A Journey to Exascale Computing

William Harrod
Director – Research Division
Advanced Scientific Computing Research
William.Harrod@Science.DOE.gov

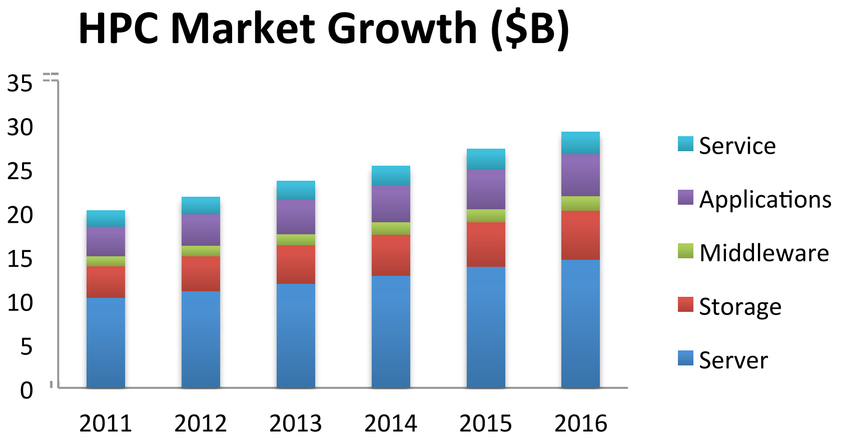
SC12
November 15, 2012

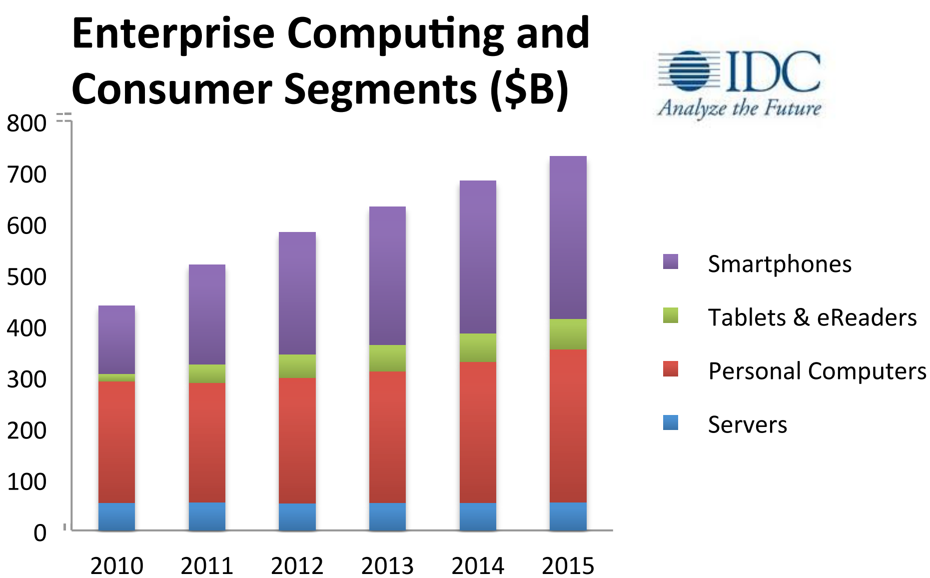
Extreme Scale Science

- **Next generation scientific breakthroughs will require major, novel advances in computing technology - Exascale computing**
- **Causing a data explosion – a natural component of exascale computing**
  - DOE experimental facilities face exponentially burgeoning data caused by technology advances
- **Challenges are well-documented in numerous studies**
  - Can no longer scale up from or modify existing solutions
- **Requires a major paradigm shift**

"There are no more magic bullets" – Mark Horowitz

Revenue & The Power Wall





IDC
Analyze the Future

High Utilization

- DOE HPC centers utilization is approximately 95%

On demand response

- Most enterprise data centers waste 90% of their power consumption – utilization is approximately 10%
- Worldwide enterprise centers consume 30 GW

Source: New York Times

U.S. DEPARTMENT OF ENERGY | Office of Science

Exascale Computing: The Vision

The Grand Challenge Problem

- **Exaflops sustained performance**
  - Approximate peak power 20 MW
- **Real-world mission-critical applications**
  - Extreme Scale Science
- **Productive system**
  - Usable by a wide variety of scientists
  - "Easier" to develop software & management of the system
- **Based on marketable technology**
  - Not a "one off" system
  - Scalable, sustainable technology, exploiting economies of scale and trickle-bounce effect
- **Deployed in 2020+**

Achieving the Exascale Vision:
Key Issues / Requirements

- **Exascale Challenges[1]**
  - Power & Energy
  - Memory & Storage
  - Concurrency & Locality
  - Resiliency

