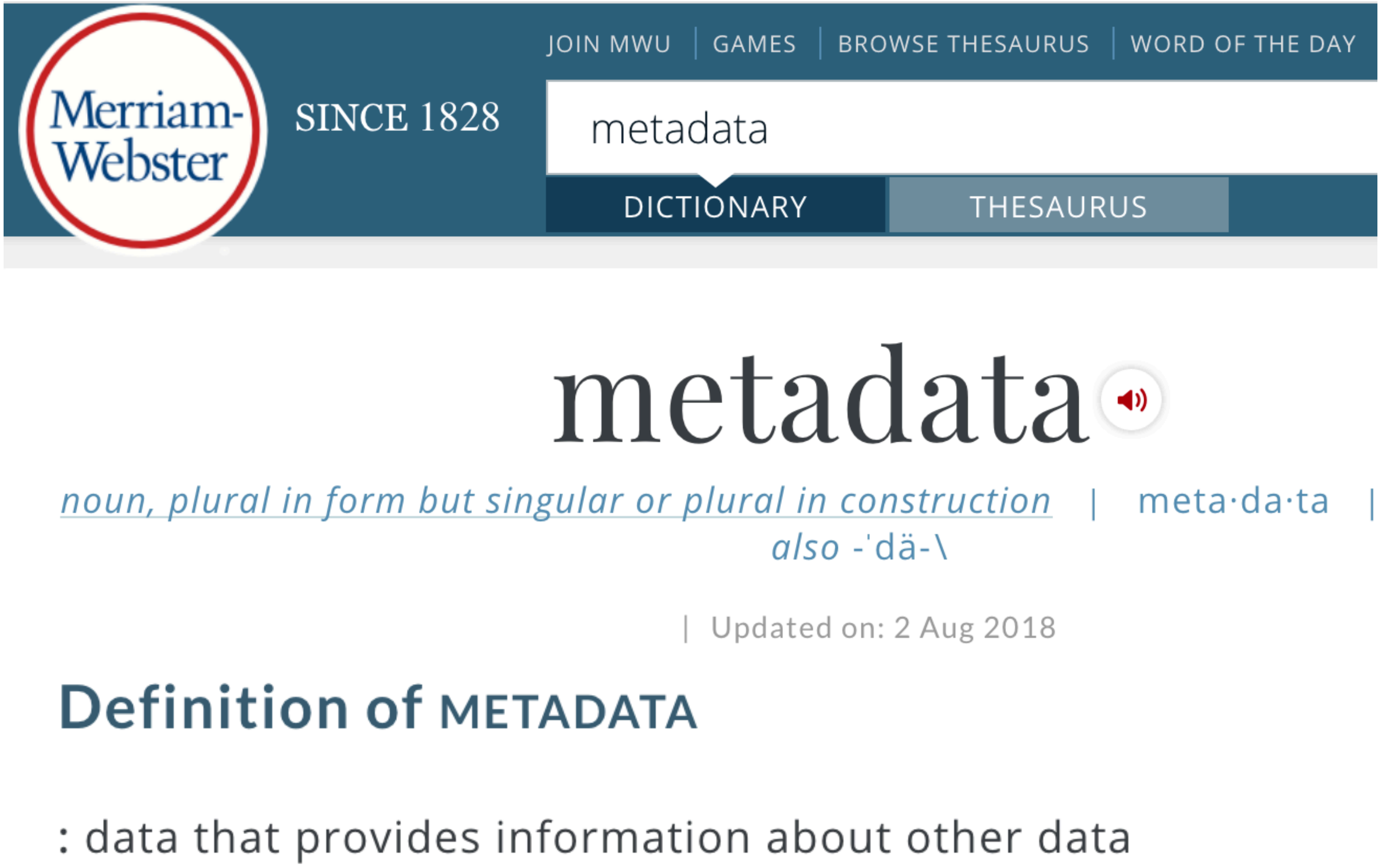
A Cryptographic Airbag for Metadata

Protecting Business Records Against Unlimited Search and Seizure

Charles V. Wright
Portland State University
@hackermath

Mayank Varia
Boston University
@mvaria



metadata

metadata

noun, plural in form but singular or plural in construction | meta·da·ta | *also* -ˈdä-\

| Updated on: 2 Aug 2018

Definition of METADATA

: data that provides information about other data

Communications Metadata

- **WHO** did you talk to?
- **WHEN** did you talk to them?
- **HOW LONG** did you talk?
- **HOW MUCH** did you send/receive?
- **WHERE** did you talk from/to?

