



# Security & Privacy

... and why you should care.

Sarah Harvey

{CrySP, Information Retrieval} group

University of Waterloo

sharvey@cs.uwaterloo.ca

# Why this talk?



# Why this talk?

- Lots of media coverage on Snowden, NSA, and other government agencies
- Media coverage on sec/privacy policies companies
- Does the general public actually understand what all of this means?
- What are the implications of these findings?
  - Not just technological, but social context

# Who am I?

- ...mostly just some PhD student at UW
- Interested in Sec/Pri problems in IR systems
  - User profiling, user behavior
  - Privacy implications of profiling, linking
  - Improving privacy of large IR systems
- Interested promoting awareness of security, privacy systems

# Outline

---

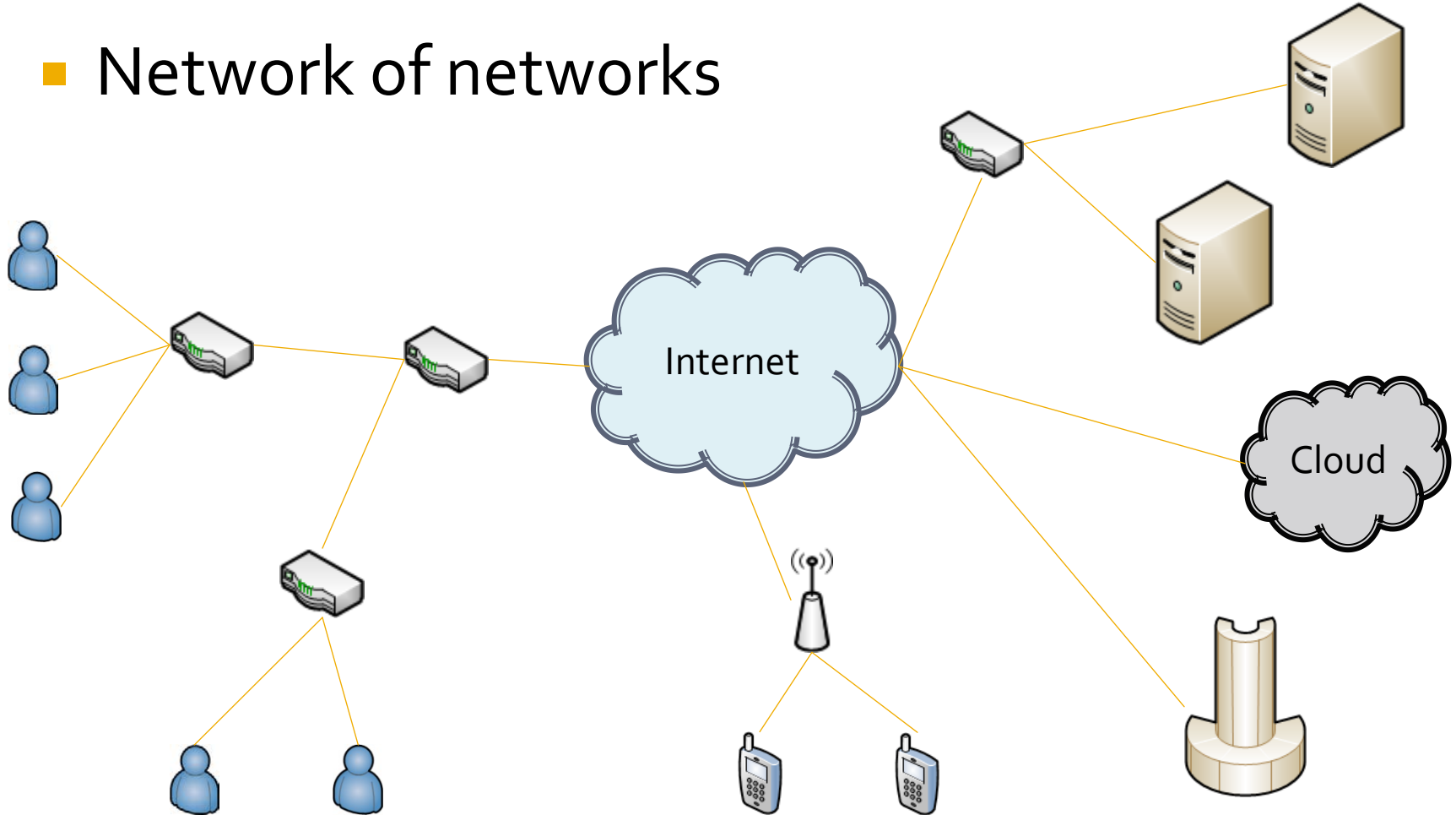
- Motivation
- What defines security? privacy?
- Who gets to see your stuff?
  - Who is the “bad guy”?
- Snowden and friends
- The issue of trust