Computers will never be the same

Game Change

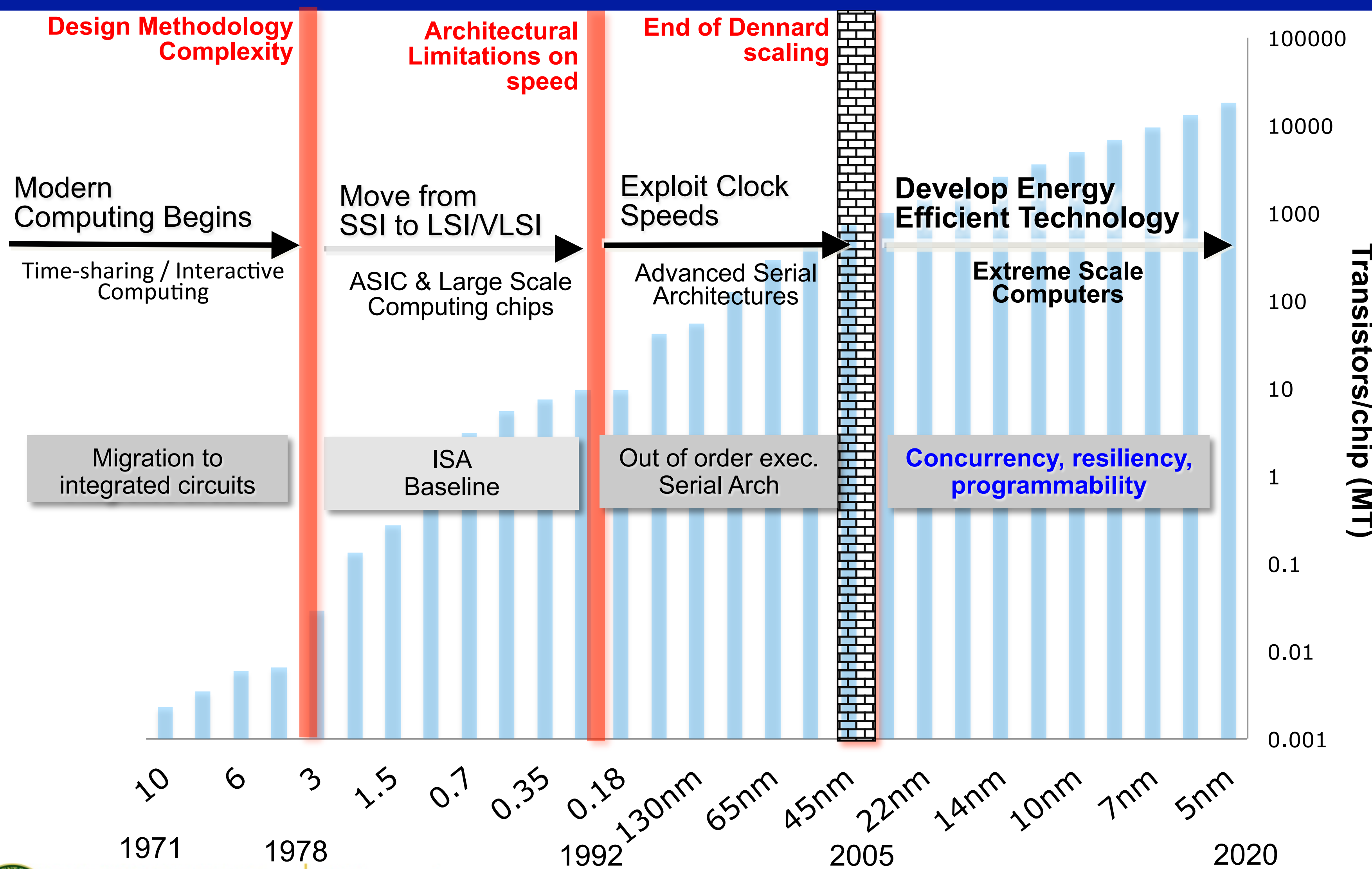
- **Significant research and development investments are required at all levels**
- **Exascale computing requires new system designs**
- **Must enable an economically viable ecosystem**

1 DARPA ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems

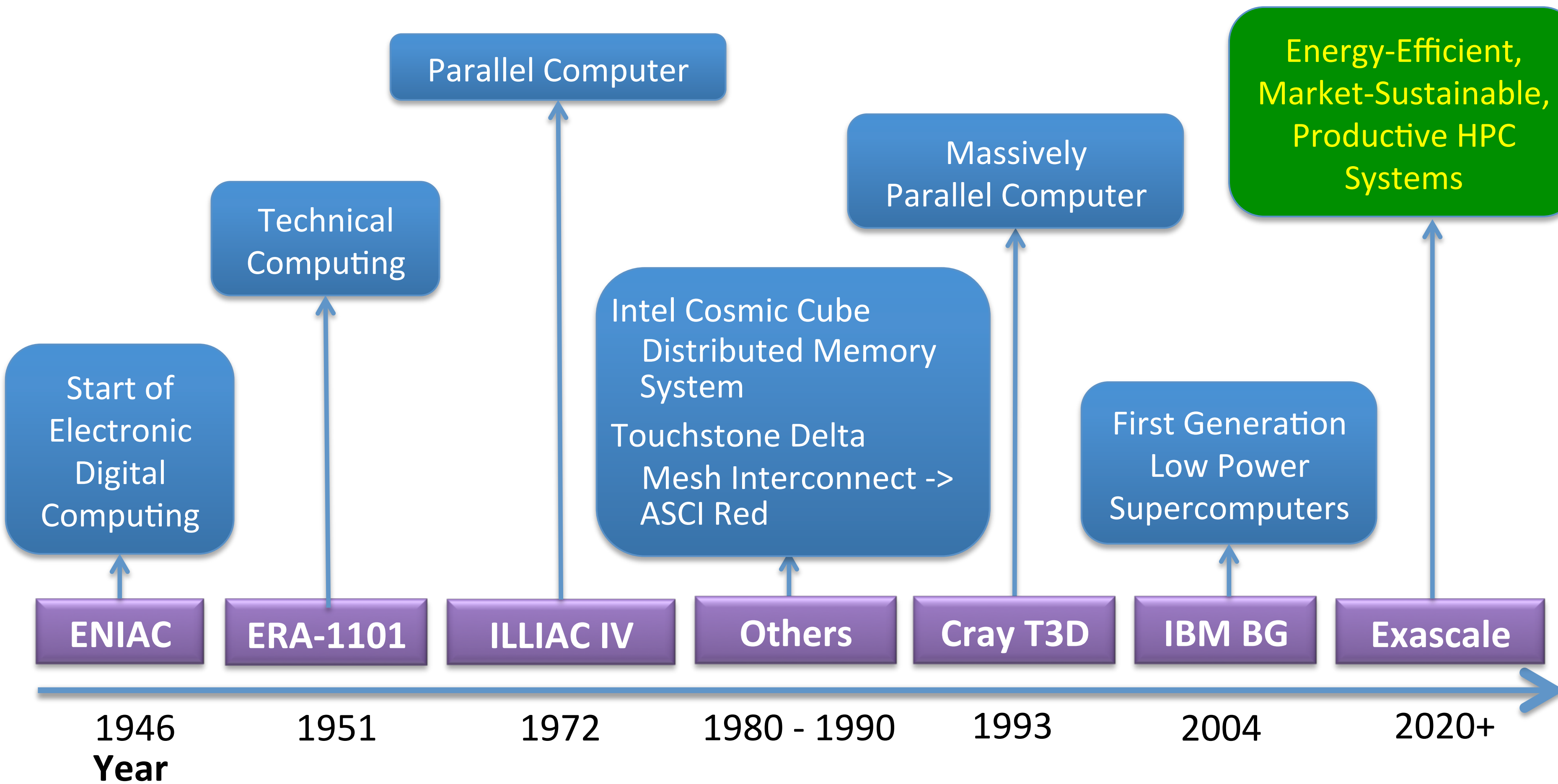
HPC “Movers & Shakers”

