XKEYSCORE: Technology, Jamming

Robert Graham
robert_david_graham@yahoo.com
@ErrataRob

Code

- Two recently released bits of source code
  - January New York Times piece that improperly redacted a PDF
  - July ARD/NDR piece that targets Tor http://daserste.ndr.de/panorama/aktuell/nsa230_page-1.html

whoami = coder

- Created first IPS (BlackICE Guard)
  - A DPI product
- Did the first sidejacking
  - A major DPI hack
- https://github.com/robertdavidgraham
  - “masscan” port scanner
  - “ferret” DPI tool
  - exploits
  - other

1. NSA XKEYSCORE
2. DPI (Deep Packet Inspection)
3. Code walkthrough
4. Jamming

NSA and XKEYSCORE

UNITED STATES

FOREIGN INTELLIGENCE SURVEILLANCE COURT

WASHINGTON, D.C.

IN RE APPLICATION OF THE
FEDERAL BUREAU OF INVESTIGATION
FOR AN ORDER REQUIRING THE
PRODUCTION OF TANGIBLE THINGS
FROM VERIZON BUSINESS NETWORK SERVICES,
INC. ON BEHALF OF MCI COMMUNICATION
SERVICES, INC. D/B/A VERIZON
BUSINESS SERVICES.

Docket Number: BR 13-80

SECONDARY ORDER

This Court having found that the Application of the Federal Bureau of Investigation (FBI) for an Order requiring the production of tangible things from **Verizon Business Network Services, Inc. on behalf of MCI Communication Services Inc., d/b/a Verizon Business Services (individually and collectively "Verizon")** satisfies the requirements of 50 U.S.C. § 1861,

IT IS HEREBY ORDERED that, the Custodian of Records shall produce to the National Security Agency (NSA) upon service of this Order, and continue production

Derived from: Pleadings in the above-captioned docket
Declassify on: 12 April 2038

“People have unfairly demonized the NSA to a point that's too extreme. These are good people trying to do hard work for good reasons”

“collection”

“tasking”

“efficiency”


SUBJECTE:
THX 1138
CURRENT POSITION:
VAC SHAFT
LEVEL ONE
PROJECT:
OVERBUDGET
3410 UNITS

~~smarts~~

money

access

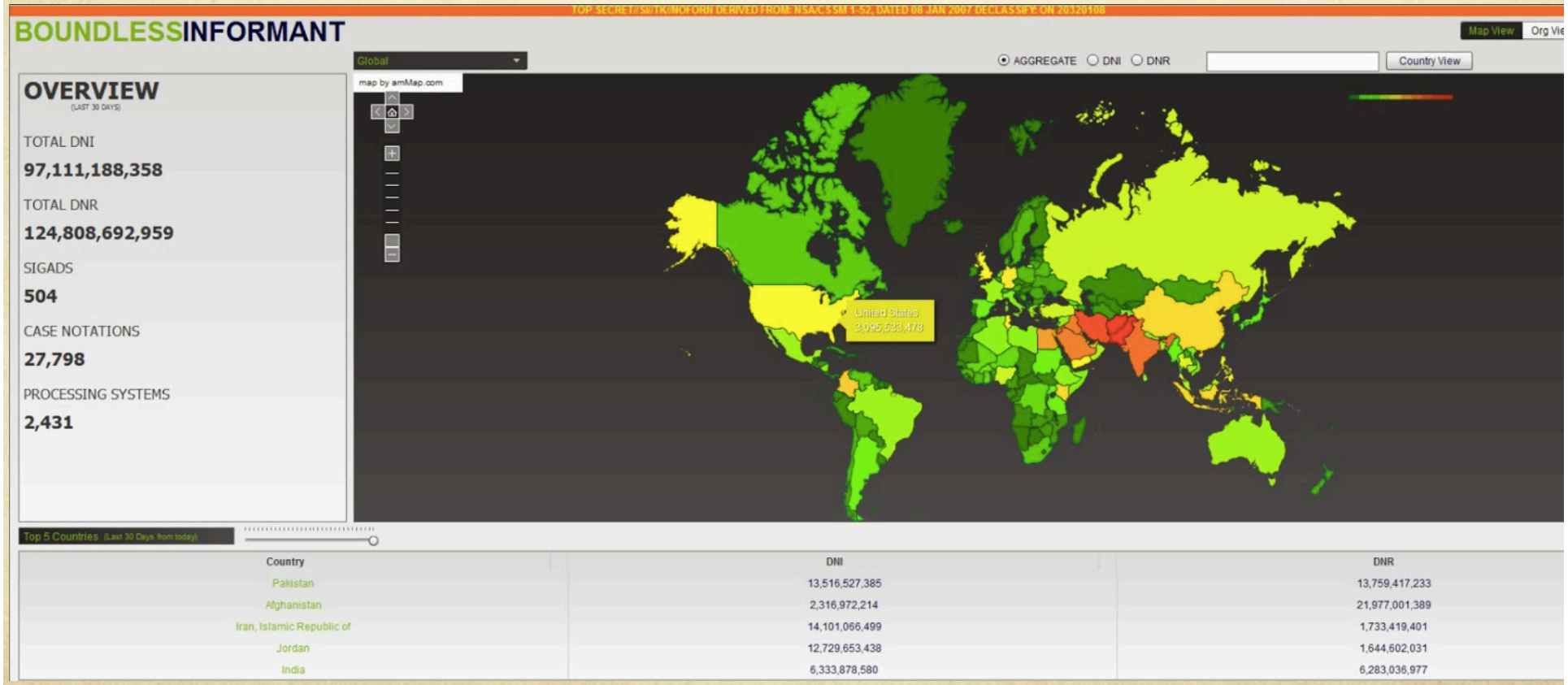
- CHANGELING: spoof email sender
- SHADOWCAT: SSH through Tor

