A Post-Moore’s Law World

Presented at the Looking Beyond CMOS Technology for Future HPC Workshop

Jeffrey A. Nichols
Associate Laboratory Director for Computing and Computational Sciences

Hanover, MD
April 5, 2016
Computing technology is at a disruptive inflection point

• Power limitation
  – Fixed clock rate
  – Reduced memory/compute ratio
  – Increased failure rate

• New architectures and new business model(s)
  – Integration of CPU, accelerator(s), memory controller, network interface on one chip
  – Need and ability for increased customization
  – Possible IP reuse (e.g., ARM) for noncritical components
  – Requires skills and business model of System on Chip manufacturers

Snapdragon, Qualcomm
President Obama issued an executive order creating the National Strategic Computing Initiative

- Collaborating agencies: DOE, DoD, NSF
- DOE is the lead for advanced simulation through a capable exascale computing program
- July 27th ASCAC meeting: Head of ASCR announced that a project office for the Exascale Computing Initiative was established at ORNL
- President’s FY17 budget request includes funding to prepare for post-Moore’s Law era

NSCI has 5 strategic themes:

1. Create systems that can apply Exaflops of computing power to Exabytes of data (Exa = $10^{18}$)
2. Keep the United States at the forefront of High Performance Computing (HPC) capabilities
3. Improve HPC application developer productivity
4. Make HPC readily available
5. Establish hardware technology for future HPC systems even after the limits of current semiconductor technology are reached (the “post-Moore’s Law era”)
President Obama’s FY17 budget request mentions *Beyond Moore’s Law*

- As noted in the NSCI, the era of silicon-based microchips advancing in accordance with Moore’s *Law* ....is nearing an end due to limits imposed by fundamental physics. ASCR will invest $12 million across research and facilities to understand the impacts these technologies may have on our applications. Beginning in FY 2017, the computer science and computational partnerships activities will invest $7 million to initiate new research efforts on technologies “Beyond Moore’s Law,” responding to the NSCI and recommendations made by the Secretary of Energy Advisory Board, to understand the challenges that these dramatically different technologies pose to DOE mission applications and to identify the hardware, software and algorithms that will need to be developed for DOE mission applications to harness these developing technologies.
Current state of supercomputing

- Pflops (> $10^{15}$ Flop/s) computing fully established with 81 systems.
- Three technology architecture possibilities or “swim lanes” are thriving.
  - Commodity (e.g. Intel)
  - Commodity + accelerator (e.g. GPUs) (104 systems)
  - Special purpose lightweight cores (e.g. IBM BG, ARM, Knights Landing)
- Interest in supercomputing is now worldwide, and growing in many new markets (over 50% of Top500 computers are in industry).
- Exascale ($10^{18}$ Flop/s) projects exist in many countries and regions.
- Intel processors largest share, 89% followed by AMD, 4%.
Computing architectures: Family tree

1980s
- Cray 1
- Cray YMP
- Cray C90

1990s
- Cray YMP
- ASCI Blue Mountain
- ASCI Blue Pacific
- ASCI Q
- ASCI White

2000s
- Earth Simulator
- ASCI Red
- Jaguar
- Red Storm
- BG/L
- BG/P
- Ascii Purple
- Roadrunner
- ASCI Blue
- ASCI Purple
- ASCI White
- ASCI Mountain
- ASCI Pacific
- ASCI Purple

CPU generations:
- GF
- TF
- PF
**Functional materials design**

**Objectives and Strategies**
- Design a carbon-based material with 3D macroscale structure and large surface area
- Theory and modeling to guide experiment
- Large-scale predictive electronic structure calculations

**Results and Impact**
- A macroscopic 3D porous nanotube material created by boron substitutional doping: *delivering new materials for cleaning oil spills, storing energy, ...*

---

*Computational Discovery*

*Synthesis/Characterization*

*Application/Demonstration*

**References**

Meeting challenges

“Equipping the materials community with the advanced tools and techniques to work across materials classes”

QMCPACK for high-accuracy electronic structure

LSMS for electronic structure calculations of $>10^4$ atoms

Emerging microscopies

Multiple spectra at each point in a map
Data growth

- Better decisions, better science, need better (or more) data

- The growth of data:
  - The rate exceeds “Moore’s Law”
  - > 80% of new data sources are unstructured
  - The proportion of unstructured data is increasing
Merging of HPC and data analytics

