Opportunistic IPv6 Insight via Abusive Traffic

Robert Beverly, Geoffrey Xie

Naval Postgraduate School
{rbeverly,xie}@nps.edu
February 8, 2012

CAIDA Workshop on Active Internet Measurements

Outline

What we can all (sort of) agree on

Crying Wolf Again? (U.S. perspective)

- Exhaustion of v4 addresses finally exerting (economic) pressure on providers to use IPv6
- More and more devices (e.g. mobile)
- Widespread OS support, auto-tunneling
- Carrier-grade NAT is bad (viz. E2E)
- U.S. government mandates

Err....

```
; <<>> DiG 9.8.1 <<>> AAAA www.disa.mil
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 63718
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0

;; QUESTION SECTION:
;www.disa.mil.                  IN      AAAA
```

What we can all (sort of) agree on

Crying Wolf Again? (U.S. perspective)

- Exhaustion of v4 addresses finally exerting (economic) pressure on providers to use IPv6
- More and more devices (e.g. mobile)
- Widespread OS support, auto-tunneling
- Carrier-grade NAT is bad (viz. E2E)
- U.S. government mandates

Err....

```
; <<>> DiG 9.8.1 <<>> AAAA www.disa.mil
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 63718
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 0

;; QUESTION SECTION:
;www.disa.mil.                  IN      AAAA
```

IPv6 Measurements

Many independent IPv6 measurement efforts:

- Multiple web-bug / javascript
- Passive traffic analysis
- Active probing
- Dark/Grey nets

Our Hypothesis:

Our Hypothesis:

- Opportunistically utilize abusive IPv6 traffic
- Abusive traffic has been productive in other measurement efforts
- Suggests at a means to obtain (a large number of) samples from the IPv6 edge, with *different* sample bias
- Additionally, reveal properties/prevalence of IPv6 as emergent attack vector

This talk: initial experiments to test the opportunistic abusive IPv6 traffic hypothesis (read as: ongoing effort).

Outline

IPv6 Abusive Traffic

What do we mean by “abusive?”

- Many IPv6 protocol-specific attacks, not in scope here
- Instead: Traditional abusive traffic (DoS, messaging, worm propagation, etc) using IPv6 transport

Why might we expect abusive IPv6 traffic?

- Bad guys will exploit any possible attack vector
- Easy: incestuous abusive/malicious code libraries permit widespread adoption
  - e.g. THC-IPV6
- Near zero cost to test for IPv6 connectivity
- Newly adopted protocols often rife with vulnerabilities
- All old security problems in IPv4 are new again...

IPv6 Abusive Traffic

Fly under the radar of monitoring, or evade blocking:

- Firewalls, filters, IDS, DPI, etc rarely configured to support IPv6
- Tunnels and auto-tunnel mechanisms (e.g. 6to4, Teredo) subvert administrative security policies and protection/detection
- E.g. residential outbound TCP SMTP blocked only for IPv4
- Address agility, IPv6 RBLs not as well-maintained:
  - `http://www.ipv6whitelist.eu`
- Lots of buggy implementations:
  - Ask us about our IDS fuzz testing where we can throw snort into infinite recursion via crafted IPv6 packets!

IPv6 Attacks

Bad stuff is IPv6 connected:

Database	**Entries w/** A	**Entries w/** AAAA
malwaredomainlist.com	2095	35 (1.7%)
malwaredomains.com	845	10 (1.2%)
phishtank.com	3318	16 (0.5%)

- Coincidentally or intentionally on IPv6?
- (Collected and probed February, 2012)

IPv6 Attacks

Unsurprisingly, bad stuff is IPv6 connected:

Database	**Entries w/** AAAA	**Unique ASN**	**RIPE ASN**
malwaredomainlist.com	35	10	8
malwaredomains.com	10	5	5
phishtank.com	16	10	9

- Not all in one AS
- Mostly in Europe (none in US)
- (Collected and probed February, 2012)