[edit] **Techniques**

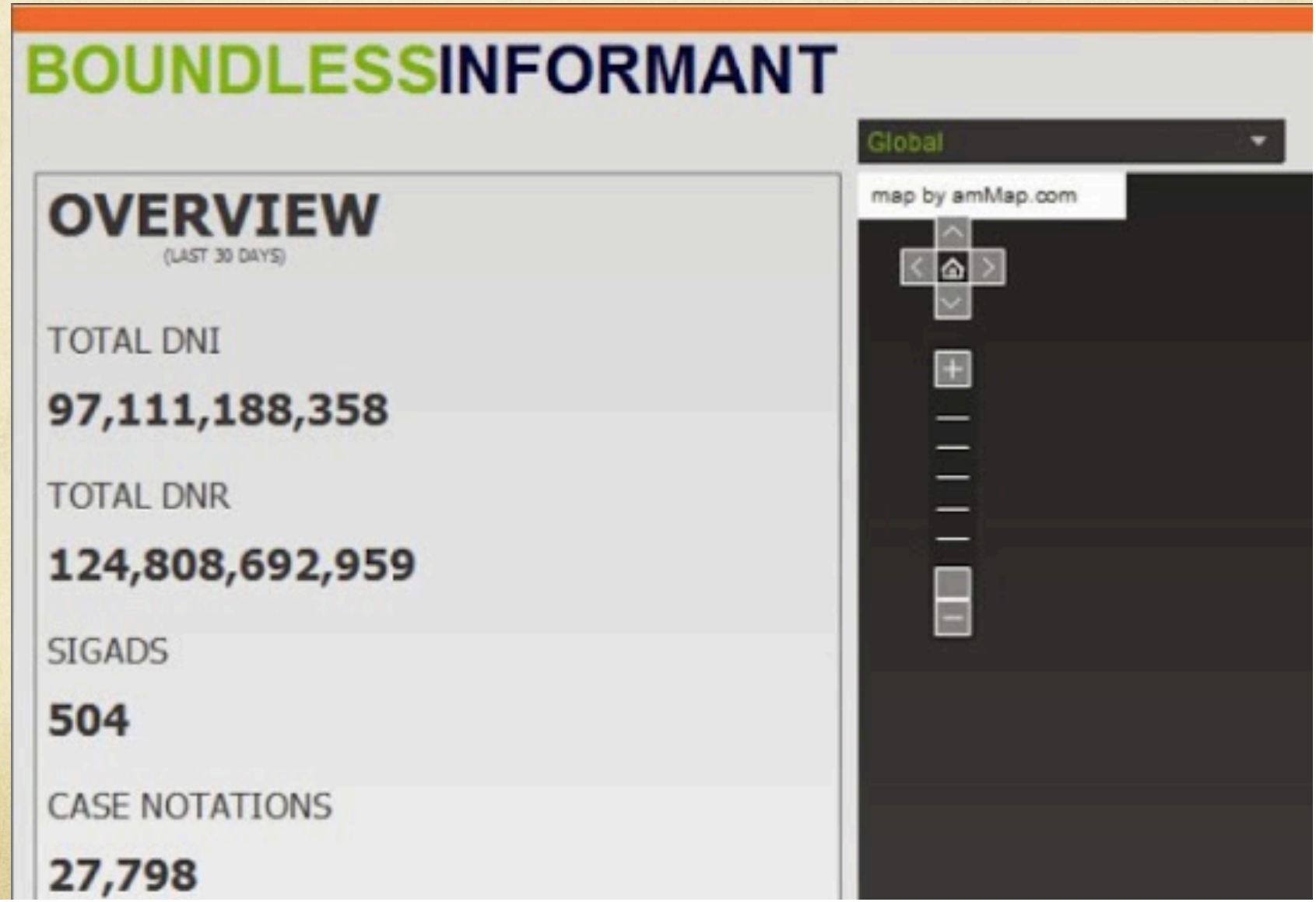
Tool	Description
CHANGELING	Ability to spoof any email address and send email under that identity
HAVOK	Real-time website cloning technique allowing on-the-fly alterations
MIRAGE	
SHADOWCAT	End-toEnd encrypted access to a VPS over SSH using the TOR network