Why Metadata?

"We kill people based on metadata."
-- Gen. Michael Hayden, former director, NSA and CIA

"Metadata absolutely tells you everything about somebody's life."

"If you have enough metadata you don't really need content…. [It's] sort of embarrassing how predictable we are as human beings."
-- Stewart Baker, former general counsel, NSA

Unlimited Search and Seizure?

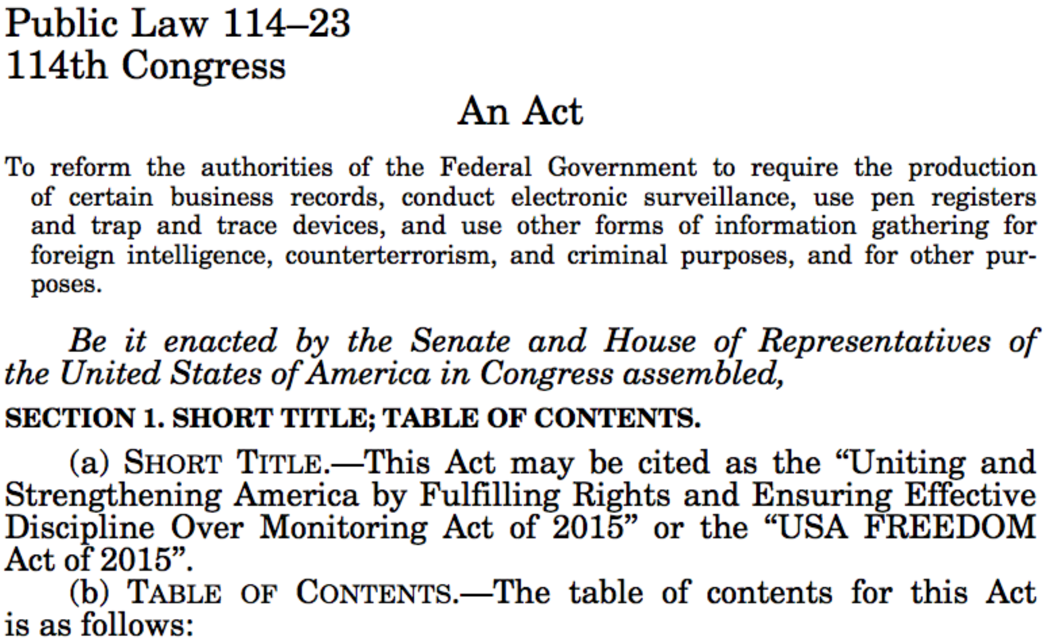
"Big Brother" is watching. And he is monitoring the phone calls and digital communications of every American, as well as of any foreigners who make or receive calls to or from the United States.

-- Rep. Jim Sensenbrenner, author of the USA PATRIOT Act
(*The Guardian*, June 2013)

https://www.theguardian.com/commentisfree/2013/jun/09/abuse-patriot-act-must-end

USA FREEDOM Act (2015)

- Providers collect metadata, not the government
- Government may still demand access with little oversight
  - Third party doctrine



Public Law 114–23
114th Congress

An Act

To reform the authorities of the Federal Government to require the production of certain business records, conduct electronic surveillance, use pen registers and trap and trace devices, and use other forms of information gathering for foreign intelligence, counterterrorism, and criminal purposes, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE; TABLE OF CONTENTS.

(a) SHORT TITLE.—This Act may be cited as the “Uniting and Strengthening America by Fulfilling Rights and Ensuring Effective Discipline Over Monitoring Act of 2015” or the “USA FREEDOM Act of 2015”.

