

## Heterogeneous Integration for HPC and Data Centers



### **TWG Chair:** Kanad Ghose, Ph.D.

Distinguished Professor of Computer Science, SUNY-Binghamton  
PhD (CS), M.Tech (EE)  
and Site Director of Center for Energy-Smart Electronic Systems,  
a NSF Industry/University Collaborative Research Center



### **TWG Co-Chair:** Dale Becker, Ph.D. (*emeritus*)

Formerly Chief Engineer of Electronic Packaging Integration, IBM Systems  
PhD (EE), MS (EE)  
Fellow IEEE  
Chair, IEEE EPS TC-EDMS



### **TWG Co-Chair:** John Shalf (*newbie*)

Department Head for Computer Science and Computer Engineering  
Lawrence Berkeley National Laboratory  
MS (EE/CE)



<http://eps.ieee.org/technology/heterogeneous-integration-roadmap.html>

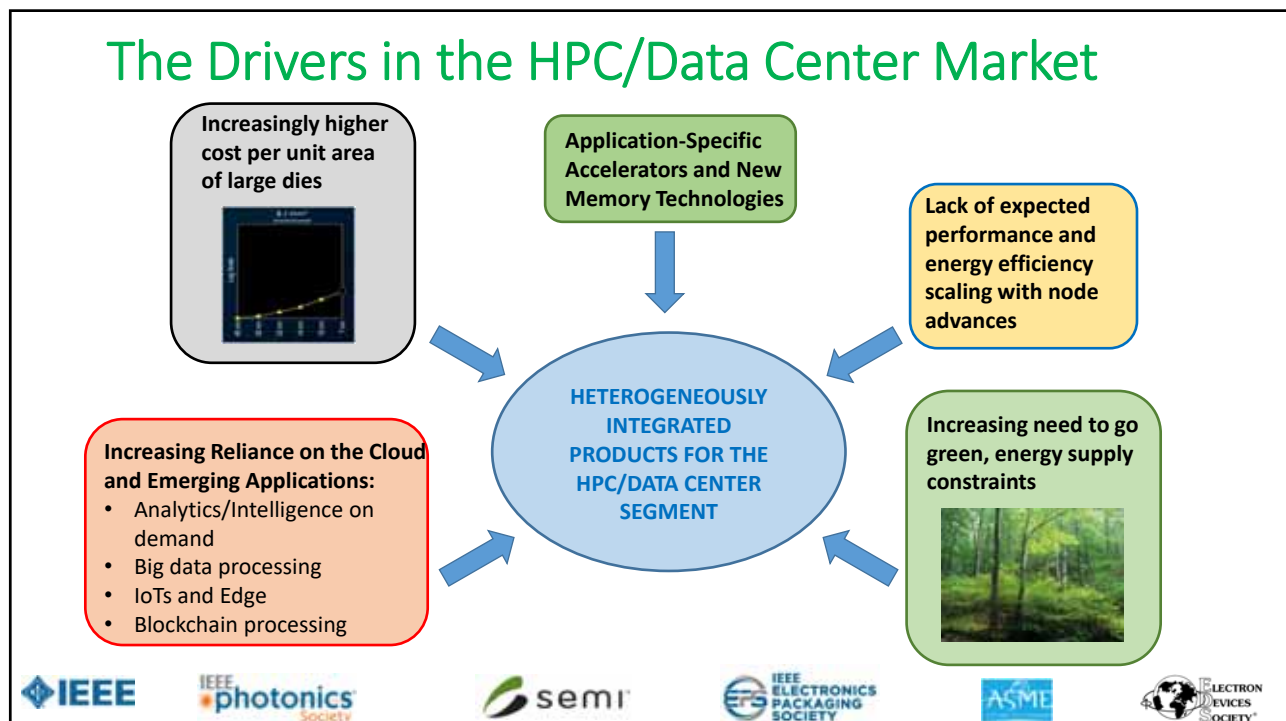
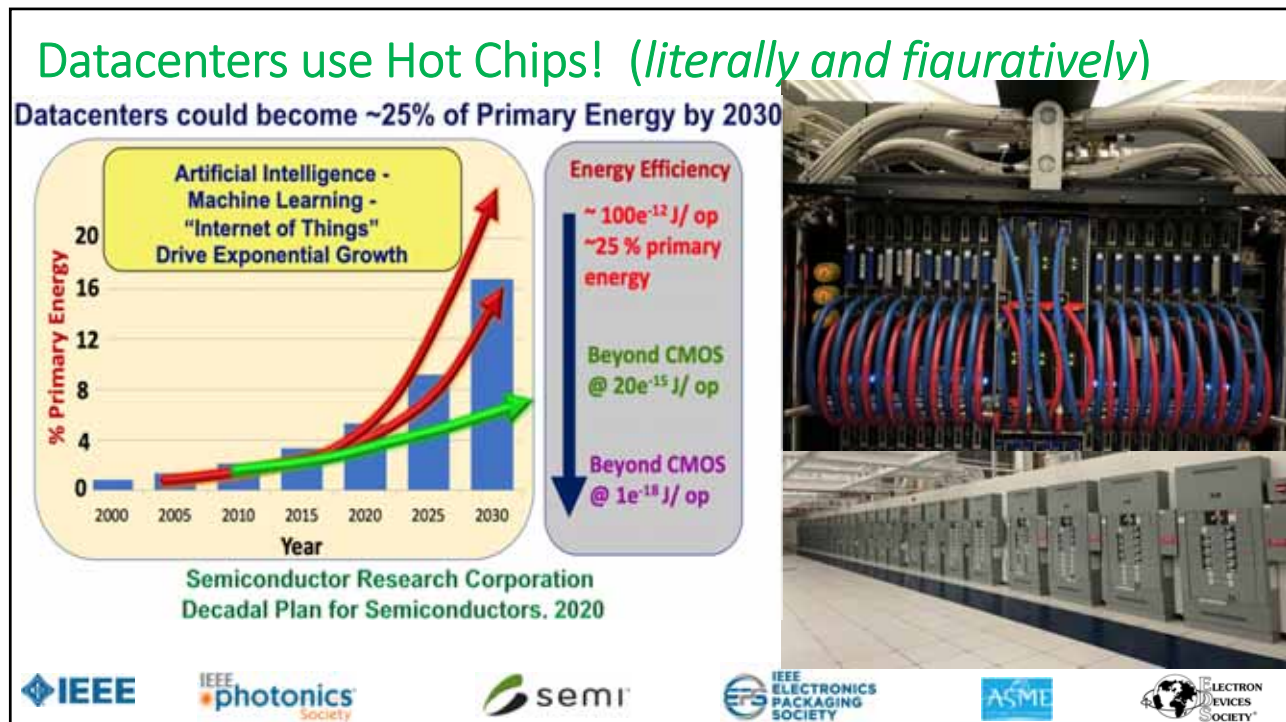
[eps.ieee.org/hir-2021](http://eps.ieee.org/hir-2021)