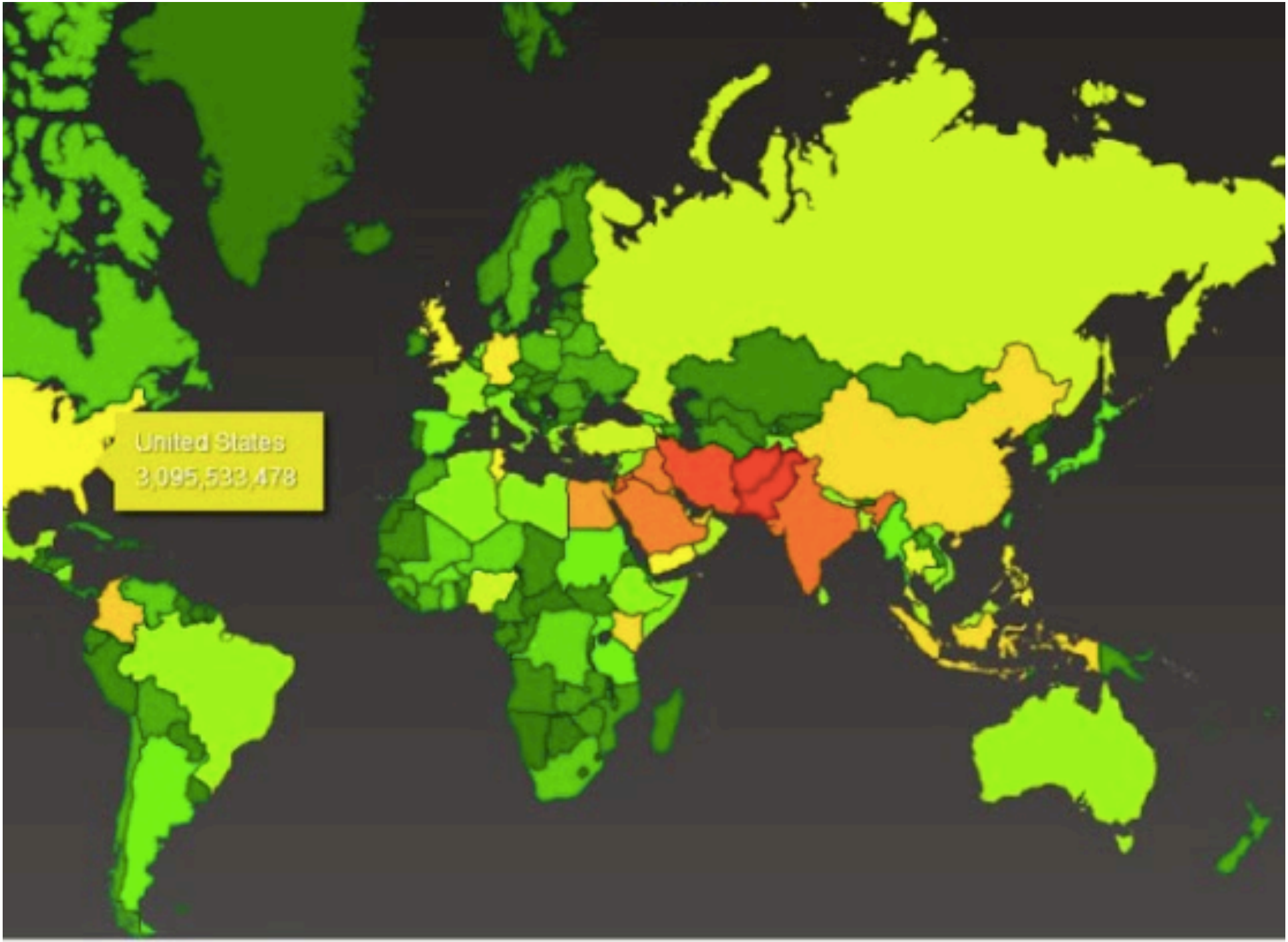
BOUNDLESSINFORMANT

- Counts pieces of metadata

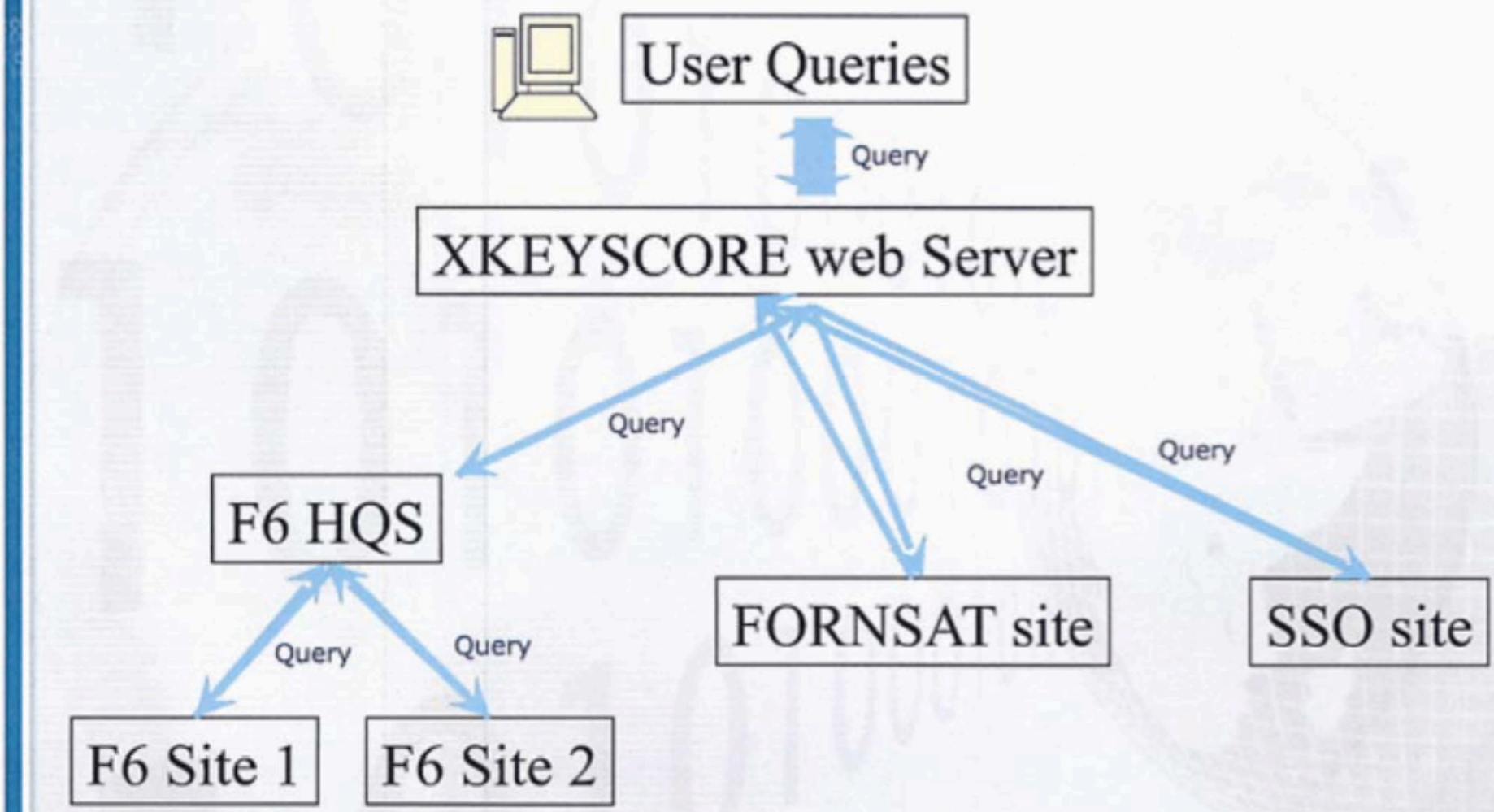


97 giga metadata per month



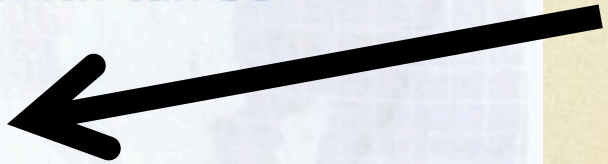

United States
3,095,533,478

Query Hierarchy



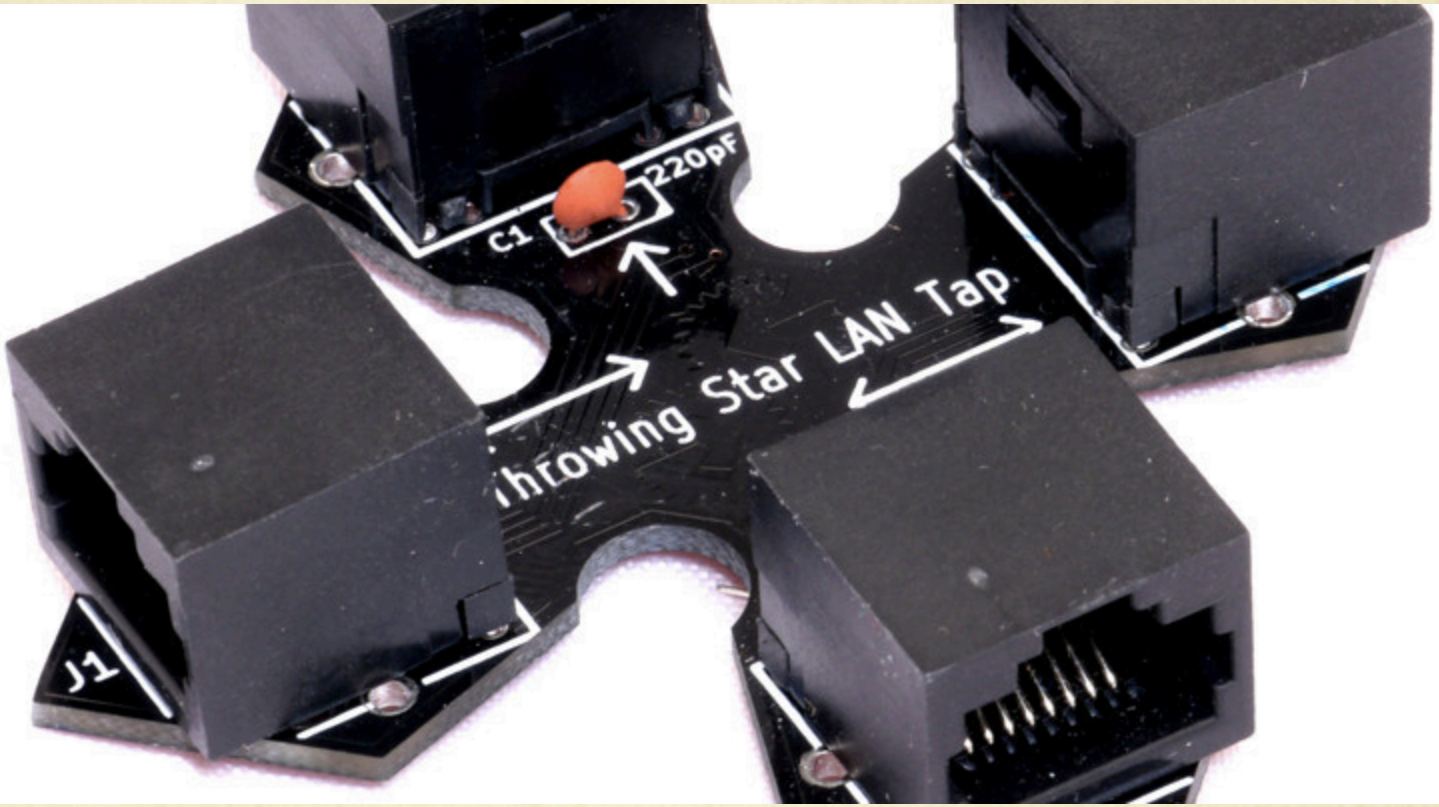
Google Maps

- My target uses Google Maps to scope target locations – can I