Codesign for Energy Efficient Computing

A few thoughts about post Exascale Supercomputing

John Shalf

Department Head for Computer Science

Lawrence Berkeley National Laboratory

ISC EEHPC Workshop 2025

Hamburg, Germany

The fundamental problem with Wires (and data movement):

Moore's Law undermined by data movement (when smaller is not better)

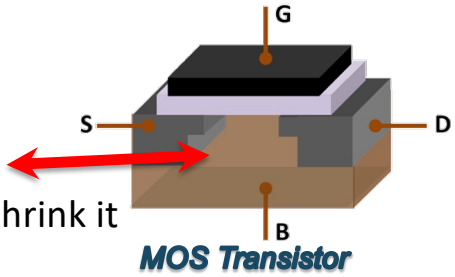
Energy Efficiency of copper wire:

- **Power = Frequency* Length / cross-section-area**

- Wire efficiency *does not improve* as feature size shrinks

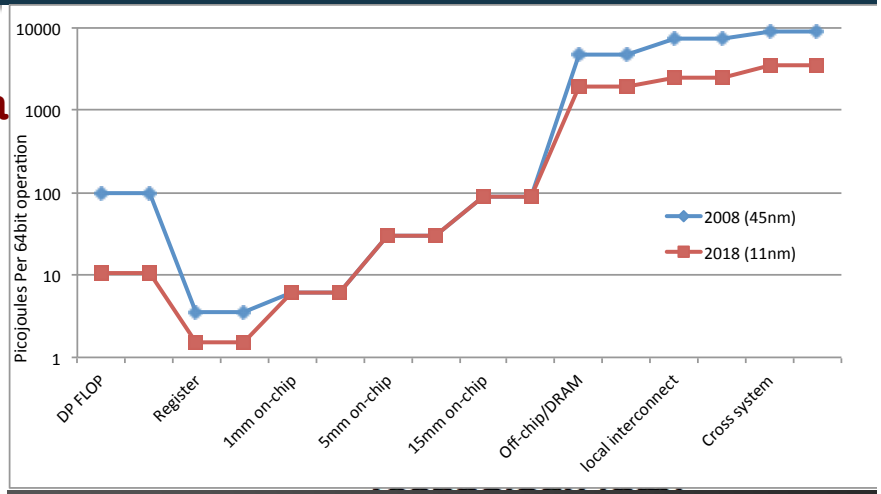
Energy Efficiency of a Transistor:

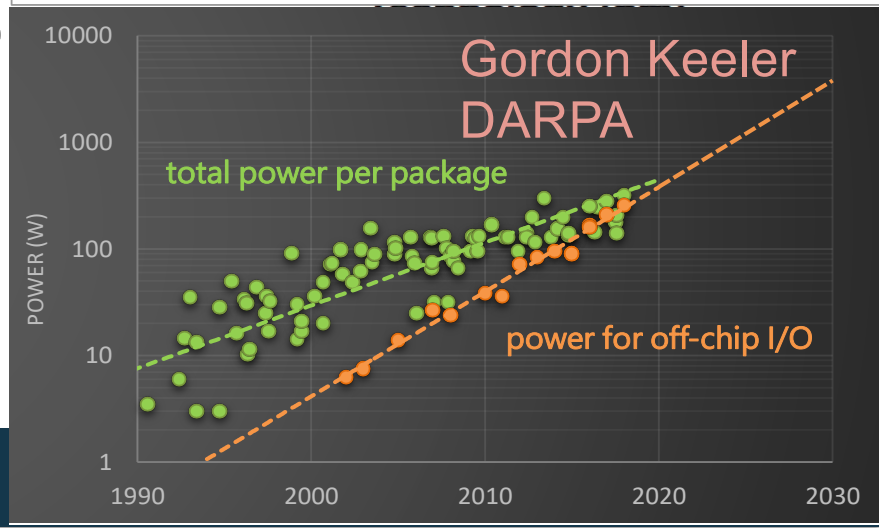
- Power = V^2 * frequency * Capacitance
- Capacitance ~= Area of Transistor
- Transistor efficiency improves as you shrink it



Net result is that moving data on wires is starting to cost more energy than computing on said data

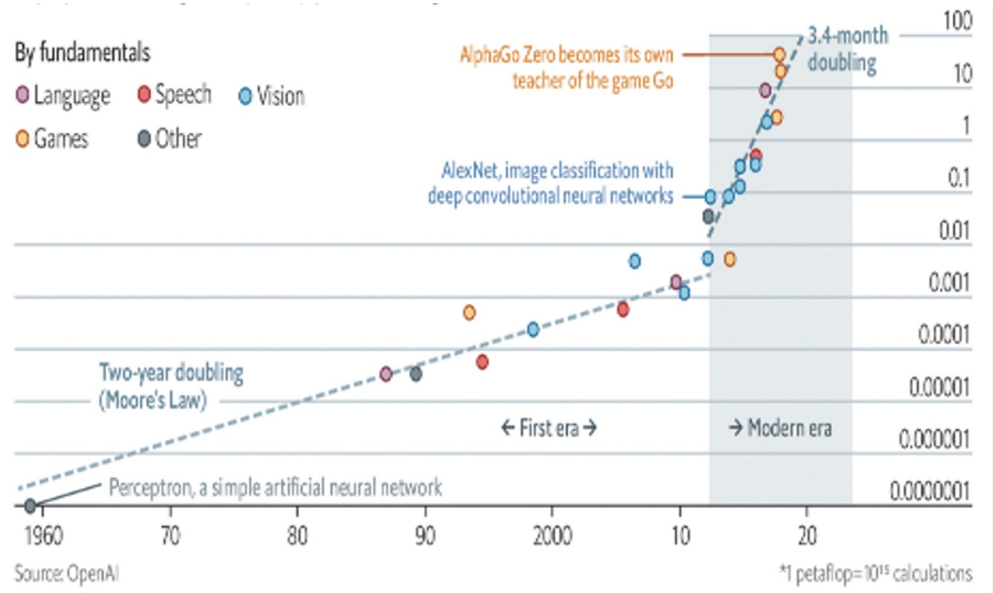
(see also Silicon Photonics)



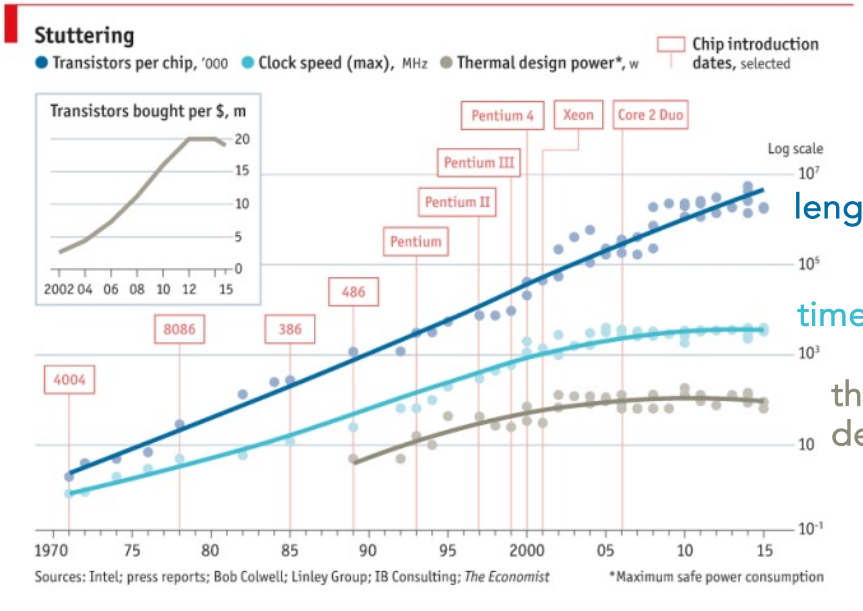


Explosion of Computing Demand: Driving Need for Hyper-exponential Improvement in Performance, Energy Efficiency and Integration

Demand for Computing



Supply for Computing



NVIDIAnomics

- **NVIDIA A100 has a 250W TDP**
- **NVIDIA H100 SXM consumes 700W TDP**
- **Next Generation B100 is projected to consume 1400-2000W TDP! (>100% increase)**
- **Street price $20k-$30k**
  - That's an 800% profit
  - Prices lower @ volume
- **And still supply cannot keep up with demand**

Google Sustainability Page in Late 2023

Google Sustainability | Empowering individuals | Working together | Operating sustainably | Reports

Overview | Net-zero carbon | Water stewardship | Circular economy | Nature & biodiversity | Stories

Net-zero carbon

What a difference a year makes!

Google's AI ambitions threaten carbon neutrality pledge

Data center expansion leads to 48% Increase in emissions since 2019

Omer Kabir 09:53, 04.07.24

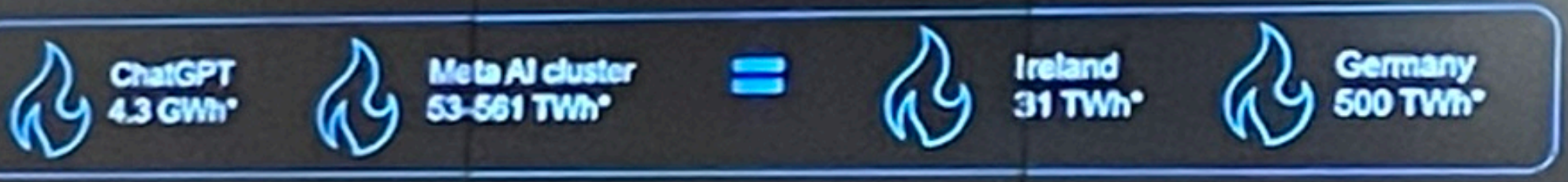
TAGS: Climate Change Data Center Google Emissions

