  5 198.51.100.1 [AS 64497] 0 msec 0 msec 4 msec
R1#
```

It should be noted that some solutions to this problem have been proposed in [RFC5837], which provides extensions to ICMP to allow the identification of interfaces and their components by any combination of the following: ifIndex, IPv4 address, IPv6 address, name, and MTU. However, at the time of this writing, little or no deployment was known to be in place.

4. Effects on Path MTU Discovery

The Path MTU Discovery (PMTUD) process was designed to allow hosts to make an accurate assessment of the maximum packet size that can be sent across a path without fragmentation. Path MTU Discovery is utilised by IPv4 [RFC1191], IPv6 [RFC1981], and some tunnelling protocols such as Generic Routing Encapsulation (GRE) and IPsec.

The PMTUD mechanism requires that an intermediate router can reply to the source address of an IP packet with an ICMP reply that uses the router’s interface address. If the router’s interface address is a private IP address, then this ICMP reply packet may be discarded due to unicast reverse path forwarding (uRPF) or ingress filtering,

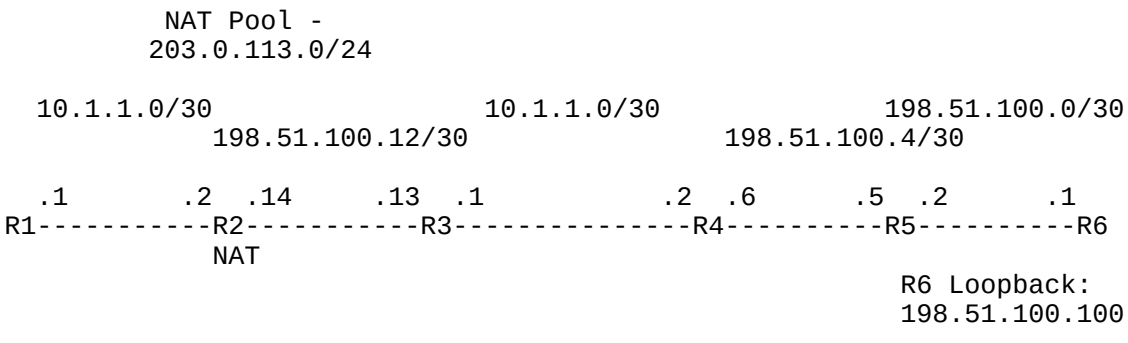
thereby causing the PMTUD mechanism to fail. If the PMTUD mechanism fails, this will cause transmission of data between the two hosts to fail silently due to the traffic being black-holed. As a result, the potential for application-level issues may be created.

5. Unexpected Interactions with Some NAT Implementations

Private addressing is legitimately used within many enterprise, corporate, or government networks for internal network addressing. When users on the inside of the network require Internet access, they will typically connect through a perimeter router, firewall, or network proxy that provides Network Address Translation (NAT) or Network Address Port Translation (NAPT) services to a public interface.

Scarcity of public IPv4 addresses is forcing many service providers to make use of NAT. CGN (Carrier-Grade NAT) will enable service providers to assign private [RFC1918] IPv4 addresses to their customers rather than public, globally unique IPv4 addresses. NAT444 will make use of a double NAT process.

Unpredictable or confusing interactions could occur if traffic such as traceroute, PMTUD, and possibly other applications were launched from the NAT IPv4 'inside address', and it passed over the same address range in the public IP core. While such a situation would be unlikely to occur if the NAT pools and the private infrastructure addressing were under the same administration, such a situation could occur in the more typical situation of a NATed corporate network connecting to an ISP. For example, say 10.1.1.0/24 is used to internally number the corporate network. A traceroute or PMTUD request is initiated inside the corporate network from say 10.1.1.1. The packet passes through a NAT (or NAPT) gateway, then over an ISP core numbered from the same range. When the responses are delivered back to the originator, the returned packets from the privately addressed part of the ISP core could have an identical source and destination address of 10.1.1.1.




```
R1#traceroute 198.51.100.100

Type escape sequence to abort.
Tracing the route to 198.51.100.100

  1 10.1.1.2 0 msec 0 msec 0 msec
  2 198.51.100.13 0 msec 4 msec 0 msec
  3 10.1.1.2 0 msec 4 msec 0 msec        <<<<
  4 198.51.100.5 4 msec 0 msec 4 msec
  5 198.51.100.1 0 msec 0 msec 0 msec
R1#
```

This duplicate address space scenario has the potential to cause confusion among operational staff, thereby making it more difficult to successfully debug networking problems.

Certainly a scenario where the same [RFC1918] address space becomes utilised on both the inside and outside interfaces of a NAT/NAPT device can be problematic. For example, the same private address range is assigned by both the administrator of a corporate network and its ISP. Some applications discover the outside address of their local Customer Premises Equipment (CPE) to determine if that address is reserved for special use. Application behaviour may then be based on this determination. "IANA-Reserved IPv4 Prefix for Shared Address Space" [RFC6598] provides further analysis of this situation.

