U.S. DEPARTMENT OF ENERGY | Office of Science

ASCR

February 8, 2011

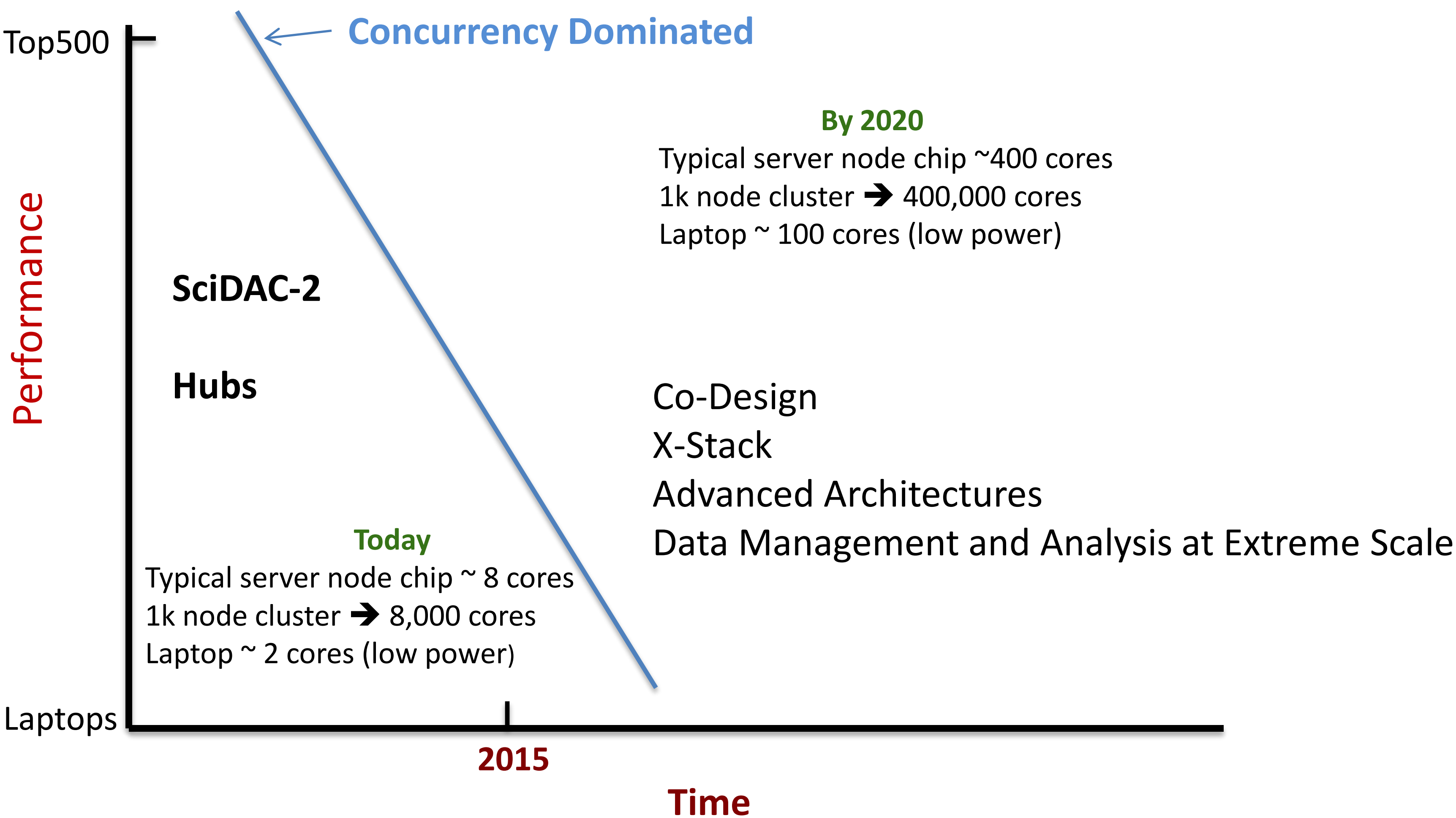
The Challenge of Exascale and Exabyte

Daniel Hitchcock

Acting Associate Director

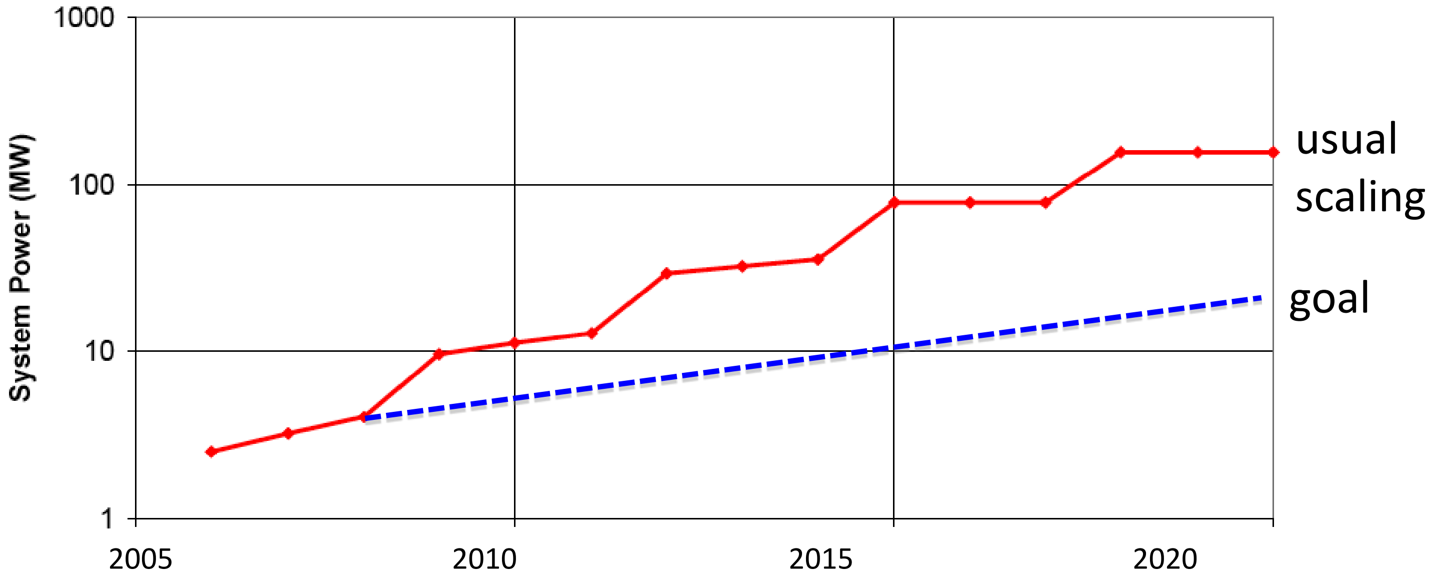
Advanced Scientific Computing Research

Phase Change in Computing

