# Accelerating the Pace of Discovery

John Hengeveld
Director of High Performance Computing Strategy
Intel Corporation

# The Compute Continuum

Cloud Service Data Delivery

Enterprise Transactional Computing

Technical Computing

Desktops | Laptops | Netbooks | Personal Devices | Smartphones | Smart TVs & Displays | Intelligent Devices

Technical Computing builds on ideas
to develop new products, experiences and science



Copyright © 2011, Intel Corporation. All rights reserved.

# Today's Agenda

**Accelerating the pace of discovery:**

- Technical computing's contribution to solving the world's most pressing challenges

**Crossing the Gap:**

- Enabling broader access to high performance computing

**A look to the horizon of insight:**

- Reaching Exascale and expanding HPC



Copyright © 2011, Intel Corporation. All rights reserved.

# The World's Need for Scientific Insight Has Never Been More Acute

Weather/Climate

New Forms of Energy

Healthcare



Copyright © 2011, Intel Corporation. All rights reserved.

# Driving Discovery in Addressing Global Climate Change

## The Challenge

- World population: 9.22B by 2075
- Managing weather, climate, water and environment prediction is critical to the long term health of our planet

## Scientific Progress

- Multi-day weather prediction weather prediction achieved
- Ocean current simulation enables sea current insight and study



Copyright © 2011, Intel Corporation. All rights reserved.

# Creating New Paths to Fuel Our World

## The Challenge

- Accurate exploration of reservoirs
- Oil and gas production from complex reservoirs
- Reduced cost of hardware and maintenances?

## Scientific Progress

- Found new oil reserves and developing new drilling techniques to access them
- Modeling Oil Diffusion and environmental risks
- Breakthroughs in fusion research



Copyright © 2011, Intel Corporation. All rights reserved.

# Delivering New Insights in Curing Disease and Forging New Paths to Health

## The Challenge

Cost of first AML genome was 1.6Million

Today cost is ~30K – imagine when cost drops under 1K. Personalized medicine?

## Scientific Progress

- Genome mapping delivered
- Full body medical imaging enabling visualization to improved diagnosis



Copyright © 2011, Intel Corporation. All rights reserved.


# Big Data, HPC, and Cancer

Alex Bayen, Armando Fox, Michael Franklin,
Michael Jordan, Anthony Joseph, Randy Katz,
**David Patterson**, Ion Stoica, Scott Shenker

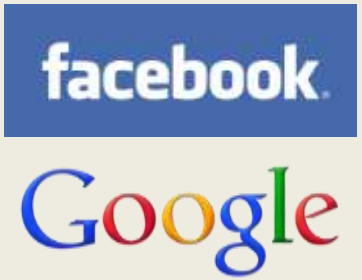
UC Berkeley

September, 2011

# Big Data is …

**Massive**

- Facebook: 200-400TB/day:
  83 million pictures
- Google: > 25 PB/day processed data




**Growing**

- More devices (cell phones),
  More people (3rd world),
  Bigger disks (2TB/$100)

**Dirty**

- Diverse, No Schema, Uncurated,
  Inconsistent Syntax and Semantics



amplab

# “Big Data”: Working Definition

When the normal application of current technology doesn’t enable users to obtain **timely** and **cost-effective** answers of sufficient **quality** to data-driven questions

Challenge: Use machine learning
**A**lgorithms, HPC/cloud computing
**M**achines, and crowd-sourced
**P**eople to extract value from Big Data
while decreasing the cost of maintaining it


## Interesting Big Data for Academic Research?

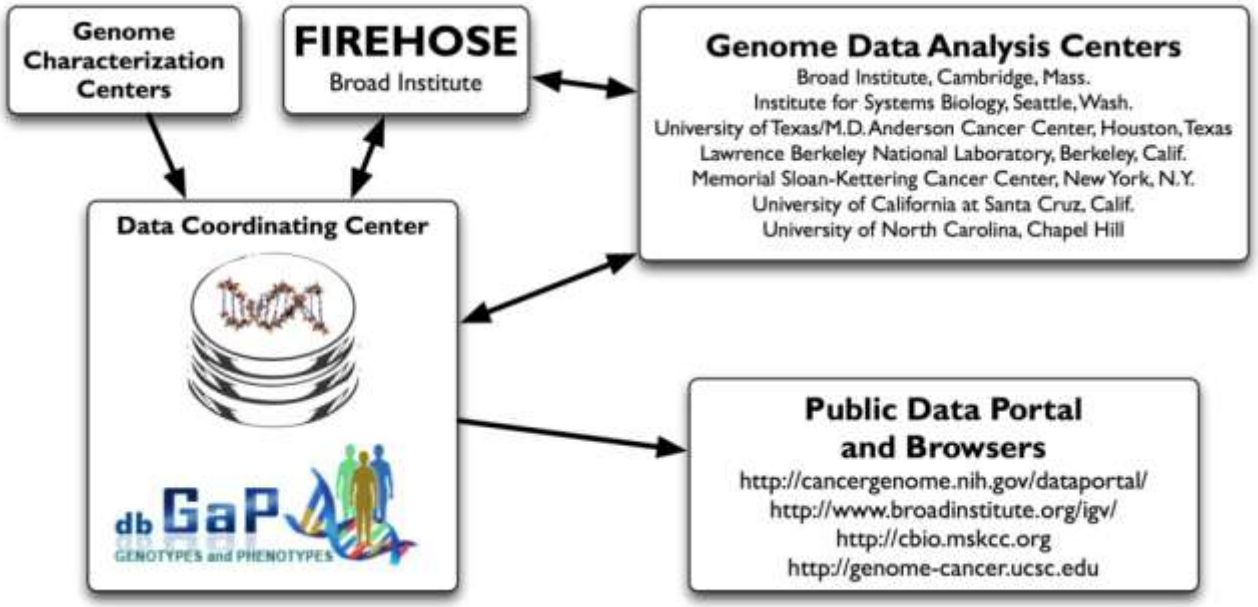


amplab

# Big Data Opportunity

## The Cancer Genome Atlas (TCGA)

- 20 cancer types, 500 tumors each: 5 petabytes
- David Haussler (UCSC) Datacenter online Oct 1
- OK to place Berkeley cluster next to 5 PB cluster
- Novelty: Academic Access yet Important Big Data






*Slide from David Haussler, UCSC, "Cancer Genomics," AMP retreat, 5/24/11*

# TCGA Potential Impact?

"We fully expect that 10 years from now, each cancer patient is going to want to get a genomic analysis of their cancer and will expect customized therapy based on that information."

Brad Ozenberger
TCGA program director
"Cracking Cancer's Code"
*Time Magazine*
June 2, 2011