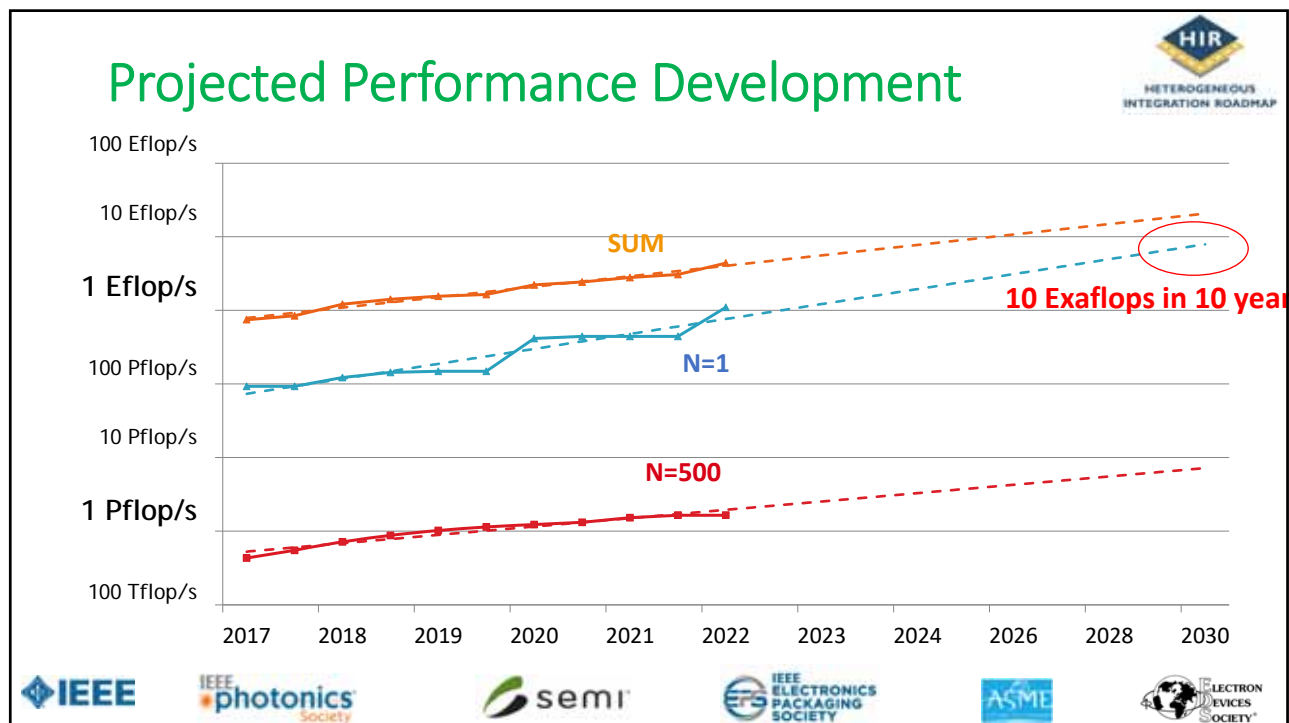
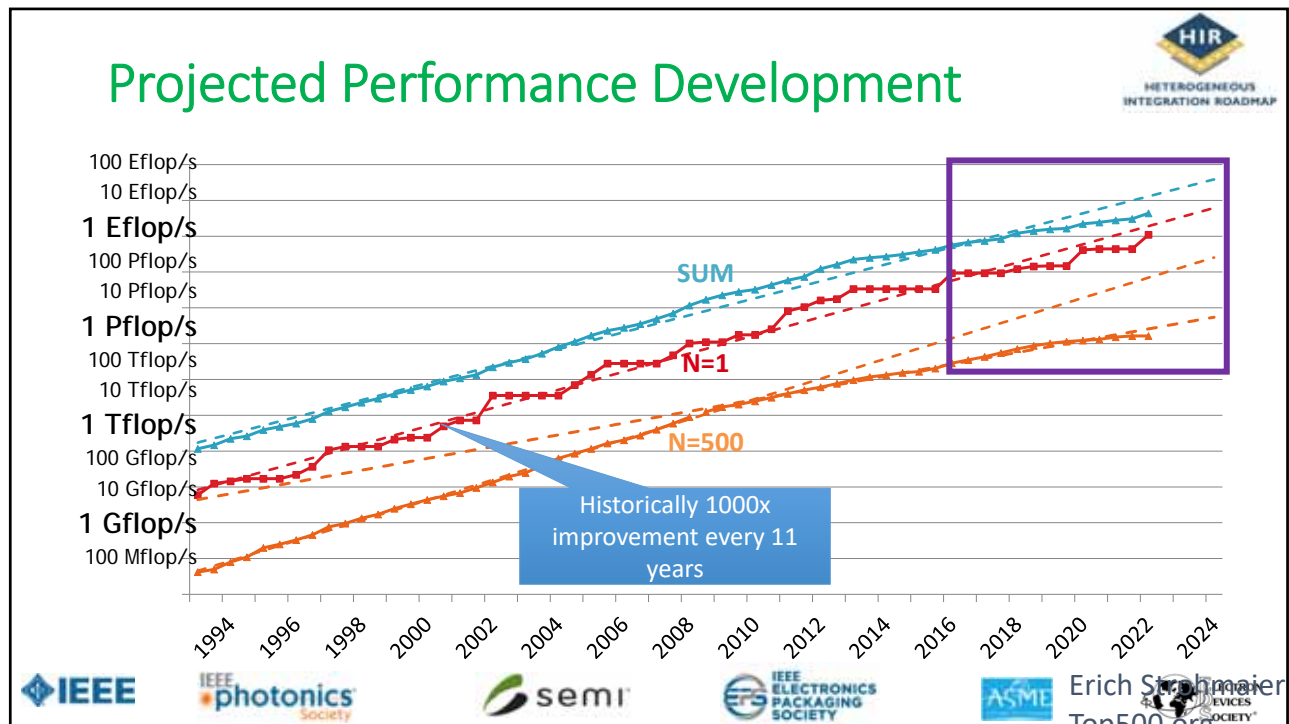
## Intent and Notes