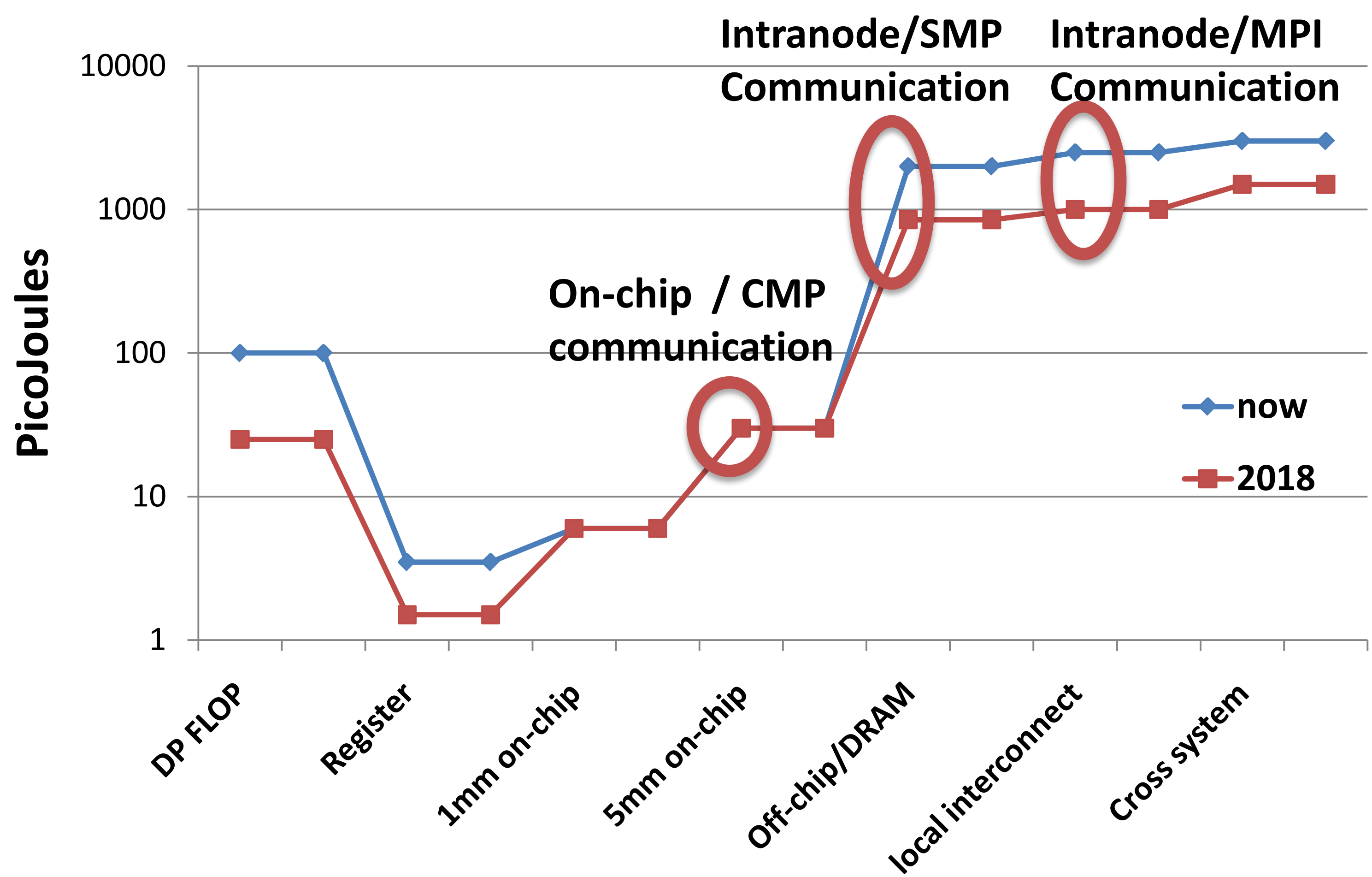


Exascale is about Energy Efficient Computing

- **At $1M per MW, energy costs are substantial**
- **1 petaflop in 2010 will use 3 MW**
- **1 exaflop in 2018 at 200 MW with “usual” scaling**
- **1 exaflop in 2018 at 20 MW is target**



The Fundamental Issue: Where does the Energy (and Time) Go?