# Big Data, Genomics, and Cancer

**¼ US deaths, 7M/year worldwide**

**⅓ US women will get cancer**

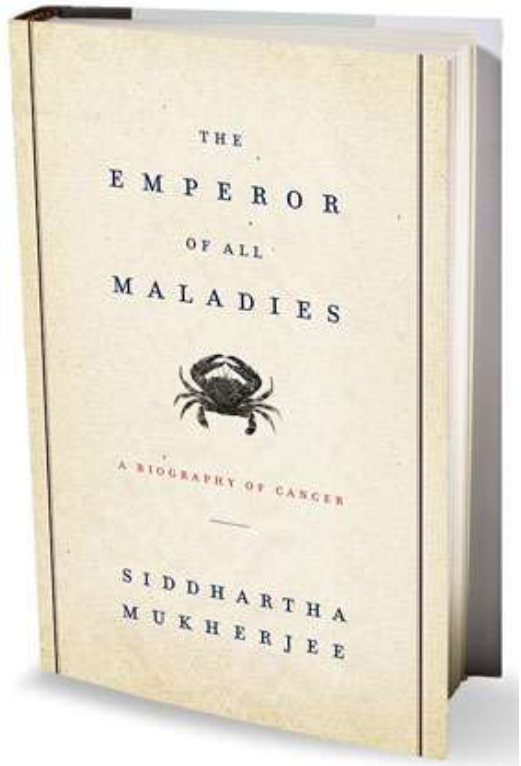
**½ US men will get cancer**

**Cancer: perversion of normal cell**

Limitless growth, evolves, spreads

**Cancer is a genetic disease**

Accidental DNA cell copy flaws + carcinogen-based mutations lead to cancer






amplab

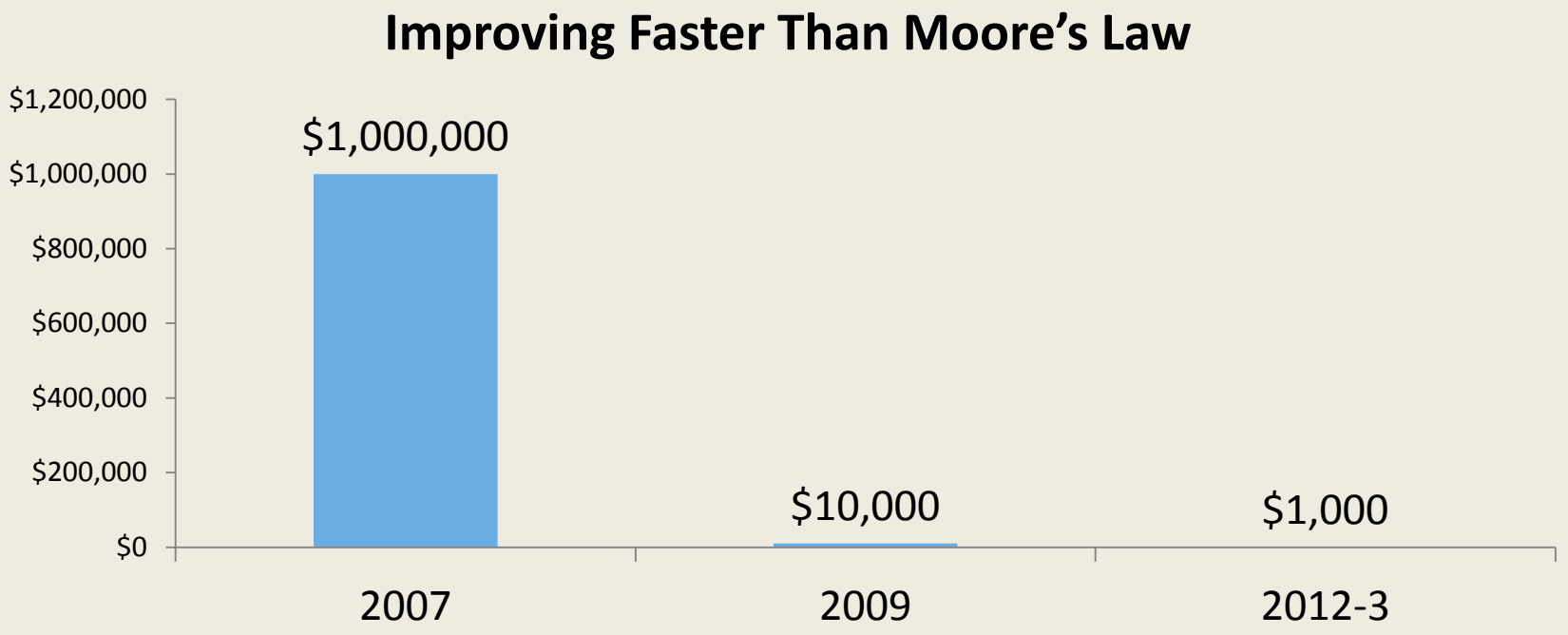
# 5 Steps to Customized Therapy

1. Sequencing machine that identifies many 300 base pair segments from a cancer tumor
2. Create full genome sequence of the cancer tumor from many segments (“alignment”)
3. Verify correctness of this sequence
4. Insights from comparing many other sequences and treatments to tumor genome
5. Diagnose and suggest of therapeutic targets for cure or non-progression based on tumor genome comparisons and patient records



amplab

# 1. Sequencing Machine Costs




2007 wet lab processing problem => 2012 digital processing problem

**Looks like sequencing machines not the bottleneck in speed *or* cost**



amplab

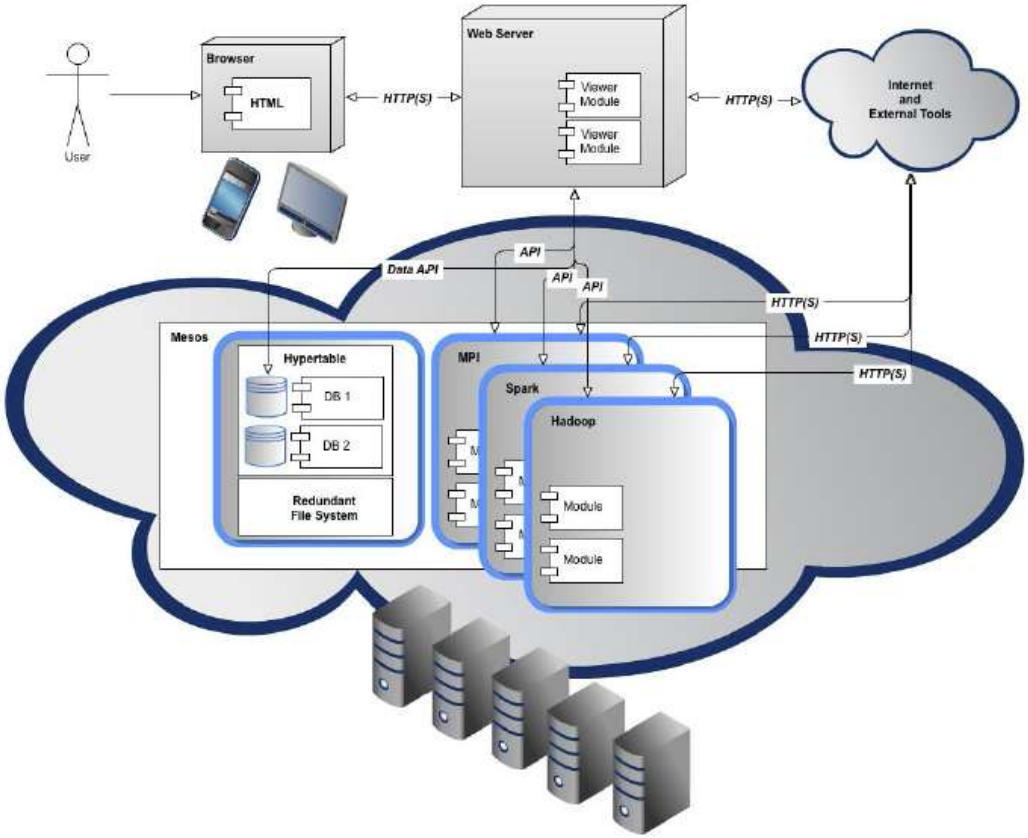
# Information Technology Obstacle?

"There is a growing gap between the generation of massively parallel sequencing output and the ability to process and analyze the resulting data. New users are left to navigate a bewildering maze of base calling, alignment, assembly and analysis tools with often incomplete documentation and no idea how to compare and validate their outputs. **Bridging this gap is essential, or the coveted $1,000 genome will come with a $20,000 analysis price tag**."

John D McPherson
"Next-generation gap,"
*Nature Methods*,
October 15, 2009

# 2 & 3. Full sequence and verification

UCSF prototype using open source SW, Cloud, Hadoop, Hypertable, Berkeley tools




>1 year on PC to <1 day in the cloud



amplab

# 4. Compare Sequences

**Machine Learning + Data Analytics**

**Cloud Programming Frameworks and Storage**

**(AMP Lab strengths)**

# 5. Clinical Diagnosis

**Suggest effective therapeutic targets for cure or stabilization**

- Often events are relatively rare mutations, and not identifiable by traditional statistical methods (solutions lie in long tail of rare mutations: tiny needle in huge haystack)

**Use Artificial Artificial Intelligence: (Crowd Sourcing)?**

- Imitate success of Astronomy's Galaxy Zoo or Biology's FoldIt?