Future architectures will need to combine HPC and big data analytics into a single box

Apollo: Urika-GD
Graph Analytics

Helios: Urika-XA
BDAS (Hadoop, Spark)

OLCF’s Titan
Cray XK7

CADES Pods
Compute & Storage

Metis
Cray XK7

BEAM’s “BE Analyzer” tool displaying interactive 2D and 3D views of analyzed multi-dimensional data generated at ORNL’s Center for Nanophase Materials Sciences (CNMS)
Exascale applications will address key questions of national significance

<table>
<thead>
<tr>
<th>Domain</th>
<th>Key Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climate (BER)</td>
<td>What will the earth’s sea level and global temperature be in 2100? How will extreme weather patterns change? How can we best prepare to adapt? What mitigation scenarios are most influential and also still implementable?</td>
</tr>
<tr>
<td>Combustion (BES)</td>
<td>Can engine efficiency and pollutant emission be predicted for conventional and new fuels? Can new engines that exhibit a better compromise between maximal energy efficiency and minimal emissions be designed more quickly?</td>
</tr>
<tr>
<td>Nuclear Energy (NE)</td>
<td>What steps are possible to improve the efficiency, economics, and safety of the existing fleet? Can nuclear fuel be safely burned longer while withstanding severe accidents? What advanced nuclear fuel and reactor design concepts hold the most promise for deployment on the power grid?</td>
</tr>
<tr>
<td>Carbon Capture and Storage (FE)</td>
<td>What is necessary for safe and permanent carbon capture, storage, and utilization technologies to become technically and commercially feasible? How can the translation of these technologies to market be accelerated?</td>
</tr>
<tr>
<td>Wind Energy (EERE)</td>
<td>How can wind plant energy losses be reduced (e.g., by a few percent)? What performance improvements can be identified and implemented that will enable the economic viability of unsubsidized wind plants?</td>
</tr>
<tr>
<td>Magnetic Fusion Energy (FES)</td>
<td>Can tokamak plasma disruption physics be understood well enough to predict/avoid/mitigate their occurrence and associated deleterious effects? Can viable candidate plasma-facing components and technologies be identified? What is the optimum magnetic confinement configuration?</td>
</tr>
</tbody>
</table>
Exascale computing: Vision and plans

• Continuing leadership
  – Secure first benefits of sustained \((O)10^{18}\) operations per second and \((O)10^{18}\) bytes of data
  – Address next generation of scientific, engineering, and large-data problems
  – Enable extreme-scale computing: 1,000× capabilities of today’s computers with similar size and power footprint
  – **Set the US on a new trajectory of progress** toward a broad spectrum of computing capabilities over the succeeding decade

• Productive system
  – Usable by wide variety of scientists and engineers
  – “Easier” to develop software and manage system

• Based on marketable technology
  – Not a “one-off” system
  – Scalable, sustainable technology, exploiting economies of scale and trickle-bounce effect

• Deploy exascale computers in early 2020s
• Prepare for “beyond exascale”
Exascale Computing Project (ECP): Target system characteristics

- 20 pJ per average operation
  - 40× improvement over today's efficiency
- Billion-way concurrency (current systems have million-way)
- Ecosystem to support new application development and collaborative work, enable transparent portability, accommodate legacy applications
- High reliability and resilience through self-diagnostics and self-healing
- Programming environments (high-level languages, tools, …) to increase scientific productivity

<table>
<thead>
<tr>
<th>Operation</th>
<th>Energy (pJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>64-bit integer operation</td>
<td>1</td>
</tr>
<tr>
<td>64-bit floating-point operation</td>
<td>20</td>
</tr>
<tr>
<td>256 bit on-die SRAM access</td>
<td>50</td>
</tr>
<tr>
<td>256 bit bus transfer (short)</td>
<td>26</td>
</tr>
<tr>
<td>256 bit bus transfer (1/2 die)</td>
<td>256</td>
</tr>
<tr>
<td>Off-die link (efficient)</td>
<td>500</td>
</tr>
<tr>
<td>256 bit bus transfer(across die)</td>
<td>1,000</td>
</tr>
<tr>
<td>DRAM read/write (512 bits)</td>
<td>16,000</td>
</tr>
<tr>
<td>HDD read/write (32k bits)</td>
<td>( O(10^6) )</td>
</tr>
</tbody>
</table>