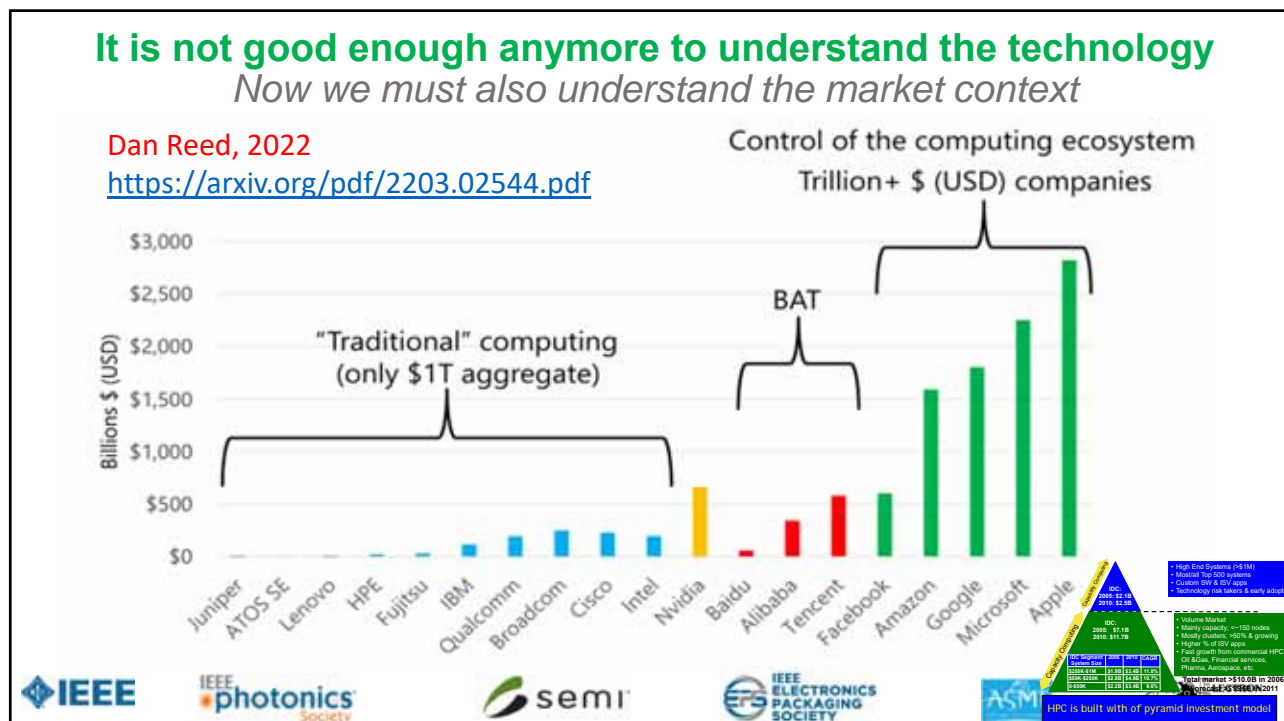
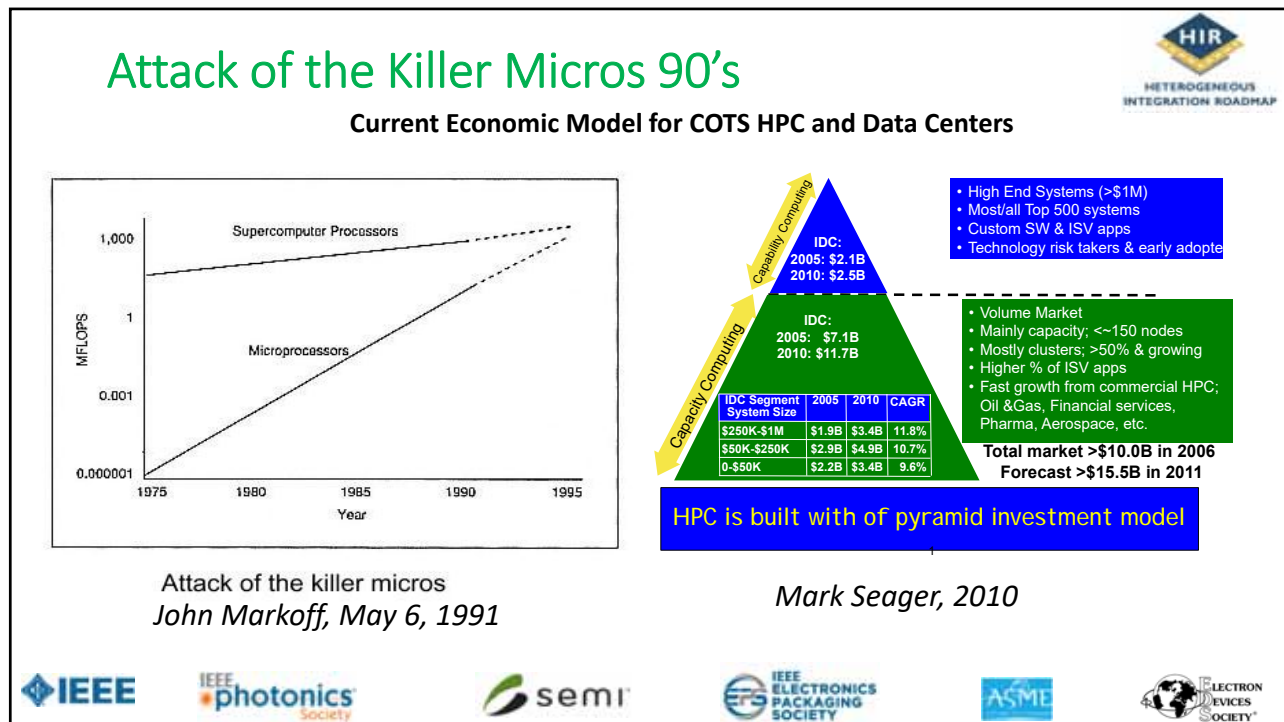


- **Impress on readers that heterogeneous integration is not just about coming up with a packaging solution to house connected chiplets**  
**Its all about systems integration**
- There are **many crosscutting issues** that need to be considered as part of the packaging solution:
  - Diversity of chiplets
  - Interconnections
  - Power conversion and delivery
  - Security issues
  - Other considerations, including QC systems
  - Verification/test, Design automation
- **Not possible to come up with generations, quantification of trends of various factors**, but some order of magnitude trends can be made in many cases
  - System architectures/ components evolving continuously
  - Tying trends to a timeline is difficult





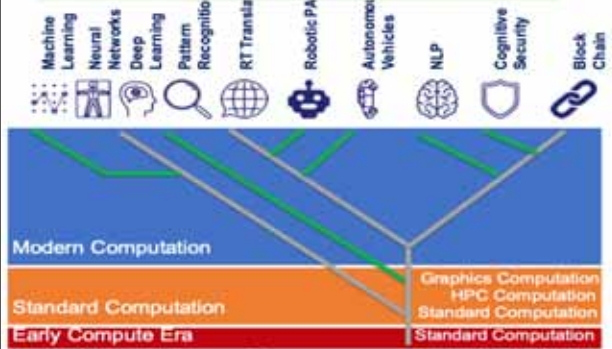




## What is next? Codesign of Domain Specific Architectures

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

Dharmesh Jani, Facebook –  
ODSA Workshop, Regional Summit, Amsterdam, Sep. 2019



AI/ML/data workload explosion needs DSAs

- Hyperscalers are imaging a future of domain Specialization built on a foundation of advanced packaging for modularity
- HPC is discovering it has to follow a similar approach to stay aligned with industry forces.