# An Opportunity or Obligation?

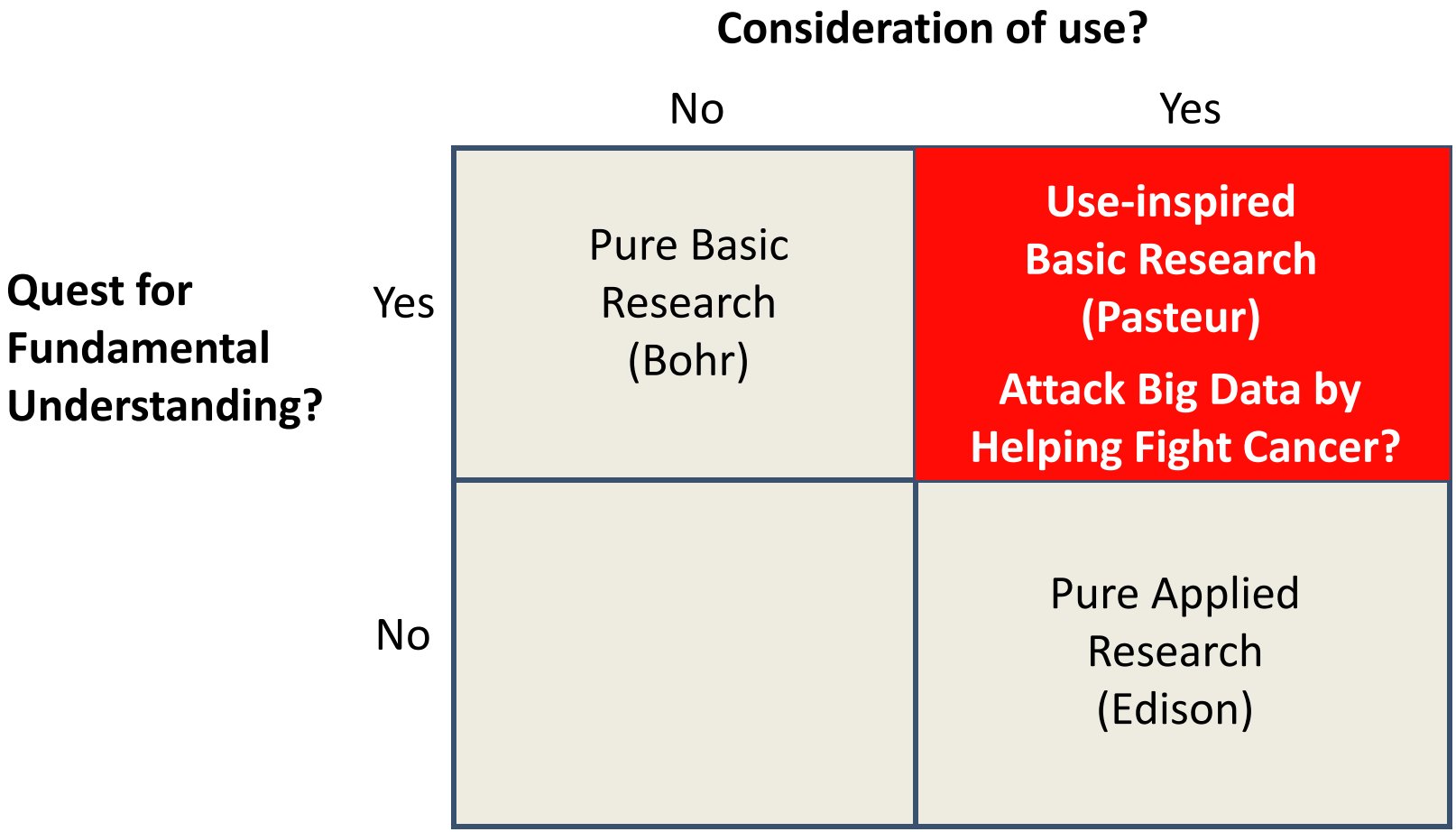
Given increasing genomic databases, next breakthroughs in cancer fight as likely to come from computer scientists as from biological scientists?

If it is plausible that CS could help millions of cancer patients live longer and better lives, as moral people, aren't we obliged to try?

(If need more motivation, huge future industry for customized, genome based medicine?)

# Comp Sci and Pasteur's Quadrant

**Research is inspired by:**






*From Pasteur's Quadrant: Basic Science and Technological Innovation, Donald E. Stokes, 1997*
*Slide from "Engineering Education and the Challenges of the 21st Century," Charles Vest, 9/22/09*

amplab

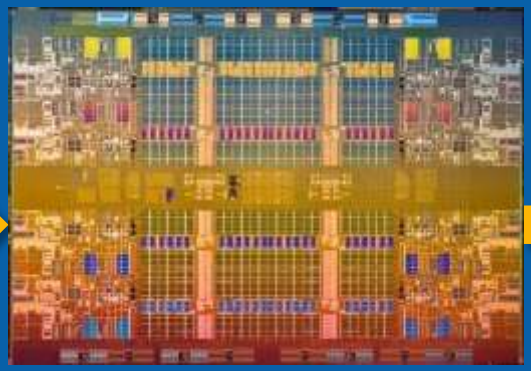
# Driving Future Insight

## Ecosystem Collaboration

Standards Based Solution Optimization and Delivery

Accelerate Time to Insight

## Products and Technology



Optimize Intel® Solutions to Address Emerging Workloads

Intel Xeon, MIC, Common Tools & Compilers

## Research Community

Optimize Computing to serve major Scientific and Research Challenges

Supercomputing to move forward Science

**Intel: Full Engagement on Delivery of Technical Compute Workload Optimization**

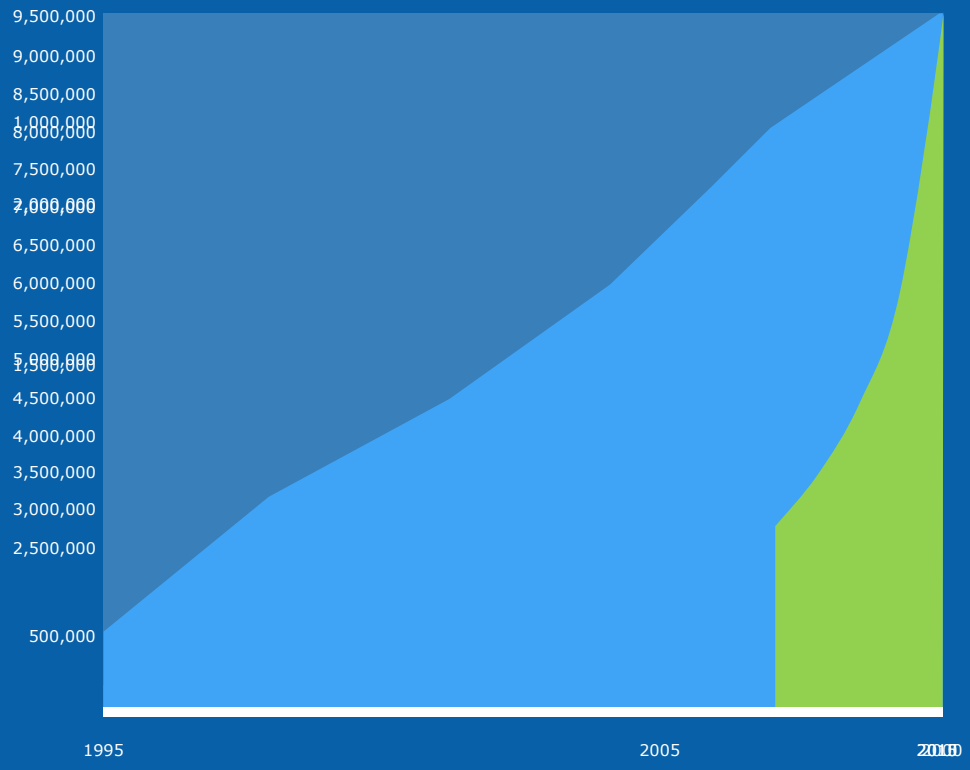


Copyright © 2011, Intel Corporation. All rights reserved.

# Transforming The Path to Discovery:

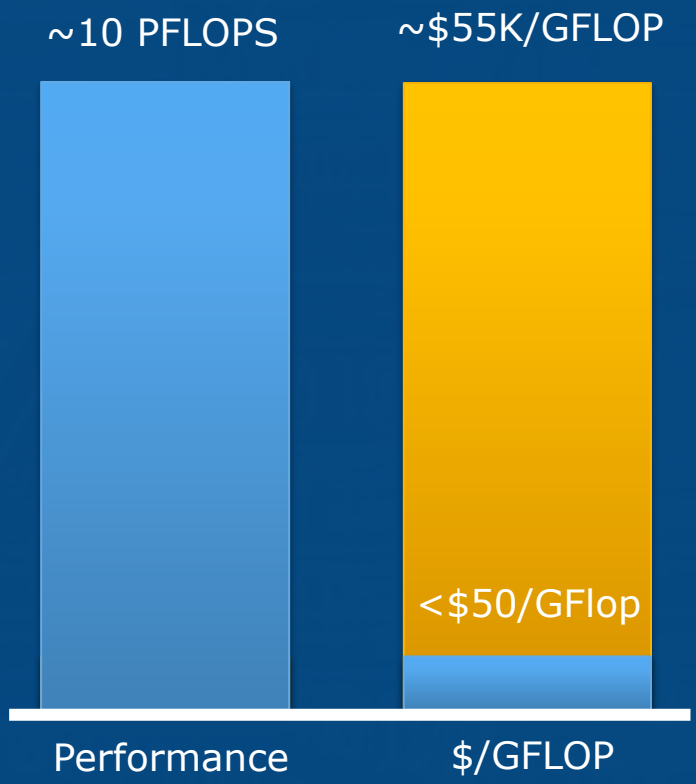
*The Growth of High Performance Computing*

## Annual Server Processor Shipments




Source: IDC

## Supercomputing in 2010






Copyright © 2011, Intel Corporation. All rights reserved.

