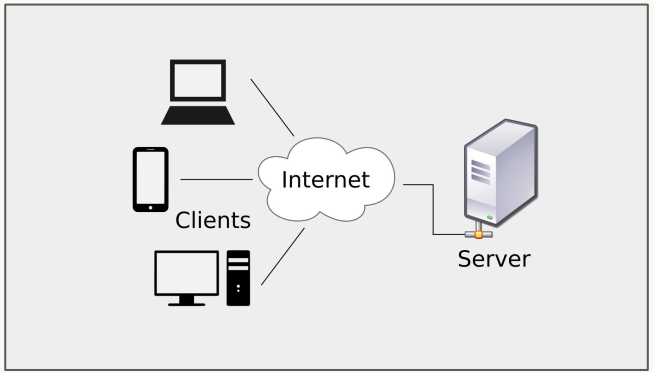

ACSL
Accelerated Computing
Systems Lab

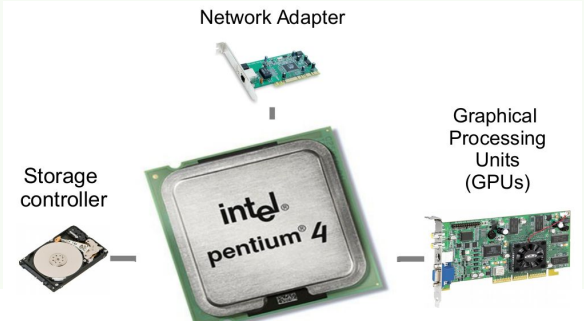
מפגש

ACSL
Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

What do computer systems look like?

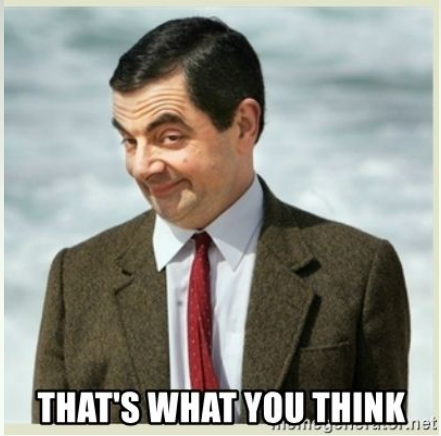


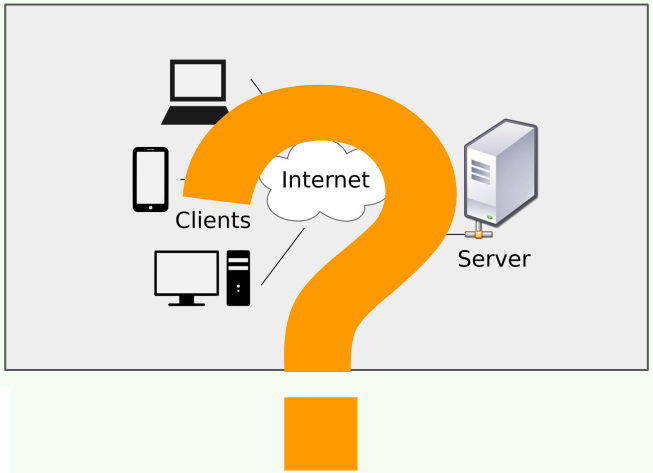


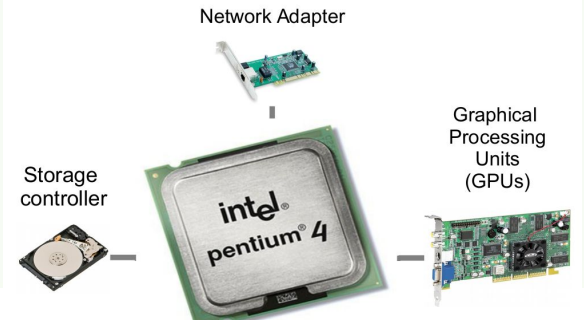
Applications

Operating System

Hardware
CPU, Disk, Keyboard..







Applications

Operating System

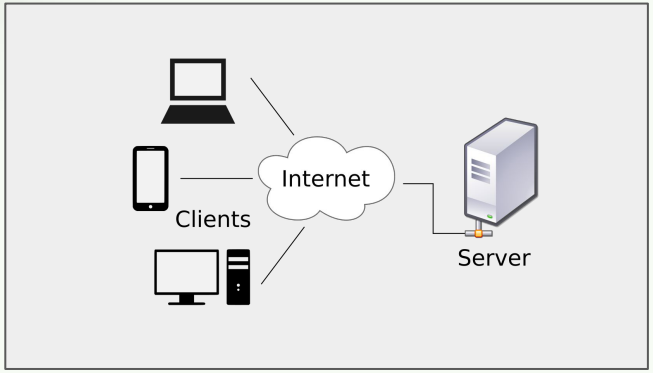
Hardware
CPU, Disk, Keyboard..

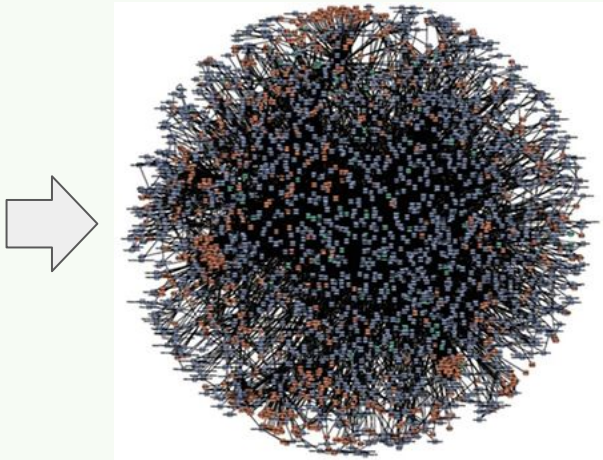
מפגש

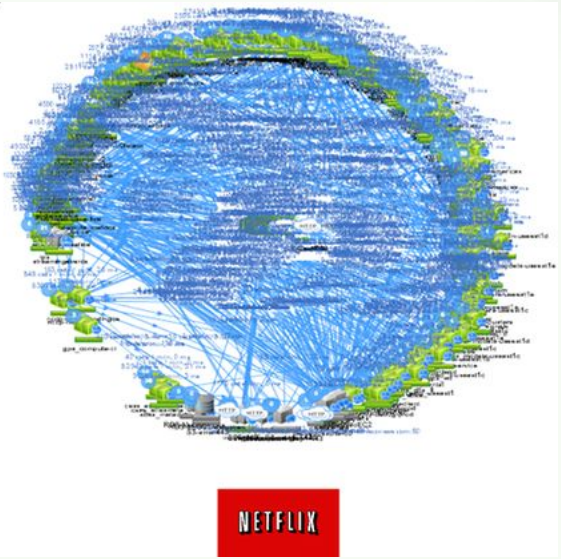
ACSL
Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