# This, Jen... is the internet

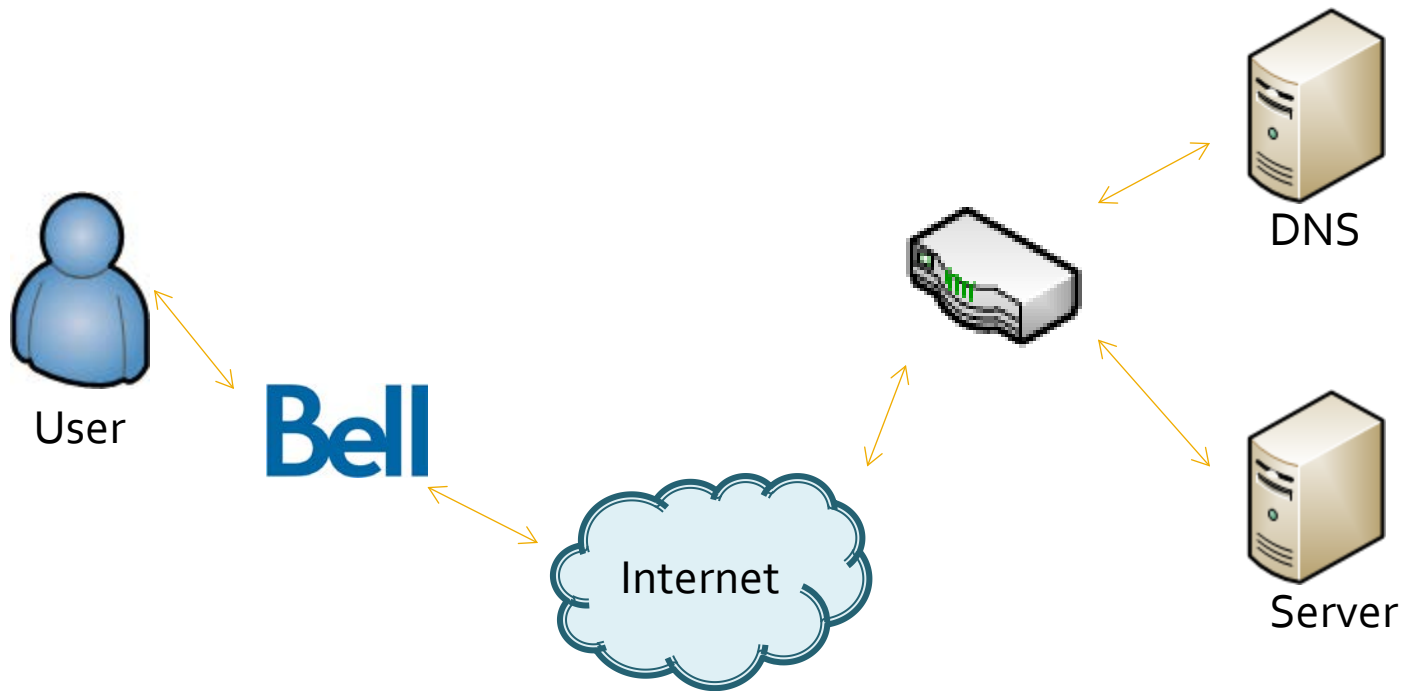


# What is the internet?

- Network of networks



# A typical internet pathway





# What's on the wire?

- Any of those people can see my stuff

No. ↓	Time	Source	Destination	Protocol	Info
25	2.232123	192.168.1.165	192.168.1.178	HTTP	GET /hiding.php HTTP/1.0
26	2.233117	192.168.1.178	192.168.1.165	TCP	http > 63801 [ACK] Seq=1 Ack=404 win=6432 Le
27	2.297124	192.168.1.178	192.168.1.165	HTTP	HTTP/1.1 200 OK (text/html)
28	2.297127	192.168.1.178	192.168.1.165	TCP	http > 63801 [FIN, ACK] Seq=377 Ack=404 win=
29	2.298120	192.168.1.165	192.168.1.178	TCP	63801 > http [ACK] Seq=404 Ack=378 win=63864