# Improving HPC Solutions for Continued Transformation


**Power, Performance, Parallelism: The Three Pillars of HPC**



Copyright © 2011, Intel Corporation. All rights reserved.

# The Future Intel® Xeon® Processor E5

*Codenamed Sandy Bridge-EP*


Powerful.
Intelligent.

**Growing Performance**

- Up to 8 cores per socket
- Up to 2X FLOPS with Intel® Advanced Vector Extensions

**Efficient I/O**

- Integrated PCIe reduces latency and power
- Platform includes integrated 6Gb SAS for high performance local memory

**Advanced Security**

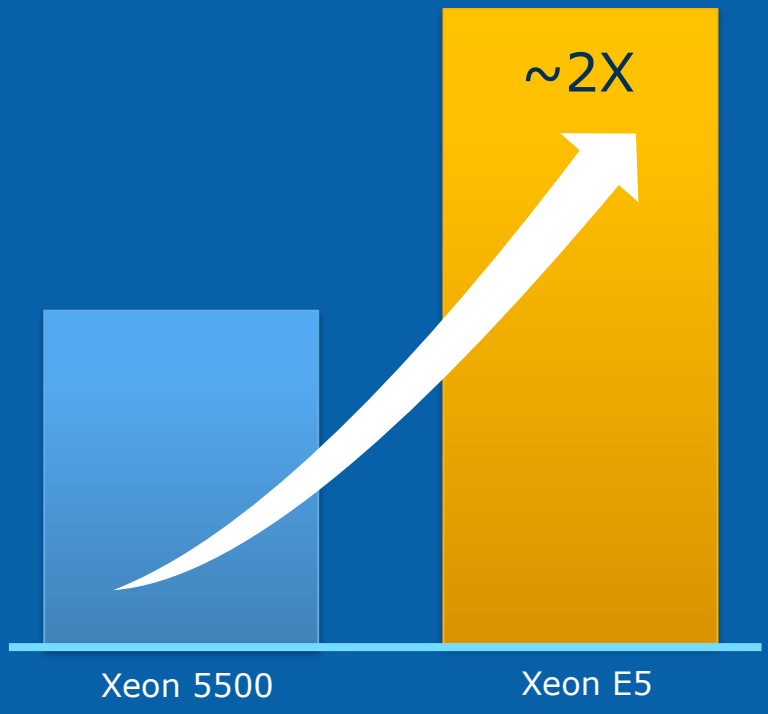
- Support for the latest Intel security features like Intel® Trusted Execution Technology and Intel® AES New instructions

**The Foundation of the Next Generation Datacenter**

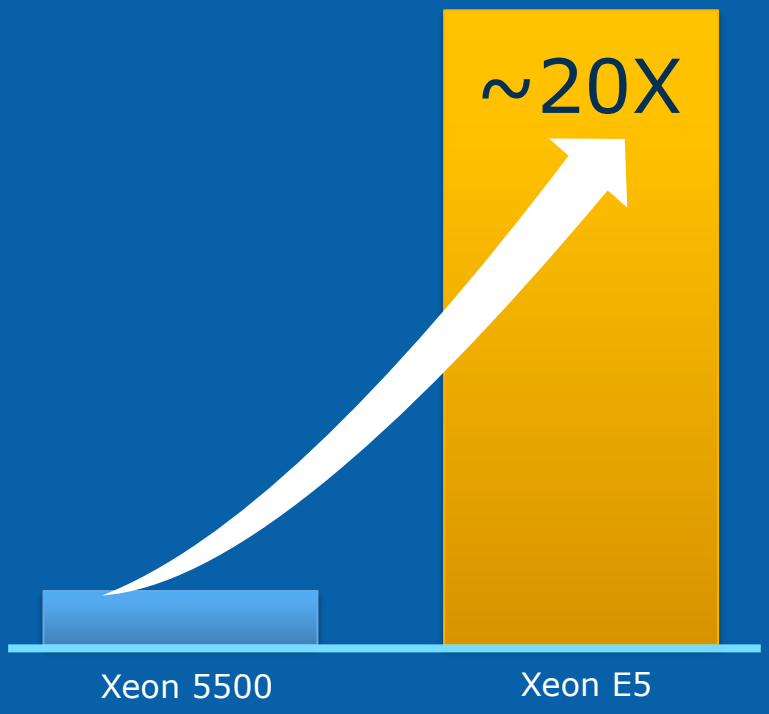


Copyright © 2011, Intel Corporation. All rights reserved.

# Intel® Xeon® E5: Broadest Xeon Product Line




Expect to Launch Almost 2X the Designs of Xeon 5500/5600

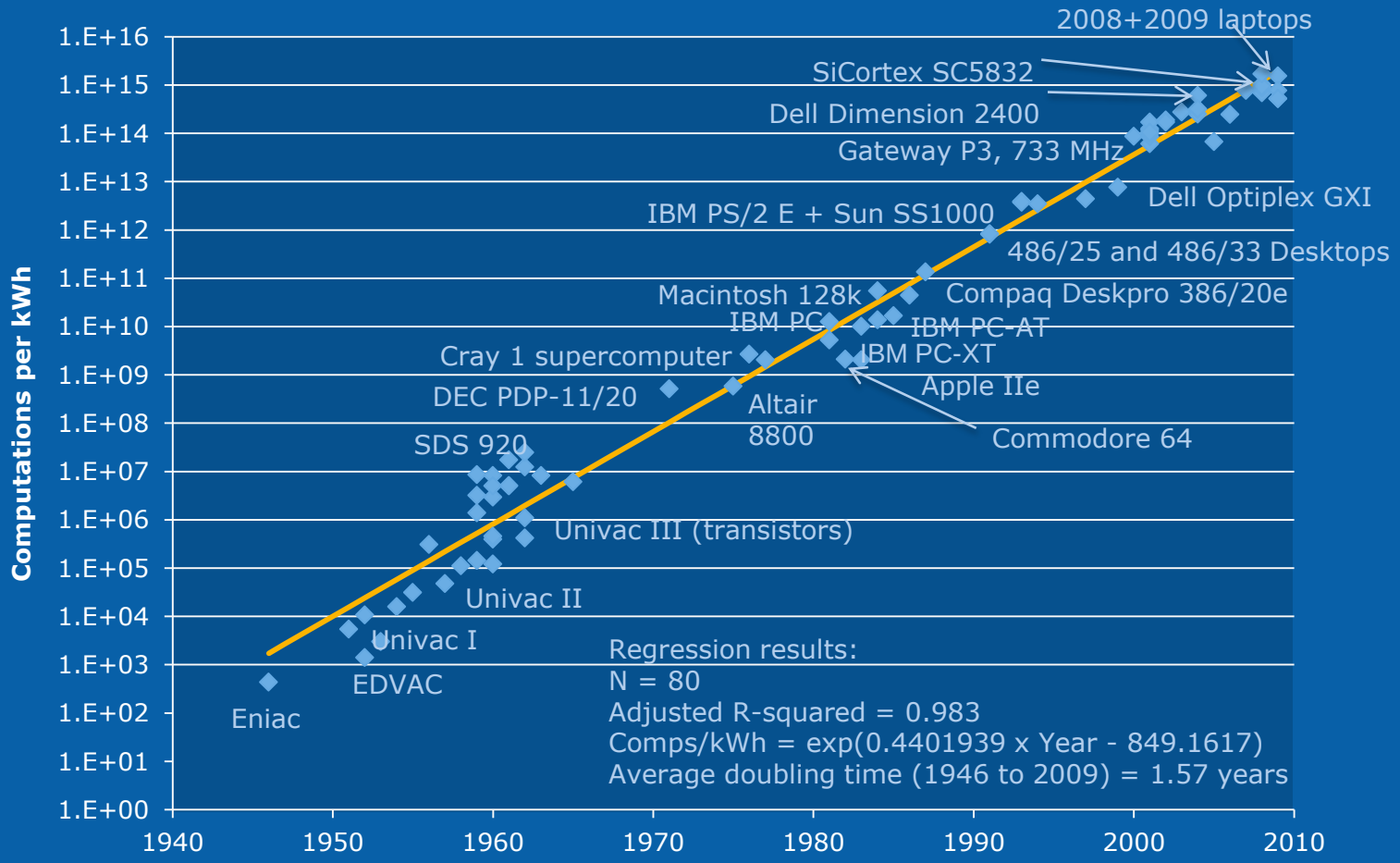



Cloud and Supercomputing Drive Unprecedented Initial Demand



Copyright © 2011, Intel Corporation. All rights reserved.