Thousands of inter-operating distributed services



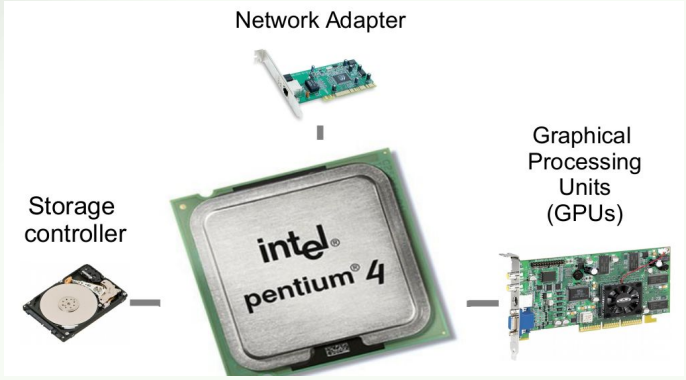




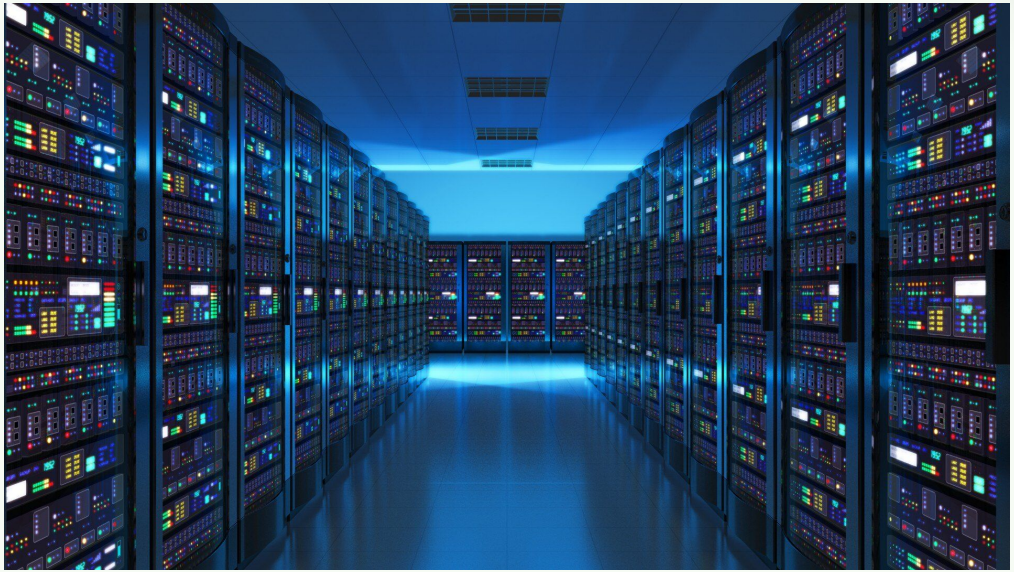
מפגש

ACSL
Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי



Data centers with millions of servers, accelerators, networks



מפגש

ACSL
Accelerated Computing Systems Lab

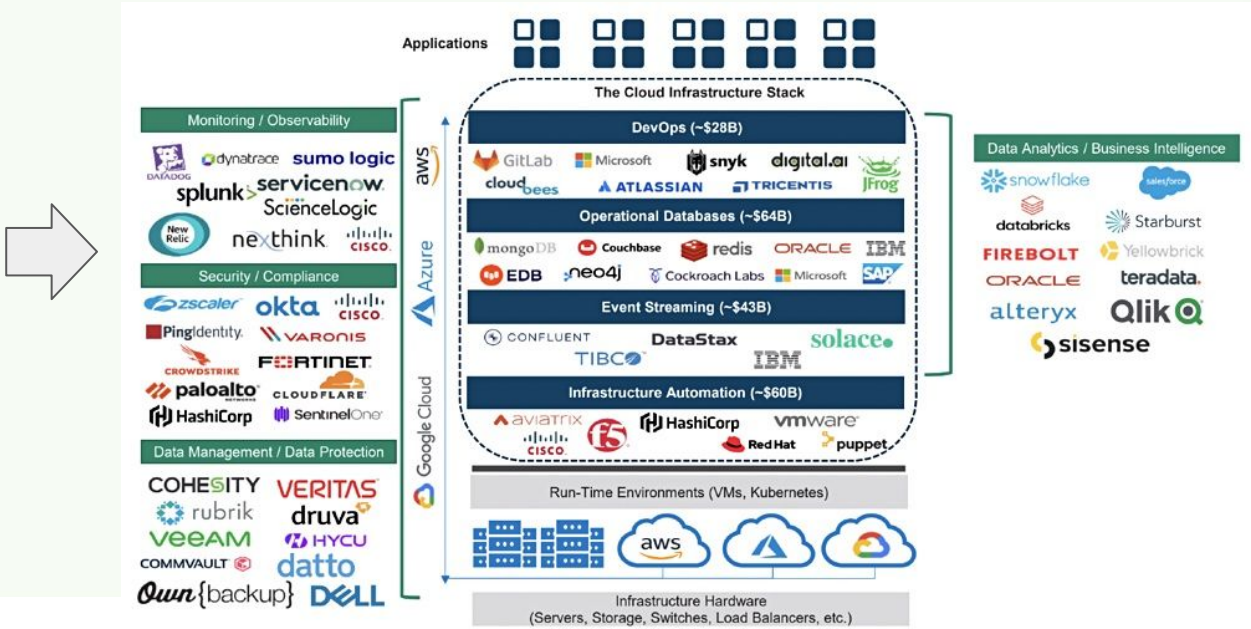
הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

A multi-layer infrastructure with complex interactions

Applications

Operating System

Hardware
CPU, Disk, Keyboard..



Source: William Blair Equity Research

מפגש 360°

ACSL
Accelerated Computing Systems Lab

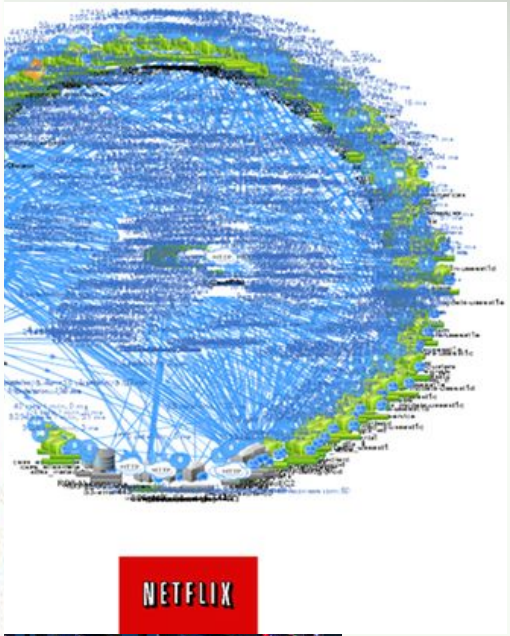
הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