(b) TABLE OF CONTENTS.—The table of contents for this Act is as follows:

Outline for the Talk

- Motivation
- Background
  - Crypto Crumple Zones
  - US Legal Situation & Uncertainty
- Proposed Constructions
- Discussion

Background: Crypto Crumple Zones

Crypto Crumple Zones

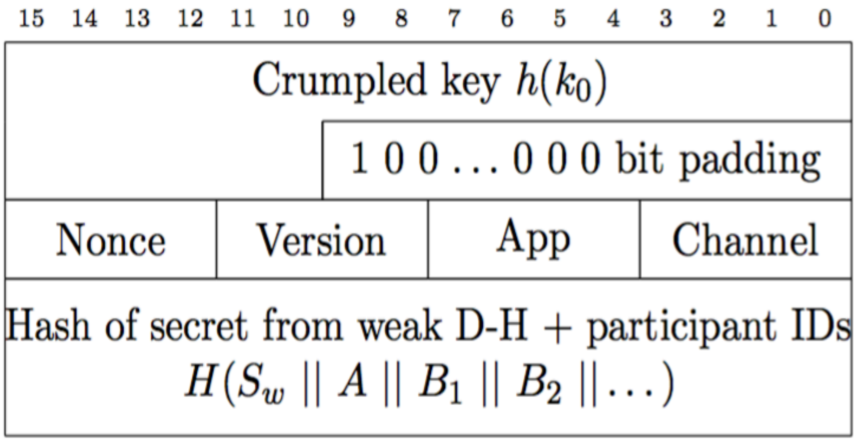
(Wright & Varia, EuroS&P 2018)

- **Problem:** Law enforcement demands access to encrypted message contents

- **Idea**: Make brute force key recovery **possible** but extremely **expensive.** Cost *C* to decrypt *n* messages:

$$C = I + M * n$$

> “2^{80} security is interesting.
> 2^{128} security is boring.”
>
> – Dan Bernstein

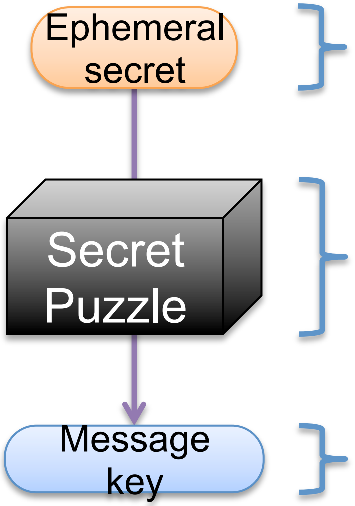


<table>
<tr><td>15</td><td>14</td><td>13</td><td>12</td><td>11</td><td>10</td><td>9</td><td>8</td><td>7</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td><td>1</td><td>0</td></tr>
<tr><td colspan="16">Crumpled key $h(k_0)$</td></tr>
<tr><td colspan="6"></td><td colspan="10">$1\ 0\ 0 \ldots 0\ 0\ 0$ bit padding</td></tr>
<tr><td colspan="4">Nonce</td><td colspan="4">Version</td><td colspan="4">App</td><td colspan="4">Channel</td></tr>
<tr><td colspan="16">Hash of secret from weak D-H + participant IDs
$H(S_w \mid\mid A \mid\mid B_1 \mid\mid B_2 \mid\mid \ldots)$</td></tr>
</table>

$$G(x) = \texttt{SHA256}(x) >> 128$$
$$h(x) = (\texttt{SHA256}(x) >> (256 - \ell)) << (256 - \ell)$$
$$k_1 = \texttt{SHA256}(\texttt{SHA256}(\text{header}))$$

Symmetric crumpling

- Create a puzzle for each ciphertext
- Solution unlocks the message key
- Government searches for solution