The Top

```
01010011 01100011
01101001 01100101
01101110 01100011
01100101 00000000
```



Software

Software performance engineering  
Removing software bloat  
Tailoring software to hardware features

Algorithms

New algorithms  
New problem domains  
New machine models

Hardware architecture

Hardware streamlining  
Processor simplification  
Domain specialization

The Bottom

for example, semiconductor technology

Papers from Neil Thompson ([neil-t.com](http://neil-t.com))

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

## Specialization:

Natures way of Extracting More Performance in Resource Limited Environment

Powerful General Purpose



Xeon, Power

Many Lighter Weight  
(post-Dennard scarcity)



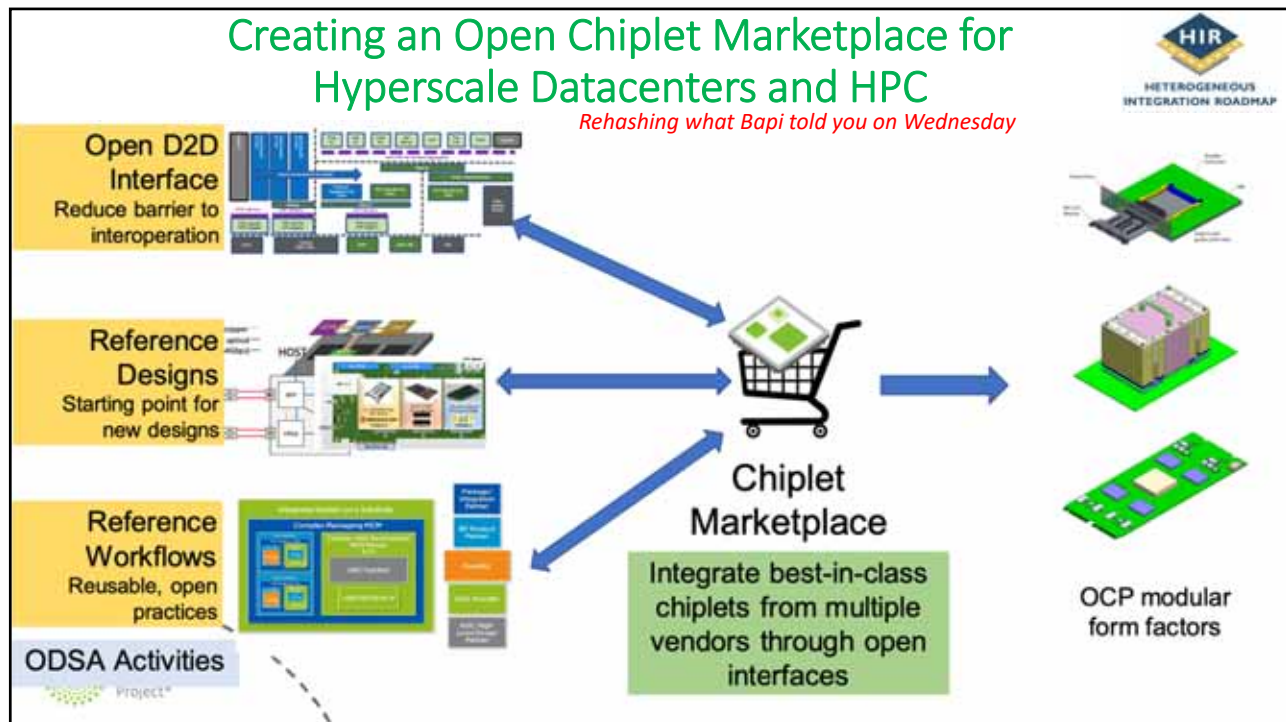
KNL, AMD, Cavium/Marvell, GPU

Many Different Specialized  
(Post-Moore Scarcity)



Apple, Google, Amazon





## Opportunity for New Economic Model for Affordable Specialization

### Open Chiplets Marketplace is forming (ODSA and UClexpress)

- Licensable IP and assembly by 3<sup>rd</sup> party lowers that barrier
- Leverage the economic model being created by HyperScale


### Leverage this baseline and extend to support HPC

- Smaller incremental cost for HPC to “play”
- *HPC has become “too small to attack the city”*

### 80:20 Rule: Focus open efforts on what uniquely benefits HPC

- Build up a library of reusable accelerators for HPC.
- **Interoperability for sustainability:** Interoperate with Arm IP for commercially supported IP where it exists and focus Open on the 20% that doesn't make commercial sense to license












## HPC/Data Center TWG Members

- Tawfik Arabi, AMD
- Ivor Barber, AMD
- Dale Becker, IBM
- Bill Bottoms, 3MTS
- Tahir Cader, HPE
- Don Draper,
- William Chen, ASE fellow
- Luke England, Marvell
- Eric Eisenbraun, SUNY
- Kanad Ghose, Binghamton
- Ali Heydari, NVIDIA

- Rockwell Hsu, Cisco
- Madhu Iyengar, Google
- Sam Karikalan, Broadcom
- Michael Liehr, Leih Consulting
- Ravi Mahajan, Intel
- Gamal Refai-Ahmed, SRC
- Tom Salmon, SEMI
- Lei Shan, IBM
- Bahgat Sammakia, SUNY
- John Shalf, LBNL
- Raja Swaminathan, AMD
- Jin Y. Kim, Google (*just joined*)