To address this scenario and others, "IANA-Reserved IPv4 Prefix for Shared Address Space" [RFC6598] allocated a dedicated /10 address block for the purpose of Shared CGN (Carrier Grade NAT) Address Space: 100.64.0.0/10. The purpose of Shared CGN Address Space is to number CPE (Customer Premise Equipment) interfaces that connect to CGN devices. As explained in [RFC6598], [RFC1918] addressing has issues when used in this deployment scenario.

6. Interactions with Edge Anti-Spoofing Techniques

Denial-of-Service (DOS) attacks and Distributed Denial-of-Service (DDoS) attacks can make use of spoofed source IP addresses in an attempt to obfuscate the source of an attack. Network Ingress Filtering [RFC2827] strongly recommends that providers of Internet connectivity implement filtering to prevent packets using source addresses outside of their legitimately assigned and advertised prefix ranges. Such filtering should also prevent packets with private source addresses from egressing the AS.

Best security practices for ISPs also strongly recommend that packets with illegitimate source addresses should be dropped at the AS perimeter. Illegitimate source addresses includes private [RFC1918] IP addresses, addresses within the provider's assigned prefix ranges, and bogons (legitimate but unassigned IP addresses). Additionally, packets with private IP destination addresses should also be dropped at the AS perimeter.

If such filtering is properly deployed, then traffic either sourced from or destined for privately addressed portions of the network should be dropped, hence the negative consequences on traceroute, PMTUD, and regular ping-type traffic.

7. Peering Using Loopbacks

Some ISPs use the loopback addresses of Autonomous System Border Routers (ASBRs) for peering, in particular, where multiple connections or exchange points exist between the two ISPs. Such a technique is used by some ISPs as the foundation of fine-grained traffic engineering and load balancing through the combination of IGP metrics and multi-hop BGP. When private or non-globally reachable addresses are used as loopback addresses, this technique is either not possible or considerably more complex to implement.

8. DNS Interaction

Many ISPs utilise their DNS to perform both forward and reverse resolution for infrastructure devices and infrastructure addresses. With a privately numbered core, the ISP itself will still have the capability to perform name resolution of its own infrastructure. However, others outside of the autonomous system will not have this capability. At best, they will get a number of unidentified [RFC1918] IP addresses returned from a traceroute.

It is also worth noting that in some cases, the reverse resolution requests may leak outside of the AS. Such a situation can add load to public DNS servers. Further information on this problem is documented in "AS112 Nameserver Operations" [RFC6304].

9. Operational and Troubleshooting Issues

Previous sections of this document have noted issues relating to network operations and troubleshooting. In particular, when private IP addressing within an ISP core is used, the ability to easily troubleshoot across the AS boundary may be limited. In some cases, this may be a serious troubleshooting impediment. In other cases, it may be solved through the use of alternative troubleshooting techniques.

The key point is that the flexibility of initiating an outbound ping or traceroute from a privately numbered section of the network is lost. In a complex topology, with multiple paths and exit points from the AS, the provider may be restricted in its ability to trace paths through the network to other ASes. Such a situation could be a severe troubleshooting impediment.

For users outside of the AS, the loss of the ability to use a traceroute for troubleshooting is very often a serious issue. As soon as many of these people see a row of "* * *" in a traceroute they often incorrectly assume that a large part of the network is down or inaccessible (e.g., behind a firewall). Operational experience in many large providers has shown that significant confusion can result.

With respect to [RFC1918] IP addresses applied as loopbacks, in this world of acquisitions, if an operator needed to merge two networks, each using the same private IP ranges, then the operator would likely need to renumber one of the two networks. In addition, assume an operator needed to compare information such as NetFlow / IP Flow Information Export (IPFIX) or syslog, between two networks using the same private IP ranges. There would likely be an issue as the unique ID in the collector is, in most cases, the source IP address of the UDP export, i.e., the loopback address.

10. Security Considerations

One of the arguments often put forward for the use of private addressing within an ISP is an improvement in the network security. It has been argued that if private addressing is used within the core, the network infrastructure becomes unreachable from outside the provider’s autonomous system, hence protecting the infrastructure. There is legitimacy to this argument. Certainly, if the core is

privately numbered and unreachable, it potentially provides a level of isolation in addition to what can be achieved with other techniques, such as infrastructure Access Control Lists (ACLs), on their own. This is especially true in the event of an ACL misconfiguration, something that does commonly occur as the result of human error.

There are three key security gaps that exist in a privately addressed IP core.

1. The approach does not protect against reflection attacks if edge anti-spoofing is not deployed. For example, if a packet with a spoofed source address corresponding to the network's infrastructure address range is sent to a host (or other device) attached to the network, that host will send its response directly to the infrastructure address. If such an attack was performed across a large number of hosts, then a successful large-scale DoS attack on the infrastructure could be achieved. This is not to say that a publicly numbered core will protect from the same attack; it won't. The key point is that a reflection attack does get around the apparent security offered in a privately addressed core.

