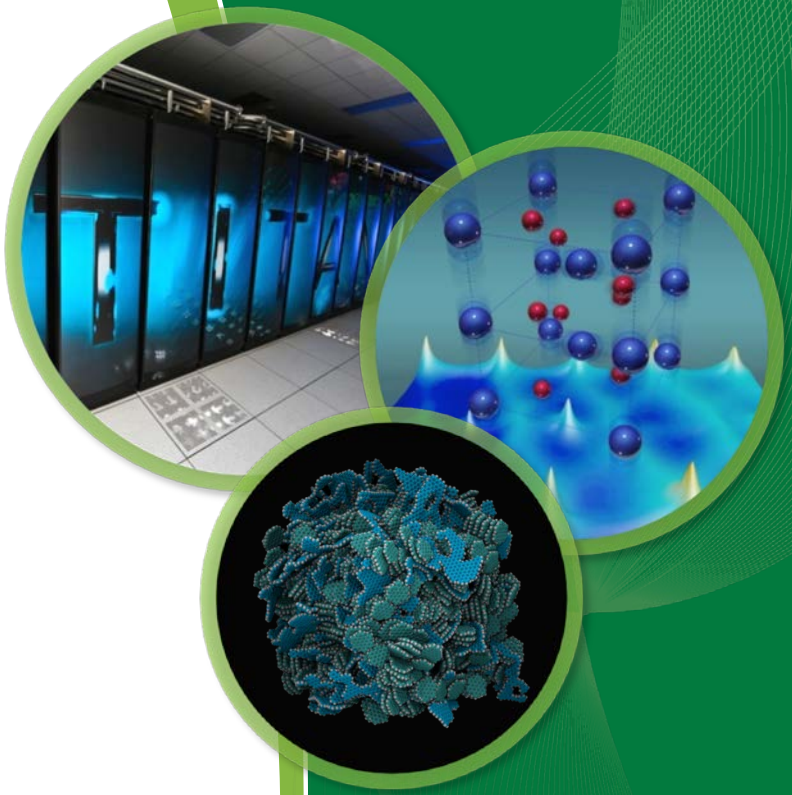
A Post-Moore's Law World

Presented at the
Looking Beyond CMOS Technology for Future HPC Workshop



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

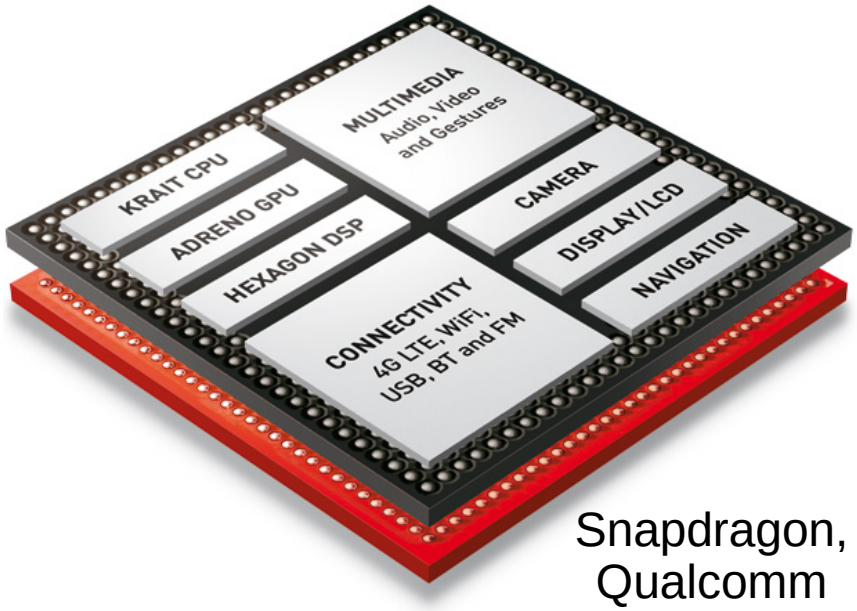
Hanover, MD
April 5, 2016

ORNL is managed by UT-Battelle
for the US Department of Energy

OAK RIDGE National Laboratory

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

OAK RIDGE National Laboratory

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

NSCI has 5 strategic themes:

Create systems that can apply Exaflops of computing power to Exabytes of data (Exa = 10^{18})

Keep the United States at the forefront of High Performance Computing (HPC) capabilities

Improve HPC application developer productivity

Make HPC readily available

Establish hardware technology for future HPC systems even after the limits of current semiconductor technology are reached (the “post-Moore’s Law era”)

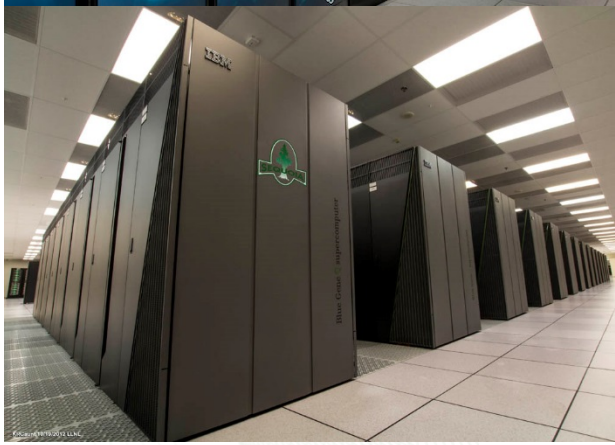
- Collaborating agencies: DOE, DoD, NSF
- DOE is the lead for advanced simulation through a capable exascale computing program
- July 27th ASCAC meeting: Head of ASCR announced that a project office for the **Exascale Computing Initiative** was established at ORNL
- President’s FY17 budget request includes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

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.

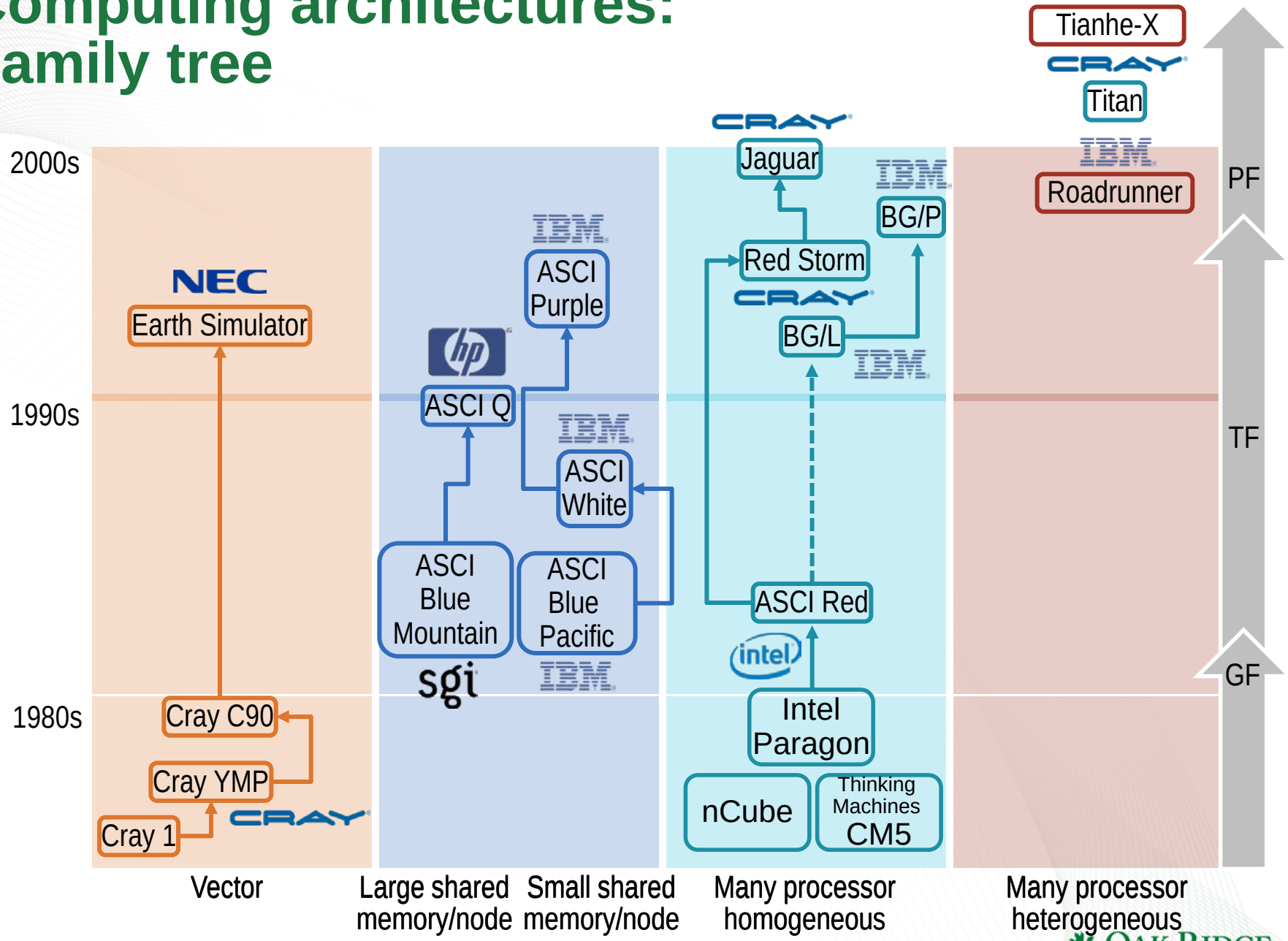
Current state of supercomputing

- 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%.



