2021 MTU Women in Computer Science (WiCS) Summer Program

Security Track

An Introduction of Cybersecurity and Flash Memory Security Research

Bo Chen

Assistant Professor, Department of Computer Science

bchen@mtu.edu

https://cs.mtu.edu/~bchen

https://snp.cs.mtu.edu

About Me



Bo Chen

Assistant Professor, Computer Science

bchen@mtu.edu

906-487-3149

Rekhi 301

Links of Interest

Faculty Website

MTU Security and Privacy (SnP) lab

Areas of Expertise

- Mobile Device Security
- Cloud Computing Security
- Named Data Networking Security
- Big Data Security
- Blockchain

https://cs.mtu.edu/~bchen

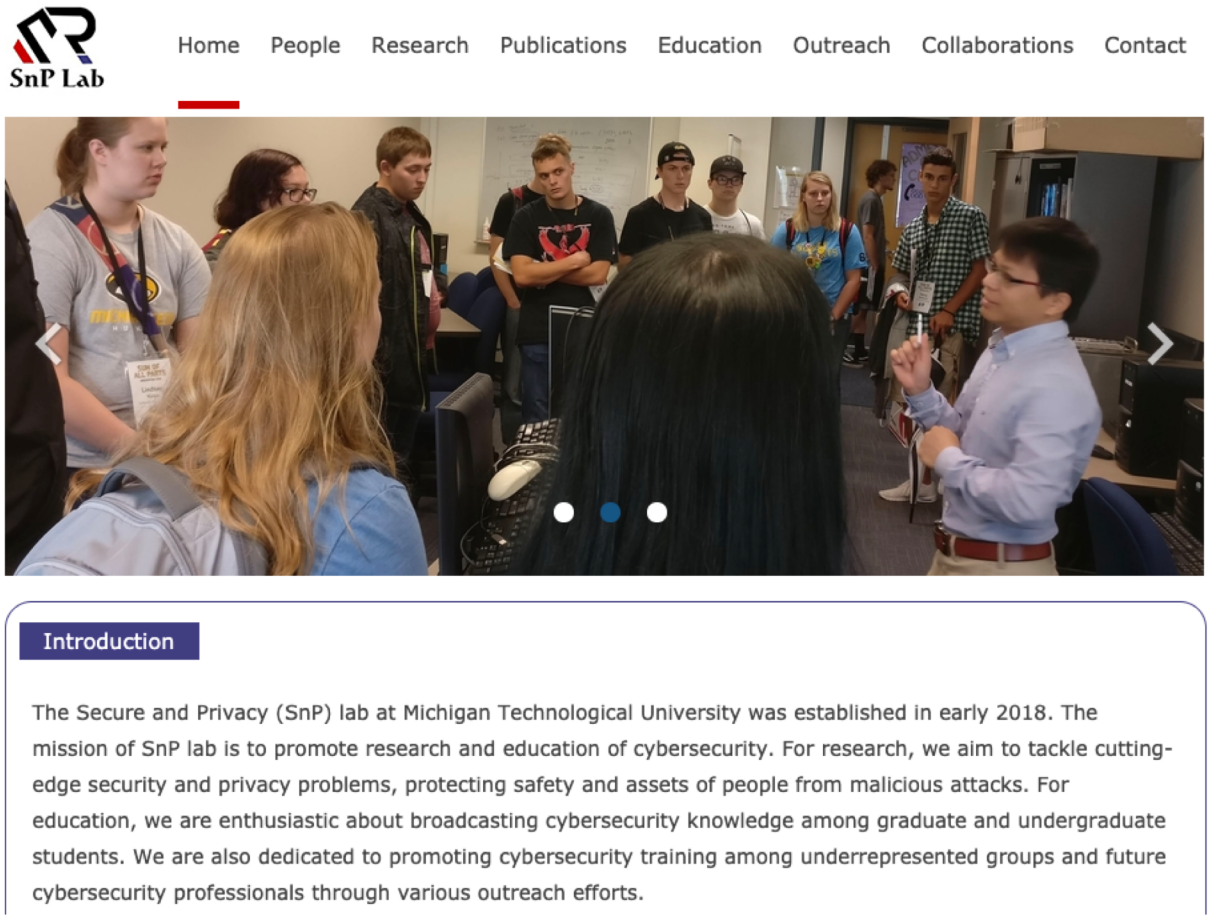
Lab director: MTU Security and Privacy (SnP) Lab

SnP Lab

Co-advisor: MTU CS cybersecurity reading group, MTU RedTeam

Faculty coach: MTU NCL cyber competition team (our team ranked 3rd out of 920+ teams across US in Spring 2021)

About MTU Security and Privacy (SnP) Lab



https://snp.cs.mtu.edu

PhD Students

Niusen Chen

Wen Xie

Weijing You (UCAS)

Master Students

Nagaharshitha Chinta (co-advised

Shashank Reddy Danda

Deepthi Tankasala

Sai Akhil Vankamamidi

Undergraduate Students

Dominika Bobik

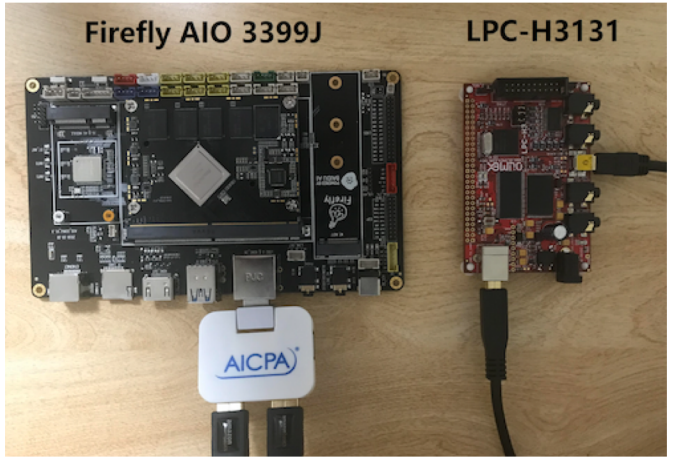
Sam Breuer

Thomas Grifka

Gary Watson

About MTU Security and Privacy (SnP) Lab

- Projects are currently funded by national science foundation, national security agency, etc.
  - Protecting sensitive data in mobile devices, IoT devices
  - Protecting critical data/ infrastructures outsourced to public clouds
  - Blockchain and information centric networking
  - Leveraging mobile devices for COVID-19 mitigation (recently)







A Starting Video

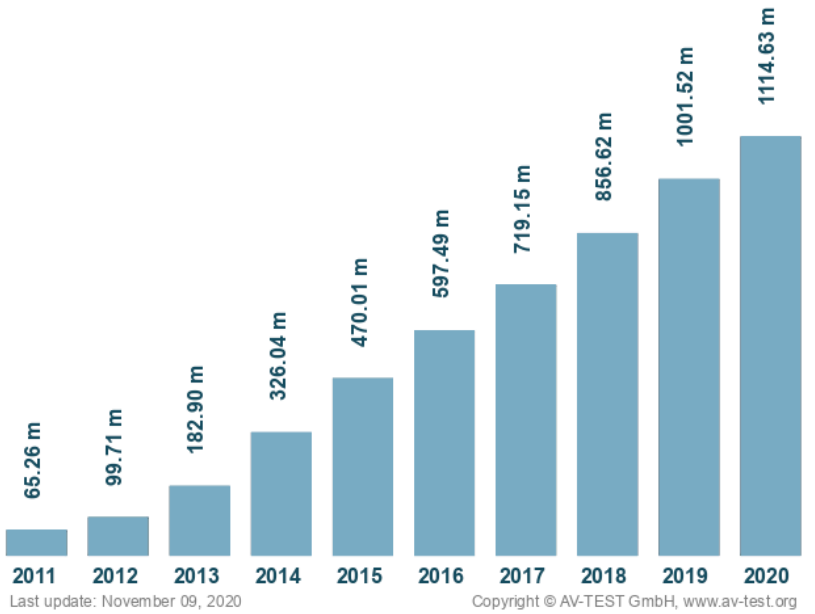
- https://www.youtube.com/embed/ThBpRBpyxLI?start=0&end=50&version=3

All Starts from Malware and Hacks



- On November 2, 1988, a graduate student at Cornell University, Robert Morris, unleashed what became known as the Morris worm
  - Morris worm disrupted a large number of computers then on the Internet, guessed at the time to be 10% of all those connected
- Malware and Hacks are here and there today





How to Combat Malware and Hacks?

- The answer is cybersecurity

- Ensure our systems and networks are well protected, such that any intruders can be detected, identified, and blocked
  - Make sure the software (code) we build is free of vulnerabilities
    - The attackers cannot exploit the vulnerabilities to intrude into our systems and networks

  - Make sure our data are protected
    - Not disclosed to unauthorized parties
    - Not modified by unauthorized parties
    - Always available for use
    - Always recoverable

Outline

- Recent cybersecurity instances
- What is cybersecurity
- Why learning cybersecurity
- Mobile devices and flash memory
- Flash memory security research

Recent Cybersecurity Instances

Hacking Instances Just Happen



https://www.nytimes.com › U.S. › Politics

Cyberattack Forces a Shutdown of a Top U.S. Pipeline - The ...

May 13, 2021 — The operator, **Colonial Pipeline**, said it had halted systems for its 5500 miles of pipeline after being hit by a ransomware attack.





It was reported that within hours after the attack the company paid a ransom of nearly 75 Bitcoins ($5 million) to the hackers in exchange for a decryption tool which proved so slow that Colonial's own backups were used to bring the system back online

Hacking Instances in 2020

HEALTHCARE & PHARMA MAY 8, 2020 / 1:20 PM / UPDATED 6 MONTHS AGO

Exclusive: Iran-linked hackers recently targeted coronavirus drugmaker Gilead - sources



REUTERS World Business Markets Breakingviews Video More

U.S. LEGAL NEWS MARCH 23, 2020 / 3:08 PM / UPDATED 7 MONTHS AGO

Exclusive: Elite hackers target WHO as coronavirus cyberattacks spike

Early 2020



EMAIL SECURITY, SECURITY NEWS

Texas school district phished for $2.3 million

By **Doug Olenick** January 13, 2020

The Manor Independent School District fell victim to an apparent phishing scam to the



DATA BREACH, EMAIL SECURITY, HEALTH CARE, SECURITY NEWS

Breach of email accounts impacts 50,000 patients of Minnesota hospital

By **Bradley Barth** January 7, 2020

Data Breach Instances in 2019

Entity	Year	Records	Organization type	Method
Adobe Inc.	2019	7.5 million	tech	poor security
Amazon Japan G.K.	2019	unknown	web	accidentally published
2019 Bulgarian revenue agency hack	2019	over 5,000,000	government	hacked
Canva	2019	140,000,000	web	hacked
Capital One	2019	106,000,000	financial	unsecured S3 bucket
Desjardins	2019	2,900,000	financial	inside job
DoorDash	2019	4,900,000	web	hacked
Facebook	2019	540,000,000	social network	poor security
Facebook	2019	1,500,000	social network	accidentally uploaded
Facebook	2019	267,000,000	social network	poor security
First American Corporation	2019	885,000,000	financial service company	poor security
Health Sciences Authority (Singapore)	2019	808,000	healthcare	poor security
Justdial	2019	100,000,000	local search	unprotected api
LifeLabs	2019	15,000,000	healthcare	hacked
Ministry of Health (Singapore)	2019	14,200	healthcare	poor security/inside job
Mobile TeleSystems (MTS)	2019	100,000,000	telecommunications	misconfiguration/poor security
Quest Diagnostics	2019	11,900,000	Clinical Laboratory	poor security
StockX	2019	6,800,000	retail	hacked