National Academies Report on Computing Performance

PREPUBLICATION COPY—SUBJECT TO FURTHER EDITORIAL CORRECTION

The Future of Computing Performance

Game Over or Next Level?

Samuel H. Fuller and Lynette I. Millett, Editors

Committee on Sustaining Growth in Computing Performance

Computer Science and Telecommunications Board

Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

i

- **Report makes same points**
  - Past performance increases have driven innovation
  - Processor speeds stalled
  - Energy is limitation
- **Impacts everything**
  - Data
  - Simulations
  - Control

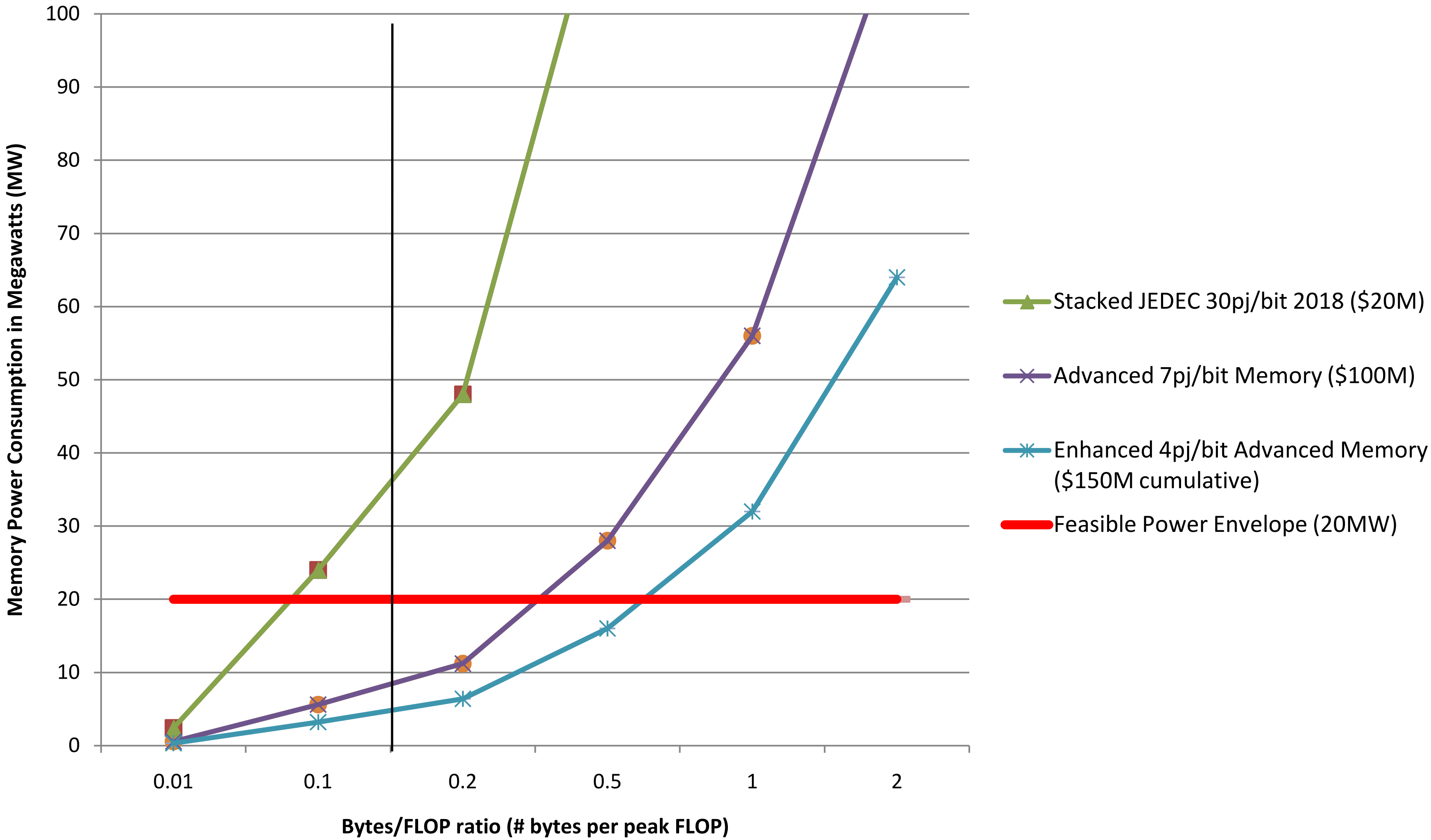
Challenges to ~~Exascale~~

Performance Growth

1) **System power** is the primary constraint
2) **Memory** bandwidth and capacity are not keeping pace
3) **Concurrency** (1000x today)
4) **Processor** architecture is an open question
5) **Programming model** heroic compilers will not hide this
6) **Algorithms** need to minimize data movement, not flops
7) **I/O bandwidth** unlikely to keep pace with machine speed
8) **Reliability and resiliency** will be critical at this scale
9) **Bisection bandwidth** limited by cost and energy

Unlike the last 20 years most of these (1-7) are equally important across scales, e.g., 100 10-PF machines

