Internet Engineering Task Force (IETF)
Request for Comments: 6018
Category: Informational
ISSN: 2070-1721

F. Baker
Cisco Systems
W. Harrop
G. Armitage
Swinburne University of Technology

September 2010

IPv4 and IPv6 Greynets

Abstract

This note discusses a feature to support building Greynets for IPv4 and IPv6.

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/rfc6018.

Copyright Notice

Table of Contents

1. Introduction

Darknets, also called "Network Telescopes" among other things, have been deployed by several organizations (including CAIDA, Team Cymru, and the University of Michigan) to look at traffic directed to addresses in blocks that are not in actual use. Such traffic becomes visible by either direct capture (it is routed to a collector) or by virtue of its backscatter (its resulting in ICMP traffic or transport-layer resets).

Darknets, of course, have two problems. As their address spaces become known, attackers stop probing them, so they are less effective. Also, the administrators of those prefixes are pressured by Regional Internet Registry (RIR) policy and business requirements to deploy them in active networks.

[Harrop] defines a ’Greynet’ by extension, in these words:

> Darknets are often proposed to monitor for anomalous, externally sourced traffic, and require large, contiguous blocks of unused IP addresses - not always feasible for enterprise network operators. We introduce and evaluate the Greynet - a region of IP address space that is sparsely populated with "darknet" addresses interspersed with active (or "lit") IP addresses. Based on a small sample of traffic collected within a university campus network we saw that relatively sparse greynets can achieve useful levels of network scan detection.

In other words, instead of setting aside prefixes that an attacker might attempt to probe and in so doing court discovery, Harrop proposed that individual (or small groups of adjacent) addresses in subnets be set aside for the purpose, using different host identifiers in each subnet to make it more difficult for an address scan to detect them. The concept has value in the sense that it is harder to map the addresses or prefixes out of an attacker’s search pattern, as their presence is more obscure. Harrop’s research was carried out using IPv4 [RFC0791] and yielded interesting information.

1.1. History and Experience

The research supporting this proposal includes two prototypes, one with IPv4 [RFC0791] and one with IPv6 [RFC2460]. Both have limitations, being research experiments as opposed to deployment of a finished product.

The original research was done by Warren Harrop and documented in [Harrop]. This was IPv4-only. His premise was that one would put a virtual or physical machine on a LAN that one was not otherwise using, and use it to identify scans of various kinds. As reported in his paper, the concept worked effectively in a prototype deployment at the Centre for Advanced Internet Architectures (CAIA), Swinburne University of Technology. The basic reason was that there was a reasonable expectation on the part of a potential attacker that a given address might be represented, and there was no pattern that would enable the attacker to predict which addresses were being used in this way. CAIA developed and released a prototype FreeBSD-based Greynet system in 2008 built around this premise [Armitage].

Baker’s addition to his concept started from the router, the idea that the router would be highly likely to encounter any such scan if it came from off-LAN, and the fact that the router would have to use Address Resolution Protocol (ARP) or Neighbor Discovery (ND) to identify -- or fail to identify -- the machine in question. In effect, any address that is not currently instantiated in the subnet acts as a Greynet trigger address. This clearly also works for any system that would implement ARP or ND, but the router is an obvious focal point in any subnet.

Tim Chown, of the School of Electronics and Computer Science, University of Southampton, offered privately to do some research on it, and had Owen Stephens do a Linux prototype in spring 2010. They demonstrated that the technology was straightforward to implement and in fact worked in a prototype IPv6 implementation.

The question that remains with IPv6 address scanning is the likelihood that the attack would occur at all. Chown originally argued in [RFC5157] that address scans were impossible due to the sheer number of possibilities. However, in September 2010 a report was made to NANOG of an IPv6 address scan. Additionally, there are ways to limit the field; for example, one can observe that a company buys a certain kind of machine or network interface card (NIC), and therefore its probable EUI-64 addresses are limited to a much smaller range than 2^64 -- more like 2^24 addresses on a given subnet -- or one can observe DNS, SMTP envelopes, Extensible Messaging and Presence Protocol (XMPP) messages, FTP, HTTP, etc., that carry IP addresses in other ways. Such attacks can be limited by the use of Privacy Addresses [RFC4941], which periodically change, rendering historical information less useful, but the fact is that such analytic methods exist.