Listen to this article now 4 min listen

00:00 10 10 1.0x -03:48

ENERGY EFFICIENCY – UNSUSTAINABLE

AI Training Energy footprint on par with entire industrial nations

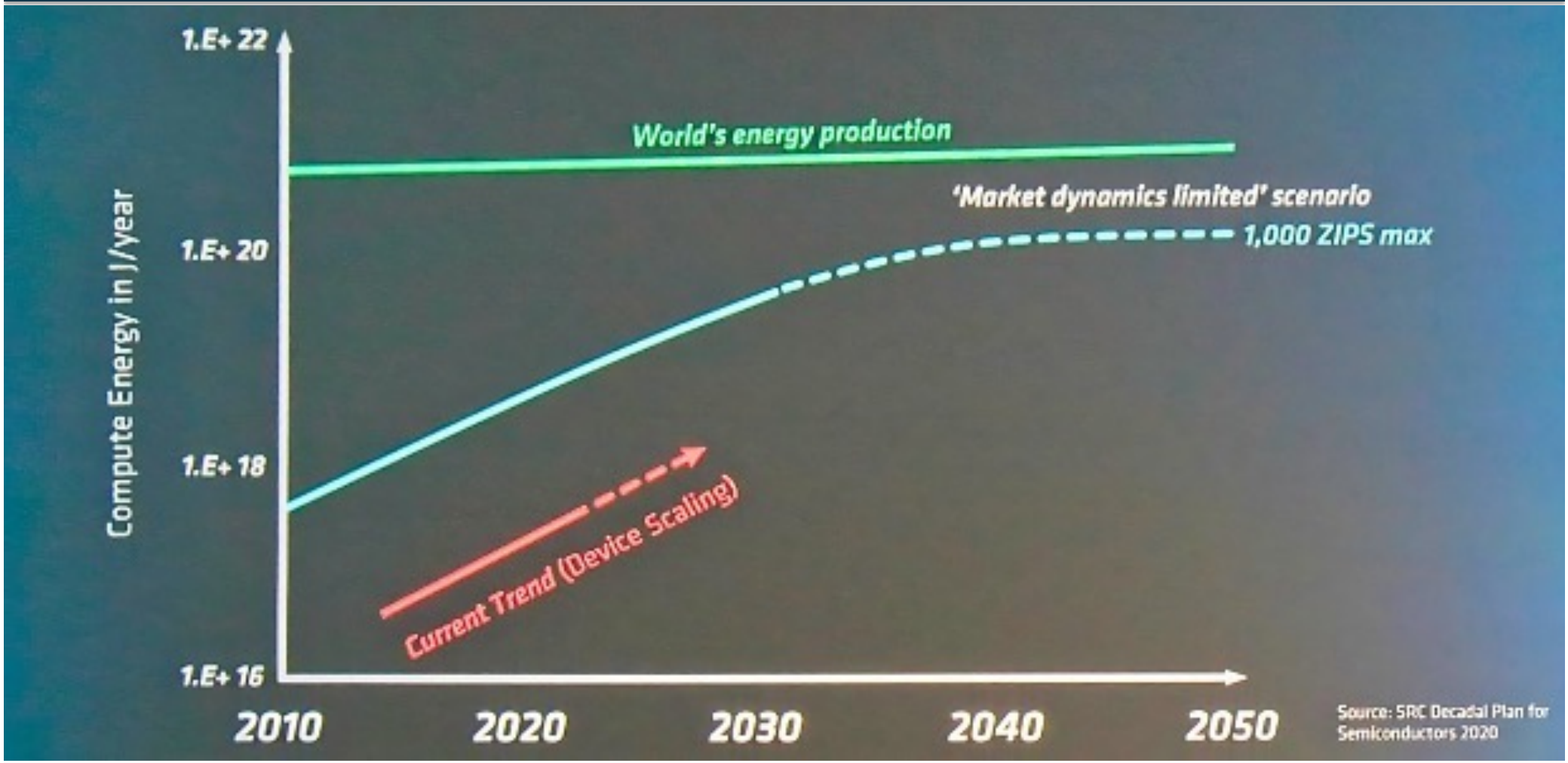


Performance Per Watt: A Critical Metric in the Age of AI

The carbon footprint of ChatGPT, An estimate of the carbon emissions... | by Chris Pelosin | Dec, 2022 | Medium
https://www.semianalysis.com/p/the-inconvenient-truth-ai-training
*Germany - Energy consumption in Germany (worldbank.org), April 2024
*Ireland - Energy consumption in Ireland (worldbank.org), Apr 9 2024
**Yu Wang, Tsinghua University, Feb 2018 https://www.nextplatform.com/blog/2022/05/24/ai-is-harming-our-planet/

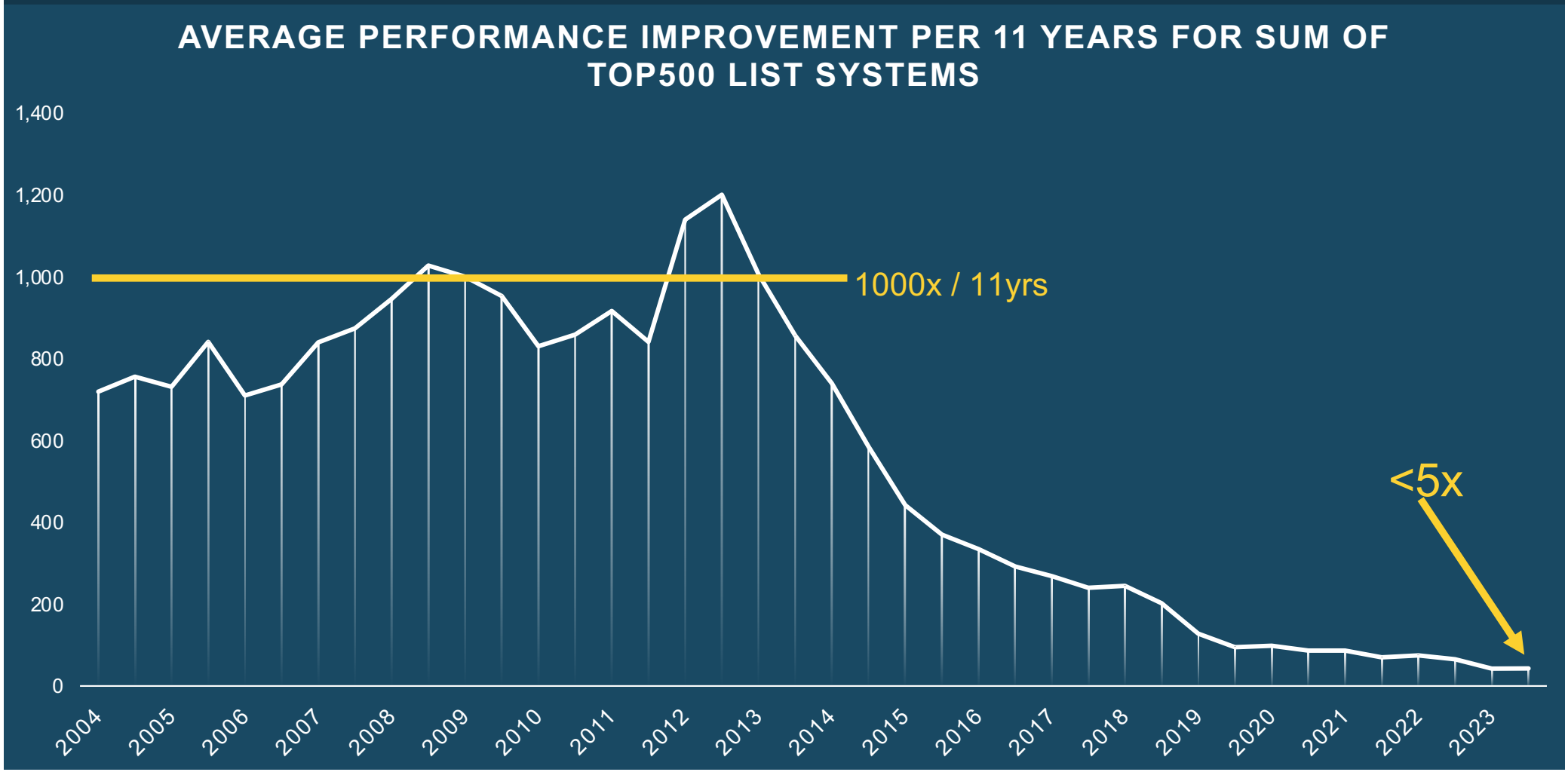
AMD

AI Energy Consumption On Pace to Surpass Supply



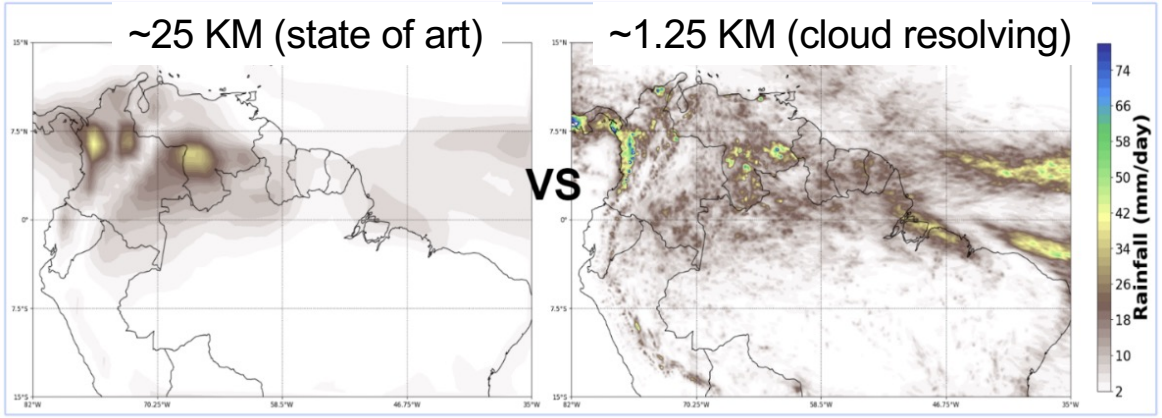
This is HPCs future if we continue business as usual!

… and scale alone is just power and capital cost…

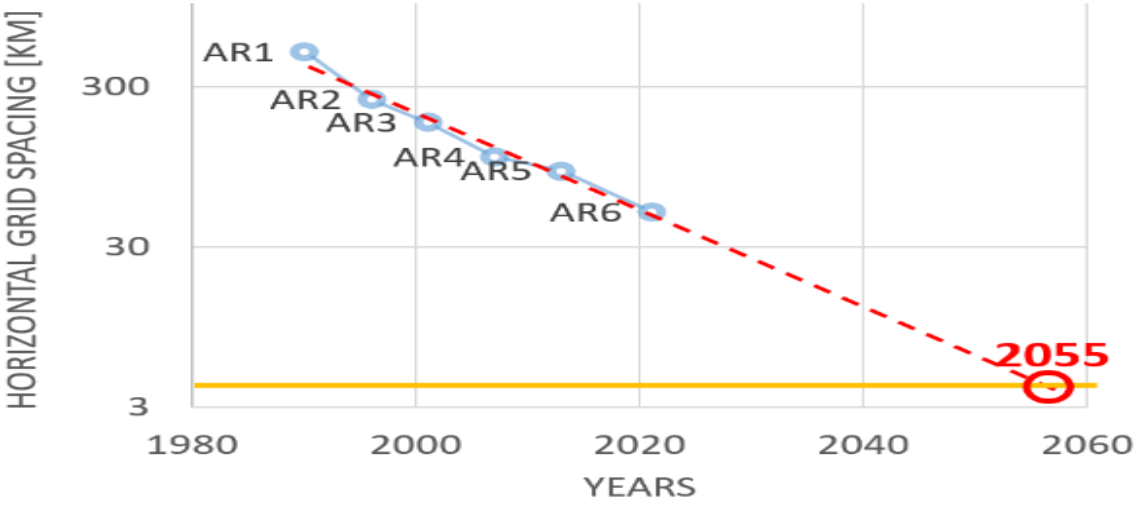


Example: Kilometer Scale Climate Modeling

Warning: ***Power alone is not a scientific imperative***



Earth System Models



Landmark 3.5KM Simulation on Frontier (exascale) achieved 1.5 simulated years per day performance.