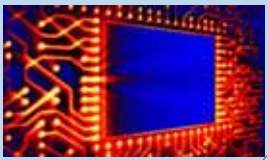

## Table of Contents for Chapter 2







### *High Performance Computing and Data Centers*

<p>Chapter 1: Heterogeneous Integration Roadmap: Driving Force and Enabling Technology for Systems of the Future</p> <p>Chapter 2: High Performance Computing and Data Centers</p> <p style="padding-left: 20px;">Introduction: The Need for Heterogeneous Integration .....</p> <p style="padding-left: 20px;">Analyzing the Future Demands of SiPs .....</p> <p style="padding-left: 20px;">Demands of Future SiPs and Solutions for the HPC/DC Market .....</p> <p style="padding-left: 20px;">Chiplet Standards for Heterogeneous Integration Targeting HPC and Data Centers .....</p> <p style="padding-left: 20px;">Heterogeneous Integration and its Role in Quantum Computing .....</p> <p style="padding-left: 20px;">Applicable Tracking Metrics .....</p> <p>Chapter 3: Heterogeneous Integration for the Internet of Things (IoT)</p> <p>Chapter 4: Medical, Health and Wearables</p> <p style="border: 1px solid green; padding: 2px;">Chapter 5: Automotive</p> <p>Chapter 6: Aerospace and Defense</p> <p>Chapter 7: Mobile</p> <p>Chapter 8: Single Chip and Multi Chip Integration</p> <p>Chapter 9: Integrated Photonics</p>	<p style="border: 1px solid red; padding: 2px;">Chapter 10: Integrated Power Electronics</p> <p>Chapter 11: MEMS and Sensor Integration</p> <p>Chapter 12: 5G, RF and Analog Mixed Signal</p> <p style="border: 1px solid red; padding: 2px;">Chapter 13: Co-Design for Heterogeneous Integration</p> <p>Chapter 14: Modeling and Simulation</p> <p>Chapter 15: Materials and Emerging Research Materials</p> <p>Chapter 16: Emerging Research Devices</p> <p>Chapter 17: Test Technology</p> <p style="border: 1px solid red; padding: 2px;">Chapter 18: Supply Chain</p> <p style="border: 1px solid red; padding: 2px;">Chapter 19: Cyber Security</p> <p style="border: 1px solid red; padding: 2px;">Chapter 20: Thermal</p> <p style="border: 1px solid red; padding: 2px;">Chapter 21: SiP and Module</p> <p style="border: 1px solid red; padding: 2px;">Chapter 22: Interconnects for 2D and 3D Architectures</p> <p style="border: 1px solid red; padding: 2px;">Chapter 23: Wafer-Level Packaging, Fan-in and Fan-out</p> <p>Chapter 24: Reliability</p> <p style="color: red;">Chapter 25: Quantum?</p>
--	---

## Ongoing Crosscuts









- **Interconnects:** [See Chapter 22, 23](#)
  - **Functions:** Die2Die (in package) and escape bandwidth
  - **Challenges:** reach, lower latency, signal integrity/power, simplified transceiver
  - **Emerging Solutions:** high density omni-dimensional interconnects, CoBO, backside power delivery, advanced symbol encoding, Open D2D stds. (UCle/BoW)
- **Power Conversion/Delivery:** [See Chapter 10](#)
  - **Functions:** Power delivery and maintaining power quality (low EMI)
  - **Challenges:** Insane current densities, noise sources, delivery in 3DI
  - **Emerging Solutions:** Higher voltage feed to package and local/in-package conversion, new power devices, new converter topologies, advanced magnetics, dielectrics, backside delivery, power routing within interposer, embedded converters
- **Thermal Management:** [See Chapter 20](#)
  - **Functions:** Heat removal, heat spread, active thermal mgmt.
  - **Challenges:** Cost/reliability, heat removal in 3D stacks, mech. stress, nonuniform heat
  - **Solutions:** software managed thermal mgmt., Conformal lids, dummy dies with thermal vias, advanced TIMs, water cooling, 2-phase cooling, immersion cooling, low-temperature CVD deposition diamond thin films

## Security: [See Chapter 19](#)

- **Specific Needs:** **Trust No-one security (Trusted Execution Environments)**
  - **Address package-level compromises:** compromised chiplets, **side channels** on shared interconnections
  - **Address chiplet-scale compromises, particularly with firmware in chiplets:** *isolation at chiplet boundaries*
  - **Address compromises in the supply chain**
  - **Simultaneously address reliability needs on chiplet failures**
- **Challenges:**
  - Lack of consensus on needs and threat model
  - Lack of standardization at interfaces for security needs (even root of trust)
  - Tradeoffs among performance, power, area etc. – deciding what is optimal
- **Some Solutions:**
  - **Root of trust, Tamper proofing, chiplet watermarking, boundary isolation, package-internal TPM-like certification, embedded security co-processor, side-channel elimination techniques**



## Emerging area Quantum Computing

- **Specific Needs:**

- Support general-purpose computations
- Reduction of physical dimensions
- Improvement of energy efficiency: Cooling needs to zero degrees Kelvin in *many* existing solutions get in the way of this
- Support high reliability

- **Challenge areas:**

- Qubit decoherence
- Qubit manipulations: complexity in terms of size, electronics, precision, power
- Qubit measurements
- Qubit count (= system size) scaling
- Cooling (for many QC systems, **not all**)

- **Some Solutions:**

- See Section 5 of CHAPTER 2



## Planned Updates for 2023

- **Ongoing:** *updates charts and figures*

- **Shorten chapter text!!!!**

- **Modular HPC**

- **Chiplet D2D interfaces**

- Update on role of Open Chiplets Ecosystems on hyperscale + HPC
- Add extended section on open ecosystem D2D link standards **UCle** and **BoW**

- **Memory**

- Update section on memory devices with recent developments
- Need for open Unified Virtual Memory (**UVM**) designs/standards/APIs

- **Codesign: *Simulation and Modeling for early design assessment***

- Add role of co-design and need for rapid design space exploration
- Hardware innovations by themselves are no good till the software to use them exist!

- **Enhance discussions on analog accelerators with trend data to support power scaling.**

- **Enhance discussions on sensors for failure/reliability tracking etc.**

- *Crosscut with automotive*





## Cross-TWG Collaborations

- Need for cross-TWG coordination and collaborations with many chapters:
  - Interconnections for 2D/3D
  - Single chip/multi-chip integration
  - Thermal
  - Integrated Photonics / CoBO
  - Integrated Power Electronics
  - **Automotive** (*Vika Gupta contacted me yesterday about this*)
  - Security
  - **New QC group??**
  - Test
  - Co-Design
  - Others



## Highlights of New Content in 2021 Edition



- Substantial updates were made in 2020; 2021 updates were *mostly* limited to addressing reviewer comments. **Next updates were planned for 2023.**
- Chiplet diversity: emphasized role of analog ML accelerators, GDDRx/GDDR6x as a cheaper alternative to HBM, new HBM generation
- Integration: added subsection on INFO-RDL (using TSMC variants as examples)
- Interconnections: updated several numerical data, emphasized subsection on face-to-face microbump connections, updated mitigation of 3D power delivery challenge, added trends
- Power delivery: added subsection on PowerVIAs/Backside power delivery
- Added small section on cold atom qubits
- Updated trend table data (at end of chapter) in a few cases

## Interconnections



- **Key functions:**
  - Inter-chiplet: **tight coupling to reap full benefits of package-level integration**
  - IO: **Overcome physical package level IO limits** to effectively integrate SiP into rest of the platform
- **Challenges/Needs:**
  - **Low latency:** *mostly for inter-chiplet*
  - **High bandwidth:** *for both inter-chiplet and IO; practical photonics technologies for IO connections*
  - **Low end-to-end power/bit:** *function of reach*
  - **Low error rate**
  - **Reduced physical footprint:** *simplified data transceivers (e.g., clock-forwarded links to eliminate PLLs), smaller physical dimensions of interconnections*
  - **Minimize interference of data routing and power routing**
- **Some Solutions:**
  - **High-density omni-dimensional interconnections, microbumps and face-to-face bonding, integrated photonics transceivers, advanced symbol encoding, back-side power delivery, standards, standards, standards...**
  - **SEE CHAPTERS 22 and 23**