2018 Data Breaches and Hacks

Facebook - Cambridge Analytica data scandal: 87 million user profiles were disclosed



- Various political organizations used information from Cambridge Analytica to attempt to influence public opinion:
  - 2015 and 2016 campaigns of United States politicians Donald Trump and Ted Cruz
  - 2016 Brexit (British exit from the European Union) vote
  - 2018 Mexican general election, 2018 for Institutional Revolutionary Party
- Successors: a company run by former officials at Cambridge Analytica, Data Propria, has been quietly working for President Donald Trump's 2020 re-election effort

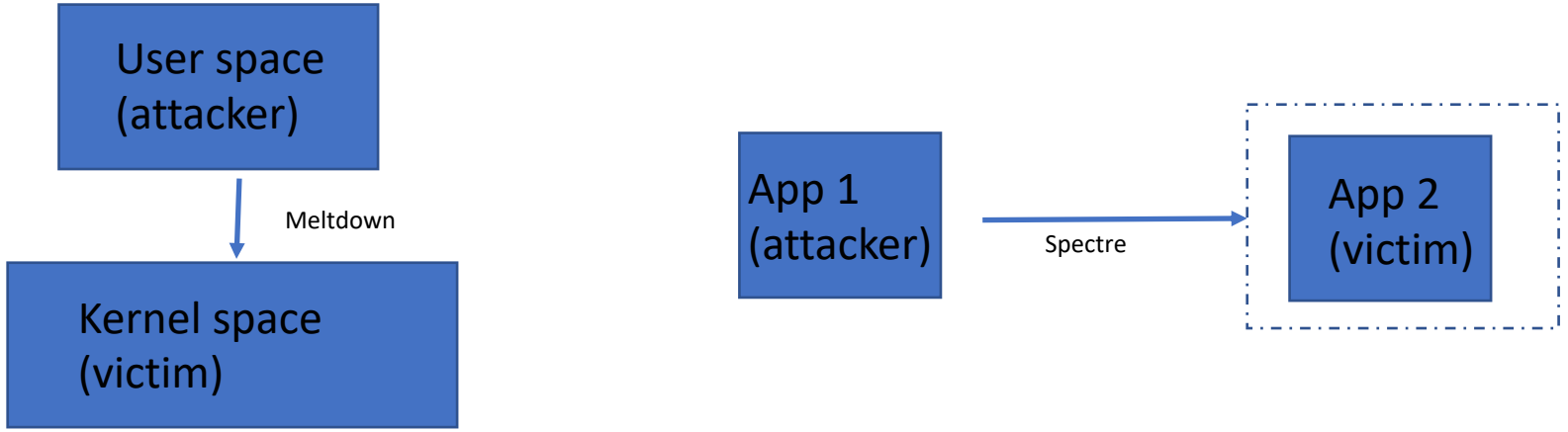


2018 Data Breaches and Hacks (cont.)

Under Armour: a data breach of 150 million accounts, with compromised data consisting of user names, the users' e-mail addresses and hashed passwords



Intel x86 microprocessors hardware vulnerabilities Meltdown and Spectre





More Data Breaches in 2018

- Saks Fifth Avenue / Lord & Taylor, 5 million credit card holders compromised
- British Airways, a data theft of about 380,000 customer records
- US Centres for Medicare & Medicaid Services (CMS), a data breach that exposed files of 75,000 individuals
- SingHealth, 1.5 million personal data compromised
- Quora reported a data breach that affected its 100 million users data
- …

More in Years Prior to 2018 …

- https://en.wikipedia.org/wiki/List_of_data_breaches

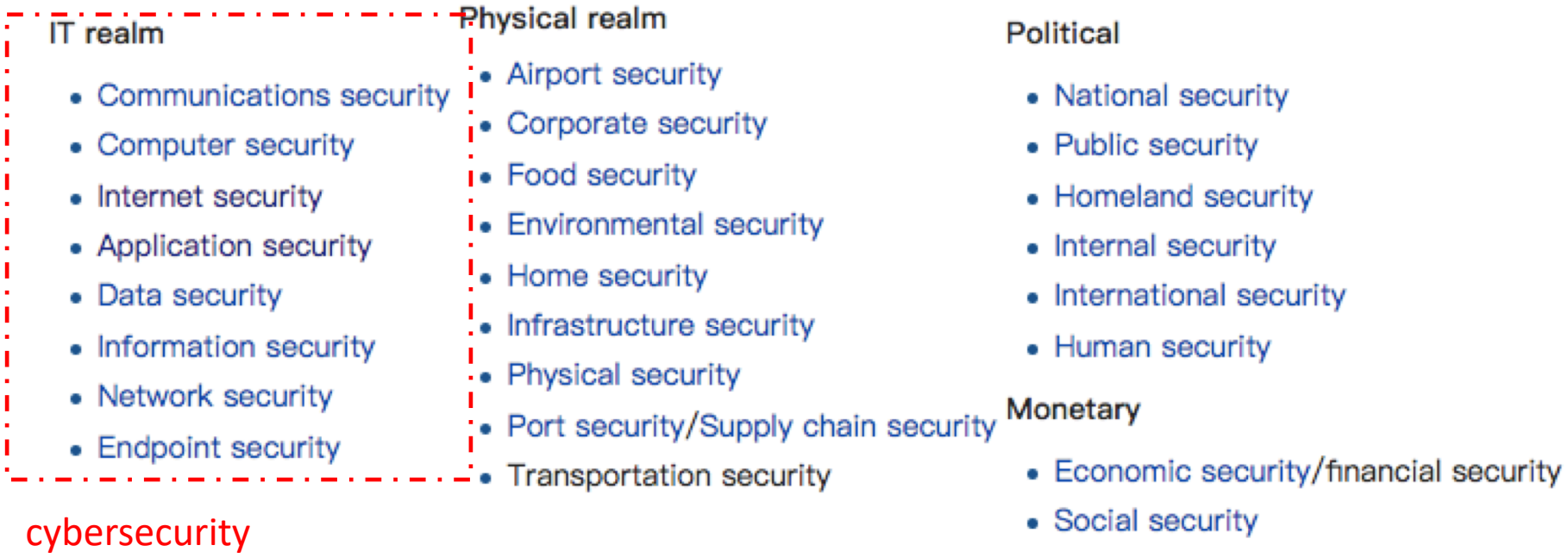
Cox Communications	2016	40,000	telecoms	hacked	[90]
Democratic National Committee	2016	19,252	political		[98]
US Department of Homeland Security	2016	30,000	government	poor security	[99][100]
EyeWire	2016	unknown	tech	lost / stolen computer	[130]
Friend Finder Networks	2016	412,214,295	web	poor security / hacked	[144][145]
Funimation	2016	2,500,000	web	hacked	[146][147]
Gyft	2016	unknown	web	hacked	[164][165]
Inuvik hospital	2016	6,700	healthcare	inside job	[188]
KM.RU	2016	1,500,000	web	hacked	[196]
Nival Networks	2016	1,500,000	gaming	hacked	[241]
Ofcom	2016	unknown	telecom	inside job	[244]
Rosen Hotels	2016	unknown	hotel	hacked	[261]
Taobao	2016	20,000,000	retail	hacked	[293]
TaxSlayer.com	2016	8,800	web	hacked	[297][298][299]
University of California, Berkeley	2016	80,000	academic	hacked	[330]
University of Central Florida	2016	63,000	academic	hacked	[332]
Verizon Communications	2016	1,500,000	telecoms	hacked	[344]
Weebly	2016	43,430,316	web	hacked	[354][355]
Bell Canada	2017	1,900,000	telecoms	poor security	[46]
Defense Integrated Data Center (South Korea)	2017	235 GB	military	hacked	[95]
Deloitte	2017	350 clients emails	consulting, accounting	poor security	[96][97]
Equifax	2017	163,119,000	financial, credit reporting	poor security	[120][121]
Grozio Chirurgija	2017	25,000	healthcare	hacked	[159][160][161]
Heathrow Airport	2017	2.5GB	transport	lost / stolen media	[175][176][177]
Taringa!	2017	28,722,877	web	hacked	[294]
Uber	2017	57,000,000	transport	hacked	[320]

What is Cybersecurity?

Security 101

The Definition of Security

- Security: freedom from, or resilience against, potential harm (or other unwanted coercive change) from external forces (*wikipedia*) – **in physical space**
- Cybersecurity: the protection of computer systems from theft or damage to their hardware, software or electronic data, as well as from disruption or misdirection of the services they provide – **in cyber space**



Cybersecurity Objectives: CIA

Main security objectives:

- **C**onfidentiality: unauthorized users cannot read information
- **I**ntegrity: unauthorized users cannot alter information
- **A**vailability: the information must be available when it is needed

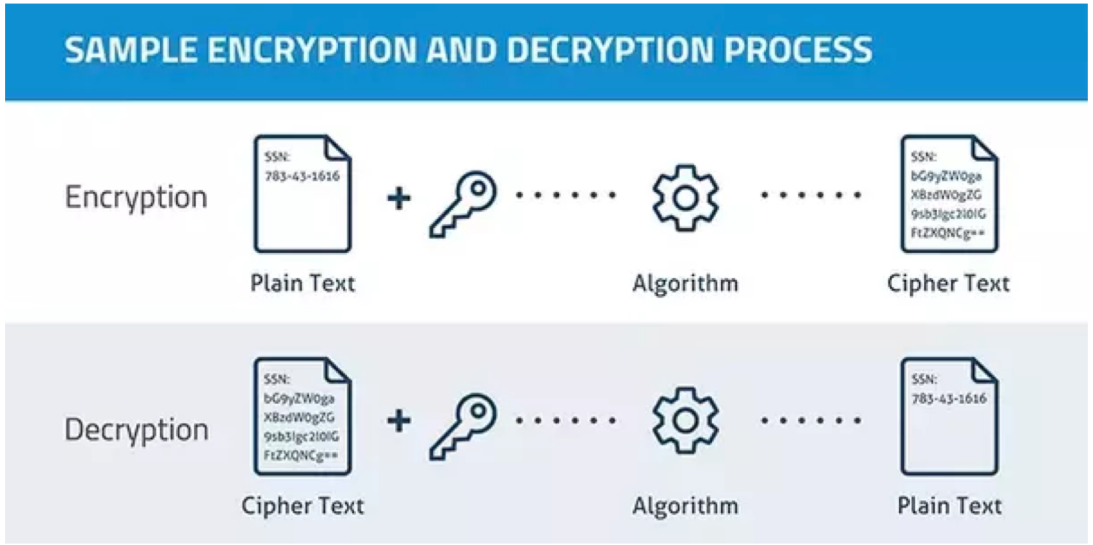
Other security objectives:

- Authentication and identification
- Access control
- Anonymity
- Non-repudiation: users cannot deny actions
- Privacy
- …