OAK RIDGE National Laboratory

Computing architectures: Family tree



OAK RIDGE National Laboratory

Functional materials design

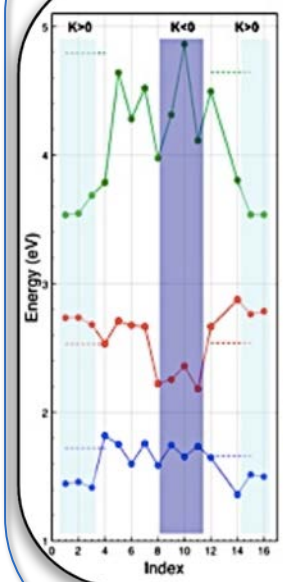
Objectives and Strategies

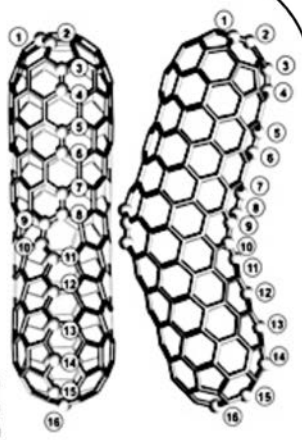
- Design a carbon-based material with 3D macroscale structure and large surface area
- Theory and modeling to guide experiment
- Large-scale predictive electronic structure calculations

Results and Impact

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

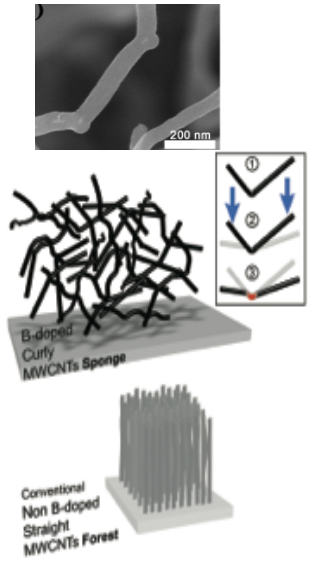
Computational Discovery



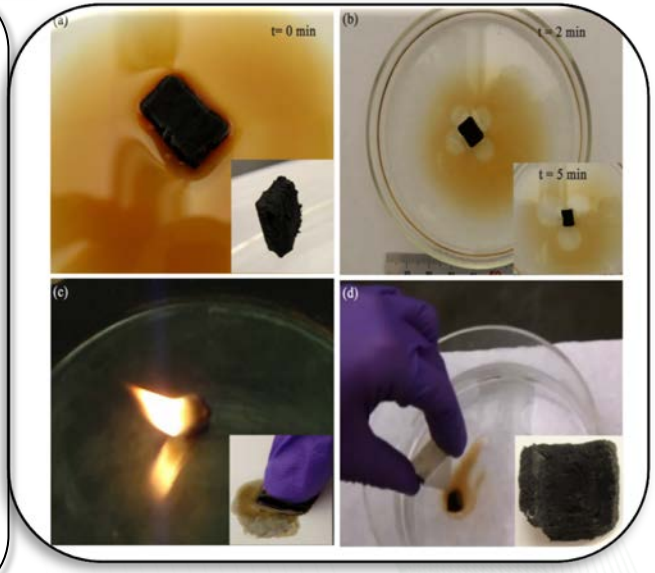


	Boron	Nitrogen	Sulfur
Pentagon	0.48 eV	-0.55 eV	-0.44 eV
Heptagon	-0.67 eV	0.03 eV	-0.28 eV

Synthesis/Characterization



Application/Demonstration



Rice University & Penn State

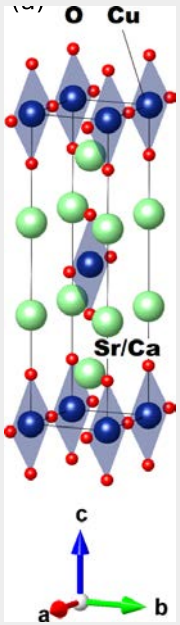
ACS Nano **4**, *369 (*2007*); IJQC* **109**, 97 (2009*); Nature Sci. Rep* **2**, 363 (2012)

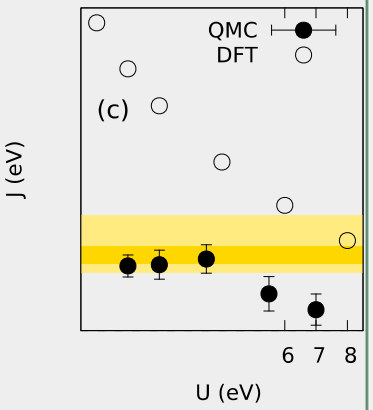
OAK RIDGE National Laboratory

Meeting challenges

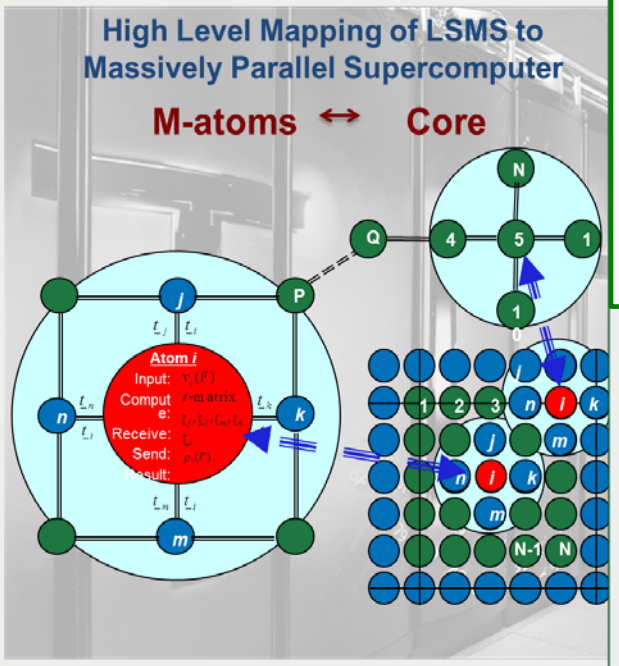
“Equipping the materials community with the advanced tools and techniques to work across materials classes”

QMCPACK for high-accuracy electronic structure

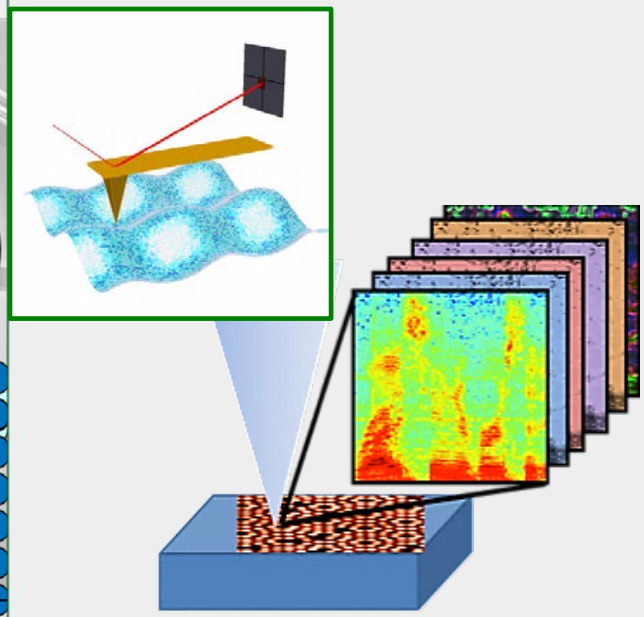




LSMS for electronic structure calculations of >10^4 atoms



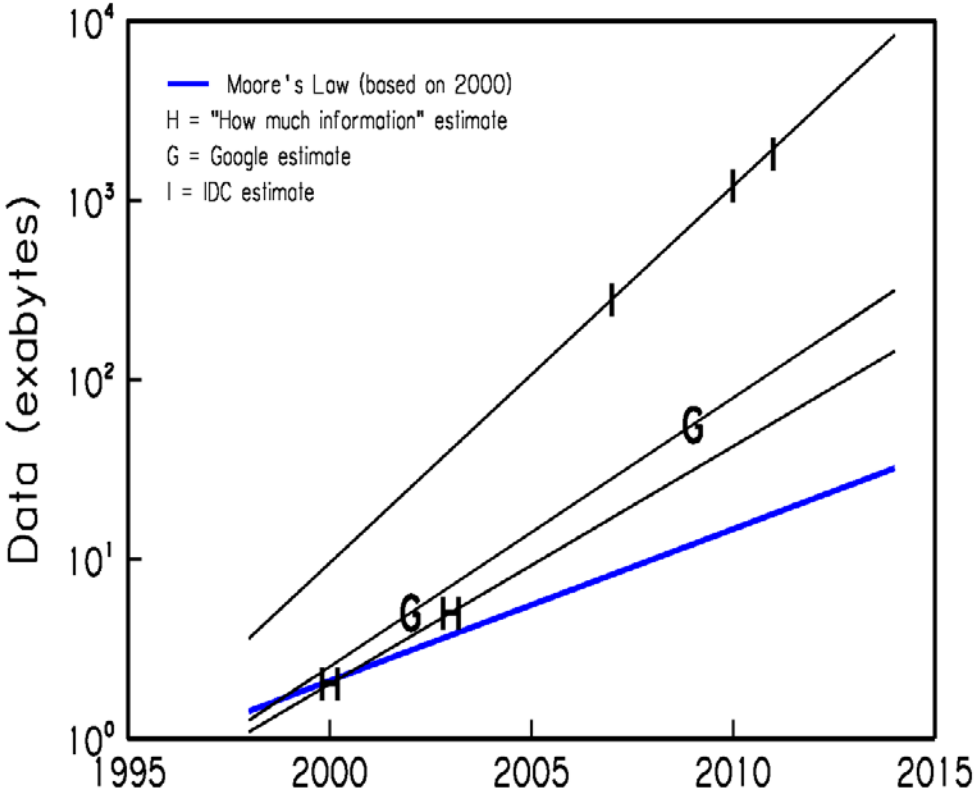
Emerging microscopies



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



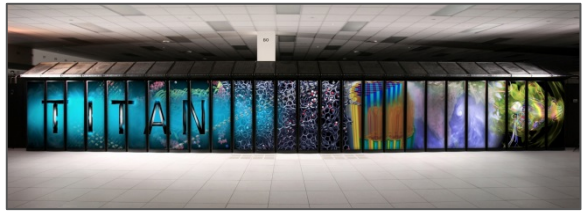
Merging of HPC and data analytics

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

Apollo: Urika-GD
Graph Analytics

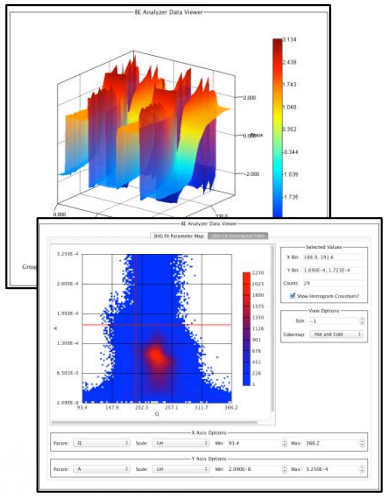
Helios: Urika-XA
BDAS
(Hadoop, Spark)