IPv6 Attacks

Lots of anecdotal evidence:

- Trojans: Troj/LegMir-AT IPv6 IRC (public reference)
- Worms: W32/VB-DYF (public reference)
- Wordpress malware using IPv6 site-scraping (private conversation with CDN, 2011)

Take-away:

- There exist sources of abusive IPv6 traffic
- Even if traffic is small relative to v4, still interesting
- Exploit abusive IPv6 traffic for measurement of the IPv6 Internet

IPv6 Attacks

Lots of anecdotal evidence:

- Trojans: Troj/LegMir-AT IPv6 IRC (public reference)
- Worms: W32/VB-DYF (public reference)
- Wordpress malware using IPv6 site-scraping (private conversation with CDN, 2011)

Take-away:

- There exist sources of abusive IPv6 traffic
- Even if traffic is small relative to v4, still interesting
- Exploit abusive IPv6 traffic for measurement of the IPv6 Internet

Outline

IPv6 Honeypot

Initial experiment: IPv6 Spam Honeypot

- Easy and popular method to attract abusive traffic: spam honeypot
- We built and instrumented an IPv6 spam honeypot

Prior Work

- `ripe.net:` Not a honeypot; 3.5% of IPv6 emails spam (2010)
- `cert.br:` Total of 6 IPv6 HTTP hits over 3 months (2009)
- `soton.ac.uk:` Not a honeypot; “roughly half of IPv6 email is spam.” (2008)

Idea: run a IPv6 spam honeypot before/after World IPv6 day

IPv6 Honeypot

Initial experiment: IPv6 Spam Honeypot

- Easy and popular method to attract abusive traffic: spam honeypot
- We built and instrumented an IPv6 spam honeypot

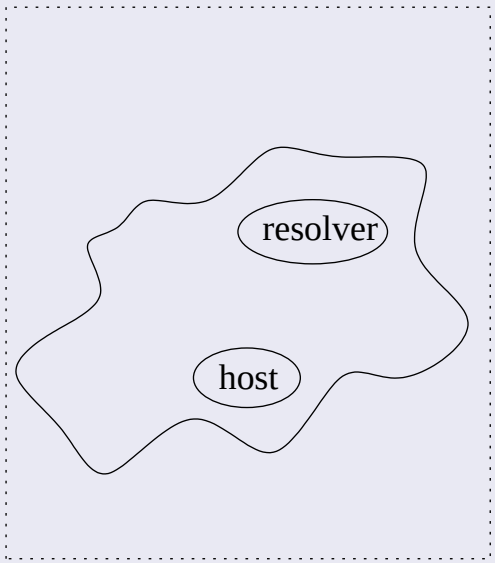
Prior Work

- `ripe.net:` Not a honeypot; 3.5% of IPv6 emails spam (2010)
- `cert.br:` Total of 6 IPv6 HTTP hits over 3 months (2009)
- `soton.ac.uk:` Not a honeypot; “roughly half of IPv6 email is spam.” (2008)

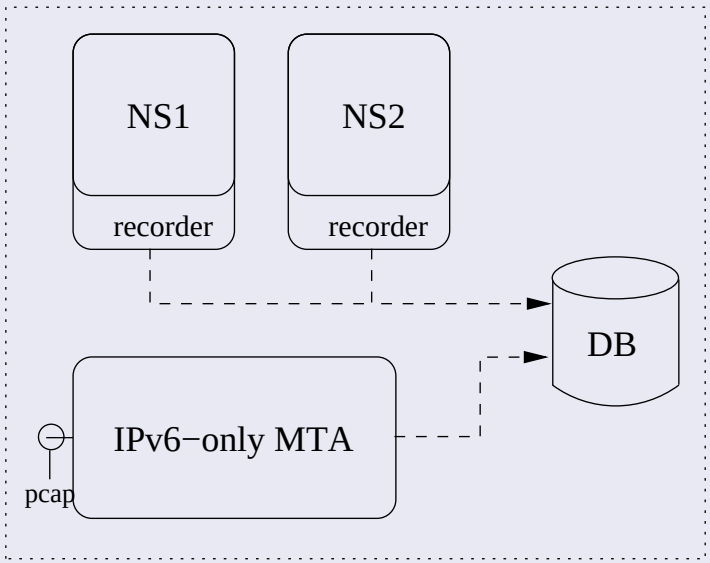
Idea: run a IPv6 spam honeypot before/after World IPv6 day