28 nm CMOS, DDR3

- Computation is almost “free” relative to communication (wrt energy)
**Exascale system architecture with a cap of $200M and 20MW**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>System peak</td>
<td>55 Pflop/s</td>
<td>150-300 Pflop/s</td>
<td>1 Eflop/s</td>
<td>~20x</td>
<td>?</td>
</tr>
<tr>
<td>Power</td>
<td>18 MW (3 Gflops/W)</td>
<td>10-20 MW (15-30 Gflops/W)</td>
<td>~20 MW (50 Gflops/W)</td>
<td>O(1)</td>
<td>?</td>
</tr>
<tr>
<td>System memory</td>
<td>1.4 PB (1.024 PB CPU + .384 PB CoP)</td>
<td>2.8-5.6 PB (NVRAM) 1.8-3.6 PB (DRAM+HBM)</td>
<td>32 - 64 PB</td>
<td>~50x</td>
<td>?</td>
</tr>
<tr>
<td>Node performance</td>
<td>3.43 TF/s (.4 CPU +3 CoP)</td>
<td>&gt; 40 TF/s</td>
<td>1.2 or 15TF/s</td>
<td>O(1)</td>
<td>?</td>
</tr>
<tr>
<td>Node concurrency</td>
<td>24 cores CPU + 171 cores CoP</td>
<td>TBD</td>
<td>O(1k) or 10k</td>
<td>~5x - ~50x</td>
<td>?</td>
</tr>
<tr>
<td>Node Interconnect BW</td>
<td>6.36 GB/s</td>
<td>23 GB/s (EDR)</td>
<td>200-400GB/s</td>
<td>~40x</td>
<td>?</td>
</tr>
<tr>
<td>System size (nodes)</td>
<td>16,000</td>
<td>~3400-6800</td>
<td>O(100,000) or O(1M)</td>
<td>~6x - ~60x</td>
<td>?</td>
</tr>
<tr>
<td>Total concurrency</td>
<td>3.12 M 12.48M threads (4/core)</td>
<td>TBD</td>
<td>O(billion)</td>
<td>~100x</td>
<td>?</td>
</tr>
<tr>
<td>MTTF</td>
<td>Few / day</td>
<td>TBD</td>
<td>Many / day</td>
<td>O(?)</td>
<td>?</td>
</tr>
</tbody>
</table>
Exascale Computing Project Roadmap

DOE Milestones

- CD-0
- CD-1/3a
- CD-2/3b
- CD-4

- Application Development
- Software Technology
- Hardware Technology
- System Build NRE
- Site Prep
- System expansion
- Testbeds & Prototypes
- Exascale Systems

Is this post-Moore's Law?
Roadblocks to reaching exascale

• Major Challenges are ahead for extreme computing
  – Parallelism
  – Hybrid
  – Fault Tolerance
  – Power
  – … and many others not discussed here

• Short term
  – Evolutionary CMOS solutions
  – 3D integration/packaging, deep memory hierarchy, photonics integration
  – Merge of HPC and big data

• Longer term
  – New architectures..
Performance development of Top 500

Data source: Jack Dongarra
Moore’s Law

- Moore’s Law refers to the observation by Gordon Moore that the number of transistors on a microprocessor doubles every 18-24 months.

- Moore’s Law has held up for past fifty years.
- Moore’s Law is slowing down due to limitations imposed by fundamental physics.