amazon.com

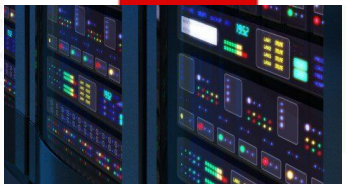
NETFLIX

Source: William Blair Equity Research

WELCOME TO REALITY

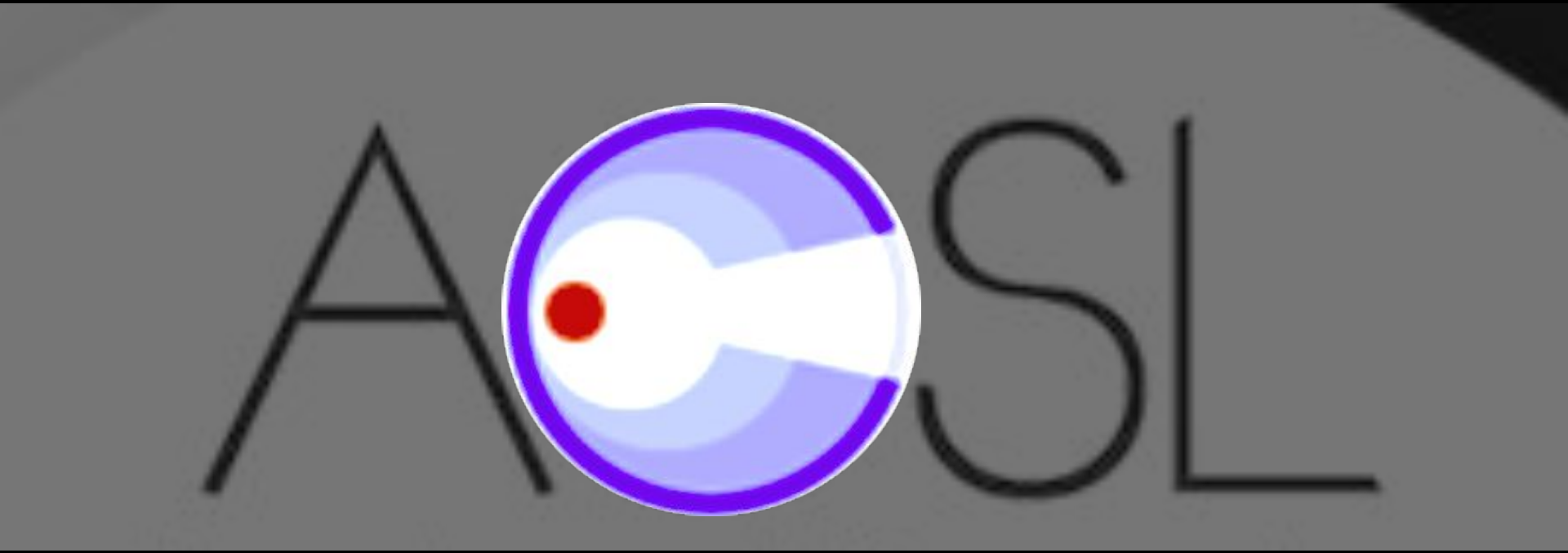


NETFLIX



Source: William Blair Equity Research

Your window into the world of



computer systems research

מפגש

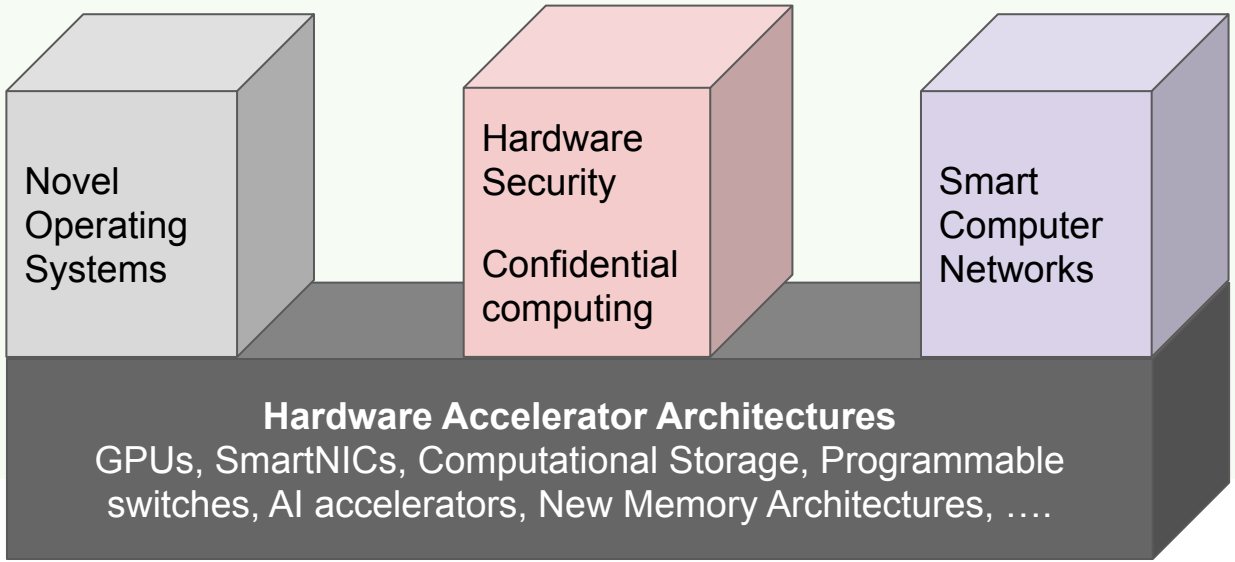
ACSL
Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים
ע"ש אנדרו וארנה ויטרבי

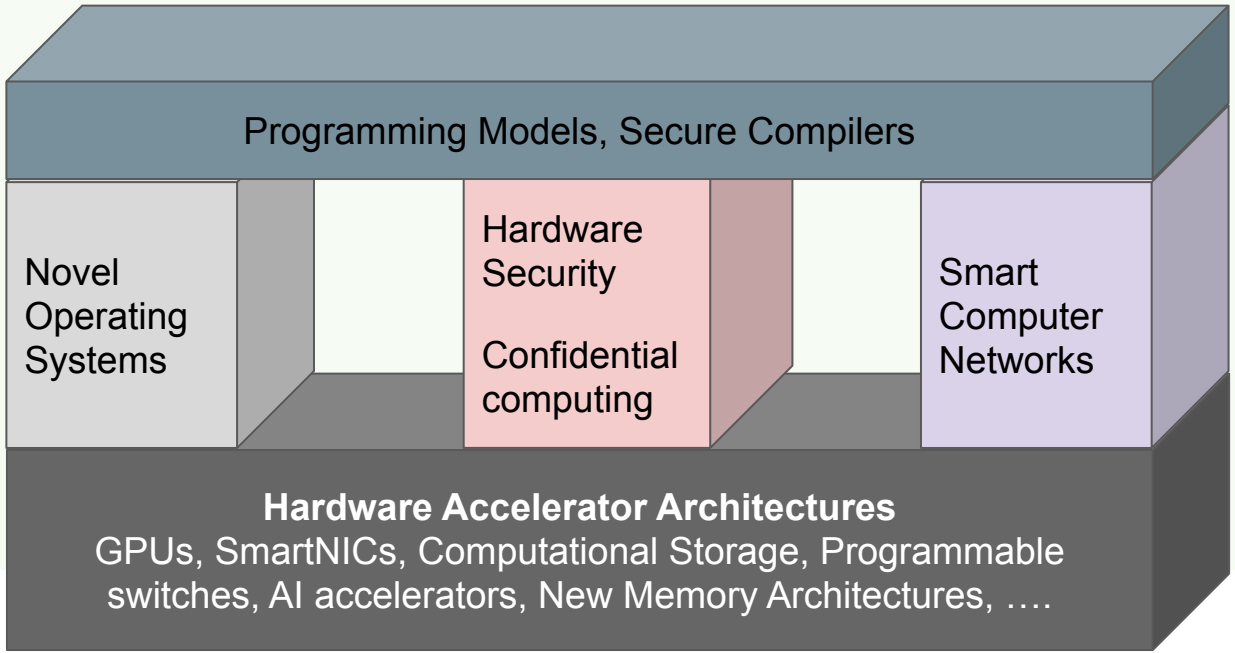
Holistic research of computer systems

Hardware Accelerator Architectures
GPUs, SmartNICs, Computational Storage, Programmable switches, AI accelerators, New Memory Architectures, ….