IPv6 Pot

IPv6 Pot:



Abusive Network

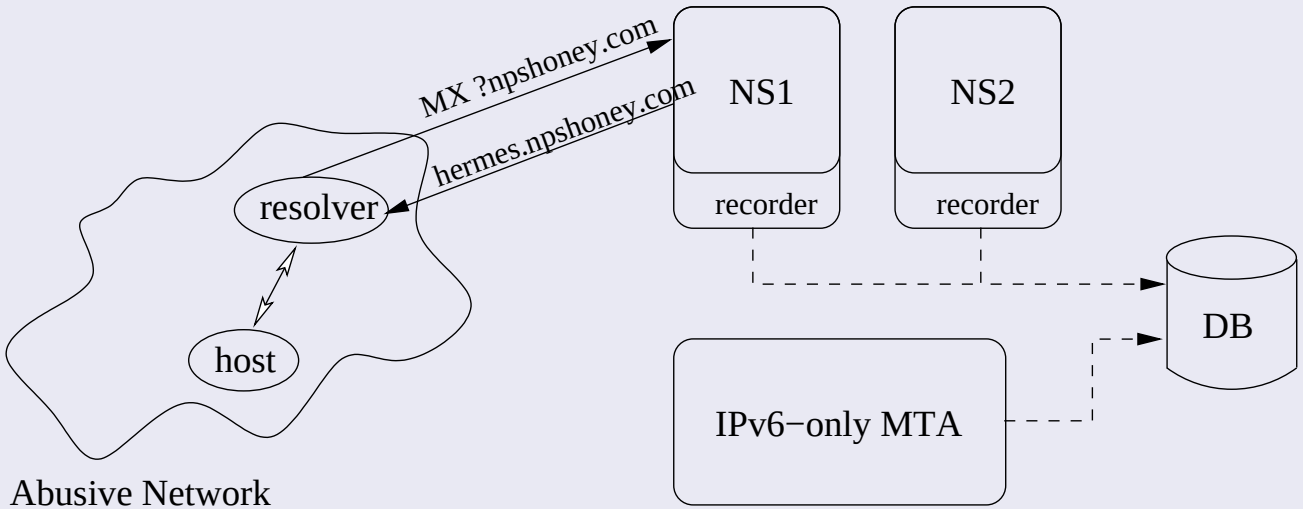


NPS IPv6 Honeypot

Run instrumented authoritative name servers and *IPv6-only* spam sink (RFC3974)