Confidentiality

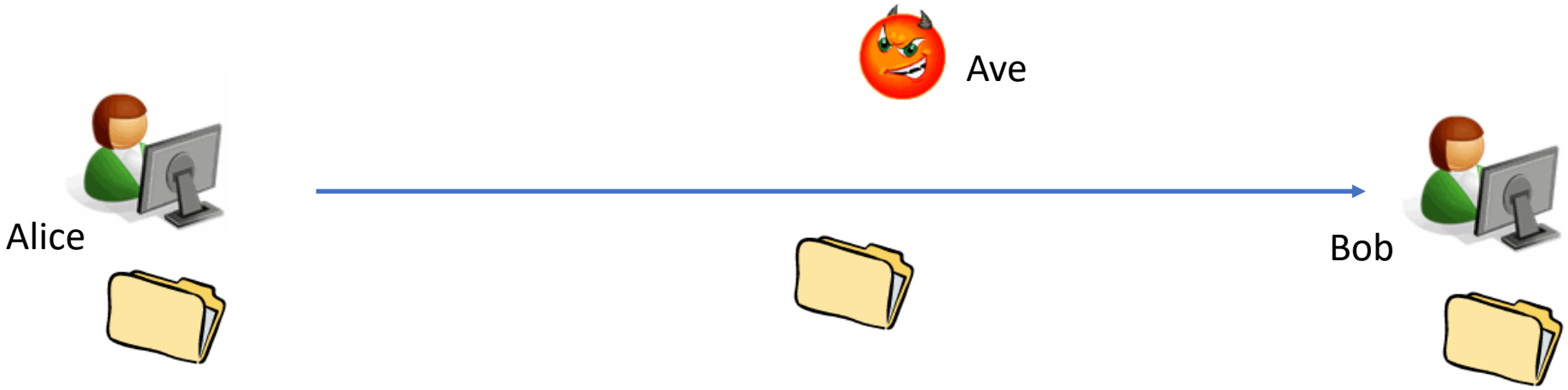
- *The concealment of information or resources*
  - Information is not made available or disclosed to unauthorized individuals, entities, or processes
  - E.g., your bank accounts, private photos, etc

- How to achieve confidentiality? Encrypt the data using a secret key, and only the authorized entities can obtain the secret key to decrypt the data
  - Symmetric encryption: AES, DES, 3DES
  - Asymmetric encryption: RSA



Integrity

- Maintaining and assuring the accuracy and completeness of data over its entire lifecycle
  - Data cannot be modified in an unauthorized or undetected manner
  - E.g., your emails, your electronic homework



Do to Ensure Integrity?

- Generate digest and perform integrity checking



Availability

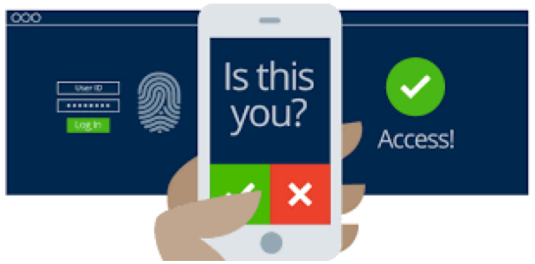
- For any information system to serve its purpose, the service/ information must be available when it is needed
  - This means the computing systems used to store and process the information, the security controls used to protect it, and the communication channels used to access it must be functioning correctly

- High availability systems aim to remain available at all times
  - Preventing service disruptions due to power outages, hardware failures, and system upgrades
  - Preventing denial-of-service attacks, such as a flood of incoming messages to the target system, essentially forcing it to shut down

Authentication and Identification

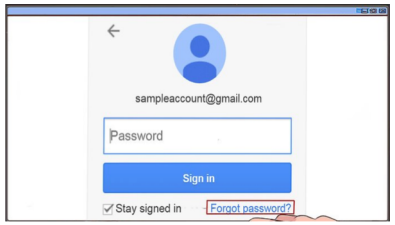
- Authentication in physical world: are you really who you claim?
  - Confirm the identity of a person by validating his/her identity document (e.g., driver license, passport, student ID card)

- Authentication in computers:
  - Confirm whether a person is the owner of a smartphone
  - Confirm whether a person is a user of online banking
  - Confirm whether a website is authentic



How to Do Authentication?

- Four general means for authenticating user's identity
  - Something the individual knows
    - Passwords



  - Something the individual possesses, a *token*
    - Memory card, smart card

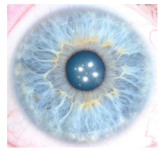


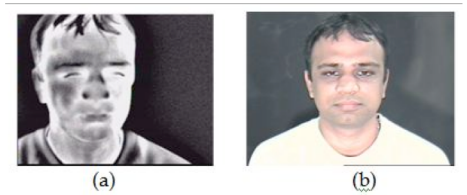


  - Something the individual is
    - Fingerprint, iris, retina, face



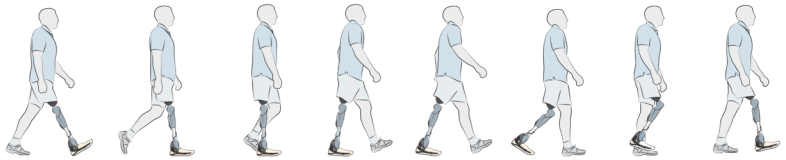






  - Something the individual does (behavior pattern)
    - Typing rhythm, gait, and voice







How to Do Authentication (cont.)?

- Multi-factor authentication (MFA) – used in our own IT systems in MTU



Access Control

- Access control in physical world: the selective restriction of access to a place. It is a matter of who, where, and when.
  - Historically, this was partially accomplished through mechanical keys and locks





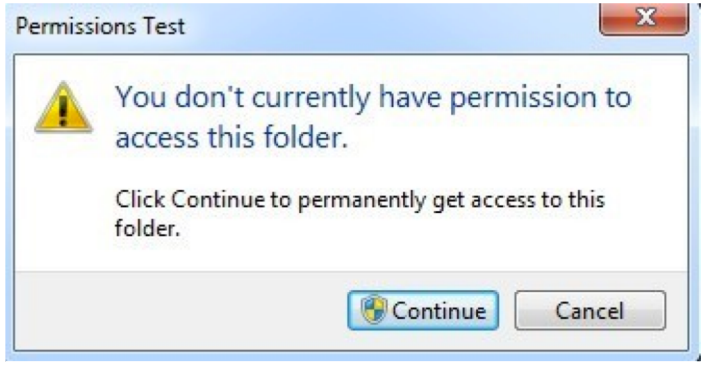


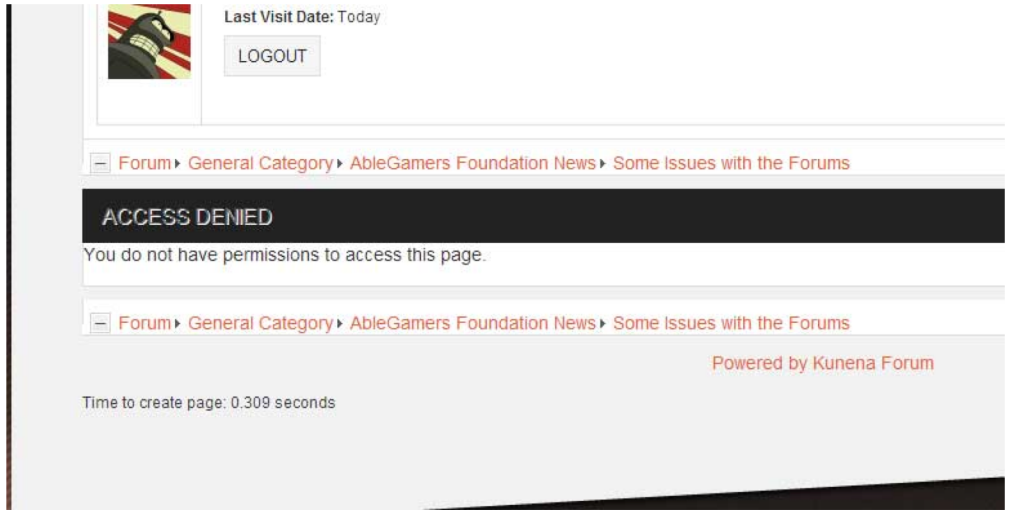
- Access control in computers: the selective restriction of access to computing resources (who, what, and how)
  - Who: users, programs, processes, etc.
  - What: computing resources like files, memory, I/O ports, etc.
  - How: how the computing resources can be “touched”

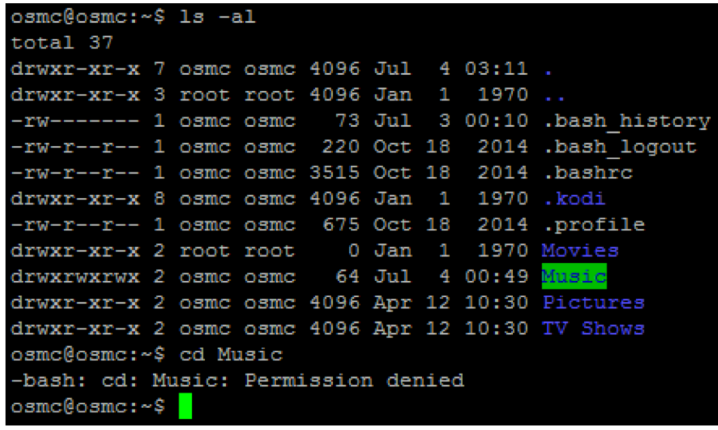
How to Do Access Control?

- Encrypting the protected computing resources using secret keys, and only disclose keys to those who are authorized

- The access control is enforced by systems (operating systems, database management systems, etc.) following permissions







Why Learning Cybersecurity?

Great Job Market

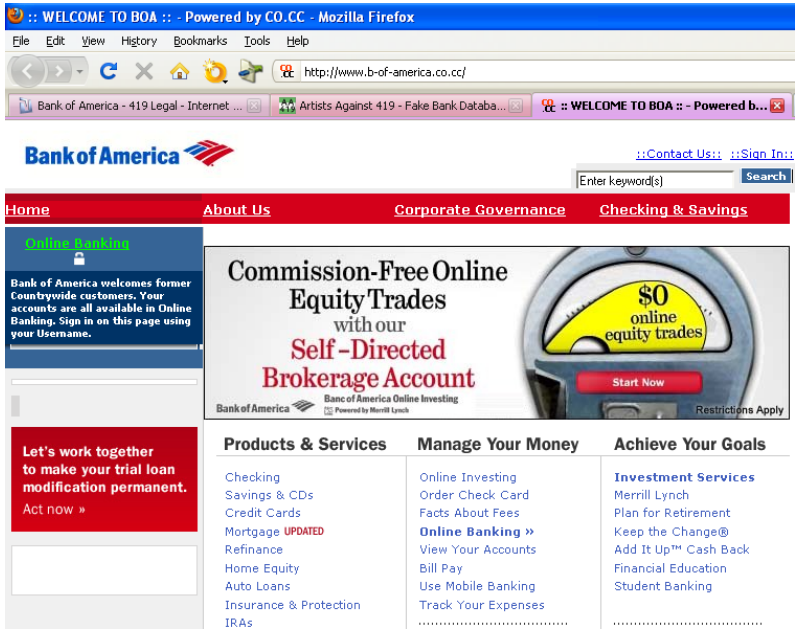
- There will be 3.5 million unfilled cybersecurity positions by 2021
  - According to Cybersecurity Jobs Report, sponsored by Herjavec Group
- The rate of growth for jobs in information security is projected at 37% from 2012 to 2022
  - According to the Bureau of Labor Statistics
  - Much faster than the average for all other occupations



