Foundstone
A division of McAfee

Passive Host Characterization

Matthew Wollenweber

Shmoocon 2008

Agenda

- Passive Host Characterization
  - Core Principals
  - Existing Technology
  - Current Uses
  - Research

Background

- Matthew Wollenweber
  - Sr. Consultant at Foundstone
  - Specializes in Penetration Testing
  - Former developer for DoD on Trickler Project
  - Former member of various commercial and DoD Red Teams

Credit/Thanks

- Experience with Passive Host Characterization initially developed while at G2, www.g2secure.com.
- Ron Gula at Tenable for general advice and for use of PVS
- The Government (despite being a bit difficult)

Passive Host Characterization

- Why PHC is good:
  - It’s passive so it doesn’t cost your network anything
  - The basic technology is simple
  - Active scanning can be a political nightmare
  - PHC watches over time; scans are snapshots
  - PHC can detect problems that active scanning and traditional IDS systems can’t

Passive Host Characterization

- Basic Concepts:
  - Passively tap networks
  - Observe traffic
    - Server Versions
    - Client Versions
    - TCP/IP Fingerprints
    - DNS Queries
    - HTTP Traffic – special emphasis as http tends to leak loads of information.

Fundamentals

- Data
  - TCP/IP Fingerprints
    - P0f
    - SynFP
  - Server Strings
    - SSH
    - FTP
    - HTTP
    - Proxies
    - SMTP

Fundamentals

- Data
  - Client Strings
    - USER-AGENT
    - HTTP-REFERER
    - Limewire
    - Email clients
  - DNS
    - Simple protocol – very interesting data
    - More later

Fundamentals

- Basic Concepts -- Continued
  - Aggregate/Reduce/Process Data
  - Correlate to known vulnerable applications
  - Datamine (manually or through automated scripts)

Fundamentals

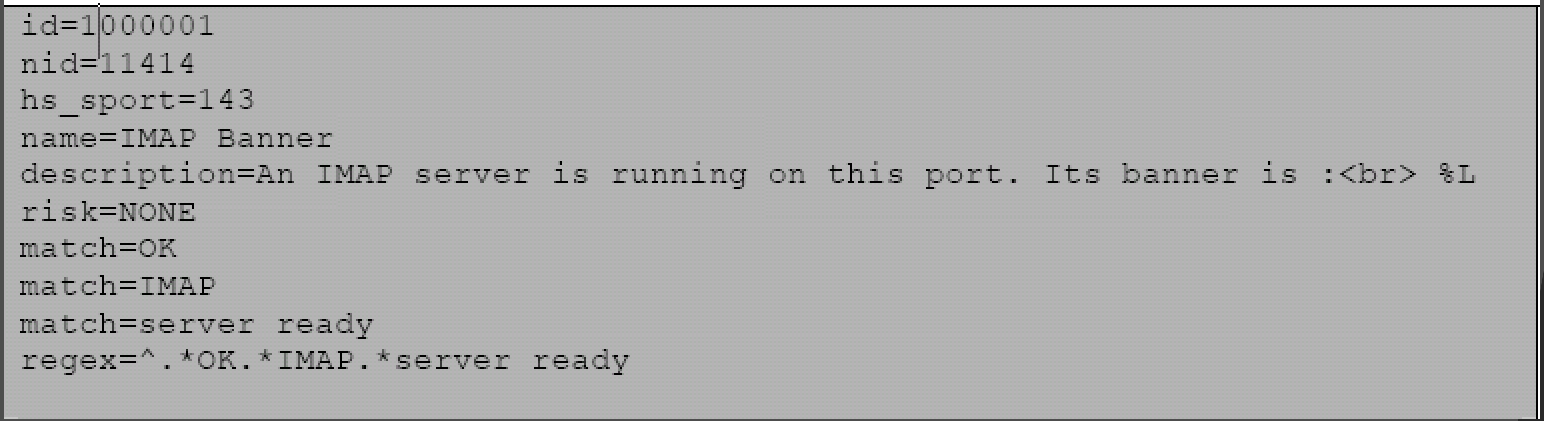
- Data collected at network pipes
- Commodity hardware (I prefer Linux)
- Libpcap
- Not necessary to keep state
  - Memory is a key limitation on many IDS
- Data can be processed AFTER collection

Known Projects: PVS

- Passive Vulnerability System
  - Tenable Project (Makers of Nessus)
  - Signature based
    - Tied to Nessus NASL scripts
    - Regularly updated
  - ~GigE throughput
  - Very good at detecting vulnerabilities
  - Backend not readily accessible for custom queries
  - Flexible Rule language – similar to most IDS systems