## HTTP/1.1 200 OK\r\n

Date: Mon, 18 May 2009 01:48:43 GMT\r\n

Server: Apache/2.2.8 (Unix) mod\_ssl/2.2.8 OpenSSL/0.9.8g DAV/2 PHP/5.2.6\r\n

X-Powered-By: PHP/5.2.6\r\n

Content-Encoding: gzip\r\n

Vary: Accept-Encoding\r\n

Content-Length: 109

Connection: close\r\n

Content-Type: text/html\r\n

\r\n

Content-encoded entity body (gzip): 109 bytes -> 100 bytes

## Line-based text data: text/html

<html>\n

<body>\n

<p>You can't read the content of this page while sniffing on wire.</p>\n

</html>\n

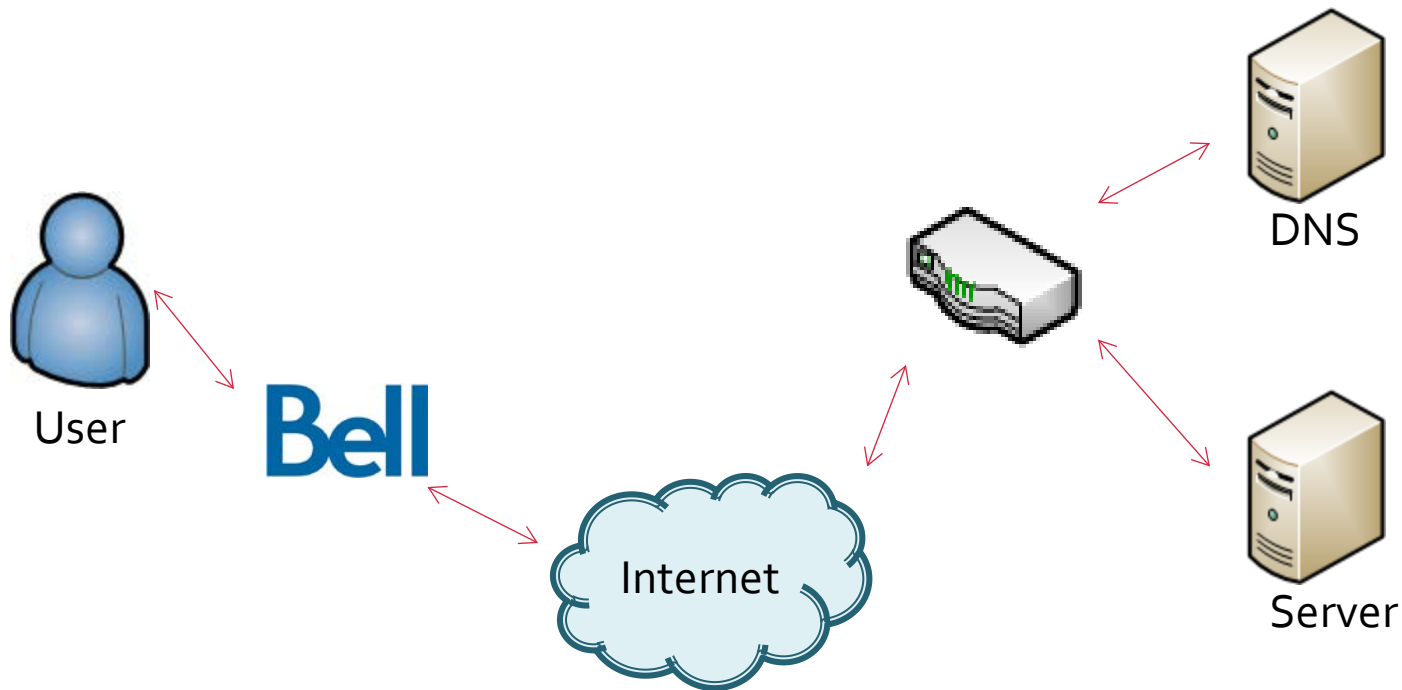
<body>\n

# Cryptography in 30 seconds

- So we have this wonderful technology called **cryptography**.
  - Encryption protects **confidentiality**.
  - MACs/digital signatures protect **integrity** and **authenticity**.
- Types of cryptographic systems:
  - Symmetric-key systems
  - Public-key systems