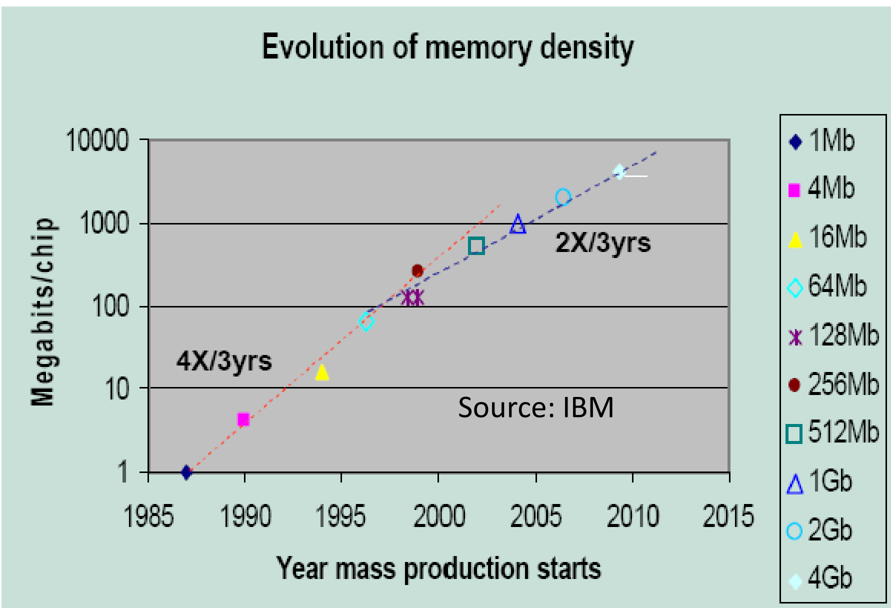
Memory Technology: Bandwidth costs power

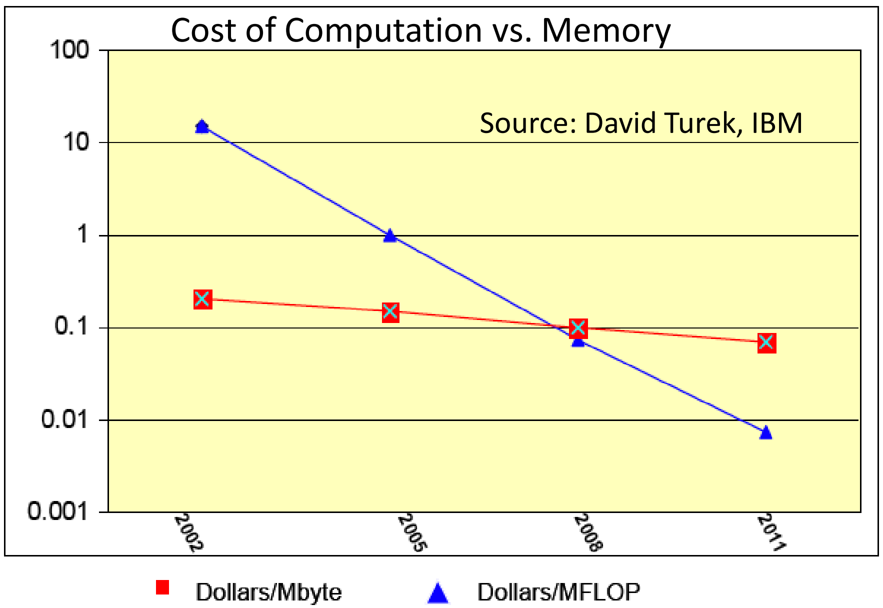


Memory is Not Keeping Pace

Technology trends against a constant or increasing memory per core

- Memory density is doubling every three years; processor logic is every two
- Memory costs are dropping gradually compared to logic costs

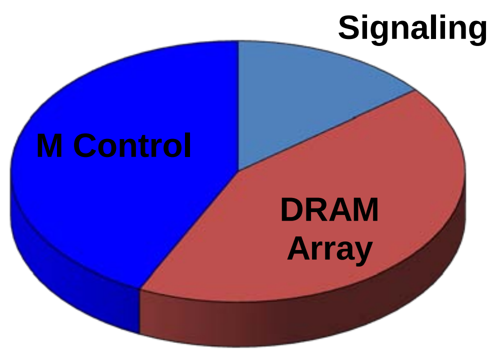




The cost to sense, collect, generate and calculate data is declining much faster than the cost to access, manage and store it

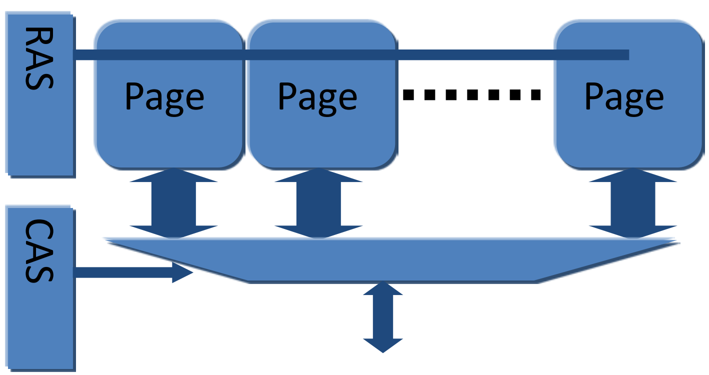
Question: *Can you double concurrency without doubling memory?*