Source: https://www.elektormagazine.com/articles/moores-law
Near term solutions may be Silicon based

- Three dimensional integration and packaging
  - Intel’s Knights Landing, integrated memory, integrated fabric, parallel performance
  - Vertically integrated ecosystem: System Scalable framework

- Silicon photonics
  - for fast data movement: use silicon as an optical medium, optical and electronic components are connected on a single chip

“Today, optics is a niche technology. Tomorrow, it's the mainstream of every chip that we build”, --Pat Gelsinger, former Intel senior vice president , 2006.
Deeper memory and storage hierarchy

Node memory moving on package

Cold storage moving to disk
Primary storage moving to Flash
New technologies coming to bridge memory-Flash gap

PCM
ReRAM
STT-MRAM

3D Xpoint

On Node
Off Node

CPU
Near Memory (HBM)
Mid Memory (DRAM DIMMs)
Far Memory (NVDIMM)
Near Storage (SSD, NVM)
Far Storage (HDD)

Storage only?
The International Technology Roadmap for Semiconductors (ITRS)

End of Moore’s Law will spur innovation and lead to new architectures

- We still use Von-Neumann architecture
- Sustained growth of transistor technology allowed us to ignore architecture so far
- Most “Beyond CMOS” technology are not ready for near-term deployment..

### Technology vs. Architecture Diagram

- **CMOS**: Today
- **Beyond CMOS**: Beyond CMOS
- **Non-Von Neumann**
  - Neuromorphic
  - Quantum
- **Von Neumann**
Quantum computing

- Quantum computing codified in 1990’s to harness capabilities of quantum physics
  - Use “inherent parallelism” of quantum systems to provide exponential speed ups over select classical algorithms
- For 20 years, most quantum technologies have remained in the proof of concept phase
  - R&D with significant basic research
  - Diverse technology base with promise
- Current research addresses system-level concerns
  - **Microarchitecture**: instruction sets, layout
  - **Programming**: logical, physical
  - **Macroarchitecture**: technology, integration
  - **Performance**: costs, efficiency, stability

Superconducting chip from D-Wave Systems
Linear optical chip from Univ. Bristol
Diamond chip from Delft Univ. /UCSB
Superconducting chip Google/UCSB
Ion trap chip from NIST
Photonic QKD system from Id Quantique
Quantum computing: Expected payoffs

• Solve important problems faster
  – Although factoring has dominated the discussion, it is not the only example: Database search, discrete optimization, quantum simulation..

• Outstanding need to demonstrate speed ups
  – State of the art examples are still trivial
  – Implementation details *may* undermine speed ups
  – Crossover points *may* be onerous, e.g., too large
  – Engineering costs *may* be prohibitive

• Current research is addressing these concerns
  – Estimating computational resource costs
  – Designing strategies to reduce overhead
  – Extending algorithms to application domains

Performance comparison for computational chemistry algorithms,
Evaluating post-CMOS alternatives

- There are no evident replacement technologies yet
- Post-CMOS alternatives need to meet
  - Scalability criteria: must allow density increases and corresponding energy reduction
  - Signal to noise immunity (e.g. quantum computing)
  - Scalable manufacturability: implementation at industrial scale (e.g. carbon based materials, carbon nanotubes)
  - Cost efficiency: some solutions based on III-V semiconductors, optical computing, are too costly
  - Large scale demonstration: need to demonstrate end to end solution (e.g. lack of large scale cryogenic memory for superconducting computing)
HPC does, indeed matter

- Gravitational waves detected 100 years after Einstein's prediction
- Experiment aided by High Performance Computing
Summary

• The computing world is changing rapidly.

• Unpredictable dynamics will emerge with the end of Moore’s law.

• HPC may evolve towards a hybrid model, integrating emerging non-Von Neumann architectures, data analysis, and new applications.

• Delivering an ecosystem focused on the integration of computing and data into instruments of science and engineering.

• This ecosystem delivers important, time-critical science with enormous impacts.
Questions?
Backup Slides
## ASCR computing upgrades at a glance