- **Emerging implementation technologies**
- **Partnerships led by government**
- **Academic research**
- **Computer industry**

Impact of Implementation Technologies on Commercial Processors



US Federal Government Partnerships
Selected Milestones along the Journey



Computing Industry Landscape

Common Challenges

- Parallelism
- Energy / Power
- Memory / Storage
- Reliability
- Ease of Use

Major Differences

- Applications
- Scale
- Market Space

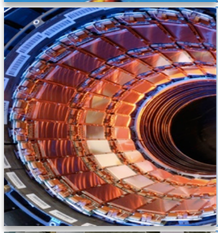
Technical Computing Contributions

- Energy efficient Un-core
- Minimizing data transport cost
- Memory capacity
- Scalability

Extreme Scale Science is Causing a Data Explosion

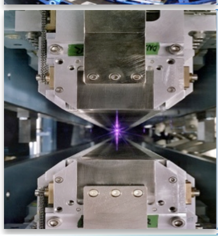
Genomics

Data Volume increases to 10 PB in FY21



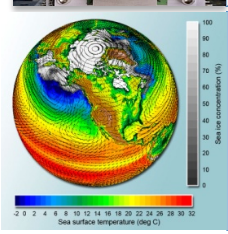
High Energy Physics (Large Hadron Collider)

15 PB of data/year



Light Sources

Approximately 300 TB/day



Climate

Data expected to be hundreds of 100 EB

Driven by exponential technology advances

Data sources

- Scientific Instruments
- Scientific Computing Facilities
- Simulation Results

Big Data is part of Big Compute

- Using Big Data requires processing (e.g., search, transform, analyze, …)
- Exascale computing will enable timely and more complex processing of increasingly large Big Data sets

"Very few large scale applications of practical importance are NOT data intensive." – Alok Choudhary, IESP, Kobe, Japan, April 2012

