Public Clouds Will Subsume (Most of) HPC

Babak Falsafi
ecocloud.ch

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

an EPFL research center

EUROLAB-4-HPC

Disclaimer:
HPC = Supercomputing

Wiki's page for HPC redirects:

"A **supercomputer** is a computer with a high level of computing performance compared to a general-purpose computer."

"Supercomputers play an important role in the field of computational science."

A Brief History of IT

Mobile Era

Consumer Era

1970s- | 1980s | 1990s | Today+

Mainframes

PC Era

- From computing-centric to data-centric
- Consumer Era: Internet-of-Things in the Cloud

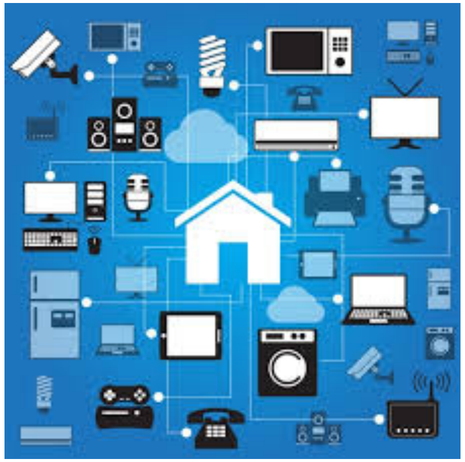
IT (as we know it) is Changing

1. Big Data
   - Scales faster than platforms
2. Silicon Scaling
   - End of conventional scaling
3. Warehouse-Scale Computers
   - Centralization exploits economies of scale
   - Whither future of HPC?

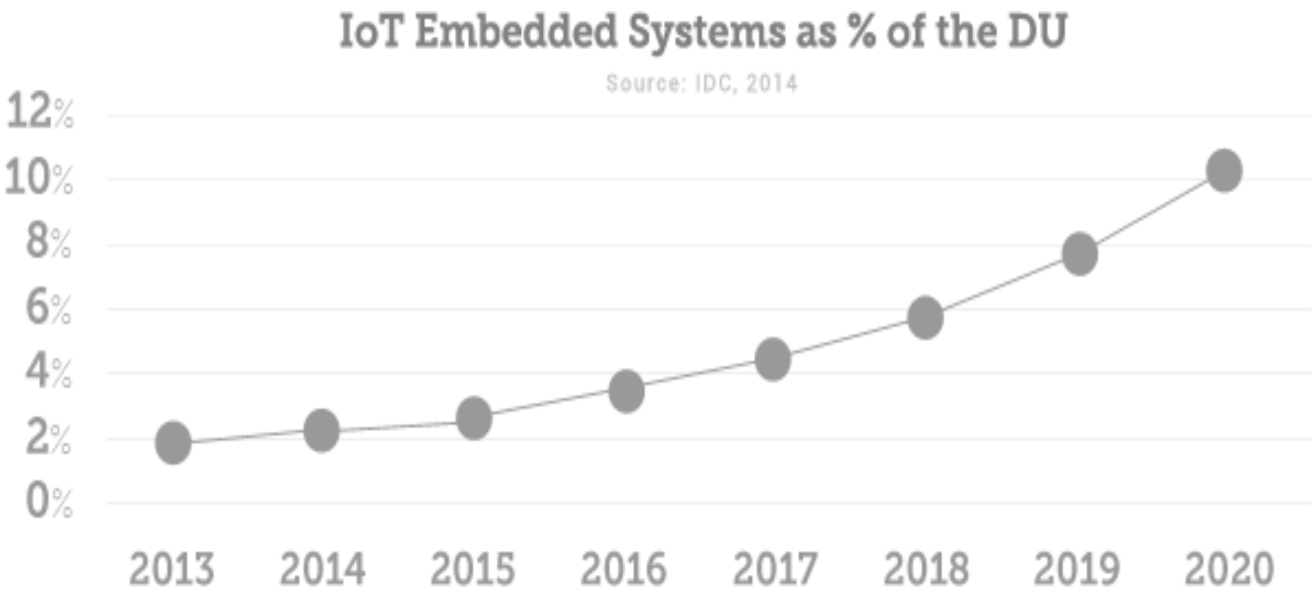
The Future of IT is Data

- Data growth (by 2015) = 100x in ten years [IDC 2012]
  - Population growth = 10% in ten years
- Monetizing data for commerce, health, science, services, ….
- Big Data is shaping IT & pretty much whatever we do!

Internet-of-Things (IoT) Growing Fast Too



20 Billion Connected Devices



4 Zettabytes of Data, 10% of Digital Universe

$7 Trillion
Market Revenue

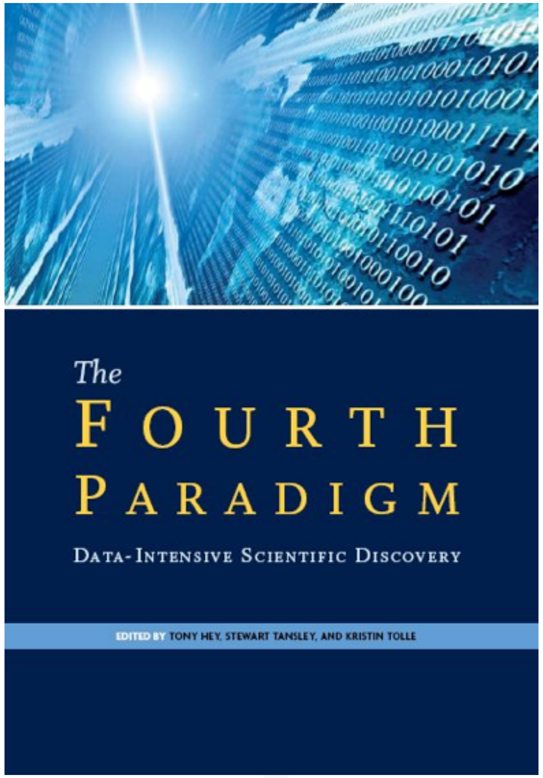
Source: IDC Worldwide and Regional IoT forecast, EMC Digital Universe with Research and Analysis by IDC

Data Shaping All Science & Technology

ecocloud
an EPFL research center

Science entering 4th paradigm

- Analytics using IT on
  - Instrument data
  - Simulation data
  - Sensor data
  - Human data
  - …



Complements theory, empirical science & simulation

Data-centric science key for innovation-based economies!

Scientific Data
[Frontiers in Massive Data Analysis, 2013]

Square Kilometer Array (SKA)

- 100's TB/s

Astronomy

- 100 PB in a decade

Genome Sequencing

- Cost per genome < $1K

Sky Server

- Trillions of particles with TB snapshots

Social Media & Sciences

- 6000 tweets/s

The Seven Computational Giants of Science

- Basic statistics
- Generalized *N*-body problem
- Graph-theoretic computations
- Linear algebraic computations
- Optimization
- Integration
- Alignment problems

Challenges for Data-Centric Science

ecocloud
an EPFL research center

- Massive data sets
- Distributed data sources
- Sampling biases & hete
- Heterogeneo
- Scalab ithms
- er architecture
- a integrity & security
- abling data discovery, integration, sharing
- Visualization

But, which of these are science specific?