# Introducing "Koomey's Law"

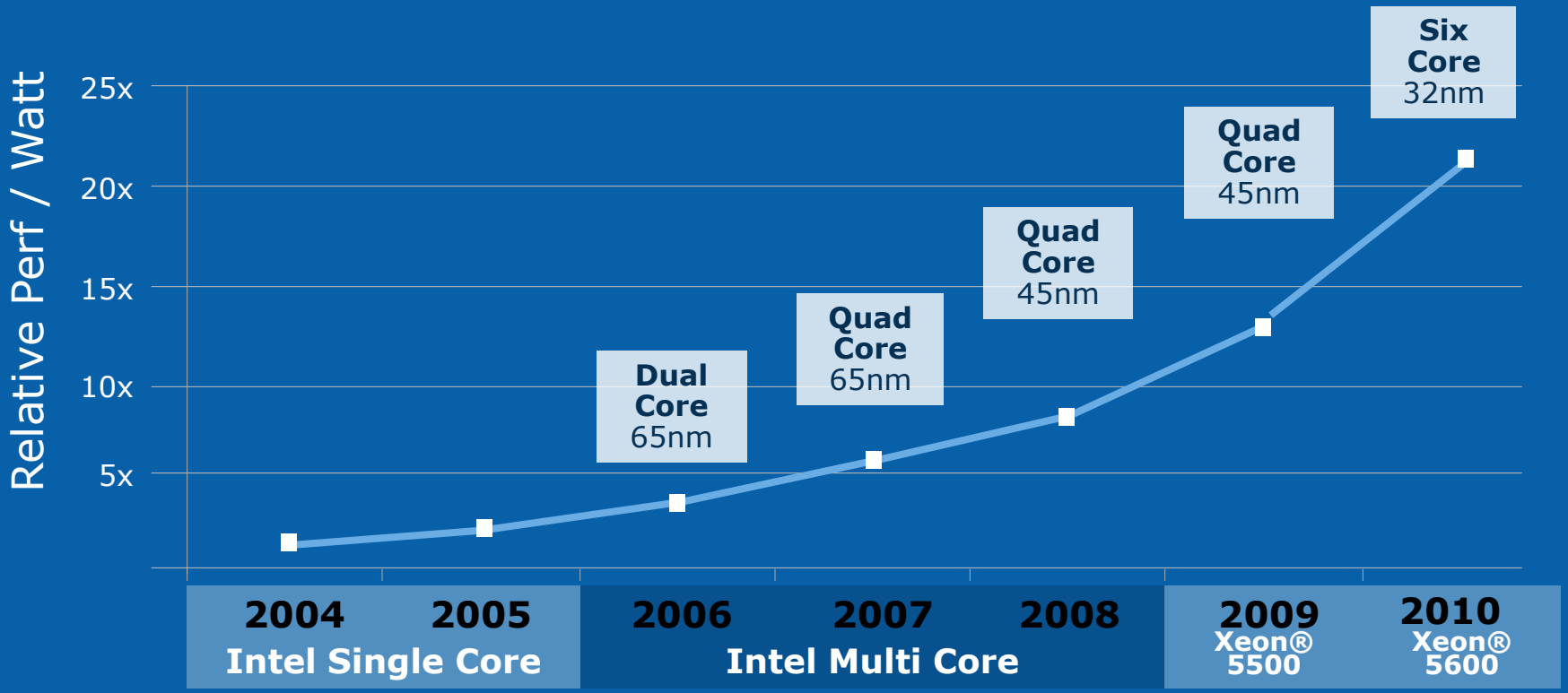





Copyright © 2011, Intel Corporation. All rights reserved.

# 2S Server Power Efficiency

*Compelling Energy Efficient Performance Gains*




**Outstanding Performance Per Watt Improvement**

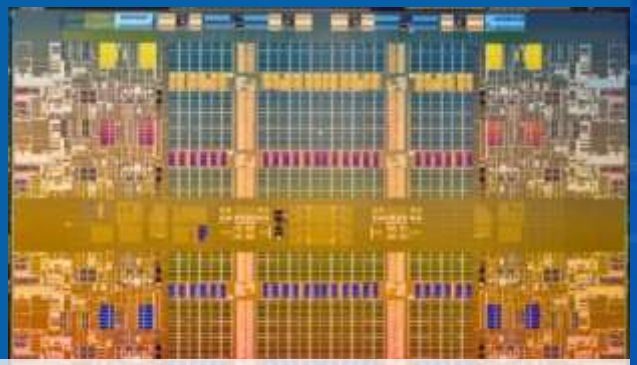
Source: Intel internal measurements using SPECjbb2005* as of Feb 2009.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit Intel Performance Benchmark Limitations

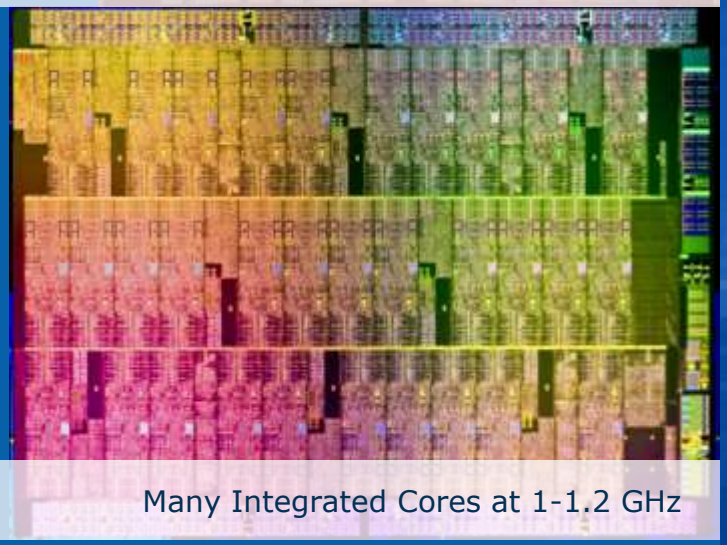


Copyright © 2011, Intel Corporation. All rights reserved.

# Intel's Many Core and Multi-core Engines



Multi-core Intel® Xeon® processor at 2.26-3.5 GHz



Many Integrated Cores at 1-1.2 GHz

Die Size not to scale

## Intel Xeon:

- Foundation of HPC Performance
- Suited for full scope of workloads
- Industry leading performance/watt for serial & highly parallel workloads.

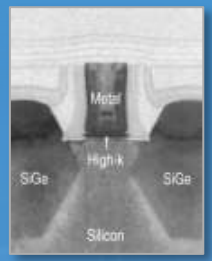
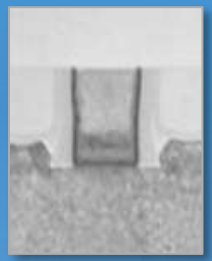
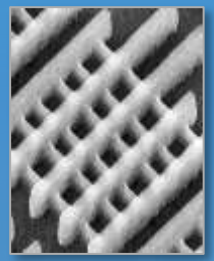
## MIC Architecture:

- Optimized for highly parallelized compute intensive workloads
- Common software tools with Xeon enabling efficient app readiness and performance tuning



Copyright © 2011, Intel Corporation. All rights reserved.

# Process Technology Leadership:
## *The Foundation for Performance Power, and Parallelism Innovation*

| 2003 | 2005 | 2007 | 2009 | 2011 |
|---|---|---|---|---|
| **90**nm | **65**nm | **45**nm | **32**nm | **22**nm |
|  |  |  |  |  |
| Invented SiGe Strained Silicon | 2nd Gen. SiGe Strained Silicon | Invented Gate-Last High-k Metal Gate | 2nd Gen. Gate-Last High-k Metal Gate | First to Implement Tri-Gate |



*Strained Silicon*

*High-k Metal Gate*

*Tri-Gate*

# 22nm

A Revolutionary Leap in Process Technology

**37%** Performance Gain at Low Voltage*

**>50%** Active Power Reduction at Constant Performance*

## The Foundation for All Computing

Source: Intel
*Compared to Intel 32nm Technology



Copyright © 2011, Intel Corporation. All rights reserved.

# Today's Agenda

**Accelerating the pace of discovery:**

- High performance computing's contribution to solving the world's most pressing challenges

**Crossing the Gap:**

- Enabling broader access to high performance computing

**A look to the horizon of insight:**

- Reaching Exascale and expanding HPC



Copyright © 2011, Intel Corporation. All rights reserved.

# Is it Enough?

**Lack of Access Remains Critical**



Copyright © 2011, Intel Corporation. All rights reserved.