IPv6 Pot

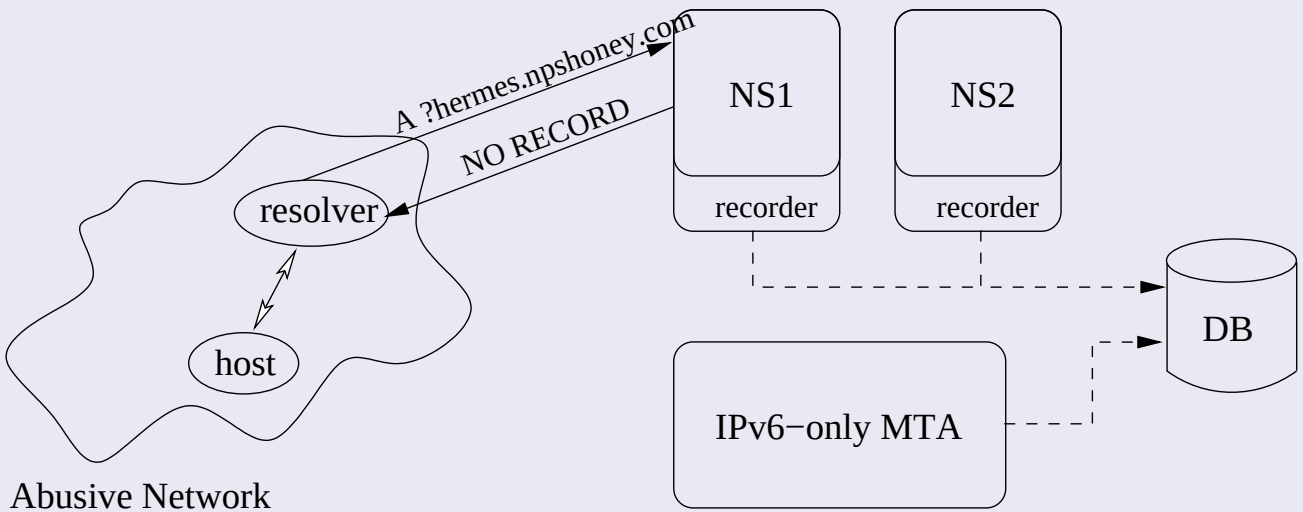
IPv6 Pot:



MX queries (via IPv4 or IPv6) returned and recorded to database

IPv6 Pot

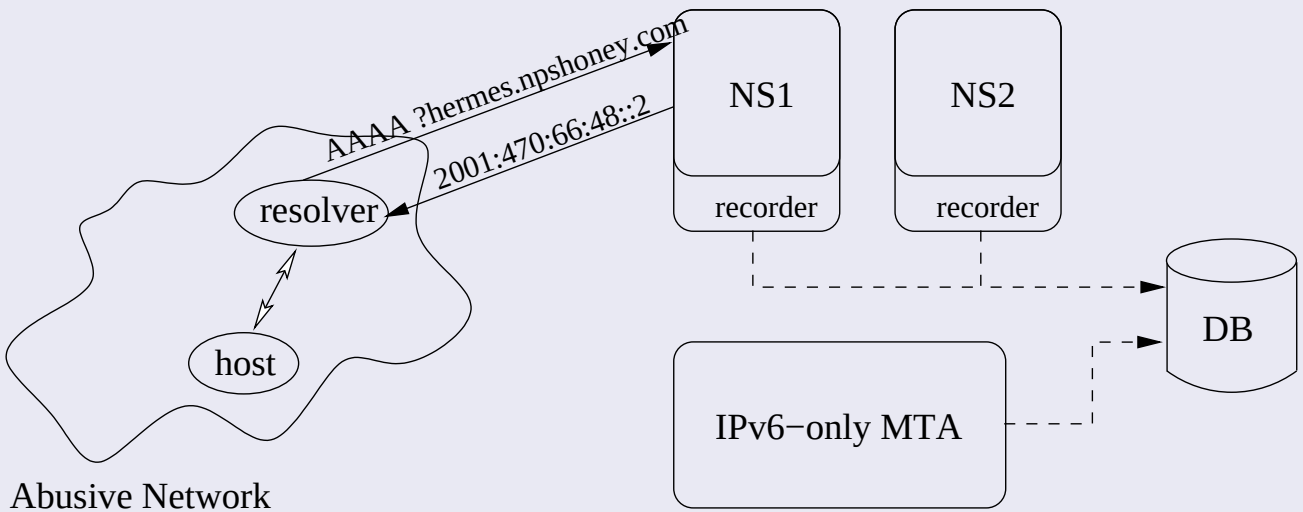
IPv6 Pot:



No associated A record (query recorded)