CADES Pods
Compute & Storage



OLCF's Titan
Cray XK7

Metis
Cray XK7



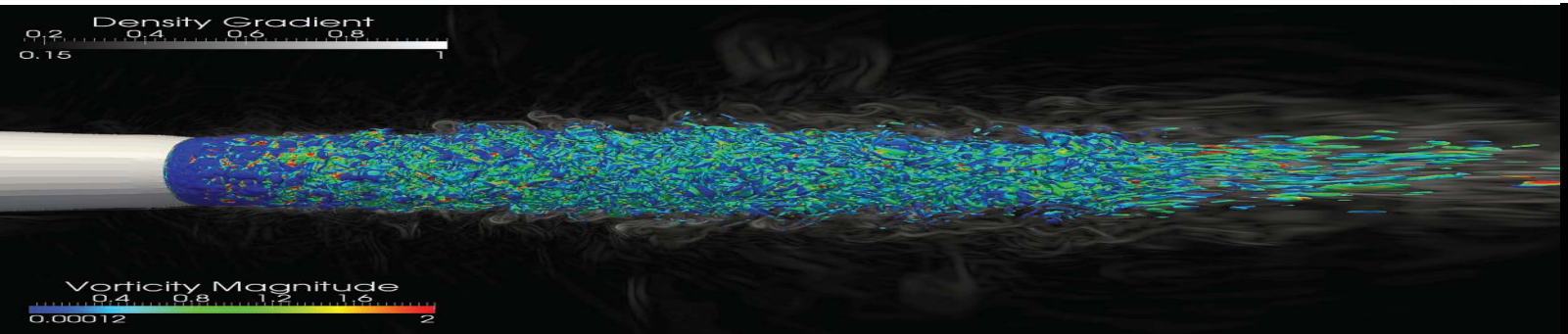
BEAM's "BE Analyzer" tool displaying interactive 2D and 3D views of analyzed multi-dimensional data generated at ORNL's Center for Nanophase Materials Sciences (CNMS)

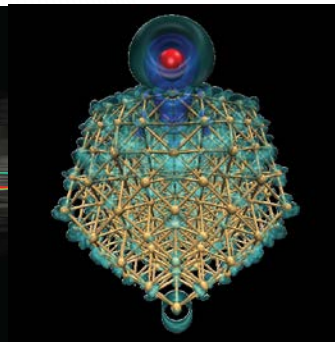
Exascale applications will address key questions of national significance

Domain	Key Questions
Climate (BER)	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?
Combustion (BES)	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 ?
Nuclear Energy (NE)	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?
Carbon Capture and Storage (FE)	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?
Wind Energy (EERE)	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?
Magnetic Fusion Energy (FES)	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?

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

Operation	Energy (pJ)
64-bit integer operation	1
64-bit floating-point operation	20
256 bit on-die SRAM access	50
256 bit bus transfer (short)	26
256 bit bus transfer (1/2 die)	256
Off-die link (efficient)	500
256 bit bus transfer(across die)	1,000
DRAM read/write (512 bits)	16,000
HDD read/write (32k bits)	$\mathcal{O}(10^6)$

28 nm CMOS, DDR3

- Computation is almost “free” relative to communication (wrt energy)

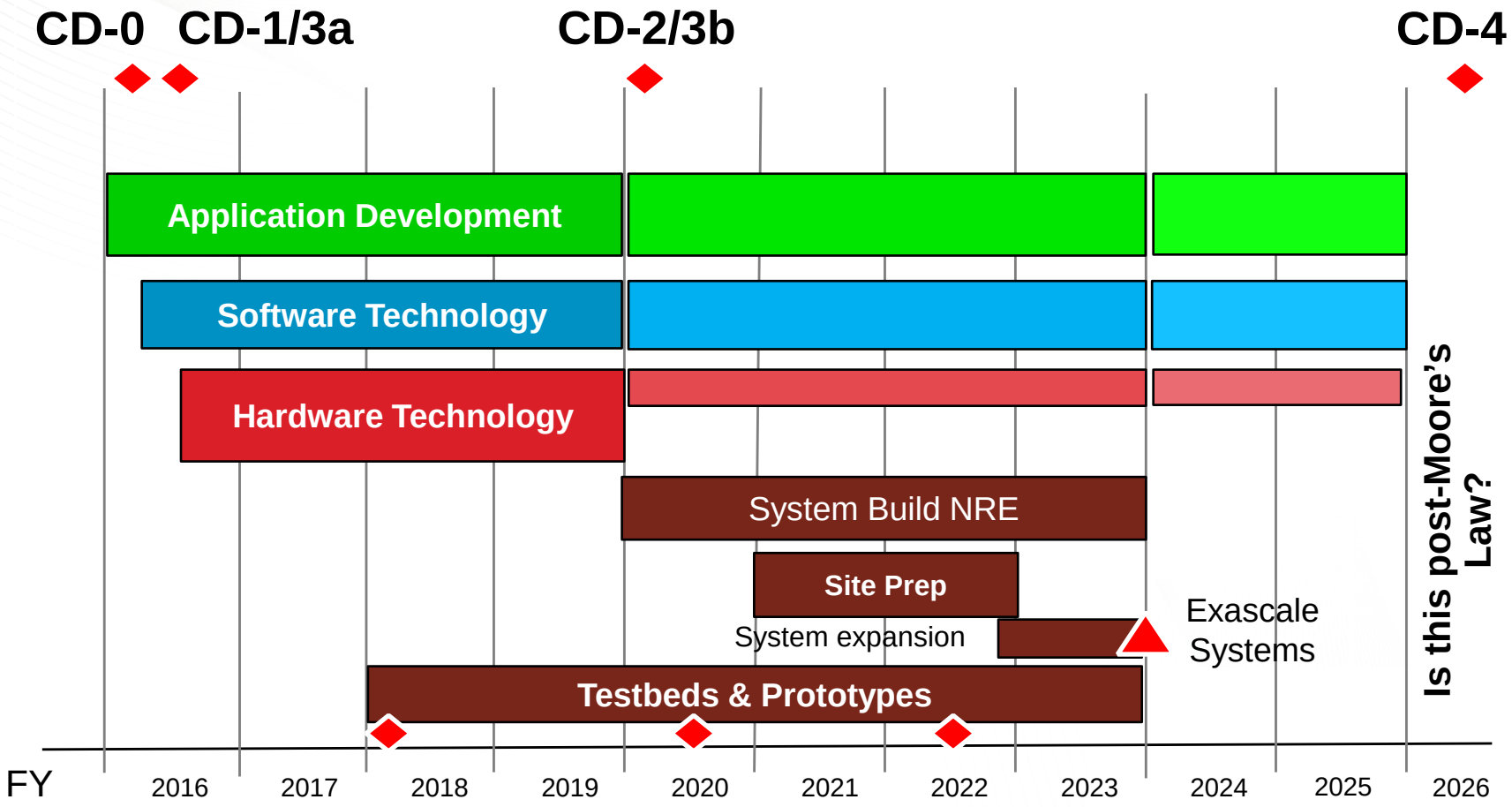
OAK RIDGE National Laboratory

Exascale system architecture with a cap of $200M and 20MW

Systems	2015 Tianhe-2	2017-2018 Summit	2020-2022	Difference Today & Exa	Beyond Exa
System peak	55 Pflop/s	150-300 Pflop/s	1 Eflop/s	~20x	?
Power	18 MW (3 Gflops/W)	10-20 MW (15 -30 Gflops/W)	~20 MW (50 Gflops/W)	O(1) ~15x	?
System memory	1.4 PB (1.024 PB CPU + .384 PB CoP)	2.8-5.6 PB (NVRAM) 1.8-3.6 PB (DRAM+HBM)	32 - 64 PB	~50x	?
Node performance	3.43 TF/s (.4 CPU +3 CoP)	> 40 TF/s	1.2 or 15TF/s	O(1)	?
Node concurrency	24 cores CPU + 171 cores CoP	TBD	O(1k) or 10k	~5x - ~50x	?
Node Interconnect BW	6.36 GB/s	23 GB/s (EDR)	200-400GB/s	~40x	?
System size (nodes)	16,000	~3400-6800	O(100,000) or O(1M)	~6x - ~60x	?
Total concurrency	3.12 M 12.48M threads (4/core)	TBD	O(billion)	~100x	?
MTTF	Few / day	TBD	Many / day	O(?)	?