# Opportunity in Industry:

*Benefits of Digital Testing & Refining of Product Designs*

Simulation

Digital Content Creation

**Digital Development Opportunity is Acute**

- Lower R&D costs[1]
  - ~½ the number of physical prototypes
  - ~½ prototyping costs
- Accelerated product delivery[2]
  - Competitive advantage

**The Cost of Dedicated Digital Development Remains a Barrier**

- Intel® Xeon® dual-processor workstations standard for professional creators
- But access to digital render farms still limited
- External cloud can provide an alternative to dedicated systems for peak load
  - Dreamworks[3] used 11m hours of external cloud rendering creating Kung Fu Panda (20% of total)

## Full Time Investment for Intermittent Use

Source: 1. Aberdeen group
Source: 2. Aberdeen 2010 Manufacturing Survey
Source: 3. Dreamworks



Copyright © 2011, Intel Corporation. All rights reserved.

# Bridging the Great Divide: Seamless Access to Efficient Compute Services on Demand

**Our Challenges to Getting There:**

- Technical & Social Barriers
- New economic Models for Shared High Performance Computing
- New Standards

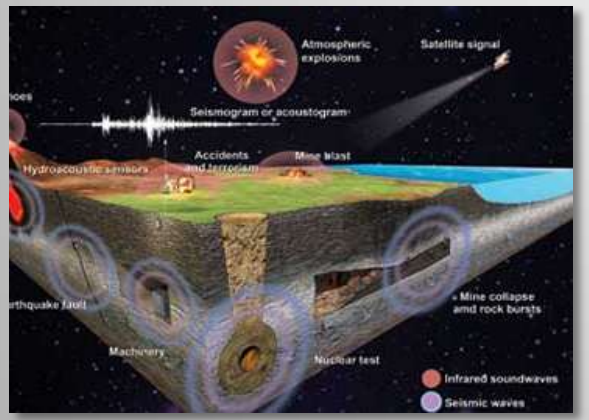


Copyright © 2011, Intel Corporation. All rights reserved.

# Solving global security challenges for the nation

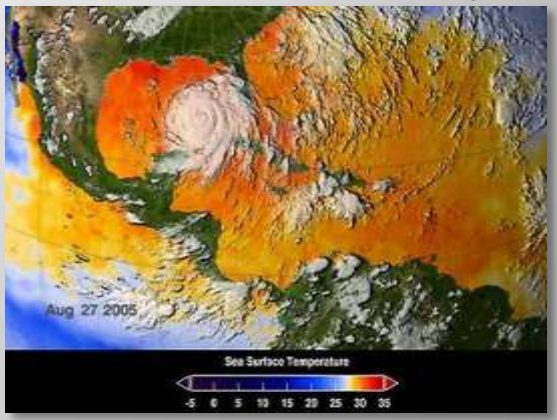
*Multidisciplinary science, technology, and engineering*

**Nuclear Security**



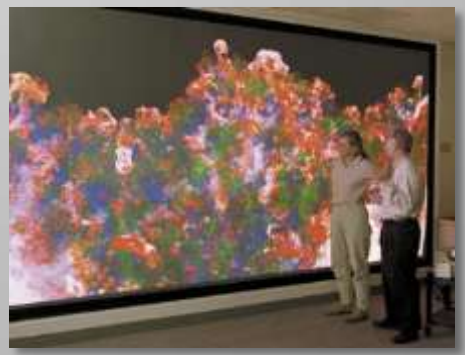

**International and Domestic Security**

**Energy and Environmental Security**



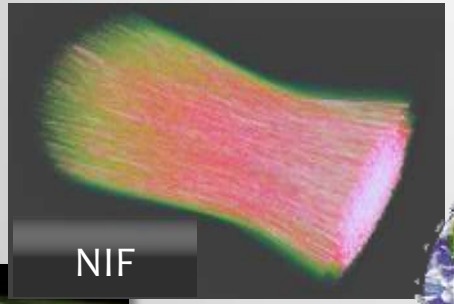

**Basic Science**

**Engineering**

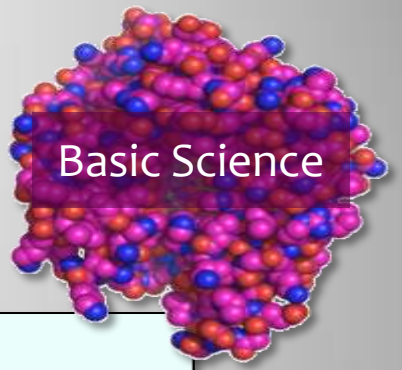


**Computing**

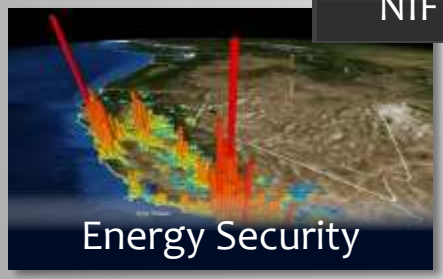
# Modeling, simulation and analysis are central to nearly every LLNL program







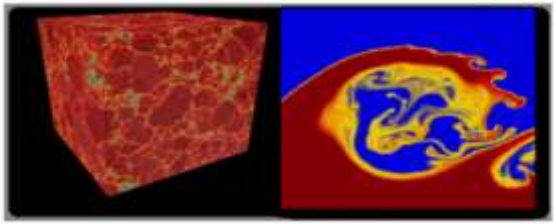






**Predictive Integrated Codes**

**Physics and Engineering Models**



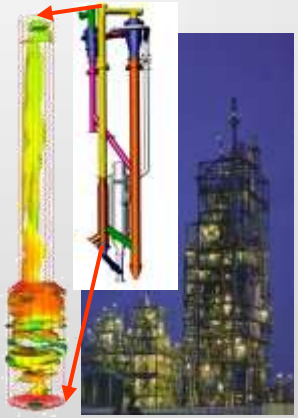
+

**Computers**

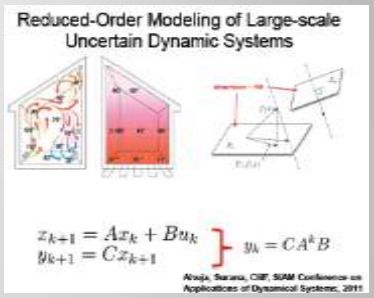
+

**Experimental Verification and Validations**

# High Performance Computing is also essential to economic competitiveness: Broadening the base

Carbon Capture




Building Energy Efficiency

Electric Grid

Addressing barriers of HPC adoption due to

- Lack of expertise
- Lack of appropriate software
- Cost

Building new programs through application of HPC and computational science

# Livermore Valley Open Campus

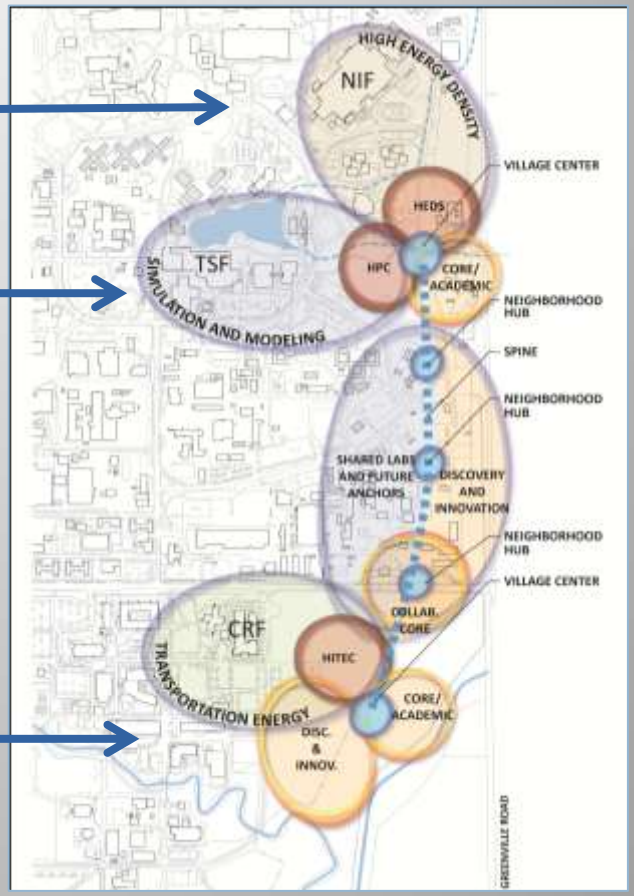
- Campus-like environment with collaborative space
- Building-level security
- Ready access for all partners, including foreign nationals
- Wireless capability and unclassified computing
- Synergy with community plans for economic growth

International Center for High Energy Density and Inertial Fusion Energy Science
***National Ignition Facility***