2. Even if anti-spoofing is deployed at the AS boundary, the border routers will potentially carry routing information for the privately addressed network infrastructure. This can mean that packets with spoofed addresses, corresponding to the private infrastructure addressing, may be considered legitimate by edge anti-spoofing techniques (such as Unicast Reverse Path Forwarding - Loose Mode) and forwarded. To avoid this situation, an edge anti-spoofing algorithm (such as Unicast Reverse Path Forwarding - Strict Mode) would be required. Strict approaches can be problematic in some environments or where asymmetric traffic paths exist.

3. The approach on its own does not protect the network infrastructure from directly connected customers (i.e., within the same AS). Unless other security controls, such as access control lists (ACLs), are deployed at the ingress point of the network, customer devices will normally be able to reach, and potentially attack, both core and edge infrastructure devices.

11. Alternate Approaches to Core Network Security

Today, hardware-based ACLs, which have minimal to no performance impact, are now widespread. Applying an ACL at the AS perimeter to prevent access to the network core may be a far simpler approach and provide comparable protection to using private addressing; such a technique is known as an infrastructure ACL (iACL).

In concept, iACLs provide filtering at the edge network, which allows traffic to cross the network core but not to terminate on infrastructure addresses within the core. Proper iACL deployment will normally allow required network management traffic to be passed, such that traceroutes and PMTUD can still operate successfully. For an iACL deployment to be practical, the core network needs to have been addressed with a relatively small number of contiguous address blocks. For this reason, the technique may or may not be practical.

A second approach to preventing external access to the core is IS-IS core hiding. This technique makes use of a fundamental property of the IS-IS protocol, which allows link addresses to be removed from the routing table while still allowing loopback addresses to be resolved as next hops for BGP. The technique prevents parties outside the AS from being able to route to infrastructure addresses, while still allowing traceroutes to operate successfully. IS-IS core hiding does not have the same practical requirement for the core to be addressed from a small number of contiguous address blocks as with iACLs. From an operational and troubleshooting perspective, care must be taken to ensure that pings and traceroutes are using source and destination addresses that exist in the routing tables of all routers in the path, i.e., not hidden link addresses.

A third approach is the use of either an MPLS-based IP VPN or an MPLS-based IP Core where the 'P' routers (or Label Switch Routers) do not carry global routing information. As the core 'P' routers (or Label Switch Routers) are only switching labeled traffic, they are effectively not reachable from outside of the MPLS domain. The 'P' routers can optionally be hidden so that they do not appear in a traceroute. While this approach isolates the 'P' routers from directed attacks, it does not protect the edge routers ('PE' routers or Label Edge Routers (LERs)). Obviously, there are numerous other engineering considerations in such an approach; we simply note it as an option.

These techniques may not be suitable for every network. However, there are many circumstances where they can be used successfully without the associated effects of privately addressing the core.

12. References

12.1. Normative References

[BCP38] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", May 2000.

[BCP84] Baker, F. and P. Savola, "Ingress Filtering for Multihomed Networks", March 2004.

[RFC1191] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[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.

[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.

12.2. Informative References

[RFC3021] Retana, A., White, R., Fuller, V., and D. McPherson, "Using 31-Bit Prefixes on IPv4 Point-to-Point Links", RFC 3021, December 2000.

[RFC5837] Atlas, A., Bonica, R., Pignataro, C., Shen, N., and JR. Rivers, "Extending ICMP for Interface and Next-Hop Identification", RFC 5837, April 2010.

[RFC6304] Abley, J. and W. Maton, "AS112 Nameserver Operations", RFC 6304, July 2011.

[RFC6598] Weil, J., Kuarsingh, V., Donley, C., Liljenstolpe, C., and M. Azinger, "IANA-Reserved IPv4 Prefix for Shared Address Space", BCP 153, RFC 6598, April 2012.

Appendix A. Acknowledgments

The author would like to thank the following people for their input and review: Dan Wing (Cisco Systems), Roland Dobbins (Arbor Networks), Philip Smith (APNIC), Barry Greene (ISC), Anton Ivanov (kot-begemot.co.uk), Ryan Mcdowell (Cisco Systems), Russ White (Cisco Systems), Gregg Schudel (Cisco Systems), Michael Behringer (Cisco Systems), Stephan Millet (Cisco Systems), Tom Petch (BT Connect), Wes George (Time Warner Cable), and Nick Hilliard (INEX).

The author would also like to acknowledge the use of a variety of NANOG mail archives as references.

Author's Address

Anthony Kirkham
Palo Alto Networks
Level 32, 101 Miller St
North Sydney, New South Wales 2060
Australia

Phone: +61 7 33530902
EMail: tkirkham@paloaltonetworks.com