Revised DRAM Architecture



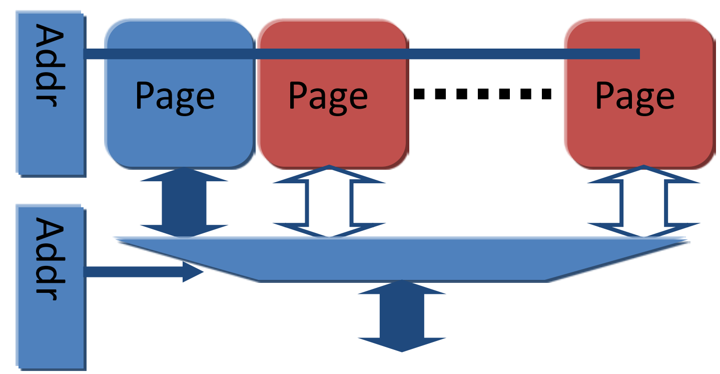
Today's energy cost :
~150 pJ/bit

Today's DRAM



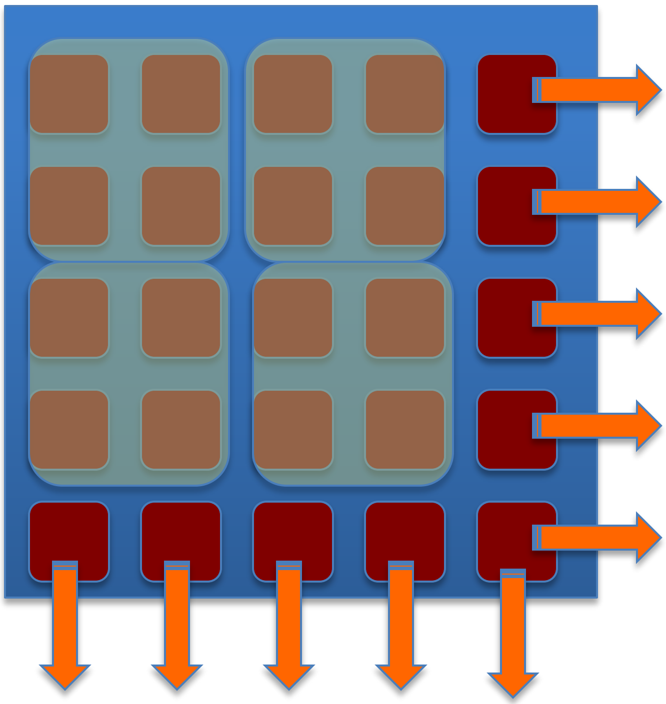
Activates many pages
Lots of reads and writes (refresh)
Small amount of read data is used
Requires small number of pins (DIPs)

Tomorrow's DRAM



Activates few pages
Read and write (refresh) what's needed
All read data is used
Requires large number of IO's (3D)

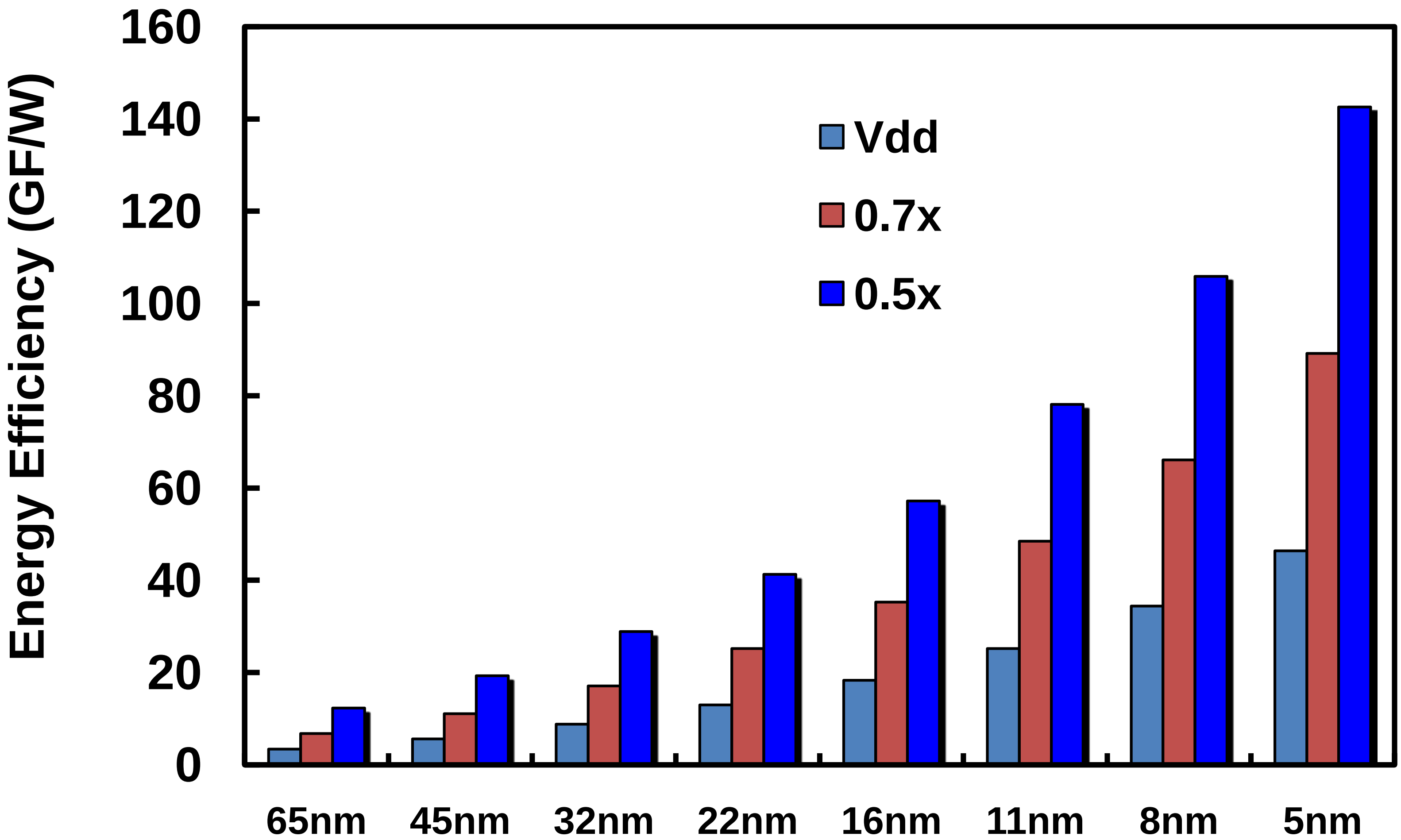
On-Chip Architecture:
different approaches to on-chip clustering