At that rate, for an ensemble calculation it would take ~20 years of dedicated computing to answer important policy questions necessary to achieve 2055 goals.

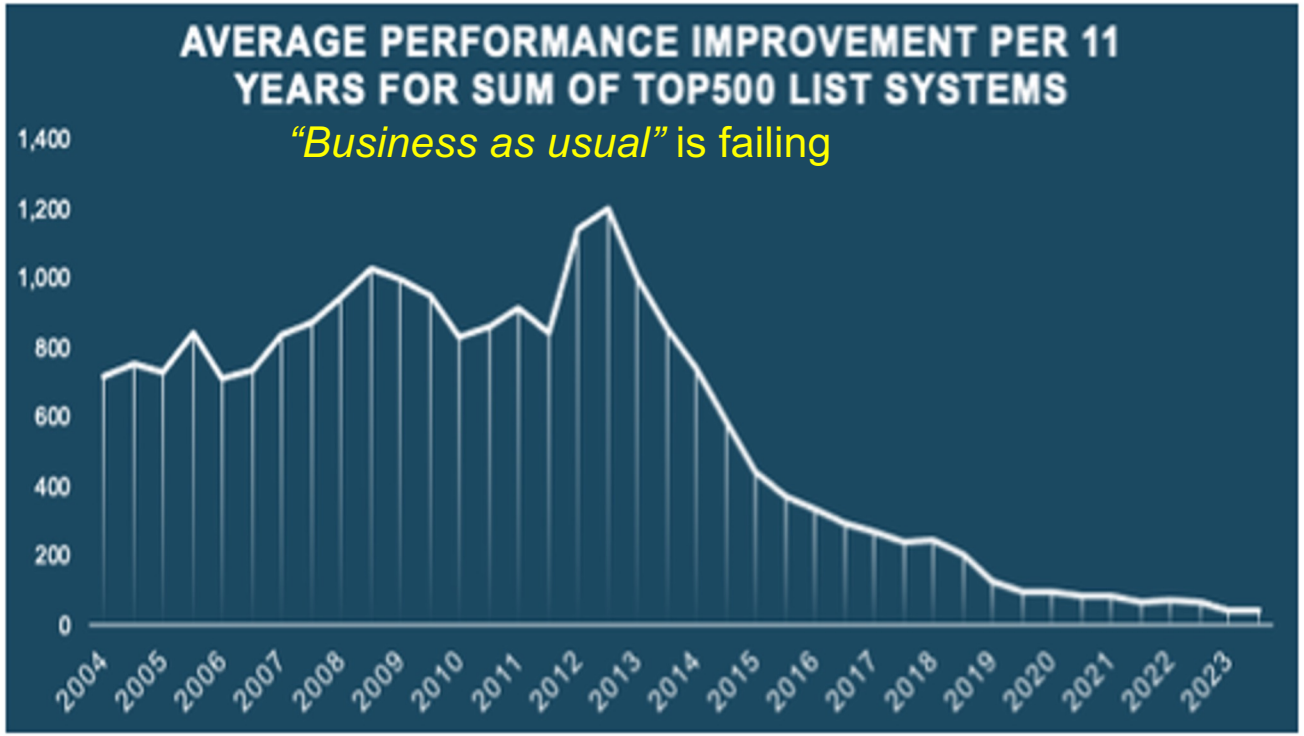
(and that is just for one policy scenario!!!)

Even if we wait for HPC performance improvements, projected 1km modeling goal will be achievable in 2055 at the current rate of progress

This is NOT an acceptable future when there are important scientific imperatives that have ***global societal consequences…***

Global Climate Computer Summit
https://climatecomputer.ncsa.illinois.edu/

Algorithm-Driven Codesign of Specialized Architectures for Energy-Efficient HPC



NASEM study on post-Exascale computing *"We must expand (and create where necessary) integrated teams that identify the key algorithmic and data access motifs in its applications and begin collaborative ab-initio hardware development of supporting accelerators,... a first principles approach that considers alternative mathematical models to account for the limitations of weak scaling."*

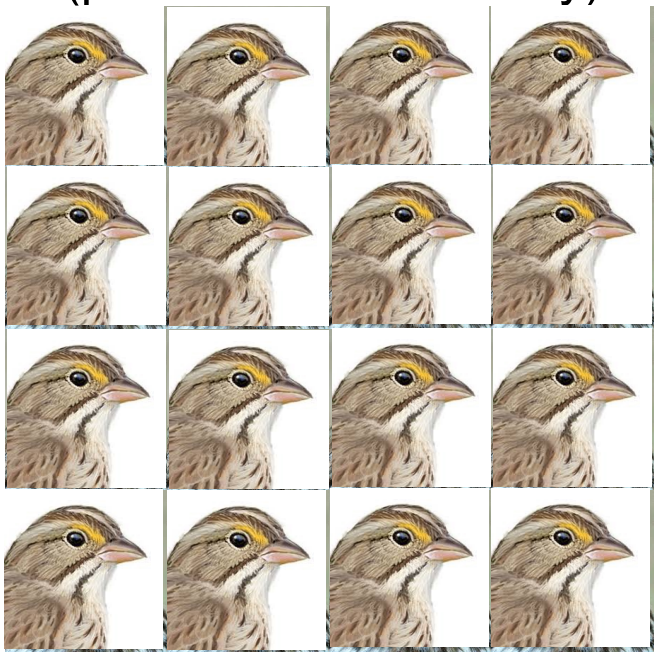
This is a call for co-design at a much deeper level than we are currently realizing

Specialization:
Natures way of Extracting More Performance in Resource Limited Environment

Powerful General Purpose

Xeon, Power

Many Lighter Weight
(post-Dennard scarcity)



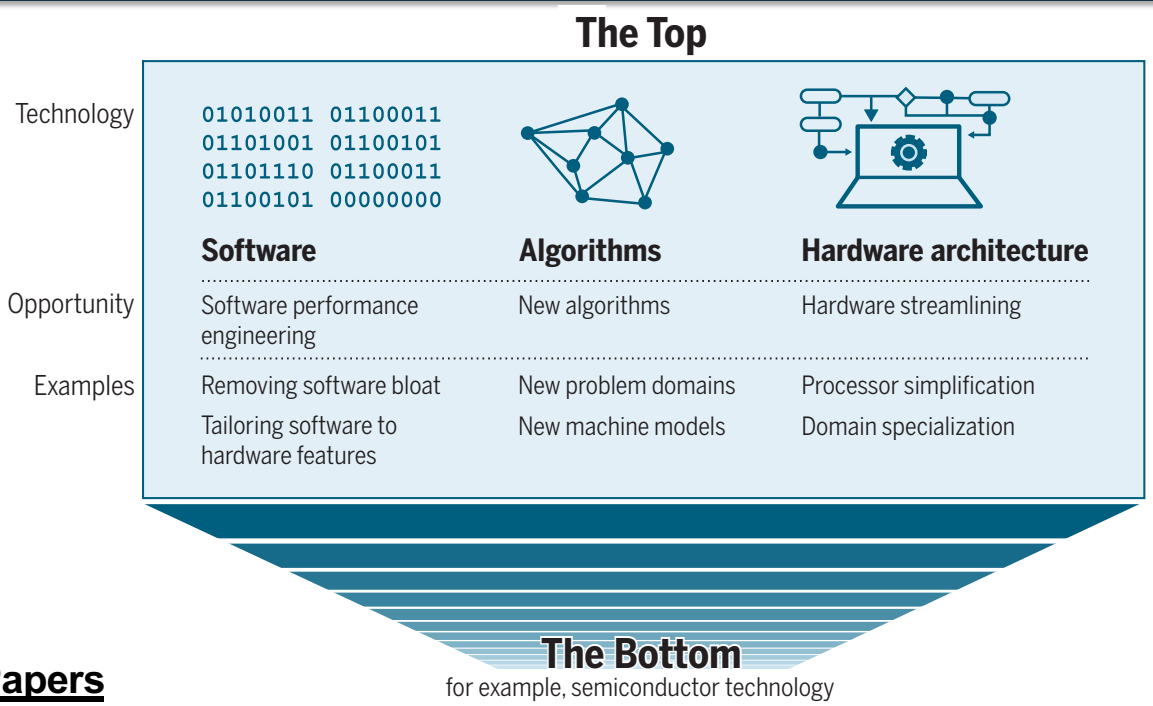
KNL AMD, Cavium/Marvell, GPU

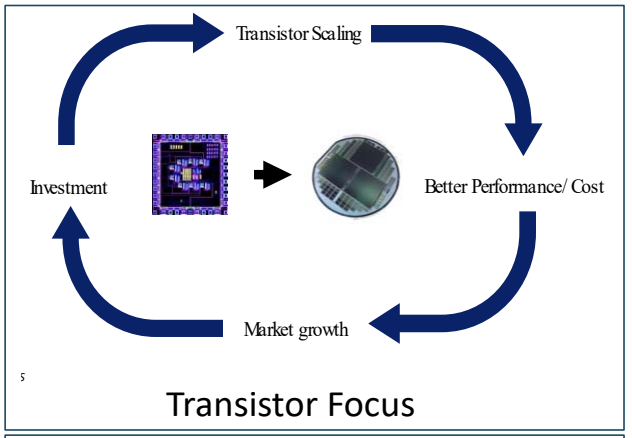
Many Different Specialized
(Post-Moore Scarcity)

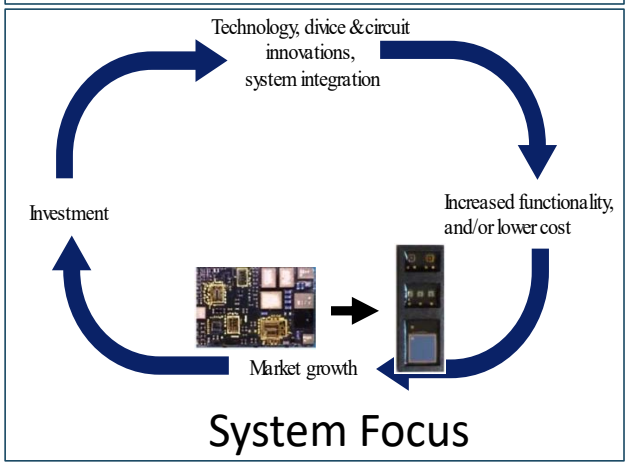
Apple, Google, Amazon, Microsoft Azure

Leiserson/Thompson: Economics of Post-Moore Electronics

http://neil-t.com, MIT CSAIL, MIT Sloan School





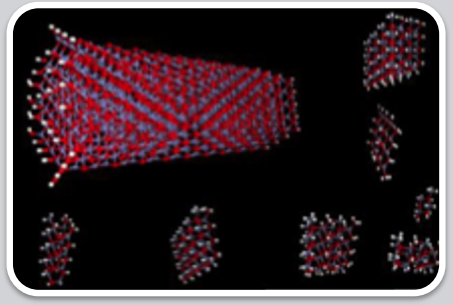


Papers

1. The Economic Impact of Moore's Law
2. There's Plenty of Room at the Top: What will drive computer performance after Moore's Law?
3. The Decline of Computers as a General Purpose Technology

