

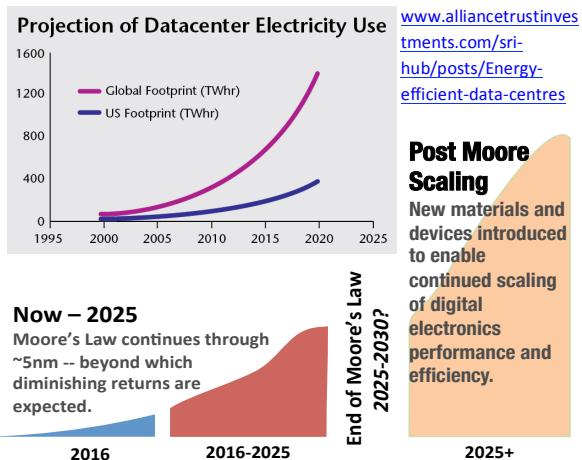
Solving the Information Technology Energy Challenge Beyond Moore's Law

Despite ever improving computing efficiency, information technology (IT) still represents the fastest growing consumer of energy. Uncontrolled, this demand would have significant implications on the U.S. energy landscape. In parallel we are facing a possible end to Moore's Law in the coming decade, which would either greatly accelerate the energy problem or significantly restrict U.S. computing growth severely threatening the nation's ability to solve important problems in science and national security. Semiconductor technology clearly has an essential role to play in future energy and technology security. In one example of projected IT energy growth: Cisco reports (Cisco Global Cloud Index, 2013–2018) that data center traffic (a useful metric for energy demand) is projected to grow at a compound annual rate (CAGR) of 23% from 2013-18. With no improvement in computing efficiency (i.e., the end of Moore's Law), one would expect this growth to be directly reflected in increased energy demand going from 91 billion kilowatt-hours in 2013 to 252 billion kilowatt-hours in 2018. Just meeting this increased demand alone would require 60 new 500-megawatt power plants, and will likely be exacerbated by the end of conventional Moore's Law technology scaling in the next decade.

DOE has a **unique opportunity** to create a new Public-Private partnership for translational basic/applied research to accelerate the development of energy efficient IT beyond the end of current roadmaps as well as maintaining an advanced U.S. manufacturing base in the economically critical semiconductor space. This partnership will allow DOE to leverage significant industry investments with the goal of enabling low-power computing impacting HPC, and broadly the Internet of Things.

To address this daunting problem both in the intermediate and long terms, a two-pronged approach is required: revolutionary for the long (10-20 year) term and evolutionary for the intermediate (10 year) term. Timing needs for the intermediate term will require an evolutionary approach based on achieving *manufacturing technology* advances allowing the continuation of Moore's Law with current CMOS technology and relying on new computing architectures. The long term solution requires *fundamental advances* in our *knowledge of materials* and pathways to control and manipulate information elements at the limits of energy flow, ultimately achieving 1 AttoJoule/operation, which would be 6 orders of magnitude smaller than today's devices. As we approach the longer term, we will require groundbreaking advances in *device technology going beyond CMOS (arising from fundamentally new knowledge of control pathways), system architecture, and programming models* to allow the energy benefits of scaling to be realized. Only with co-design covering this broad space and consideration of manufacturing challenges, can we expect to make progress in all areas cohesively to bring about real change to the IT energy outlook. In addition to containing the growth of IT related energy demand, the output of this work will provide a path to sustaining exponential growth in computing capabilities to enable new scientific discoveries, and maintain U.S. competitiveness in all segments of the computing market (from IoT, consumer electronics to datacenters, to supercomputing).