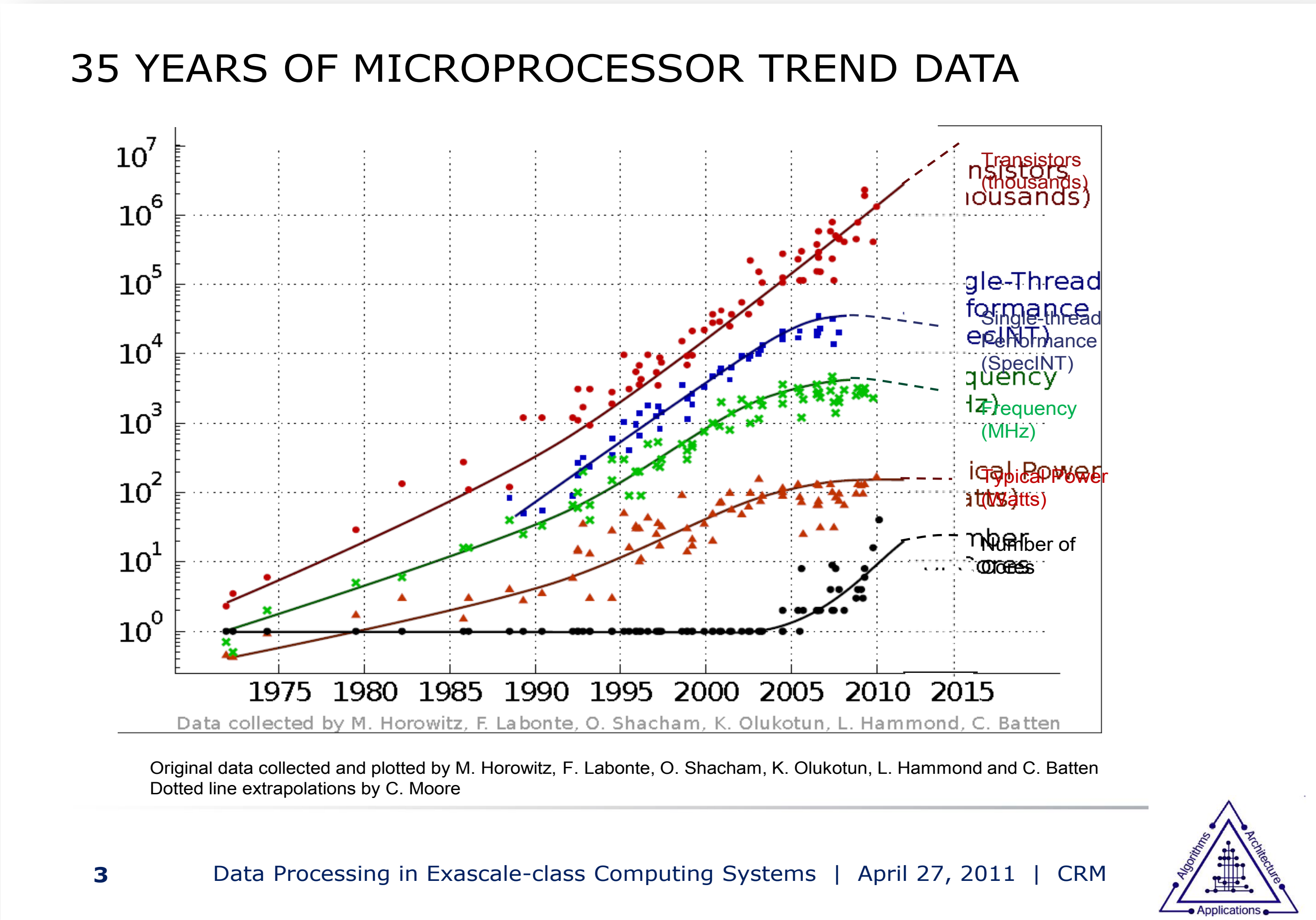
Exascale Computing

- **What isn't Exascale computing?**
  - Exaflops Linpack Benchmark Computer
  - Just a billion floating-point ALUs packaged together
- **What is Exascale computing?**
  - 1000X performance over a "Petaflops" system (exaflops sustained performance on complex, real-world applications)
  - Similar power and space requirements as a Petaflops computer
  - High programmability, generality, and performance portability
- **Required areas of investment**
  - New system designs / execution models
  - Enabling technologies (such as, processor, memory & interconnects)
  - System Software
  - Collaborative Environments

System Design Drivers

- **What is wrong with extrapolating current HPC designs?**
  - Insufficient parallelism exploited
  - Mechanisms don't exist for minimizing data movement
  - Does not exploit runtime information for efficiency, energy, reliability
  - Lacks global perspective
- **Today's bulk synchronous (BSP), distributed memory, communicating sequential processes (CSP) based execution model is approaching an efficiency, scalability, and power wall**

Enabling Technology Driver



Moore's Law continues

- Transistor count still doubles every 24 months

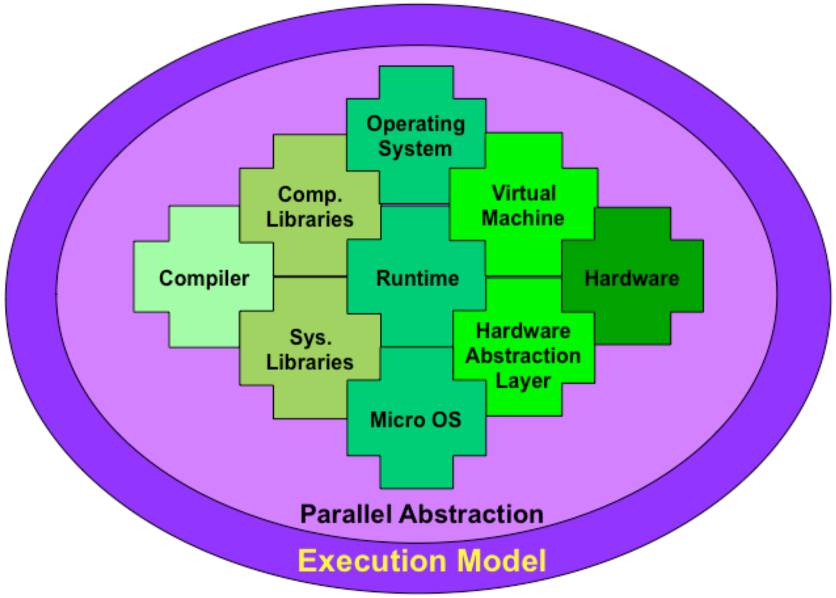
Dennard scaling stalls – key parameters flatline:

- Voltage
- Clock Speed
- Power
- Performance/ clock

U.S. DEPARTMENT OF ENERGY | Office of Science

System Software Driver

- **Current software stacks are built on 50 years of legacy system software**
- **Not designed to minimize energy and data movement**
- **Does not balance over-provisioned power and communication budgets**
- **Programming models don't allow full expression of parallelism, data locality, resiliency, …**
- **Resiliency must be cross-cutting, not just in the application layer**



Collaboration Driver

- **Current environment not optimal for:**
  - Dynamic, interactive workflow
  - Geographically distributed data, users, and facilities
  - Computational steering and real-time in-situ analysis
- **Need real-time comparison of simulation and experimental results, and control of active processes**

- **Shared real-time visualization - create unified teams from distributed users and facilities**