2. Deploying Greynets

Corporate IT departments and other network operators frequently run collectors or other kinds of sensors. A collector is a computer system on the Internet that is expressly set up to attract and "trap" nefarious attempts to penetrate computer systems. Such systems may simply record the attempt or the datagram that initiated the attempt (darknets/Greynets), or they may act as a decoy, luring in potential attacks in order to study their activities and study their methods (honeypots).

To accomplish this, we separate nefarious traffic from that which is likely normal and important, studying one and facilitating the other.

2.1. Deployment Using Routing - Darknets

One obvious way to isolate and identify nefarious traffic is to realize that it is sent to a prefix or address that is not instantiated. If a campus uses an IPv4 /24 prefix or an IPv6 /56 prefix but contains less than 100 actual subnets, for example, we might use only odd numbered subnets (128 of the 256 available in that prefix), and not quite all of those. Knowing that the active prefixes are more specific and therefore attract appropriate traffic, we might also advertise the default prefix from the collector, attracting traffic directed to the uninstantiated prefixes in that routing domain.

A second question involves mimicking a host under attack; the collector may simply record this uninvited traffic, or may reply as a honeypot system.

2.2. Deployment Using Sparse Address Space - Greynets

IPv4 subnets usually have some unallocated space in them, if only because Classless Inter-Domain Routing (CIDR) allocates O(2^n) addresses to an IP subnet and there are not exactly that many systems there.

Similarly, with active IPv6 prefixes, even a very large switched LAN is likely to use a small fraction of the available addresses. This is by design, as discussed in Section 2.5.1 of [RFC4291]. If the addresses are distributed reasonably randomly among the possible values, the likelihood of an attacker guessing what addresses are in actual use is limited. This gives us an opportunity with respect to unused addresses within an IP prefix.

Routers use IPv4 ARP [RFC0826] and IPv6 Neighbor Discovery [RFC4861] to determine the MAC (Media Access Control) address of a neighbor to which a datagram needs to be sent. Both specifications intend that when a datagram arrives at a router that serves the target prefix, but that doesn’t know the MAC address of the intended destination, it should:

- Enqueue the datagram,
- Emit a Neighbor Solicitation or ARP Request,
- Await a Neighbor Advertisement or ARP Response, and
- On receipt, dequeue and forward the datagram.

Once the host’s MAC address is in the router’s tables (and in so doing the address proven valid), the matter is not an issue.

In [Harrop], the Greynet is described as being instantiated on an end-host that replies to ARP Requests for all ’dark’ IP addresses. However, a small modification to router behavior can augment this model. As well as queuing or dropping a datagram that has triggered an ARP Request or Neighbor Solicitation, the router forwards a copy of this datagram over an independent link to the Greynet’s analytic equipment. This independent link may be a different physical interface, a circuit, VLAN, tunnel, UDP, or other encapsulation, or in fact any place such a datagram could be handled. Depending on the requirements of the receiving collector, one could also imagine summarizing information in a form similar to IP Flow Information Export (IPFIX) [RFC5101] [RFC5610].

The analytic equipment will now receive two types of datagrams. Of most interest will be those destined for ’dark’ IP addresses. Of less interest will be the irregular case where a datagram arrives for a legitimate local neighbor who has, for some temporary reason, no MAC address in the router’s tables. Datagrams arriving for an IP destination for which an ARP reply (or Neighbor Advertisement) has not yet received might also be forwarded to the analytical equipment over the independent link -- or might not, if they are considered to be unlikely to provide new analytic information.

Analytic equipment, depending on the router to recognize ’dark’ IP addresses in this manner, can easily track arrival patterns of datagrams destined to unused parts of the network. It may also optionally choose to respond to such datagrams, acting as a honeypot to elicit further datagrams from the remote source.

If the collector replies directly, the attacker may be able to identify the fact through information in or about the datagram - datagrams sent to the same IP subnet may come back with different TTL values, for example. Hence, it may be advisable for the collector to send the reply back through the tunnel and therefore as if from the same IP subnet. Naturally, the collector in this scenario should not respond to datagrams destined for ’lit’ IP addresses -- the intended destination will eventually respond to the router’s ARP or Neighbor Solicitation anyway.

One implication of this model is that distributed denial-of-service (DDoS) attacks terminate on router subnets within a network, as opposed to stopping on inter-router links.

2.3. Other Filters

An obvious extension of the concept would include traffic identified by other filters as appropriate to send to the collector. For example, one might configure the system to forward traffic that fail a unicast Reverse Path Forwarding (uRPF) check [RFC2827] to the collector via the same tunnel.

