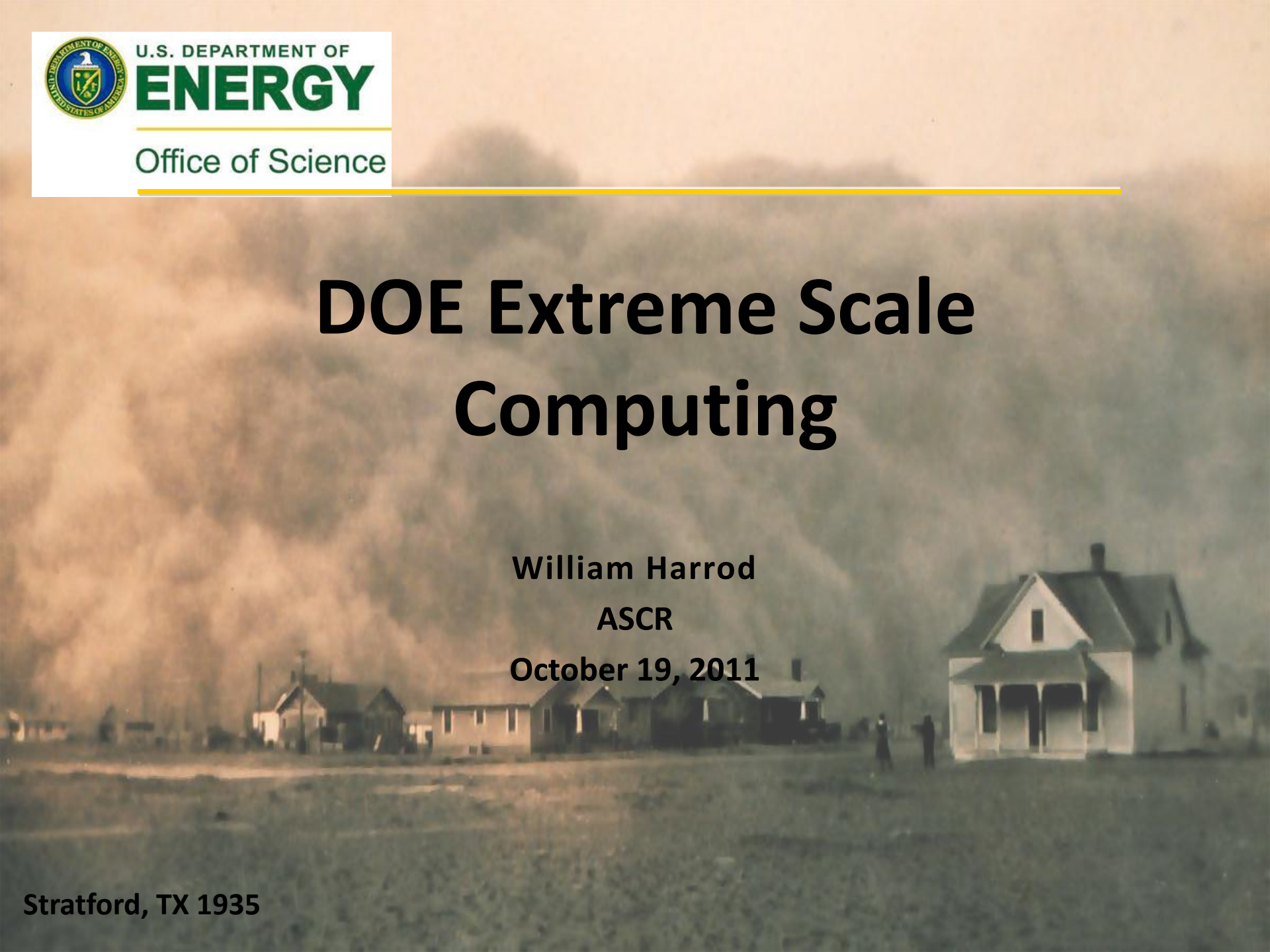


DOE Extreme Scale Computing

William Harrod

ASCR

October 19, 2011

Stratford, TX 1935

U.S. DEPARTMENT OF ENERGY
Office of Science

Outline

- View from Germantown
- Exascale Technical Challenges
- Reports to Congress
- Summary