High Performance Computing
***High Performance Computing Capabilities and Facilities***

Transportation Energy Center
***Combustion Research Facility***

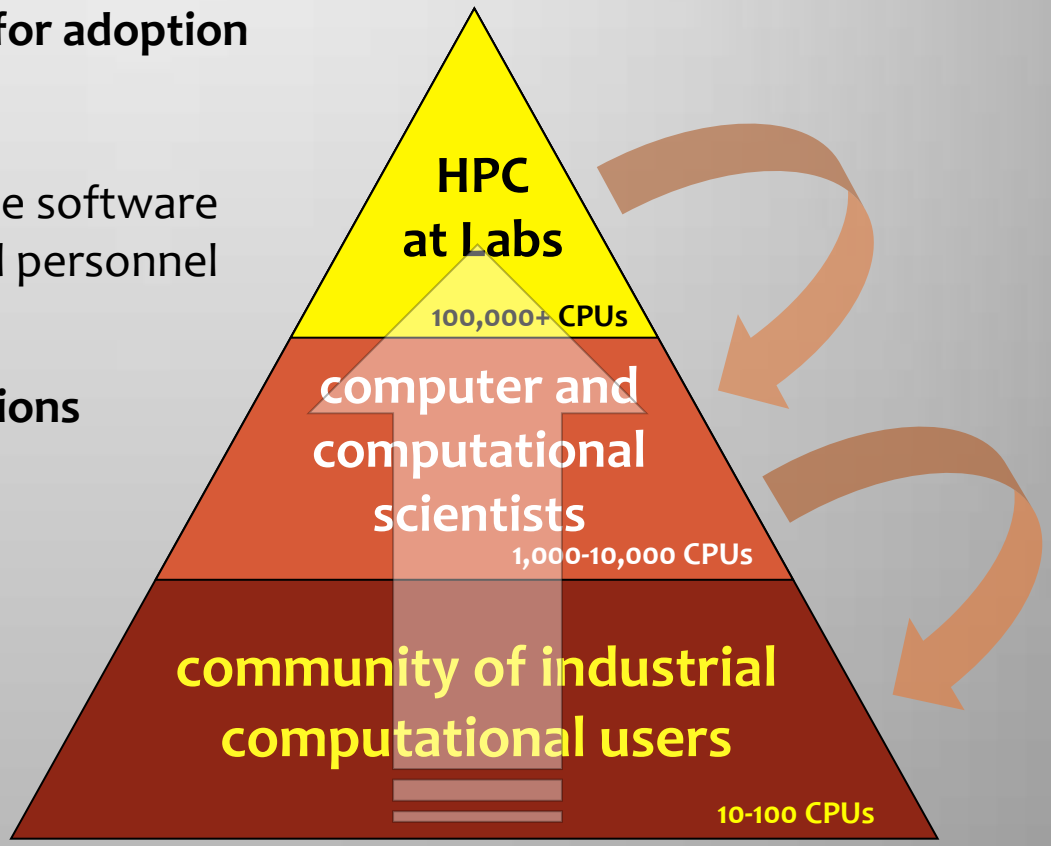
**Flad architects completed several planning scenarios; "Village centers" is the preferred option**



# LLNL HPC Innovation Center

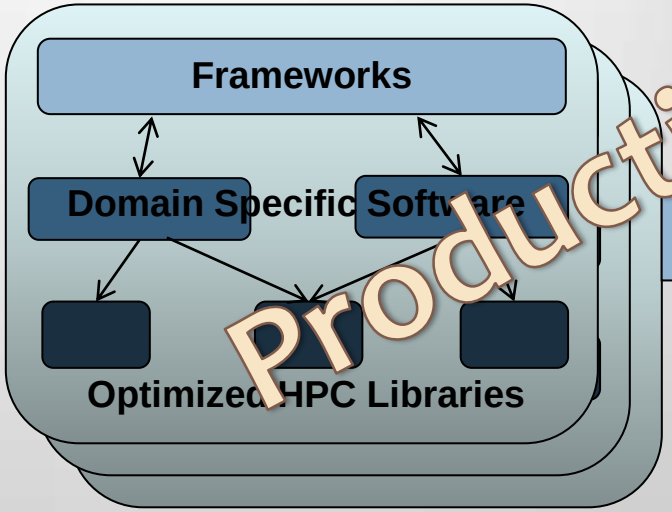
…engaging with U.S. industry to enhance American economic competitiveness by promoting the adoption of high performance computing.

- **Lowering the barriers for adoption of HPC:**
  - High cost of entry
  - Lack of appropriate software
  - Shortage of skilled personnel
- **Deliver true business solutions for our industrial partners**
- **Build and nurture an HPC innovation ecosystem in the Livermore Valley Open Campus**



# HPC-IC and LLNL will work with vendors and industry to develop HPC standards, tools, and software to ease use of HPC speed development


Ease of Use




Productivity

Ensembles of calculations

e.g. Uncertainty Quantification (UQ) Pipeline

HPC-IC partners

Vendors / co-design

LLNL programs

Leading edge development of reusable software

Innovation

Rapid prototyping, development, and deployment

Portable and scalable software

Competitive advantage

# Today's Agenda

**Accelerating the pace of discovery:**

- High performance computing's contribution to solving the world's most pressing challenges

**Crossing the Gap:**

- Enabling broader access to high performance computing

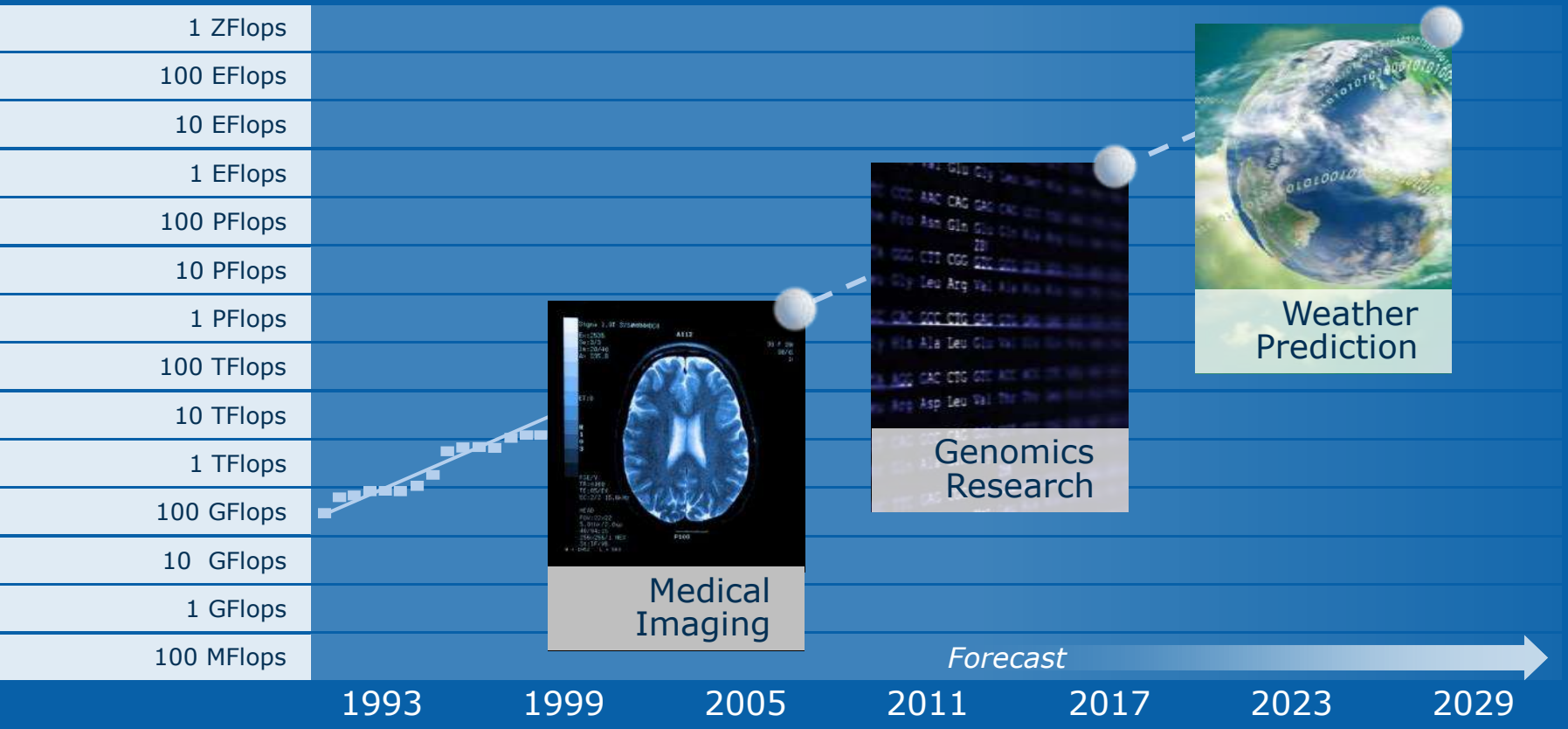
**A look to the horizon of insight:**

- Reaching Exascale and expanding HPC



Copyright © 2011, Intel Corporation. All rights reserved.