Protect Your Own Asset

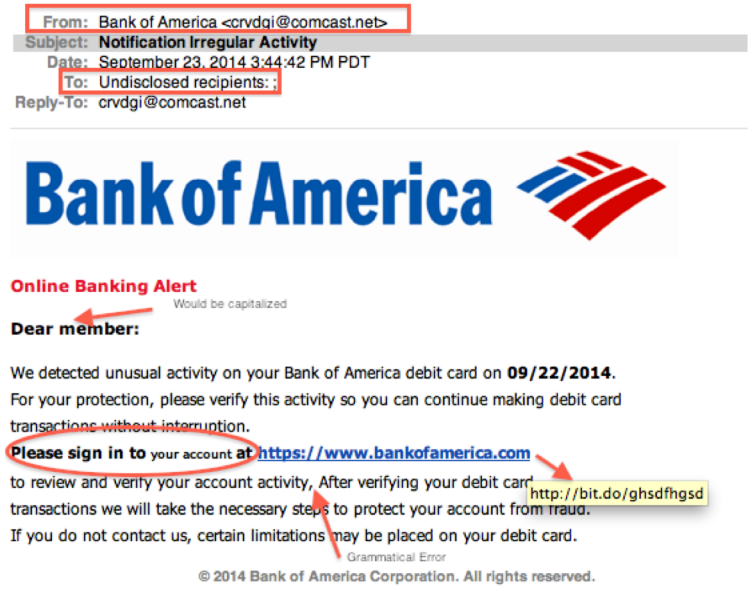
- Reduce the possibility of exposure to potential hacks
  - Malicious code is here and there (malicious java scripts, applets, etc.)
  - Make sure you trust the web sites before you go there (a lot of phishing website)
    - www.google.com is fine, but www.go0gle.com may not
    - Do you want to click the link www.facebook.net, or www.b-of-America.co.cc

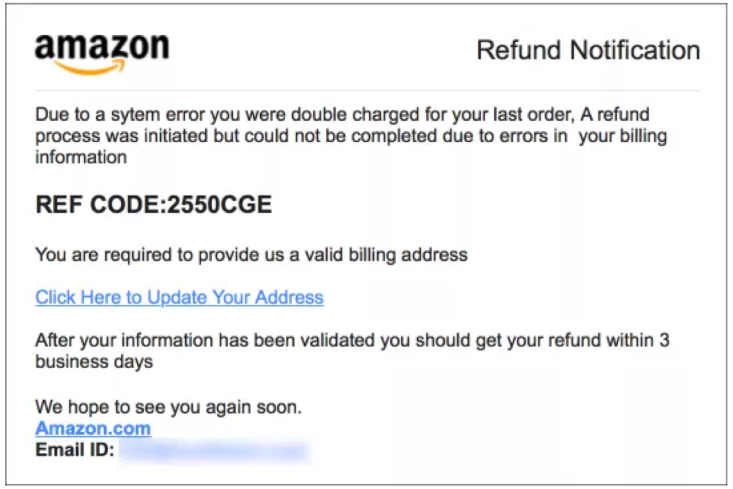


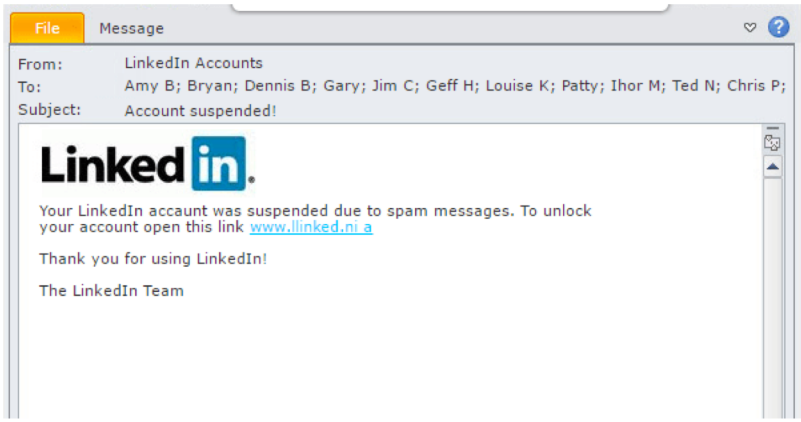


Protect Your Own Asset (cont.)

- Reduce the possibility of exposure to potential hacks
  - A lot of phishing emails