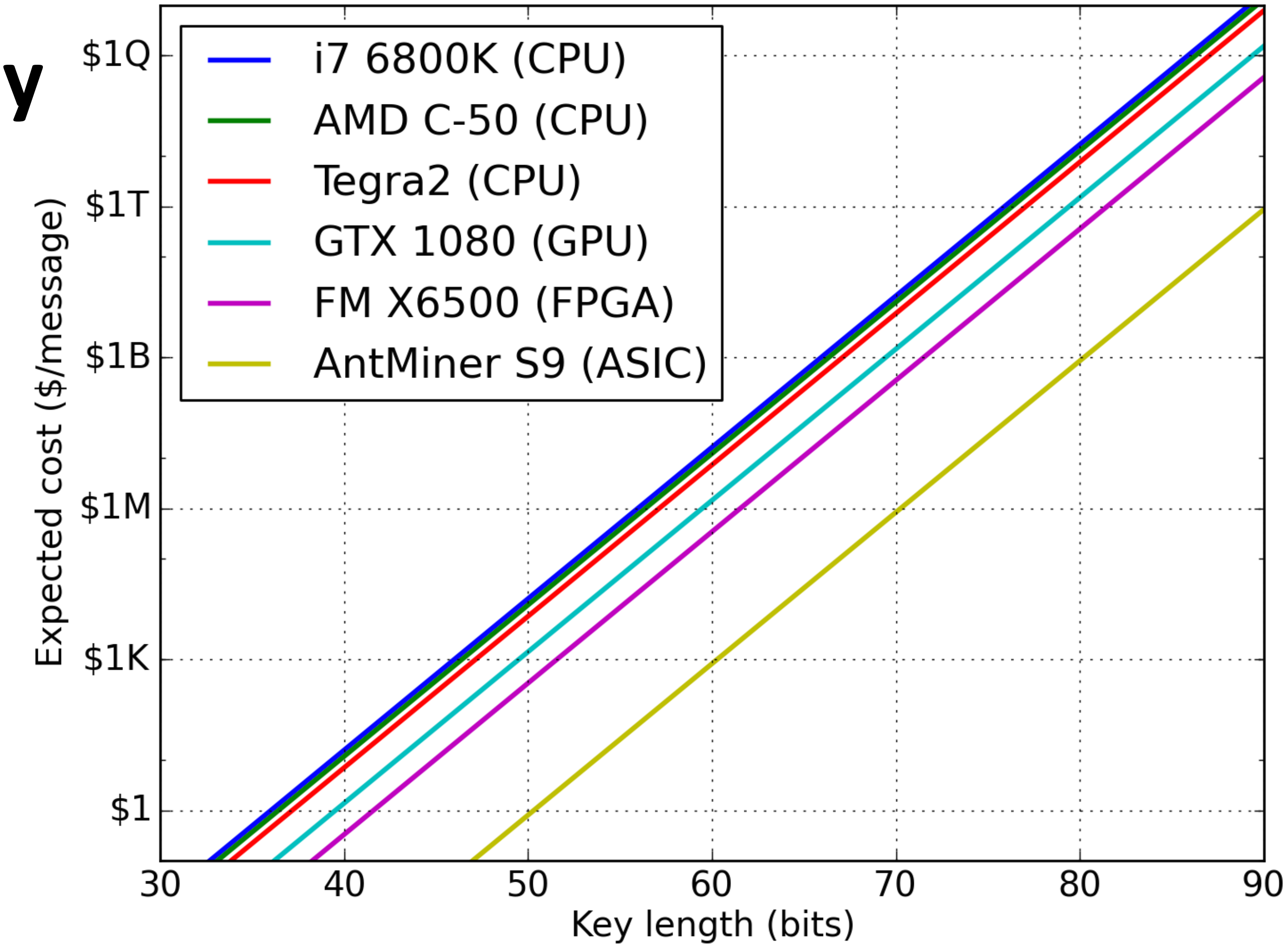


small search space think: 60-80 bits from original key k0

double SHA-256 (from Bitcoin)

large search space think: 128 or 256 bits

Electricity Cost to Decrypt ($$$)