# Applying crypto

- Revisiting communication securely



# Revisiting the wire

The image shows a Wireshark capture of a network session. The main pane displays a list of packets, with packet 11 selected. The packet list shows the following details:

No.	Time	Source	Destination	Protocol	Info
6	2009-01-23 20:50:26.711545	10.88.229.196	10.88.229.209	TLSv1	Server Hello, Certificate, Server Hello C...
7	2009-01-23 20:50:26.711616	10.88.229.209	10.88.229.196	TCP	38353 > https [ACK] Seq=121 Ack=1449 win=...
8	2009-01-23 20:50:26.711647	10.88.229.209	10.88.229.196	TCP	38353 > https [ACK] Seq=121 Ack=1530 win=...
9	2009-01-23 20:50:26.713357	10.88.229.209	10.88.229.196	TLSv1	Client Key Exchange, Change Cipher Spec, ...
10	2009-01-23 20:50:26.717451	10.88.229.196	10.88.229.209	TLSv1	Change Cipher Spec, Encrypted Handshake M...
11	2009-01-23 20:50:26.717792	10.88.229.209	10.88.229.196	TLSv1	Application Data
12	2009-01-23 20:50:26.763286	10.88.229.196	10.88.229.209	TLSv1	Application Data
13	2009-01-23 20:50:26.763288	10.88.229.196	10.88.229.209	TLSv1	Application Data
14	2009-01-23 20:50:26.802843	10.88.229.209	10.88.229.196	TCP	38353 > https [ACK] Seq=485 Ack=2140 win=...
15	2009-01-23 20:50:26.815833	10.88.229.209	10.88.229.196	TLSv1	Application Data
16	2009-01-23 20:50:26.816144	10.88.229.209	10.88.229.196	TLSv1	Application Data
17	2009-01-23 20:50:26.816538	10.88.229.196	10.88.229.209	TCP	https > 38353 [ACK] Seq=2140 Ack=1577 wir...

The packet details pane for packet 11 shows the following structure:

- Frame 11 (248 bytes on wire, 248 bytes captured)
- Ethernet II, Src: HewlettP\_c3:c6:01 (00:14:c2:c3:c6:01), Dst: Vmware\_a2:58:b1 (00:50:56:a2:58:b1)
- Internet Protocol, Src: 10.88.229.209 (10.88.229.209), Dst: 10.88.229.196 (10.88.229.196)
- Transmission Control Protocol, Src Port: 38353 (38353), Dst Port: https (443), Seq: 303, Ack: 1573, Len: 182
- Secure Socket Layer
  - TLSv1 Record Layer: Application Data Protocol: http
    - Content Type: Application Data (23)
    - Version: TLS 1.0 (0x0301)
    - Length: 177
    - Encrypted Application Data: 166813911B828BE9E3710FC1F3BDCB91D1B9840376F1521D...

The packet bytes pane shows the raw data in hexadecimal and ASCII. The ASCII column shows the beginning of the encrypted data, which appears as a series of non-printable characters.