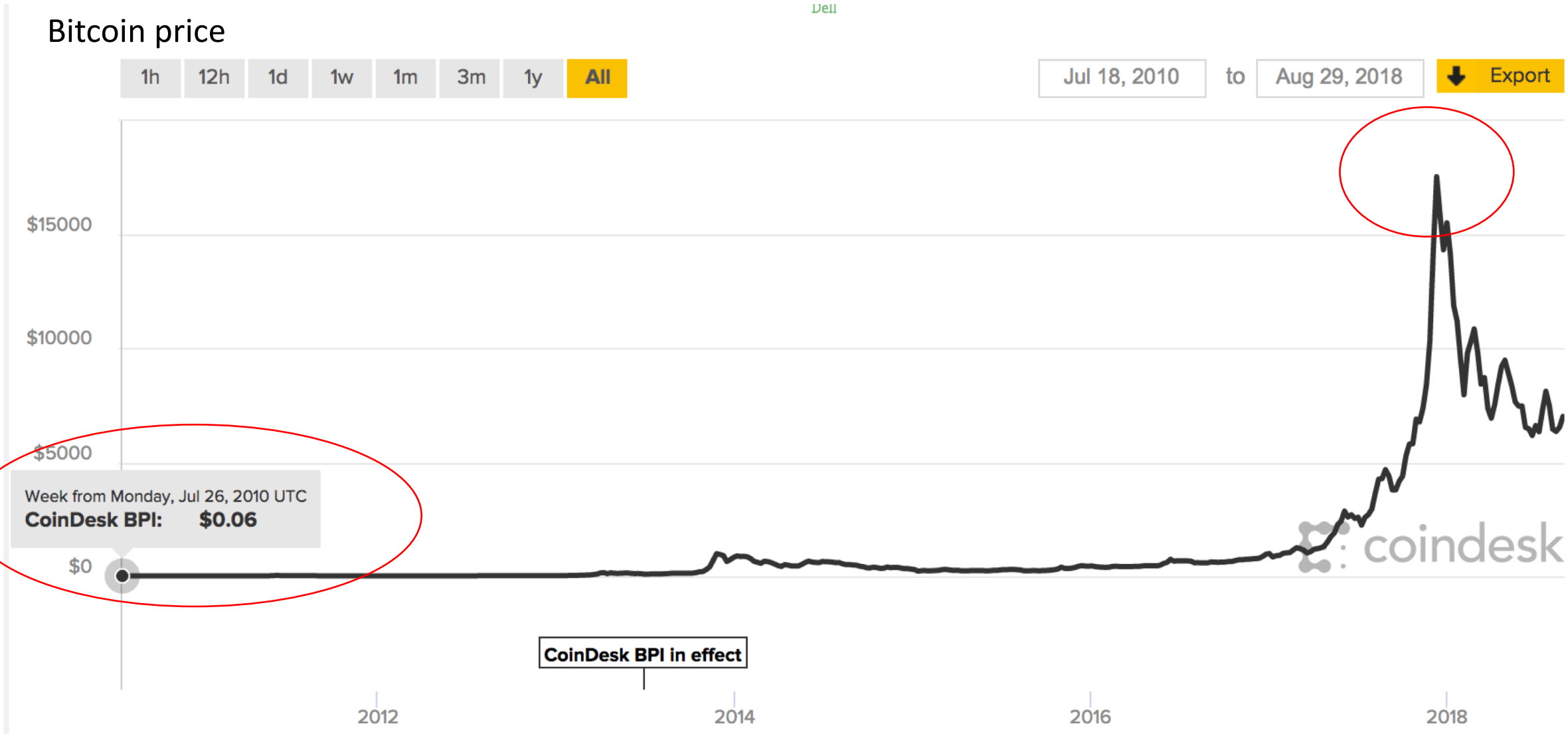






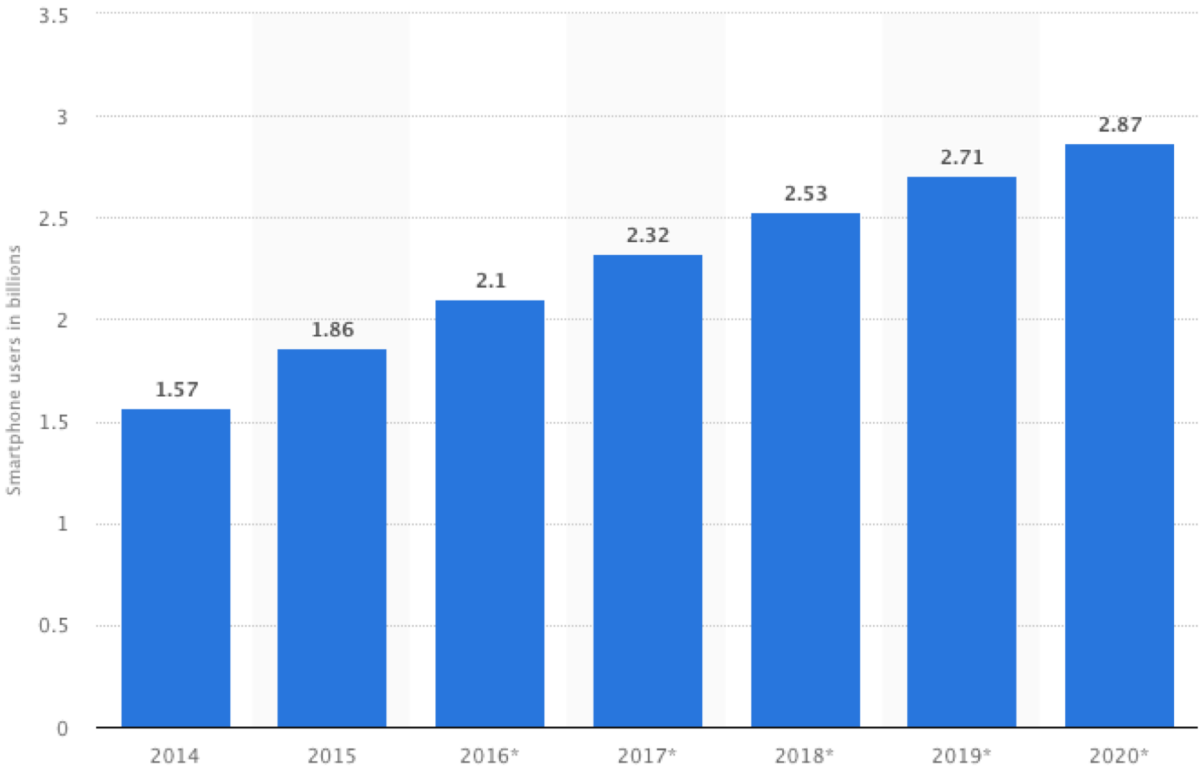
Security Technology Is Money Sometimes

Bitcoin price



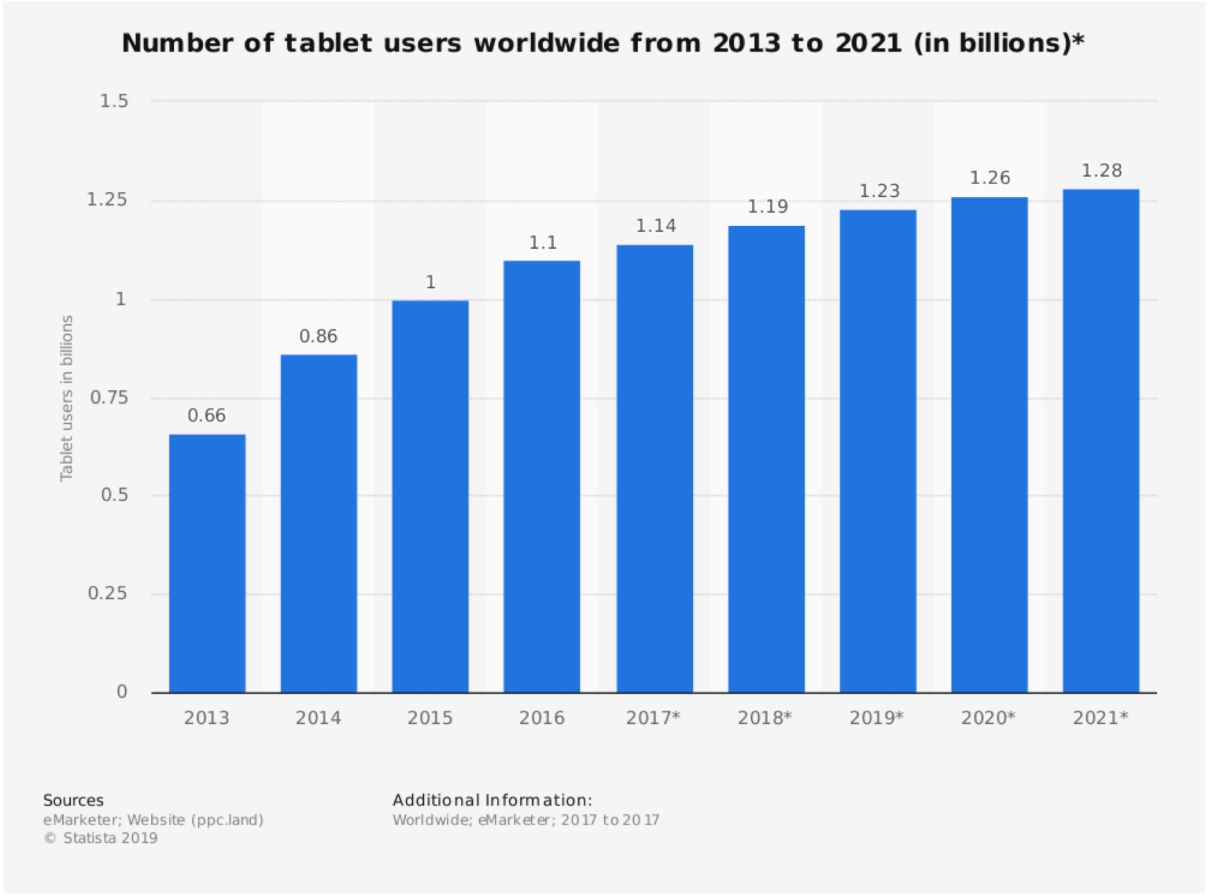
Mobile Devices and Flash Memory

Mobile Devices are Turning to Mainstream Computing Devices



Number of smartphone users worldwide from 2014 to 2020 (in billions)

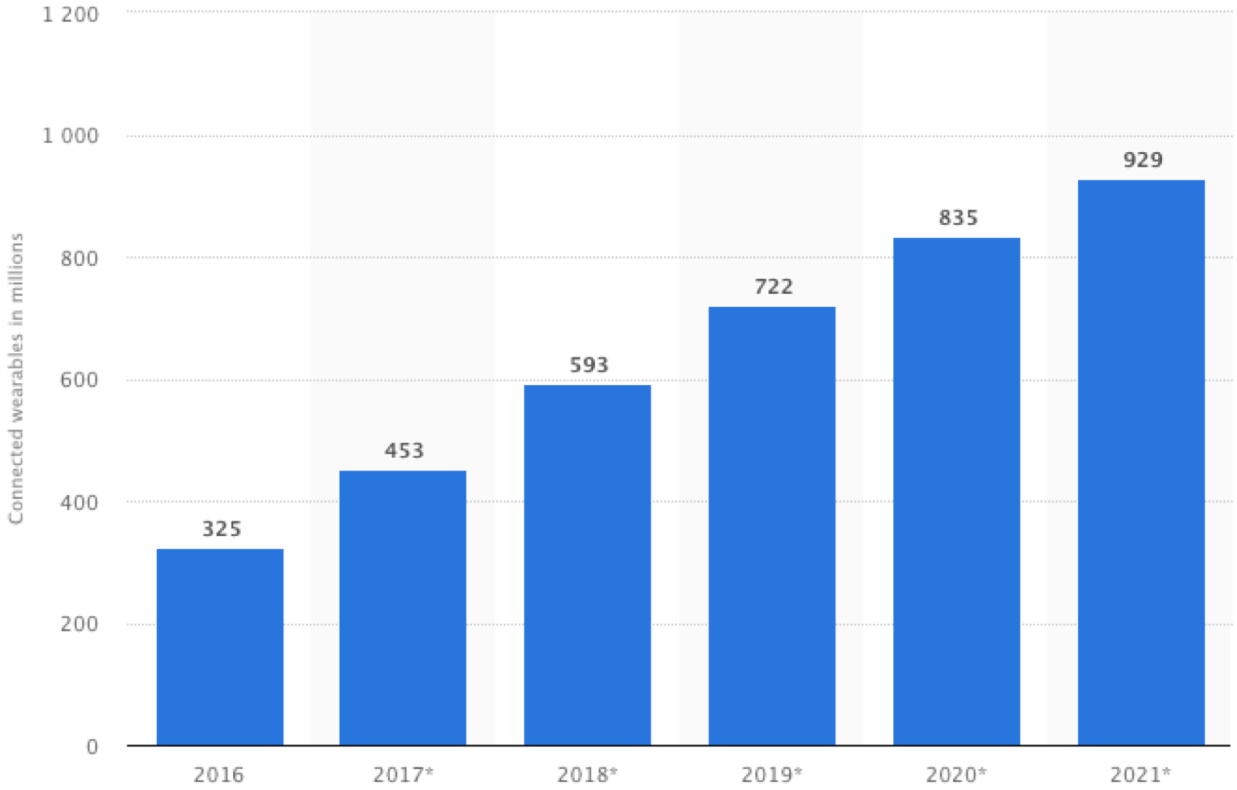




Number of tablet users worldwide from 2013 to 2021 (in billions)



Mobile Devices are Turning to Mainstream Computing Devices (cont.)



Number of connected wearable devices worldwide from 2016 to 2021 (in millions)



Mobile Devices are Used for Critical Applications

- Mobile devices are increasingly used to handle sensitive data
  - Online banking
  - Ecommerce
  - Cryptocurrency/stock trading
  - Naked photos
  - A human rights worker collects evidence of atrocities in a region of oppression
  - Etc.



- Security issues in mobile computing devices
  - Confidentiality
  - Integrity
  - Authentication
  - Access control
  - …

A Common Storage System of a Mobile Device

Applications layer

Files, APPs







Mobile file system layer

EXT4, EXT3, EXT2, etc.
Implement system calls like open, read, write, etc

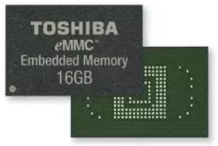
Manage the mappings between the applications' view and the block device's view

Block device layer

Flash memory layer







eMMC Chip

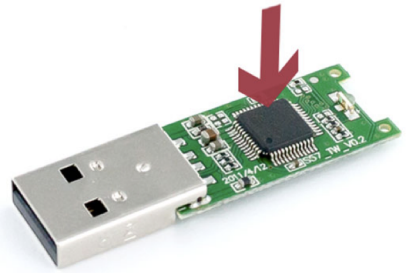


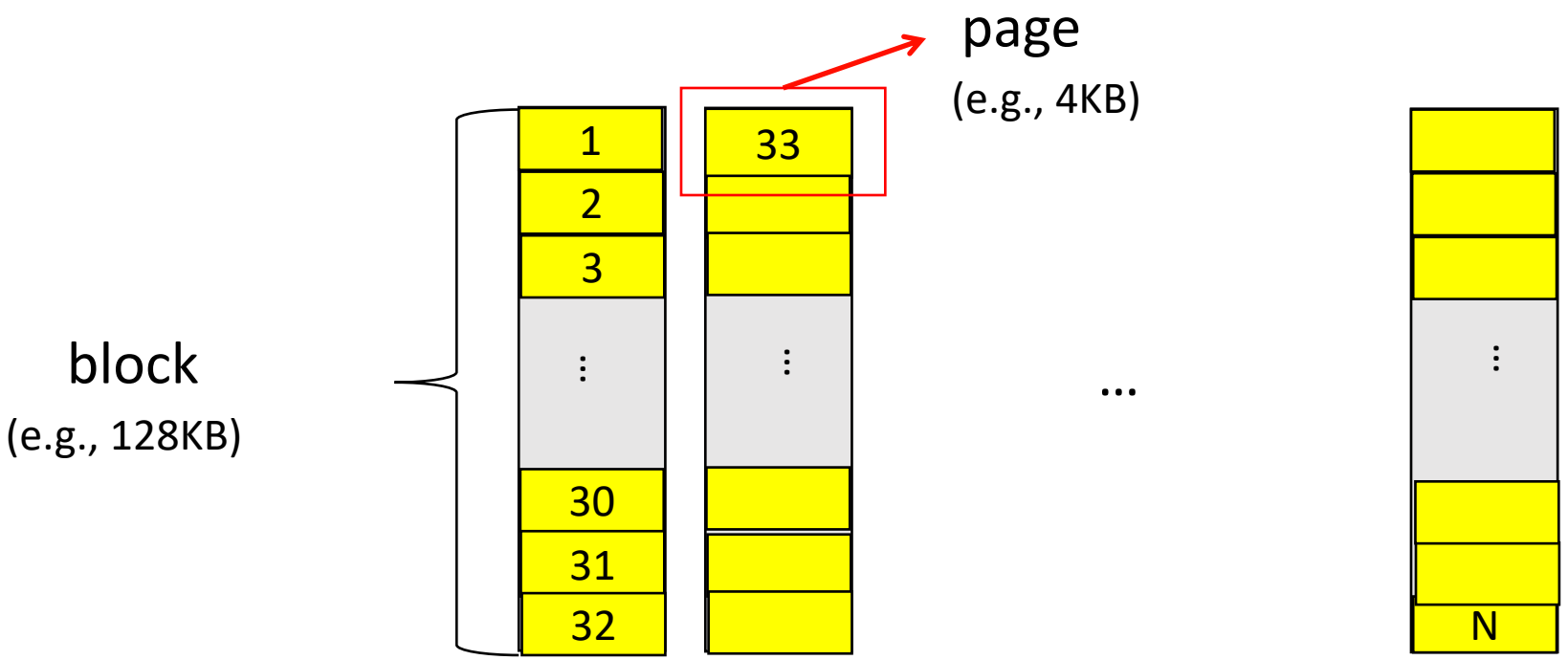
MMC Card

Flash Memory Security Research

NAND Flash is Usually Used as Storage Media

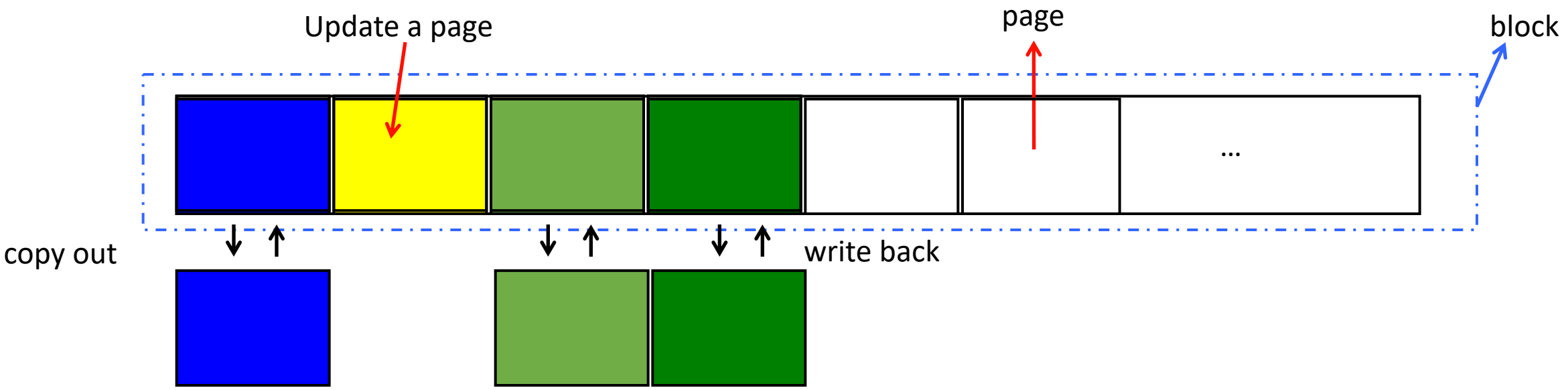
- NAND flash
  - USB sticks
  - Solid state drives (SSD)
  - SD/miniSD/microSD/eMMC