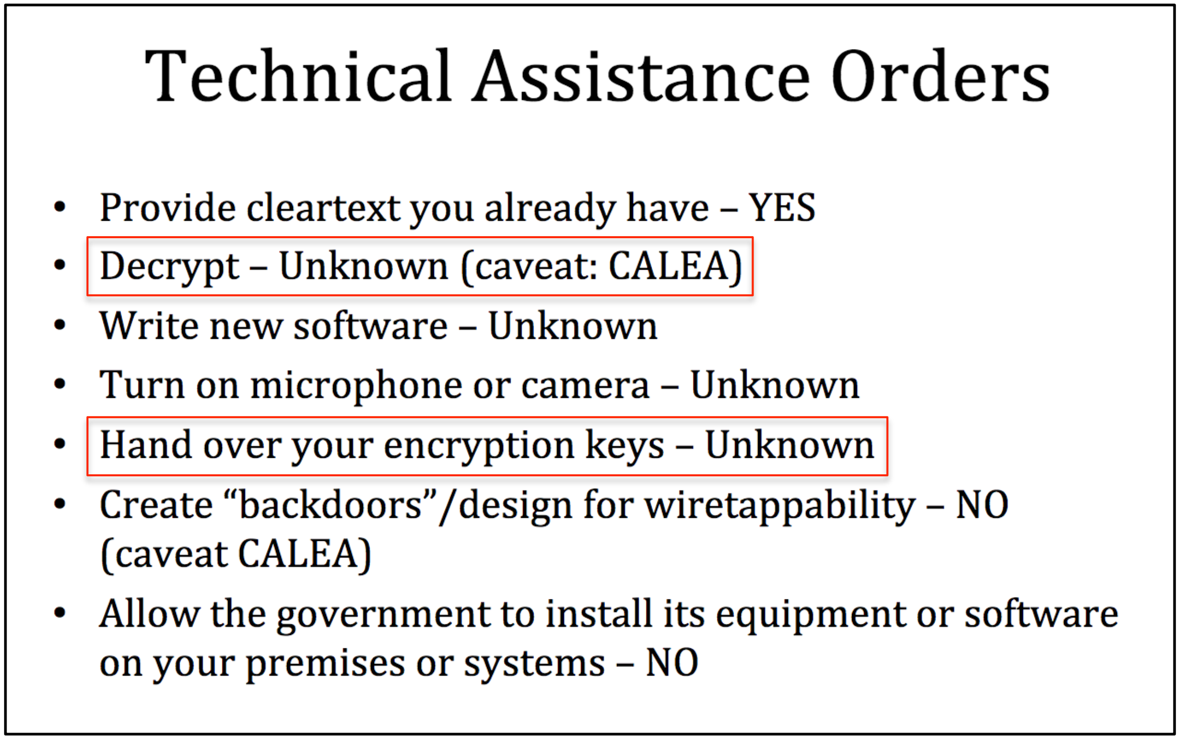


Outline for the Talk

- Motivation
- Background
  - Crypto Crumple Zones
  - US Legal Situation & Uncertainty
- Proposed Constructions
- Discussion

What Does US Law Require?

Granick & Pfefferkorn, *When the Cops Come A-Knocking: Handling Technical Assistance Demands from Law Enforcement*, Black Hat 2016



Technical Assistance Orders

- Provide cleartext you already have – YES
- Decrypt – Unknown (caveat: CALEA)
- Write new software – Unknown
- Turn on microphone or camera – Unknown
- Hand over your encryption keys – Unknown
- Create "backdoors"/design for wiretappability – NO (caveat CALEA)
- Allow the government to install its equipment or software on your premises or systems – NO

A Cryptographic Airbag for Metadata

Idea: Encrypt Metadata with Crumpled Keys!

#	FROM	TO	START	LENGTH	
1	410-555-1234	410-555-6789	8/14/18 09:00	37 min	
2	410-555-6789	410-555-0001	8/14/18 10:35	3 min	
3	410-555-1234	503-555-0002	8/14/18 11:19	5 min	

Idea: Encrypt Metadata with Crumpled Keys!

#	FROM	TO	START	LENGTH	NONCE
1	[redacted]	[redacted]	[redacted]	[redacted]	8asd8765sd
2	[redacted]	[redacted]	[redacted]	[redacted]	874s3d3s4f
3	[redacted]	[redacted]	[redacted]	[redacted]	5422024fbc

Encrypt each record with
its own unique (crumpled) key

Need a unique nonce
to generate each key

Idea: Encrypt Metadata with Crumpled Keys!