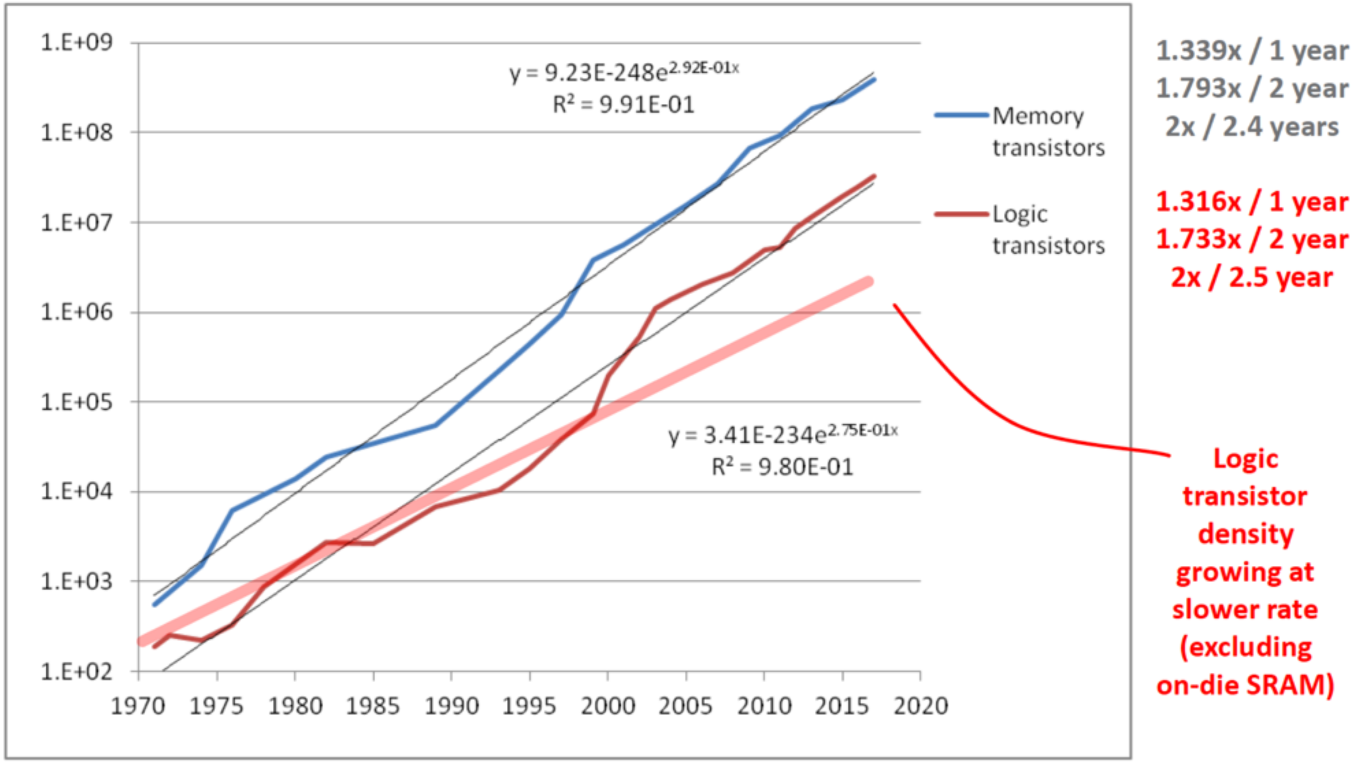
IT (as we know it) is Changing

1. Big Data
   - Scales faster than platforms
2. Silicon Scaling
   - End of conventional scaling
3. Warehouse-Scale Computers
   - Centralization exploits economies of scale
   - Whither future of HPC?

ecocloud
an EPFL research center

DRAM Array and Logic Transistors per square mm

- Suggests using memory wafers more extensively than logic, widening gap
- Logic transistor curve based only on Intel production start dates to maintain harmony of methodologies Dates rounded down to year – Note that this does not show COST



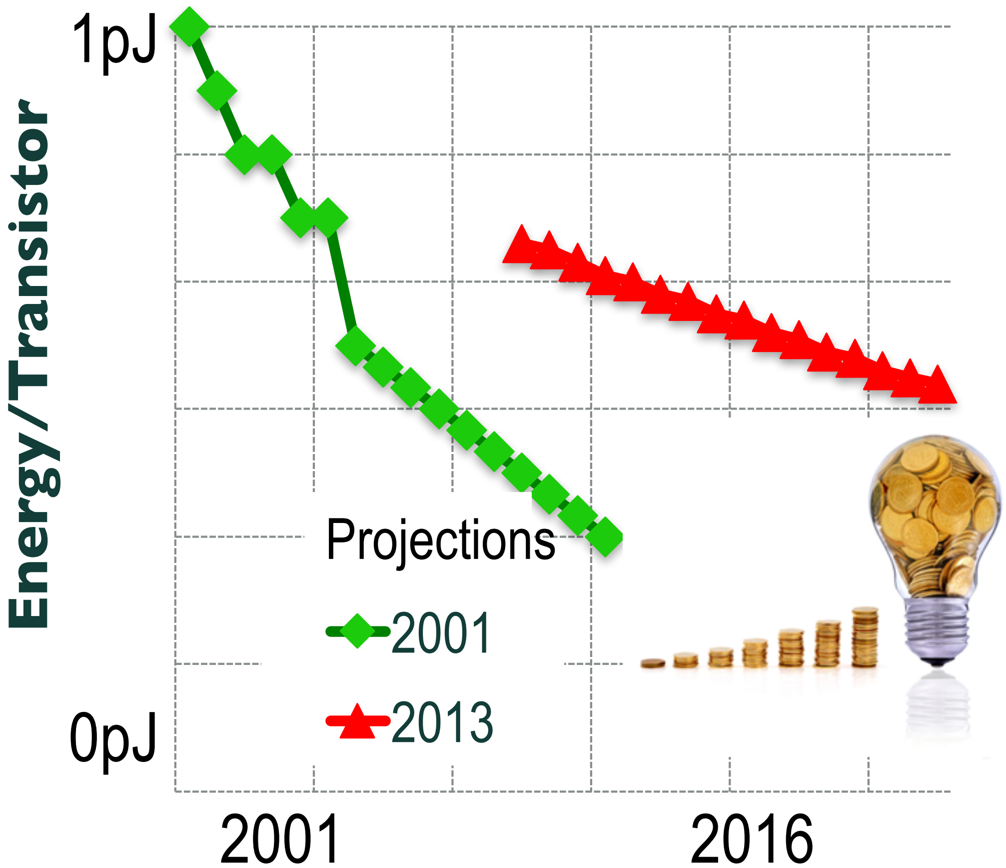
Source: JTPawlowski, Micron

December 8, 2015
Micron

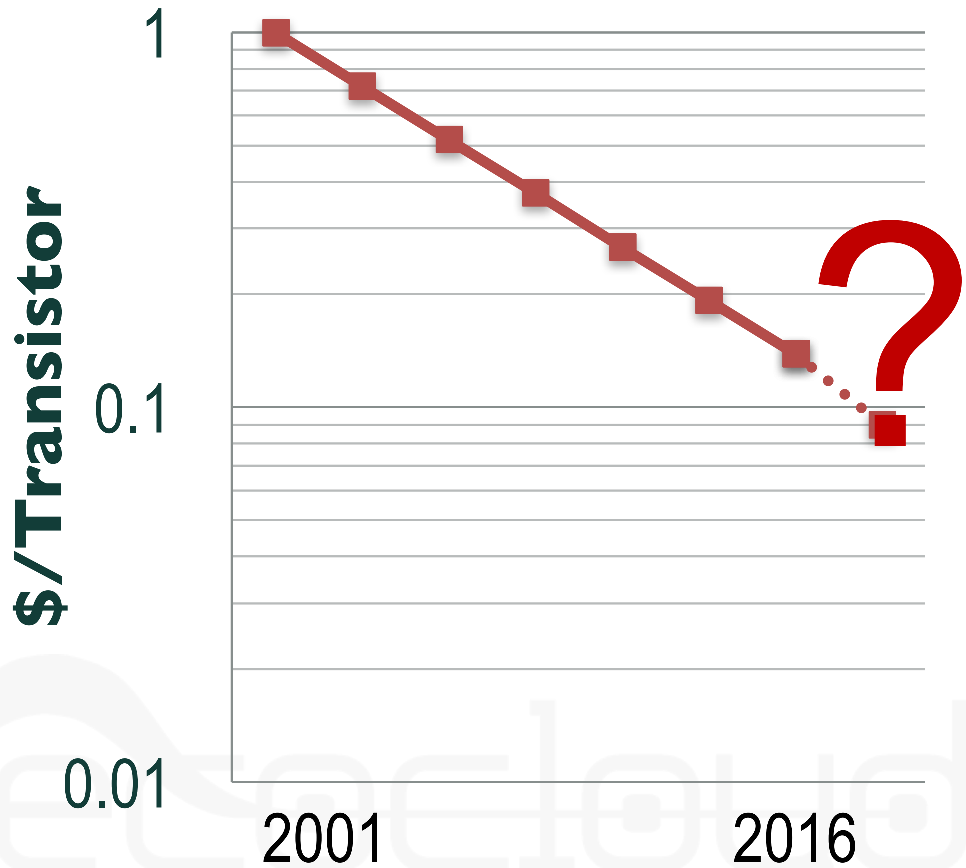
Silicon is running out of steam!

an EPFL research center

Silicon efficiency is dead (long live efficient silicon)