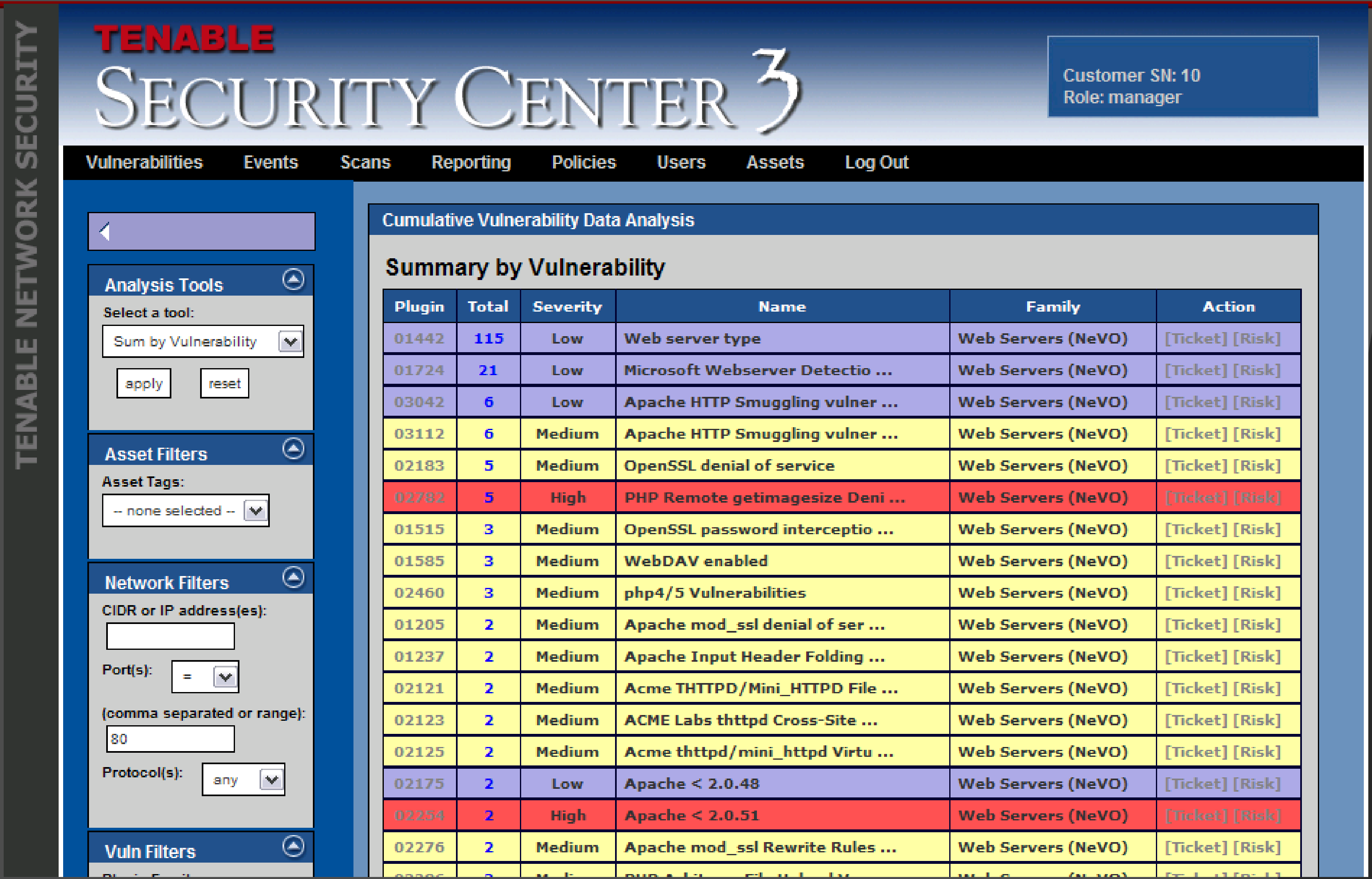
Known Projects: PVS

- Sample PVS rule, looking for IMAP servers



```
id=1000001
nid=11414
hs_sport=143
name=IMAP Banner
description=An IMAP server is running on this port. Its banner is :<br> %L
risk=NONE
match=OK
match=IMAP
match=server ready
regex=^.*OK.*IMAP.*server ready
```

Passive Vulnerability Identification



TENABLE NETWORK SECURITY

TENABLE SECURITY CENTER 3

Customer SN: 10
Role: manager

Analysis Tools

Select a tool: Sum by Vulnerability

Asset Filters

Asset Tags: -- none selected --

Network Filters

CIDR or IP address(es):