Being there *without being there*

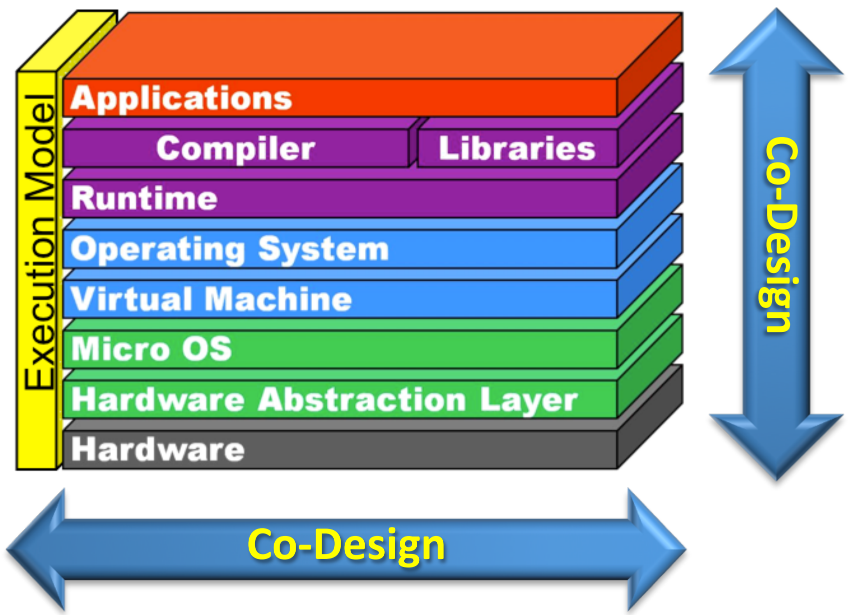
Exascale Research Imperative

- **Industry will not develop Exascale systems or technology alone**
  - Public needs (science, weather, …, not equal to large market)
  - Hand-held devices are the growing part of the market. PC's are a replacement business in the developed world. Servers growing slowly
- **Why Now?**
  - Conventional practices losing traction with future scaling needs
  - Microprocessors / Computing Systems are now at a crossroads in design space
  - Research needed now to set directions so that broadly useable systems become available. Need time to test ideas against metrics and get into design cycle
- **Effort requires a Co-design process**

Exascale Co-Design

- **Application-driven co-design is the process by which:**
  - Scientific problem requirements guide computer architecture and system software design
  - Technology capabilities and constraints inform formulation and design of algorithms and software

- **Shared global perspective across the design-space establishes shared conceptual framework for co-design and interoperability**
  - Parallelism
  - Latency
  - Overhead
  - Dependability



U.S. DEPARTMENT OF ENERGY | Office of Science

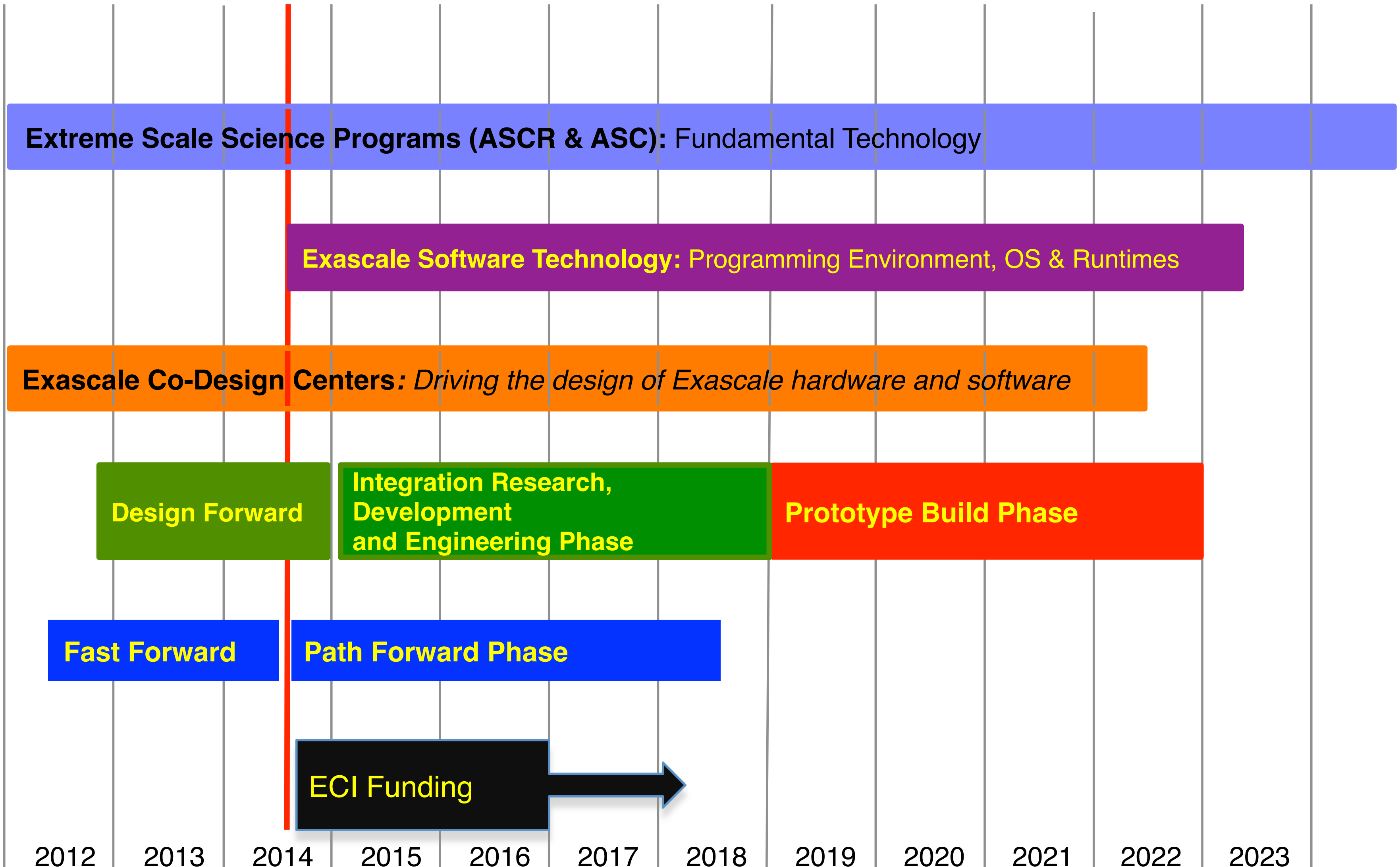
Exascale Computing Initiative (ECI)

- **Perform research, development and integration required to deploy Exascale computers in 2020+**
- **Partnership involving:**
  - Government
  - Computer industry
  - DOE laboratories
  - Academia
- **Target System Characteristics**
  - 500 to 1,000 times more **performance** than a Petaflops system
  - **1 Billion** degrees of concurrency
  - **20 MW** Power requirement
  - **500** cabinets
  - **Development and execution time productivity improvements**
- **The ECI timeframe assumes Congressional funding in the fiscal 2014 budget**