use this information to determine his email address? What about the web-searches – do any stand out and look suspicious?
  - XKEYSCORE extracts and databases these events including all web-based searches which can be retrospectively queried
  - No strong-selector
  - Data volume too high to forward



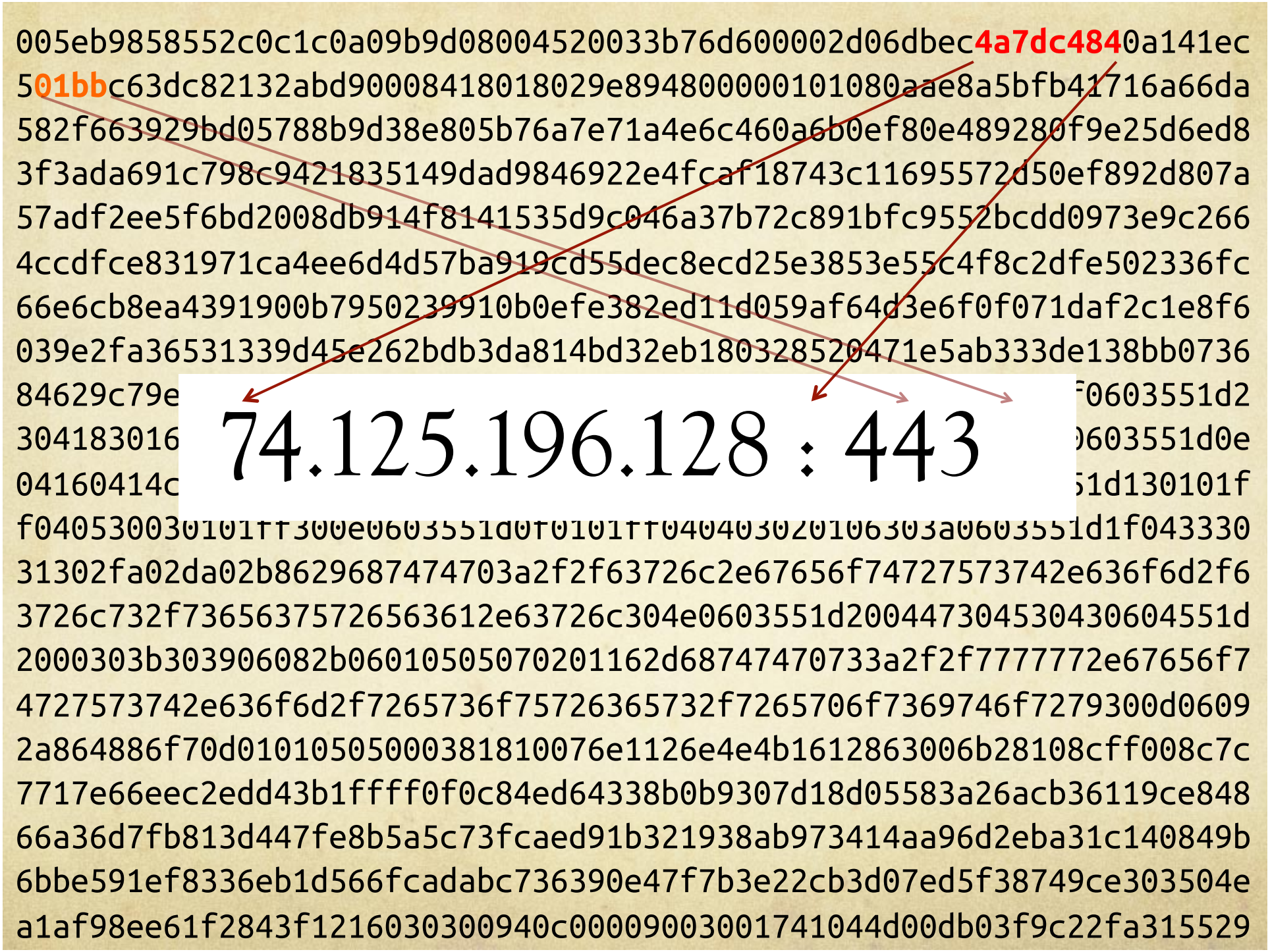
Deep Packet Inspection (DPI)