Port(s): =

(comma separated or range): 80

Protocol(s): any

Vuln Filters

Cumulative Vulnerability Data Analysis

Summary by Vulnerability

Plugin	Total	Severity	Name	Family	Action
01442	115	Low	Web server type	Web Servers (NeVO)	[Ticket] [Risk]
01724	21	Low	Microsoft Webserver Detectio ...	Web Servers (NeVO)	[Ticket] [Risk]
03042	6	Low	Apache HTTP Smuggling vulner ...	Web Servers (NeVO)	[Ticket] [Risk]
03112	6	Medium	Apache HTTP Smuggling vulner ...	Web Servers (NeVO)	[Ticket] [Risk]
02183	5	Medium	OpenSSL denial of service	Web Servers (NeVO)	[Ticket] [Risk]
02782	5	High	PHP Remote getimagesize Deni ...	Web Servers (NeVO)	[Ticket] [Risk]
01515	3	Medium	OpenSSL password interceptio ...	Web Servers (NeVO)	[Ticket] [Risk]
01585	3	Medium	WebDAV enabled	Web Servers (NeVO)	[Ticket] [Risk]
02460	3	Medium	php4/5 Vulnerabilities	Web Servers (NeVO)	[Ticket] [Risk]
01205	2	Medium	Apache mod_ssl denial of ser ...	Web Servers (NeVO)	[Ticket] [Risk]
01237	2	Medium	Apache Input Header Folding ...	Web Servers (NeVO)	[Ticket] [Risk]
02121	2	Medium	Acme THTTPD/Mini_HTTPD File ...	Web Servers (NeVO)	[Ticket] [Risk]
02123	2	Medium	ACME Labs thttpd Cross-Site ...	Web Servers (NeVO)	[Ticket] [Risk]
02125	2	Medium	Acme thttpd/mini_httpd Virtu ...	Web Servers (NeVO)	[Ticket] [Risk]
02175	2	Low	Apache < 2.0.48	Web Servers (NeVO)	[Ticket] [Risk]
02254	2	High	Apache < 2.0.51	Web Servers (NeVO)	[Ticket] [Risk]
02276	2	Medium	Apache mod_ssl Rewrite Rules ...	Web Servers (NeVO)	[Ticket] [Risk]

Known Projects: Trickler

- Trickler References:

  Source is entirely UNCLASSIFIED

  - www.truststc.org/pubs/256/Berkeley.pdf
  - www.defenselink.mil/comptroller/defbudge
  - http://www.nsa.gov/techtrans/techt00004.c

Known Projects: Trickler

- Department of Defense Project
  - Source is entirely unclassified
  - Source is publicly available (Tech Transfer)
- Not signature based
  - Grabs server/client strings
- MySQL Backend

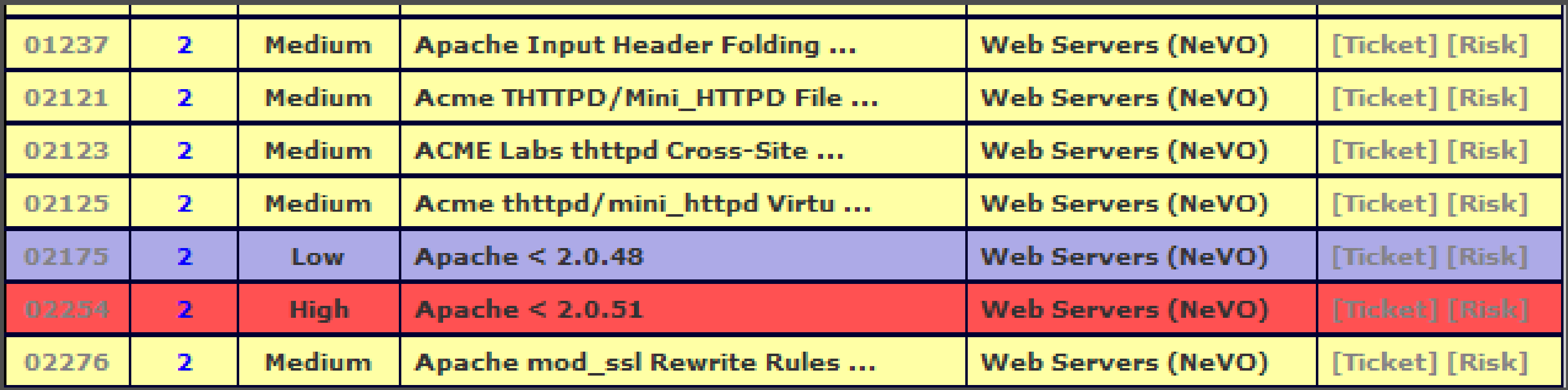
Real World Capacity

- PVS
  - GigE
  - Backbones of major organizations
- Trickler
  - Ask the government
- Endace DAG Cards:
  - OC-48+
  - Observed at >10Gbs
- Bivio
  - 10Gbs

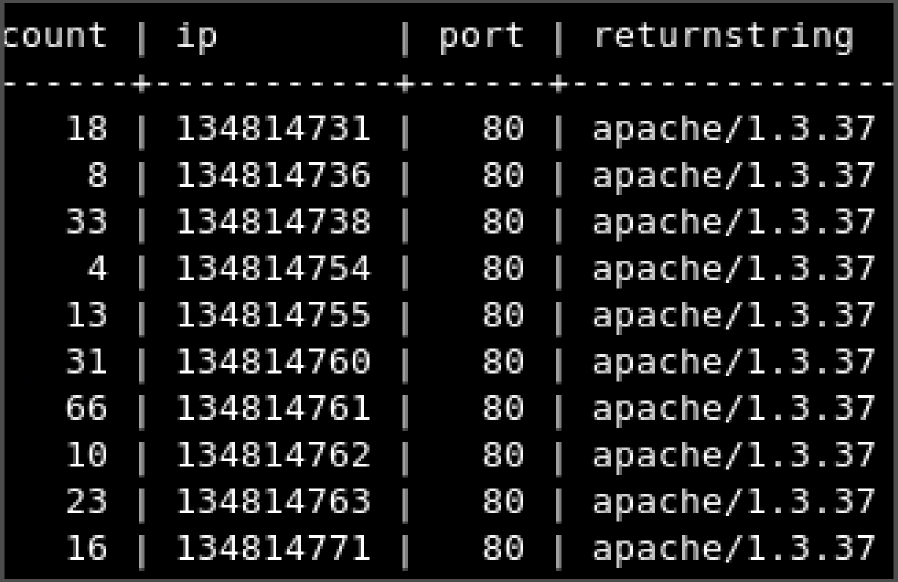
Finding Vulnerabilities

- PVS
  - Based on Nessus scripts
- Software Versioning
  - Grab Version strings
  - Compare version strings
    - CVE
    - NVD
    - Parsing/Correlating can be difficult