3. Implications for Router Design

The implication for router design applies to the IPv4 ARP and IPv6 Neighbor Discovery algorithms. It might be interesting to provide, under configuration control, the ability to forward to an analytic system the arriving datagrams that trigger an ARP Request or Neighbor Solicit, and then fail to receive the intended response, to an interface, circuit, VLAN, or tunnel.

4. Security Considerations

This note describes a tool for managing IPv4 and IPv6 network security. Like any tool, it has limitations and possible attacks. If discarding traffic under overload is a good thing, then holding and subsequently forwarding the traffic instead places a potential load on the network and the router in question, and as such represents a possible attack. Such an attack has obvious mitigations, however; one simply selects (in a manner the operator deems appropriate) a subset of the traffic to forward and discards the rest. In addition, this attack is not new; it is only changed in character. A stream that would instantiate the attack today results in a load of ARP or Neighbor Solicit messages that all listening hosts must intelligently discard. The new attack additionally consumes bandwidth that is presumably set aside specifically for that purpose.

The question of exactly what subset of traffic is interesting and economical to forward is intentionally left open. Key questions in algorithm design include what can be learned from a given sample (Are bursts happening? If so, with what data?), what the impact on the router and other equipment in question is, how that might be mitigated, etc. Possible selection algorithms dependent only on state and algorithms typically available in a router include:

- Select all datagrams that trigger an ARP Request or Neighbor Solicit.
- Select the subset of those that are not responded to within some stated interval and are therefore likely dark.
- Select the subset of those that are new; if the address is currently being solicited, forwarding redundant data may not be useful.
- Select all datagrams up to some rate.
- Select all datagrams matching (or not matching) a specified filter rule.

5. Acknowledgements

Algorithms for learning about Internet attack behavior by observing backscatter traffic have been used by CAIDA, University of Michigan, Team Cymru, and others. Harrop extended them in his research. This formulation of the notion originated in a discussion among the authors in 2005. This note grew out of a conversation with Paul Vixie and Rhette Marsh on Internet traffic sensors; they also made useful comments on it. Albert Manfredi commented on the distinction between a LAN (as defined by IEEE 802) and an IP subnet.

Tim Chown [RFC5157] has observed that, at least at the time of writing that RFC, address scanning attacks in IPv6 have not been reported in the wild. However, as mentioned in Section 1.1 above, a (partial) scanning attack was recently reported on the NANOG mailing list. Rhette Marsh has suggested the structure of such an attack, however, and Fred Baker has suggested approaches based on addressing

information exchanged by applications. Hence, we believe that such issues may be relevant to IPv6 in the future, when IPv6 is a more interesting target.

Tim Chown and Owen Stephens tested the proposal, and made useful comments that have been incorporated in this text. His fundamental comment was, however, that "it works".

6. References

6.1. Normative References

[Harrop] Harrop, W. and G. Armitage, "Greynets: a definition and evaluation of sparsely populated darknets", IEEE LCN IEEE 30th Conference on Local Computer Networks, 2005.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC0826] Plummer, D., "Ethernet Address Resolution Protocol: Or converting network protocol addresses to 48.bit Ethernet address for transmission on Ethernet hardware", STD 37, RFC 826, November 1982.

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

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

6.2. Informative References

[Armitage] Armitage, G., Harrop, W., Heyde, A., Parry, L., "Greynets: Passive Detection of Unsolicited Network Scans in Small ISP and Enterprise networks", CAIA, Swinburne University of Technology, December 2008, http://caia.swin.edu.au/greynets/.

[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.

[RFC5101] Claise, B., "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information", RFC 5101, January 2008.

[RFC5157] Chown, T., "IPv6 Implications for Network Scanning", RFC 5157, March 2008.

[RFC5610] Boschi, E., Trammell, B., Mark, L., and T. Zseby, "Exporting Type Information for IP Flow Information Export (IPFIX) Information Elements", RFC 5610, July 2009.

Authors' Addresses

Fred Baker
Cisco Systems
Santa Barbara, California 93117
USA

EMail: fred@cisco.com

Warren Harrop
Centre for Advanced Internet Architectures
Swinburne University of Technology
PO Box 218
John Street, Hawthorn,
Victoria, 3122
Australia

EMail: wazz@bud.cc.swin.edu.au

Grenville Armitage
Centre for Advanced Internet Architectures
Swinburne University of Technology
PO Box 218
John Street, Hawthorn,
Victoria, 3122
Australia

EMail: garmitage@swin.edu.au