U.S. DEPARTMENT OF ENERGY
Office of Science

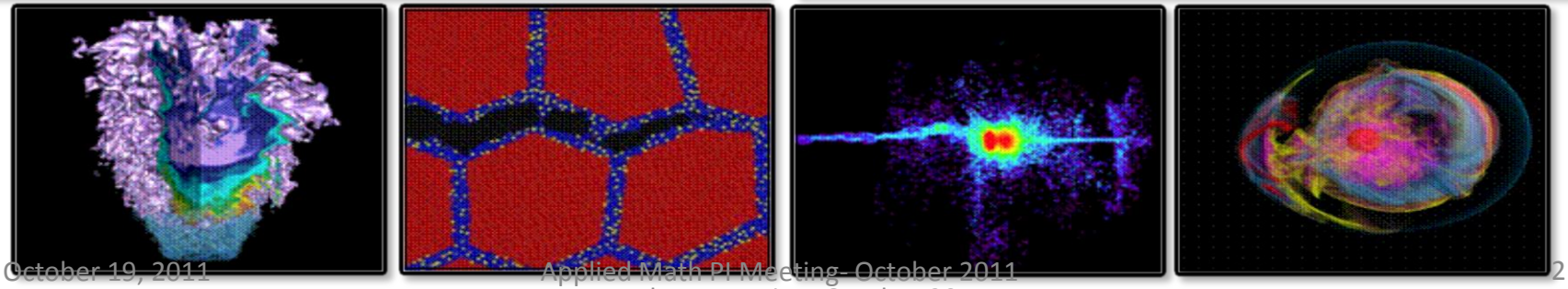
Advanced Scientific Computing Research (ASCR)

The Scientific Challenges:

- **Deliver next-generation scientific applications using today's petascale computers.**
- **Discover, develop and deploy tomorrow's Exascale computing and networking capabilities.**
- **Develop, in partnership with U.S. industry, next generation computing hardware and tools for science.**
- **Discover new applied mathematics and computer science for the ultra-low power, multicore-computing future.**
- Provide technological innovations for U.S. leadership in Information Technology to advance competitiveness.

FY 2013 Highlights:

- Investments in U.S. industry to address critical challenges in hardware and technologies on the path to Exascale.
- Co-design centers to deliver next generation scientific applications by coupling application development with formulation of computer hardware architectures and system software.
- Operation of a 10 petaflop low-power IBM Blue Gene/Q at the Argonne Leadership Computing Facility and a hybrid, multi-core prototype computer at the Oak Ridge Leadership Computing Facility.
- ESnet puts100Gbps technologies into production.



DOE Office of Science
Early Career Research Program

FY12 is 3rd year of this University & DOE Laboratory funding opportunity:

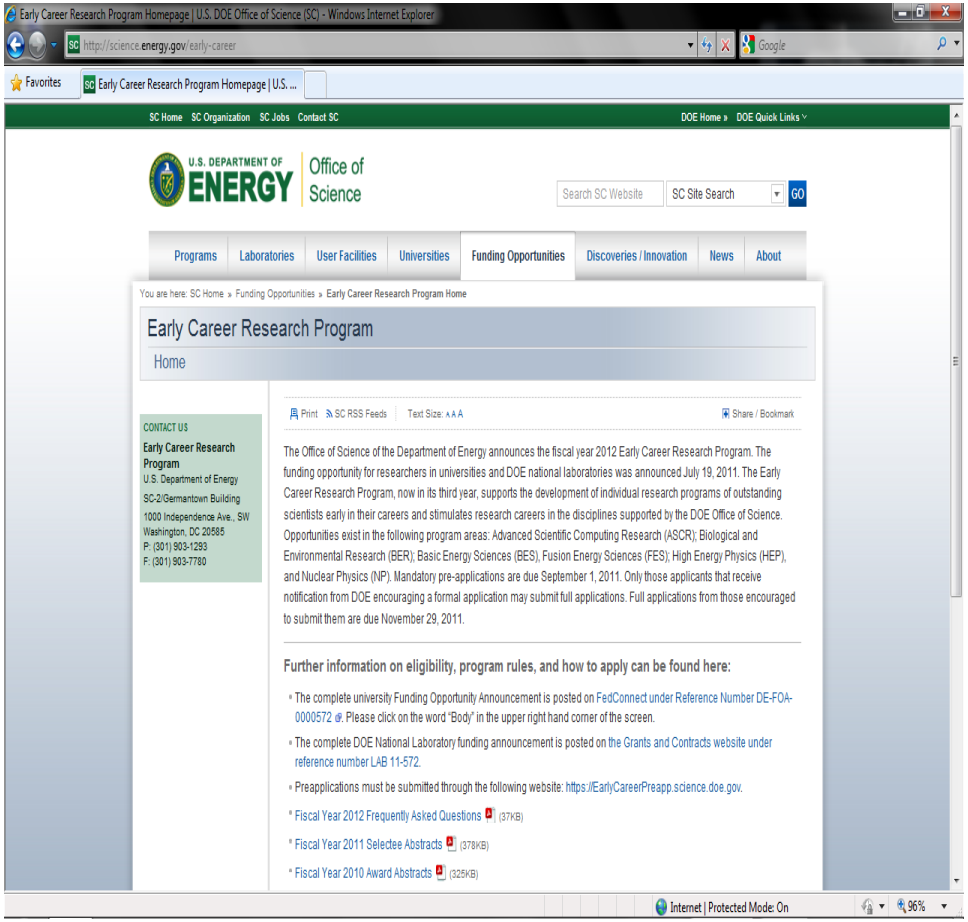
- Posted on July 19, 2011
- Pre-applications due Sep 1
- Encourage applications by Oct 3
- Full applications due Nov 29

ASCR base research programs:

- Applied Mathematics – 4 topics areas
- Computer Science – topic areas are focused on 3 key software challenges for Exascale platforms

ASCR Panel Reviews: Jan 24-25, 2012

- In Rockville/Bethesda, MD



For rules and requirements, see http://science.energy.gov/early-career

U.S. DEPARTMENT OF ENERGY

Office of Science

Exascale Reports

- Town Hall Meetings April-June 2007
- Scientific Grand Challenges Workshops (11/08, 09/09) Climate Science (11/08),
  - High Energy Physics (12/08),
  - Nuclear Physics (1/09),
  - Fusion Energy (3/09),
  - Nuclear Energy (5/09) (with NE)
  - Biology (8/09),
  - Material Science and Chemistry (8/09),
  - National Security (10/09) (with NNSA)
- Cross-cutting workshops
  - Architecture and Technology (12/09)
  - **Architecture, Applied Mathematics and Computer Science (2/10)**
- External Panels
  - Trivelpiece Panel (1/10)
  - ASCAC Exascale Charge (FACA) (11/10)

DOE ASCR Exascale Reports

"The key finding of the Panel is that there are compelling needs for Exascale computing capability to support the DOE's missions in energy, national security, fundamental sciences, and the environment. The DOE has the necessary assets to initiate a program that would accelerate the development of such capability to meet its own needs and by so doing benefit other national interests. Failure to initiate an Exascale program could lead to a loss of U. S. competitiveness in several critical technologies."

Trivelpiece Panel Report, January, 2010

U.S. DEPARTMENT OF ENERGY
Office of Science

Senate Appropriations Committee

ENERGY AND WATER DEVELOPMENT APPROPRIATIONS BILL, 2012

- ***The Committee supports the Department's initiative to develop Exascale computing***—*1,000 times more powerful than today's most powerful computer. The Committee recommends $126,000,000* **(FENCED)** *to support this initiative, which includes $90,000,000 for the Office of Science and $36,000,000 for the National Nuclear Security Administration.*

- *The Committee understands that the path* ***to Exascale computing will be extremely challenging and will require significant research and development breakthroughs****.*

- *The Committee understands that* ***the Department (of Energy) will have the lead Government role in computing research and development.*** *The Department's role in developing more advanced computing platforms is even more important with the elimination of the DARPA High Performance Computing program.*

NOT PASSED BY CONGRESS AT THIS TIME

U.S. DEPARTMENT OF ENERGY
Office of Science

Exascale Co-Design Centers

Exascale Co-Design Center for Materials in Extreme Environments (ExMatEx)