- **Cost of moving long-distances on chip motivates clustering on-chip**
  - 1mm costs ~6pj (today & 2018)
  - 20mm costs ~120 pj (today & 2018)
  - FLOP costs ~100pj today
  - FLOP costs ~25pj in 2018

- **Different Architectural Directions**
  - GPU: WARPs of hardware threads clustered around shared register file
  - CMP: limited area cache-coherence
  - CMT: hardware multithreading clusters

Extreme Voltage Scaling Efficiency



~3X compute energy efficiency with extreme Vdd Scaling

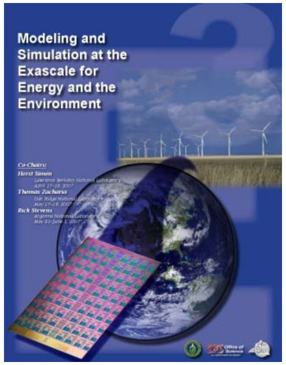
Approaches

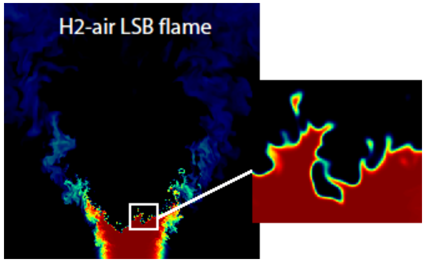
- **Locality, Locality, Locality!**
- **Billion Way Concurrency;**
- **Uncertainty Quantification (UQ) including hardware variability;**
- **Flops free data movement expensive so:**
  - Remap multi-physics to put as much work per location on same die;
  - Include embedded UQ to increase concurrency;
  - Include data analysis if you can for more concurrency
  - Trigger output to only move important data off machine;
  - Reformulate to trade flops for memory use.
- **Wise use of silicon area**

Why Exascale --
Exascale Applications and Technology

http://www.science.doe.gov/ascr/WorkshopsConferences/GrandChallenges.html

- **Town Hall Meetings April-June 2007**
- **Scientific Grand Challenges Workshops November 2008 – October 2009**
  - 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)
- **Meetings with industry (8/09, 11/09)**
- **External Panels**
  - Trivelpiece Panel (1/10)
  - ASCAC Exascale Charge (FACA) (11/10)





MISSION IMPERATIVES

"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

ASCR Advisory SubCommittee Exascale Report

http://www.science.doe.gov/ascr/ASCAC/Reports/Exascale-Subcommittee-Report.pdf

- **Charge**
  - Assess the opportunities and challenges of exascale computing for the advancement of science, technology, and Office of Science missions
  - Identify strategies that ASCR can use to address the challenges and deliver on such opportunities
- **Findings**
  - The mission and science opportunities in going to exascale are compelling
  - Making he transition to exascale poses numerous unavoidable scientfic, algorithmic, mathematical, software, and technological challenges
  - The benefits of going to exascale far outweigh the costs
  - The exascale initiative as described in workshop reports and expert testimony portends an integrated approach to the path forward
- **Recommendation**
  - DOE should proceed expeditiously with an exascale initiative so that it continues to lead in using extreme scale computing to meet important national needs.

Exascale computing will uniquely provide knowledge leading to transformative advances for our economy, security and society in general. A failure to proceed with appropriate speed risks losing competitiveness in information technology, in our industrial base writ large, and in leading-edge science. *ASCAC subcommittee report*

Exascale Progress since May, 2010

- **Proposals processed in Exascale related topic areas:**
  - Applied Math: Uncertainty Quantification
  - Computer Science: Advanced Architectures
  - Computer Science: X-Stack
  - Computer Science: Data Management and Analysis at Extreme Scales
  - Computational Partnerships: Co-Design (21 Proposals requesting ~ $160M/year)
- **Exascale Coordination meeting with DOD and DARPA – June 15 with decisions to --**
  - Have yearly coordination/planning meetings in accordance with current MOU
  - Follow-on meeting of Program Managers to identify gaps and overlaps in computer science and applied math areas currently funded
  - Address critical technologies issues, initially in memory
- **Ongoing meetings with NNSA**