## Power Conversions and Distribution



- **Needs:**
  - **Higher power demands** of high-end SiPs
  - **High power quality, slew rate, low EMI**
  - Power distribution to chiplets
- **Challenges:**
  - **Higher Ohmic ( $I^2R$ ) losses** at sub-1V and 200+ Amps: *requires conversion/regulation at or near point-of-loading (POL), with higher input voltage*
  - **Precise voltage regulation** for functional integrity of analog components/chiplets **with high efficiency across entire load range** with or without DVFS
  - **Noise mitigation** – *particularly with integration of analog chiplets, NV memory*
  - **Power delivery in deep 3D stacks**, Interference among power routing vs. signal routing
  - **Materials and devices:** magnetics, decoupling capacitors, integrable/small-footprint power devices
- **Some Solutions:**
  - **Higher voltage feed to package and local/in-package conversion, new power devices, new converter topologies, advanced magnetics, dielectrics, back-side delivery, power routing within interposer, embedded converters**
  - [SEE CHAPTER 10](#)

## Thermal Management



- **Needs:**
  - **Address high power densities** with SiP demands over 250W to 500W: *practical limits of conventional air cooling exceeded*
  - **Address mechanical stresses on smaller inter-chiplet interconnections**
  - **Cool power devices, clock drivers, optical power supplies and other photonic components** etc. effectively
- **Challenges:**
  - **Cost and reliability** of cooling solutions: *nothing too exotic*
  - Address **cooling needs of 3D stacked chiplets**
  - Address non-uniform temperature distributions: critical for reducing SiP level failures
- **Solutions:**
  - **Conformal lids, dummy dies with thermal vias, advanced TIMs, water cooling, 2-phase cooling, immersion cooling**
  - [SEE CHAPTER 20](#)

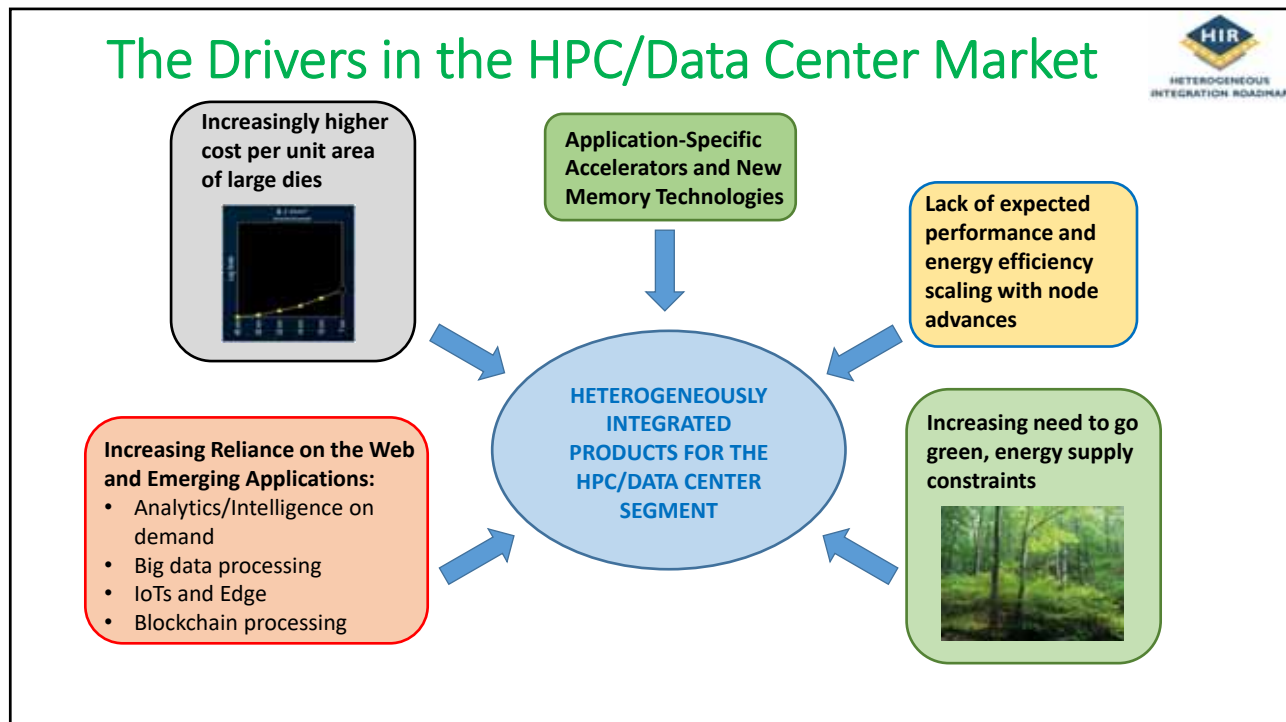




- Put up high level thoughts
- Connections to other areas slide
- Put them as horizontal slabs with requirements and terms
- Outline of presentation?
- Might more closely track the development of server class D2D.
  - Where is the “roadmap”?
  - Specialization targets (aligned and distinct)



