The Global Importance of HPC

Historical Progress, Accomplishments and Future Directions

Tilak Agerwala
VP, Systems

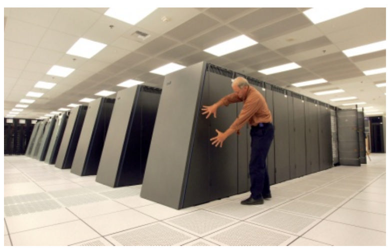
Leadership Supercomputing for Almost 60 Years

DoE and IBM scientists have shared 6 Gordon Bell Awards


IBM 701 at LLNL circa 1954


ASCI Blue Pacific, LLNL, c1999


ASCI White, LLNL 2001


Blue Gene/L, LLNL, 2005


ASC Purple, LLNL, 2005


Road Runner, LANL, 2008


Endeavor, ANL, 2008

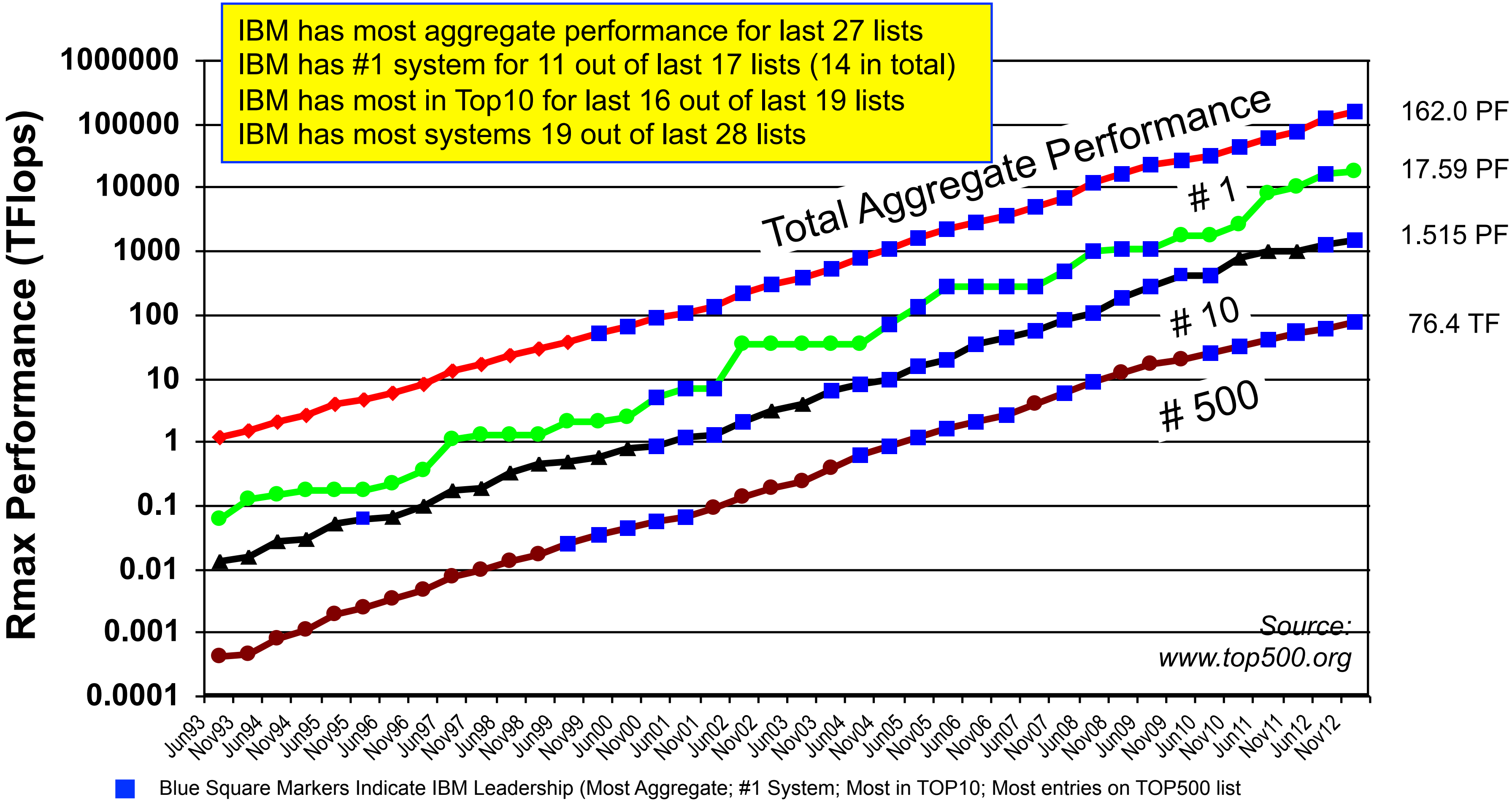

Dawn, LLNL, 2010


Mira, ANL, 2012


Sequoia, LLNL, 2012

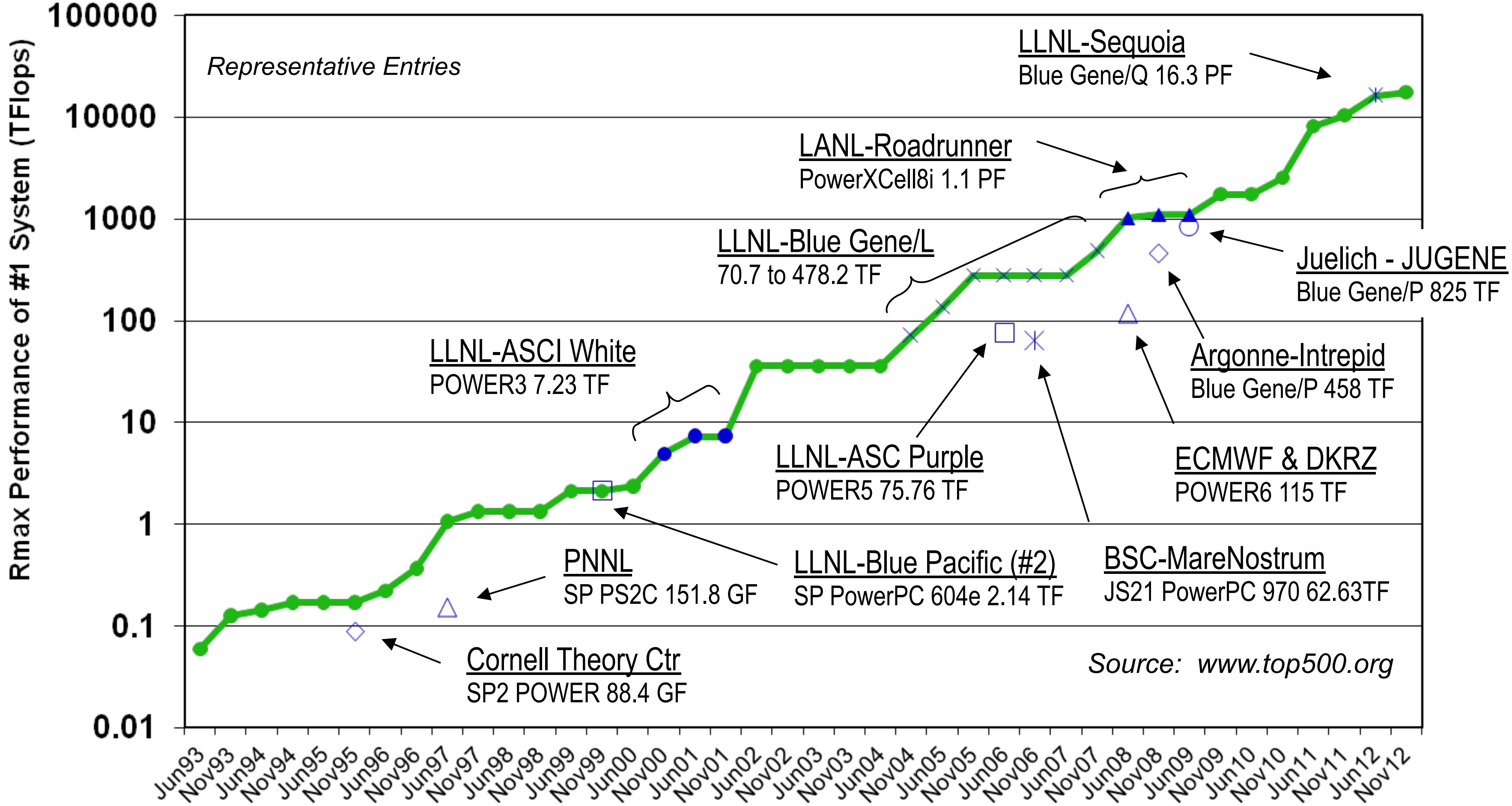
TOP500 Performance Trend

Over the long haul IBM has demonstrated continued leadership in various TOP500 metrics, even as the performance continues it's relentless growth.



IBM Technology Leadership