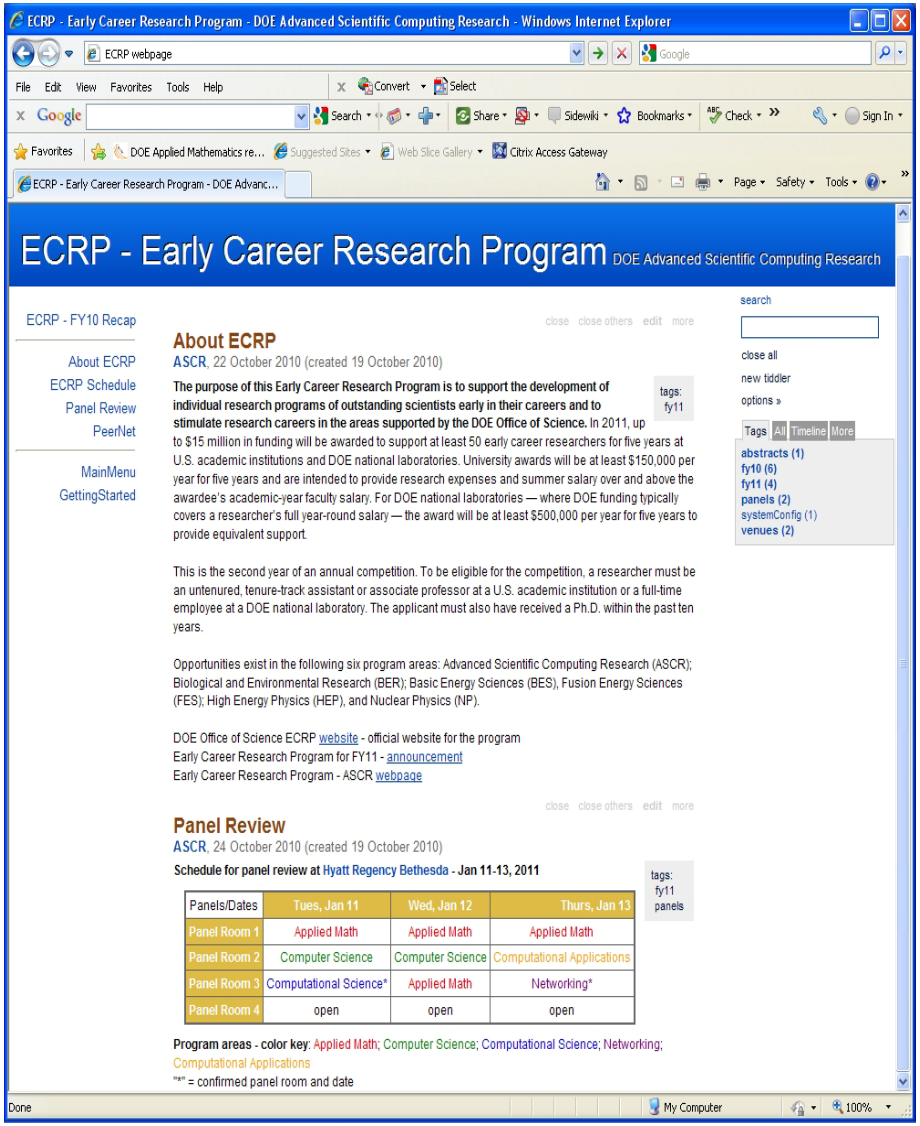
ASCR Exascale Research Kick-off Meeting

- **PI meeting for**
  - Advanced Architecture, X-Stack, Scientific Data Management and Analysis at Extreme Scales awardees
  - Co-Design planning grant recipients
- **Expected Outcomes:**
  - Awareness **within** each solicitation communities and ASCR what members are doing and areas where they can leverage and supplement their work
  - Awareness **across** solicitation communities of what is going on and where each project fits in relation to the broad spectrum
  - Identification of gaps in ASCR exascale research potfolio
  - Lay groundwork for collaboration/cooperation with NNSA Exascale Roadmapping activities
- **Tentative dates: March 7-11, 2011**

Early Career Research Program (FY11) - ECRP

FY11 – 2nd year of competition:

- Solicitation issued - July 1
- Pre-applications due - Aug 13
- Encourage applications – Sep 14
- Full applications due Nov 9



Schedule for 2011:

- ASCR Panel Reviews, Jan 11-13
  - Applied Mathematics
  - Computer Science
  - Computational Science / Biology
  - Networking

Advanced Architectures

(28 proposals requesting ~$28M/yr)

- **Blackcomb: Hardware-Software Co-design for Non-Volatile Memory in Exascale Systems**
  - ORNL, Hewlett Packard Labs, University of Michigan, Penn State University
- **Enabling Co-Design of Multi-Layer Exascale Storage Architectures**
  - ANL, RPI
- **NoLoSS: Enabling Exascale Science through Node Local Storage Systems**
  - ANL, LLNL
- **CoDEx: A Hardware/Software Co-Design Environment for the Exascale Era**
  - LBNL, SNL, LLNL, GATECH
- **Data Movement Dominates: Advanced Memory Technology to Address the Real Exascale Power Problem**
  - SNL, LBNL, Micron, Columbia, UMD
- **Thrifty: An Exascale Architecture for Energy-Proportional Computing**
  - UIUC, LLNL, UCSD, Intel

Six projects funded at $5M/yr

X-Stack Software

(55 proposals requesting ~$40M/yr)

- **A Fault-oblivious Extreme-scale Execution Environment**
  - SNL, PNNL, LLNL,OSU, IBM, Alcatel-Lucent Bell Labs, BU
- **Auto-tuning for Performance and Productivity on Extreme-scale Computations**
  - LBNL, UCSB, PPPL
- **Software Synthesis for High Productivity ExaScale Computing**
  - MIT, UC Berkeley
- **COMPOSE-HPC: Software Composition for Extreme Scale Computational Science and Engineering**
  - ORNL, Galois, Inc., LLNL, PNNL, SNL
- **Damsel: A Data Model Storage Library for Exascale Science**
  - Northwestern University, ANL), NCSU, The HDF Group
- **Vancouver: Designing a Next-Generation Software Infrastructure for Productive Heterogeneous Exascale Computing**
  - ORNL, UIUC, University of Oregon, GATECH

X-Stack Software (cont)

(55 proposals requesting ~$40M/yr)