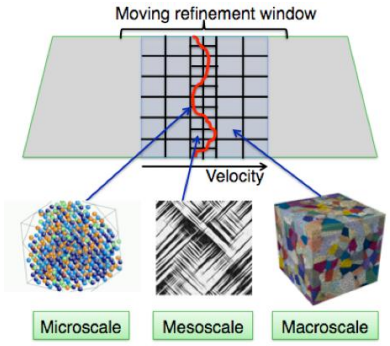
Director: Timothy Germann (LANL)

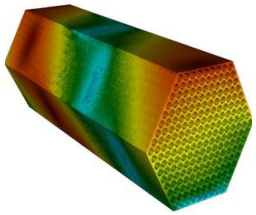
Center for Exascale Simulation of Advanced Reactors (CESAR)

Director: Robert Rosner (ANL)

Combustion Exascale Co-Design Center (EXACT)

Director: Jacqueline Chen (SNL)





	ExMatEx (Germann)	CESAR (Rosner)	EXACT (Chen)
National Labs	LANL	ANL	SNL
	LLNL	PNNL	LBNL
	SNL	LANL	LANL
	ORNL	ORNL	ORNL
		LLNL	LLNL
			NREL
University & Industry Partners	Stanford	Studsvik	Stanford
	CalTech	TAMU	GA Tech
		Rice	Rutgers
		U Chicago	UT Austin
		IBM	Utah
		TerraPower	
		General Atomic	
		Areva	

U.S. DEPARTMENT OF ENERGY
Office of Science

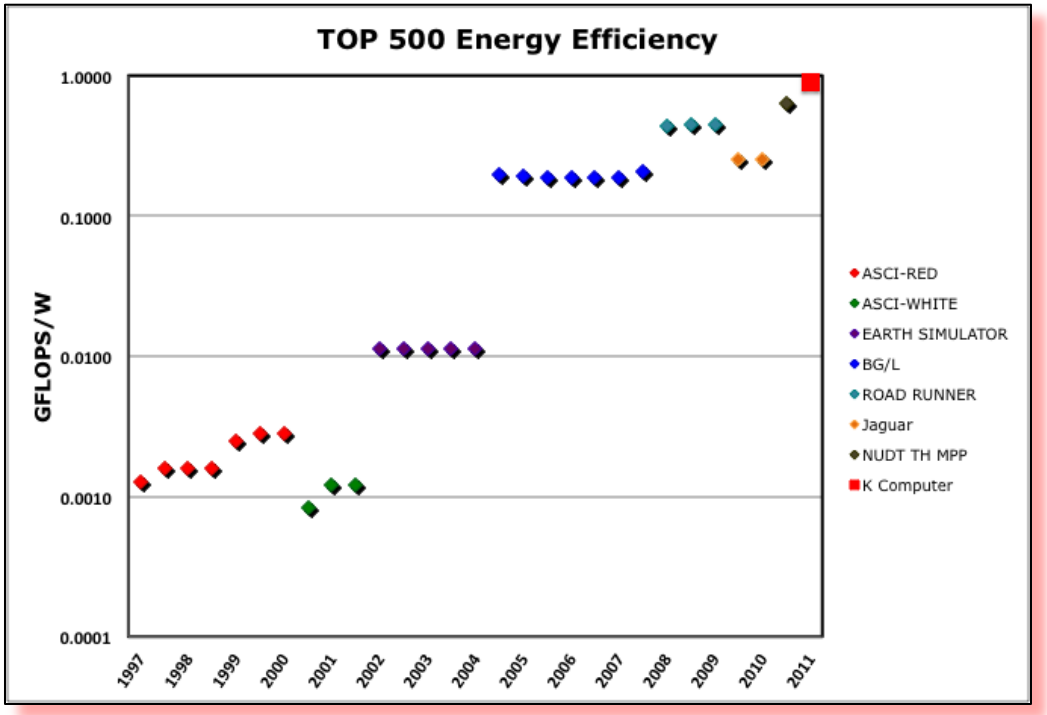
DOE/ASCR Progress Toward Exascale

- **FY 2010 Basic Research funded in Exascale related topic areas:**
  - **Applied Math:** Uncertainty Quantification
  - **Computer Science:**
    - Advanced Architectures
    - X-Stack
    - Scientific Data Management and Analysis
- **FY2011:**
  - **Computational Partnerships:** 3 Co-Design Centers Funded
  - **Exascale MOU with NNSA signed**
  - **Request for Information (released by ANL):** critical and platform technologies
- **FY2012:**
  - **Exascale Plan to Congress**
- **On-going**
  - **Exascale Coordination meetings with other Federal Agencies**

