September 2013
Geoff Huston

All IP Addresses are not the Same

One IP address is much the same as another - right? There's hardly a difference between 192.0.2.45 and 192.0.2.46 is there? They are just encoded integer values, and aside from numerological considerations, one address value is as good or bad as any other - right? So IP addresses are much the same as each other and an after-market in IP addresses should be like many other markets in undistinguished commodity goods. Right?

So one would've thought. But it seems that this is really not the case. When it comes to IP addresses, not all addresses are the same. IP addresses have a certain amount of history, and this history alters its utility value in some ways.

For example, an address may exist on one or more "blacklists". Many readers would be aware of the various forms of blacklists that have been used in the never-ending fight against spam. These blacklists enumerate the IP addresses of hosts that have been observed to emit spam, and once an IP address is listed in one of these lists, then many other mail systems will not communicate with it. It's often claimed that it's extremely easy to get an IP address into one of these blacklists, but very hard to get off them once it has been listed. Like many reputation services, once a good reputation is lost for an IP address it's often very difficult to re-establish it again. Part of the problem is that it's very easy to set up a blacklist, so many folk have done so. But you can tell when a concept has gone perhaps a little too far when aggregators enter the fray. As is claimed on the dnsbl.info web site: "DNSBL Information provides a single place where you can check that status of your mail server's IP address on more than 100 DNS based blacklists."

There are others ways in which IP addresses can be differ from each other. These include the extent to which an address acts as an attractor for unsolicited incoming traffic, and the extent to which other network operators actively filter incoming packets that use this address as a source. I'd like to explore that aspect of IP addresses in this column.

Unsolicited Incoming Traffic in 1.0.0.0/8

In early 2010 the IANA allocated to APNIC the address block 1.0.0.0/8. It was noted at the time that addresses from this address block had already been "widely used as private address space in large organizations whose needs exceed those provided for by RFC 1918," according to one commentator at the time.[1]

Indeed it seems that wherever I go there is a public wifi access point that is squatting on the IP address 1.1.1.1, an address which is, in effect, being hijacked and used without authority or permission by these Wifi base stations, as this address has been allocated to APNIC Labs for test work.[2]

[1] Leo Vegoda, "Awkward /8 Assignments", Internet Protocol Journal, September 2007. (http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_10-3/103_awkward.html)

[2] http://rdap.apnic.net/ip/1.1.1.1

How does this form of prior use manifest itself? One way in which the prior use of an address creates problems for the current user of an address is in the appearance of "unsolicited traffic." This is incoming traffic appearing at the address that is not in response to any previous transaction or request. The way in which this form of traffic can be captured is by setting up a "dark net". A dark net is a configuration where a route to the address is announced to the Internet, and any incoming traffic that is being sent towards this address is passed into a recording device. Typically, the device never responds to this incoming traffic in any way, so to the external world the network block looks like a "dark" block. It accepts and records all incoming packets, but emits nothing in return.

What does this unsolicited incoming traffic look like? Here's what the traffic profile looked like for the address block 1.0.0.0/8 in early 2010. At the time the address block had just been passed from the IANA to APNIC, and, supposedly, the address block was vacant and ready for allocation. While it would be naive to expect that this address block would attract absolutely no incoming traffic, it was nevertheless a surprise that this block attracted so much traffic.

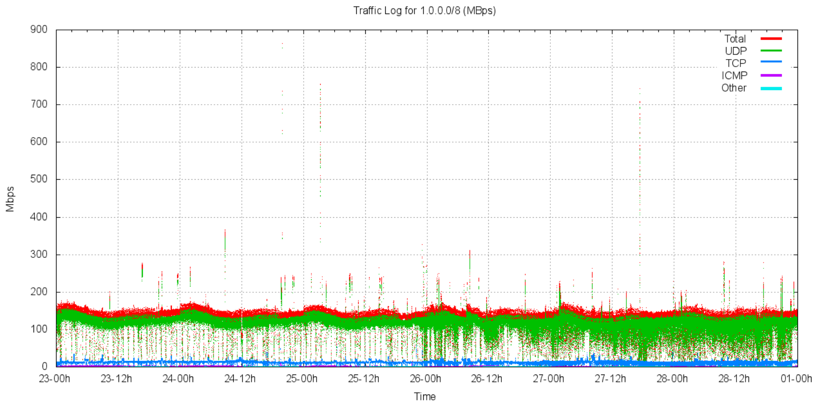

Figure 1 – Incoming Traffic profile for 1.0.0.0/8 (February 2010)