Moore's Law is Dying

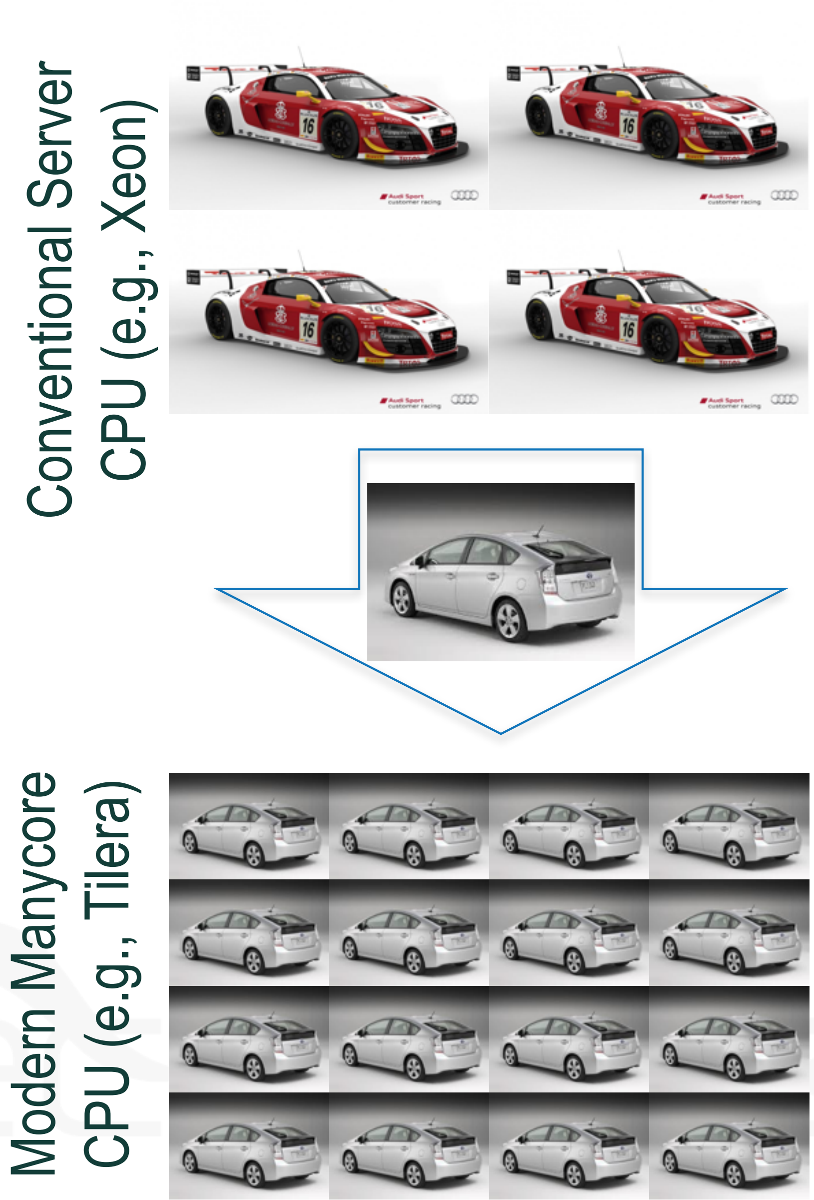


Manycore Accelerators

With voltages leveling:

- Parallelism has emerged as the only silver bullet
- Use simpler cores
  - Prius instead of Audi R8
- Restructure software
- Each core ➔

  fewer joules/op

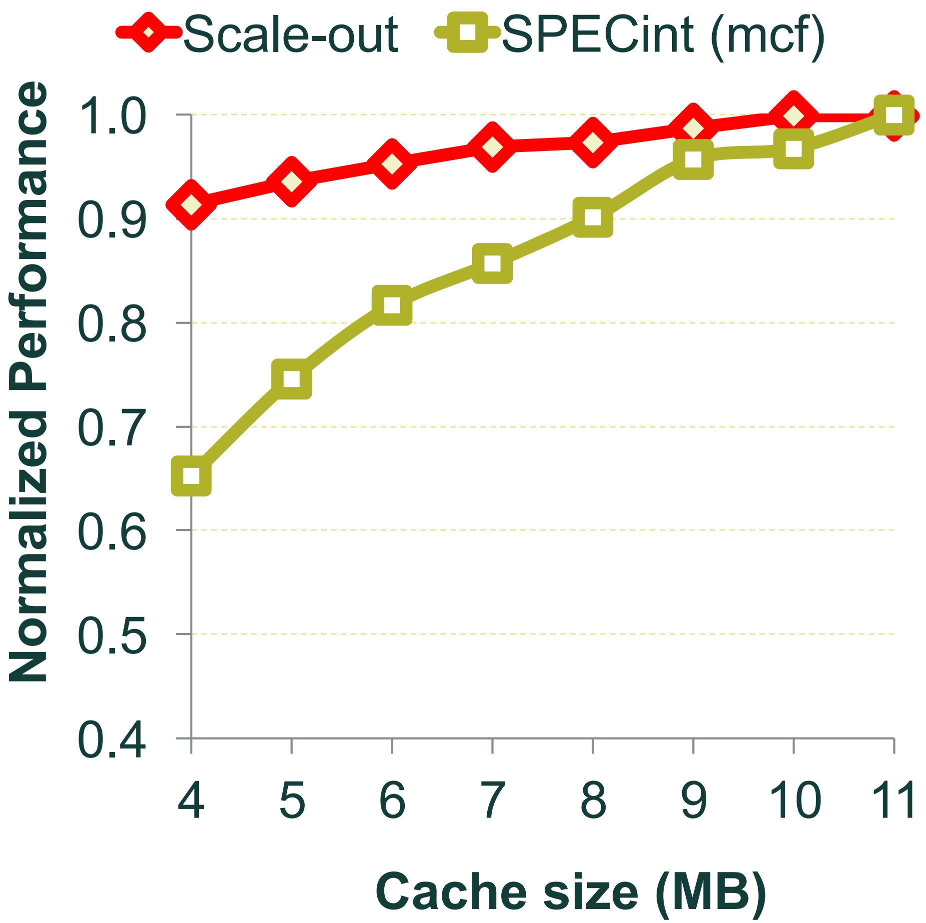


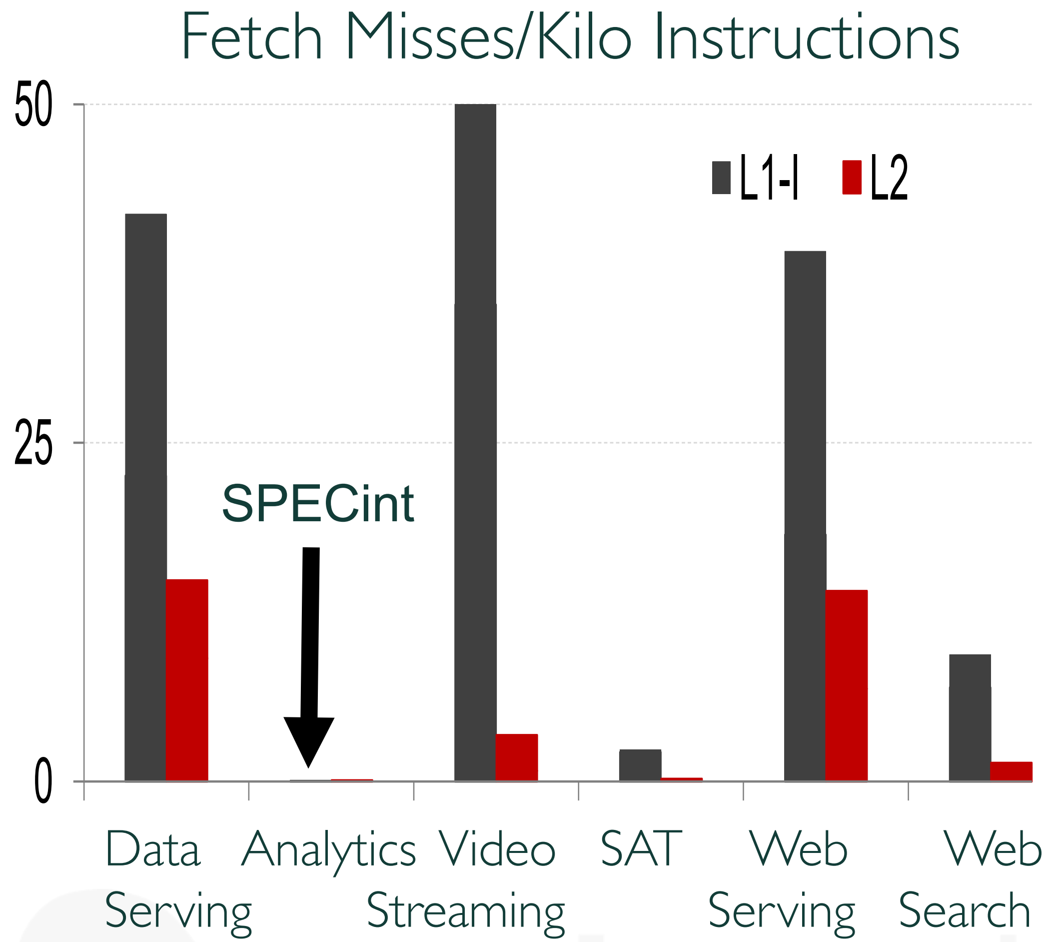
Server Benchmarking with CloudSuite 3.0 (cloudsuite.ch)

- **Data Analytics** — Machine learning
- **Data Caching** — Memcached
- **Data Serving** — Cassandra NoSQL
- **Graph Analytics** — GraphX
- **Media Streaming** — Nginx, HTTP Server
- **Web Serving** — Nginx, PHP server
- **Web Search** — Apache Solr & Nutch
- **In-Memory Analytics** — Recommendation System

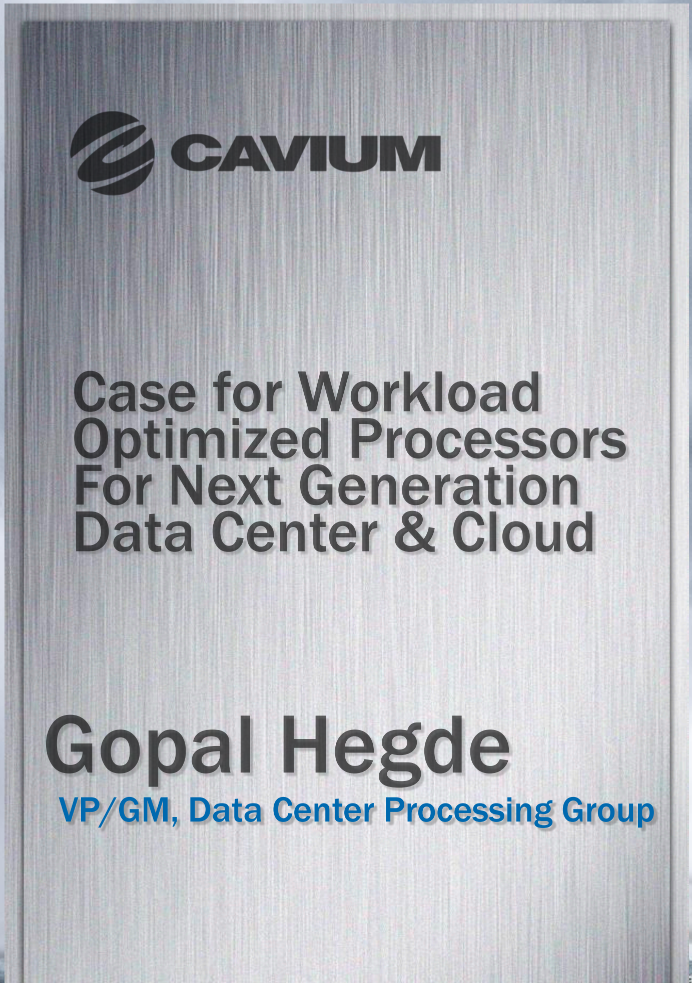
Building block for Google PerfKit, EEMBC Big Data!