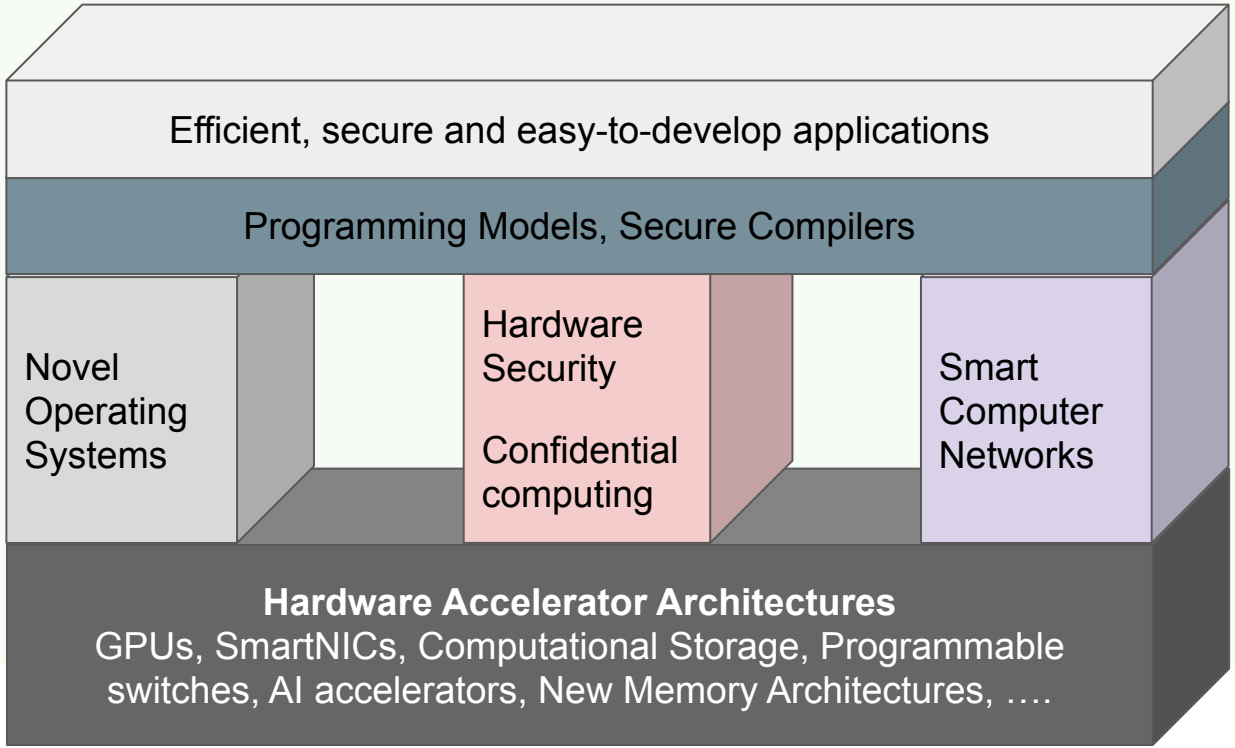
Holistic research of computer systems



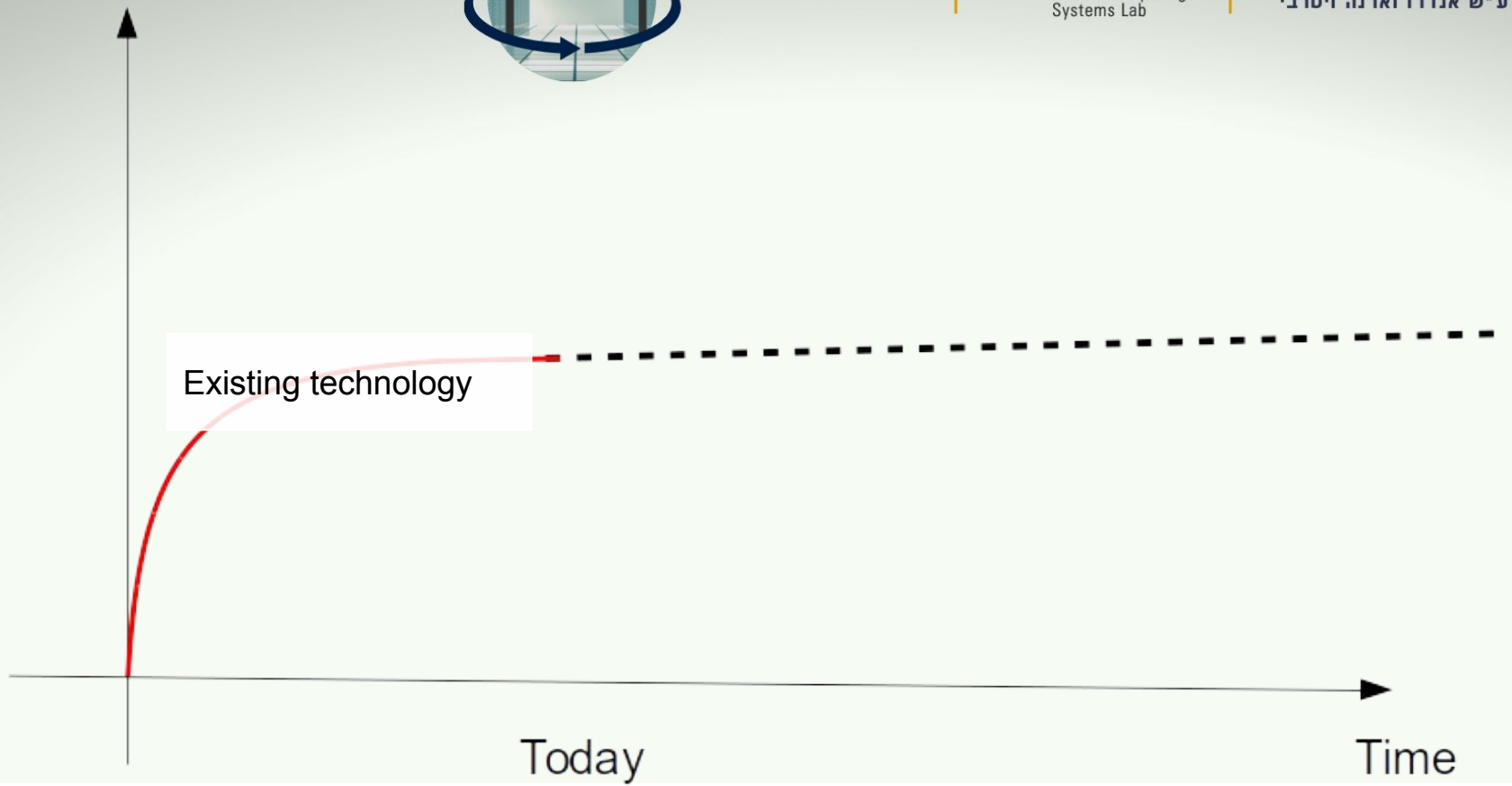
Holistic research of computer systems

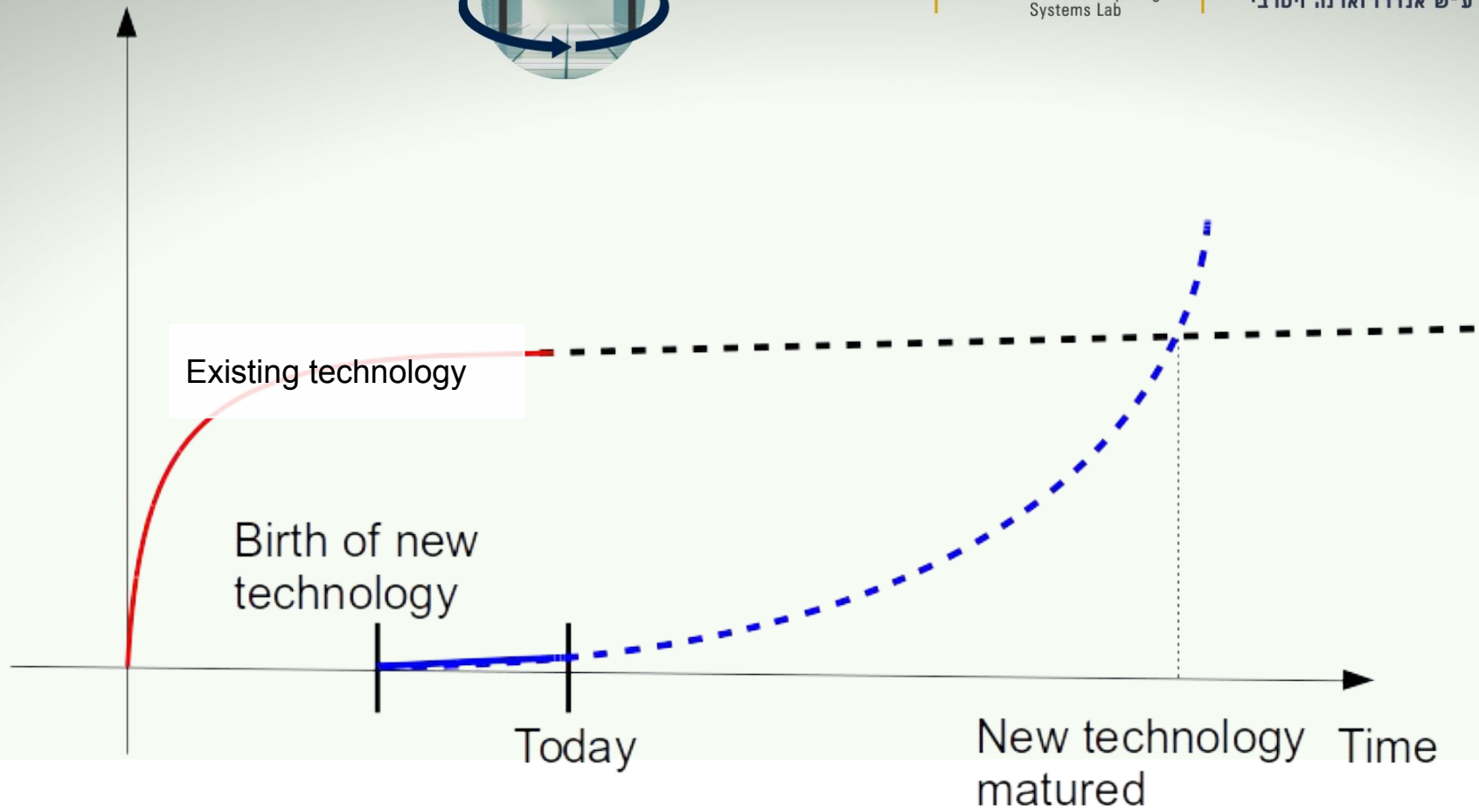


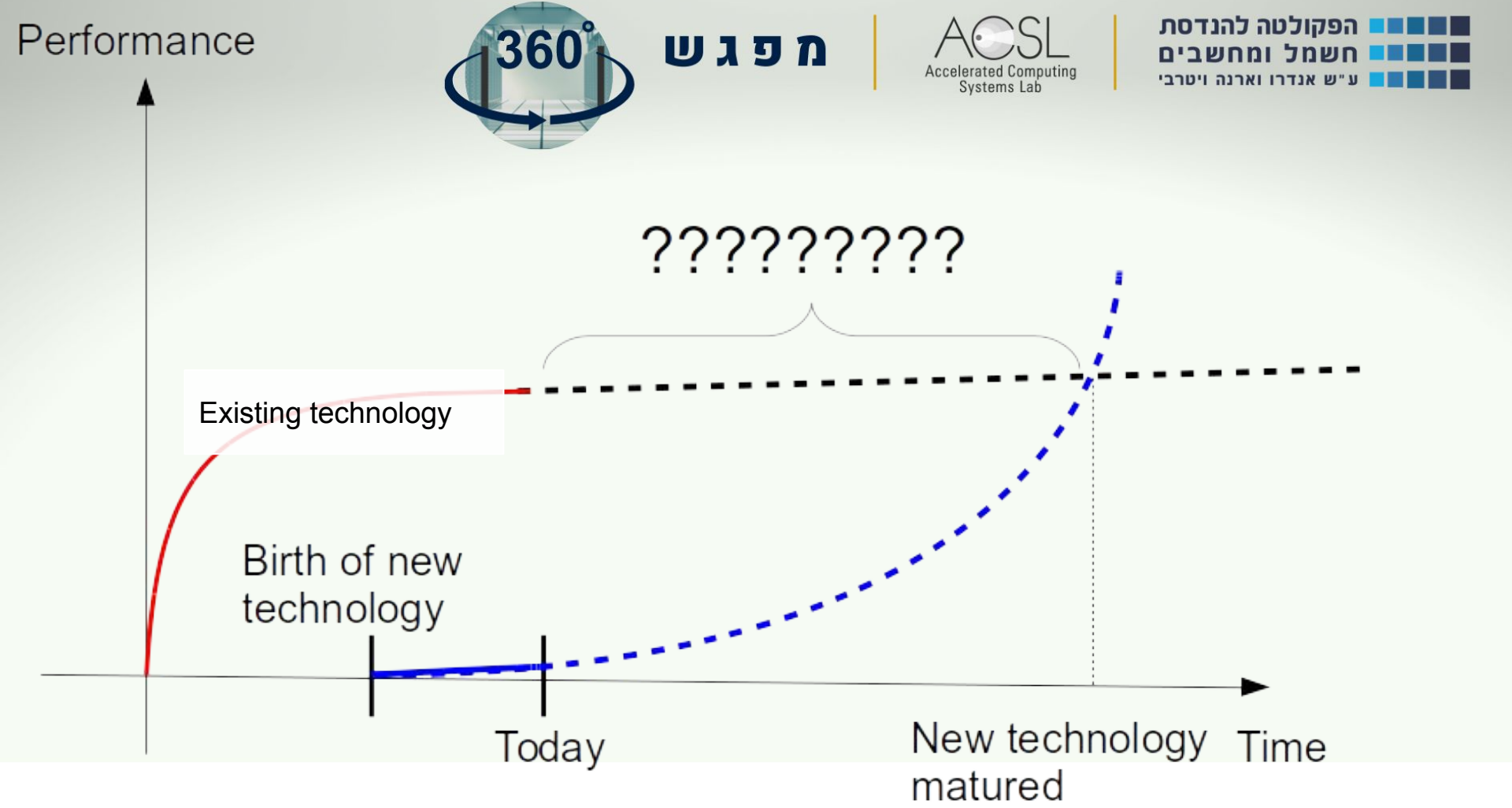