At the time of this text (February 2010) this block was consistently attracting a total of 160Mbps of incoming traffic. There are short bursts of between 1 and 30 seconds of elevated traffic levels. There are 20 or so incidents of burst traffic levels of between 200Mbps and 300Mbps in this recording. There is a 3 second isolated burst at 860Mbps and a 10 second burst at 750Mbps in this period.

Obviously 1.0.0.0/8 is a very active block, as other /8s address blocks attracted an average of between 12Mbps to 25Mbps when they were tested in 2010 and early 2011. But perhaps there is a more relevant question here. Was this incoming traffic evenly spread across addresses drawn from the entire address block, or was the traffic profile such that individual addresses were being hammered by extremely high volumes of incoming traffic?

If we look at the incoming traffic segmented into each of the 256 /16 address blocks, then it is evident that there is some considerable variation across each of the /16 blocks.

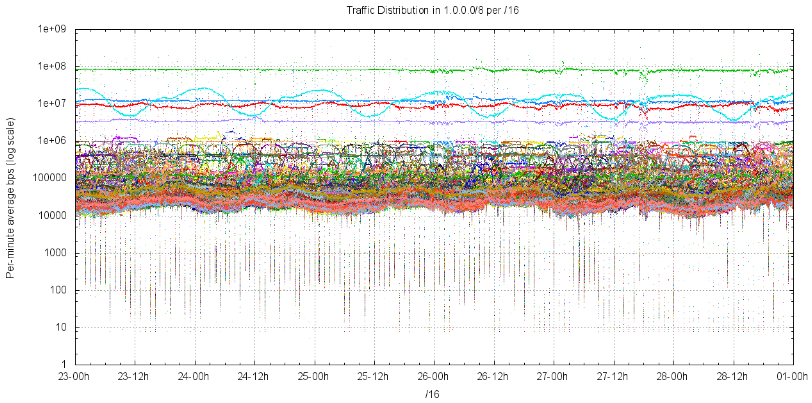

Figure 2 – Incoming Traffic profile for 1.1.1.0/8, per /16 (February 2010)

This figure uses a log scale, so its pretty evident that some /16s attract far more traffic than others. In this period one /16 recorded a sustained level of 100Mbps of incoming traffic, a small number of other /16s saw more than 1Mbps, and the majority of /126s experienced average incoming traffic levels of between 10Kbps and 100Kbps.

Causes of Unsolicited Incoming Traffic

Obviously, the /16s with average incoming traffic rates in excess of 100Kbps are abnormal in some manner, but what is "normal" anyway? During 2010 and 2011 we performed the same dark net tests for a number of the /8 address blocks that we received from the IANA prior to using them for allocations, and it is interesting to compare the traffic profile for the /16 blocks in 1.0.0.0/8 with 49.0.0.0/8.

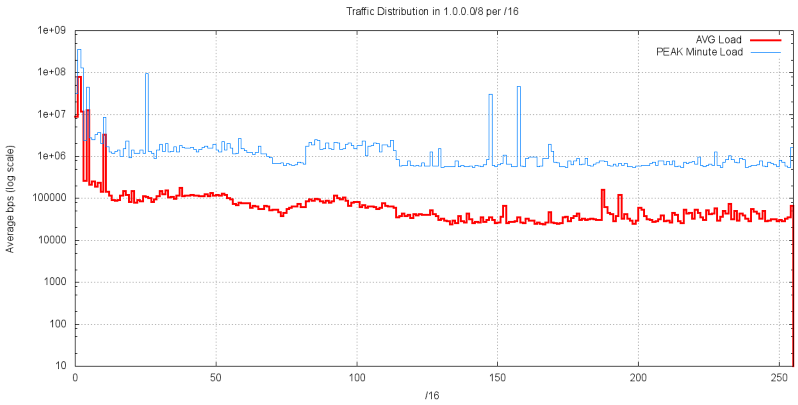

Figure 3 – Average Incoming Traffic profile for 1.0.0.0/8, per /16 (February 2010)