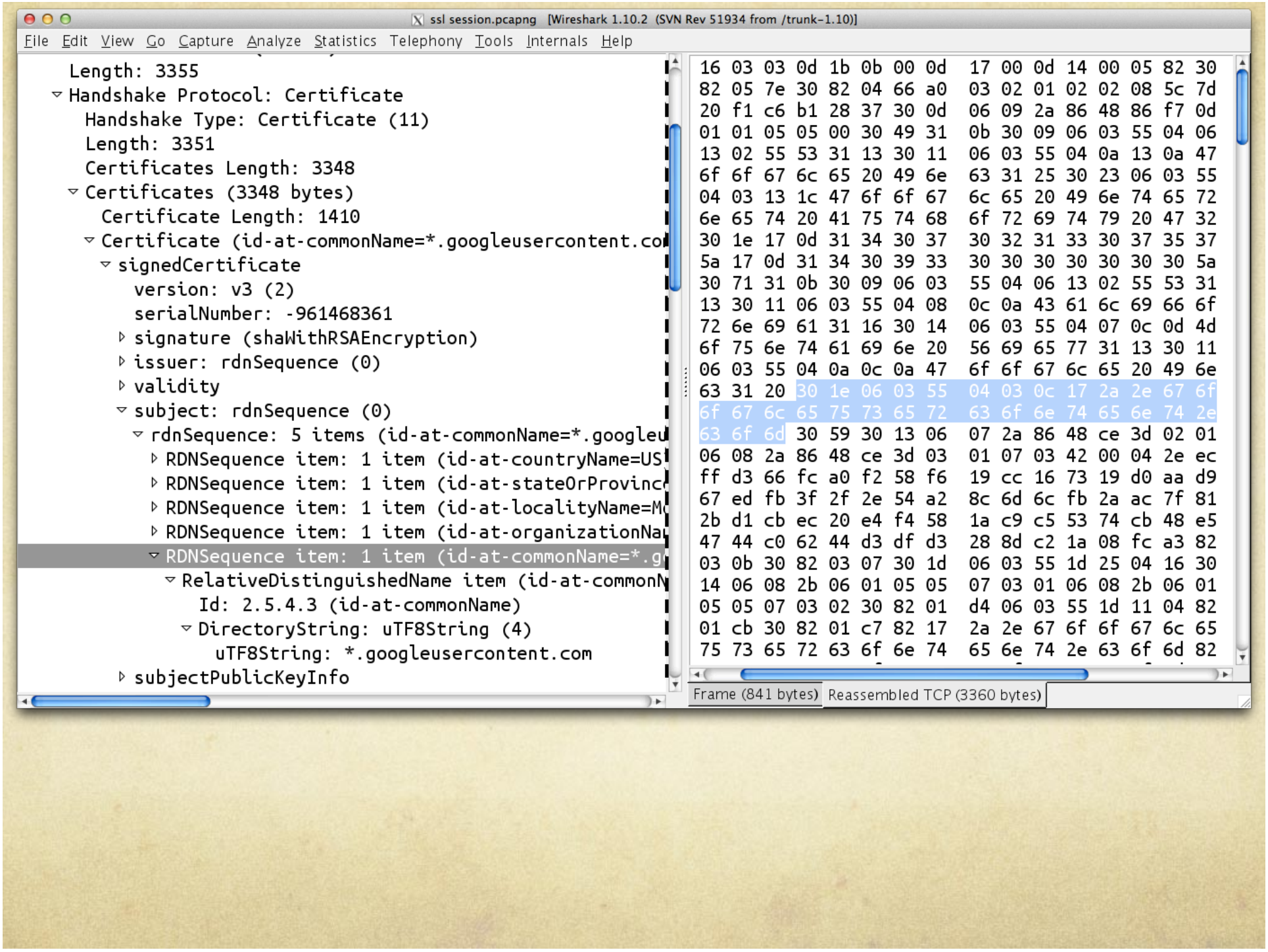
tap

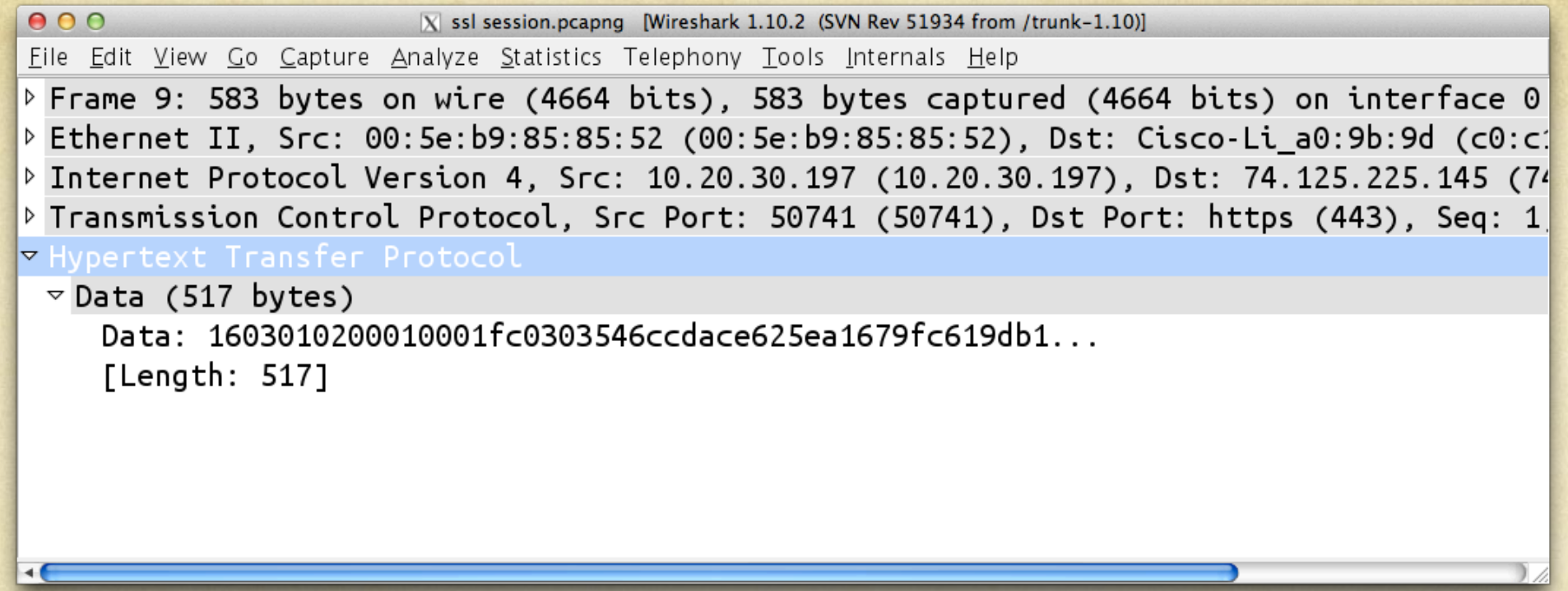


005eb9858552c0c1c0a09b9d08004520033b76d600002d06dbec4a7dc4840a141ec
501bbc63dc82132abd90008418018029e8948000001010 80aae8a5bfb41716a66da
582f663929bd05788b9d38e805b76a7e71a4e6c460a6b0ef80e489280f9e25d6ed8
3f3ada691c798c9421835149dad9846922e4fcaf18743c11695572d50ef892d807a
57adf2ee5f6bd2008db914f8141535d9c046a37b72c891bfc9552bcdd0973e9c266
4ccdfce831971ca4ee6d4d57ba919cd55dec8ecd25e3853e55c4f8c2dfe502336fc
66e6cb8ea4391900b7950239910b0efe382ed11d059af64d3e6f0f071daf2c1e8f6
039e2fa36531339d45e262bdb3da814bd32eb180328520471e5ab333de138bb0736
84629c79ea1630f45fc02be8716be4f90203010001a381f03081ed301f0603551d2
3041830168014 48e668f92bd2b295d747d82320104f3398909fd4301d0603551d0e
04160414c07a98688d89fbab05640c117daa7d65b8cacc4e300f0603551d130101f
f040530030101ff300e0603551d0f0101ff040403020106303a0603551d1f043330
31302fa02da02b8629687474703a2f2f63726c2e67656f74727573742e636f6d2f6
3726c732f736563726563612e63726c304e0603551d20044730453043060455 1d
2000303b303906082b06010505070201162d68747470733a2f2f7777772e67656f7
4727573742e636f6d2f7265736f75726365732f7265706f7369746f7279300d0609
2a864886f70d01010505000381810076e1126e4e4b1612863006b28108cff008c7c
7717e66eec2edd43b1ffff0f0c84ed64338b0b9307d18d05583a26acb36119ce848
66a36d7fb813d447fe8b5a5c73fcaed91b321938ab973414aa96d2eba31c140849b
6bbe591ef8336eb1d566fcadabc736390e47f7b3e22cb3d07ed5f38749ce303504e
a1af98ee61f2843f1216030300940c0000900300174 1044d00db03f9c22fa315529

..I don't see the code, I see blonde, brunette, redhead


005eb9858552c0c1c0a09b9d08004520033b76d600002d06dbec4a7dc4840a141ec
501bbc63dc82132abd90008418018029e8948000001010 80aae8a5bfb41716a66da
582f663929bd05788b9d38e805b76a7e71a4e6c460a6b0ef80e489280f9e25d6ed8
3f3ada691c798c9421835149dad9846922e4fcaf18743c11695572d50ef892d807a
57adf2ee5f6bd2008db914f8141535d9c046a37b72c891bfc9552bcdd0973e9c266
4ccdfce831971ca4ee6d4d57ba919cd55dec8ecd25e3853e55c4f8c2dfe502336fc
66e6cb8ea4391900b7950239910b0efe382ed11d059af64d3e6f0f071daf2c1e8f6
039e2fa36531339d45e262bdb3da814bd32eb18032852 0471e5ab333de138bb0736
84629c79e
304183016
04160414c
74.125.196.128 : 443
f0603551d2
9603551d0e
51d130101f
f04053003 0101ff300e0603551d0f0101ff040403020106303a0603551d1f043330
31302fa02da02b8629687474703a2f2f63726c2e67656f74727573742e636f6d2f6
3726c732f73656375726563612e63726c304e0603551d20044730453043060 4551d
2000303b303906082b06010505070201162d68747470733a2f2f7777772e67656f7
4727573742e636f6d2f7265736f75726365732f7265706f7369746f7279300d0609
2a864886f70d01010505000381810076e1126e4e4b1612863006b28108cff008c7c