CloudSuite Stuck in Memory
[ASPLOS'12]





- On-chip memory overprovisioned
- Instruction supply is bottlenecked



Cavium Thunder X

- Based on SOP @ EPFL
- Designed to serve data
- Optimized code supply
- Trade off SRAM for cores
- Runs stock software
- 10x faster than Xeon for CloudSuite

GPU's Come to Rescue

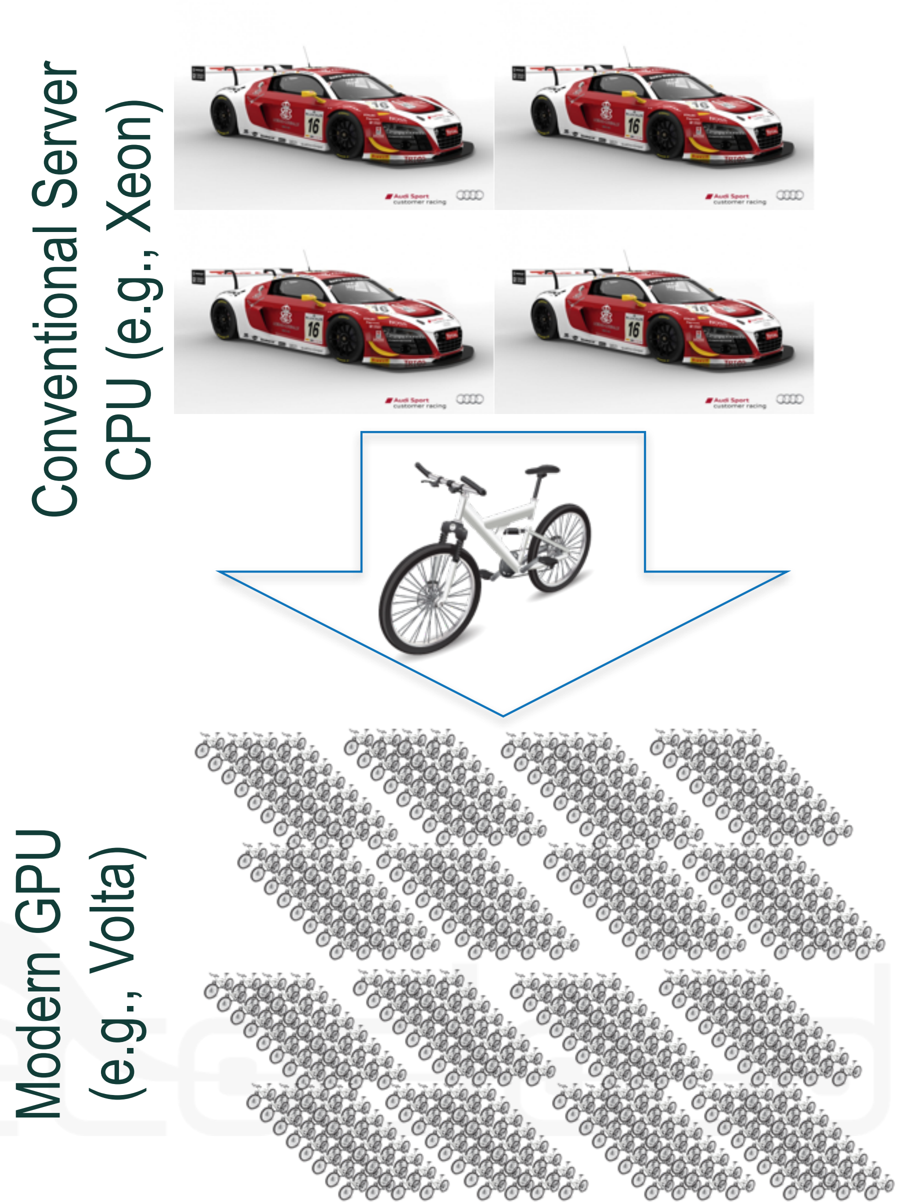
Massively parallel cores

- Data parallel
- Higher memory b/w

Super simple cores

- Shared front end
- 10x slower clocks

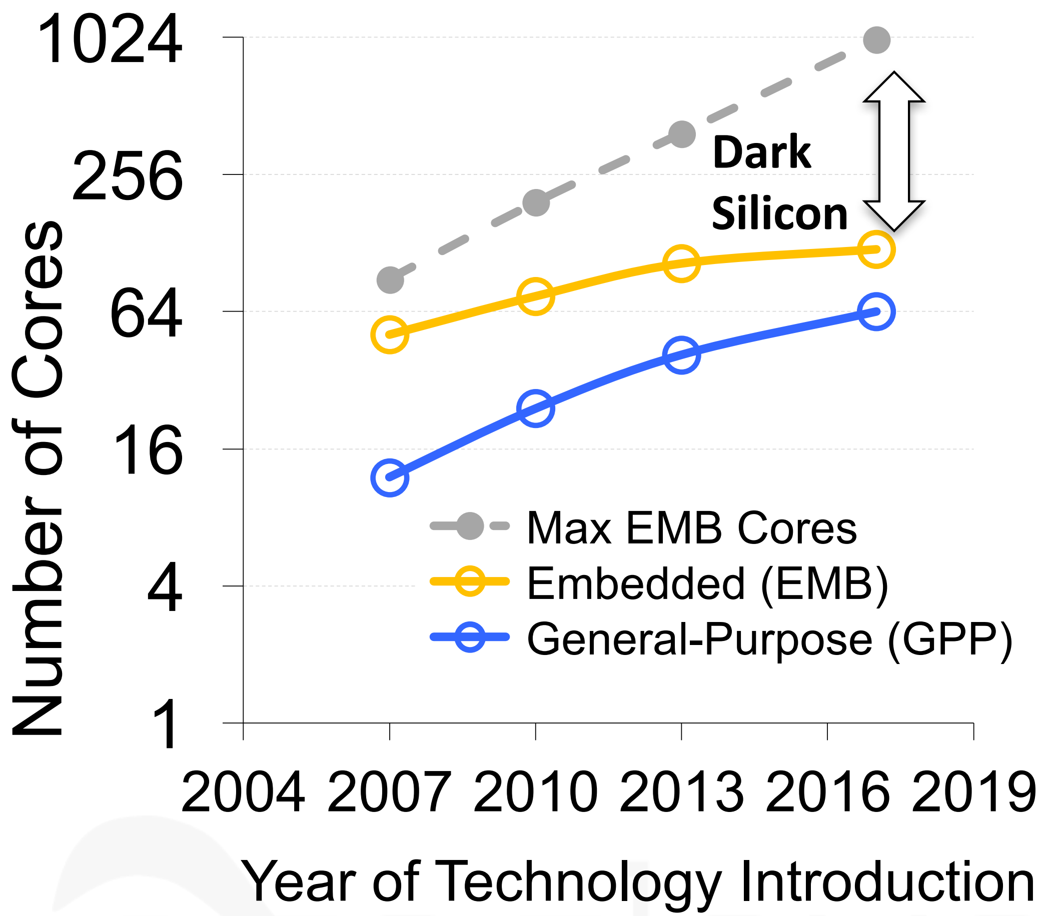
Great for dense parallel computation



Parallelism Alone Can't Help

Look for **ISA** opportunities

- **I**ntegration
  - Use less energy moving
  - Work closer to memory
- **S**pecialization
  - Customize work
  - Less work/computation
- **A**pproximation
  - Adjust precision