Passive Vulnerability Identification



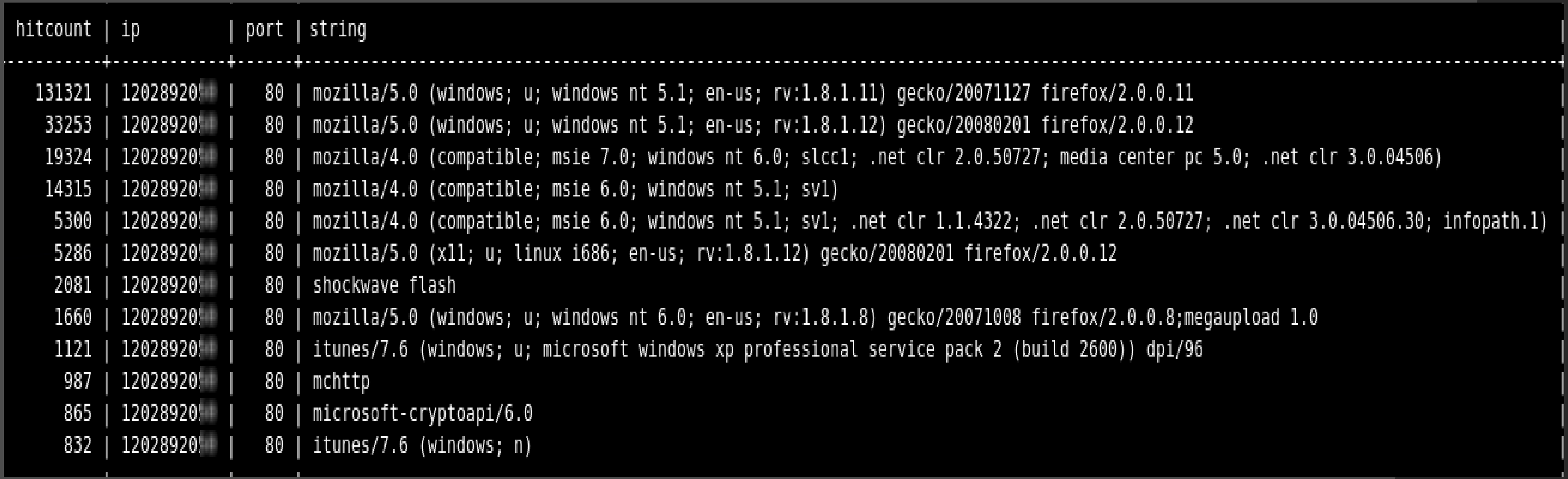
01237	2	Medium	Apache Input Header Folding ...	Web Servers (NeVO)	[Ticket] [Risk]
02121	2	Medium	Acme THTTPD/Mini_HTTPD File ...	Web Servers (NeVO)	[Ticket] [Risk]
02123	2	Medium	ACME Labs thttpd Cross-Site ...	Web Servers (NeVO)	[Ticket] [Risk]
02125	2	Medium	Acme thttpd/mini_httpd Virtu ...	Web Servers (NeVO)	[Ticket] [Risk]
02175	2	Low	Apache < 2.0.48	Web Servers (NeVO)	[Ticket] [Risk]
02254	2	High	Apache < 2.0.51	Web Servers (NeVO)	[Ticket] [Risk]
02276	2	Medium	Apache mod_ssl Rewrite Rules ...	Web Servers (NeVO)	[Ticket] [Risk]



```
count | ip        | port | returnstring
------+-----------+------+--------------
   18 | 134814731 |   80 | apache/1.3.37
    8 | 134814736 |   80 | apache/1.3.37
   33 | 134814738 |   80 | apache/1.3.37
    4 | 134814754 |   80 | apache/1.3.37
   13 | 134814755 |   80 | apache/1.3.37
   31 | 134814760 |   80 | apache/1.3.37
   66 | 134814761 |   80 | apache/1.3.37
   10 | 134814762 |   80 | apache/1.3.37
   23 | 134814763 |   80 | apache/1.3.37
   16 | 134814771 |   80 | apache/1.3.37
```

Host Characterization: Knowing Your Network

- What’s the most common client traffic on your network?