Special Characteristics of NAND Flash

- Update unfriendly
  - Over-writing a page requires first erasing the entire block
  - Write is performed in pages (e.g., 4KB), but erase is performed in blocks (e.g., 128KB)



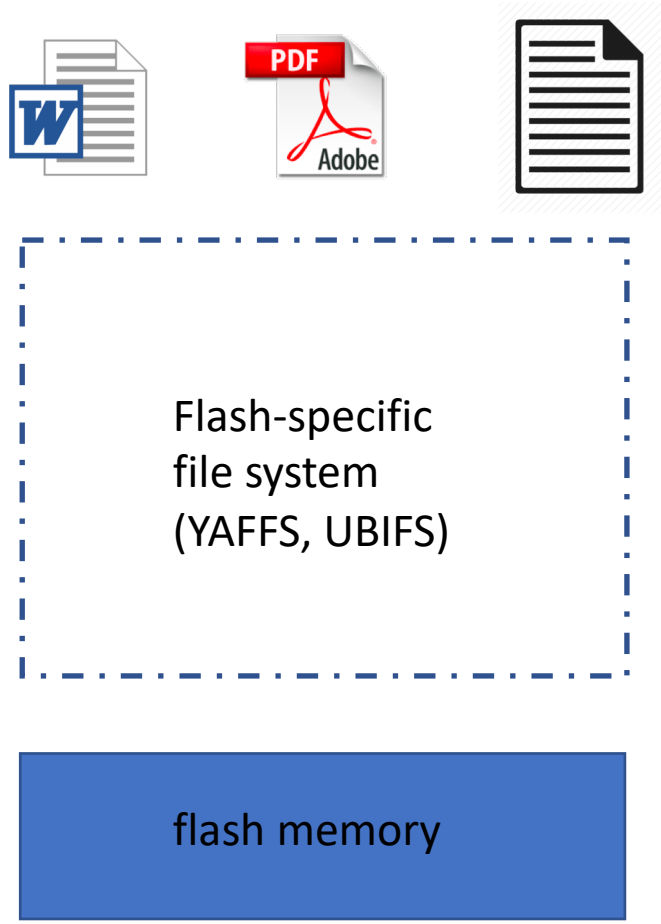
  - Over-write may cause significant write amplification

Special Characteristics of NAND Flash (cont.)

- Support a finite number of program-erase (P/E) cycles
  - Each flash block can only be programmed/erased for a limited number of times (e.g., 10K)
  - Data should be placed evenly across flash (wear leveling)

How to Manage NAND Flash

- Flash-specific file systems, which can handle the special characteristics of NDND flash
  - YAFFS/YAFFS2, UBIFS, F2FS, JFFS/JFFS2
  - Less popular



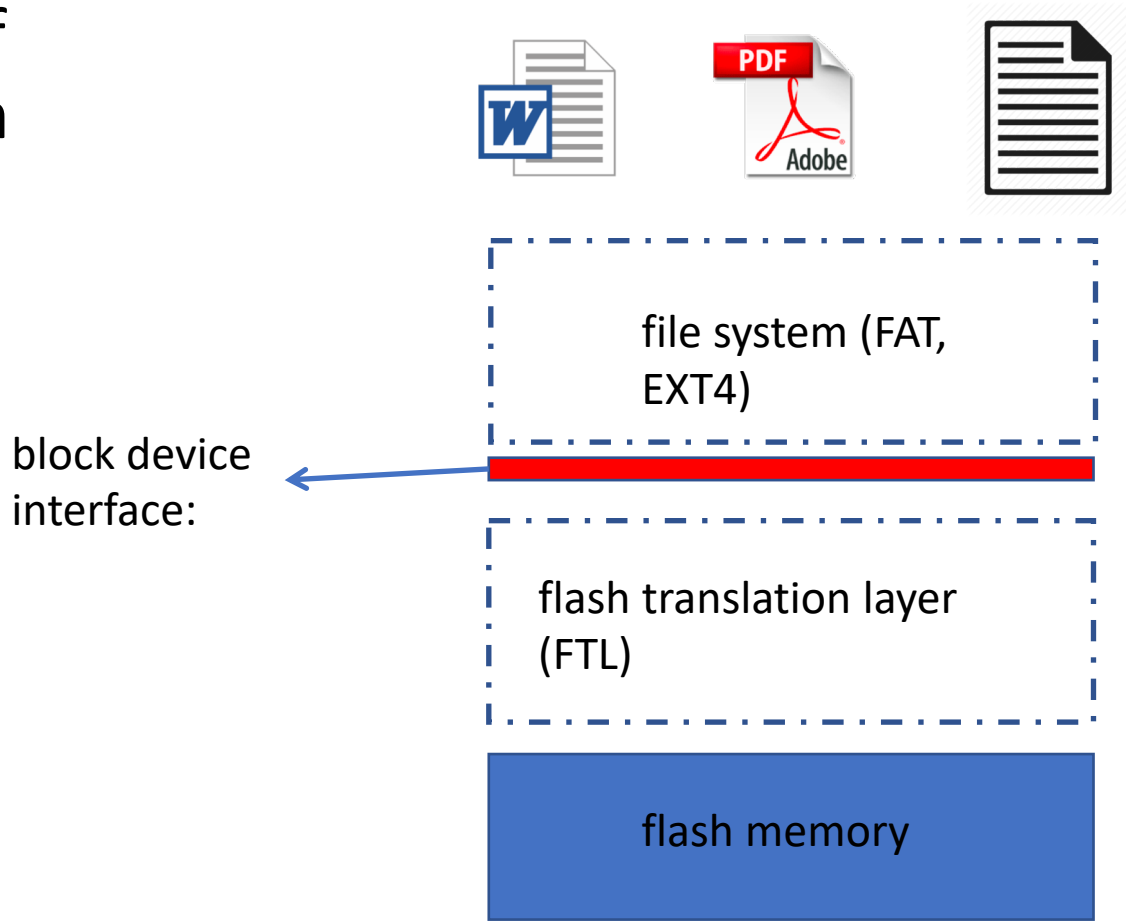
How to Manage NAND Flash (cont.)

- Flash translation layer (FTL) – a piece of flash firmware embedded into the flash storage device, which can handle the special characteristics of NAND flash and emulate the flash storage as a regular block device (most popular)
  - SSD
  - USB
  - SD



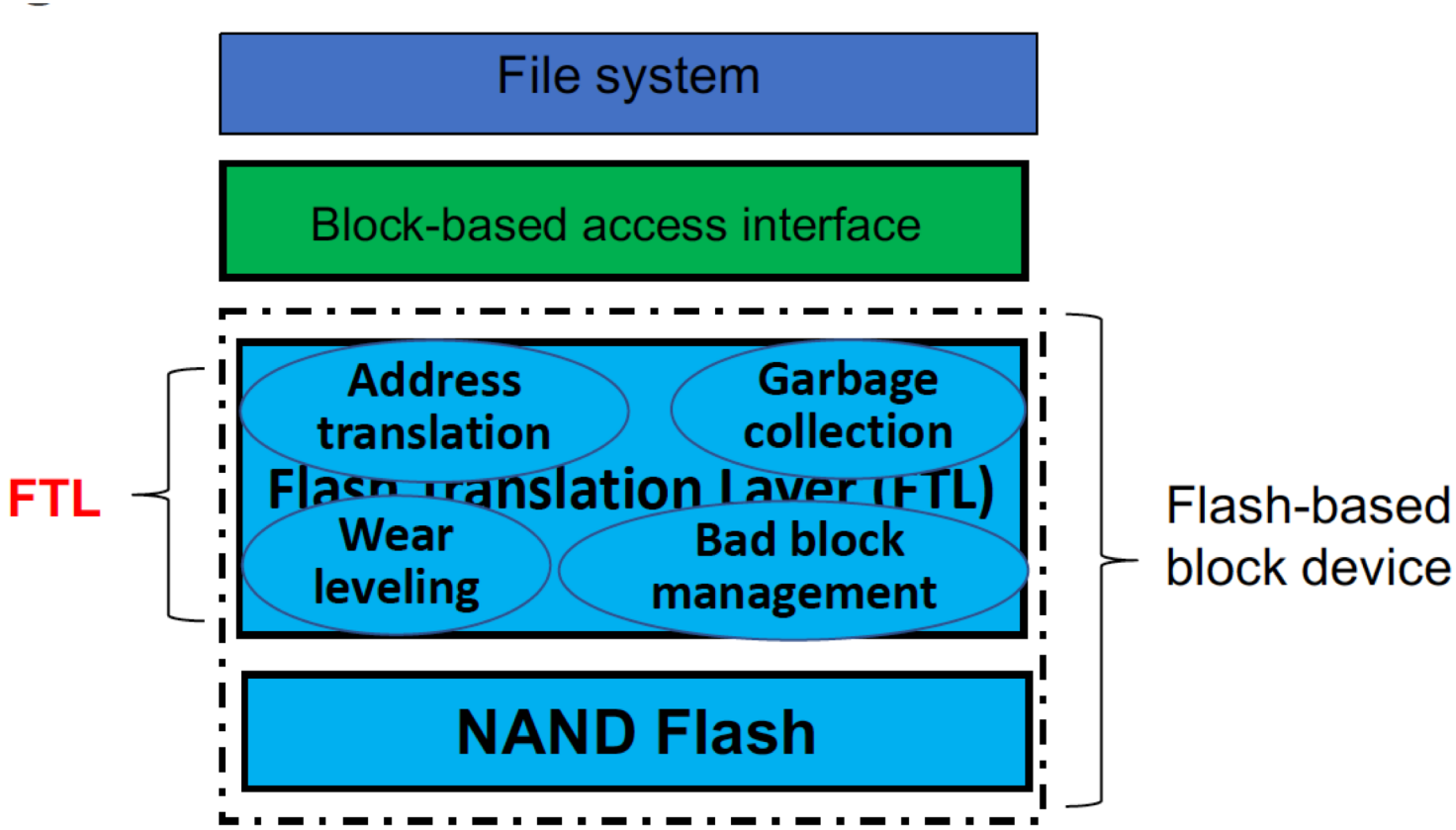






Flash Translation Layer (FTL)

- FTL usually provides the following functionality:
  - Address translation
  - Garbage collection
  - Wear leveling
  - Bad block management