0040 0e 63 17 03 01 00 b1 16 68 13 91 1b 82 8b e9 e3 .C.....h.....  
0050 71 0f c1 f3 bd cb 91 d1 b9 84 03 76 f1 52 1d 7f q.....v.R..  
0060 d8 bd c1 76 1a 1c 0e d6 f3 4b 84 7c a5 04 38 71 ..v.....K|.8q  
0070 a5 50 1f 07 be 0c 5e d2 f0 36 9c 4c fb 36 18 6c .P.....6.L.6.l  
0080 93 19 6a a3 9e 04 f4 e5 2c 75 ad b0 7e d3 9b 86 .j.....u.....  
0090 94 b3 67 87 f7 af f4 2a 32 7d fc b9 0b 5d 4b 15 .g....\*2}...K.  
00a0 46 df 44 7c cb 08 2b 4a 53 c2 e6 23 24 62 c7 53 F.D[...+J S.##b.S  
00b0 cc c0 49 51 b5 b8 59 6a bd 6a f8 27 0f 95 c1 41 .IQ..Yj..j....A  
00c0 40 ca 28 1e b7 3e a1 08 aa ca b1 38 42 6f d2 c3 @.(...>...8Bo..  
00d0 6f 86 d8 77 fc 9f a6 40 e4 6d dc fe 82 0b 02 2a o..w...@..m....w  
00e0 63 6a 76 04 dc 97 95 95 b4 e8 ac 31 65 0a fb 55 cJv.....1e..U  
00f0 45 d6 6d 63 df 2a 7e 93 E.mc.\*~..

Payload is encrypted application data (ssl.app\_data), 177 bytes

Packets: 33 Displayed: 33 Marked: 0

# Crypto in practice

- HTTPS (the green padlock in your browser)
  - HTTP with SSL
  - Doesn't hide endpoints
- SSH (host keys, transport, pub/pri keys)
  - Doesn't hide endpoints
- Mail (STARTTLS, PGP/GPG)
  - PGP/GPG doesn't protect mail headers

# So...

- We're transmitting our data **securely**, but that doesn't mean our communication is necessarily **private**.
- Metadata is still being leaked:
  - Who we're talking to
  - When
  - What method (e.g. thing ports/protocol)

# Definitions in 30 seconds

- Security is the **practice of** defending information from unauthorized parties
  - Prevent use, tampering, duplication, destruction
- Privacy is the **ability to** seclude one's information from unauthorized parties

# Is this really of concern?

---

- The communication itself is protected.
- Is the metadata really that useful?
- Is it possible to record all that information?



# A different view

- Let's take a look at what other things we may inadvertently reveal:
  - Search/click habits (tied to a Google/Bing account)
  - Purchase habits (tied to a credit card, account)
  - Location habits (GPS, PRESTO card, etc.)
  - Etc.
- We are living in an age where any and all information is collected about us.

# Do we need to be concerned?

- It depends on who the bad guy is.
- In security/privacy circles, we have a notion of identifying **who/what is our adversary**.
- We then make certain security assurances about **what we can secure/hide against the defined adversary**.

# Recall how we're usually told to secure our systems:

---

- Don't go to the super sketchy websites
- Use antivirus
- Use firewall
- Don't reuse passwords
- Never put out personal information about yourself
- We're totally cool here, right guys?

# Consider who the adversary is

- Case 1: scriptkiddies and co.
  - Target: home machines/routers
  - Purpose: Pwn ur PC (for fun-and-profit)
  - Purpose: create botnets, zombie PCs, etc.
  - Method: various scripts/packages readily available (e.g. Metasploit)

# Consider who the adversary is

- Case 2: identity thieves
  - Target: accounts of specific users
  - Purpose: look for personal information for financial gain
  - Method: OSInt, specific backdoors, phishing

# Consider who the adversary is

- Case 3: government agencies
  - Target: whistleblowers (the physical person)
  - Purpose: prevent highly classified/sensitive information from being revealed
  - Method: <CLASSIFIED>

# Consider who the adversary is

- Case 4: corporations
  - Target: everyone
  - Purpose: improve services for all users; research
  - Method: marketing, lax policies, privacy guarantees
  - Method: scanning through consumed content

# Why does this all matter?

- We knowingly or unknowingly end up **providing a large amount of information** about ourselves
- We now have systems that are capable of both **storing** and **analyzing** this data
  - (This is the focus of information retrieval systems)
- We often **trust** major third parties to **do the right thing** in order to provide us with **useful services**



# Speaking of third parties...

## Former NSA Honcho Calls Corporate IT Security "Appalling"

Posted by [samzenpus](#) on Thursday October 03, 2013 @12:35AM  
from the [is-that-better-than-terrible?](#) dept.