To meet the goals of broad societal impact, we must ensure transition of basic research to high volume manufacturing and even more fundamentally shape basic research from the start with an eye to manufacturability. This will be achieved through the development of a multi-lab ecosystem serving as a facility that can evaluate and demonstrate the manufacturing and energy savings feasibility of next generation technology options. Technologies will be rigorously evaluated for potential benefits on energy and implications on architecture, programming paradigms. The most promising technologies will be evaluated for issues around high volume manufacturing followed by ramp-up demonstration and getting them to deliver on the energy promises. This phase will depend heavily on identifying specific sets of materials, and pathways by which they can be scaled up where we will leverage the capabilities of the Materials Project and current HPC capabilities to accelerate the development through modeling and "virtual cycles of learning". Manufacturing feasibility would also include demonstration of whatever patterning technology would be needed to support the various technologies and scaling of those technologies. Delivering on this vision will require the integration of five major thrusts as follows:



1. The ***Fundamental Materials Science*** thrust will explore, discover, and design elemental excitations in condensed matter that can in new information channels. Those include system where the principles of quantum mechanical and competing spin-orbital and lattice interactions will enable new states of correlated matter and ordered phases. This will allow an exquisitely fine energy scale control of the states, such as reducing the energy consumption in a nanoscale switch from 1pJoule to 1AttoJoule. This thrust will use the Basic Energy Sciences “Grand Challenges” report as the guiding principle leading to the discovery of new materials, phenomena and processes that will ultimately lead us to a totally new paradigm of information processing, storage and communication.
2. The ***devices and CMOS Technology*** thrust will explore, identify, model, and demonstrate the new materials and devices for ultra-efficient computing. Examples include low voltage transistor concepts such as the TFET, photonic devices, spintronics, and novel memory devices. It requires advances in BES material modeling capability together with advances in algorithms from applied mathematics and data processing from ASCR research. *The goal is to use Materials Project and HPC high-throughput search for new materials to increase throughput for discovering new electronic materials and devices by a factor of 1000x over current methods.*
3. The ***Advanced Manufacturing and Integration*** thrust would leverage DOE’s expertise in EUV lithography and materials to develop novel nanomanufacturing methods, including EUV lithography, heterogeneous integration of advanced photonics and wide bandgap devices, and 3D stacking, are increasing density enabling memory layers on top of logic layers, and even multiple memory and logic layers interleaved. This radical change challenges assumptions embedded in current architectures, and would provide a new dimension to extend Moore’s Law scaling.
4. The ***Architecture*** thrust applies DOE ASCR/ASC expertise in advanced computing to exploit new device and materials systems and packaging technologies developed in the first two thrusts. Components include accelerators, on-chip wide-bandgap devices, photonic blocks, and emerging memory devices. The goal of these architectures is to remove overheads in current designs, as well as offer hardware and thus more efficient support for important functionality for security and resiliency.
5. The ***Programming models*** thrust seeks to create new paradigms integrated with the new systems that define how application designers interact with the machine. Existing programming models are designed with old architectures in mind. New programming models and runtimes are necessary that both expose the fundamental changes in relative costs of each operation, as well as break abstraction barriers such that the heterogeneity of future machines can be both exposed and exploited.

We envision a significant fundamental Materials discovery component, guided by the “Grand Challenges in Directing Matter and Energy” that will drive the program. This is expected to extensively utilize BES capabilities including user facilities such as the nanocenters, synchrotron, electron and neutron facilities, and the network of supercomputing facilities. ASCR provides capabilities in advanced computing facilities, applied mathematics (new algorithms), data processing (to support materials search), programming models and system architecture. This project also leverages NNSA’s microelectronics expertise and advanced computing capabilities. EERE will demonstrate how low power computing can be extended to other platforms, vastly reducing the nations computing power consumption. Advanced manufacturing capabilities will be developed in the U.S. and provide high tech jobs. Semiconductor manufacturing in the U.S currently employs 250,000 skilled workers domestically and more than 1 million domestic jobs through associated activities.

Within 10 years, we will invent new energy efficient electronics materials systems, devices, manufacturing systems, and architectures to continue exponential technology scaling of digital electronics performance and energy efficiency beyond 2035-2040.