# An Insatiable Need For Computing




**Exascale Problems Cannot Be Solved Using the Computing Power Available Today**

Source: www.top500.org



Copyright © 2011, Intel Corporation. All rights reserved.

# Intel Labs & HPC

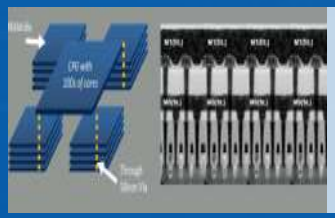
## Strong Research Partnerships

Industry

Government

Universities

## World Class Research in HPC

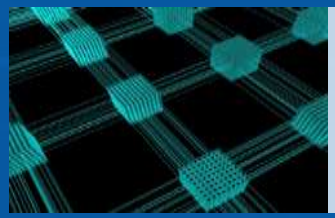


Memory Stacking & Technologies

Security


Interconnect Technologies

Silicon Photonics



Programm-ability

Power Reduction

**Delivering Breakthrough Technologies to Fuel Innovation**

* Other names, logos and brands may be claimed as the property of others.



Copyright © 2011, Intel Corporation. All rights reserved.

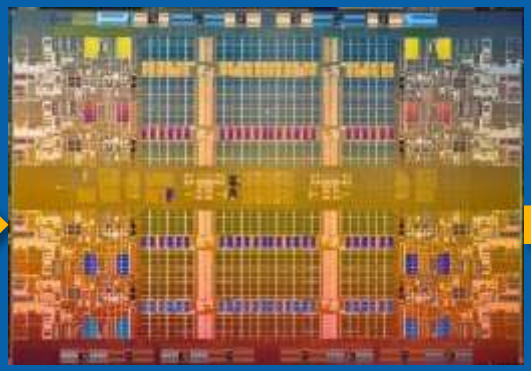
# Driving Future Insight

## Ecosystem Collaboration

Standards Based Solution Optimization and Delivery

Accelerate Time to Insight

## Products and Technology



Optimize Intel® Solutions to Address Emerging Workloads

Intel Xeon, MIC, Common Tools & Compilers

## Research Community

Identify Computing's Role in Scientific and Research Challenges

Top 500 Exascale Insight

**Intel: Full Engagement on Delivery of Technical Compute Workload Optimization**



Copyright © 2011, Intel Corporation. All rights reserved.

# In Summary

**The world's most pressing challenges continue to be solved with the help of Intel based technical computing**

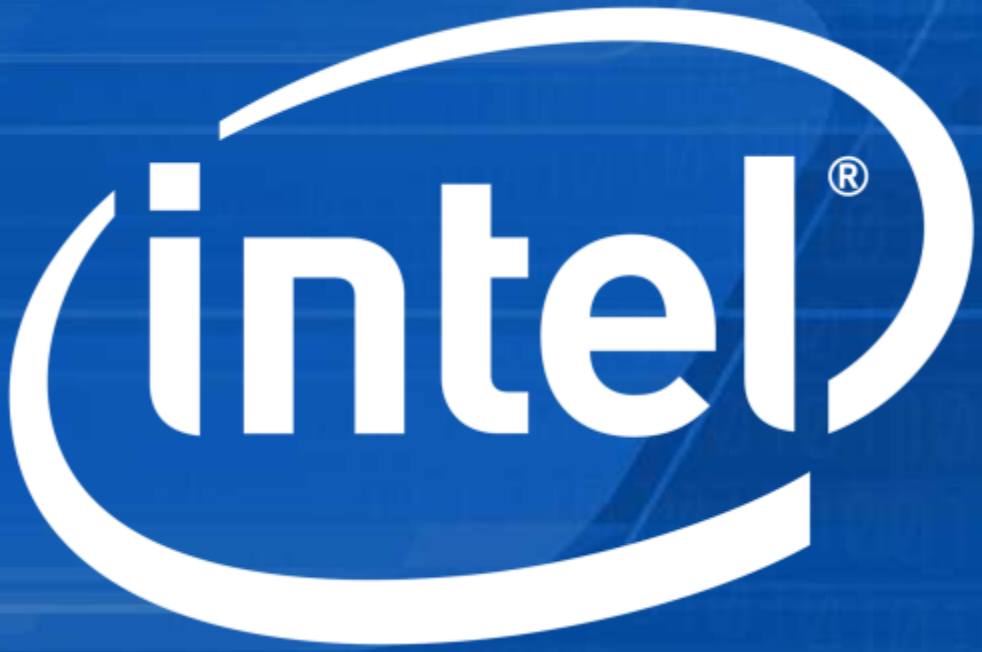
**The path ahead offers new opportunities in both continued performance transformation and creating gateways for new access**

**Intel is committed to help deliver solutions to meet these new opportunities**

**Intel is collaborating across the world to help achieve a new era of insight .. transforming scientific discovery once again**



Copyright © 2011, Intel Corporation. All rights reserved.


intel

# Legal Disclaimers

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Intel does not control or audit the design or implementation of third party benchmarks or Web sites referenced in this document. Intel encourages all of its customers to visit the referenced Web sites or others where similar performance benchmarks are reported and confirm whether the referenced benchmarks are accurate and reflect performance of systems available for purchase.

Relative performance is calculated by assigning a baseline value of 1.0 to one benchmark result, and then dividing the actual benchmark result for the baseline platform into each of the specific benchmark results of each of the other platforms, and assigning them a relative performance number that correlates with the performance improvements reported.

SPEC, SPECint, SPECfp, SPECrate. SPECpower, SPECjAppServer, SPECjbb, SPECjvm, SPECWeb, SPECompM, SPECompL, SPEC MPI, SPECjEnterprise* are trademarks of the Standard Performance Evaluation Corporation. See http://www.spec.org for more information. TPC-C, TPC-H, TPC-E are trademarks of the Transaction Processing Council. See http://www.tpc.org for more information.

Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, virtual machine monitor (VMM) and, for some uses, certain platform software enabled for it. Functionality, performance or other benefits will vary depending on hardware and software configurations and may require a BIOS update. Software applications may not be compatible with all operating systems. Please check with your application vendor.

Hyper-Threading Technology requires a computer system with a processor supporting HT Technology and an HT Technology-enabled chipset, BIOS and operating system. Performance will vary depending on the specific hardware and software you use. For more information including details on which processors support HT Technology, see here

Intel® Turbo Boost Technology requires a Platform with a processor with Intel Turbo Boost Technology capability. Intel Turbo Boost Technology performance varies depending on hardware, software and overall system configuration. Check with your platform manufacturer on whether your system delivers Intel Turbo Boost Technology. For more information, see http://www.intel.com/technology/turboboost

No computer system can provide absolute security under all conditions. Intel® Trusted Execution Technology (Intel® TXT) requires a computer system with Intel® Virtualization Technology, an Intel TXT-enabled processor, chipset, BIOS, Authenticated Code Modules and an Intel TXT-compatible measured launched environment (MLE). Intel TXT also requires the system to contain a TPM v1.s. For more information, visit http://www.intel.com/technology/security. In addition, Intel TXT requires that the original equipment manufacturer provides TPM functionality, which requires a TPM-supported BIOS. TPM functionality must be initialized and may not be available in all countries.

Intel® AES-NI requires a computer system with an AES-NI enabled processor, as well as non-Intel software to execute the instructions in the correct sequence. AES-NI is available on Intel® Core™ i5-600 Desktop Processor Series, Intel® Core™ i7-600 Mobile Processor Series, and Intel® Core™ i5-500 Mobile Processor Series. For availability, consult your reseller or system manufacturer. For more information, see http://software.intel.com/en-us/articles/intel-advanced-encryption-standard-instructions-aes-ni/

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor series, not across different processor sequences. See http://www.intel.com/products/processor_number for details. Intel products are not intended for use in medical, life saving, life sustaining, critical control or safety systems, or in nuclear facility applications. All dates and products specified are for planning purposes only and are subject to change without notice

Copyright © 2011 Intel Corporation. All rights reserved. Intel, the Intel logo, Xeon and Intel Core are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. All dates and products specified are for planning purposes only and are subject to change without notice



Copyright © 2011, Intel Corporation. All rights reserved.