Exascale Computing Project Roadmap

DOE Milestones

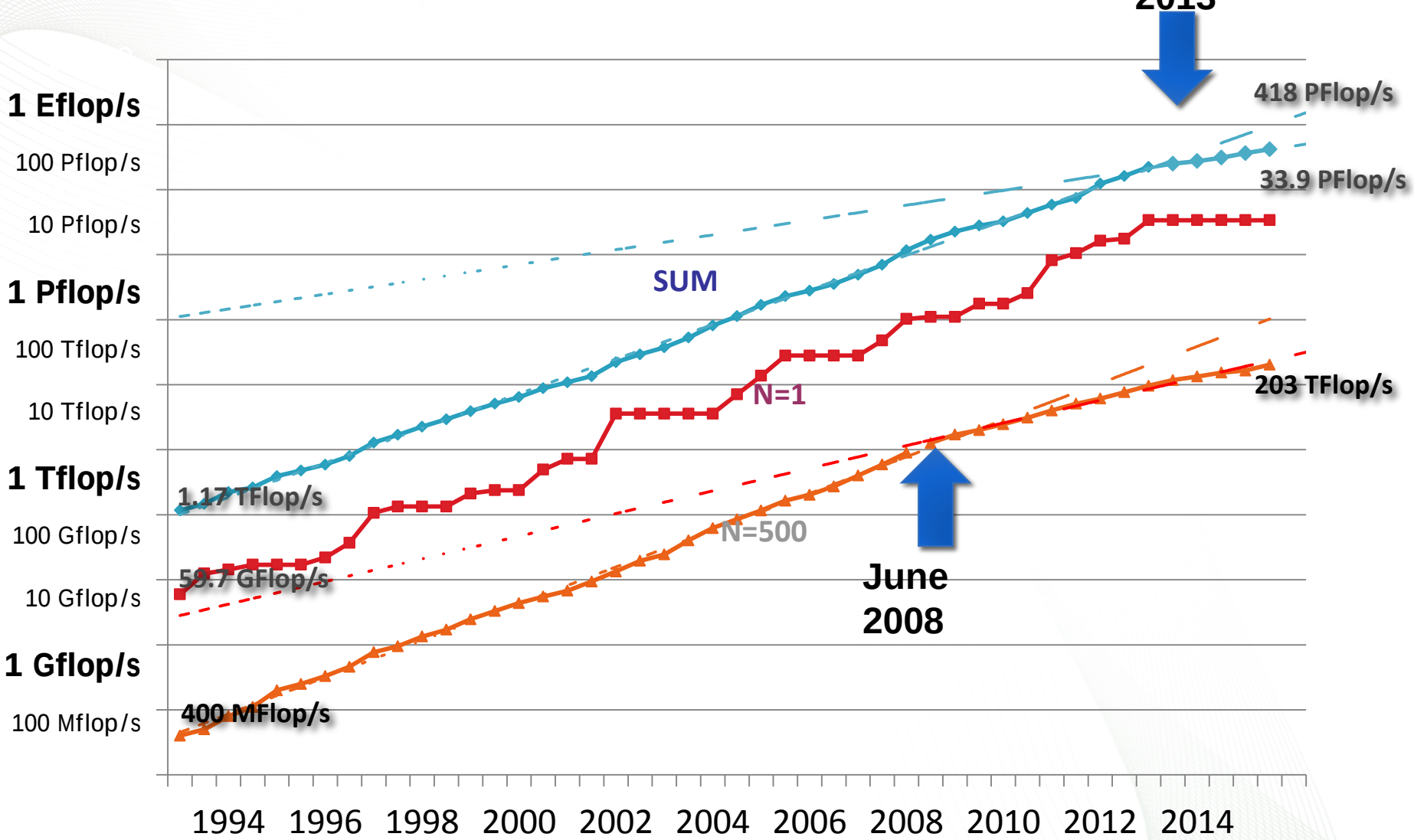


Roadblocks to reaching exascale

- 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..

OAK RIDGE National Laboratory

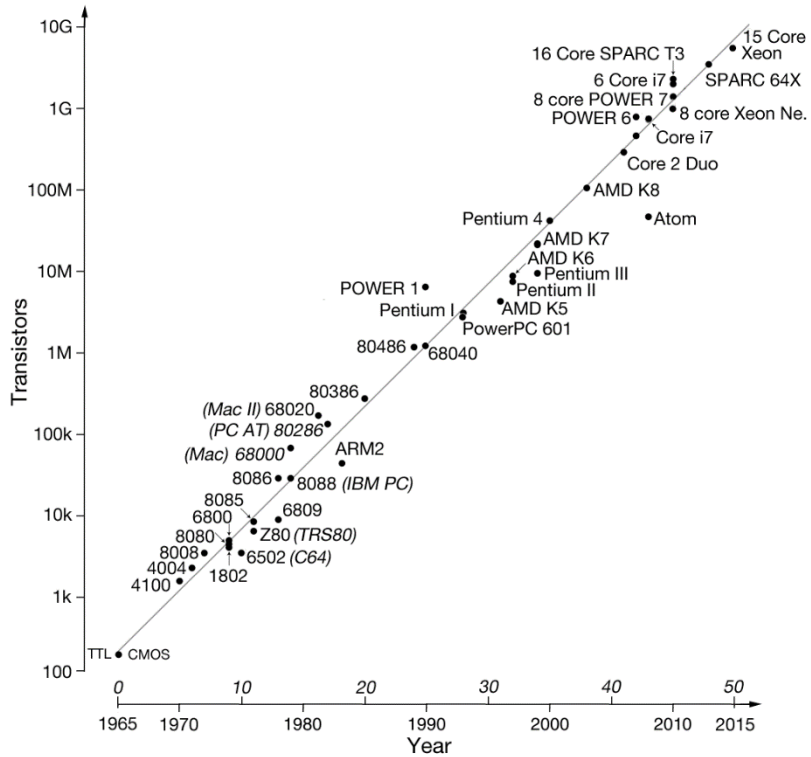
Performance development of Top 500



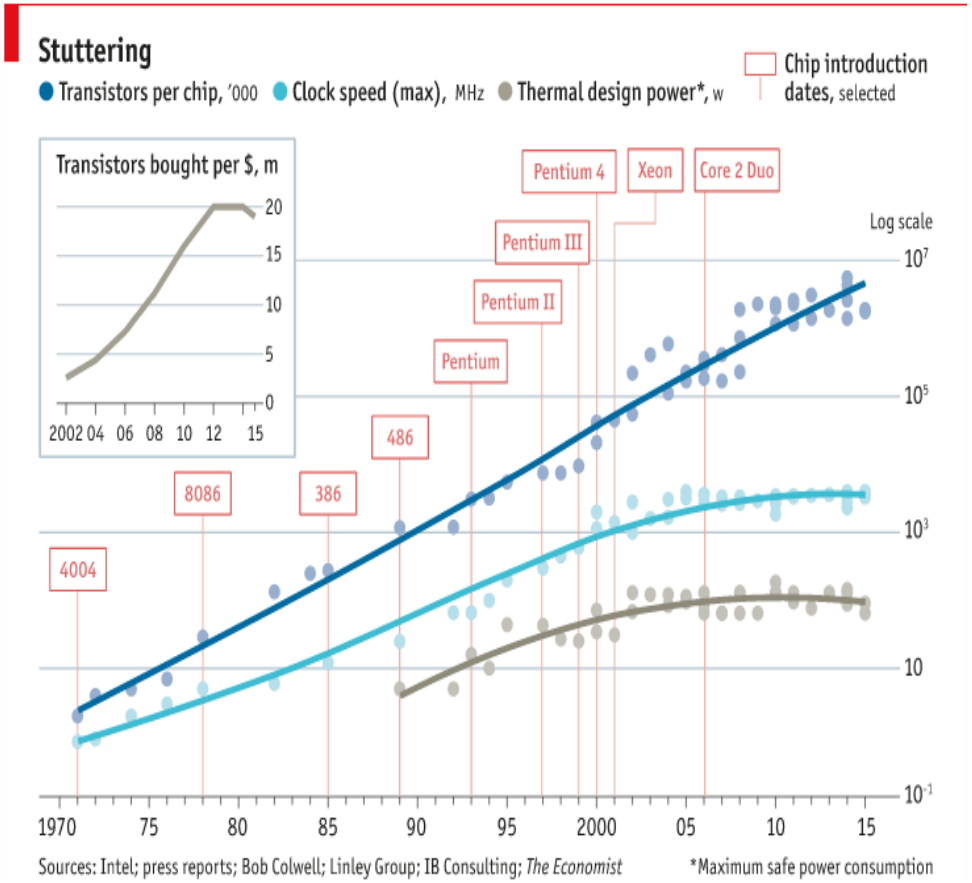
Data source: Jack Dongarra

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



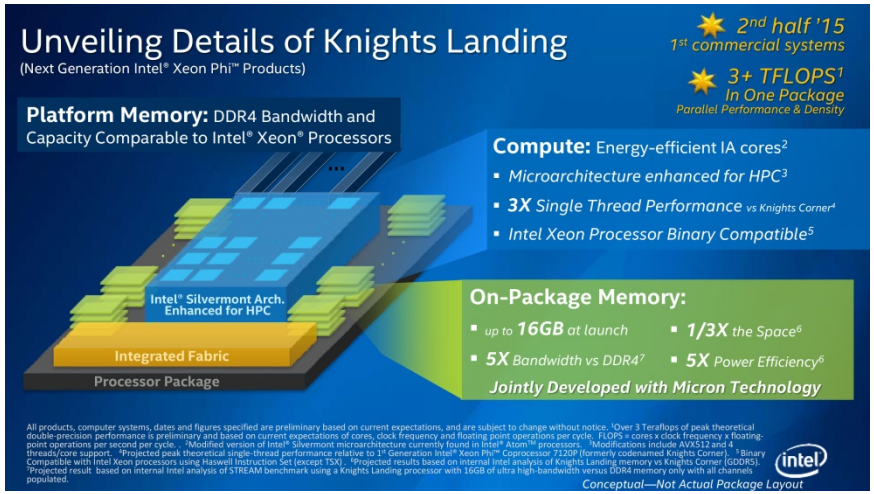
Source: https://www.elektormagazine.com/articles/moores-law