- **Enabling Exascale Hardware and Software Design through Scalable System Virtualization**
  - UNM, Northwestern University, SNL, ORNL
- **ZettaBricks: A Language, Compiler and Runtime Environment for Anyscale Computing**
  - MIT
- **Compiled MPI: Cost-Effective Exascale Application Development**
  - UIUC, LLNL, Indiana
- **ExM: System support for extreme-scale, many-task applications**
  - ANL, Chicago, University of British Columbia
- **An Open Integrated Software Stack for Extreme Scale Computing**
  - ANL, LBNL, ORNL, SNL, UTK

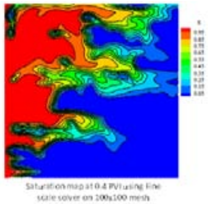
Eleven projects funded at $8.5M/yr

Summary of UQ Projects

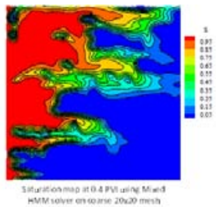
(90 projects received requesting ~ $45M/yr)

- **Modeling and Simulation of High-Dimensional Stochastic Multiscale PDE Systems at the Exascale**
  - Guang Lin (PNNL), Nicholas Zabaras (Cornell), and Ioannis Kevrekidis, (Princeton)
- **Advanced Dynamically Adaptive Algorithms for Stochastic Simulations on Extreme Scales**
  - Dongbin Xiu (Purdue), Richard Archibald, Ralf Deiterding, and Cory Hauck (ORNL)
- **A High-Performance Embedded Hybrid Methodology for Uncertainty Quantification with Applications**
  - Charles Tong (LLNL), Barry Lee (PNNL), Gianluca Iaccarino (Stanford)
- **Enabling Predictive Simulation and UQ of Complex Multiphysics PDE Systems by the Development of Goal-Oriented Variational Sensitivity Analysis and a-Posteriori Error Estimation Methods**
  - John Shadid (SNL), Don Estep (CSU), Victor Ginting (UWyoming)
- **Bayesian Uncertainty Quantification in Predictions of Flows in Highly Heterogeneous Media and its Application to CO_2 Sequestration**
  - Yalchin Efendiev (Texas A&M), Panayot Vassilevski (LLNL)
- **Large-Scale Uncertainty and Error Analysis for Time-Dependent Fluid/Structure interactions in Wind Turbine Applications**
  - Juan Alonso (Stanford) and Michael Eldred, et al (SNL)

Six projects funded at $3M/yr

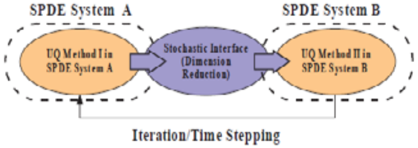


Fine mesh



Coarse-graining

Saturation profile of oil-water porous media flow using fine and multiscale coarse-graining solver



Sketch of hybrid UQ method between multi-physics stochastic PDE systems

Sci. Data Mgmt. & Analysis at Extreme Scale

(37 proposals requesting ~$22/yr)

- **Dynamic Non-Hierarchical File Systems for Exascale Storage**
  - University of California, Santa Cruz
- **Runtime System for I/O Staging in Support of In-Situ Processing of Extreme Scale Data**
  - Oak Ridge National Laboratory; Lawrence Berkeley National Laboratory; Georgia Institute of Technology
- **Bringing Exascale I/O Within Science's Reach: Middleware for Enabling & Simplifying Scientific Access to Extreme Scale Data**
  - Lawrence Berkeley National Laboratory; Pacific Northwest National Laboratory
- **Adding Data Management Services to Parallel File Systems**
  - University of California, Santa Cruz; Lawrence Livermore National Laboratory
- **Scalable and Power Efficient Data Analytics for Hybrid Exascale Systems**
  - Northwestern University; Lawrence Berkeley National Laboratory; Oak Ridge National Laboratory

Sci. Data Mgmt. & Analysis at Extreme Scale

- **A Pervasive Parallel Processing Framework for Data Visualization & Analysis at Extreme Scale**
  - Sandia National Laboratories; Kitware, Inc.; University of California, Davis
- **An Information-Theoretic Framework for Enabling Extreme-Scale Science Discovery**
  - Ohio State University; New York Institute of Technology; Argonne National Laboratory
- **Enabling Scientific Discovery in Exascale Simulations**
  - Lawrence Livermore National Laboratory; University of Minnesota
- **Graph-Based 3D Flow Field Visual Analysis**
  - Pacific Northwest National Laboratory
- **Topology-Based Visualization & Analysis of Multidimensional Data & Time-Varying Data at the Extreme Scale**
  - Lawrence Berkeley National Laboratory

Ten projects funded at ~$5M/yr

SciDAC Recompetition Planning

- **DOE SciDAC program manager working group reconstituted**
  - **Representation from all Office of Science programs and the NNSA**
  - **Activities: planning, ‘lessons learned’ and ‘gap’ analysis**

- **Strategy and Schedule**
  - **January-February, 2011: Issue Funding Opportunity Announcement for SciDAC Institutes**
    - **Coordinated investments in applied math, computer science, algorithms, code development, data analytics, visualization, software development and outreach aligned with SciDAC goals and cross-cutting strategic needs identified by ASCR’s SciDAC funding partners.**
    - **ASCR funded**
    - **Anticipated award date- late FY2011**

  - **Spring-Summer, 2011: Issue the first in a series of focused FOAs for domain science components of SciDAC**
    - Jointly issued and funded by ASCR with one, or more, SC program offices and/or the NNSA
    - Requesting proposals that address topics of strategic importance to the domain science funding office(s) and ASCR.
    - Tangible links to the SciDAC Institutes
    - Anticipated award date- early FY2012

BACKUP