U.S. DEPARTMENT OF ENERGY
Office of Science

Challenges to Achieving Exascale Computing

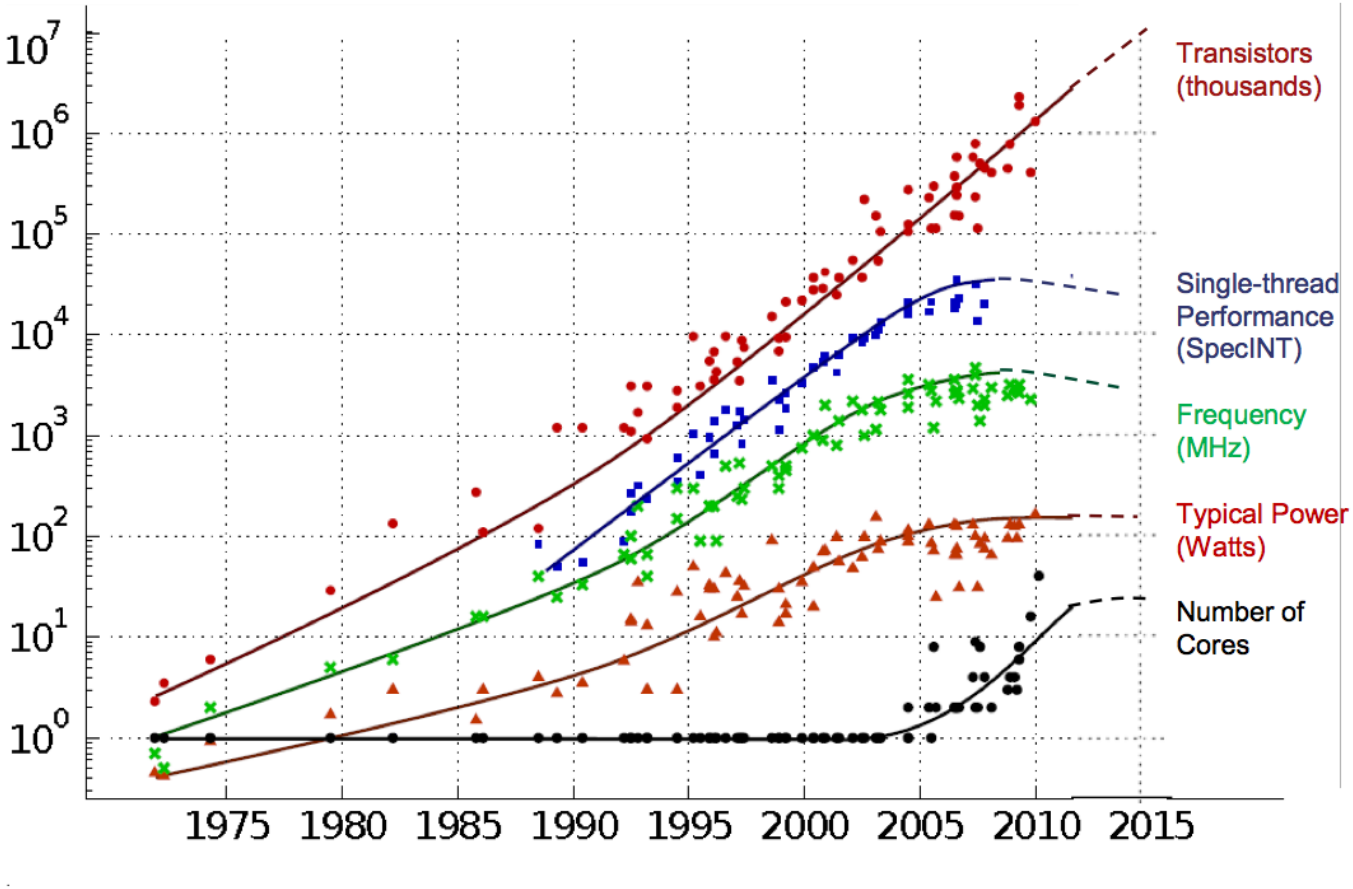
- **It has been estimated that an Exascale computer will consist of one billion cores.**
- **Power**
  - 1-2 nJ/operation today
  - 20 pJ/flop (or 50 GFLOPS/W system) required for ExaScale (decrease by 1000x)
  - Dominated by data movement and overhead
- **Programmability**
  - Writing an efficient parallel program is hard
  - Locality required for efficiency
  - System complexity greatly inhibits programmability