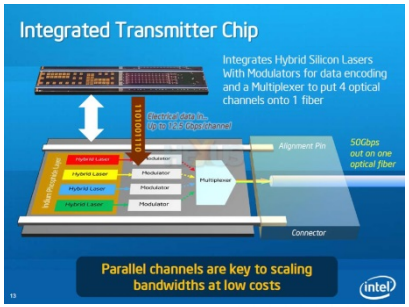


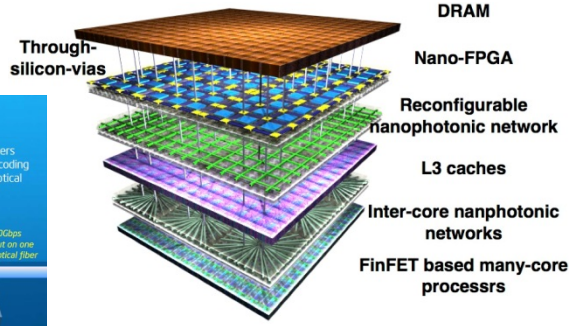
- Moore's Law has held up for past fifty years
- Moore's Law is slowing down due to limitations imposed by fundamental physics

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



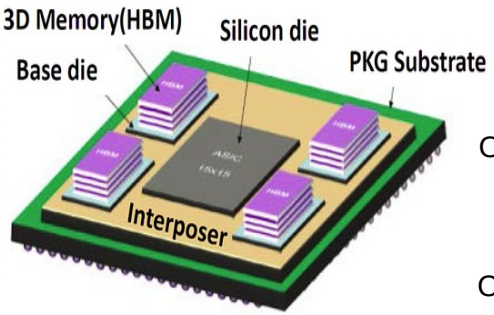


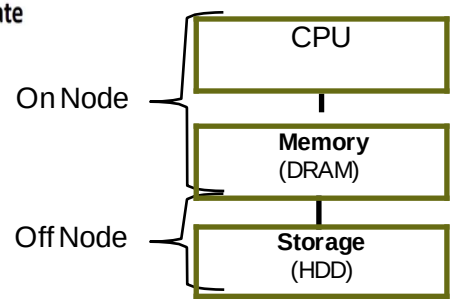


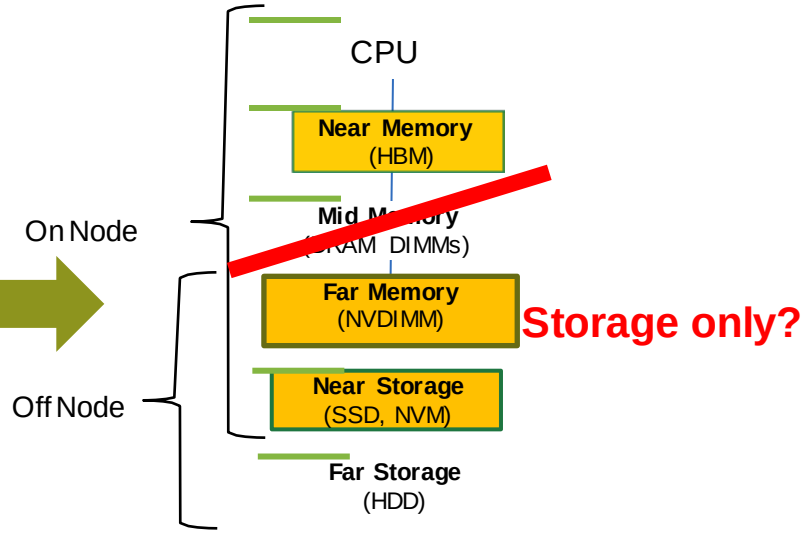
"Today, optics is a niche technology. Tomorrow, it's the mainstream of every chip that we build", --Pat Gelsinger, former Intel senior vice president , 2006.

OAK RIDGE National Laboratory

Deeper memory and storage hierarchy

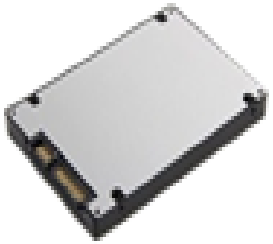




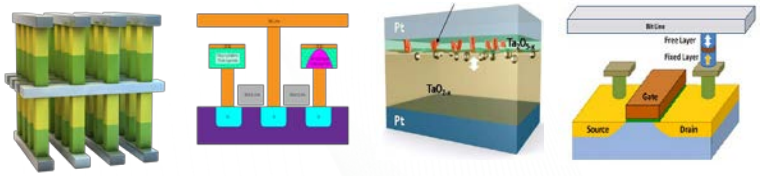


Node memory moving on package

Cold storage moving to disk



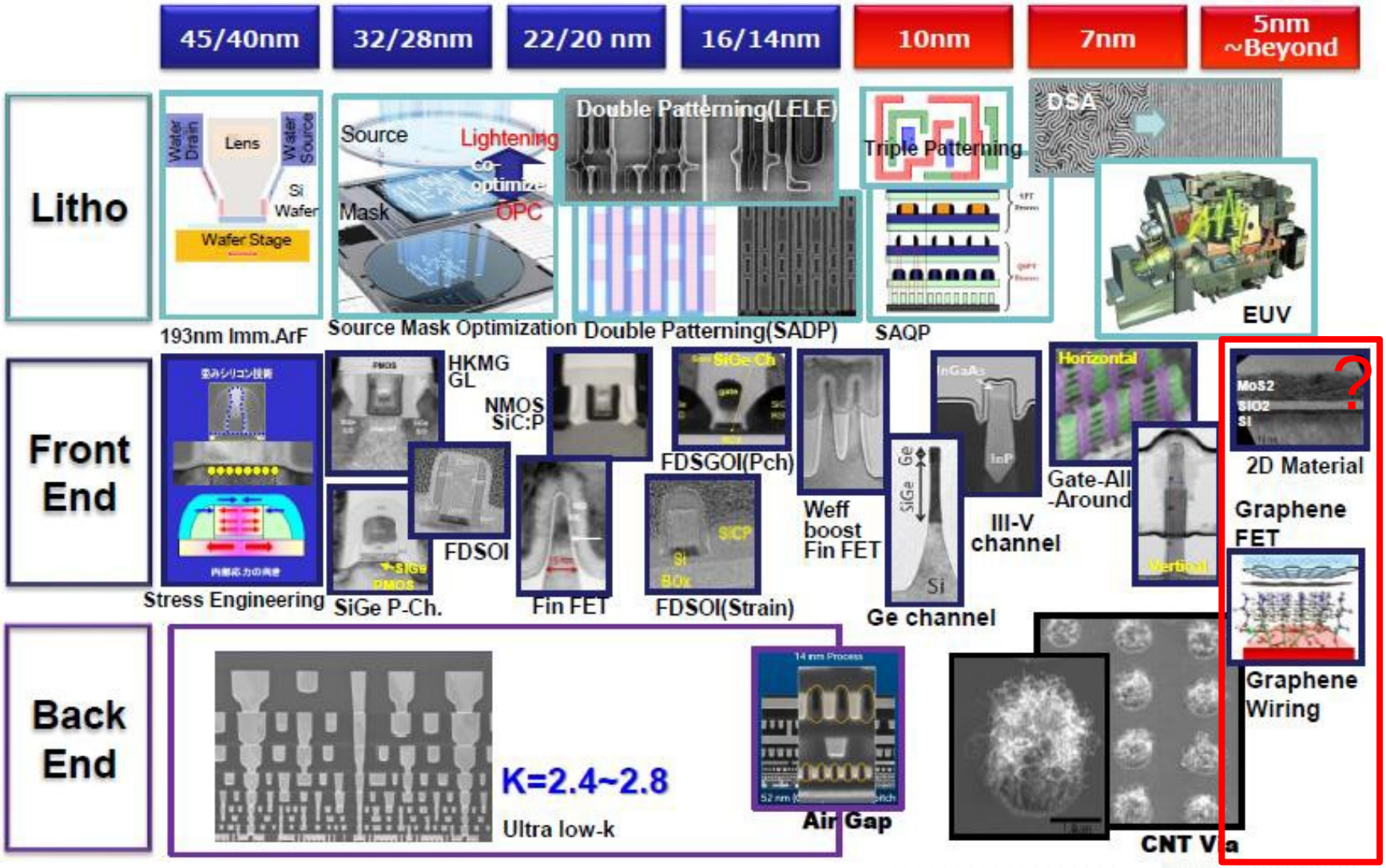
Primary storage moving to Flash



3D Xpoint **PCM** **ReRAM** **STT-MRAM**

New technologies coming to bridge memory-Flash gap

The International Technology Roadmap for Semiconductors (ITRS)