Hardavellas et. al.,
"Toward Dark Silicon in Servers",
IEEE Micro, 2011

Custom Computing

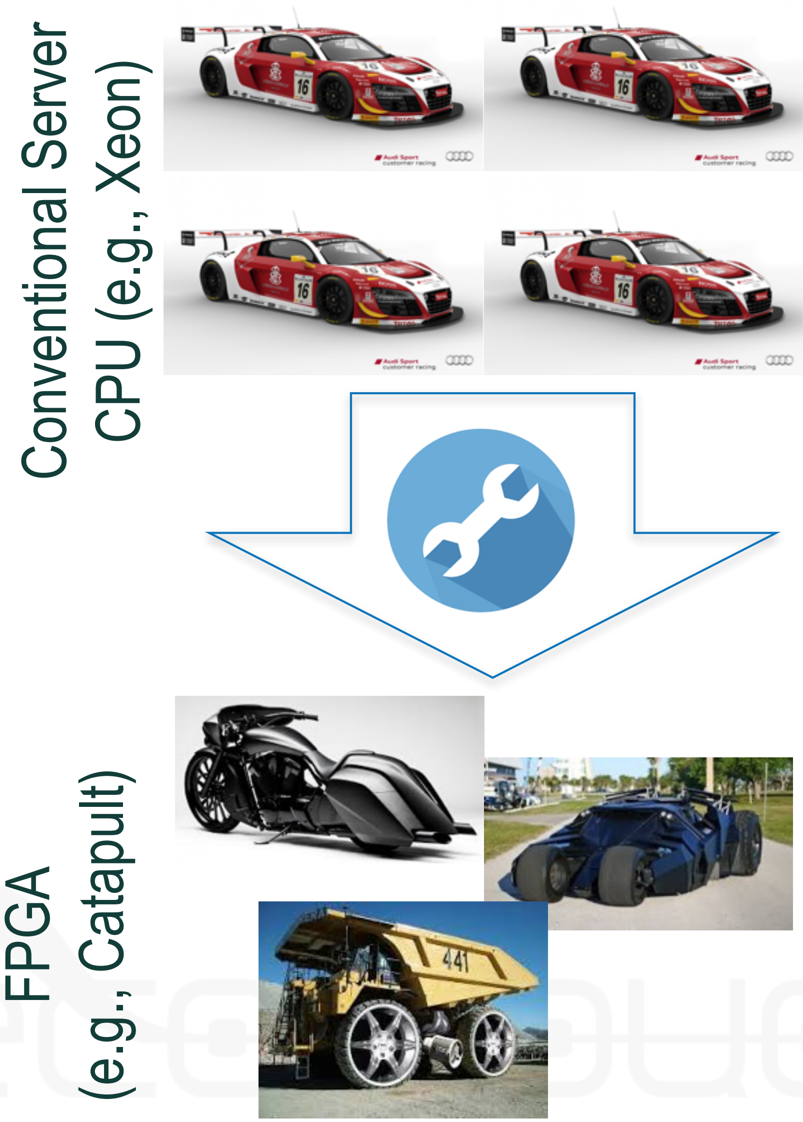
[FPGA's vs. GPU's in Data centers, IEEE Micro'17]

ecocloud
an EPFL research center

Reconfigurable
- Best for spatial computing
- Not caching/reuse

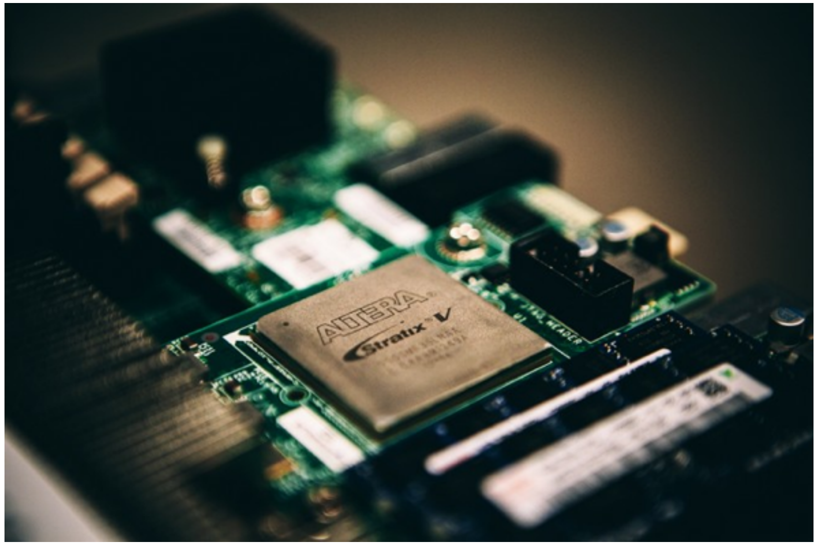
Parallel, dataflow
- 10x slower clocks
- Better for sparse arithmetic

Both Microsoft & Amazon



FPGA's in Servers

ecocloud
an EPFL research center



Microsoft Unveils Catapult to Accelerate Bing!
[EcoCloud Annual Event, June 5th, 2014]

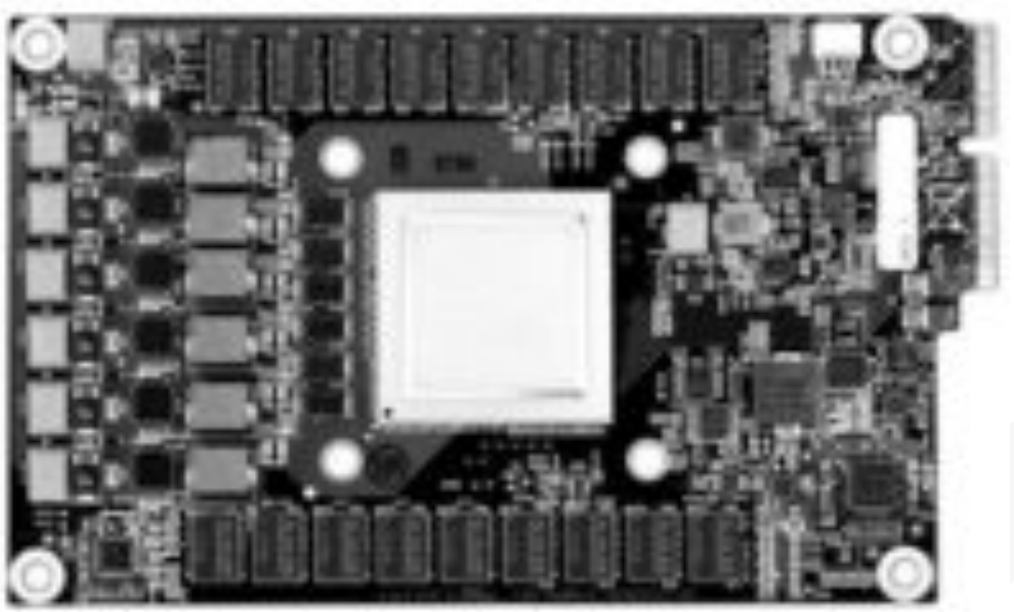
Latest version [MICRO'16]:

- High-end Altera FPGAs
- One FPGA per blade
- Sits on the network
- Backend connected to CPU/NI
- Originally to accelerate Bing
- Now deployed in Azure

Google's TPU

Custom array of arithmetic units:

- Linear algebra for ML/NN
- Currently memory bound
- 10x over GPU
- ML as a service



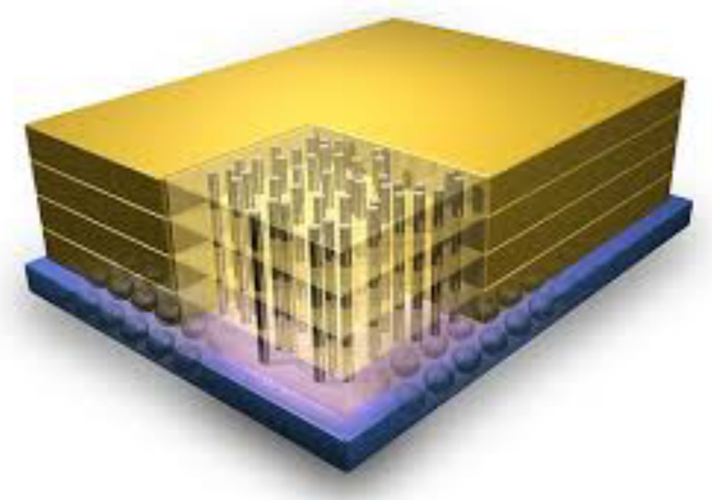
Near-Memory Computing
[IEEE Micro'16]

Why in-memory?

- Minimize data movement & energy
- Leverage DRAM's massive internal BW

Basic data services:

- Scan, Join, GroupBy, Filter
- Best for sequential access
- Accelerators must co-exist with conventional memory semantics



Opportunities for algorithm/hardware co-design

Near-Memory Commandments
[IEEE Micro issue on Big Data'16]

Not (CPU) business as usual

1. DRAM favors sequential vs. random access
   - CPU's leverage reuse & locality in cache hierarchy
2. DRAM favors wide slow cores vs. many fast cores
   - Both data and thread-level parallelism to match DRAM b/w
3. Memory must maintain semantics relative to CPU
   - Shared memory + coherence between near-memory & CPU

Need data-parallel streaming algorithms co-designed with HW!