Exascale Goal 50 GFLOPS/W

It's a Flat World

35 YEARS OF MICROPROCESSOR TREND DATA

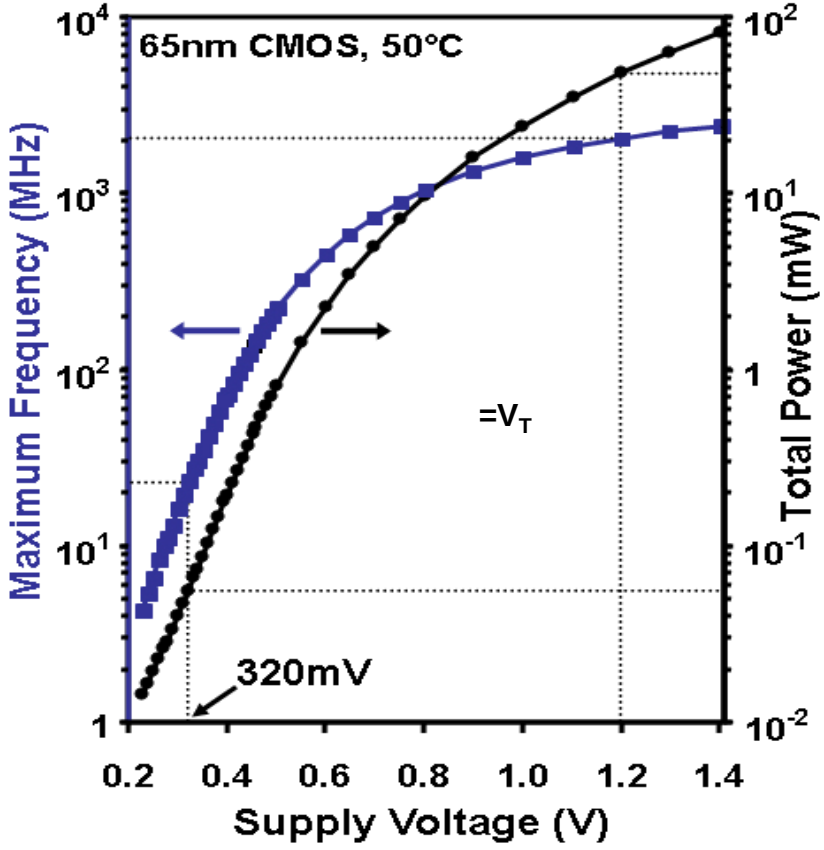


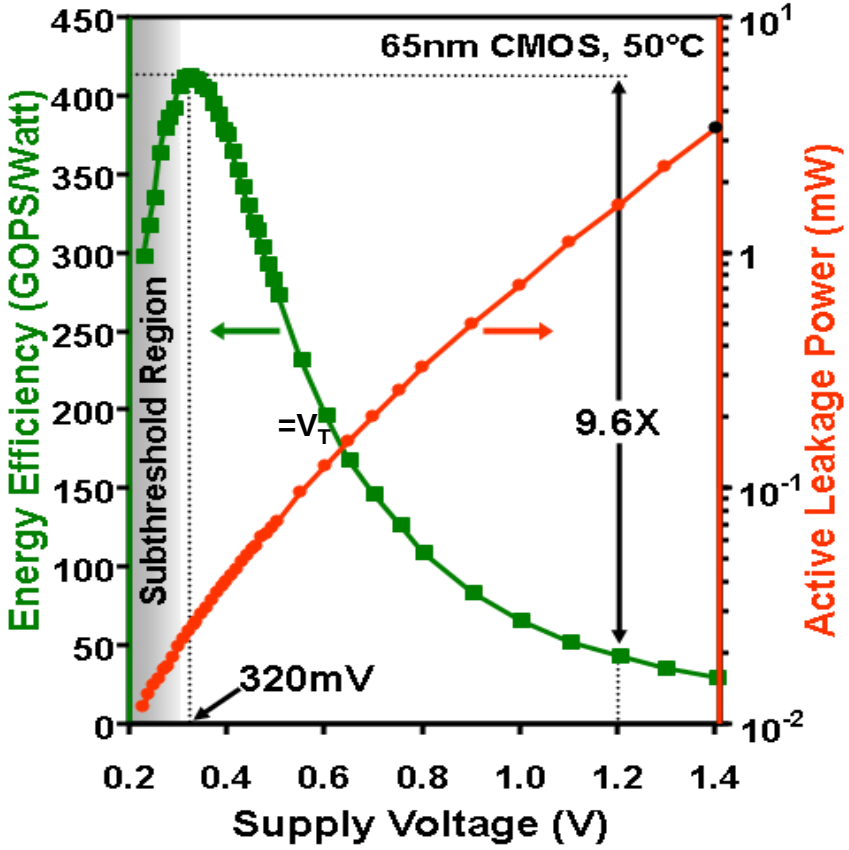
Original data collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond and C. Batten
Dotted line extrapolations by C. Moore

U.S. DEPARTMENT OF ENERGY
Office of Science

Optimized Power Performance

(Data from Intel: Shekhar Borkar)