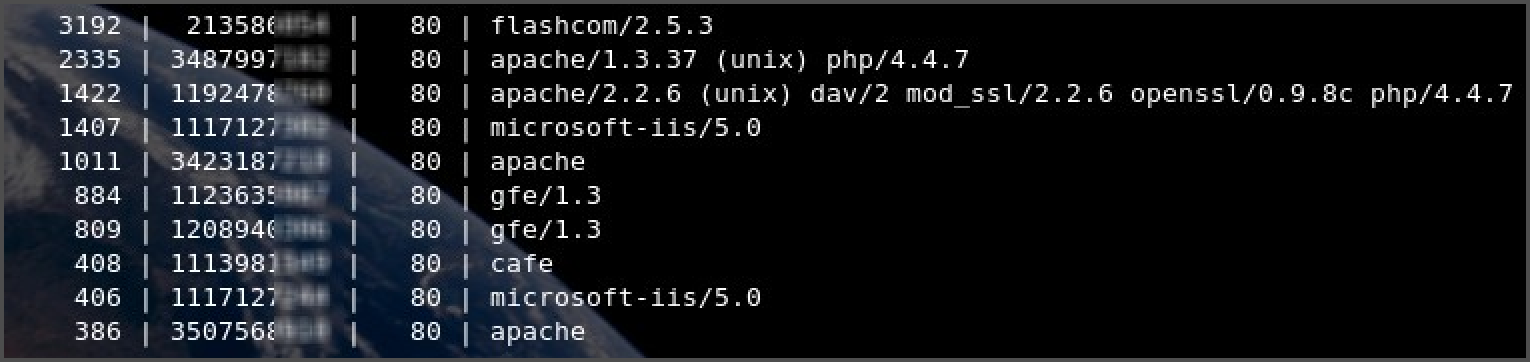


```
 hitcount | ip                   | port | string                                                                                                  |
----------+----------------------+------+---------------------------------------------------------------------------------------------------------+
   131321 | 12028920[redacted]   |   80 | mozilla/5.0 (windows; u; windows nt 5.1; en-us; rv:1.8.1.11) gecko/20071127 firefox/2.0.0.11            |
    33253 | 12028920[redacted]   |   80 | mozilla/5.0 (windows; u; windows nt 5.1; en-us; rv:1.8.1.12) gecko/20080201 firefox/2.0.0.12            |
    19324 | 12028920[redacted]   |   80 | mozilla/4.0 (compatible; msie 7.0; windows nt 6.0; slcc1; .net clr 2.0.50727; media center pc 5.0; .net clr 3.0.04506) |
    14315 | 12028920[redacted]   |   80 | mozilla/4.0 (compatible; msie 6.0; windows nt 5.1; sv1)                                                 |
     5300 | 12028920[redacted]   |   80 | mozilla/4.0 (compatible; msie 6.0; windows nt 5.1; sv1; .net clr 1.1.4322; .net clr 2.0.50727; .net clr 3.0.04506.30; infopath.1) |
     5286 | 12028920[redacted]   |   80 | mozilla/5.0 (x11; u; linux i686; en-us; rv:1.8.1.12) gecko/20080201 firefox/2.0.0.12                    |
     2081 | 12028920[redacted]   |   80 | shockwave flash                                                                                         |
     1660 | 12028920[redacted]   |   80 | mozilla/5.0 (windows; u; windows nt 6.0; en-us; rv:1.8.1.8) gecko/20071008 firefox/2.0.0.8;megaupload 1.0 |
     1121 | 12028920[redacted]   |   80 | itunes/7.6 (windows; u; microsoft windows xp professional service pack 2 (build 2600)) dpi/96          |
      987 | 12028920[redacted]   |   80 | mchttp                                                                                                  |
      865 | 12028920[redacted]   |   80 | microsoft-cryptoapi/6.0                                                                                 |
      832 | 12028920[redacted]   |   80 | itunes/7.6 (windows; n)                                                                                 |
```