#	FROM	TO	START	LENGTH	NONCE
1	[redacted]	[redacted]	[redacted]	[redacted]	8asd8765sd
2	410-555-6789	410-555-0001	8/14/18 10:35	3 min	874s3d3s4f
3	[redacted]	[redacted]	[redacted]	[redacted]	5422024fbc

Authorities can recover selected messages via brute-force search for the key. (Moderately expensive)

Better Idea: Add an Encrypted Index

- Inverted index gives the lists of records that contain each keyword

KEYWORD	LIST OF RECORDS
410-555-1234	1, 3, 5, 6, 7, 13
410-555-6789	1, 2, 6, 8, 9, 10, 13
410-555-0001	2, 4, 11
503-555-0002	3

Better Idea: Add an Encrypted Index

- Encrypt each list with crumpled keys
- Authorities must spend $$$ to recover the list

KEYWORD	LIST OF RECORDS
410-555-1234	[redacted]
410-555-6789	[redacted]
410-555-0001	[redacted]
503-555-0002	[redacted]

Better Idea: Add an Encrypted Index

- Encrypt each list with crumpled keys
- Authorities must spend $$$ to recover the list

KEYWORD	LIST OF RECORDS
410-555-1234	[redacted]
410-555-6789	1, 2, 6, 8, 9, 10, 13
410-555-0001	[redacted]
503-555-0002	[redacted]

Costs – For the Provider

Storage Space

- For Netflow records, the storage space is **doubled**

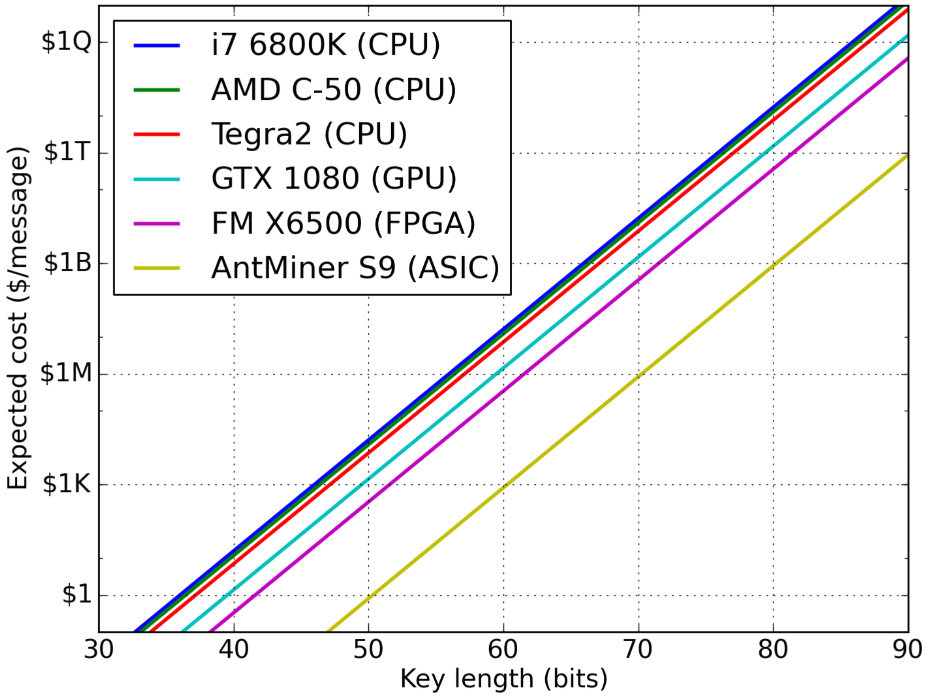
Computation

- Minimal increase on platforms with HW support for AES and SHA-2

Costs – For the Authorities

Hypothetical Telephone Example

- **Assumptions:**
  - 1 key per call detail record
  - 53 bit keys ($8 to break)
  - 100 calls per person-month
- **Costs:**
  - To monitor 1 suspect: **$10k/yr**
  - To monitor all 300 million Americans: **$2.88 trillion / yr**



Postscript: Carpenter v. United States

- June 2018 – The US Supreme court ruled that access to **cell site location data** requires a search warrant under the 4th Amendment

- Implications for other data are **still unknown**

- Do we **need the airbag** after all?

Discussion