IBM has a long history of systems at or near the top of the TOP500 Worldwide Supercomputer List based on evolving technology.



Key: Blue markers are IBM machines.
Solid blue markers indicate #1 IBM machines.
Grouped markers of the same shape indicate the same architecture.

SP Design Principles

Principle 1: "Ride the technology curve"
High-performance scalable system must utilize standard microprocessors, packaging and OS

Principle 2: Time-to-market
Time-to-market with the latest technology is critical to achieving leadership performance and price performance

Principle 3: Communication is critical
Latency and bandwidth requirements demand custom interconnects and communication subsystems

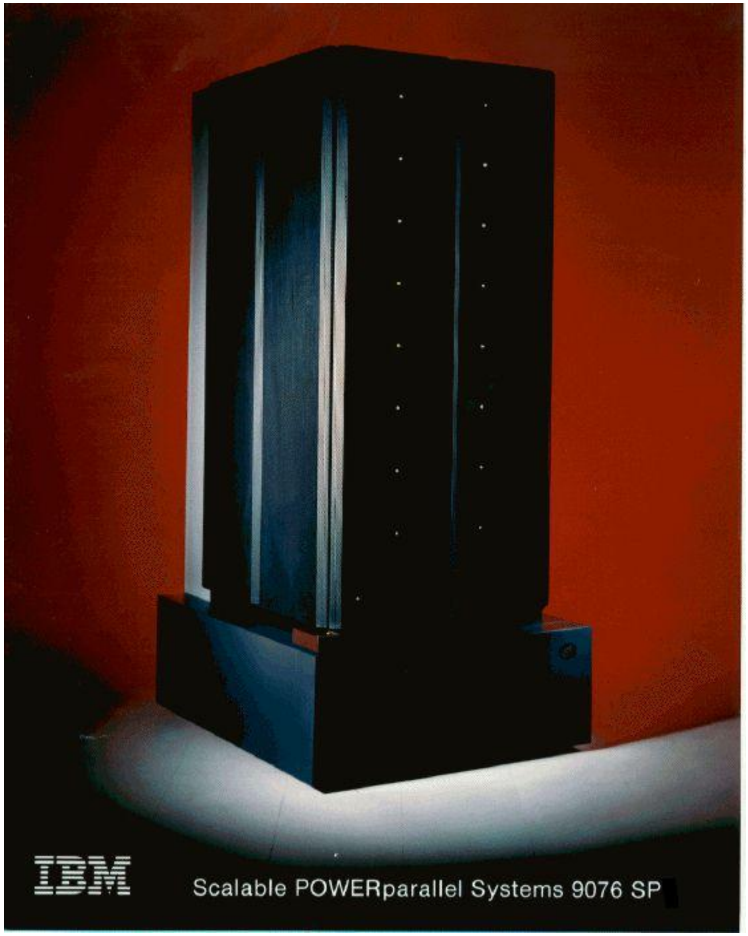
Principle 4: Standard UNIX
The system must support a programming and execution environment identical to a standard, open, distributed UNIX environment

