Post-Moore introduction

Jim Kowalkowski

μRetreat

19 April 2018

Post-Moore (After Moore's Law falls apart)

- When is this (or are we in it)?
  - Doomsday has been predicted for a long time
    - 2017 was one of the predictions.
  - Most realistic: Hits in 2020, but start seeing real affects around 2022-2023
    - Based on 4-5nm technology/fabrication limits
    - Will manufacturers even want to go to this level?
- How do we know about this?
  - Workshops like https://sites.google.com/view/pmes17/program
  - DOE Office of Science documents like this https://science.energy.gov/~/media/ascr/ascac/pdf/meetings/201612/ASCAC_BMoore_Susut.pdf

If you forgot Moore's Law: "Moore's law is the observation that the number of transistors in a dense integrated circuit doubles about every two years… Moore's law is an observation and projection of an historical trend and not a physical or natural law." - Wikipedia

Post-Moore (2)

- What does it mean? (not mutually exclusive categories)
  - Energy efficient computing
  - Exotic technology
  - Extreme heterogeneous computing
  - Processing in close proximity to peripheral systems
  - FPGAs everywhere
- Already see evidence of this depending on definition you like best
- Well-known contenders in the exotic technology realm
  - Quantum computers (the latest craze, includes D-Wave)
  - Neuromorphic Computing (C. Shuman gave a few talks here on the subject)
  - Micron's automata processors (Practically dead)
  - **Shared property:** *Very much unconventional programming here*

What is driving changes in computing architecture?

- DOE's ASCR program looks to be driving a good amount of future technology
- This recent slide provides a good summary of computing drivers

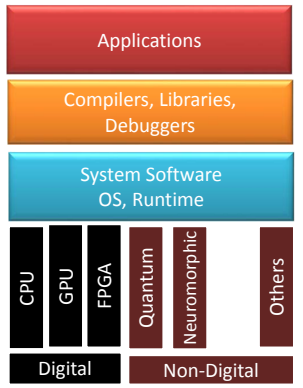
What does the Future Hold: Strategic Vision for ASCR's Research Program

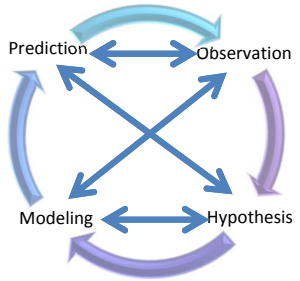
Emerging trends are pointing to a future that is increasingly

1. **Instrumented**: Sensors, satellites, drones, offline repositories
2. **Interconnected**: Internet of Things, composable infrastructure, heterogeneous resources
3. **Automated**: Complexity, real-time, machine learning
4. **Accelerated**: Faster &flexible research pathways for science & research insights

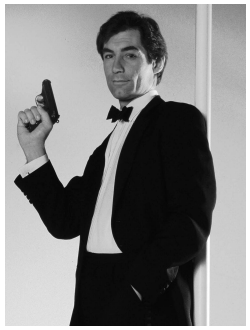
What is the role of ASCR's Research Program in transforming the way we carry out energy & science research?

1. **Post-Moore technologies**: Need basic research in new algorithms, software stacks, and programming tools for quantum and neuromorphic systems
2. **Extreme Heterogeneity**: Need new software stacks, programming models to support the heterogeneous systems of the future
3. **Adaptive Machine Learning, Modeling, & Simulation for Complex Systems**: Need algorithms and tools that support automated decision making from intelligent operating systems, in situ workflow management, improved resilience and better computational models.
4. **Uncertainty Quantification:** Need basic research in uncertainty quantification and artificial intelligence to enable statistically and mathematically rigorous foundations for advances in science domain-specific areas.
5. **Data Tsunami:** Need to develop the software and coordinated infrastructure to accelerate scientific discovery by addressing challenges and opportunities associated with research data management, analysis, and reuse.





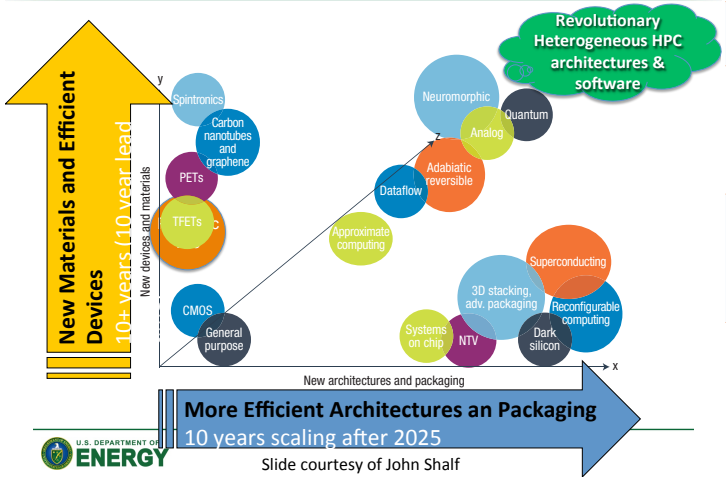
Post-Moore directions



Not that post-Moore …

Numerous Opportunities to Continue Moore's Law Technology!