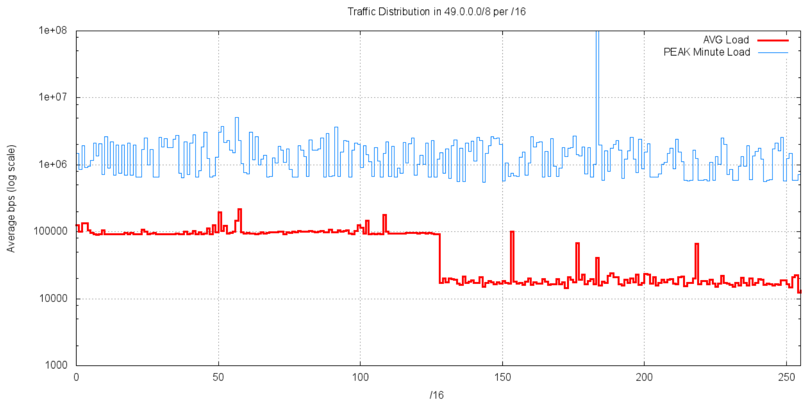

Figure 4 – Average Incoming Traffic profile for 49.0.0.0/8, per /16 (September 2010)

It seems that the unsolicited incoming traffic level for a /16 is "normally" either 100Kbps or 20Kbps. And the high number is found for all /16 nets in the low half of a /8 address block and the lower number in the higher half. This in itself is a very curious result. What is the average incoming traffic rate differ between these two halves by a factor of 5?

Taking a one hour sample period for 223.0.0.0/8 from May 2010, we saw 1 million TCP SYN packets directed to port 445 in the high /9 of 223.128.0.0/9 (or 0.8% of the packet count for the hour), as compared to 92 million in the low /9 (or 60% of the packet count). TCP port 445 is used by Microsoft systems to support the Server Message Block (SMB) protocol, used for file sharing. It is also a very common vector for attacks on Microsoft Windows systems, and the most virulent virus that attempts to exploit port 445 is the Conficker virus. Reports of the behaviour of the Conficker virus point to a outcome of the virus' random IP generation routine for port 445 scanning where bit 9 of the randomly generated IP address is always 0, as is bit 24.[3] The outcome of bit 9 being clear is that Conficker will only scan the half of any /8 network block using the random IP generator. Some 100Kbps of traffic per /16 in the bottom half of each of the /8 address blocks we tested is attributable to Conficker's port 445 scanning activity. The total traffic component of Conficker is some 40% of the total traffic load directed to these network blocks, and 60% of the total packet count, indicating the significant extent to which unpatched Windows XP systems continue to be vulnerable to this particular virus.

Aside from Conficker scanning, what else do we see?

Certainly, scanning is very commonplace in IPv4. There are a number of active viruses that force their host to probe the entire IPv4 address space to find new hosts to infect. There are also a number of what appears to be "manual" scanners, where a scan of the address space is performed in a single pass. Indeed, scanning is now so common that there is an open source IPv4 address scanning tool![4] Aside from Conficker, most of these scanners distribute their traffic evenly across the entire IPv4 space. This implies an address in the high /9 of each /8 is more likely to be slightly "cleaner" than one in the low /9, due mainly to the Conficker scanning activity.

There are also individual anomalies associated with individual addresses, where individual addresses are the targets of high volumes of incoming traffic. In some cases we have observed point-to-point traffic, where a single source address is sending a high volume of unsolicited traffic to the target address. This is often in the form of UDP packets to port 80, but we have also observed sustained high traffic rates on other UDP ports, and also seen very high TCP SYN packet streams. We have had some success in cleaning up this form of traffic by tracing back the source address and contacting the relevant network operator and requesting assistance, but this has only been effective in a small number of isolated cases.