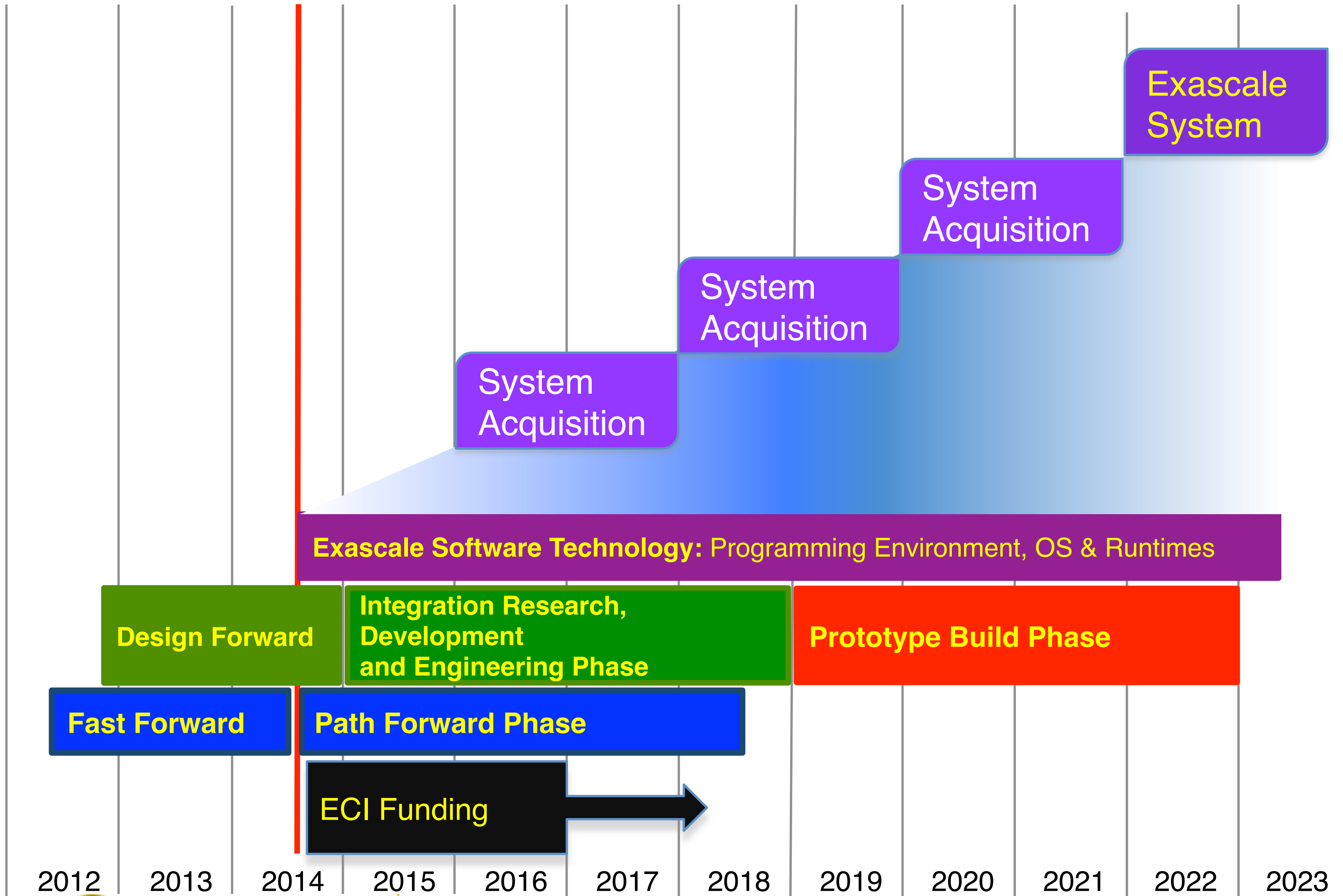
Exascale Computing Initiative Strategy

- Conduct critical supporting R&D efforts
- Develop Exascale software stacks
- Fund vendors to move technology from research to product space
- Fund the design and development of Exascale computer systems
- Effort driven by full workflow
- Collaborate with other government agencies and other countries
- Shared initiative with DOE / NNSA

Exascale Computing Initiative Timeline
(Anticipated)



Exascale Computing Initiative Timeline
System Perspective



First steps on a Path Towards Exascale:
Three Exascale Co-Design Centers Awarded

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

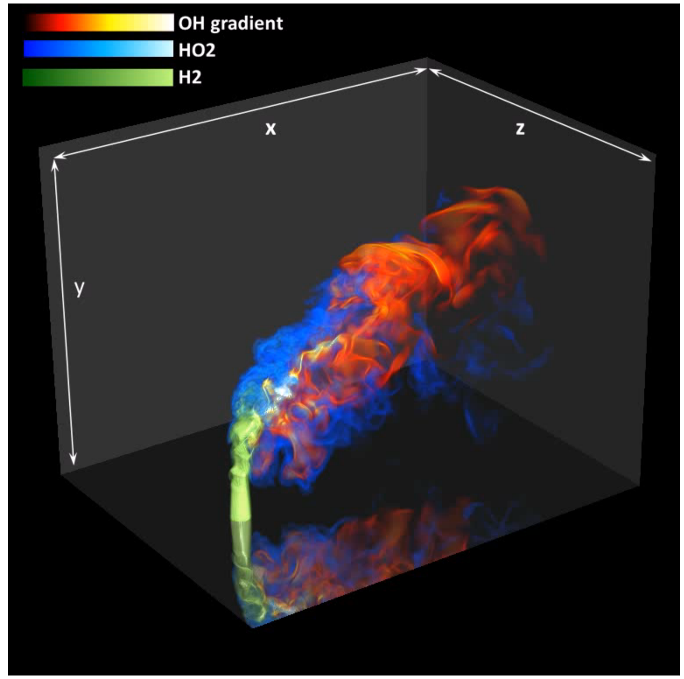
Director: Timothy Germann (LANL)