Principle 5: High-performance services
The system should provide a judiciously chosen set of high-performance services

Principle 6: High Availability
Achieve high-availability by systematically removing single points of failure and provide very fast recovery from all failures

Principle 7: Single-System Image Flexibility
Selected support for a single-system image through globalization of key resources and commands; and single-point of control for system management and administration

SP2 Impact



- **Government: Science Based Stockpile Stewardship** (SBSS, 1994)
  - Enabled ASCI success
  - Created
    - Simulation Applications
    - Computing Platforms
    - Support Environment
- **Science**
  - Dramatic new level of simulation accuracy
    - Numerical error < physical effects
  - Examples
    - Compressible turbulence
    - Weather forecasting
- **Industry**
  - Development of general purpose platform
  - Drove parallel database adaption: DB2, SAP, Oracle
  - Aerospace, Automotive, Chemistry, Database, Electronics, Finance, Geophysics, Information Processing, Manufacturing, Mechanics, Pharmaceuticals, Telecom, Transportation, etc.

Blue Gene Design Principles – Optimized for power efficiency

Principle 1: Trade clock speed of lower power consumption
- Low Voltage & Low Frequency
- High Scale-out
- Problem size limited by FLOPS/W
- Better energy ratio = better performance

Principle 2: Use integration to lower power
- EDRAM vs SRAM
- Embedded processor
- Embedded controller
- Embedded network
- Embedding reduces system power

Principle 3: Focus on network performance
- Massive scale-out drives massive partitioning drives communication
- BW and latency grow in importance

Principle 4: Reduce OS jitter
- Importance grows with scale-out
- Allows reproducibility for performance enhancement, optimization & debugging

Principle 5: Co-Design
- Close lab/IBM collaboration

Blue Gene/Q

Sequoia

- # of cores: 1,572,864
- # of nodes: 98,304
- R_{max}: 16.32 PF (20.1 PF peak)
- Power: 7.89 MW (network excl.)
- Sustained perf: 81.08%
- GF/W: 2.07

Performance Gains on ANL Applications

DNS3D	**16.9x**	**Held-Suarez**	**11.9x**
FLASH	**5.9x**	**LS3DF**	**8.1x**
GFMC	**10.5x**	**MILC**	**6.1x**
GPAW	**8.5x**	**NAMD**	**8.5x**
GTC	**11.2x**	**NEK**	**7.4x**

2012 AWARDS

- Top500: #1 (June)
- Green500
  - June: #1 thru #20
  - Nov: 25 of the top 29 slots

Highest performance on:
- Heart beat simulation: ~ real time 11.83 PF
- Cosmology: dark matter/dark energy simulation: 12 PF

Blue Gene Dominates Graph500

Blue Gene
- Developed for Physics simulations NOT graph algorithms
- Performs remarkable well on the Graph500
- Won 1st Place five times in a row --- 3x faster than best non-BG/Q winner

Attests to
- **1) The quality of the Blue Gene network**
- **2) The criticality of efficient data motion in Physics and Analytics application**

Rank	Name	Vendor	Machine Type	Installation Site	# of Nodes	# of Cores	GTEPS
1	Sequoia	IBM	BlueGene/Q	Lawrence Livermore National Laboratory	65536	1048576	15363
2	Mira	IBM	BlueGene/Q	Argonne National Laboratory	32768	524288	10461
3	JUQUEEN	IBM	BlueGene/Q	Forschungszentrum Juelich (FZJ)	16384	262144	5848
4	K Computer	Fujitsu	custom	RIKEN Advanced Institute for Computational Science (AICS)	65536	524288	5524
5	Fermi	IBM	BlueGene/Q	CINECA	8192	131072	2567
6	Turing	IBM	BlueGene/Q	CNRS/IDRIS-GENCI	4096	65536	1427
6	Blue Joule	IBM	BlueGene/Q	Science & Technology Facilities Council - Daresbury Laboratory	4096	65536	1427
6	DIRAC	IBM	BlueGene/Q	University of Edinburgh	4096	65536	1427
6	Zumbrota	IBM	BlueGene/Q	EDF R&D	4096	65536	1427
6	Avoca	IBM	BlueGene/Q	Victorian Life Sci. Comp. Initiative	4096	65536	1427