<table>
<thead>
<tr>
<th>System attributes</th>
<th>NERSC Now</th>
<th>OLCF Now</th>
<th>ALCF Now</th>
<th>NERSC Upgrade</th>
<th>OLCF Upgrade</th>
<th>ALCF Upgrades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Edison</td>
<td>TITAN</td>
<td>MIRA</td>
<td>Cori 2016</td>
<td>Summit 2017-2018</td>
<td>Theta 2016</td>
</tr>
<tr>
<td>Planned Installation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System peak (PF)</td>
<td>2.6</td>
<td>27</td>
<td>10</td>
<td>&gt; 30</td>
<td>150</td>
<td>&gt;8.5</td>
</tr>
<tr>
<td>Peak Power (MW)</td>
<td>2</td>
<td>9</td>
<td>4.8</td>
<td>&lt; 3.7</td>
<td>10</td>
<td>1.7</td>
</tr>
<tr>
<td>Total system memory</td>
<td>357 TB</td>
<td>710TB</td>
<td>768TB</td>
<td>~1 PB DDR4 +</td>
<td>&gt; 1.74 PB</td>
<td>&gt;480 TB DDR4 +</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>High Bandwidth</td>
<td>DDR4 + HBM +</td>
<td>High Bandwidth On-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Memory (HBM) +1.5PB persistent memory</td>
<td>2.8 PB</td>
<td>Package Memory Local Memory and Persistent Memory</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>persistent memory</td>
<td></td>
</tr>
<tr>
<td>Node performance (TF)</td>
<td>0.460</td>
<td>1.452</td>
<td>0.204</td>
<td>&gt; 3</td>
<td>&gt; 40</td>
<td>&gt; 3</td>
</tr>
<tr>
<td>Node processors</td>
<td>Intel Ivy Bridge</td>
<td>AMD Opteron</td>
<td>64-bit PowerPC</td>
<td>Intel Knights Landing many core CPUs</td>
<td>Multiple IBM Power9 CPUs &amp; multiple Nvidia Voltas GPUs</td>
<td>Intel Knights Landing Xeon Phi many core CPUs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nvidia Kepler</td>
<td>A2</td>
<td>Intel Haswell CPU in data partition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System size (nodes)</td>
<td>5,600 nodes</td>
<td>18,688 nodes</td>
<td>49,152</td>
<td>9,300 nodes</td>
<td>1,900 nodes in data partition</td>
<td>~3,500 nodes</td>
</tr>
<tr>
<td>System Interconnect</td>
<td>Aries</td>
<td>Gemini</td>
<td>5D Torus</td>
<td>Aries</td>
<td>Dual Rail EDR-IB</td>
<td>Aries</td>
</tr>
<tr>
<td>File System</td>
<td>7.6 PB 168 GB/s, Lustre®</td>
<td>32 PB 1 TB/s, Lustre®</td>
<td>26 PB 300 GB/s GPFS™</td>
<td>28 PB 744 GB/s Lustre®</td>
<td>120 PB 1 TB/s GPFS™</td>
<td>10PB, 210 GB/s Lustre initial</td>
</tr>
</tbody>
</table>
The CNMS - Nanomaterials Theory Institute

Developing a fundamental understanding of control over physical and chemical properties to build responsive matter

- Understanding and rational tuning of transport (electron, spin, ion, molecule), reactivity and electronic structure
- Developing methodologies for theoretical and computational nanoscience to establish new capabilities and to enhance links with experiment
- Building the scientific foundation to study and design functional correlated electronic materials (such as superconductors)
- Advancing soft matter theory and simulation for understanding morphology, stability, dynamics, and properties of topologically complex multiblock and charged copolymers, brushes,

The NTI builds a tight connection to computational sciences
Nanomaterials Theory Institute: Thomas Maier, Paul Kent, Gonzalo Alvarez, P. Ganesh, Miguel Fuentes-Cabrera, Tom Berlijn\(^1\), Rajeev Kumar, Xiaoguang Zhang, Mina Yoon, Jingsong Huang, Bobby G. Sumpter, Ariana Beste\(^2\), Mike Summers\(^3\)

\(^1\) Wigner Fellow, \(^2\) JICS associate, \(^3\) CSMD supported

- DMRG++
- DCA+
- QMCPACK
- SCFT
- Quantum Espresso
- Siesta
- VASP
- NWChem
- PSI4
- DFTB
- LAMMPS
- GROMACS
Computational Nanoscience End-Station:
Extreme scale scientific computing tools for nanomaterials
discovery/design/optimization

Scientific challenges and opportunities

Neutron scattering and other facilities

CNMS user projects

EFRCs/HUBS/and other Experimental theme areas

Computational Nanoscience End-Station

NERSC

OLCF

CNMS theme science

Advanced visualization