IPv6 Pot

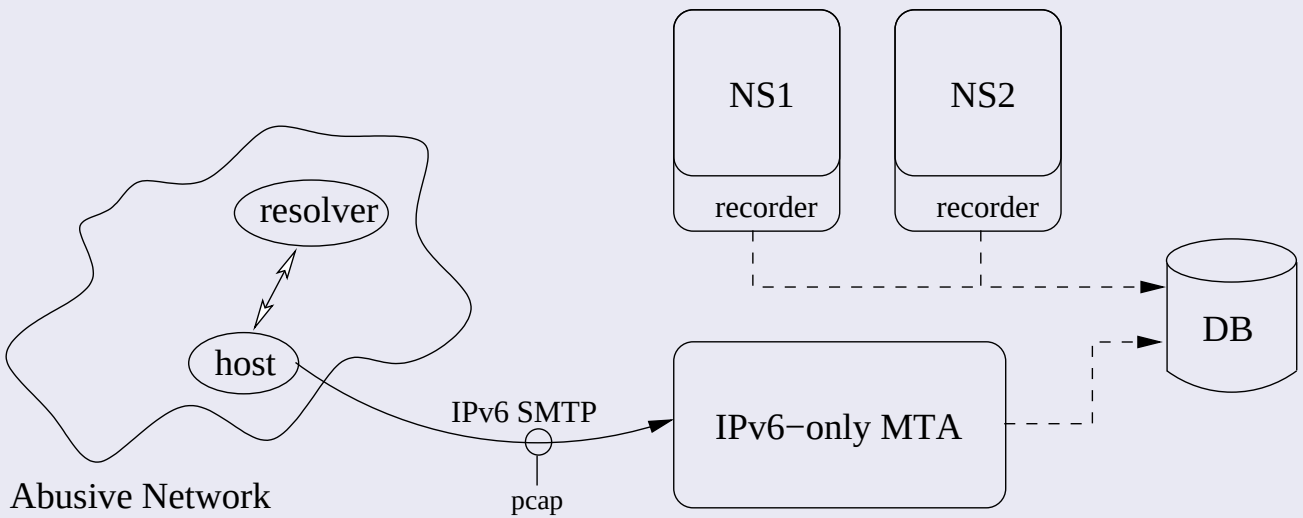
IPv6 Pot:



AAAA record available (query recorded)

IPv6 Pot

IPv6 Pot:



Spam sink catchall for any IPv6 SMTP.

Attracting Traffic

Attracting Traffic

- Dynamic HTML text at bottom of our group web pages generates: `nonce@npshoney.com`
- Records: IPv4/v6 source, browser, resource to database
- Additionally, manually visited several spam URLs and entered our email

Honeypot Analysis

What can we learn:

- How many *attempted* spam SMTP connections resulted in an email?
- Do abusive spam (hosts/bots) use IPv6 when it's the *only* transport available?
- Reconstruct how mined email addresses get to IPv6-capable spammers
- IPv6 edge:
  - Addresses for tracing
  - Prevalence of auto-tunneling
  - Mapping of IPv4 to IPv6

Validation Surprisingly Difficult

- None of: gmail, yahoo, NPS, MIT, UCSD worked
- Ended up using `mailman.nanog.org` to validate

gmail

```
Delivery to the following recipient failed permanently:

     valid@npshoney.com

Technical details of permanent failure:
The recipient server did not accept our requests to connect. Learn more at
http://mail.google.com/support/bin/answer.py?answer=7720
[hermes.npshoney.com. (10): Destination address required]
```

NPS

```
Delivery has failed to these recipients or groups:

valid@npshoney.com
A problem occurred during the delivery of this message to this e-mail address.
Try sending this message again. If the problem continues, please contact your
helpdesk.
```

Outline

Caveat

Caveat

- Started our honeypot just before World IPv6 day
- Unfortunately, we had a bug in our DNS instrumentation :(
- Now fixed
- Results still interesting