Medal of Technology – September 18, 2009

- IBM, which earned the National Medal of Technology and Innovation on seven other occasions, is the only company recognized with the award this year.
- **Blue Gene's speed and expandability have enabled business and science to address a wide range of complex problems and make more informed decisions -- not just in the life sciences, but also in astronomy, climate, simulations, modeling and many other areas. Blue Gene systems have helped map the human genome, investigated medical therapies, safeguarded nuclear arsenals, simulated radioactive decay, replicated brain power, flown airplanes, pinpointed tumors, predicted climate trends, and identified fossil fuels -- all without the time and money that would have been required to physically complete these tasks.**

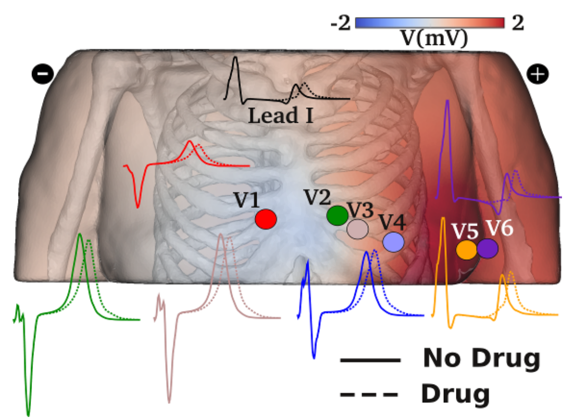
Toward Real-time Modeling of the Human Heart

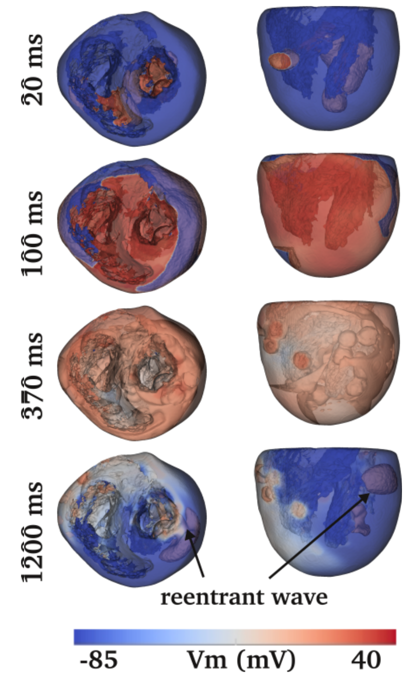
Collaboration of IBM Research and Lawrence Livermore National Laboratory

- Sudden cardiac arrest is the leading cause of death worldwide
- Improved cardiac models could improve diagnosis and therapy and could pave the way for interventions based on patient-specific simulations

High-resolution reconstruction of human heart ventricles with up to 370M mesh cells, roughly the size of heart muscle cells

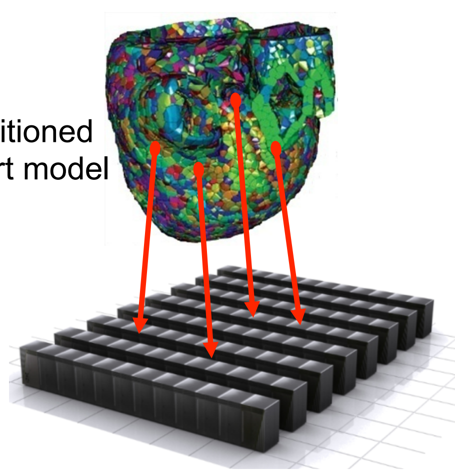
Simulated ECG showing the effect of a drug





Simulated activation of the whole heart with administration of a drug

Partitioned heart model



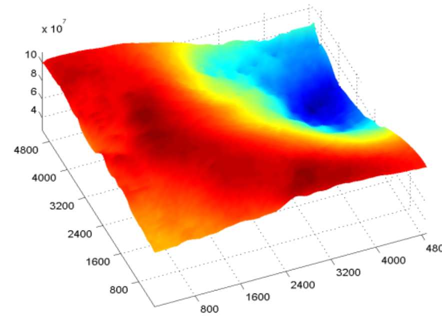
Sequoia: 96 racks of Blue Gene/Q

- Advanced algorithms and threading strategy use all 5 levels of parallelism available on Blue Gene/Q
- 58.8% of peak performance on full Sequoia machine
- 1200X faster than nearest competitor

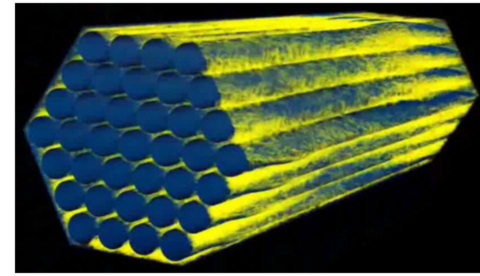
Blue Gene Impact

Renewable Energy

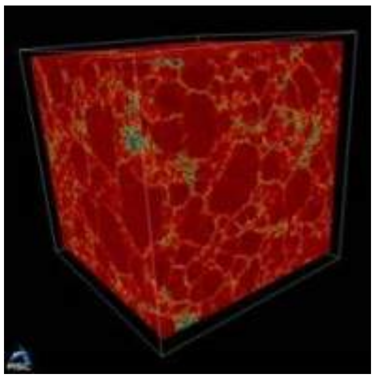
Geophysical Data Processing



Nuclear Energy

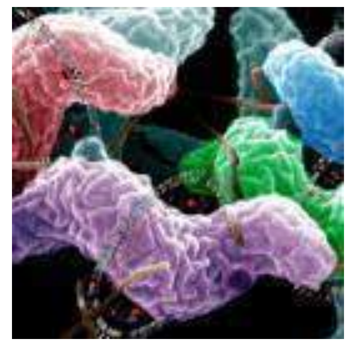


Materials Science Molecular Dynamics



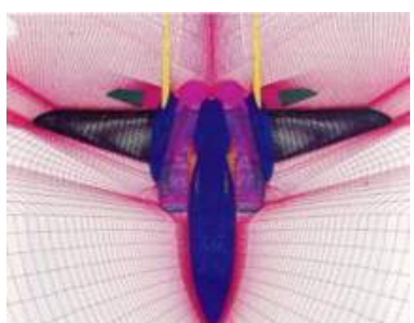
Geophysical Data Processing Upstream Petroleum