Holistic research of computer systems



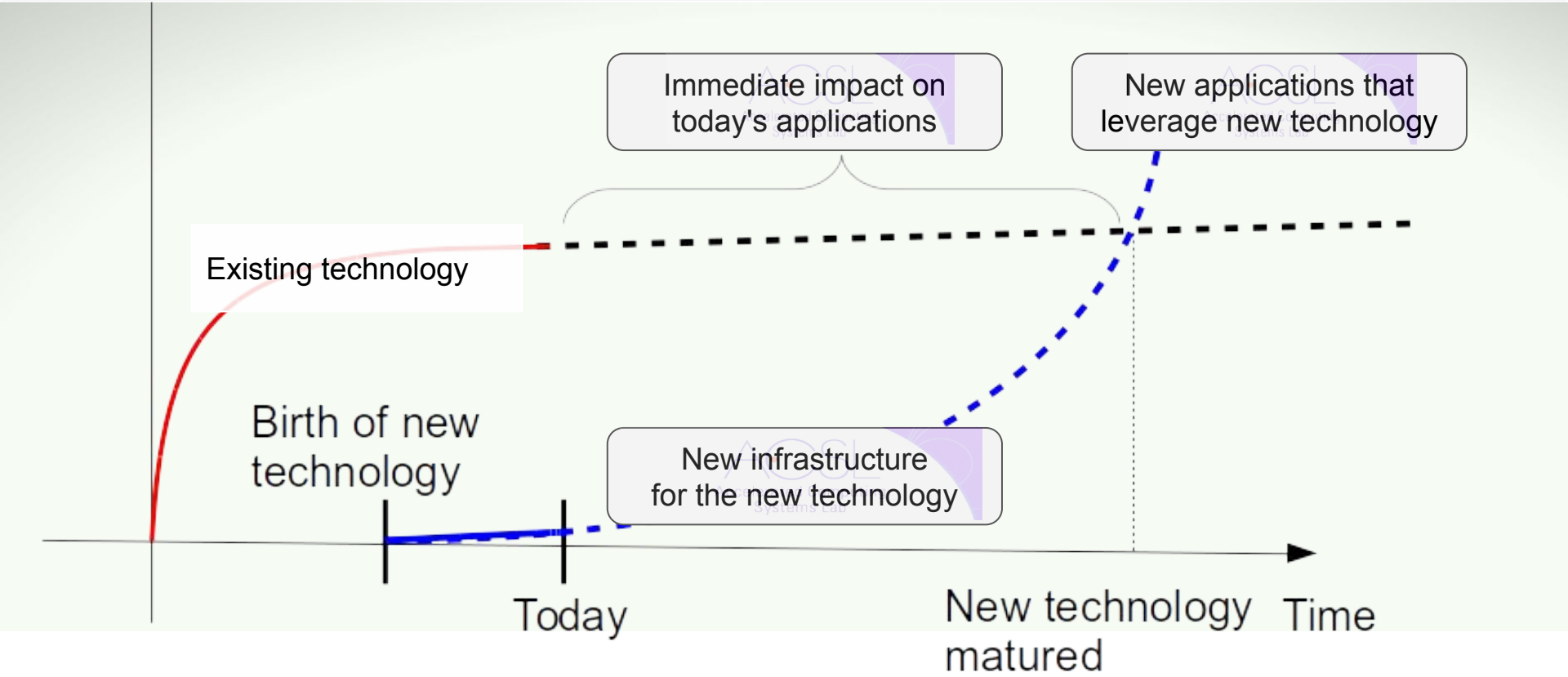
Why research?


Performance
Existing technology
Today
Time

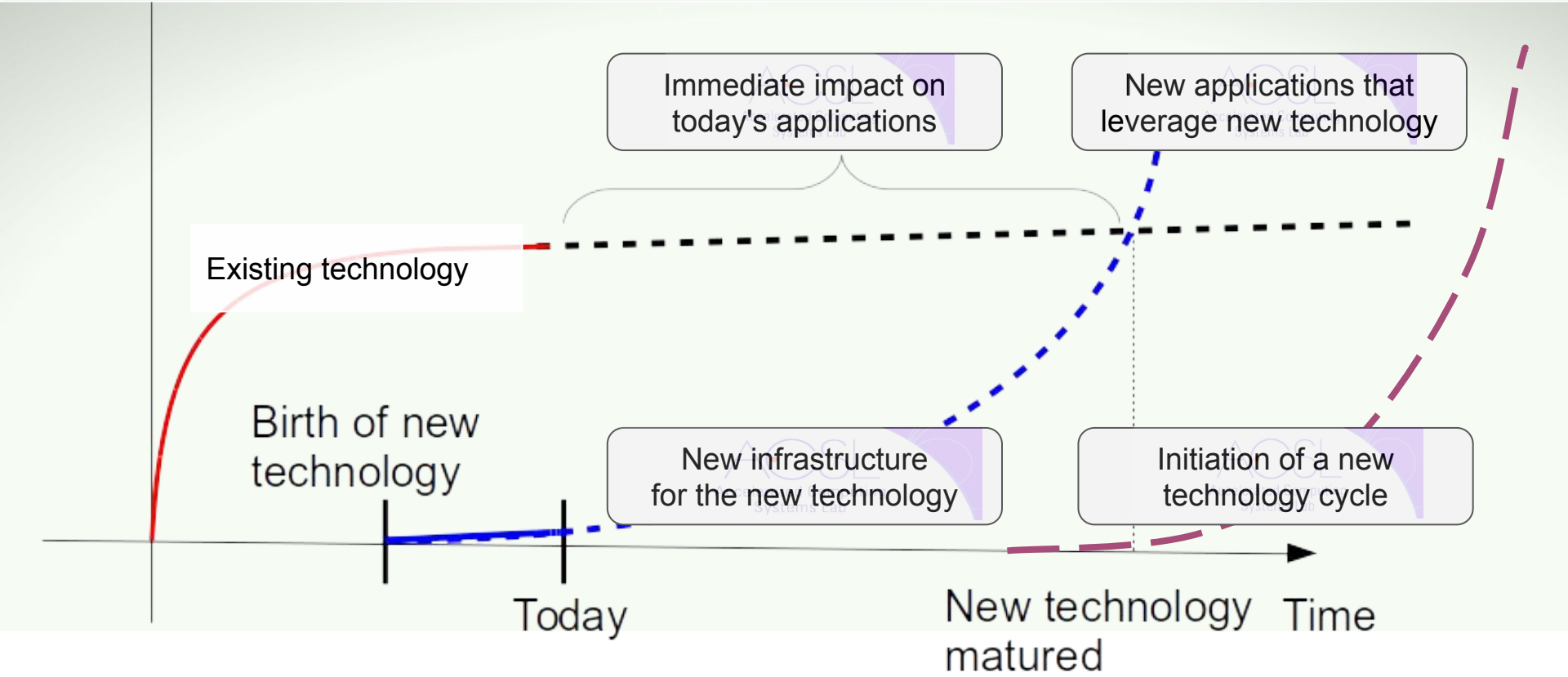

Performance
Existing technology
Birth of new technology
Today
New technology matured
Time