Architecture Specialization for Science

(hardware is design around the algorithms) can't design effective hardware without math

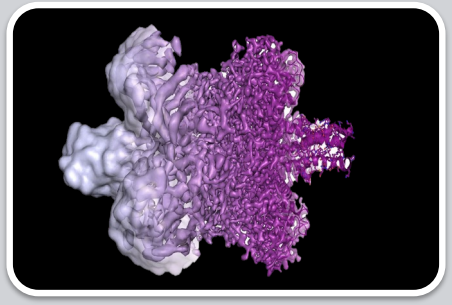


Materials

Density Functional Theory (DFT)

Use O(n) algorithm

Dominated by FFTs

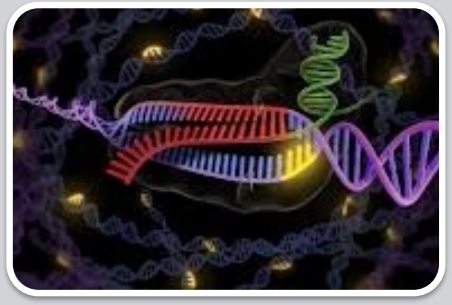


Smart Sensors

CryoEM detector

750 GB / sec

Custom compute near detector

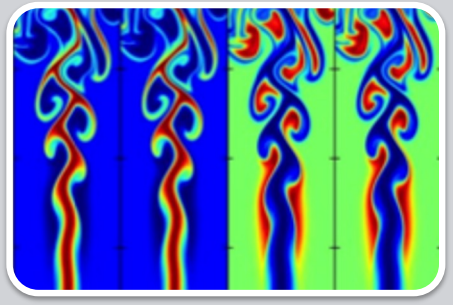


Genomics

String matching

Hashing

2-8bit (ACTG)



PDEs on Block Struct. Grids

3D integration

Petascale *chip*

1024-layers

Analogous Computing

Technology Insertion into Mainstream Platforms

AMD, Intel, Arm offer integration path for 3rd party accelerator "chiplets"

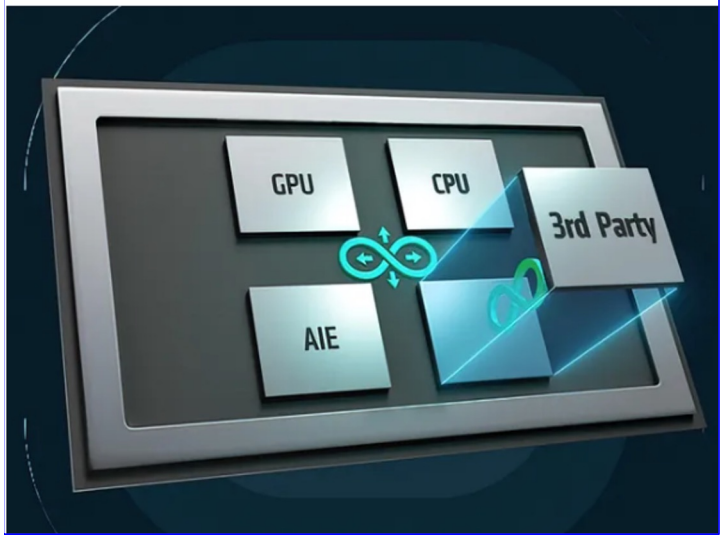
Modular AMD Chips to Embrace Custom 3rd Party Chiplets

News By Francisco Pires last updated June 20, 2022

Supercharging learnings - and earnings - from the console space.

Comments (2)

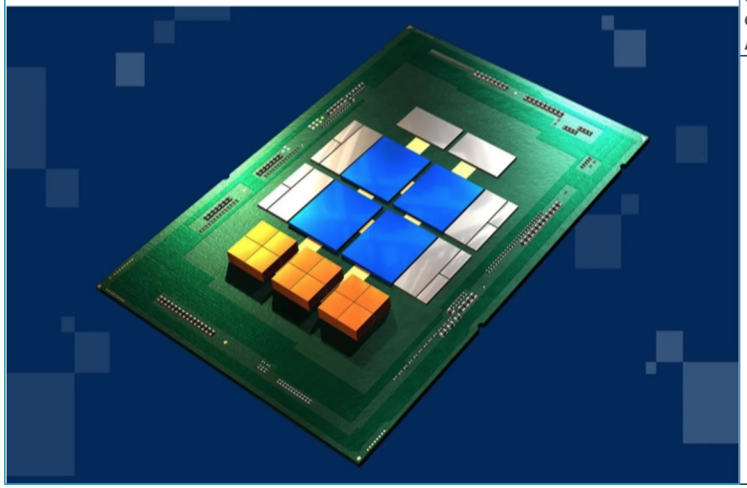
When you purchase through links on our site, we may earn an affiliate commission. Here's how it works.



To 'Meteor Lake' and Beyond: How Intel Plans a New Era of 'Chiplet'-Based CPUs

At the Hot Chips 2022 conference, Intel teased its upcoming 'Meteor Lake' and 'Arrow Lake' processor families, which will use multiple tiny tiles fused together in an attempt to break free of the limits of monolithic chip design. Here's why little tiles are a big deal.

By Michael Justin Allen Sexton August 24, 2022

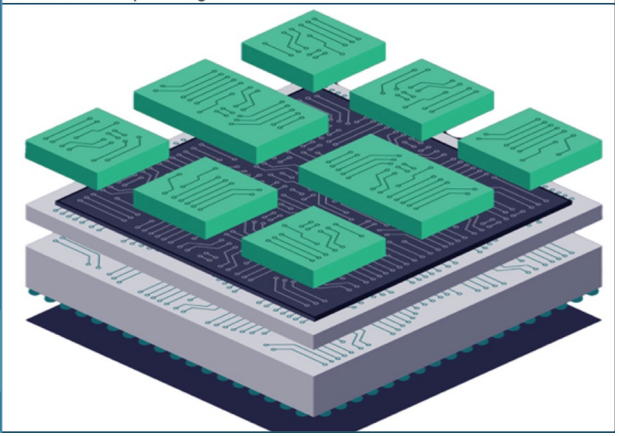


ARM Opens Door to Make Custom Chips for HPC, AI

By Doug Eadline

October 19, 2023

It is safe to say that ARM isn't a scrappy startup that was once the pride of the UK. The US-based IPO made the chip designer a big-game chip player, and the new capital is kickstarting some major initiatives to find more customers for its products. A new effort called Total Design aims at making it easier for companies looking to design chips in-house, an idea gaining ground with the AI boom and chip shortages.



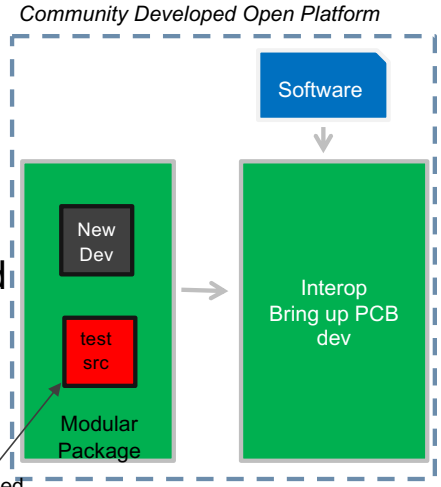
BERKELEY LAB

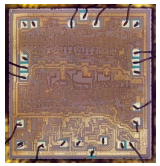
http://chiplets.lbl.gov/

More Efficient Chiplet Development and Integration Path

Platform for open development with path into commercial platform

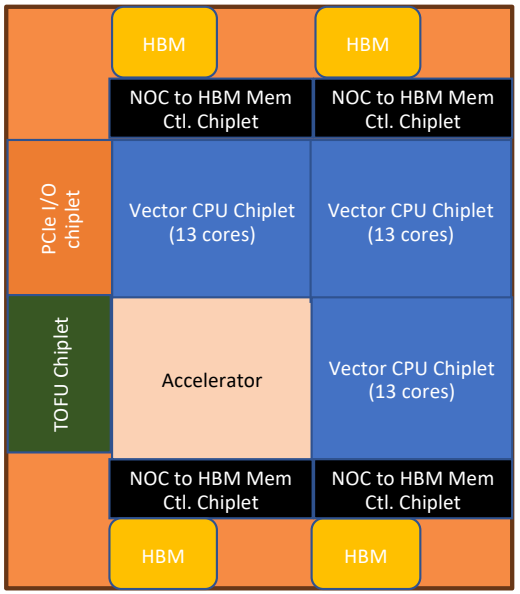
Common Open Chiplet Eval Board for Bring-up



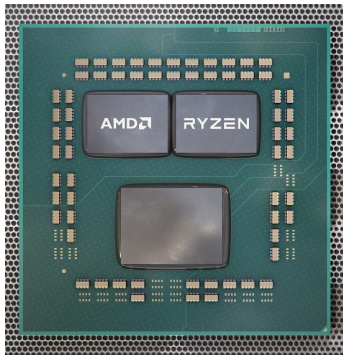


Chiplet

Reused Gold Dev



Prototype on Modular Architecture for function and performance integration



Commercial product integration

Analogous Computing

Build systems that are analogous to the problem they are solving

Analog Computing

Definition: Analog computing refers to a type of computation that uses continuously variable physical quantities to represent and solve problems. Instead of using discrete binary values (0s and 1s) like in digital computing, analog computers work with continuous data.

Analogous Computing

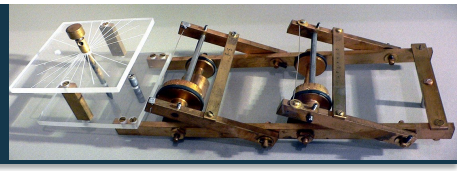
Definition: Analogous computing looks at how one system can mimic the behavior of another in terms of how problems are addressed. This can be algorithmic, physical, and even structural/topological analogies! For example, neuromorphic computing that mimics biological process is a form of analogous computing.