Life Sciences: Sequencing



Life Sciences: In-Silico Trials, Drug Discovery

Enviroment and Climate Modeling



Computational Fluid Dynamics

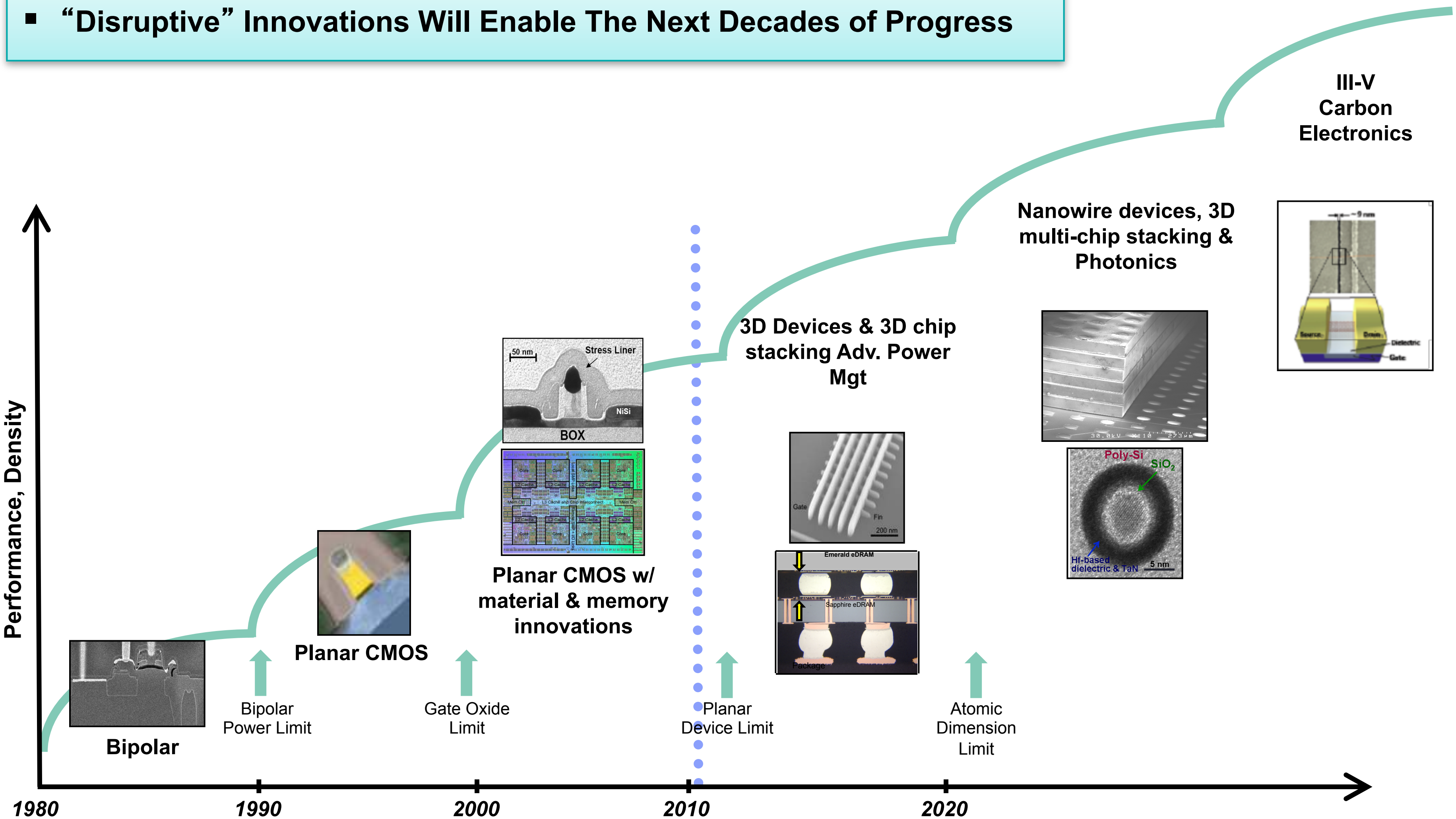
Biological Modeling Brain Science

Financial Modeling Streaming Data Analysis

Future Systems

Historical Silicon Technology Scaling

- Major Technology Innovations "Saturate" After About A Decade
- "Disruptive" Innovations Will Enable The Next Decades of Progress



Evolving Technical Computing: A Continuum

Business Analytics
- Unstructured data
- Giga- to petascale data
- Primarily data mining algs

Technical Computing
- Structured data
- Exascale data sets
- Primarily scientific calculations