Foundstone
A division of McAfee

Servers

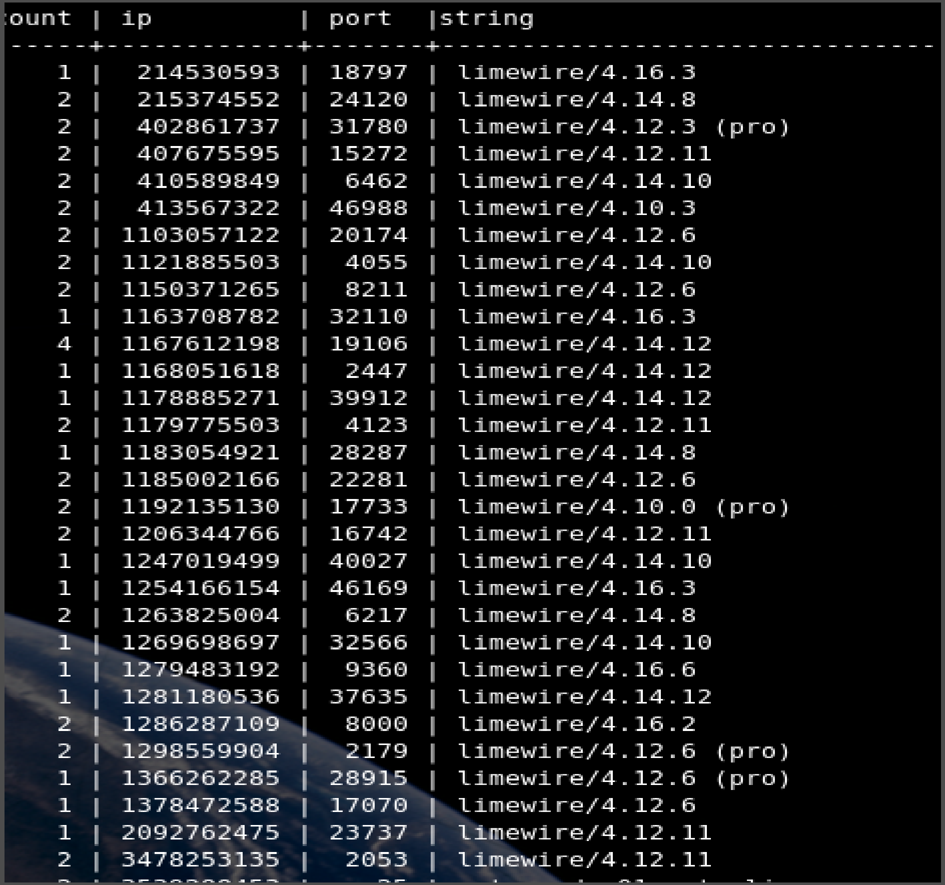
Server Traffic



```
   3192 |  213586[redacted] |    80 | flashcom/2.5.3
   2335 | 3487997[redacted] |    80 | apache/1.3.37 (unix) php/4.4.7
   1422 | 1192478[redacted] |    80 | apache/2.2.6 (unix) dav/2 mod_ssl/2.2.6 openssl/0.9.8c php/4.4.7
   1407 | 1117127[redacted] |    80 | microsoft-iis/5.0
   1011 | 3423187[redacted] |    80 | apache
    884 | 1123635[redacted] |    80 | gfe/1.3
    809 | 1208940[redacted] |    80 | gfe/1.3
    408 | 1113981[redacted] |    80 | cafe
    406 | 1117127[redacted] |    80 | microsoft-iis/5.0
    386 | 3507568[redacted] |    80 | apache
```

Practical Uses: System management

- What’s on your network that maybe shouldn’t be?



```
ount | ip         | port  |string
-----+------------+-------+-----------------------------
   1 |  214530593 | 18797 | limewire/4.16.3
   2 |  215374552 | 24120 | limewire/4.14.8
   2 |  402861737 | 31780 | limewire/4.12.3 (pro)
   2 |  407675595 | 15272 | limewire/4.12.11
   2 |  410589849 |  6462 | limewire/4.14.10
   2 |  413567322 | 46988 | limewire/4.10.3
   2 | 1103057122 | 20174 | limewire/4.12.6
   2 | 1121885503 |  4055 | limewire/4.14.10
   2 | 1150371265 |  8211 | limewire/4.12.6
   1 | 1163708782 | 32110 | limewire/4.16.3
   4 | 1167612198 | 19106 | limewire/4.14.12
   1 | 1168051618 |  2447 | limewire/4.14.12
   1 | 1178885271 | 39912 | limewire/4.14.12
   2 | 1179775503 |  4123 | limewire/4.12.11
   1 | 1183054921 | 28287 | limewire/4.14.8
   2 | 1185002166 | 22281 | limewire/4.12.6
   2 | 1192135130 | 17733 | limewire/4.10.0 (pro)
   2 | 1206344766 | 16742 | limewire/4.12.11
   1 | 1247019499 | 40027 | limewire/4.14.10
   1 | 1254166154 | 46169 | limewire/4.16.3
   2 | 1263825004 |  6217 | limewire/4.14.8
   1 | 1269698697 | 32566 | limewire/4.14.10
   1 | 1279483192 |  9360 | limewire/4.16.6
   1 | 1281180536 | 37635 | limewire/4.14.12
   2 | 1286287109 |  8000 | limewire/4.16.2
   2 | 1298559904 |  2179 | limewire/4.12.6 (pro)
   1 | 1366262285 | 28915 | limewire/4.12.6 (pro)
   1 | 1378472588 | 17070 | limewire/4.12.6
   1 | 2092762475 | 23737 | limewire/4.12.11
   2 | 3478253135 |  2053 | limewire/4.12.11
```

Practical Uses: Penetration Testing

- Pen Tests vary – but some customers want testers to represent a stealthy attacker such as an insider or sophisticated corporate espionage
- Not possible to go slow on typical time/budget
- A tool like PHC gives you insider information or what you'd learn if you went slow for a long period

Practical Uses: DNS Exfiltration Detection

- Outbound DNS requests are generally allowed outbound in every enterprise
- Data can be exfiltrated without breaking the protocol.
- Ozymandns is publicly available tool
- Other commercial tools exists