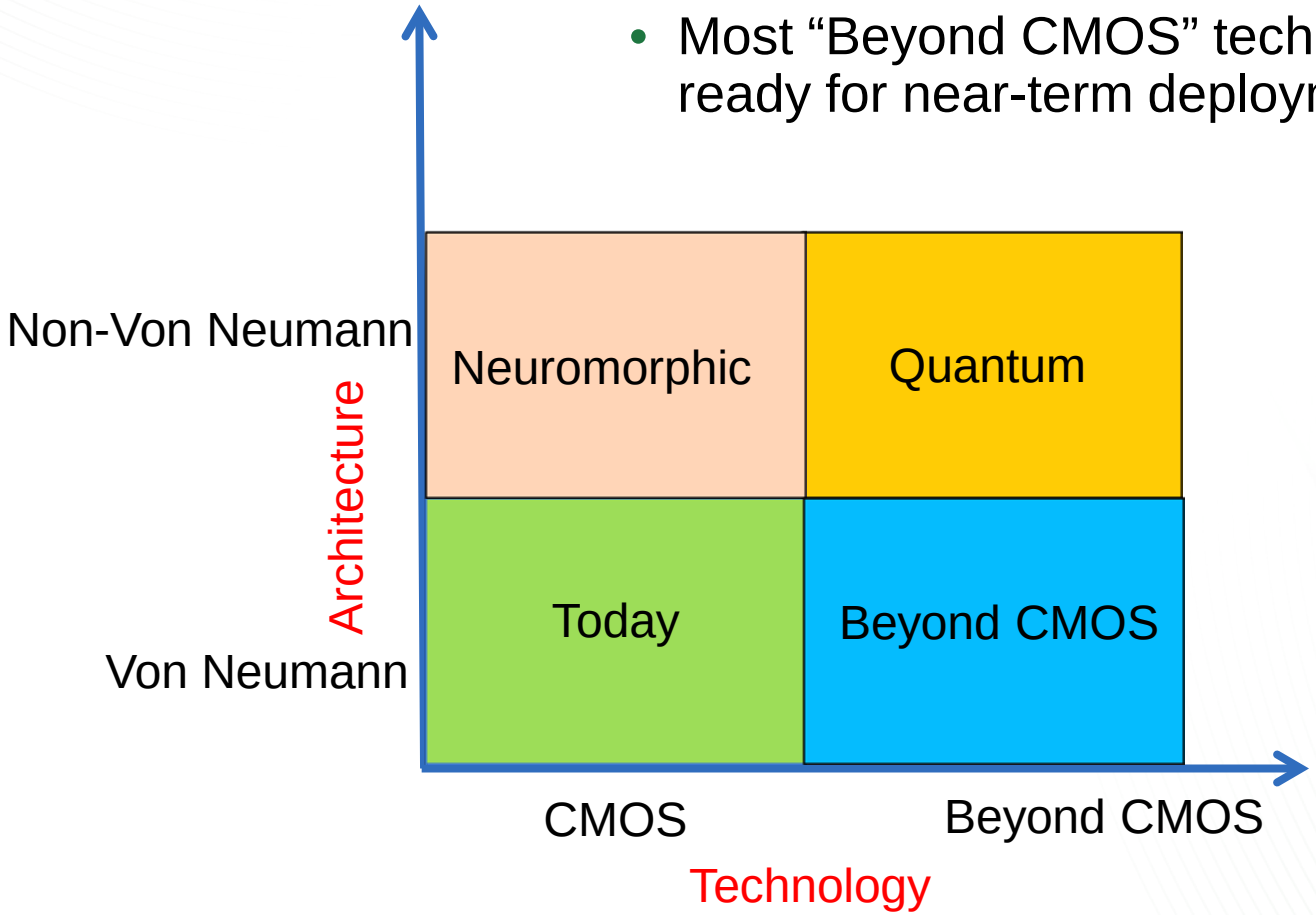


Courtesy of: Yuzo Fukuzaki, cited from M. Badaroglu, "More Moore scaling: opportunities and inflection points," ERD Meeting: Bridging Research Gap between Emerging Architectures and Devices, Feb 27, 2015

OAK RIDGE National Laboratory

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

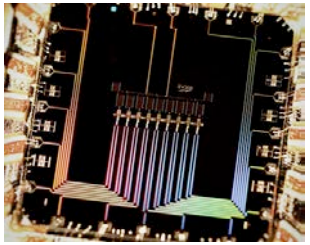
- We still use Von-Neumann architecture
- Sustained growth of transistor technology allowed us to ignore architecture so far
- Most "Beyond CMOS" technology are not ready for near-term deployment..



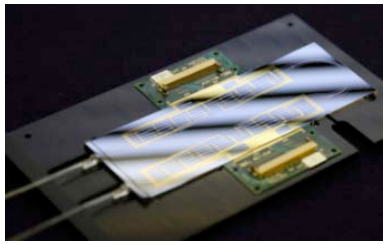
OAK RIDGE National Laboratory

Quantum computing

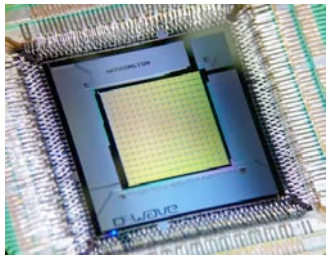
- **Quantum computing codified in 1990's to harness capabilities of quantum physics**
  - Use ''inherent parallelism'' of quantum systems to provide exponential speed ups over select classical algorithms
- **For 20 years, most quantum technologies have remained in the proof of concept phase**
  - R&D 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



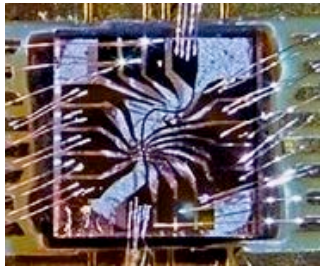
Superconducting chip Google/UCSB



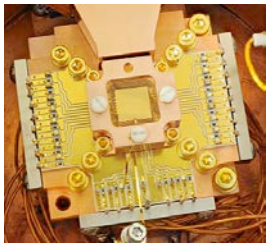
Linear optical chip from Univ. Bristol



Superconducting chip from D-Wave Systems



Diamond chip from Delft Univ. /UCSB



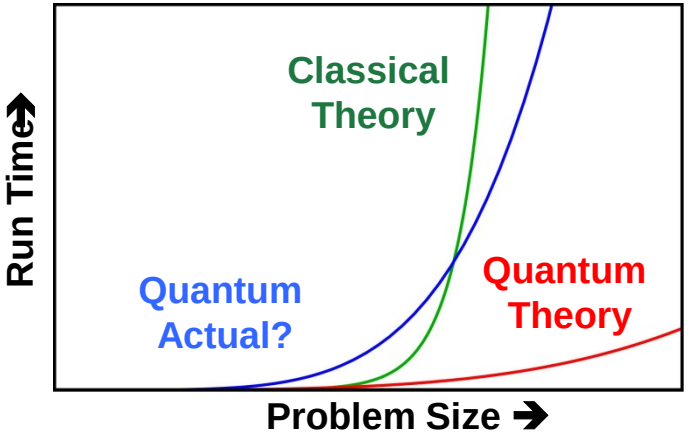
Ion trap chip from NIST

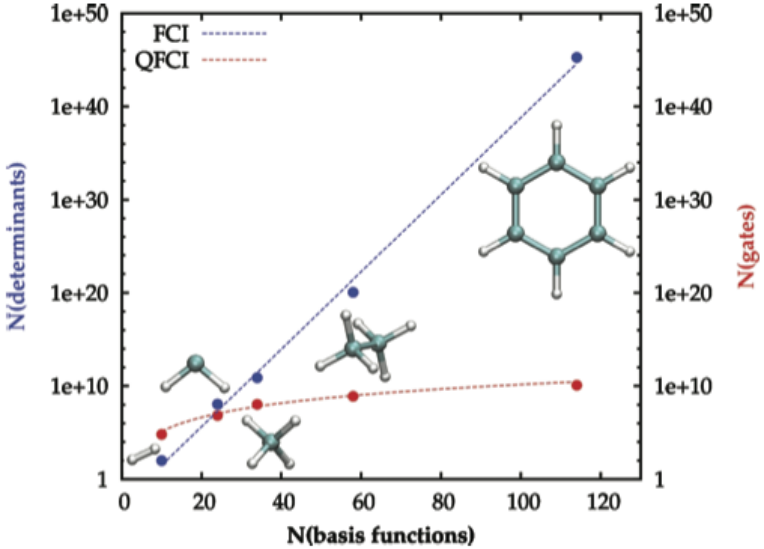
Photonic QKD system from Id Quantique

OAK RIDGE National Laboratory

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

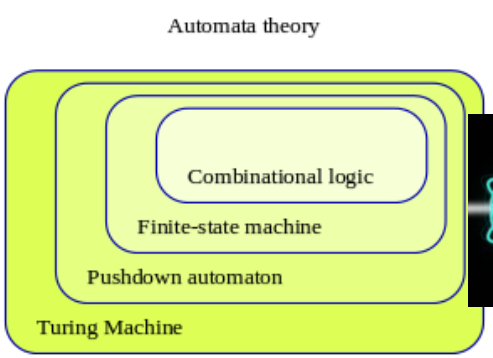


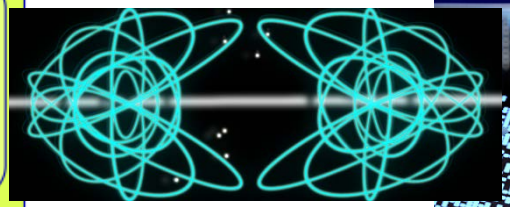


Performance comparison for computational chemistry algorithms,
Veis & Pittner, J. Chem. Phys. (2010)

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)





OAK RIDGE National Laboratory

HPC does, indeed matter

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

Selected for a Viewpoint in *Physics*

PRL **116,** 061102 (2016) PHYSICAL REVIEW LETTERS week ending 12 FEBRUARY 2016

Observation of Gravitational Waves from a Binary Black Hole Merger

B. P. Abbott *et al.*[*]

(LIGO Scientific Collaboration and Virgo Collaboration)

(Received 21 January 2016; published 11 February 2016)

On September 14, 2015 at 09:50:45 UTC the two detectors of the Laser Interferometer Gravitational-Wave Observatory simultaneously observed a transient gravitational-wave signal. The signal sweeps upwards in frequency from 35 to 250 Hz with a peak gravitational-wave strain of 1.0×10^{-21}. It matches the waveform predicted by general relativity for the inspiral and merger of a pair of black holes and the ringdown of the resulting single black hole. The signal was observed with a matched-filter signal-to-noise ratio of 24 and a false alarm rate estimated to be less than 1 event per 203 000 years, equivalent to a significance greater than 5.1σ. The source lies at a luminosity distance of 410^{+160}_{-180} Mpc corresponding to a redshift $z = 0.09^{+0.03}_{-0.04}$. In the source frame, the initial black hole masses are $36^{+5}_{-4} M_\odot$ and $29^{+4}_{-4} M_\odot$, and the final black hole mass is $62^{+4}_{-4} M_\odot$, with $3.0^{+0.5}_{-0.5} M_\odot c^2$ radiated in gravitational waves. All uncertainties define 90% credible intervals. These observations demonstrate the existence of binary stellar-mass black hole systems. This is the first direct detection of gravitational waves and the first observation of a binary black hole merger.