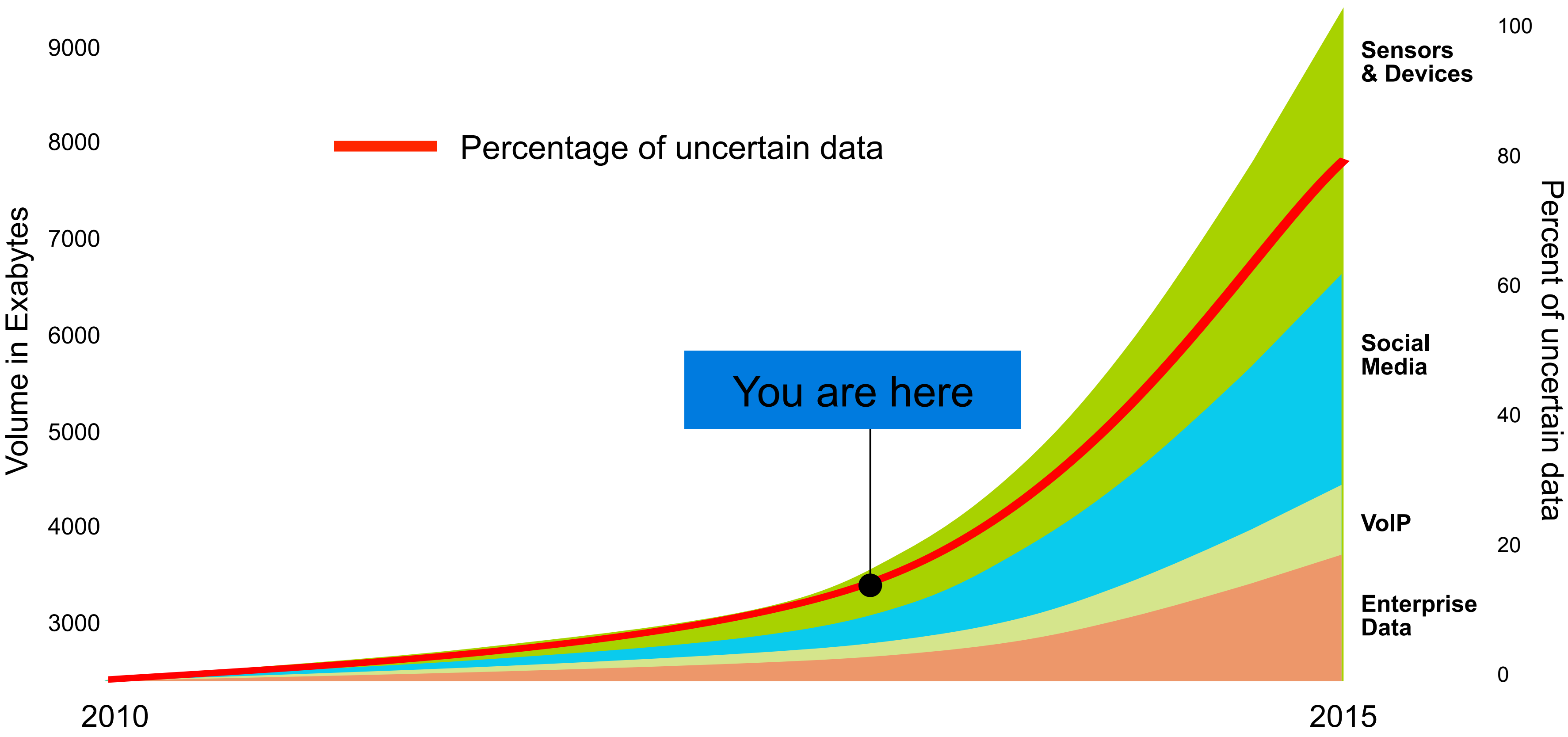
Evolving requirements

Driver: **Enhanced context Improves decision making**
- Incorporate Modeling and Simulation
- Incorporate sensor data

Driver: **Doing more with models**
- Real-time decision making
- Uncertainty Quantification
- Sensitivity Analysis
- Metadata extraction

Opportunities
- ***City planning*** enhanced to include real-time weather and resource modeling
- ***Uncertainty Quantification*** for nuclear reactor core simulations
- ***Sensitivity analysis*** for power grid simulations and optimization