Performance
?????????
Existing technology
Birth of new technology
Today
New technology matured
Time

We explore emerging technologies to leverage them for applications of tomorrow



We explore emerging technologies to leverage them for applications of tomorrow



Big questions

- How to build faster systems by pushing computations ***into the network and memory***
- How to build efficient Operating Systems ***for future computer architectures***
- How to build/verify/test systems that are ***provably*** secure
- How to break technology **limits** using ***deterministic*** ML models

Success story 1: Weaknesses in Intel SGX

- Technological driver
  - **Confidential computing** technology introduced in 2016
  - **Trusted execution** (data is encrypted in memory)
- Hypothesis (2018)
  - **SGX suffers from hardware secure vulnerabilities** of regular CPUs despite encryption

Read more here: https://foreshadowattack.eu

Success story 1: Weaknesses in Intel SGX

- Technological driver
  - **Confidential computing** technology introduced in 2016
  - **Trusted execution** (data is encrypted in memory)
- Hypothesis (2018)
  - **SGX suffers from hardware secure vulnerabilities** of regular CPUs despite encryption
- Result (2018)
  - Discovery of a **Foreshadow bug in Intel CPUs** that allowed to **leak data from SGX**
- Impact – immediate
  - Disclosure to vendors who spent man-years on fixing it
  - Understanding fundamental limits of the confidential computing hardware design
  - Reported by **Wired, Arstechnica, Hacker News**, and many others..
- Acquired/required skills
  - OS internals, advanced computer architecture, crypto/cyber security, hacking

Read more here: https://foreshadowattack.eu

מפגש

ACSL Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים ע"ש אנדרו וארנה ויטרבי

Success story 2: Automatic detection of CPU vulnerabilities

- Technological driver
  - In 2018 a new concept of **CPU speculative execution vulnerabilities** was discovered
  - All major CPUs are vulnerable
  - Correct programs leak secrets, broken isolation
- Problem (2018)
  - **No systematic way to find new vulnerabilities**

Success story 2: Automatic detection of CPU vulnerabilities

- Technological driver
  - In 2018 a new concept of **CPU speculative execution vulnerabilities** was discovered
  - All major CPUs are vulnerable
  - Correct programs leak secrets, broken isolation
- Problem (2018)
  - **No systematic way to find new vulnerabilities**
- Result (2022)
  - **A tool to automatically find new vulnerabilities**
- Impact – immediate and counting
  - Open Source project led by Microsoft Research, several new vulnerabilities found in X86 CPUs, AMD already fixed
- Acquired/Required skills
  - Advanced computer architecture, cyber security

מפגש

ACSL Accelerated Computing Systems Lab

הפקולטה להנדסת חשמל ומחשבים ע"ש אנדרו וארנה ויטרבי

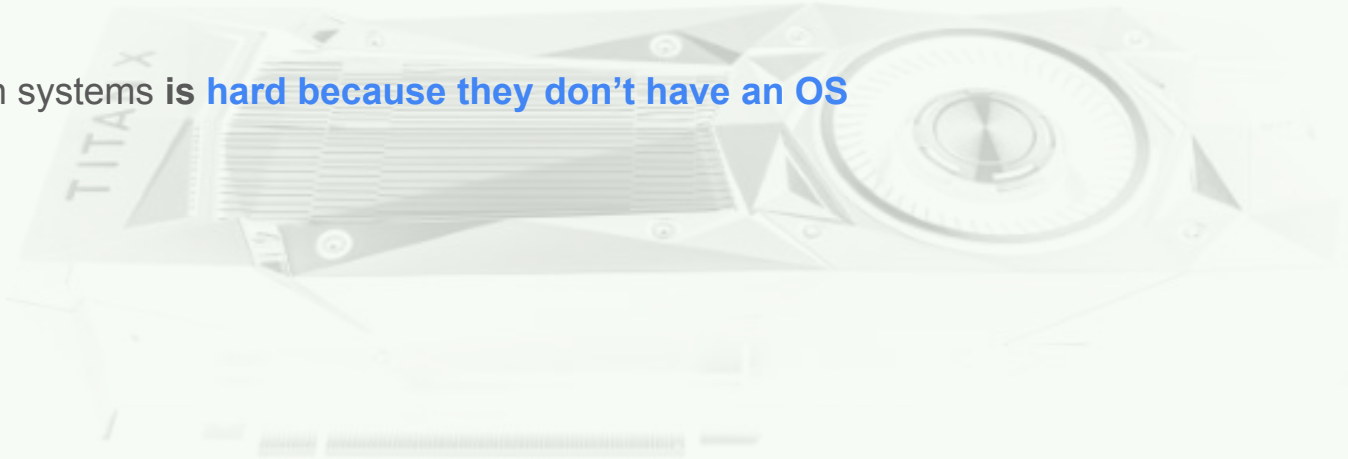
Success story 2: Automatic detection of CPU vulnerabilities

- Technological driver
  - In 2018 a new concept of **CPU speculative ex… vulnerabilities** was discovered
  - All major CPUs are vulnerable
  - Correct programs leak secrets, broken i…
- Problem (2018)
  - **No systematic way to find new …abilities**
- Result (2022)
  - **A tool to automatically f… vulnerabilities**
- Impact – immediate … counting
  - Open Source proj… by Microsoft Research, several new vulnerabilities found in X86 CPUs, AMD al… ed
- Acquired/Re… skills
  - Advanced co…puter architecture, cyber security

ACTIVE Project

Success story 3: Operating System for GPUs

- Technological driver (circa 2013)
  - **Graphical Processing Units (GPUs)** speed up machine learning by 10x
- Problem (2011)
  - Using GPUs in systems **is** **hard because they don't have an OS**

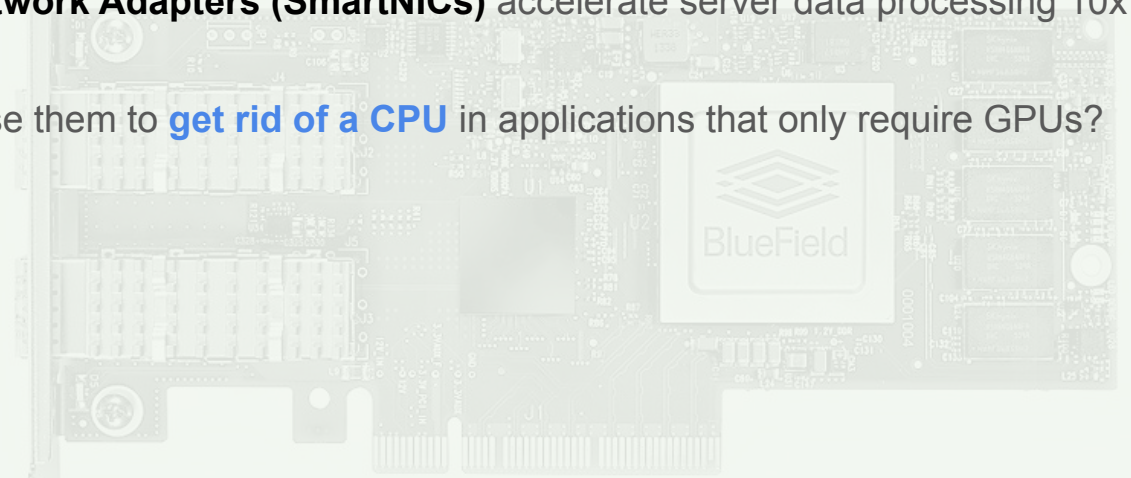


Success story 3: Operating System for GPUs

- Technological driver (circa 2013)
  - **Graphical Processing Units (GPUs)** speed up machine learning by 10x
- Problem (2011)
  - Using GPUs in systems **is hard because they don't have an OS**
- Result (2013)
  - **New OS that runs on GPUs and allows efficient access to files and network**
- Impact – 9 years and counting
  - **Conceptual change** in the way of thinking about computational **accelerators** (adopted by recent startups)
  - More than 1000 citations by other research papers
  - Some parts of the OS adopted by NVIDIA (in 2021), AMD and Mellanox in their products
- Acquired/Required skills
  - OS internals, parallel processing, GPU programming, network programming, kernel programming, driver development

Success story 4

- Technological driver (circa 2015)
  - **Smart Network Adapters (SmartNICs)** accelerate server data processing 10x
- Idea (2017)
  - Can we use them to **get rid of a CPU** in applications that only require GPUs?