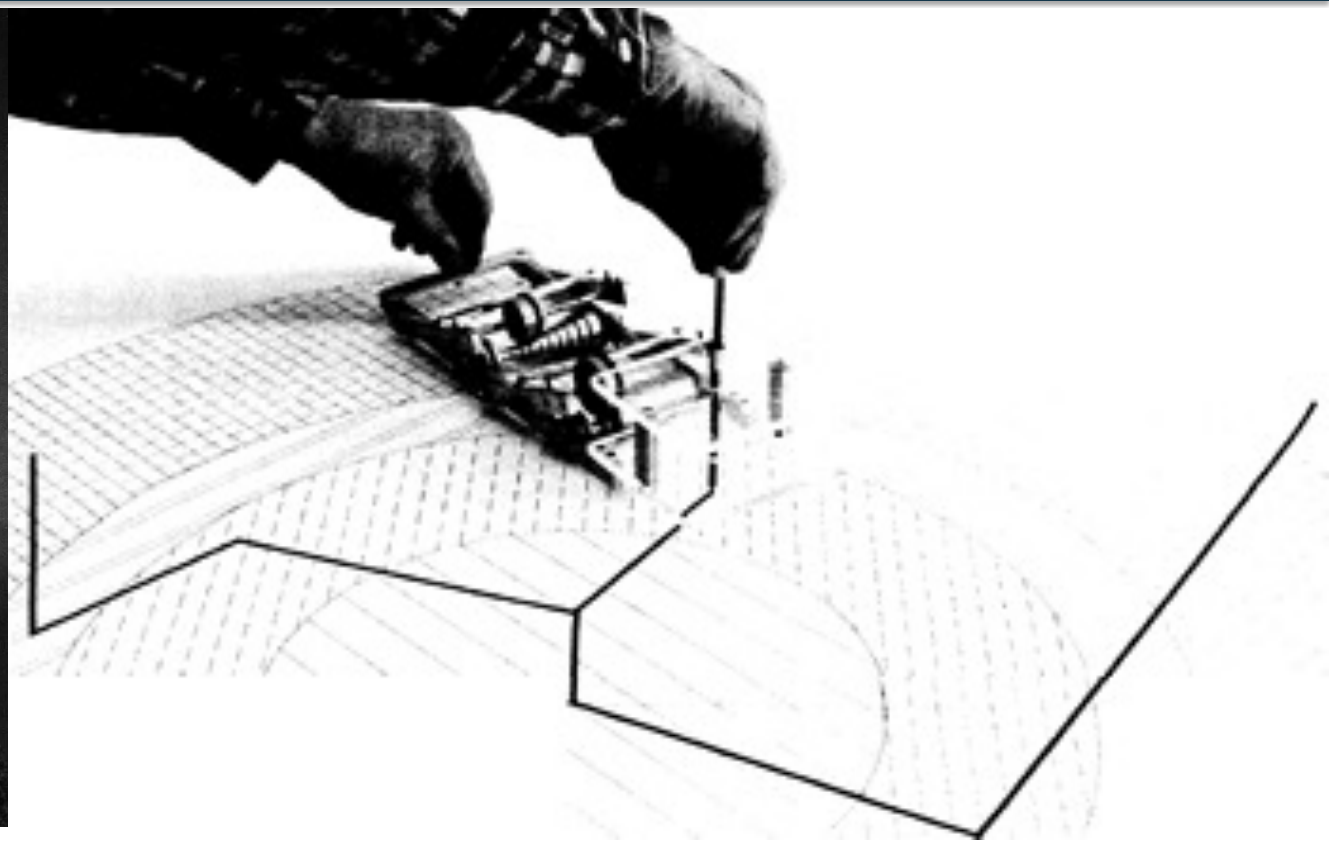
Fundamental efficiency benefits can be realized by embracing the structure of the physics being solved

For example using quantum computing to create "artificial atoms" to solve for ground state of atom.

DOE's Rich History of Analogous Computing for Science

Monte Carlo Method at LANL using FERMIAC



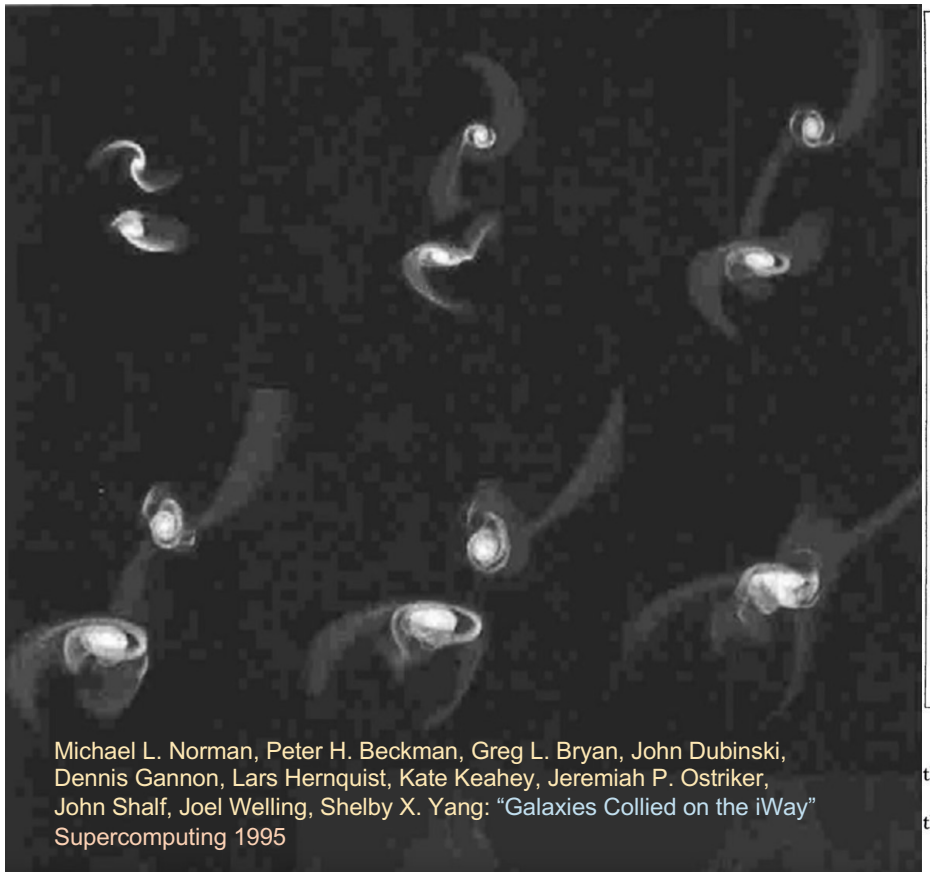


FERMIAC analogous machine for Neutronics Calculations

Not saying we go back to Monte Carlo robots with pencils,
but it is an interesting way to think about building energy efficient computing

Analogous Computing: Simulating Colliding Galaxies with Light Bulbs

Note: *Not suggesting that we go back to light-bulb computing. But it is pretty cool for 1939!*



Michael L. Norman, Peter H. Beckman, Greg L. Bryan, John Dubinski, Dennis Gannon, Lars Hernquist, Kate Keahey, Jeremiah P. Ostriker, John Shalf, Joel Welling, Shelby X. Yang: "Galaxies Collied on the iWay" Supercomputing 1995

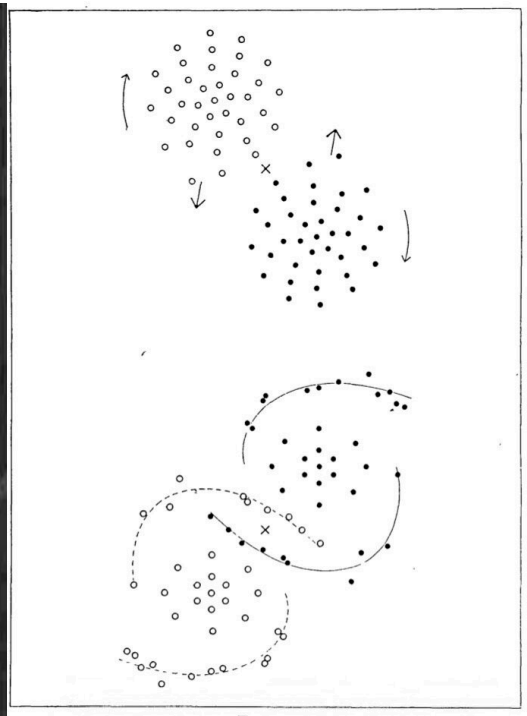


FIG. 4a

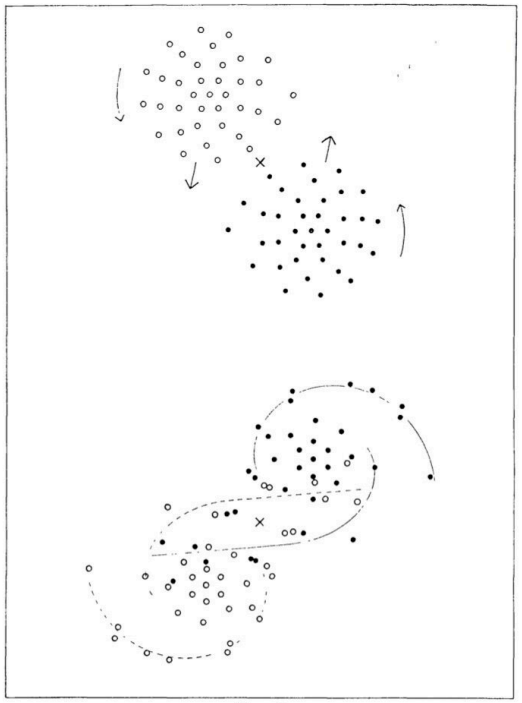


FIG. 4b

FIG. 4a.—Tidal deformations corresponding to parabolic motions, clockwise rotations, and a distance of closest approach equal to the diameters of the nebulae. The spiral arms point in the direction of the rotation.