Practical Uses: DNS Exfiltration Detection

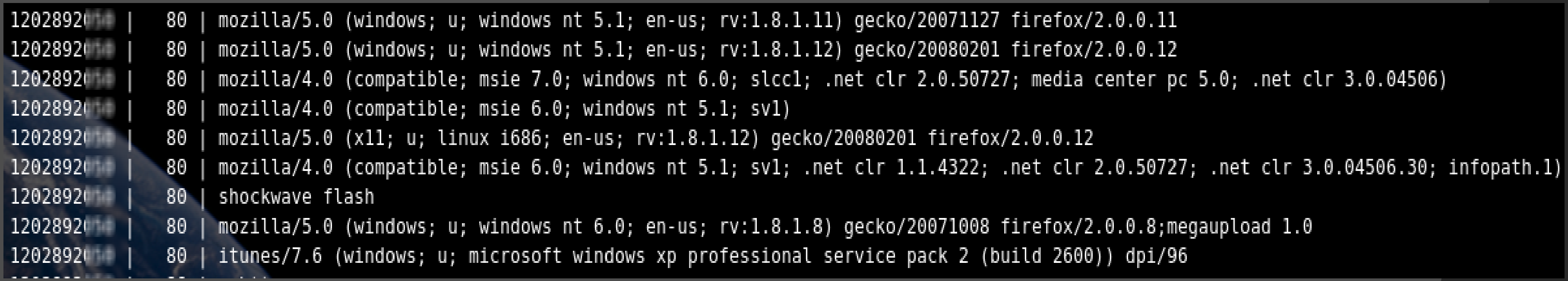
- Inspecting individual DNS messages is difficult to determine abusive content
- Communication is has identifiable characteristics
  - Messages tend to be longer
  - Messages tend to be more frequent
  - Messages have high entropy (nightmare to store in db)

Practical Uses: NAT Detection

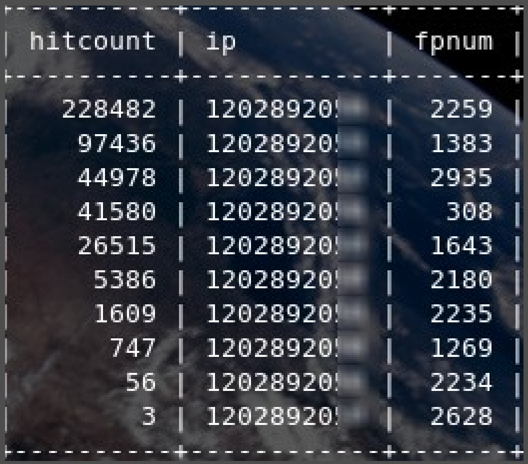
- Wireless NATs are a significant and present risk to many enterprises
- Port security is difficult across an enterprise
- NATs have identifiable characteristics
  - More traffic
  - Multiple OS identification
  - Cross platform services (MS IIS and SSH)
  - Cross platform browsers

Practical Uses: NAT Detection

- Example:



```
1202892[redacted] |   80 | mozilla/5.0 (windows; u; windows nt 5.1; en-us; rv:1.8.1.11) gecko/20071127 firefox/2.0.0.11
1202892[redacted] |   80 | mozilla/5.0 (windows; u; windows nt 5.1; en-us; rv:1.8.1.12) gecko/20080201 firefox/2.0.0.12
1202892[redacted] |   80 | mozilla/4.0 (compatible; msie 7.0; windows nt 6.0; slcc1; .net clr 2.0.50727; media center pc 5.0; .net clr 3.0.04506)
1202892[redacted] |   80 | mozilla/4.0 (compatible; msie 6.0; windows nt 5.1; sv1)
1202892[redacted] |   80 | mozilla/5.0 (x11; u; linux i686; en-us; rv:1.8.1.12) gecko/20080201 firefox/2.0.0.12
1202892[redacted] |   80 | mozilla/4.0 (compatible; msie 6.0; windows nt 5.1; sv1; .net clr 1.1.4322; .net clr 2.0.50727; .net clr 3.0.04506.30; infopath.1)
1202892[redacted] |   80 | shockwave flash
1202892[redacted] |   80 | mozilla/5.0 (windows; u; windows nt 6.0; en-us; rv:1.8.1.8) gecko/20071008 firefox/2.0.0.8;megaupload 1.0
1202892[redacted] |   80 | itunes/7.6 (windows; u; microsoft windows xp professional service pack 2 (build 2600)) dpi/96
```



```
+----------+------------+-------+
| hitcount | ip         | fpnum |
+----------+------------+-------+
|   228482 | 12028920[redacted] |  2259 |
|    97436 | 12028920[redacted] |  1383 |
|    44978 | 12028920[redacted] |  2935 |
|    41580 | 12028920[redacted] |   308 |
|    26515 | 12028920[redacted] |  1643 |
|     5386 | 12028920[redacted] |  2180 |
|     1609 | 12028920[redacted] |  2235 |
|      747 | 12028920[redacted] |  1269 |
|       56 | 12028920[redacted] |  2234 |
|        3 | 12028920[redacted] |  2628 |
+----------+------------+-------+
```

Research Uses: Detecting Network Bridges

- Consider a host connected to an enterprise network and then has an additional unauthorized network connection – say EVDO.
- Secondary connection (EVDO) is default gateway
  - Normal for bypassing corporate policy
- Host will have notably different characterization:
  - No observed external traffic except maybe DNS lookups
  - Internal Traffic (corporate web/etc)
  - IE is latest and greatest (it’s patched)