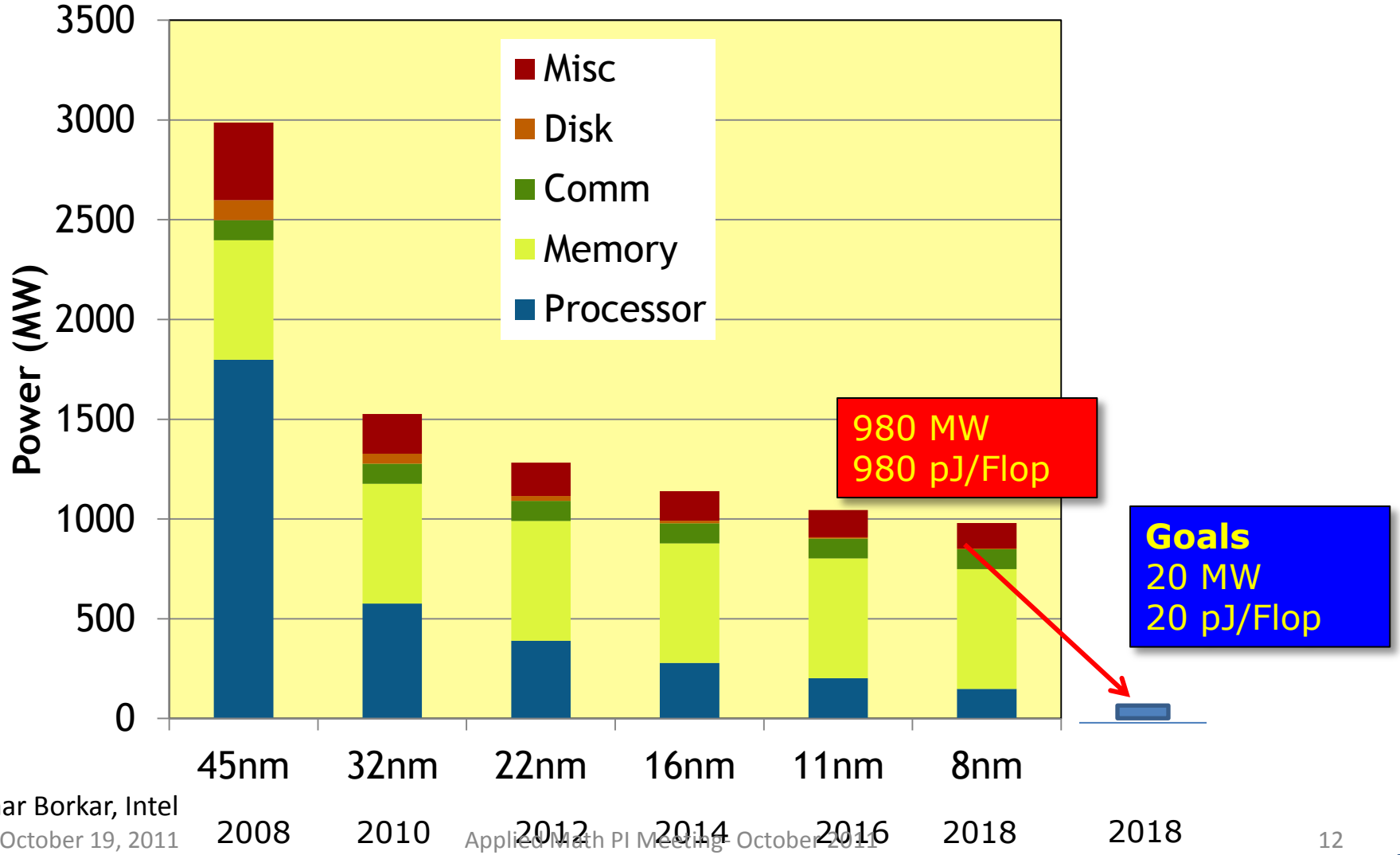


Data Path Insights

- **The cost of data storage and access – *Dominates energy consumption***
- **Computer design must keep or make data effectively local to processing**
- **Cost to access increasingly distant data – energy and latency – may appear to be "linear"**
  - It is not, data access costs increases dramatically with distance to data
  - Interconnect, latency, and overhead costs grow disproportionately with distance to data
- **Need to redesign the entire data access path:**
  - Memory technologies – reduce access cost
    - Transfer energy: DDR3 DRAM 40 pJ/bit ➔ 1-3 pJ/bit
  - Packaging costs – reduce bit movement cost

U.S. DEPARTMENT OF ENERGY
Office of Science

Peak Exascale Machine Power



Shekhar Borkar, Intel

U.S. DEPARTMENT OF ENERGY
Office of Science

Applied Math Projects

- **ASCR is preparing for Extreme Scale computing and continuing to support current generation HPC systems.**
- **Applied Mathematics projects should have DOE relevance across a broad range of science domains:**
  - SC offices (BES, BER, FES, …)
  - Applied Offices (Nuclear Energy, Fossil Energy, Office of Electricity Delivery and Energy Reliability, …)
  - NNSA

Anticipated Future Programs

(more to come)

- **X-Stack: Runtimes, Programming Models, Languages, Compilers, and Tools**
  - Minimize exposure of system complexity
  - Extreme concurrency
  - Heterogeneous system