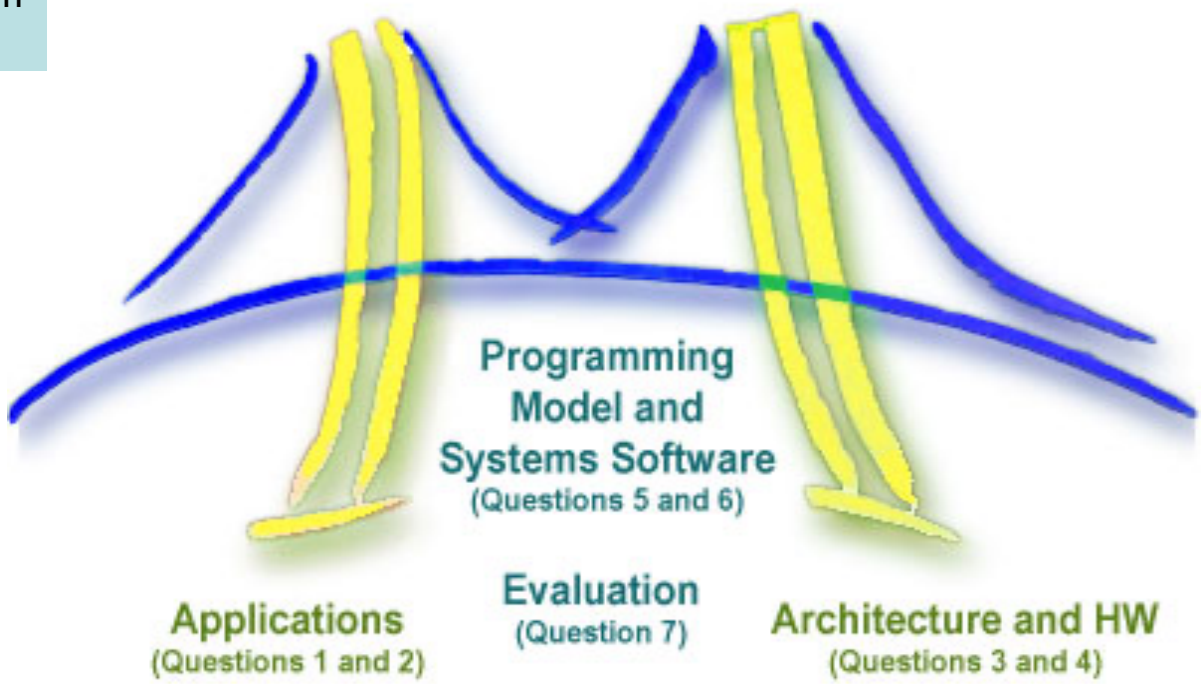
FIG. 4b.—Same as above, with the exception of counterclockwise rotations. The spiral arms point in the direction opposite to the rotation.

Phil Colella's 7 Dwarfs of Scientific Computing

High-end simulation in the physical sciences = 7 numerical methods:

Exploit the mathematical structure of the problem
design principle for "analogous" computing

1. Structured Grids
2. Unstructured Grids
3. Fast Fourier Transform
4. Dense Linear Algebra
5. Sparse Linear Algebra
6. Particles
7. Monte Carlo



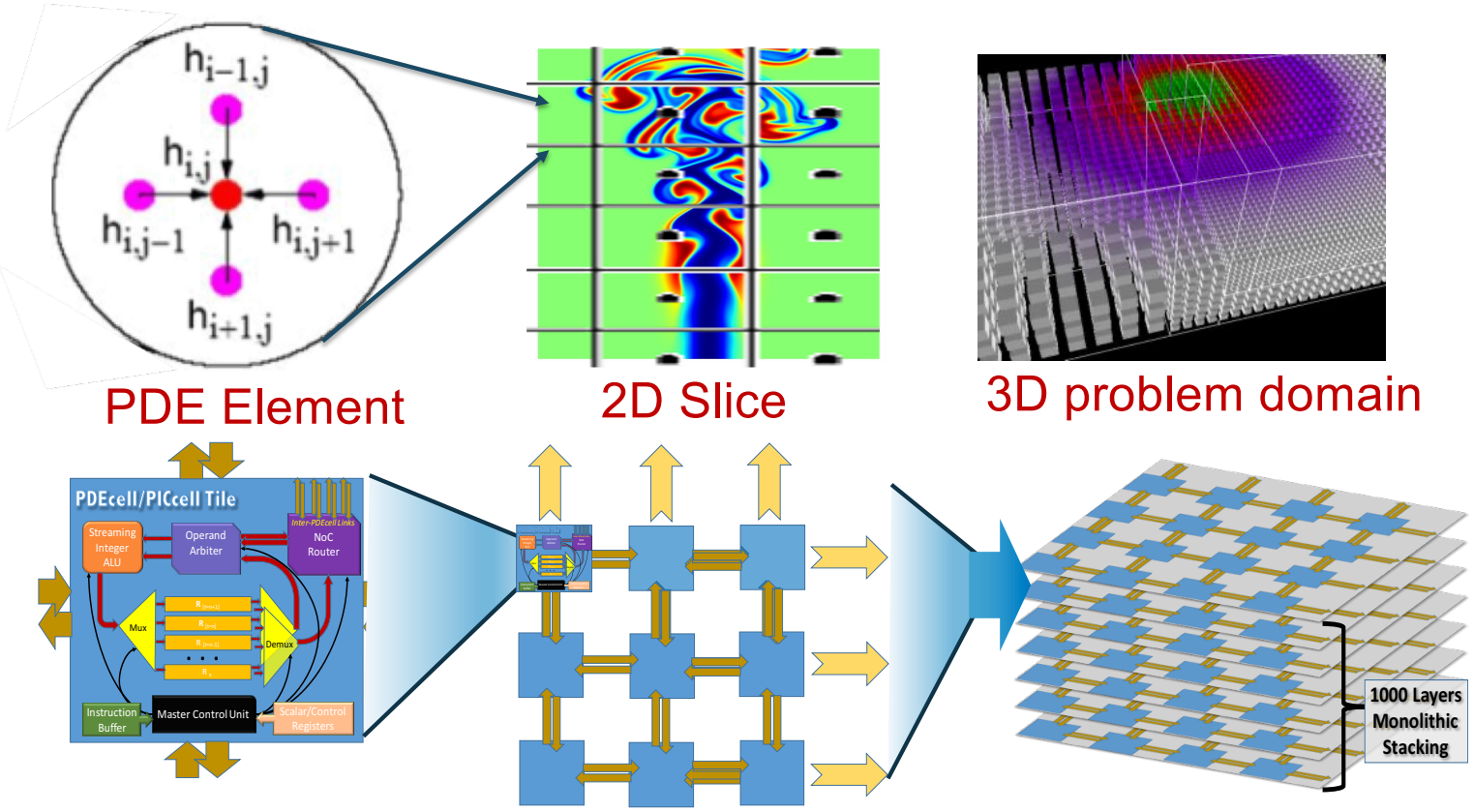
Slide from "Defining Software Requirements for Scientific Computing", Phillip Colella, 2004
Also in "The Landscape of Parallel Computing Architecture: A view from Berkeley" 2008

http://www2.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-183.pdf

Office of Science
U.S. DEPARTMENT OF ENERGY

PDE Solvers on Block Structured Grid

"Solid State Digital Fluid" digital analogous computing concept



Concept: Solid State Virtual Fluid for CFD, PIC and QMC

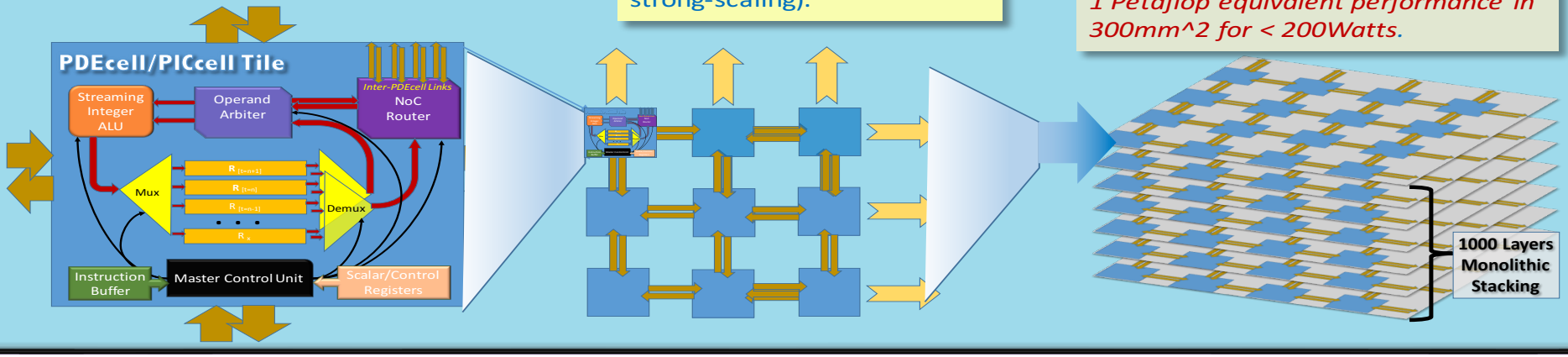
Lawrence Berkeley National Laboratory, John Shalf
Stanford University, Subashish Mitra

DARPA-BAA-16-38 ACCESS
Attachment 1 – Proposal Concept

PDEcell / PICcell: Ultra-simple compute engine (50k gates) calculates finite-difference updates, and particle forces from neighbors. Microinstructions specify the PDE equation, stencil, and PIC operators. ***Novel features:*** *variable length streaming integer arithmetic and novel PIC particle virtualization scheme.*

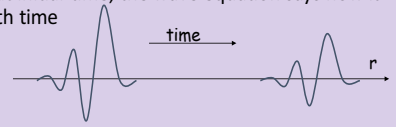
Computational Lattice: PDECells are tiles in a lattice/array on each 2D planar chip layer. Target 120x120 tiles per mm² @28nm lithography. Novel Features: each tile represents single cell of computational domain (pushes to limit of strong-scaling).

Monolithic 3D Integration: Integrate layers of compute elements using emerging monolithic 3D chip stacking. ***Novel Features:*** *1000 layer stacking (20x more than current practice). Area efficient inter-layer connectivity and new energy efficient transistor logic (ncFET).* *1 Petaflop equivalent performance in 300mm^2 for < 200Watts.*



Scalar waves in 3D are solutions of the hyperbolic wave equation: $-\phi_{,tt} + \phi_{,xx} + \phi_{,yy} + \phi_{,zz} = 0$

Initial value problem: given data for ϕ and its first time derivative at initial time, the wave equation says how it evolves with time



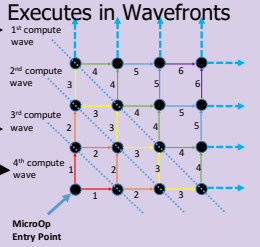
Discretized PDE Representation in DSL

$$\phi^{n+1}_{i,j,k} = 2\phi^{n}_{i,j,k} - \phi^{n-1}_{i,j,k}$$
$$+ \Delta t^2/\Delta x^2(\phi^{n}_{i+1,j,k} - 2\phi^{n}_{i,j,k} + \phi^{n}_{i-1,j,k})$$
$$+ \Delta t^2/\Delta y^2(\phi^{n}_{i,j+1,k} - 2\phi^{n}_{i,j,k} + \phi^{n}_{i,j-1,k})$$
$$+ \Delta t^2/\Delta z^2(\phi^{n}_{i,j,k+1} - 2\phi^{n}_{i,j,k} + \phi^{n}_{i,j,k-1})$$