Nerval's Lobster writes

"Former NSA technology boss Prescott Winter has a word for the kind of security he sees

# Snowden affair



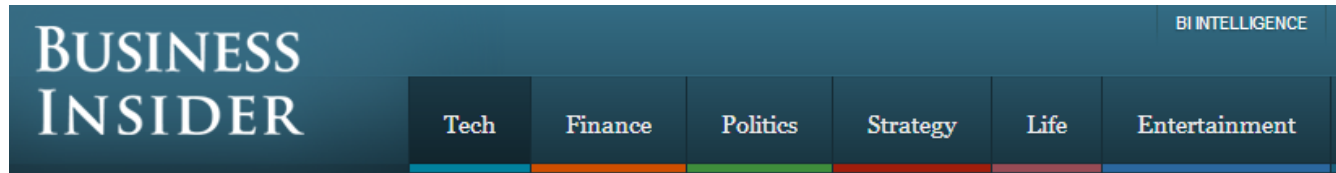
# Snowden affair

- Leaked a number of documents suggesting government surveillance programs in place:
  - PRISM
  - XKeyscore
  - Tempora
- Called the most significant leak in US history

# Companies response:

- Nope.
- There is no way that **Microsoft, Google, Facebook, Apple, etc.** would willingly provide the NSA with information.
- Policies exist to protect the user, right?

# Haha no.



[TECH](#)

More: [Yahoo](#) [Marissa Mayer](#) [NSA](#) [PRISM](#)

## Marissa Mayer: 'It's Treason' For Yahoo To Disobey The NSA

■ JULIE BORT | SEP. 11, 2013, 6:06 PM | 🔥 32,764 | 💬 105

[f Recommend](#) 1.7k [in Share](#) 27 [t Tweet](#) 646 [+1](#) 79 [EMAIL](#) [+ MORE](#)

Marissa Mayer was on stage on Wednesday at the TechCrunch Disrupt conference when Michael Arrington asked her about NSA snooping.



# Haha no.

He wanted to know what would happen if Yahoo just didn't cooperate. He wanted to know what would happen if she were to simply talk about what was happening, even though the government had forbidden it.

"Releasing classified information is treason. It generally lands you incarcerated," she said, clearly uncomfortable with the turn of the conversation.



Yahoo CEO Marissa Mayer

YHOO Oct 08 04:39PM

**32.93**

# Clearly we can trust crypto

PCWorld

Macworld

TechHive



**SECURITY** security, privacy

## Silent Circle ditches NIST cryptographic standards to thwart NSA spying

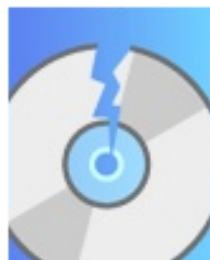
Lucian Constantin, IDG News Service

Oct 2, 2013 6:36 AM |

The U.S. National Security Agency's reported efforts to weaken encryption standards have prompted an encrypted communications company to move away from cryptographic algorithms sanctioned by the U.S. National Institute of Standards and Technology (NIST).

Silent Circle, a provider of encrypted mobile Voice over Internet Protocol (VoIP) and text messaging apps and services, will stop using the Advanced Encryption Standard (AES) cipher and Secure Hash Algorithm 2 (SHA-2) hash functions as default cryptographic algorithms in its products.

# ...or trust Linux



(Mis)Uses of  
Technology

by Glyn Moody

Thu, Sep 19th 2013  
1:44pm

## Linus Torvalds Admits He Was Approached By US Government To Insert Backdoor Into Linux -- Or Does He?

from the *who-can-you-trust?* dept

At the LinuxCon meeting in New Orleans, Linus Torvalds was asked if he had ever been approached by the US government to insert a backdoor into the Linux kernel. Here's his characteristic answer:

*Torvalds responded "no" while shaking his head "yes," as the audience broke into spontaneous laughter.*