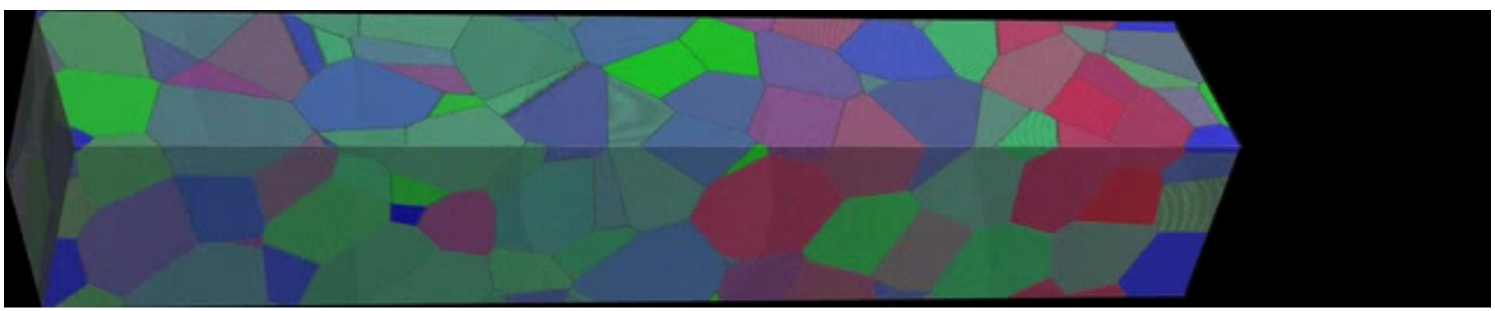
Center for Exascale Simulation of Advanced Reactors (CESAR)

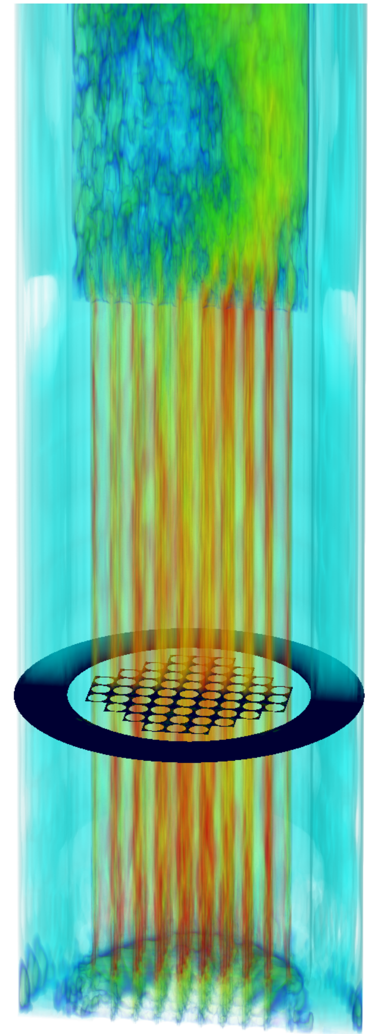
Director: Robert Rosner (ANL)

Combustion Exascale Co-Design Center (ExaCT)

Director: Jacqueline Chen (SNL)

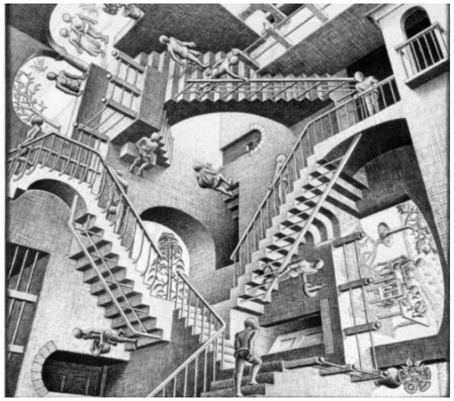


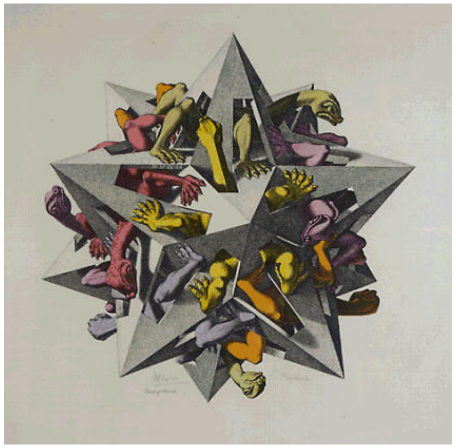




First Steps on a Path Towards Exascale:
X-Stack - Programming Environment

- **Motivation**
  - Scalability challenges, demanding unprecedented parallelism for Exascale
  - Energy efficiency, resilience, and heterogeneity challenges,
- **Goals and Objectives on an initial budget**
  - Initiate Exascale community prior to establishment of long-term Exascale program
  - Runtime systems and programming interfaces for efficiency, scalability, resilience, and performance
  - Power and resilience aware

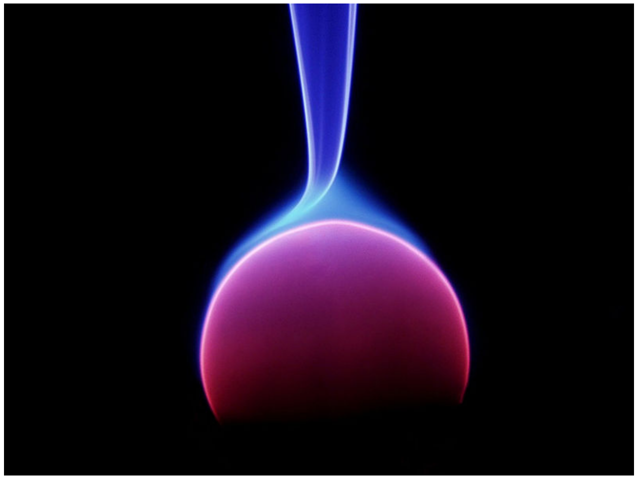




All M.C. Escher works © 2010 The M.C. Escher Company - the Netherlands. All rights reserved. Used by permission. www.mcescher.com