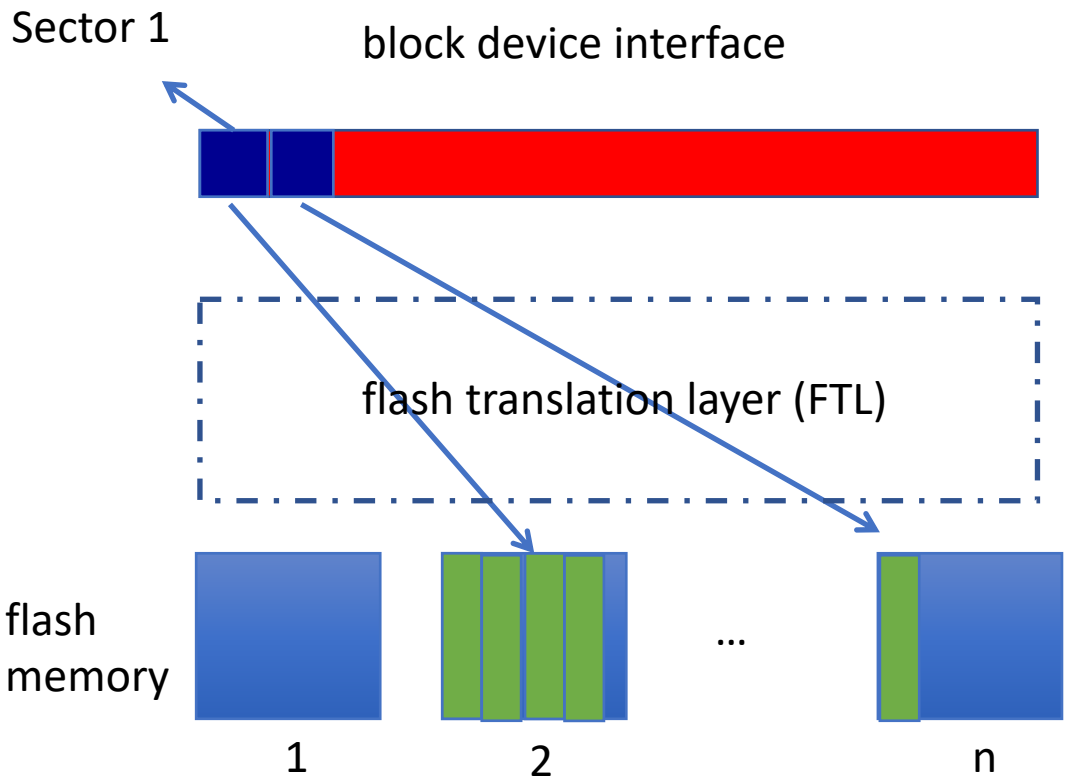


Flash Translation Layer (cont.)

- Address translation
  - Translate address between block addresses and flash memory addresses
  - Need to keep track of mappings between Logical Block Address (LBA) and Physical Block Address (PBA)

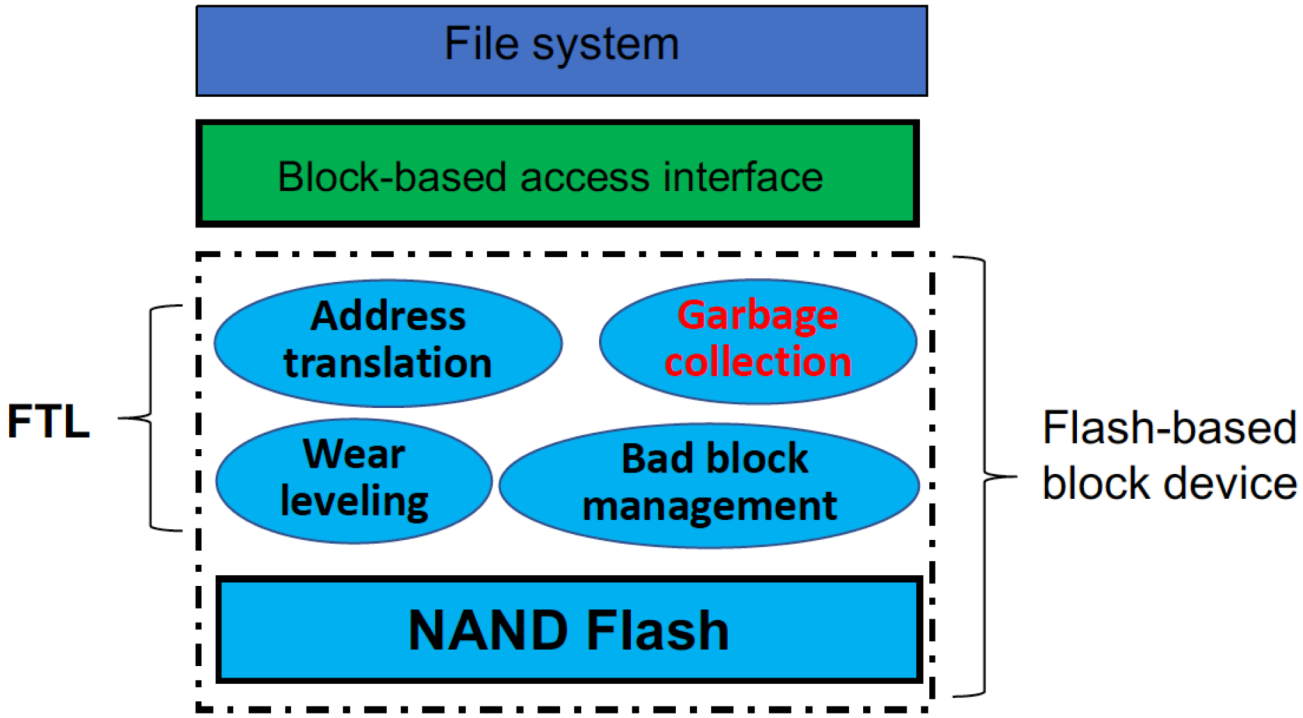
FTL should maintain a mapping table

Block device location	Flash location
Sector 1	(2,3)
Sector 2	(n,1)



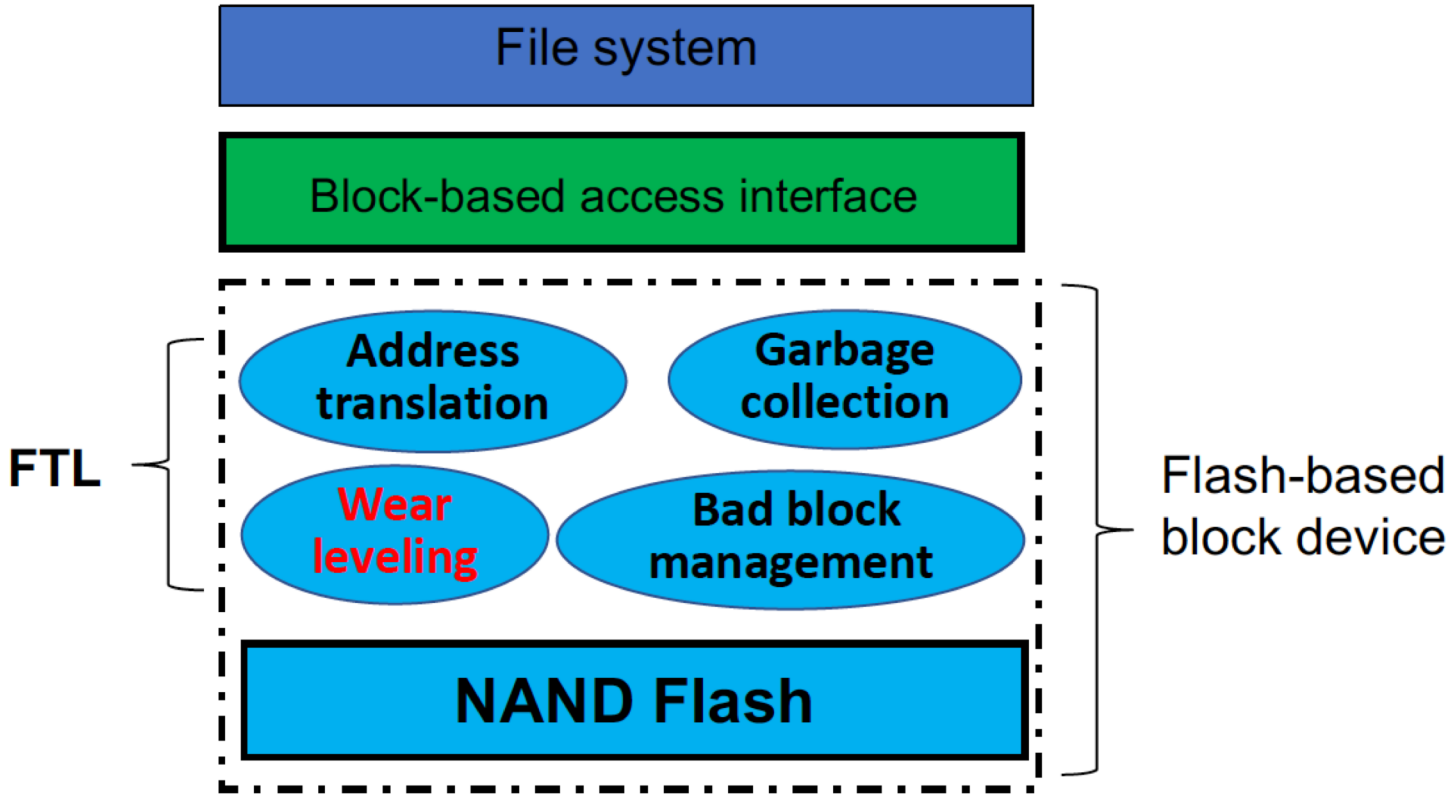
Flash Translation Layer (cont.)

- Garbage collection
  - Flash memory is update unfriendly
  - Not prefer in-place update, but prefer out-of-place update
  - The blocks storing obsolete data should be reclaimed periodically by garbage collection



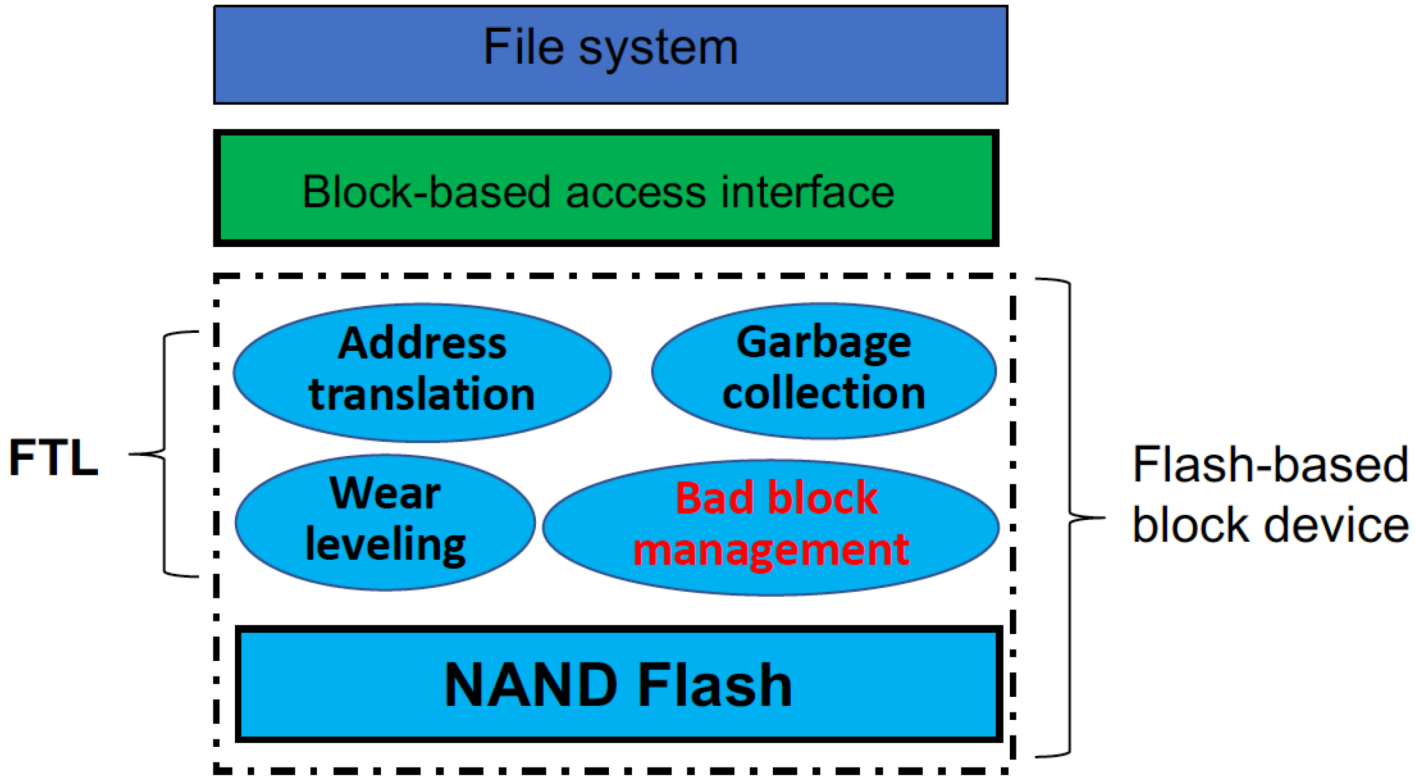
Flash Translation Layer (cont.)