Received IPv6 Spam

June 8, 2011 – July 8, 2011

- Received a total of 14 spam email messages via IPv6
- Variety of spam (Nigerian, phishing, products, backscatter)
- Variety of languages (English, Russian, Chinese)
- One 6to4 source
- All sources “server” hosts; did not observe bot/hacked “edge”

Received IPv6 Spam

June 8, 2011 – July 8, 2011 (chronologically listed)

SMTP name	IPv6
smtp.softcloud.ru	2002:c2be:6b0::c2be:6b0
vwp4845.webpack .hosteurope.de	2a01:488:42::53a9:1b45
mo-p07-ob6.rzone.de	2a01:238:20a:202:53f7::1
nb24.sierhuis.com	2a02:348:47:61e9::1
sl4.sahara.net.sa	2a02:d70:10:0:250:56ff :feae:1bde
ncu.edu.cn	2001:250:6c00:f02:230:48ff:feba:69d2
aruana2.ufscar.br	2001:12f0:503:100::22
s11.usassh.com	2607:fd70:0:6::563f:13e3
cf10.hc.ru	2a01:d8:4:4:230:48ff:feb8:36e8
re02.hc.ru	2a01:d8:4:1:230:48ff:fe67:9c0
cf5.hc.ru	2a01:d8:4:1:230:48ff:fed2:e722

Received IPv6 Spam

June 8, 2011 – July 8, 2011 (chronologically listed)

IPv6	ASN	Cntry	Type
2002:c2be:6b0::c2be:6b0	6to4	RU	backscatter
2a01:488:42::53a9:1b45	20773	DE	Product?
2a01:238:20a:202:53f7::1	6724	DE	Nigerian
2a02:348:47:61e9::1	35470	NL	Phish
2a02:d70:10:0:250:56ff:feae:1bde	41176	SA	Phish
2001:250:6c00:f02:230:48ff:feba:69d2	4538	CN	Product?
2001:12f0:503:100::22	1916	BR	Phish
2607:fd70:0:6::563f:13e3	1426	US	Nigerian
2a01:d8:4:4:230:48ff:feb8:36e8	5537	RU	backscatter
2a01:d8:4:1:230:48ff:fe67:9c0	5537	RU	backscatter
2a01:d8:4:1:230:48ff:fed2:e722	5537	RU	backscatter

World IPv6 Month Experiment

World IPv6 Month Experiment

- We received IPv6 spam
- Variety of sources, AS's, countries, and types encouraging
- Warrants keeping the infrastructure up and running during (the assured) IPv6 adoption
- We started a new (on-going) experiment in February, 2012 with bugs and kinks worked out

Name Server Hits

- New experiment thus far: Jan 29, 2012 – Feb 4, 2012

Name Server Activity for npshoney.com

Query	NS1	NS2
MX	28 (28%)	39 (27%)
A	56 (56%)	81 (56%)
AAAA	8 (8%)	6 (4%)
Other	8 (8%)	18 (12.5%)
Total	100	144

Name Server Hits

- New experiment thus far: Jan 29, 2012 – Feb 4, 2012

Name Server Activity for npshoney.com

Record	Queries for MTA	Distinct
MX	28 (100%)	28 (100%)
A	77 (56%)	31 (40%)
AAAA	1 (8%)	1 (100%)

Observations:

- One AAAA lookup for our MTA, but no connection attempt!?
- MX query rate of $\simeq$ 7/day too low. Need to attract more spam.
- Surprising number of A queries not for our MTA (who is querying?)
- Even ANY and AXFR requests!

Outline

Summary

- Existence proof of abusive IPv6 traffic:
  - $\simeq 1 - 2\%$ of malware, phishing web sites are IPv6 reachable
  - Our IPv6-only honeypot received IPv6 spam!
- IPv6 abusive traffic may yield interesting measurement insights we cannot otherwise obtain
- Other opportunistic measurement opportunities (e.g. BitTorrent to avoid blocking)
- More (hopefully) to come...

Thanks! Questions?