Compiles to MicroOps

```
R[n+1](0,0,0) = 0
R[n+1](0,0,0) += 2*R[n](0,0,0)
R[n+1](0,0,0) -=  R[n-1](0,0,0)
R[n+1](0,0,0) += C * R[n+1](+1,0,0)
R[n+1](0,0,0) -= C * 2 * R[n](0,0,0)
R[n+1](0,0,0) += C * R[n](-1,0,0)
R[n+1](0,0,0) += C * R[n+1](0,+1,0)
    ....
```

Executes in Wavefronts



Source Selection Information – see FAR 2.101 & 3.104

LBNL Beyond Moore Microelectronics CoDesign Framework

Recasting as capabilities that span the entire LBNL Complex

	Systems	Circuits	Device Physics	Junction Physics	Materials Physics
Length Scales	Processor/System: ~10k-100k Circuits	Circuit/Std. Cell: 10-100 Devices	1 Device: ~1M Atoms	1 Junction: ~100k Atoms	Bulk Material: ~100 Atoms
Measure/Metrology	NERSC	LBNL ASIC Design Grp.	CXRO, ADVANCED LIGHT SOURCE		ADVANCED LIGHT SOURCE, MOLECULAR FOUNDRY
Synthesize/Prototype	OPEN Compute Project		Met5+/CHiiP, EUREKA, MOLECULAR FOUNDRY		Chan Group
Modeling and Simulation Capabilties (e.g. Theory)	Systems	Circuits	ARTEMIS Device Simulation	LS3DF DFT	MatProj

BERKELEY LAB

U.S. DEPARTMENT OF ENERGY

Final Thoughts

- **Analogous computing is a broader term for exploiting the structure of the problem – algorithmically, but also topologically, and even using materials that mimic the physics of the problem being solved.**

- **Analogous Computing unifies why we would use analog, quantum and brain-inspired computers to solve specific scientific problem domains**
  - These are specializations… its not a general purpose
  - But that is OK!

- **Specialization is inevitable: The broader industry is adopting it, but HPC is resisting. *Attack of the killer micros lesson is "follow the industry trends"***

- **One last cautionary tale about the *danger* of focusing too much on lowering power consumption *(from the photonics community)***

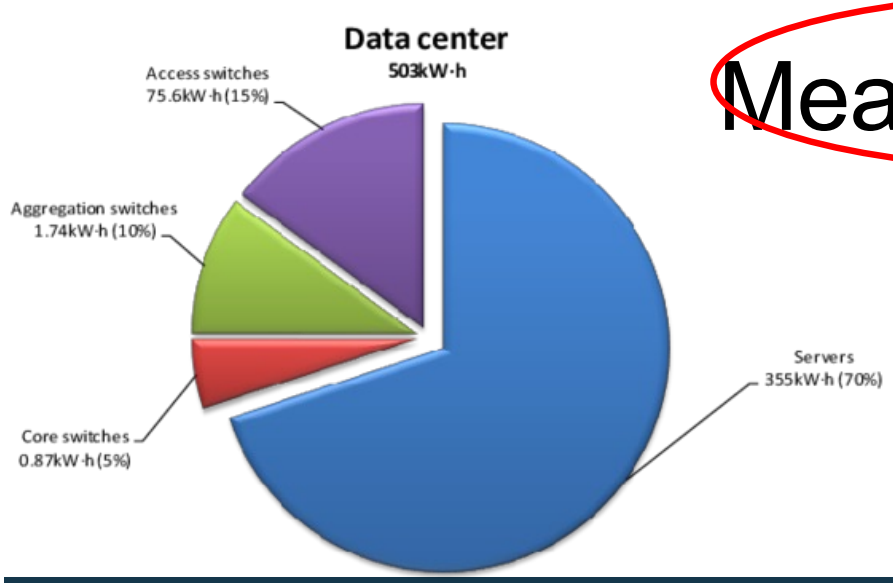
Anatomy of a “Value” Metric

$$\frac{\text{Good Stuff}}{\text{Bad Stuff}}$$

Anatomy of a “Value” Metric

$$\frac{\text{Performance}}{\text{Measured Watt}}$$

Anatomy of a "Value" Metric



Performance

Measured Watt

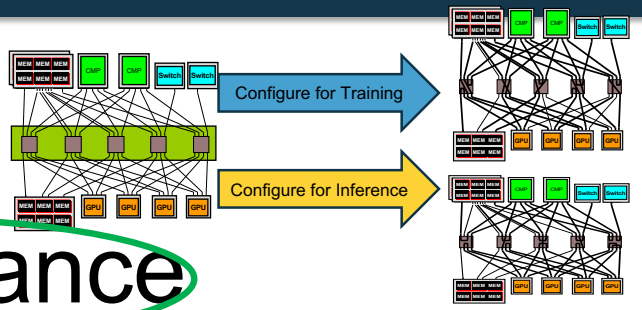
30% of datacenter power goes to network

So max savings by creating perfectly efficient (0 pJ/bit) optical interconnect is ONLY 30%!

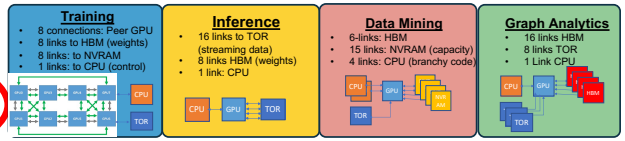
Anatomy of a "Value" Metric

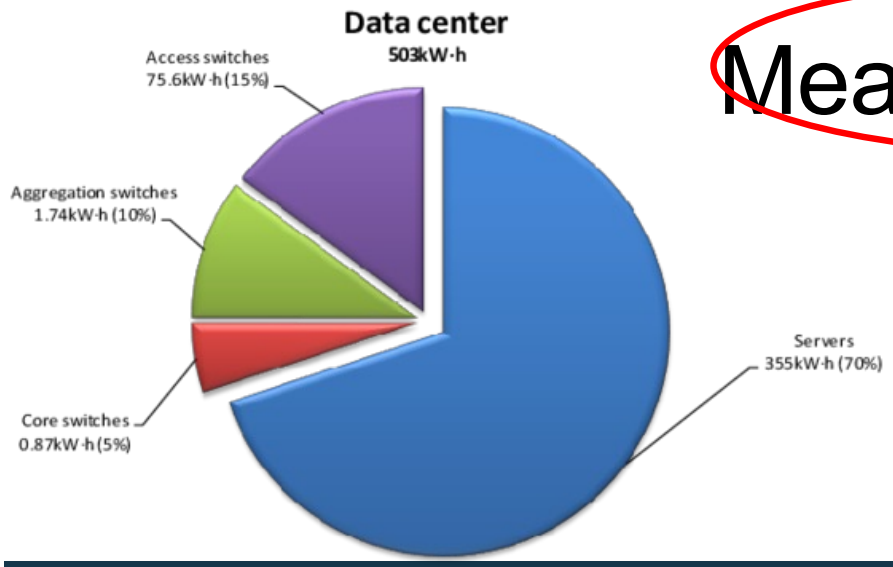
Increase performance with Disaggregation
And Bandwidth Steering!

Deliver bandwidth and resources to where it is needed
By taking it from where it isn't



$$\frac{\text{Performance}}{\text{Measured Watt}}$$





30% of datacenter power goes to network

So max savings by creating a perfectly efficient (0 pJ/bit) optical interconnect is ONLY 30%!