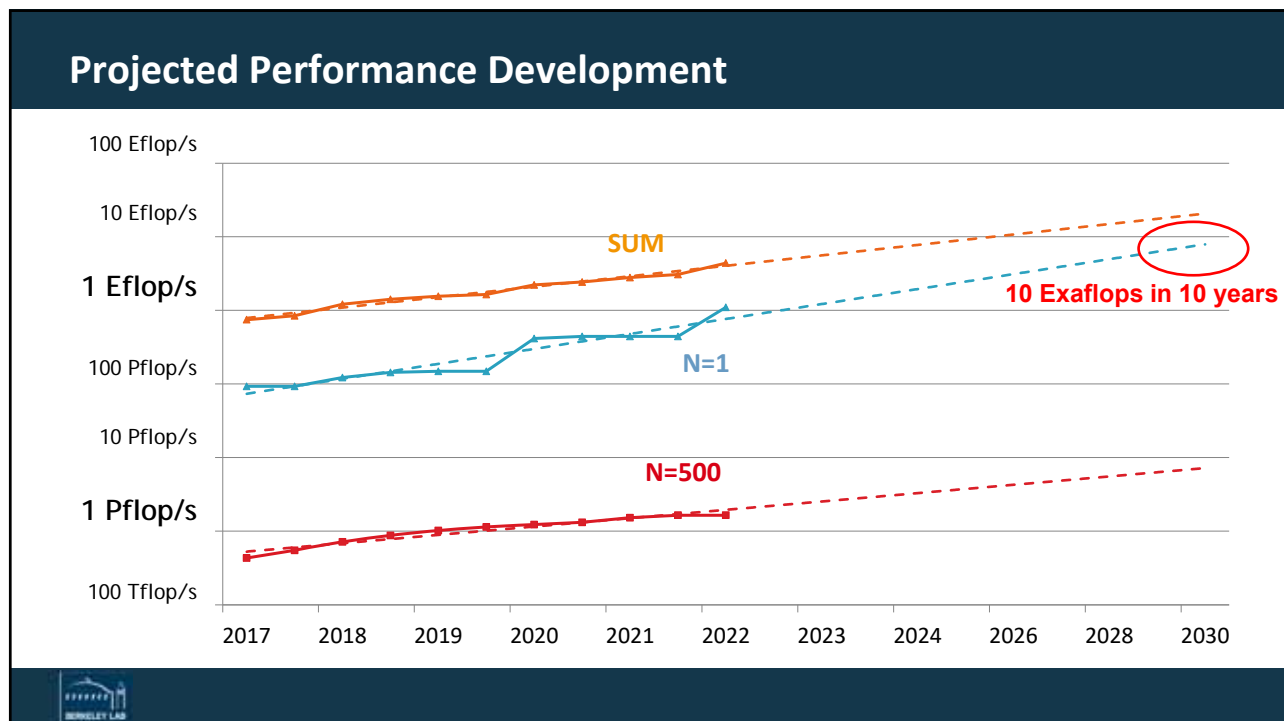
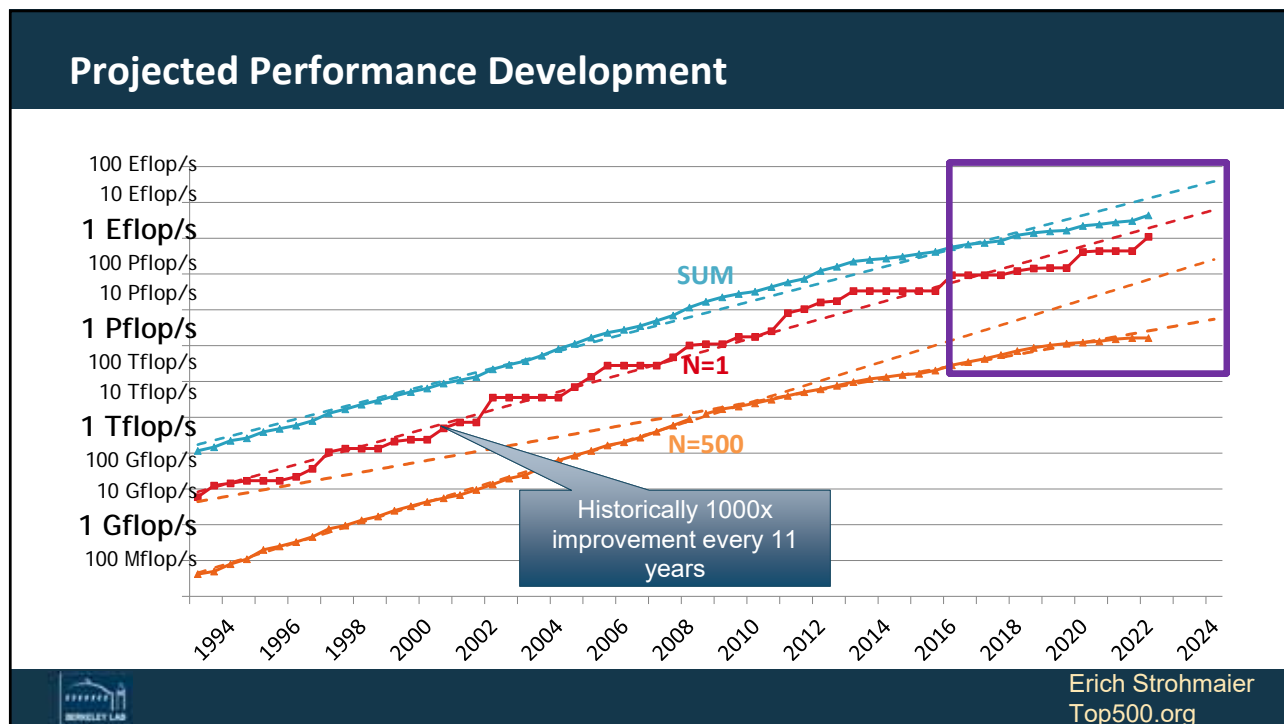
- Chapter Contributors?
- Table of Contents for the Chapter
  - With color coding areas that did get updates and will get updates
- Modular HPC and Datacenters using chiplets
- Hottest chip on the market (H100)
- And a picture of our latest supercomputer



## The HPC Data Center Segment

- **Specific needs:**
  - Often very **application-specific**
  - **Integration of diverse processing and storage technologies:** *inevitable*
  - **Scalability:** *exploit technology advances, scale within and across SiP for supporting existing and emerging applications for data size and processing needs, support **disaggregated architectures***
  - **Improved performance:** *wide diversity of compute and memory chiplets helps!*
  - **Improved energy-efficiency:** *accelerators and near/in-memory computing, analog accelerators for machine learning etc.*

***Its all about moving data from Point A to Point B with low latency, high bandwidth and low end-to-end energy cost!***



## Specialization: Nature's 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)

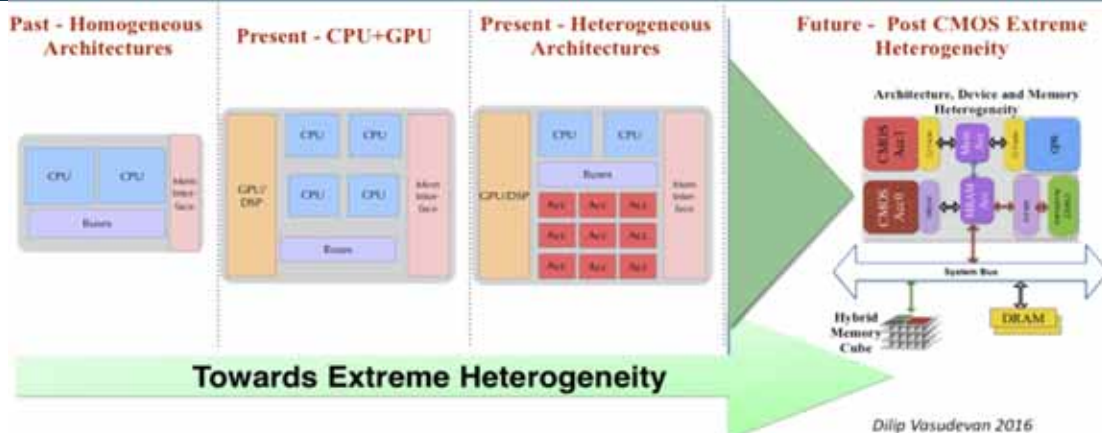


Apple, Google, Amazon



3  
1

## The Future Direction for Post-Exascale Computing



But what are the right specializations to include?