DOI: 10.1103/PhysRevLett.116.061102

I. INTRODUCTION

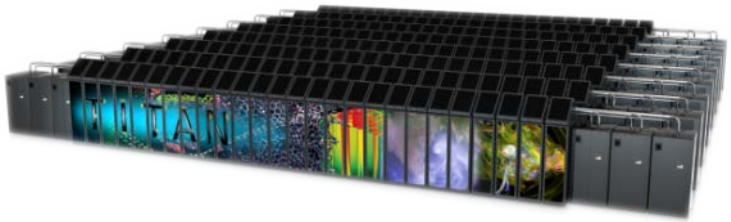
In 1916, the year after the final formulation of the field equations of general relativity, Albert Einstein predicted the existence of gravitational waves. He found that the linearized weak-field equations had wave solutions: transverse waves of spatial strain that travel at the speed of light, generated by time variations of the mass quadrupole moment of the source [1,2]. Einstein understood that gravitational-wave amplitudes would be remarkably

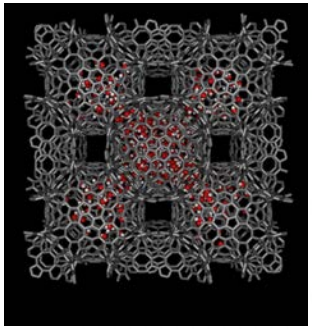
The discovery of the binary pulsar system PSR B1913+16 by Hulse and Taylor [20] and subsequent observations of its energy loss by Taylor and Weisberg [21] demonstrated the existence of gravitational waves. This discovery, along with emerging astrophysical understanding [22], led to the recognition that direct observations of the amplitude and phase of gravitational waves would enable studies of additional relativistic systems and provide new tests of general relativity, especially in the dynamic strong-field regime.

OAK RIDGE National Laboratory

Summary

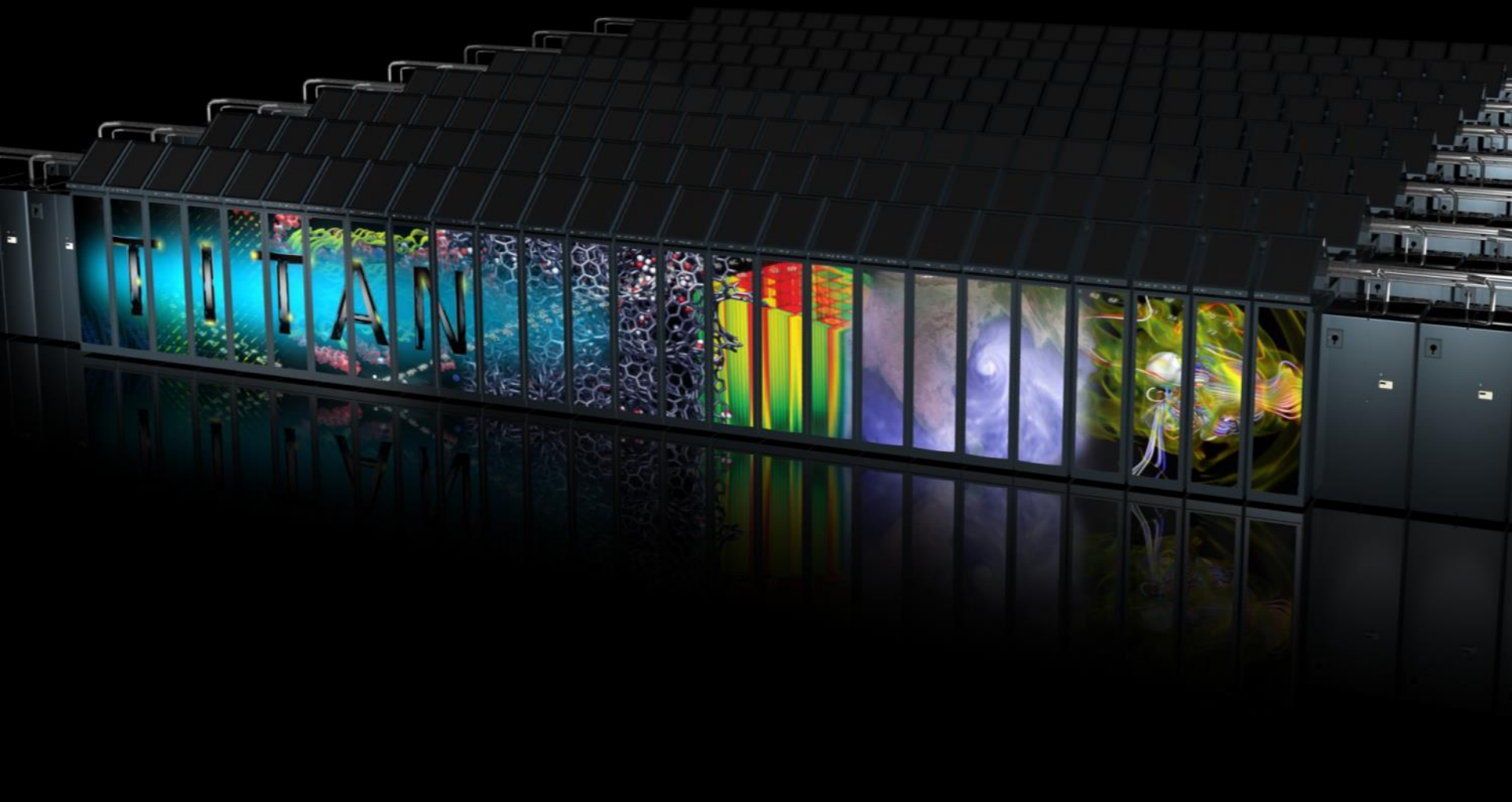
- The computing world is changing rapidly.
- Unpredictable dynamics will emerge with the end of Moore's law.
- HPC may evolve towards a hybrid model, integrating emerging non-Von Neumann architectures, data analysis, and new applications.
- Delivering an ecosystem focused on the integration of computing and data into instruments of science and engineering.
- This ecosystem delivers important, time-critical science with enormous impacts.





OAK RIDGE National Laboratory

Questions?