7717e66eec2edd43b1ffff0f0c84ed64338b0b9307d18d05583a26acb36119ce848
66a36d7fb813d447fe8b5a5c73fcaed91b321938ab973414aa96d2eba31c140849b
6bbe591ef8336eb1d566fcadabc736390e47f7b3e22cb3d07ed5f38749ce303504e
a1af98ee61f2843f1216030300940c0000900300174 1044d00db03f9c22fa315529


ssl session.pcapng [Wireshark 1.10.2 (SVN Rev 51934 from /trunk-1.10)]
File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help
Length: 3355
Handshake Protocol: Certificate
Handshake Type: Certificate (11)
Length: 3351
Certificates Length: 3348
Certificates (3348 bytes)
Certificate Length: 1410
Certificate (id-at-commonName=*.googleusercontent.co
signedCertificate
version: v3 (2)
serialNumber: -961468361
signature (shaWithRSAEncryption)
issuer: rdnSequence (0)
validity
subject: rdnSequence (0)
rdnSequence: 5 items (id-at-commonName=*.googleu
RDNSequence item: 1 item (id-at-countryName=US
RDNSequence item: 1 item (id-at-stateOrProvinc
RDNSequence item: 1 item (id-at-localityName=M
RDNSequence item: 1 item (id-at-organizationNa
RDNSequence item: 1 item (id-at-commonName=*.g
RelativeDistinguishedName item (id-at-commonN
Id: 2.5.4.3 (id-at-commonName)
DirectoryString: uTF8String (4)
uTF8String: *.googleusercontent.com
subjectPublicKeyInfo
16 03 03 0d 1b 0b 00 0d 17 00 0d 14 00 05 82 30
82 05 7e 30 82 04 66 a0 03 02 01 02 02 08 5c 7d
20 f1 c6 b1 28 37 30 0d 06 09 2a 86 48 86 f7 0d
01 01 05 05 00 30 49 31 0b 30 09 06 03 55 04 06
13 02 55 53 31 13 30 11 06 03 55 04 0a 13 0a 47
6f 6f 67 6c 65 20 49 6e 63 31 25 30 23 06 03 55
04 03 13 1c 47 6f 6f 67 6c 65 20 49 6e 74 65 72
6e 65 74 20 41 75 74 68 6f 72 69 74 79 20 47 32
30 1e 17 0d 31 34 30 37 30 32 31 33 30 37 35 37
5a 17 0d 31 34 30 39 33 30 30 30 30 30 30 30 5a
30 71 31 0b 30 09 06 03 55 04 06 13 02 55 53 31
13 30 11 06 03 55 04 08 0c 0a 43 61 6c 69 66 6f
72 6e 69 61 31 16 30 14 06 03 55 04 07 0c 0d 4d
6f 75 6e 74 61 69 6e 20 56 69 65 77 31 13 30 11
06 03 55 04 0a 0c 0a 47 6f 6f 67 6c 65 20 49 6e
63 31 20 30 1e 06 03 55 04 03 0c 17 2a 2e 67 6f
6f 67 6c 65 75 73 65 72 63 6f 6e 74 65 6e 74 2e
63 6f 6d 30 59 30 13 06 07 2a 86 48 ce 3d 02 01
06 08 2a 86 48 ce 3d 03 01 07 03 42 00 04 2e ec
ff d3 66 fc a0 f2 58 f6 19 cc 16 73 19 d0 aa d9
67 ed fb 3f 2f 2e 54 a2 8c 6d 6c fb 2a ac 7f 81
2b d1 cb ec 20 e4 f4 58 1a c9 c5 53 74 cb 48 e5
47 44 c0 62 44 d3 df d3 28 8d c2 1a 08 fc a3 82
03 0b 30 82 03 07 30 1d 06 03 55 1d 25 04 16 30
14 06 08 2b 06 01 05 05 07 03 01 06 08 2b 06 01
05 05 07 03 02 30 82 01 d4 06 03 55 1d 11 04 82
01 cb 30 82 01 c7 82 17 2a 2e 67 6f 6f 67 6c 65
75 73 65 72 63 6f 6e 74 65 6e 74 2e 63 6f 6d 82
Frame (841 bytes)
Reassembled TCP (3360 bytes)