What is the cost model (we know we cannot afford to spin our own chips from scratch)

The ARM licensable IP ecosystem : ***IP is the commodity (not the chip)***

What is the right partnership/economic model for the future of HPC?





## Neil Thompson: Economics of Post-Moore Electronics

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

### The Top

	Technology	Software	Algorithms	Hardware architecture
Technology	01010011 01100011 01101001 01100101 01101110 01100011 01100101 00000000			
Opportunity		Software performance engineering	New algorithms	Hardware streamlining
Examples		Removing software bloat Tailoring software to hardware features	New problem domains New machine models	Processor simplification Domain specialization

### The Bottom

for example, semiconductor technology

**Transistor Focus**

**System Focus**

### Papers

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

## Attack of the Killer Micros 90's

### Current Economic Model for COTS HPC

**Attack of the killer micros**  
John Markoff, May 6, 1991

- The move to COST was more about the economic model than technology alone

**High End Systems (>\$1M)**

- Most/all Top 500 systems
- Custom SW & ISV apps
- Technology risk takers & early adopters

**Volume Market**

- Mainly capacity; <=150 nodes
- Mostly clusters; >50% & growing
- Higher % of ISV apps
- Fast growth from commercial HPC; Oil & Gas, Financial services, Pharma, Aerospace, etc.

**Total market >\$10.0B in 2006**  
**Forecast >\$15.5B in 2011**

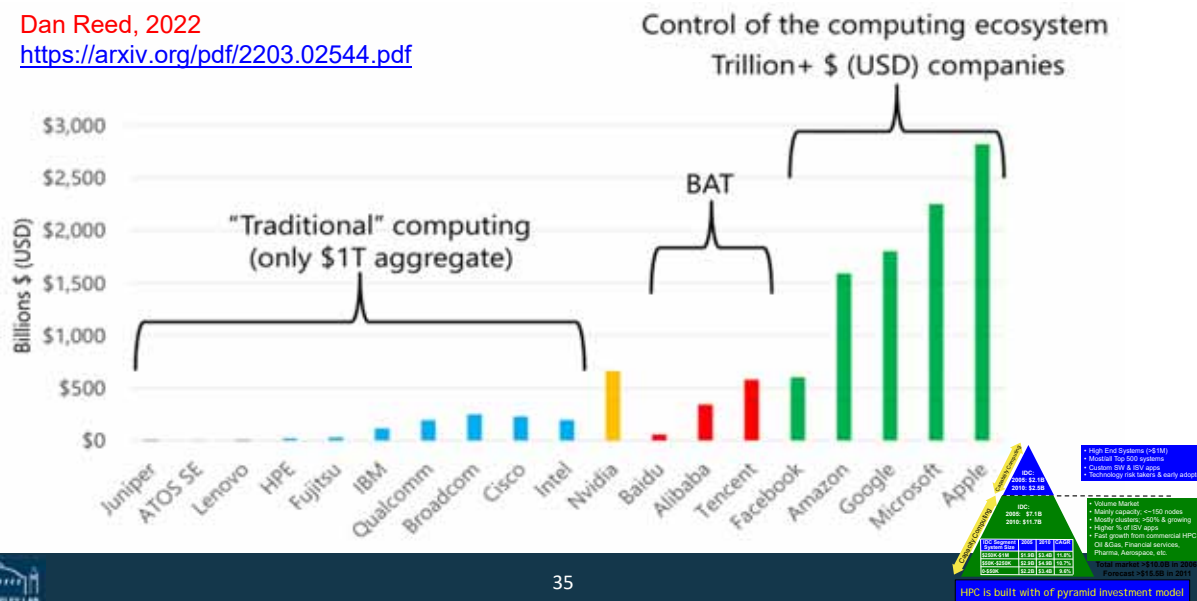
IDC Segment System Size	2005	2010	CAGR
\$250K-\$1M	\$1.9B	\$3.4B	11.8%
\$50K-\$250K	\$2.9B	\$4.9B	10.7%
0-\$50K	\$2.2B	\$3.4B	9.6%

**HPC is built with of pyramid investment model**

## It is not good enough anymore to understand the Technology Now we must also understand the market context

Dan Reed, 2022

<https://arxiv.org/pdf/2203.02544.pdf>

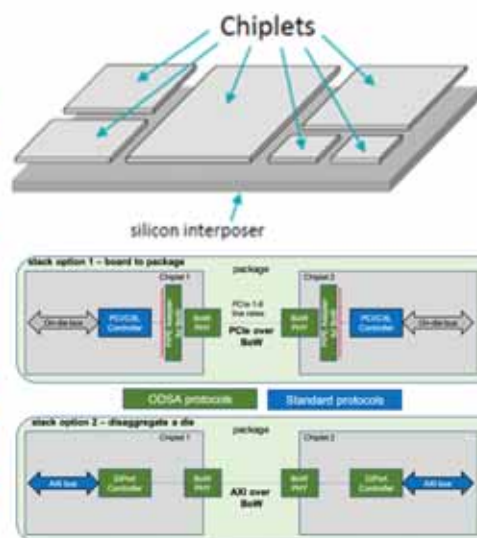
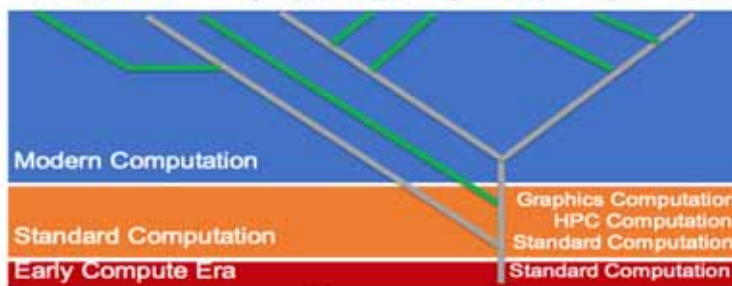


35

## Why? Domain specific Architectures driven by hyperscalers

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

Dharmesh Jani, Facebook –  
ODSA Workshop, Regional Summit, Amsterdam, Sep. 2019



3  
6

## Opportunity for HPC: New Economic Model

### Open Chiplets Marketplace is forming (ODSA and UClexpress)

- Licensable IP and assembly by 3<sup>rd</sup> party lowers that barrier
- Leverage the economic model being created by HyperScale

### Leverage this baseline and extend to support HPC

- Smaller incremental cost for HPC to “play”
- *HPC has become “too small to attack the city”*

### 80:20 Rule: Focus open efforts on what uniquely benefits HPC

- Build up a library of reusable accelerators for HPC.
- **Interoperability for sustainability:** *Interoperate with Arm IP for commercially supported IP where it exists and focus Open on the 20% that doesn't make commercial sense to license*



37