(but winning solution is unclear)



Slide courtesy of John Shalf

Computing Beyond Moore's Law

TABLE 1. Summary of techology options for extending digital electronics.

Improvement Class	Technology	Timescale	Complexity	Risk	Opportunity
Architecture and software advances	Advanced energy management	Near-Term	Medium	Low	Low
	Advanced circuit design	Near-Term	High	Low	Medium
	System-on-chip specialization	Near-Term	Low	Low	Medium
	Logic specialization/dark silicon	Mid-Term	High	High	High
	Near threshold voltage (NTV) operation	Near-Term	Medium	High	High
3D integration and packaging	Chip stacking in 3D using thru-silicon vias (TSVs)	Near-Term	Medium	Low	Medium
	Metal layers	Mid-Term	Medium	Medium	Medium
	Active layers (epitaxial or other)	Mid-Term	High	Medium	High
Resistance reduction	Superconductors	Far-Term	High	Medium	High
	Crystaline metals	Far-Term	Unknown	Low	Medium
Millivolt switches (a better transistor)	Tunnel field-effect transistors (TFETs)	Mid-Term	Medium	Medium	High
	Heterogeneous semiconductors/strained silicon	Mid-Term	Medium	Medium	Medium
	Carbon nanotubes and graphene	Far-Term	High	High	High
	Piezo-electric transistors (PFETs)	Far-Term	High	High	High
Beyond transistors (new logic paradigms)	Spintronics	Far-Term	Medium	High	High
	Topological insulators	Far-Term	Medium	High	High
	Nanophotonics	Near/Far-Term	Medium	Medium	High
	Biological and chemical computing	Far-Term	High	High	High

U.S. DEPARTMENT OF ENERGY

Slide courtesy of John Shalf

Nowell – SSDBM, June 29, 2017

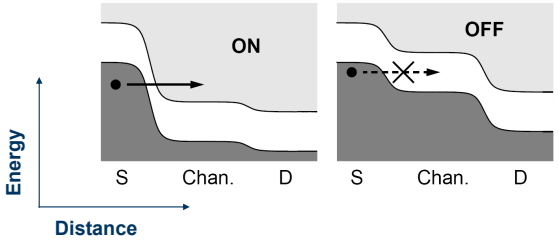
Fermilab

Memory – is it really changing?

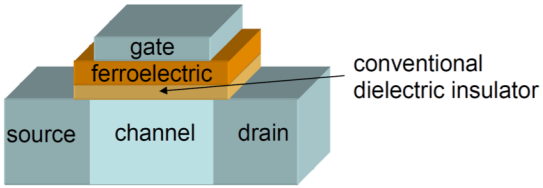
Ferroelectric (or Negative Capacitance) FET

- Initial Proposal: Salahuddin and Datta, *Nano Lett.*, 8 (2), 405-410, 2008.
- Key insight: If thickness of the insulating layers (conventional dielectric and ferroelectric) are adjusted properly, the structure acts as a step-up transformer, so the internal potential swing which gates channel current is larger than the external gate voltage swing.

The Band-to-Band Tunneling Field Effect Transistor or TFET



Conduction and valence bands are crossed in the ON state and uncrossed by the gate voltage in the OFF state. In the ON state, the energy distribution of injected carriers is limited (filtered) by the top of the valence band in the Source and by the bottom of the conduction band in the Drain.



13

New Devices and Architectures for Energy Efficient Computing

Thomas N. Theis

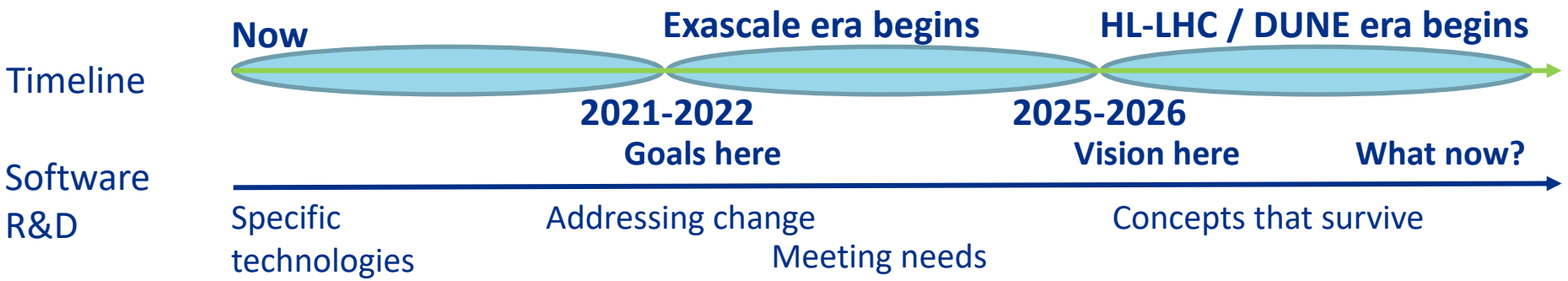
10

Fermilab

Other things to watch

- Architectures for the Post-Moore Era
  - https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=8013454
  - Near-memory acceleration
  - Processing-in-storage (Resistive CAM – Content Addressable Storage)
  - Intelligent memory

Software R&D Context



- Post-Moore: after 2022 and definitely after 2026