ssl session.pcapng [Wireshark 1.10.2 (SVN Rev 51934 from /trunk-1.10)]
File Edit View Go Capture Analyze Statistics Telephony Tools Internals Help
Frame 9: 583 bytes on wire (4664 bits), 583 bytes captured (4664 bits) on interface 0
Ethernet II, Src: 00:5e:b9:85:85:52 (00:5e:b9:85:85:52), Dst: Cisco-Li_a0:9b:9d (c0:c
Internet Protocol Version 4, Src: 10.20.30.197 (10.20.30.197), Dst: 74.125.225.145 (74
Transmission Control Protocol, Src Port: 50741 (50741), Dst Port: https (443), Seq: 1
Hypertext Transfer Protocol
Data (517 bytes)
Data: 16030102000100001fc0303546ccdace625ea1679fc619db1...
[Length: 517]

How XKeyScore works

- Collect 3-day ring-buffer of traffic
  - Like tcpdump/wireshark in ring-buffer mode
  - With BPF rules to cut down on traffic
- Post-process with a pipeline of tools
  - tshark, Snort, NetWitness
  - File extractors
  - Email parsers, image parsers, PDF parsers
  - *100s of tools*

tshark

```
smb.path contains "\\\\SERVER\\SHARE"

http and frame[100-199] contains "foobar"

x509sat.uTF8String == "*.google.com"

x509ce.dNSName == "*.android.com"
```

What's on the system

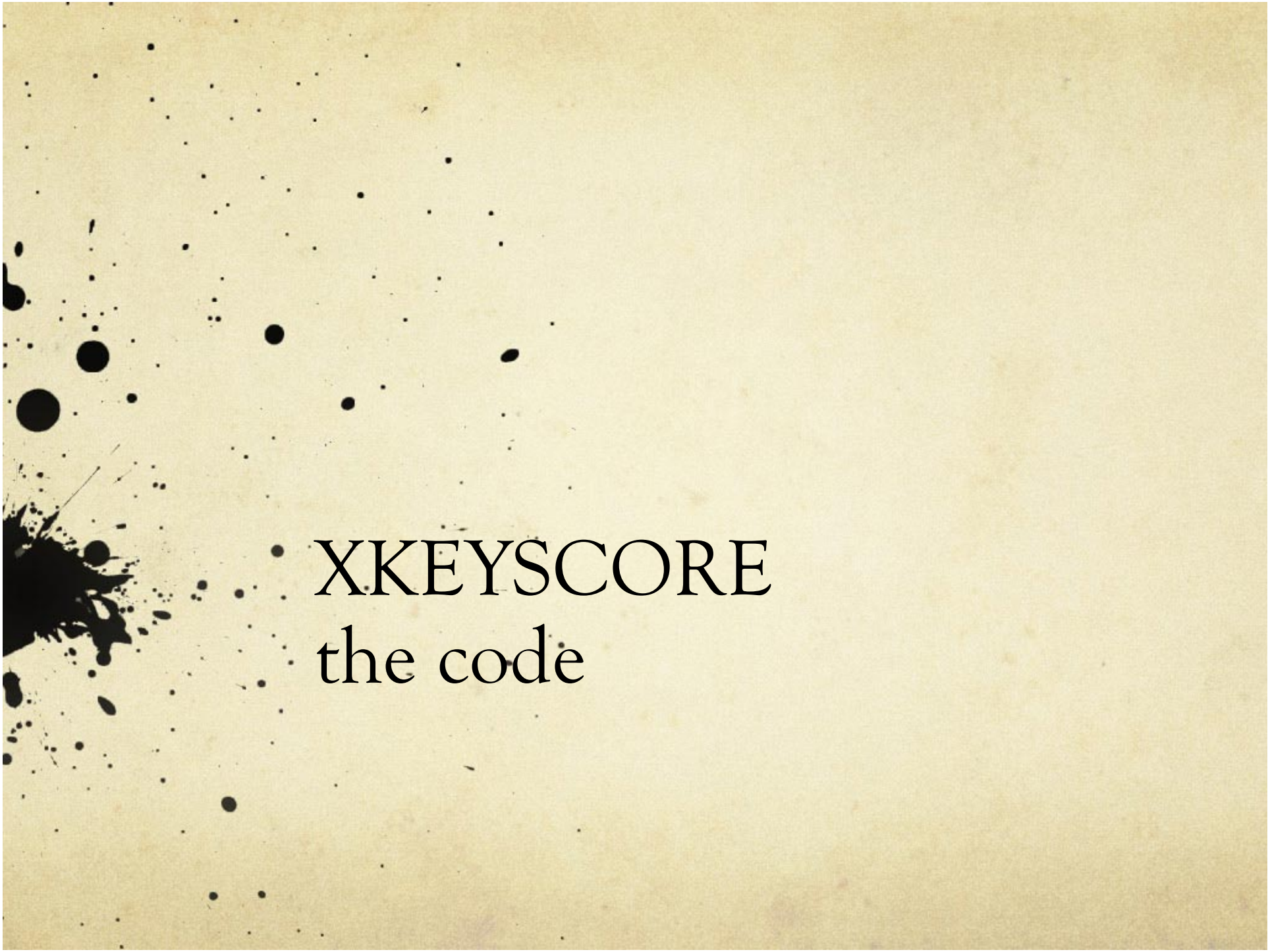
- Full capture (3 days)
- Extracted data (files, images)
- Metadata (lots of it)

How they get it

- Satellite links
- Private leased lines
- Often takes 3 days to get it back to Utah