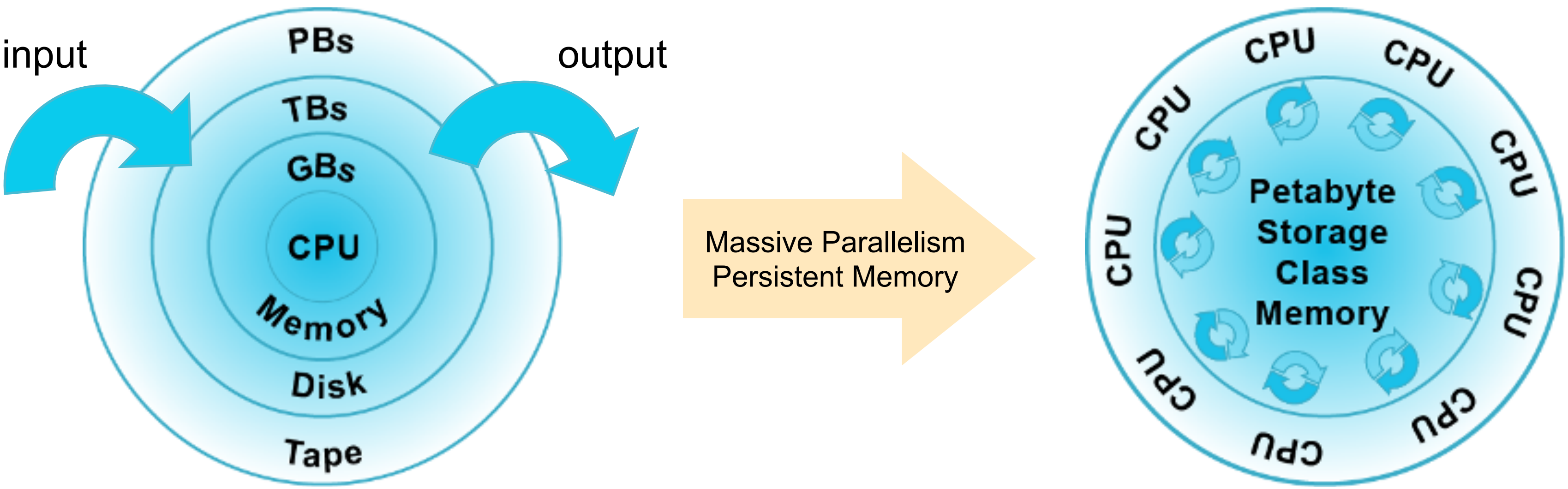
Big Data: This is just the beginning



Big Data Requires a Data-Centric Architecture for Performance

Old

New



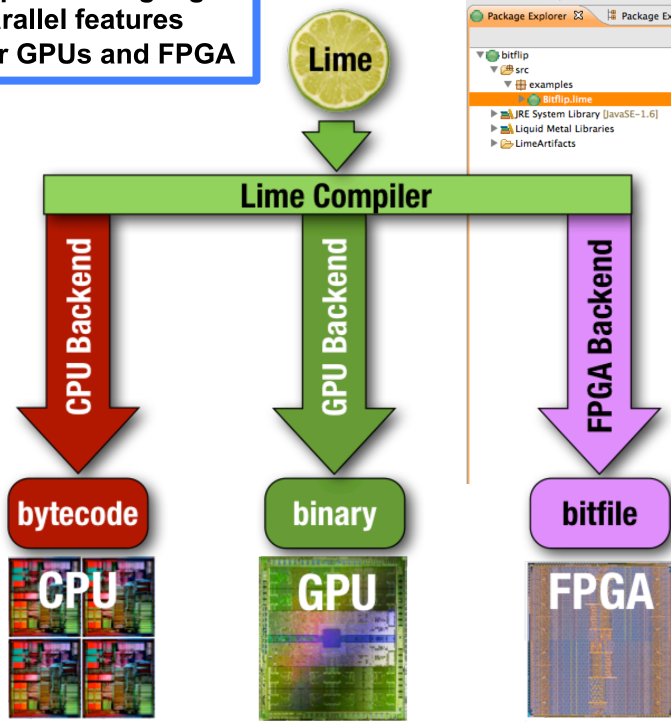
Data lives on disk and tape
Move data to CPU as needed
Deep Storage Hierarchy

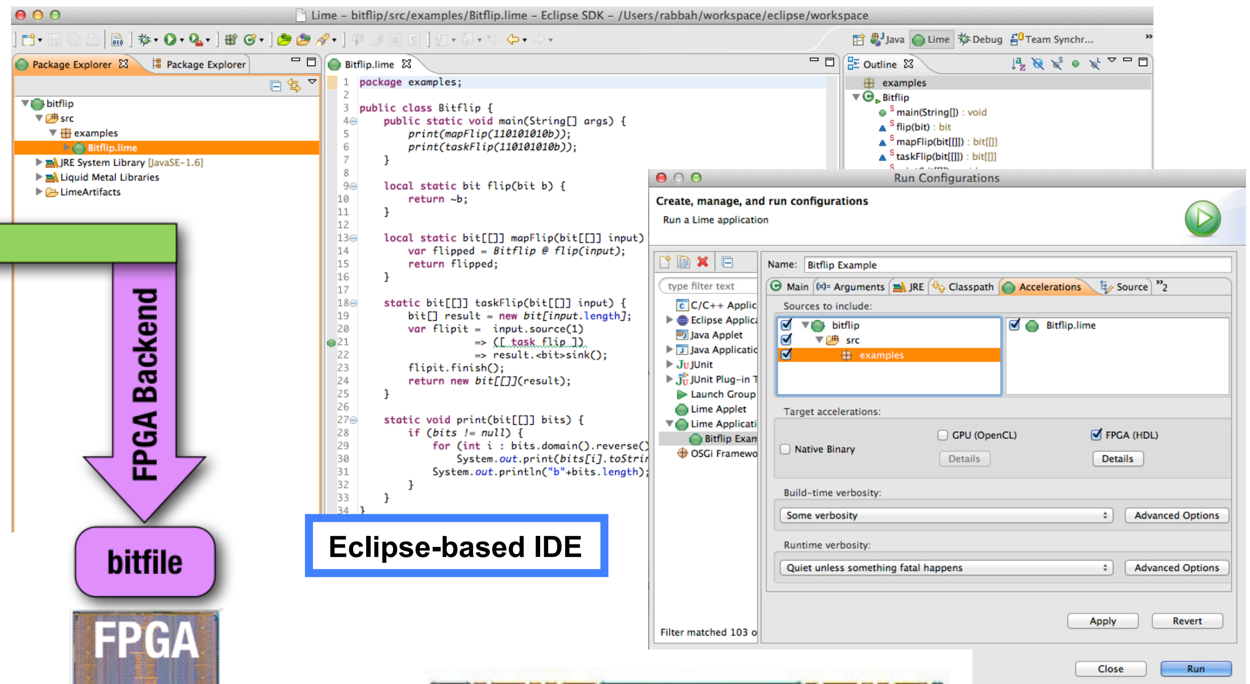
Data lives in persistent memory
Many CPU's surround and use
Shallow/Flat Storage Hierarchy