- Wear leveling
  - Each flash block can be programmed/erased for a limited number of times
  - Distribute writes evenly across the flash to prolong its lifetime



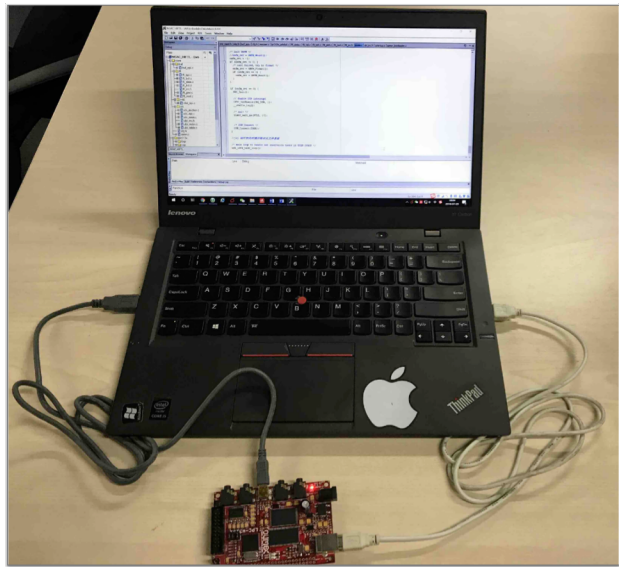
Flash Translation Layer (cont.)

- Bad block management
  - Regardless how good is the wear leveling, some flash blocks will eventually turn “bad” and cannot reliably store data
  - Bad block management is to manage these bad blocks



A Testbed for Flash Memory Security Research

- We have a flash memory testbed in MTU Security and Privacy (SnP) Lab. The lab is located in Rekhi 318



- The testbed includes:
  - Open-source flash firmware: OpenNFM
    - Implement flash translation layer (FTL) which is used to manage raw NAND flash, and provide a block access interface to upper layer
  - Embedded development environment: IAR Embedded Workbench
  - Electronic board: LPC-H3131



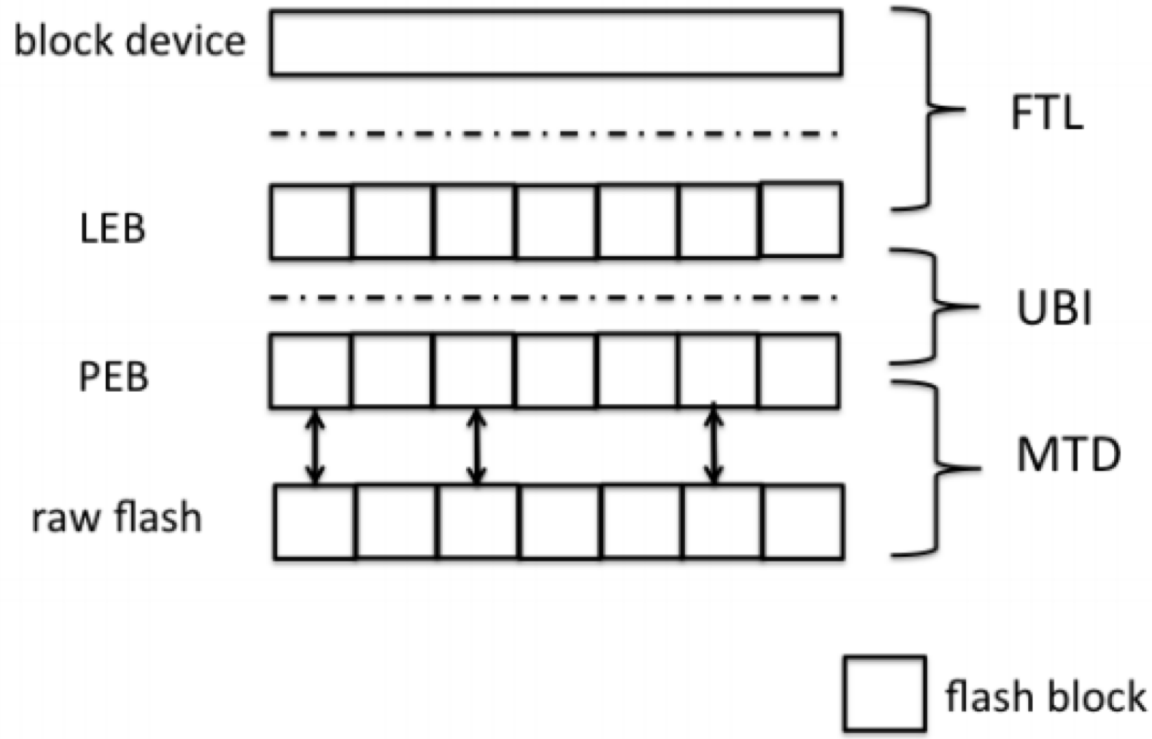
A Demo of Flash Memory Testbed

- A demo by Niusen Chen
  - Cross-compile opensource flash firmware OpenNFM using : IAR Embedded Workbench
  - Flash the binary to the electronic board LPC-H3131
  - Use the electronic board as a USB device (YOU CAN MAKE YOUR OWN USB DEVICE NOW)
  - Test throughput using benchmark tool **fio**

Opensource Flash Firmware OpenNFM







OpenNFM - MTD

- MTD: built on top of raw flash, and mainly provides three uniform APIs to allow the UBI to read, write and erase raw flash
  - MTD_Read(PEB index, offset, &data): read data from a PEB page identified by PEB index and offset
  - MTD_Write(PEB index, offset, data): write data to a PEB page identified by PEB index and offset
  - MTD_Erase(PEB index): erase the PEB identified by PEB index

OpenNFM - UBI

- UBI: built on top of MTD, and uses the APIs provided by MTD to read/write PEB pages or erase PEB blocks. Implement wear leveling, garbage collection, bad block management
  - UBI_Read(LEB index, offset, &data): read data from an LEB page identified by LEB index and offset
  - UBI_Write(LEB index, offset, data): write data to an LEB page identified by LEB index and offset
  - UBI_Erase(LEB index): erase an LEB identified by LEB index, which will cause an erasure over the corresponding PEB

OpenNFM - FTL

- FTL: build on top of UBI, and use the APIs provided by UBI to read/write LEB pages or erase LEBs. Implement address translation
  - FTL_Read(block_address, &data)
  - FTL_Write(block_address, data)

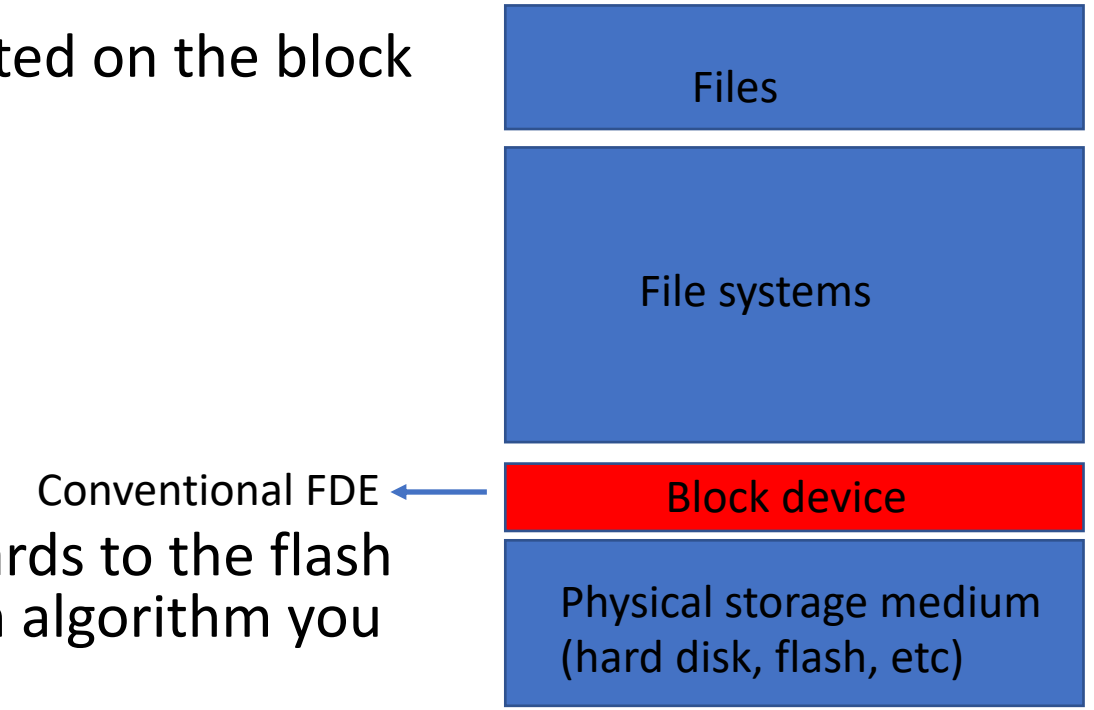
Hands-on Task 1 – Building A Flash Storage Device by Porting Open-source flash firmware to An Electronic Board

- Get familiar with the embedded development environment
- Play with the source code of flash firmware OpenNFM
- Port the OpenNFM to an electronic development board LPC-H3131

- Two students in a group
- Each group will be provided with a desktop and an electronic development board (with cables)
- Location: Rekhi 318

Hands-on Task 2 – Integrating Disk Encryption into The Flash Translation Layer

- Conventional full disk encryption (FDE) is incorporated on the block layer to transparently ensure data confidentiality
  - BitLocker (Windows)
  - FileVault (MAC OS X)
  - Android FDE
  - TrueCrypt/VeraCrypt



- We will move the FDE from the block layer downwards to the flash memory layer (implement an symmetric encryption algorithm you like)
  - Write a flash page, encrypt it
  - Read a flash page, decrypt it using the same key
  - Encryption/Decryption is completely transparent to users
  - Test the throughput using benchmark tool fio, compare the throughput with original OpenNFM firmware