Success story 4

- Technological driver (circa 2015)
  - **Smart Network Adapters (SmartNICs)** accelerate server data processing 10x
- Idea (2017)
  - Can we use them to **get rid of a CPU** in applications that only require GPUs?
- Result (2020)
  - **The first computing system that runs Deep Neural Networks on 100 GPUs without using X86 CPUs (saves 100 CPU cores)**
- Impact – 2 years and counting
  - Paves the way to a more efficient server design
  - Some parts adopted by Huawei, recent startups
- Acquired/Required skills
  - Concurrent programming, GPU programming, RDMA networking, NIC driver development, SmartNIC hardware architectures

Success story 4

- Technological driver (circa 2015)
  - **Smart Network Adapters (SmartNICs)** accele… …ver data processing 10x
- Idea (2017)
  - Can we use them to **get rid of a CPU** in… …tions that only require GPUs?
- Result (2020)
  - **The first computing system th… …Deep Neural Networks on 100 GPUs without using X86 CPUs (saves 100 CPU …**
- Impact – 2 years and co…
  - Paves the way to a m… …ent server design
  - Some parts adopte… …awei, recent startups
- Acquired/Requir… …s
  - Concurrent … …ming, GPU programming, RDMA networking, NIC driver development, SmartNIC h… …are architectures

ACTIVE PROJECT

Success story 5: network processing by using Neural Nets

- Technological driver
  - **Neural Nets** are getting faster on modern processors
  - **Network packet processing is the bottleneck in modern data centers**
- Idea (2018)
  - We can use **Neural Nets to accelerate packet processing**.

Success story 5: network processing by using Neural Nets

- Technological driver
  - **Neural Nets** are getting faster on modern processors
  - **Network packet processing is the bottleneck in modern data centers**
- Idea (2018)
  - We can use **Neural Nets to accelerate packet processing**.
- Result (2020)
  - **The first algorithm and a system that uses NNs for packet processing, up to 100x faster**
- Impact – 2 years and counting
  - Fundamentally new approach to packet processing
  - Integrated into production system, working on further adoption
  - Active collaboration with Intel on hardware-accelerated design
- Acquired/Required skills
  - CPU performance optimization, Packet processing, Networking, Neural Networks

Success story 5: network processing by using Neural Nets

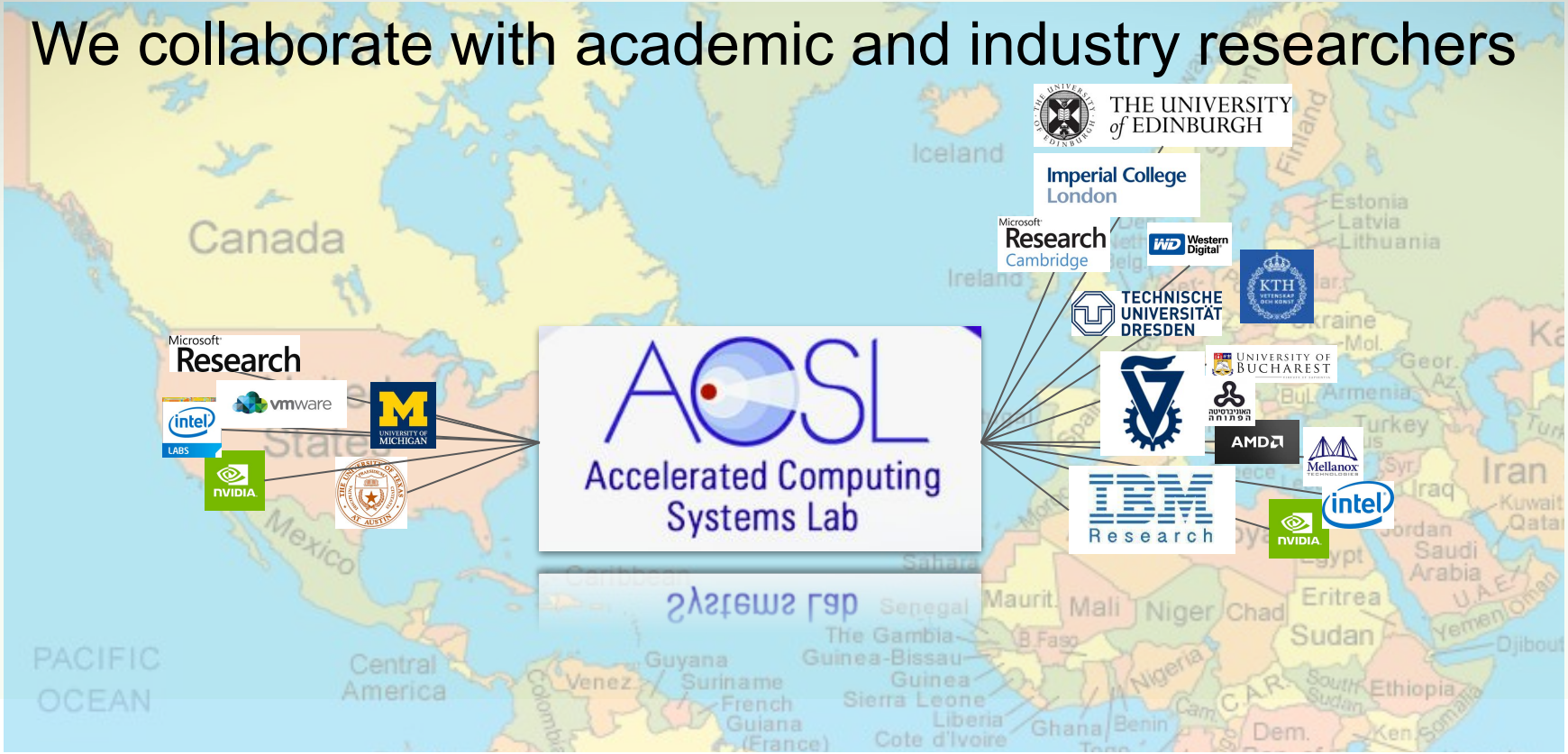
- Technological driver
  - **Neural Nets** are getting faster on modern processo
  - **Network packet processing is the bottleneck … ern data centers**
- Idea (2018)
  - We can use **Neural Nets to accelerate … rocessing**.
- Result (2020)
  - **The first algorithm and a syste… es NNs for packet processing, up to 100x faster**
- Impact – 2 years and counti…
  - Fundamentally new appro… acket processing
  - Integrated into producti… n, working on further adoption
  - Active collaboration … on hardware-accelerated design
- Acquired/Require… s
  - CPU perform… timization, Packet processing, Networking, Neural Networks

ACTIVE PROJECT

Ongoing projects

- **Emerging hardware**
  - In-memory computing systems
- **In-network computing and programmable networks**
  - In-network address translation in virtual networks
  - Accelerated Byzantine Fault Tolerance
  - Packet Programs in data centers
  - Switch-driven Congestion Control
  - Neural Net-driven packet classification
- **Microarchitectural security**
  - Verifying constant-timeness X86 instructions
  - Design-stage testing of CPU vulnerabilities
- **Trusted Hardware**
  - Provably-secure defences against untrusted OSes
- **Cloud OS**
  - Eliminating nested virtualization overheads
  - Multi-cloud networking