Who gets it

- The analyst who tasked the system
- Automatic systems that are tasked to do something with the data
  - Like TAO attacks


XKEYSCORE
the code

```
fingerprint('image/exif/gpsCoordinates') =
 file_ext('jpeg' or 'pjpeg' or 'jpg' or
'pjpg' or 'tiff' or 'gif' or 'png' or 'riff'
or 'wav') and
 'exif:GPSLatitude' or 'exif:GPSLongitude'
or 'exif:GPSDestLatitude' or
'exif:GPSDestLongitude';
```

```
fingerprint('picture') =
 file_ext('jpeg');
```

- Tags the file with additional “meta-data”
- Allows analysts to construct queries like “show me all ‘pictures’ in ‘iraq’”.

```
fingerprint('image/exif/gpsCoordinates') =
 file_ext('jpeg') and
 'exif:GPSLatitude' or 'exif:GPSLongitude';
```

- Obvious to humans what this means.
- No obvious to computers

```
fingerprint('image/exif/gpsCoordinates') =
 ('exif:GPSLatitude' and file_ext('jpeg'))
      or
 ('exif:GPSLongitude');
```

- Most programming language groups it thusly

- It probably works anyway – by accident

```
/**
 * Fingerprint Tor authoritative directories
enacting the directory protocol.
 */

fingerprint('anonymizer/tor/node/authority')
= $tor_authority
  and ($tor_directory or preappid(/
anonymizer\/tor\/directory/));
```

```
/**
 * Identify clients accessing Tor bridge
information.
 */

fingerprint('anonymizer/tor/bridge/tls') =
 ssl_x509_subject('bridges.torproject.org')
or
 ssl_dns_name('bridges.torproject.org');
```

```
fingerprint('anonymizer/tor/bridge/
email') =
 email_address('bridges@torproject.org')
 and email_body('https://
bridges.torproject.org/' : c++

...

);
```

```
main: {{

for (size_t i=0; i < bridges.size(); ++i) {
  DB[SCHEMA_NEW]["tor_ip"] = bridges[i][0];
  DB[SCHEMA_NEW]["tor_port_or"] =
bridges[i][1];
  DB.apply();
}

xks::fire_fingerprint("anonymizer/tor/
directory/bridge");

return true;
}}
```

Some things to note

- Direct database access instead of API
  - No filtering done
  - Allows attackers direct access to the database

```
/*
The fingerprint identifies sessions
visiting the Tor Project website from
non-fvey countries.
*/
fingerprint('anonymizer/tor/
torpoject_visit') =
http_host('www.torproject.org’)
and not(xff_cc('US' OR 'GB' OR 'CA' OR 'AU'
OR 'NZ'));
```

```
/*

These variables define terms and websites
relating to the TAILs (The Amnesic
Incognito Live System) software program, a
comsec mechanism advocated by
extremists on extremist forums.

*/
```

```
$TAILS_terms = word('tails' or 'Amnesiac
Incognito Live System') and word('linux’ or
' USB ' or ' CD ' or 'secure desktop' or '
IRC ' or 'truecrypt' or ' tor ');
$TAILS_websites=('tails.boum.org/') or
('linuxjournal.com/content/linux*');
```

```
/* This fingerprint identifies users
searching for the TAILs software program,
viewing documents relating to TAILs, or
viewing websites that detail TAILs.
*/

fingerprint('ct_mo/TAILS')=
fingerprint('documents/comsec/tails_doc')
or web_search($TAILS_terms) or
url($TAILS_websites) or
html_title($TAILS_websites);
```

```
/**
 * Aggregate Tor hidden service addresses
seen in raw traffic.
 */
mapreduce::plugin('anonymizer/tor/plugin/
onion') =
  immediate_keyword(/(?:([a-z]+):\/\/){0,1}
([a-z2-7]{16})\.onion(?::(\d+)){0,1}/c : c+
+

…

);
```

https://aisyd7fglasdkhjf.onion:443

- Totally different regex from bridge, such as when matching optional port
  - (?::(\d+)){0,1}
  - :?([0-9]{2,4}?[^0-9])
- Why differences?
  - Different authors?
  - Copied/pasted from different sources?
  - Different times?
  - Different sources?

```
includes: {{
  #include <boost/lexical_cast.hpp>
}}
proto: {{
  message onion_t {
    required string address = 1;
    optional string scheme = 2;
    optional string port = 3;
  }
}}
```

```
MAPPER.map(onion.address(), onion);

xks::fire_fingerprint(prefix +
onion.address());
```

```
DB["tor_onion_survey"]["onion_address"] =
iter->address() + ".onion";
DB["tor_onion_survey"]["onion_scheme"] =
iter->scheme();
DB["tor_onion_survey"]["onion_port"] =
iter->port();
DB["tor_onion_survey"]["onion_count"] =
boost::lexical_cast<std::string>(TOTAL_VALU
E_COUNT);
DB.apply();
```


Jammin’

Image bugs

```
<img src="http://www.google.com/?
q=tails+usb" />
```

Who uses Tor?

- Make Tor servers dual-use
  - Once SSL is established, either serve files or establish tunnels
  - Tor-as-Apache-plugin
-
- Same with SSL-based VPN

No defenses against too much database info

```
https://bridges.torproject.org/
bridge = 0.0.0.1:443
bridge = 0.0.0.2:443
bridge = 0.0.0.3:443
```

Badly formed data

```
goscrewyourself://o987asgia7gsdfoi.onion:443/
```

masscan

- small packets
  - average packet size is 500 bytes, TCP scan is 40 bytes
- lots of connections
  - Average tool handles 100,000 connections, masscan generates billions of connections

P2P Traffic generation

- BitTorrent
  - Generates a lot of DHT traffic in background even when not downloading
- Bitcoin Wallet
  - Transfers gigabytes per month with thousands of people

Nonsensical traffic

- Setup servers with X.509 certificates claiming to be a Tor bridge
- ...on non-standard port if you have to
- Establish connections, watch database fill up

Data roulette

- Exchange meta-data with each other
  - Exchange cookies
  - Exchange search phrases
    - http://www.googlesharing.net/
- Contact each other
  - Notify each other out-of-band of incoming call/msg
  - Works on free nighttime minutes when in “roulette” mode
  - Works with email using IMAP4 to automatically delete messages before they fill your inbox

Exploitation

- XKEYSCORE collection uses 100s of tools
  - Open and closed source
  - Network parsers and file-format parsers
- There are a zillion 0days to be found
- Before disclosing your latest TIFF 0day, send exploitable images to your friends in Iraq
  - Like known non-existent accounts with unencrypted SMTP

Questions?

- I'll be down on the main conference floor for the next hour if you have any questions