Opportunities and Challenges for Petascale Computing

Rajan Gupta
Theoretical Division
Los Alamos National Laboratory

rajan.gupta@lanl.gov  http://t8web.lanl.gov/people/rajan/  LACSI-SF06
My thoughts on solving grand challenges problems
OUTLINE

• High Performance Computing (HPC)
• Paths to discovery
• Opportunities
• Challenges
WHAT IS HPC? IS IT IT

- Squeezing the maximum performance out of a CPU?
- Using Parallel computers?
- Using the latest hardware?
- Using the latest software development tools? Debuggers?
- Using novel algorithms? Visualization?
- Incorporating the best scientific methodology?

IT IS ALL OF THE ABOVE!
The march of computer power

The future of large scale computing is in scalable parallel distributed computers
People driven to solve a very important problem
1940s

rajan.gupta@lanl.gov       http://t8web.lanl.gov/people/rajan/       LACSI-SF06

1947 Fermi Trolley
1952 Maniac

1955 Maniac II

1950s

rajan.gupta@lanl.gov  http://t8web.lanl.gov/people/rajan/  LACSI-SF06
1976 Cray I

1989 CM-2
1990s

rajan.gupta@lanl.gov  http://t8web.lanl.gov/people/rajan/  LACSI-SF06
The march of computer power
Ingredients for success

- Problems to solve
- Talented People
- Tools
- Resources

Bringing people, tools, resources, ideas, together “Virtually”
Opportunity

Modeling, Simulation, Visualization are becoming ever more powerful tools.

They will play an increasingly larger role in our efforts to understand complex systems and phenomena.

To make quantitative predictions & errors

To understand and interpret experiments

To manipulate, integrate, analyze, and synthesize large data sets
At petascale, error analysis and bounds on uncertainty reduced to linear perturbative regime

- Lattice QCD (QCDOC → BGL)
- Astrophysics and Cosmology
- Biology and Bioinformatics
- Materials
- Climate (ocean modeling)
- ........
CMB and structure formation

1965  Penzias and Wilson

1992  COBE

2003  WMAP

rajan.gupta@lanl.gov  http://t8web.lanl.gov/people/rajan/  LACSI-SF06
Analyzing large data sets
Comparing models with observations

M. Warren: Simulations of the formation of large scale structure
Bioinformatics ↔ Ecosystems

- Ecosystem
- Species
- Organism
- Physiology
- Metabolism
- Network
- Function
- Structure
- Sequence

Comparative Genomics
Pharmacogenomics
Physiome
Metabolomics
Proteomics
Functional Genomics
Structural Genomics
Genome

rajan.gupta@lanl.gov
http://t8web.lanl.gov/people/rajan/

http://t8web.lanl.gov/
Paths to Prediction/Discovery

Experiments

Modeling
Simulation
Visualization

Theory

In ever increasing number of areas
Challenges
Sweet Spots

• Different problems are most efficiently implemented on different configurations of
  – Nodes (# of Processors, clock speed, registers, …)
  – Network topology, speed, latency
  – Memory size and latency
  – I/O and Disk access latency
Space, Power, Cooling

• Need appropriately prepared space
  – Floor space
  – Breakers, alarms, emergency procedures

• Need adequate power
  – 1000 node cluster needs 200,000 watts of power

• Need adequate cooling and air circulation to remove the heat

These costs and reliability (MTF) make scaling up clusters to petascale prohibitive
Bottom line

• “It is software stupid”
  – MPI is portable but

• Optimizing performance on MPPs versus clusters versus clusters with accelerators

• Efficiency limited by memory access
  – Both on node and off node

So much has changed and yet so little has changed