We collaborate with academic and industry researchers



So why does it matter **to you?**

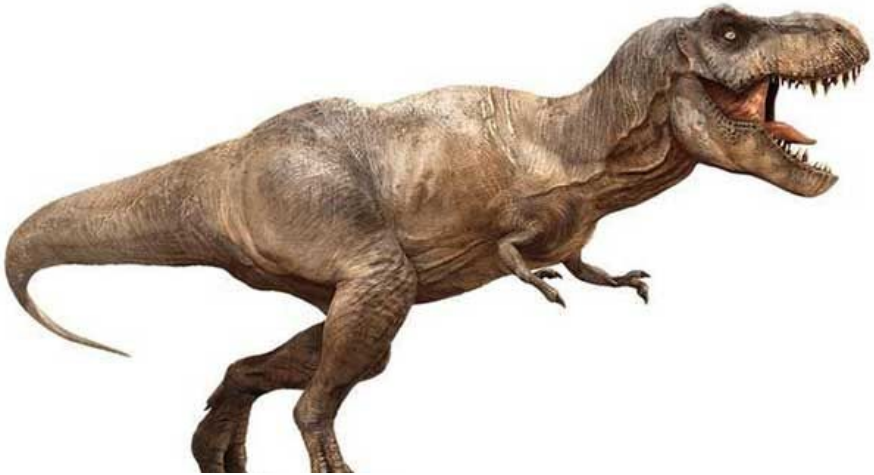
Life ***in the industry*** is often tough

Engineers

Constraints

Life in industry is often tough

Investors

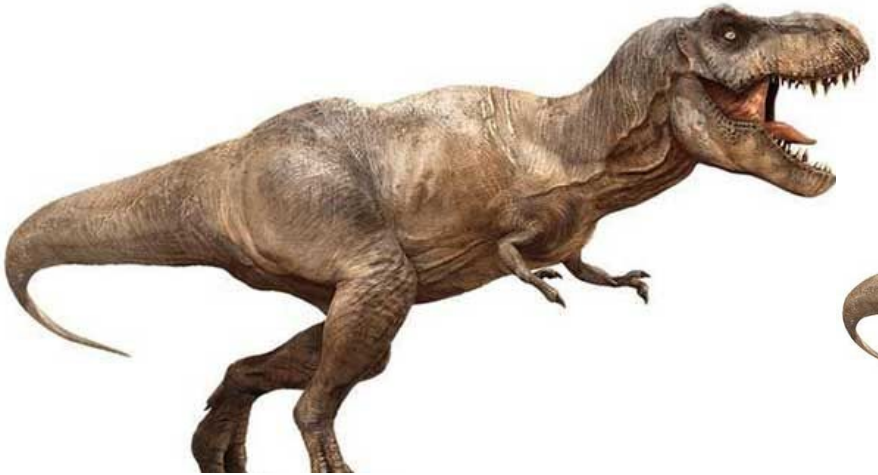


Engineers

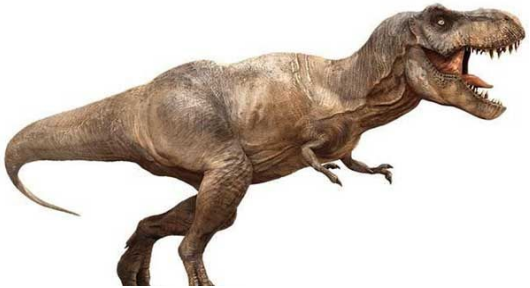
Constraints

Life in industry is often tough

Investors

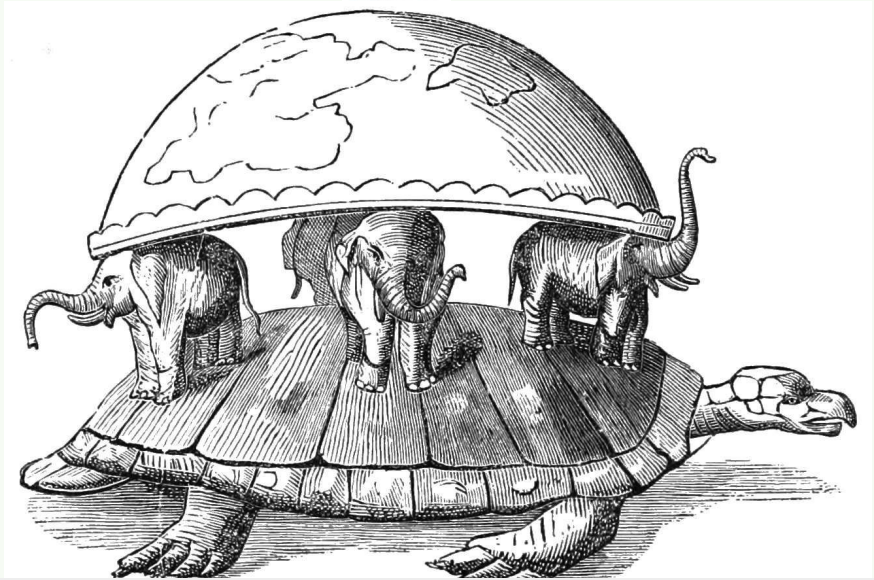


Management

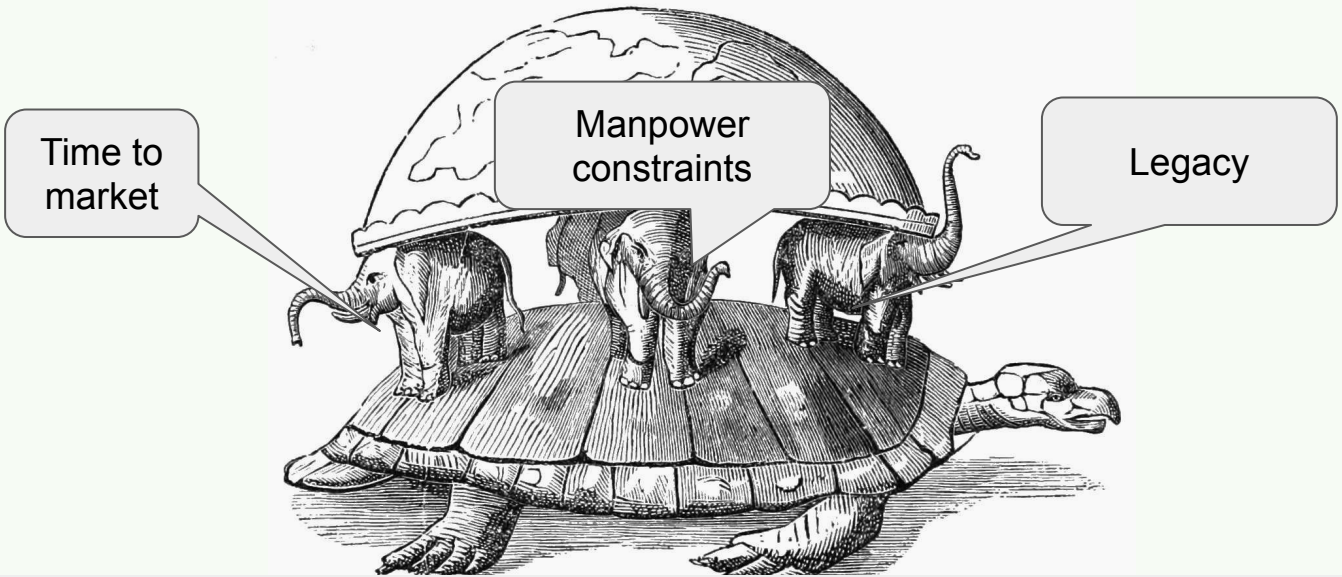


Engineers

Result: new products based on proven technology (low-risk)



But many get stuck in outdated thinking!



But many get stuck in outdated thinking!

In academia you have **freedom** to choose

You work for **yourself**, with supportive team, and great coach

In academia you have **freedom** to choose

You work for **yourself**, with supportive team, and great coach

And even if sometimes things may go wrong…

It is fun when it works!

https://acsl.group

ACSL
Accelerated Computing Systems Lab

TECHNION Israel Institute of Technology

The Andrew and Erna Viterbi Faculty of ELECTRICAL & COMPUTER ENGINEERING

Home | About Us | Research | Publications | Team | Teaching | Undergraduate Projects | News & Events

Welcome to the Accelerated Computing Systems Lab (ACSL)!

We work on a broad range of computer systems projects spanning hardware architecture, compilers, operating systems, security and privacy, high-speed networking.

All our software is open-source and free .

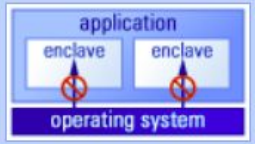
Feel passionate about building secure and fast computer systems of the future?! Check out how to apply!

RESEARCH

all research areas

Accelerator-Centric Operating System

Accelerator-centric Operating System Architecture, OmniX, enables direct interaction between accelerators and I/O devices, for example, files and network sockets for GPU kernels



Confidential Computing

New methods for detection and defense against side-channel attacks on software and hardware for Trusted Execution Environments.

Programmable Networks

New infrastructure for programmable NICs, distributed replication in data-plane using Programmable Switches, neural-network based packet classification (OVS).

GPU computing, Networking, Machine Learning, Distributed Systems

Playground for exploring interesting topics in a search for new ideas