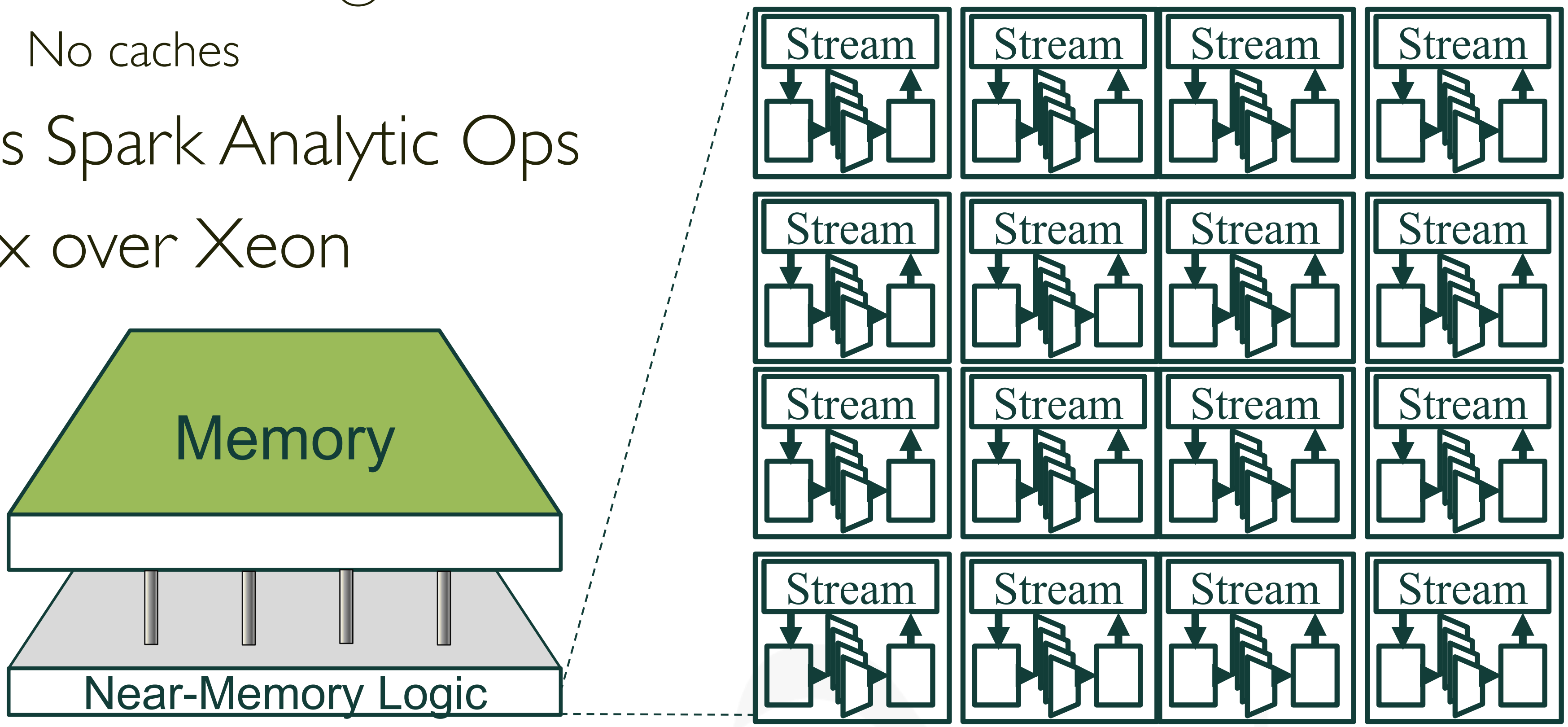
The Mondrian Data Engine [ISCA'17]

SIMD cores + data streaming

- Streams multiple sequential streams
- 1024-bit SIMD @ 1 GHz
- No caches

Runs Spark Analytic Ops

100x over Xeon



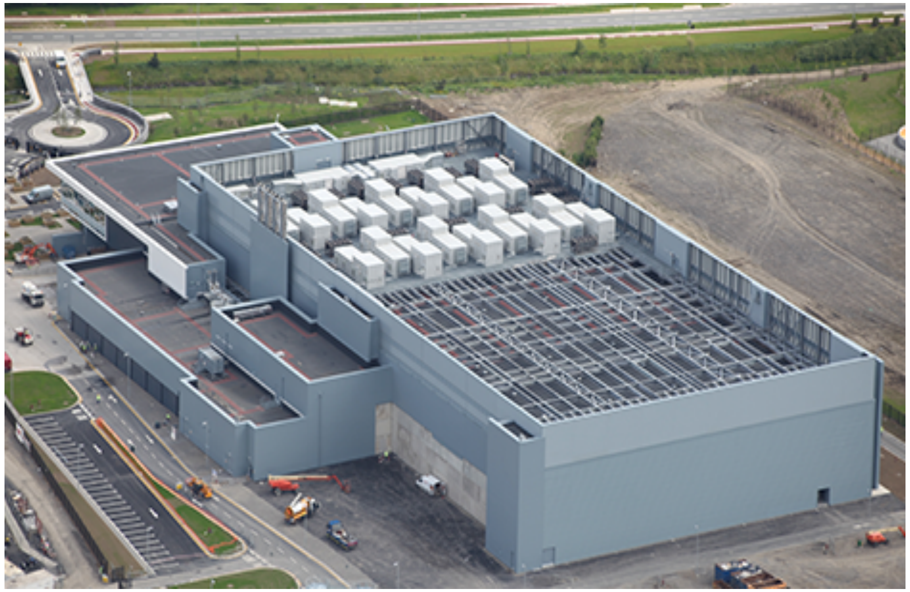
Algorithm/hardware co-design maximize near-memory performance

IT (as we know it) is Changing

1. Big Data
   - Scales faster than platforms
2. Silicon Scaling
   - End of conventional scaling
3. Warehouse-Scale Computers
   - Centralization exploits economies of scale
   - Whither future of HPC?

Modern Datacenters are Warehouse-Scale Computers

- Millions of interconnected home-brewed servers
- Centralization helps exploit economies of scale
- Network fabric provides micro-second connectivity
- At physical limits
- Need sources for
  - Electricity
  - Network
  - Cooling



20MW, 20x Football Field
$3 billion

Warning! Datacenters are not Supercomputers

- Run heterogeneous data services at massive scale
- Driven for commercial use
- Fundamentally different design, operation, reliability, TCO
  - Density 10-25KW/rack as compared to 25-90KW/rack
  - Tier 3 (~2 hrs/downtime) vs. Tier 1 (upto 1 day/downtime)
  - ……and lots more

Datacenters are the IT utility plants of the future

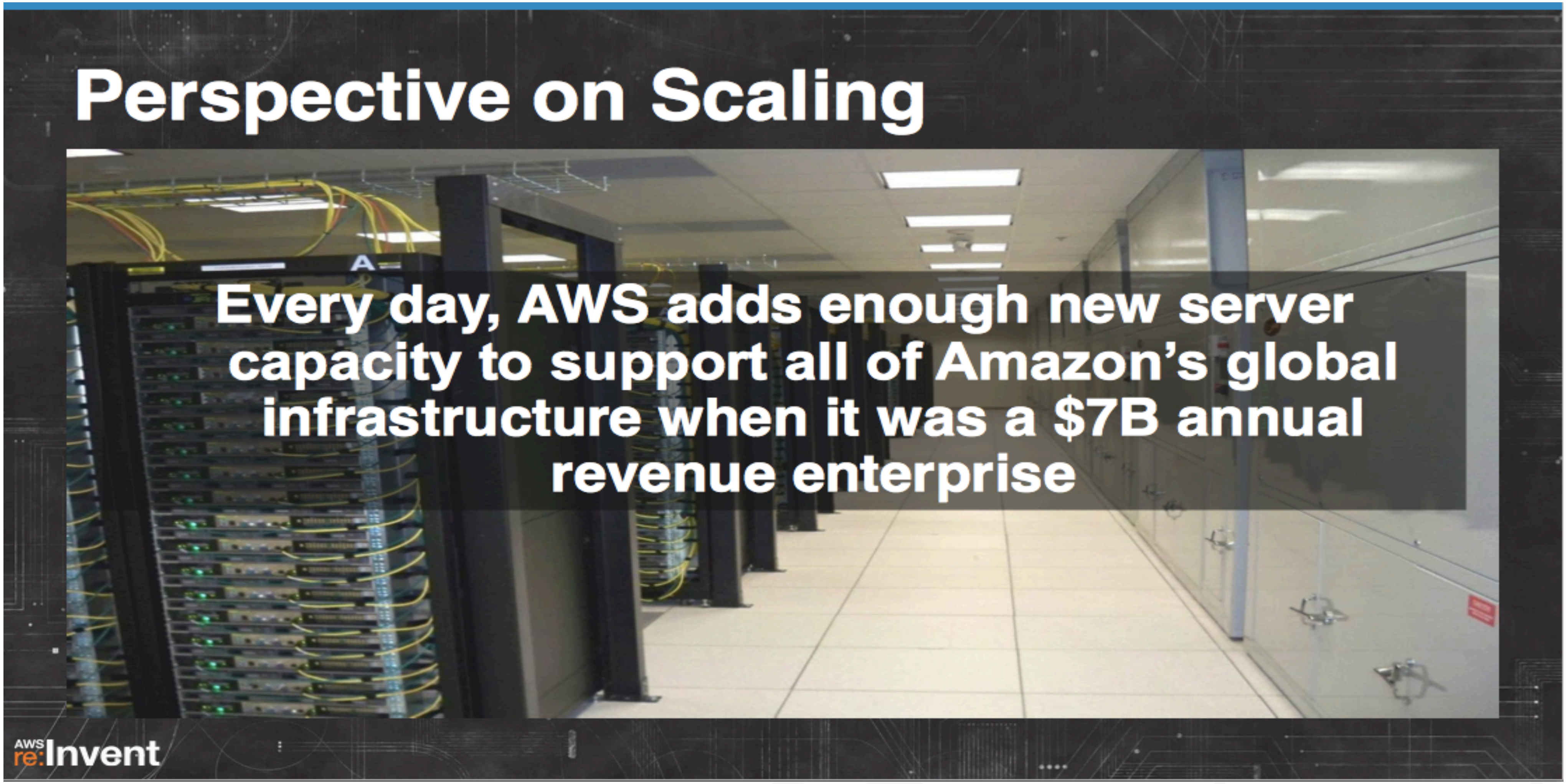
Supercomputing (HPC)