Research Uses: Fast Flux

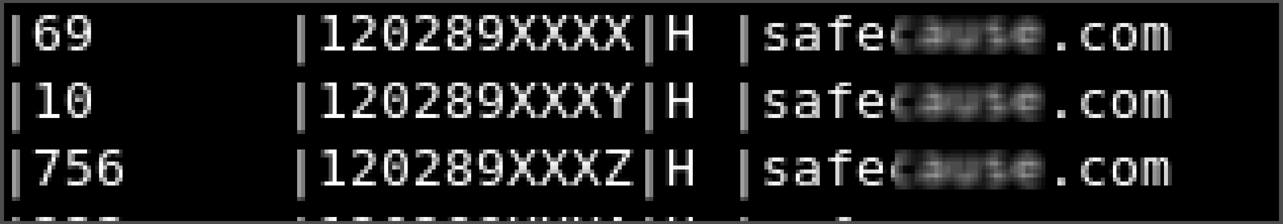
- Fast flux is a modern and effective bot tool
  - Uses short DNS TTLs to host or proxy websites across many infected machines
  - Fast flux is difficult to block because the sites are spread across many IP addresses
  - IDS/IPS need a signature or IP – thus its too late

Research Uses: Fast Flux

- Fast Flux has identifiable characteristics:
  - DNS responses with short TTL
  - FQDN with many IP addresses (though redundant hosts have this too)
  - DNS servers where they shouldn’t be
- IDS can sometimes identify same traits
  - False positives are high
  - I’ve never seen an IDS on a > GigE pipe

Research Uses: Fast Flux

- Example:



```
|69         |120289XXXX|H  |safe[redacted].com
|10         |120289XXXY|H  |safe[redacted].com
|756        |120289XXXZ|H  |safe[redacted].com
```

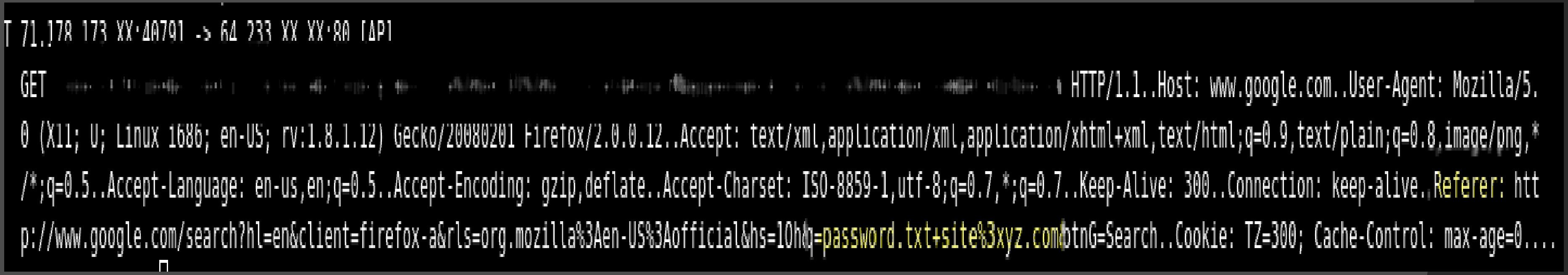
Research Uses: Threat Modeling

- Attacker’s software vulnerable just like the rest of us
- What O/S do attacker run?
- What tools are they using?

- The better you know what your attacker looks like the better you can block them
- Create rules based on characteristics rather than IPs – which change more quickly

Research Uses: Threat Modeling

- Attackers can use Google like the rest of us
  - Detect them before they even attack



```
T 71.178.173.XX:40791 -> 64.233.XX.XX:80 [AP]
  GET [redacted] HTTP/1.1..Host: www.google.com..User-Agent: Mozilla/5.
  0 (X11; U; Linux i686; en-US; rv:1.8.1.12) Gecko/20080201 Firefox/2.0.0.12..Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,image/png,*
  /*;q=0.5..Accept-Language: en-us,en;q=0.5..Accept-Encoding: gzip,deflate..Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7..Keep-Alive: 300..Connection: keep-alive..Referer: htt
  p://www.google.com/search?hl=en&client=firefox-a&rls=org.mozilla%3Aen-US%3Aofficial&hs=10h&q=password.txt+site%3xyz.com&btnG=Search..Cookie: TZ=300; Cache-Control: max-age=0....
```

Research Uses: Borrowing From Beale

- Deep document inspection
- Could we parse documents at network speed?
  - We can’t rebuild the document – too much memory
  - We can’t rebuild the document – we don’t keep state anyway
- We probably don’t need to rebuild the doc
  - Ethernet frames are usually 1500B
  - Probably big enough to grab some meta-data
  - Create a binary trigger and take snapshots
    - Enough to tie document version/author to IP (maybe?)

Future: Network Characterization

- Enterprises are often aware of “problem” networks
  - Incidents trigger identification
  - Scanning triggers identification
- Malicious networks can be characterized. For example:
  - Host O/S
  - Client Software (old IE)
  - Unneeded services running

Future: Losing The Database

- Currently the backend database is the leading limitation of large datasets
- Schema and Indexing need to be optimized to reasonably perform some queries
- G2 and Lexis Nexus are partnering to use LN’s technology
  - No indexing required
  - Some pre-processing overhead
  - Most queries complete in about the same time as an indexed DB query
  - Analysts can more easily perform complex queries in new ways

Conclusion

- PHC can be a powerful tool built on simple technology
- Can scale to any enterprise
- PoC Demo Code Available (soon) at:

  www.cyberwart.com/phc-demo.tgz

Questions