  - Minimize data movement
  - Runtimes for efficiency and resiliency
  - Self-aware OS/runtime
- **Extreme Scale Solver Algorithms**
  - Fine grain parallelism
  - Data movement & locality

The Real Challenges

- **Avoiding mediocre solutions**
  - Evolving existing systems
- **Developing a new software stack for Extreme Scale systems**
  - Not treating it as an "after thought"
  - Failure to develop programmable systems
- **Developing efficient mathematical algorithms that deliver high performance for Extreme Scale systems**
  - Minimize communications
  - Fault tolerant computations
  - Asynchronous, fine grain parallelism
- **New computers designs based on a new execution model**
  - Must be based on 2020 COTS technology
  - Exotic technology is not an option

It is a Budget Problem!

U.S. DEPARTMENT OF ENERGY
Office of Science

Summary

Voltage scaling to reduce power and energy

- Explodes parallelism
- Cost of communication vs. computation—critical balance

Its not about the FLOPS. Its about data movement.

- Algorithms should be designed to perform more work per unit data movement.
- Programming systems should further optimize this data movement.
- Architectures should facilitate this by providing an exposed hierarchy and efficient communication.

System software to orchestrate all of the above

- Self aware operating system

BACKUP

U.S. DEPARTMENT OF ENERGY
Office of Science

Exascale Timeline

(from ASCR Arch I Workshop)