≠

Cloud Computing

Source: James Hamilton, 2014
mvdirona.com/jrh/TalksAndPapers/JamesHamilton_Reinvent20131115.pdf

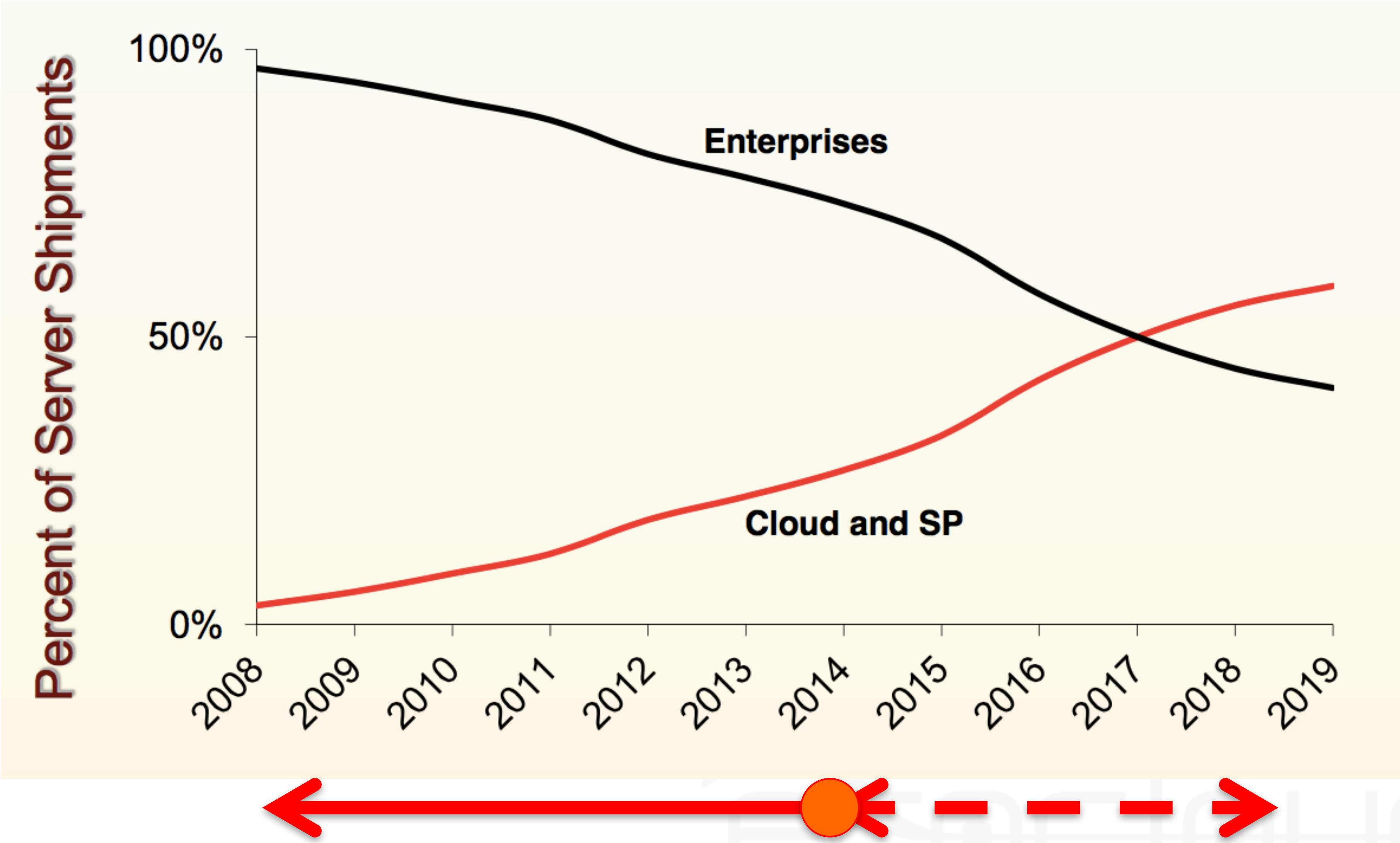
ecocloud
an EPFL research center



Daily IT growth in 2014 = All of AWS in 2004!

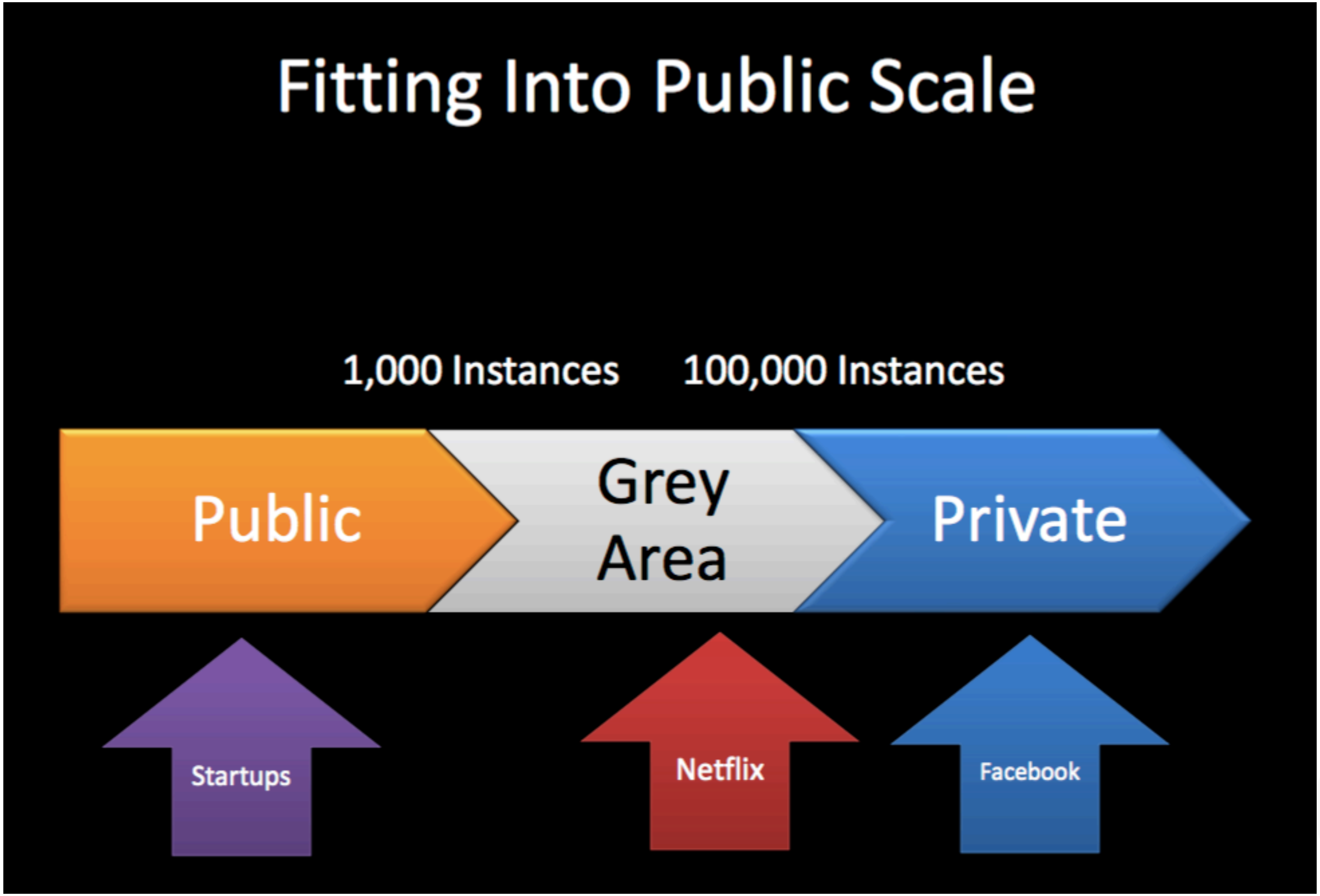
RIP Retail

Cloud Taking Over Enterprise



Source: Dell 'Oro 2Q15

Build if You Have Customers



Source: Adrian Cockcroft, NetflixOSS, 2013

Improved Services @ Scale

ecocloud
an EPFL research center

Enhanced Customer View

BigData Exploration

Security Intelligence

Operation Analysis

Warehouse Optimization

IBM

Deep Learning

ecocloud
an EPFL research center

Deep learning technology enabled speech-to-speech translation

The New York Times

Scientists See Promise in Deep-Learning Programs

November 23, 2012

Rick Rashid in **Tianjin, China**, October, 25, 2012

A voice recognition program translated a speech given by Richard F. Rashid, Microsoft's top scientist, into Mandarin Chinese.