Huge impact on hardware, systems software, and application design

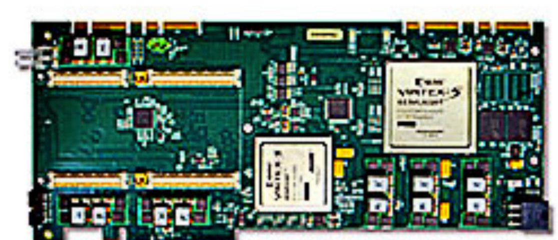
Liquid Metal - Comprehensive Programming Solution for Heterogeneous Architectures

Java-compatible language with parallel features suitable for GPUs and FPGA





Eclipse-based IDE

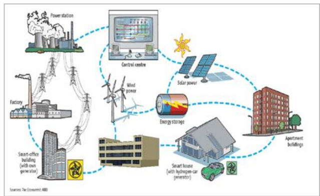


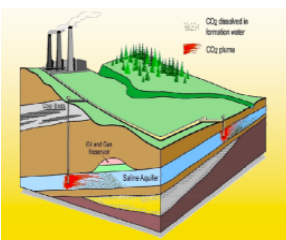
Integrated support for EDA tools and PCIe FPGA platforms

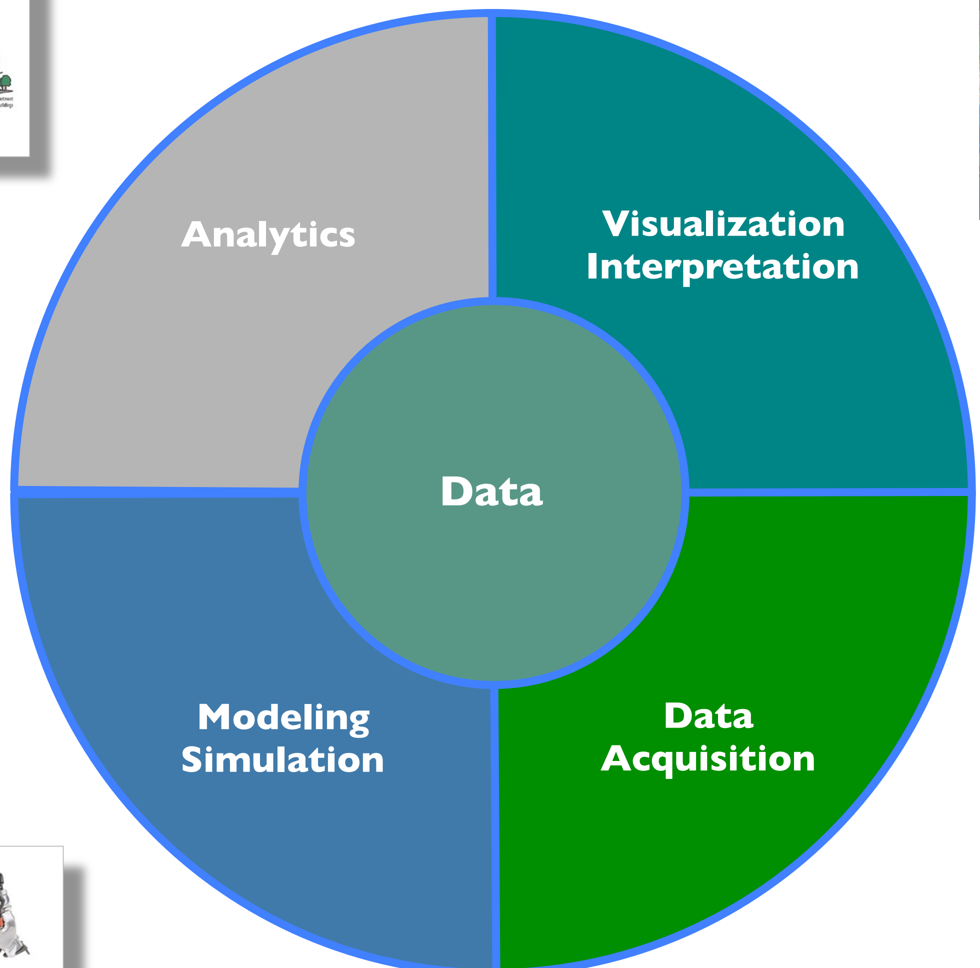
Accelerated development (12 CHStone Benchmarks)

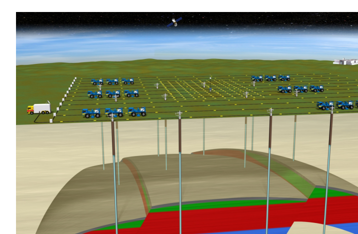
- < 1 day for most benchmark vs. many days by hand
- 3 trivial compiler bugs (less than 1 hour to fix)
- 2 non-trivial compiler bugs (1-2 days to fix)

Solutions must be complete, robust, reliable, easy to use









Evolving to Exascale Class Data, Exascale Class Compute