Where there are a large number of sources involved in sending unsolicited traffic, this form of cleanup is simply not possible, There appear to be cases of errors in DNS entries, where a name is mapped to an incorrect address. When applications attempt to connect to the service point at this address the lack of the expected response may cause the client application to retransmit. For example, the network management protocol, SNMP, was observed to cause a high query rate on some addresses. we have also observed the same with the SMTP mail protocol, and port 5060 using by VOIP applications. Other addresses appear to have been configured as Teredo relays, and attract high volumes of Teredo connection establishment messages. Finding the relevant DNS entry that triggered the consequent connection attempts is extremely challenging. There are also cases where a DNS zone's name server records are incorrectly specified, and DNS query traffic is then directed to the address. The lack of a DNS response causes queries to be repeated, further adding to the incoming traffic load.

It's not just errors in DNS zone files that can generate this form of unsolicited incoming traffic. A number of online games use a list of IP addresses to circulate the set of current game servers to potential players. This results in a high load of received UDP messages that contain what appears to be encrypted payloads, that are the initial packet in the game setup sequence.

We have also observed incoming traffic as a consequence of vendor equipment being poorly configured. The major component of unsolicited incoming traffic in 223.1.0.0/16 was directed to 223.1.1.0/24. Within this /24, the overall majority of the traffic is being directed to the single address 223.1.1.128. A web search for this address reveals that a possible cause for unsolicited traffic being directed is traffic leakage from a "secure" VPN product. It appears that this VPN product uses 223.1.1.128 as a default network adapter interface. What is being observed here appears to be leakage of traffic into the public network from this default configuration state where VPN traffic is being directed to the address 223.1.1.128. The traffic level of this leakage of VPN traffic into the public Internet is between 300Kbps and 500Kbps at the time of the test.

[3] http://www.caida.org/research/security/ms08-067/conficker.xml

[4] https://zmap.io

Mitigation?

What can be done about this?

In some cases the incoming traffic flow ceases without any intervention. Presumably a periodic examination of the DNS zone configurations, or an audit of the addresses listed in an online game's rendezvous listing identified that the address in question is an error, and one the address is removed, the associated traffic also ceases.

In other cases, where the incoming traffic is sourced from a single remote address, or a small group of addresses, it may be possible to perform a form of traceback to the sender's address, and with the assistance of the relevant security response teams and the network operator it may be possible to get the problem corrected at source. In this particular situation, mitigation would be effective.

However, this is perhaps the exception rather than the general case. In most other cases the large number of different sources imply that it is simply not possible to intervene to mitigate the traffic levels. It is not possible to trawl through the entire set of DNS zone files to locate instances of mis-typed addresses. The task of eradicating all forms of malicious malware and the creation of bot armies that spew out unsolicited traffic has equally proved to be beyond our best efforts to date. Perhaps there is simply a common background level of traffic that is normal, and each Ip address can reasonably be expected to attract its share of this traffic.

In our work in looking at the traffic profile in 2010 and 2011 we suggested that as a high threshold an average incoming traffic level per address of 32bps, or approximately 1 incoming packet every 24 seconds, would be reasonable. If an address attracted a sustained level of traffic that was higher than this threshold then it could reasonably be considered to be anomalous in some manner.[5] In APNIC's case these address blocks were withheld from allocation in some cases, and in other cases the recipient of the encompassing address block was informed of these anomalous address within the block, and advised not to allocated them to end users.

While all IP addresses are subject to a certain level of unsolicited incoming packets, due to a constant level of address scanning across the IPv4 address space, some IP addresses attract considerably higher levels of traffic than others. The most extreme case we've observed so far is the address 1.1.1.1, which attracts up to 1Gbps of unsolicited incoming traffic to just that address. But that's not the only address that stands out from the background. Other addresses also attract large quantities of traffic, But precisely which address and how much traffic is not possible to predict. It appears that the best way to find out just how big or small the problem may be for each addresses is to test them, to see precisely how much traffic it attracts, and whether it can be stopped.

In terms of acting as an unsolicited traffic attractor all IP addresses are not the same.

[5] http://www.potaroo.net/studies/retest/retest.pdf

Disclaimer

The views expressed are the authors' and not those of APNIC, unless APNIC is specifically identified as the author of the communication. APNIC will not be legally responsible in contract, tort or otherwise for any statement made in this publication.

About the Author

Geoff Huston B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector. He is author of a number of Internet-related books, and was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of Trustees of the Internet Society from 1992 until 2001.
www.potaroo.net