Microsoft

Applications Abound

ecocloud
an EPFL research center

JUNO

airbnb

nest

NETFLIX

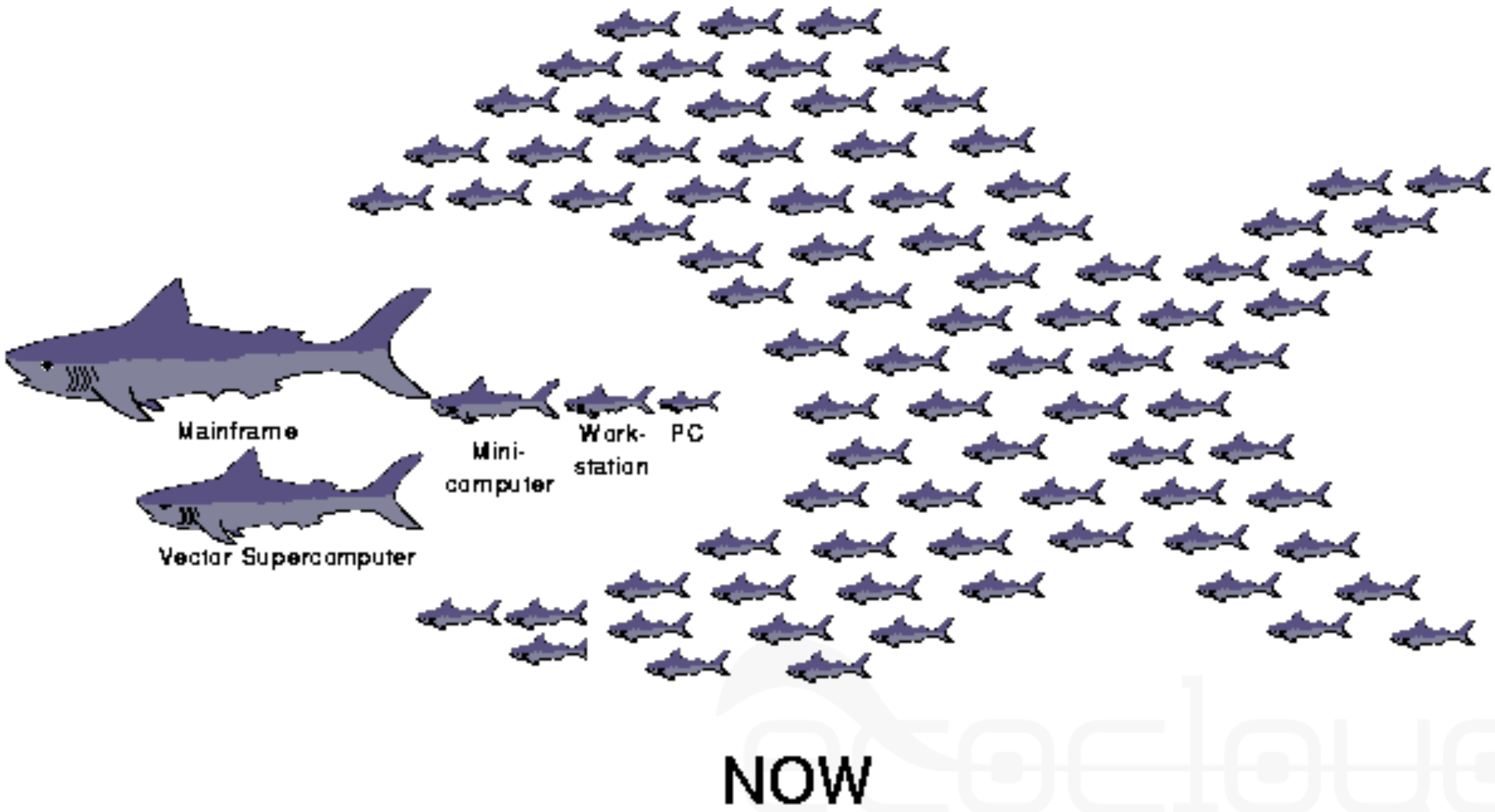
wealthfront

What does this all mean for HPC?

Requirements for HPC

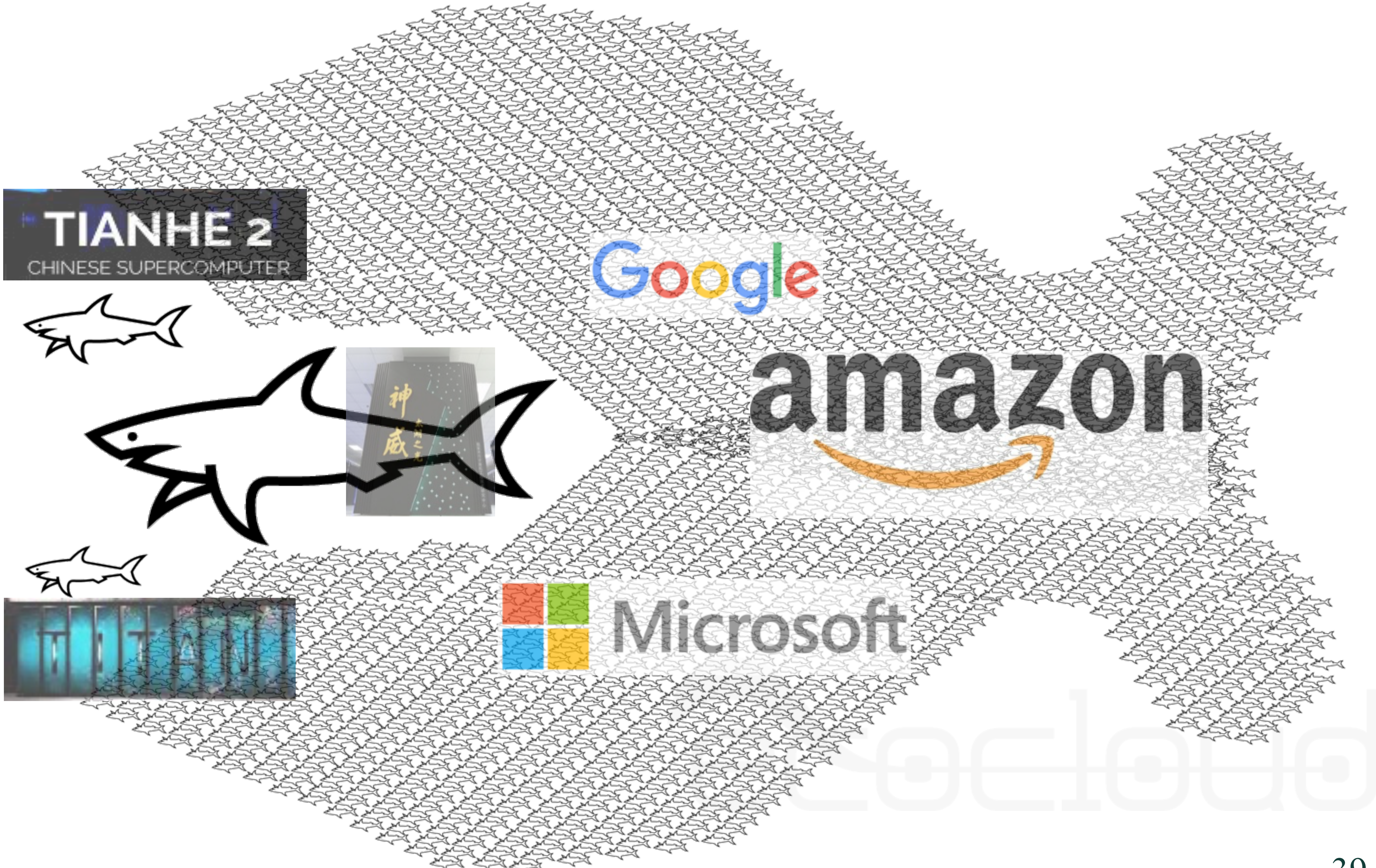
- Massively Parallel
- Accelerators
- High bandwidth/low-latency fabrics
- Massive data management
- Fault tolerance
- Security/privacy
- Lower cost/accessibility

Flashback to 90's: Network of Workstations



Welcome to 2017!

ecocloud
an EPFL research center



Public Clouds push HPC Infrastructure to Niche

- TCO/server drops 80% from 1000 to 100000 units
- Provide massive resources at low cost
- First to adopt technologies for massive computing
- Conventional HPC will no longer afford building infrastructure (short of defense, or extreme niche)
- E.g., Stanford Genomics moved to Azure in 2016

Invest in technologies to embed HPC into the cloud

Moving Forward

- Software Stacks, Tools & API for Data Science
- Support for Virtualization
- Scientific Data Management Technologies
- Topology-agnostic Simulation
- Tool chains from DSL's to Platforms for HPC
  - Need to handle both heterogeneous logic & memory
- HPC-specific IT (e.g., visualization)
- Identify niche domains for infrastructure innovation

Summary

- HPC is ever dependent on scalable IT
- Future IT & Science will be data-centric
- Clouds are the only path forward
  - Massive data analytics & management
  - Benefit from economies of scale
- Challenges
  - Build HPC infrastructure for niche
  - Invest in embedding HPC in Cloud

Thank You!

ecocloud
an EPFL research center

For more information please visit us at

ecocloud.ch

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

EUROLAB-4-HPC