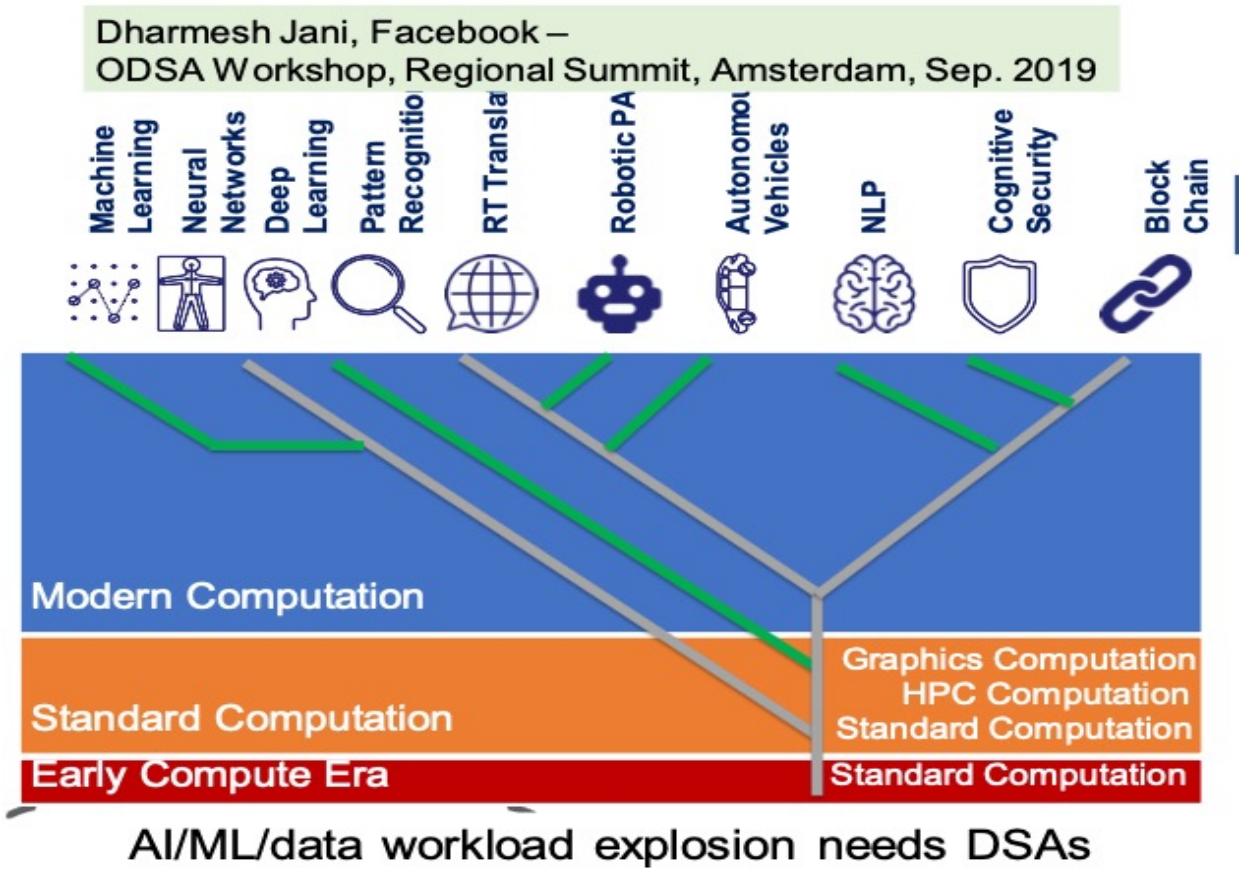
BERKELEY LAB

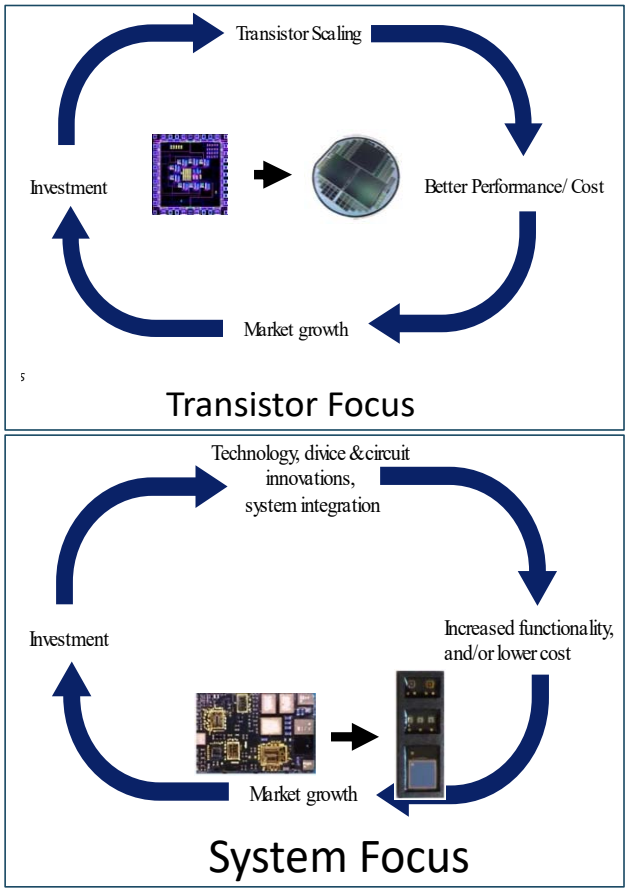
Exploit the unique capabilities to improve both numerator and denominator!

extra

Why? Domain specific Architectures driven by hyperscalers

in response to slowing of Moore's Law (switch to systems focus for future scaling)





Technology Insertion into Mainstream Platforms

AMD, Intel, Arm offer integration path for 3rd party accelerator "chiplets"

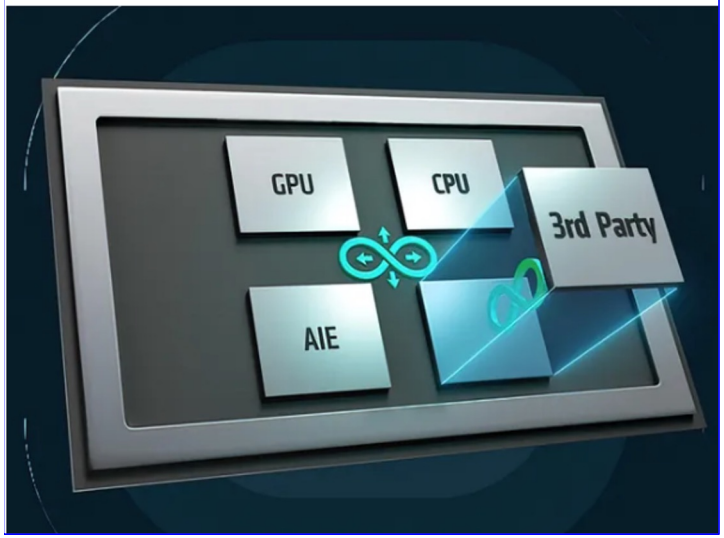
Modular AMD Chips to Embrace Custom 3rd Party Chiplets

News By Francisco Pires last updated June 20, 2022

Supercharging learnings - and earnings - from the console space.

Comments (2)

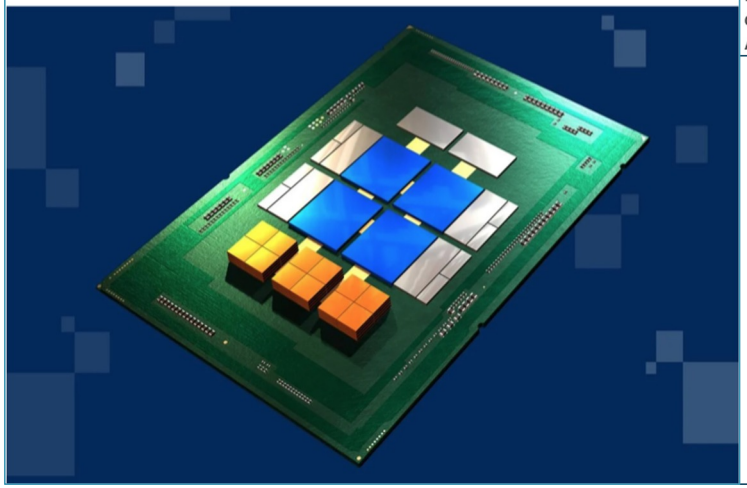
When you purchase through links on our site, we may earn an affiliate commission. Here's how it works.



To 'Meteor Lake' and Beyond: How Intel Plans a New Era of 'Chiplet'-Based CPUs

At the Hot Chips 2022 conference, Intel teased its upcoming 'Meteor Lake' and 'Arrow Lake' processor families, which will use multiple tiny tiles fused together in an attempt to break free of the limits of monolithic chip design. Here's why little tiles are a big deal.

By Michael Justin Allen Sexton August 24, 2022

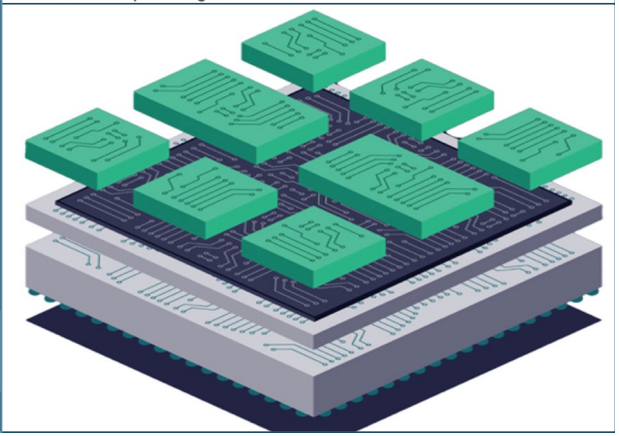


ARM Opens Door to Make Custom Chips for HPC, AI

By Doug Eadline

October 19, 2023

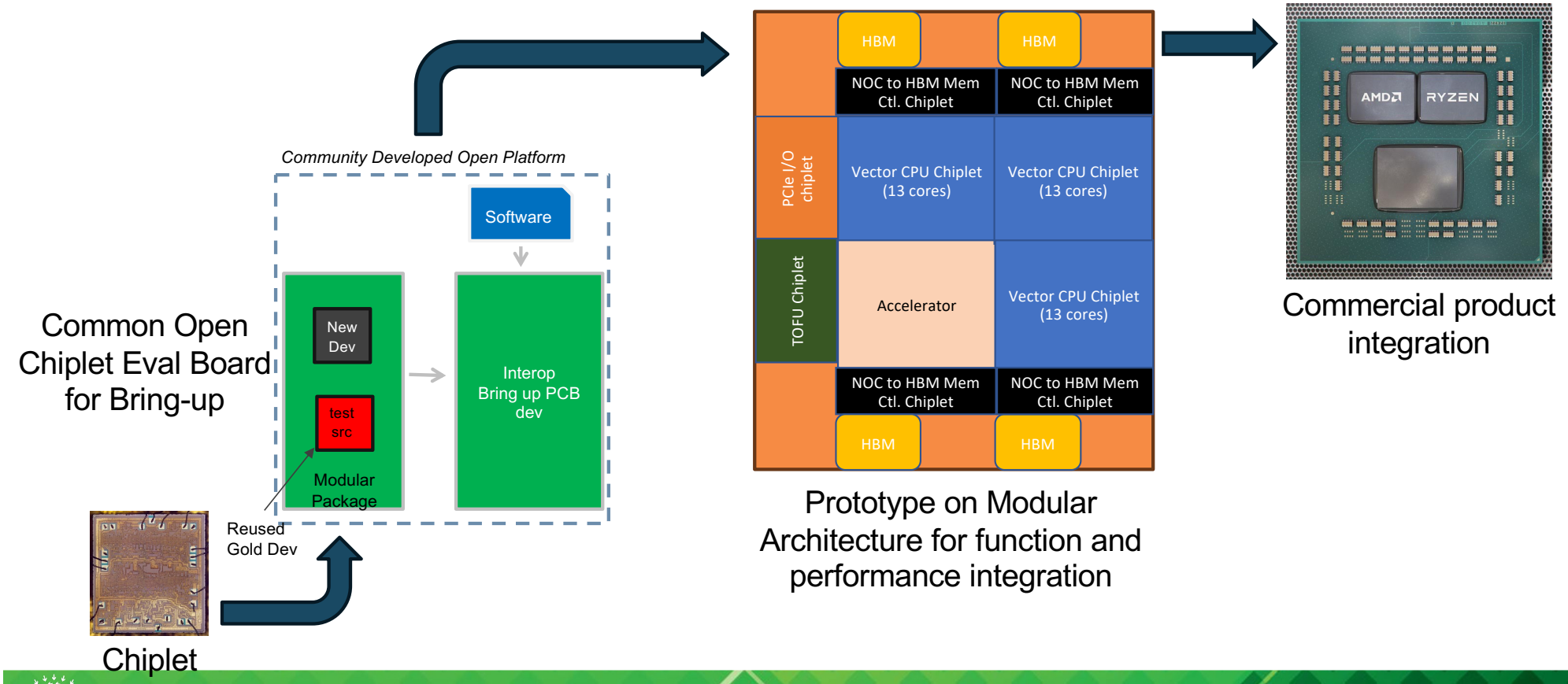
It is safe to say that ARM isn't a scrappy startup that was once the pride of the UK. The US-based IPO made the chip designer a big-game chip player, and the new capital is kickstarting some major initiatives to find more customers for its products. A new effort called Total Design aims at making it easier for companies looking to design chips in-house, an idea gaining ground with the AI boom and chip shortages.



BERKELEY LAB

http://chiplets.lbl.gov/

More Efficient Chiplet Development and Integration Path



Chiplets at LBL

Home Register Demos Meeting Venue

http://chiplets.lbl.gov

LBNL/OCP Open Chiplet Economy Experience Center

Hosted by Lawrence Berkeley National Laboratory (LBNL)

Co-organized by the Open Compute Project (OCP)

Date: June 24, 2024

Time: 12:00pm to 5:00pm

Location: Berkeley National Lab, Wang Hall Bldg. 59, Room 59-3101

OPEN Compute Project®

2024 JEDEC and OCP Standards for Chiplet Design with 3DIC Packaging Workshop

JEDEC