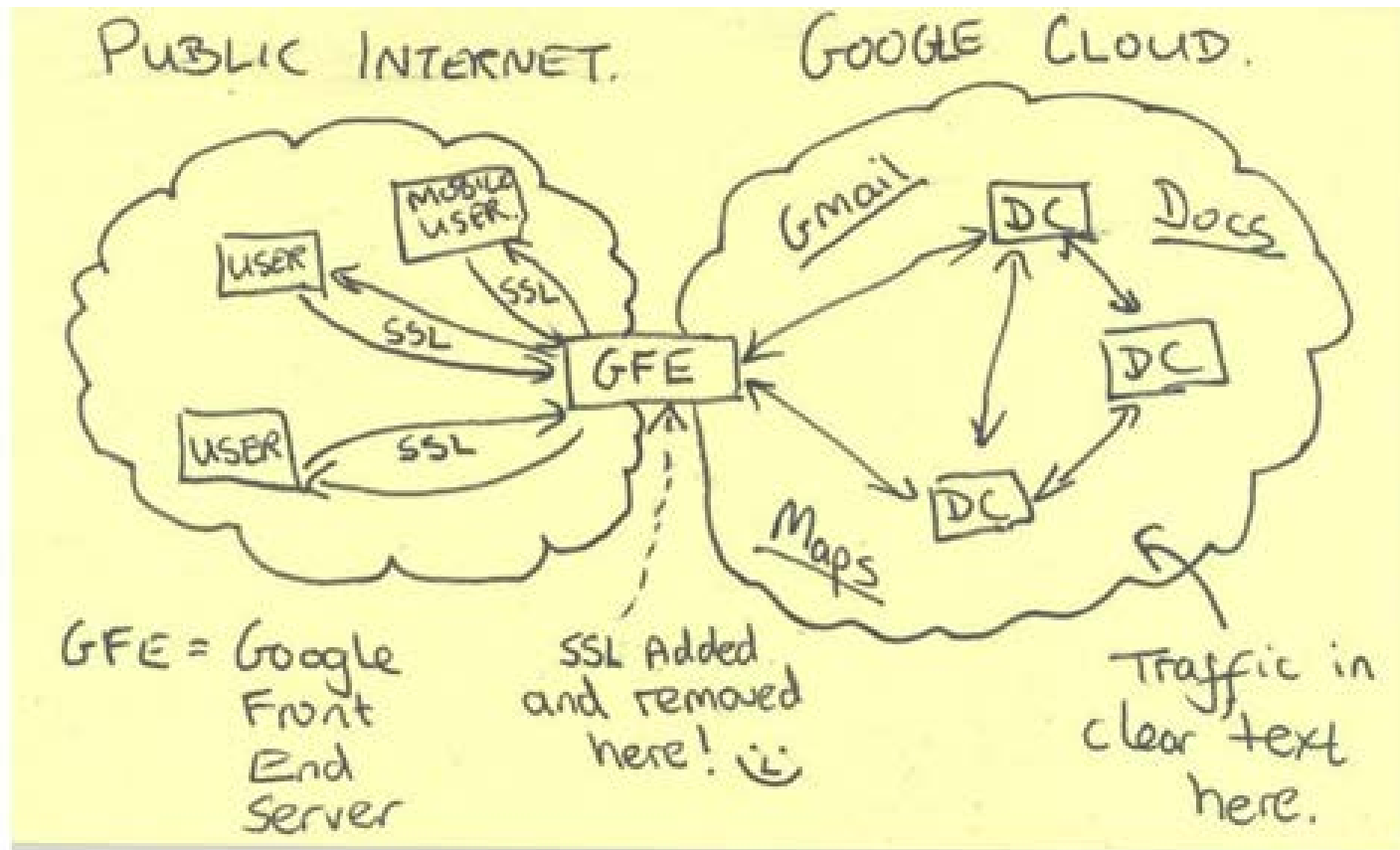
# So why do we care?

- We can't just worry about protecting explicit information
  - Lots of implicit information being leaked
- Our data is subject to... who's whims?
  - Hackers?
  - Corporations?
  - Gov't Agencies?
- We may not be threats to national security, but we should be aware that this is happening, and be guaranteed some level of privacy

# Two recent things to think about

- Adobe leak:
  - Big company = millions of users
  - Source code compromised
  - Passwords were encrypted, **not hashed**
- NSA v. The World:
  - German Chancellor Merkel's phone was **tapped**
  - NSA reveals to be monitoring the links between **users** and **corporate datacenters**

# SSL Added and Removed here!



# Questions? Open Discussion



<http://teespring.com/nsassl>