X-Stack Program Elements

- **New Programming Models:** approaches to managing parallelism and data movement through innovations in interfaces
- **Dynamic Runtime Systems:** adapt to changing application goals and system conditions
- **Locality-aware and Energy-efficient Strategies:** manage locality and minimize energy consumption
- **Interoperability:** facilitate interoperability across different HPC languages and interfaces, as well as cross-cutting execution models



First Steps on a Path Towards Exascale:
Fast Forward and Design Forward

- **Motivation**
  - **Fast Forward:** Early investments for processor, memory and storage technologies that are required to move research technology into vendor's products.
  - **Design Forward:** Early Investments for the design of Exascale systems, including integration and engineering preliminary efforts
- **Goals and Objectives on a Limited Budget**
  - Early investment in critical technologies that are required to achieve the Exascale goals. Start the design effort for Exascale systems. The two projects will work closely together to ensure that program goals are achieved.
  - Bridge funding prior to the start of the Exascale Computing Initiative.

At this time, the Design Forward RFP has not been released

Fast Forward Projects

Fast Forward

- Jointly funded by SC & NNSA, LLNL managed the RFP and contracting processes
- Two year contracts, started July 1, 2012

Project Goals & Objectives

- Initiate partnerships with multiple companies to accelerate the R&D of critical technologies needed for extreme-scale computing.
- Fund technologies targeted for productization in the 5–10 year timeframe.

Vendor	SCOPE	Value
AMD	Processor / Memory	$12,600,000
IBM	Memory	$10,476,714
Intel	Processor / Memory	$18,963,437
NVIDIA	Processor	$12,398,893
WhamCloud (Intel)	Storage & I/O	$7,996,053
Total		**$62,435,097**

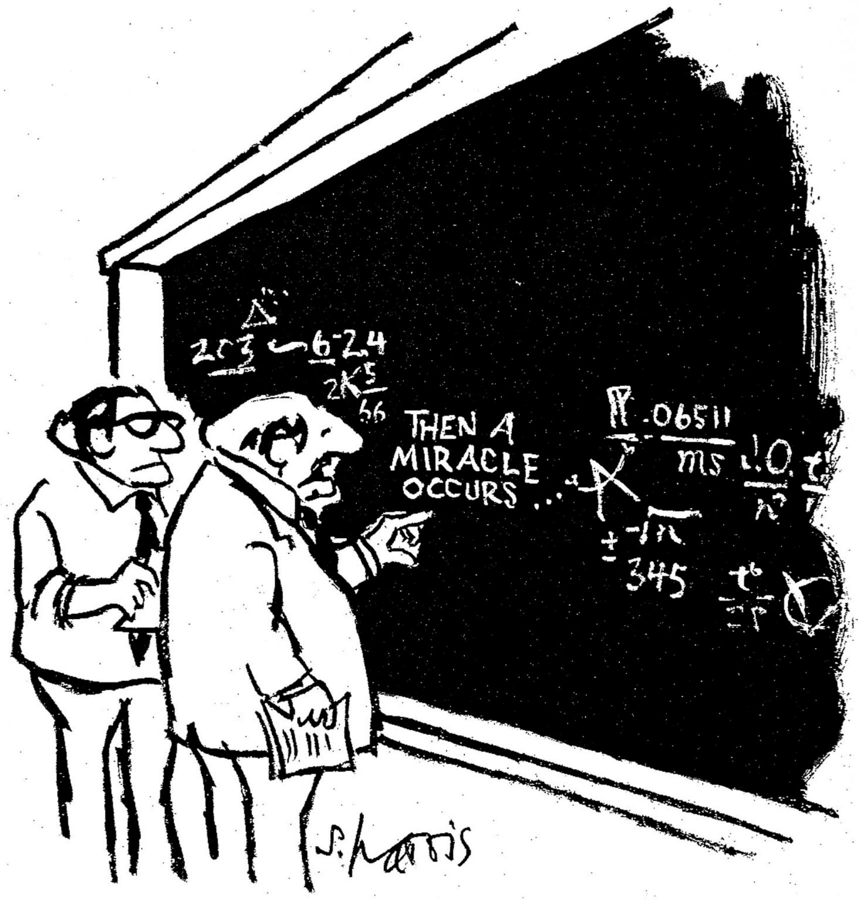
The Rest of the Story

- **New Class of Systems must have ecosystem to support future development**
- **“Legacy” applications will need to compile and execute, but may not always achieve full potential of system**
- **The current and next generation of programmers must be prepared for the new class of systems**

ScienceCartoonsPlus.com

Three "Miracles" We Need to Achieve

1. New paradigm (execution model) for technical computing that enables **scalable, energy efficient, and easily programmable** systems
2. **20 pJ per average operation** executed on the system, including the software overhead and data movement
3. Economically **viable technical computing vendor industry** and user community.



"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

U.S. DEPARTMENT OF ENERGY | Office of Science

Conclusions and Summary

- **Unique opportunity to create the future epoch of Exascale computing**
  - Empower an extraordinary suite of extreme scale science and engineering applications
- **We are setting a new direction for future generations of computing**
  - Substantial research is required and has begun
- **We have made an exciting beginning**
  - We need to work together as a cohesive community to achieve the shared goal

Luke: *All right, I'll give it a try.*

Yoda: *No. Try not. Do or do not. There is no try.*