Backup Slides

ASCR computing upgrades at a glance

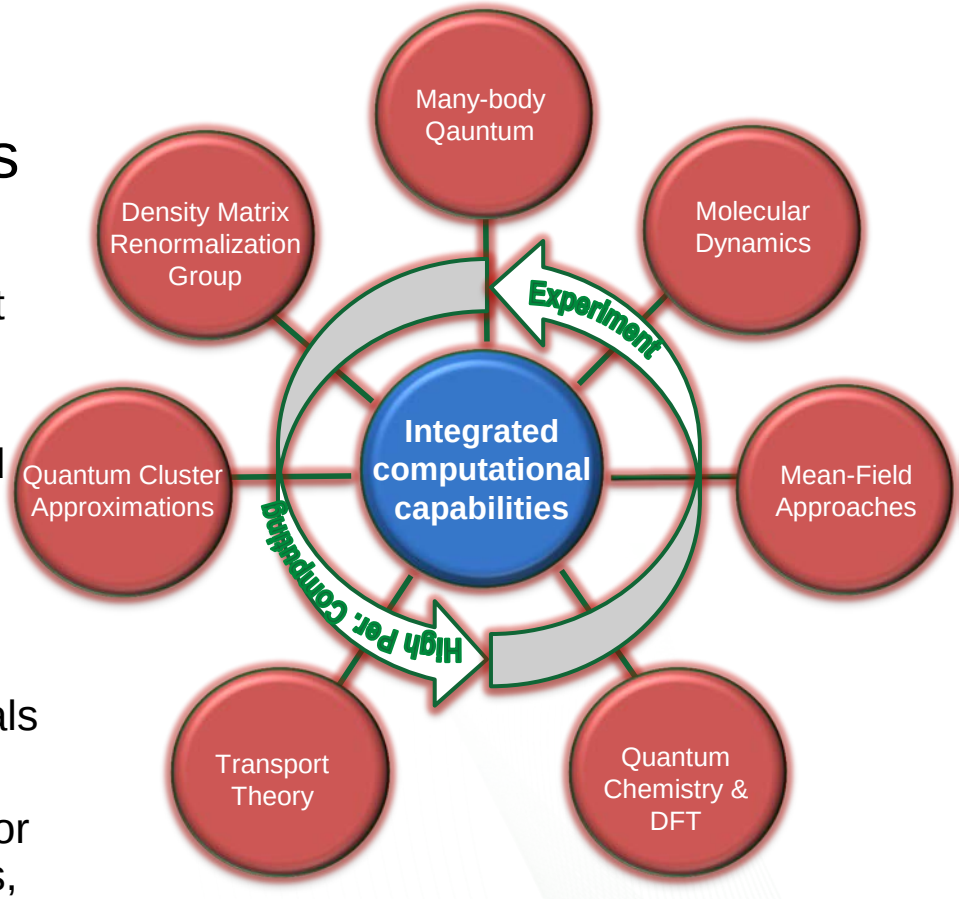
System attributes	NERSC Now	OLCF Now	ALCF Now	NERSC Upgrade	OLCF Upgrade	ALCF Upgrades	
Name Planned Installation	**Edison**	**TITAN**	**MIRA**	**Cori 2016**	**Summit 2017-2018**	**Theta 2016**	**Aurora 2018-2019**
System peak (PF)	2.6	27	10	> 30	150	>8.5	180
Peak Power (MW)	2	9	4.8	< 3.7	10	1.7	13
Total system memory	357 TB	710TB	768TB	~1 PB DDR4 + High Bandwidth Memory (HBM)+1.5PB persistent memory	> 1.74 PB DDR4 + HBM + 2.8 PB persistent memory	>480 TB DDR4 + High Bandwidth Memory (HBM)	> 7 PB High Bandwidth On-Package Memory Local Memory and Persistent Memory
Node performance (TF)	0.460	1.452	0.204	> 3	> 40	> 3	> 17 times Mira
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	Intel Knights Landing many core CPUs Intel Haswell CPU in data partition	Multiple IBM Power9 CPUs & multiple Nvidia Voltas GPUS	Intel Knights Landing Xeon Phi many core CPUs	Knights Hill Xeon Phi many core CPUs
System size (nodes)	5,600 nodes	18,688 nodes	49,152	9,300 nodes 1,900 nodes in data partition	~3,500 nodes	>2,500 nodes	>50,000 nodes
System Interconnect	Aries	Gemini	5D Torus	Aries	Dual Rail EDR-IB	Aries	2nd Generation Intel Omni-Path Architecture
File System	7.6 PB 168 GB/s, Lustre®	32 PB 1 TB/s, Lustre®	26 PB 300 GB/s GPFS™	28 PB 744 GB/s Lustre®	120 PB 1 TB/s GPFS™	10PB, 210 GB/s Lustre initial	150 PB 1 TB/s Lustre®

OAK RIDGE National Laboratory

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,



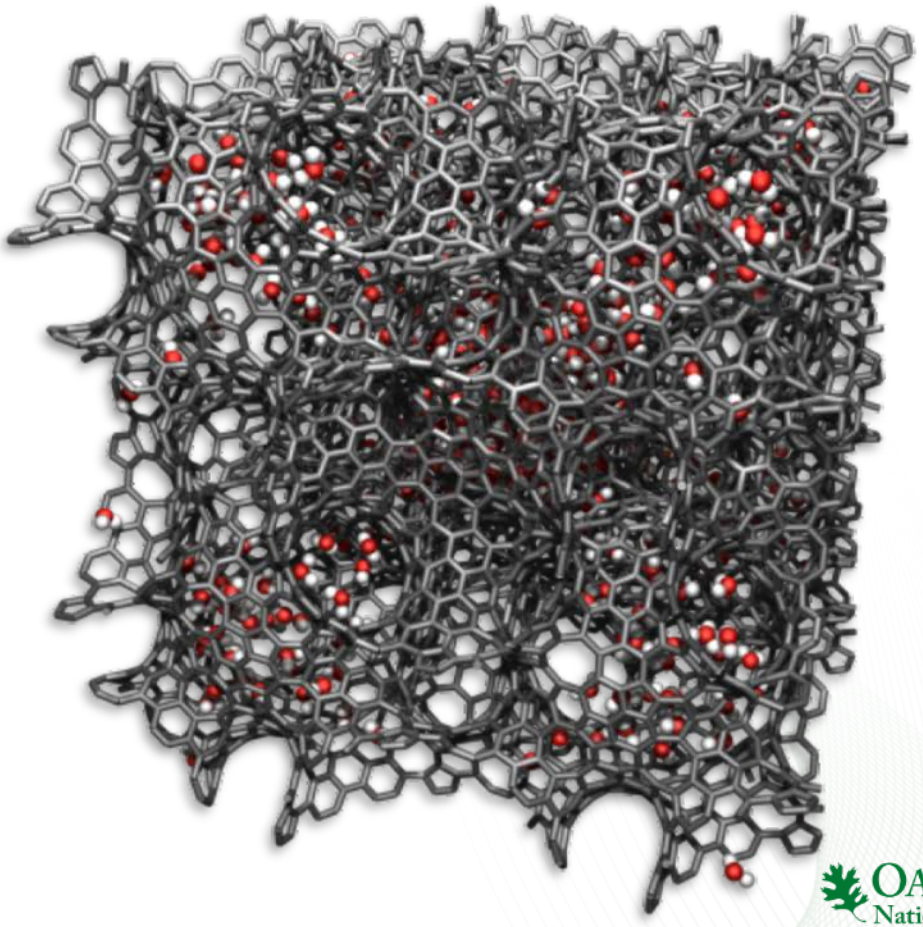
The NTI builds a tight connection to computational sciences

OAK RIDGE National Laboratory

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

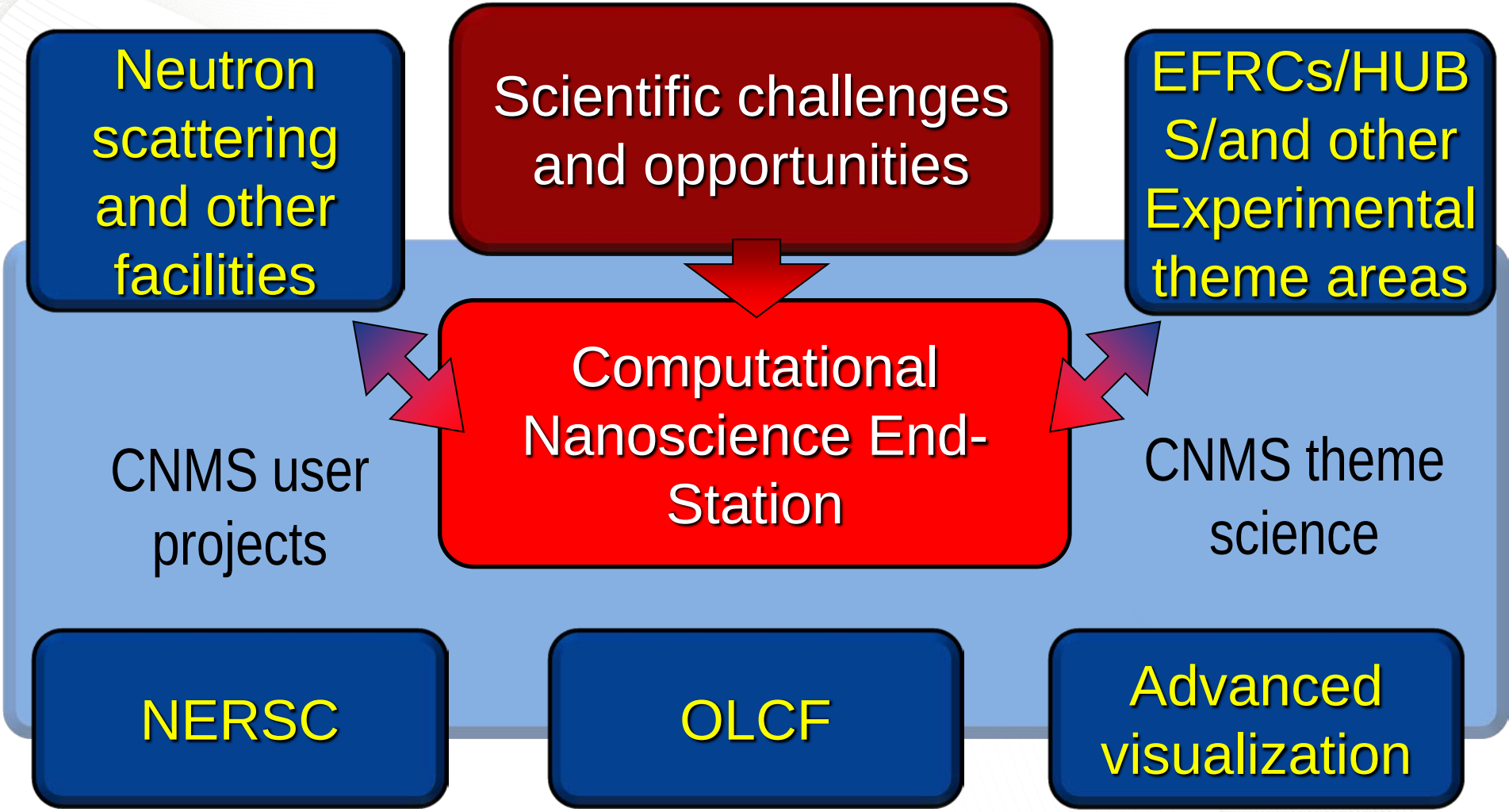
- **DMRG++**
- **DCA+**
- **QMCPACK**
- **SCFT**
- Quantum Espresso
- Siesta
- VASP
- NWChem
- PSI4
- DFTB
- LAMMPS
- GROMACS



OAK RIDGE National Laboratory

Computational Nanoscience End-Station:

Extreme scale scientific computing tools for nanomaterials discovery/design/optimization



OAK RIDGE National Laboratory