High Productivity Computing Systems

Robert Graybill
DARPA/IPTO

March 2003
High Productivity Computing Systems

Goal:
- Provide a new generation of economically viable high productivity computing systems for the national security and industrial user community (2007 – 2010)

Impact:
- **Performance** (time-to-solution): speedup critical national security applications by a factor of 10X to 40X
- **Programmability** (time-for-idea-to-first-solution): reduce cost and time of developing application solutions
- **Portability** (transparency): insulate research and operational application software from system
- **Robustness** (reliability): apply all known techniques to protect against outside attacks, hardware faults, & programming errors

Applications:
- Intelligence/surveillance, reconnaissance, cryptanalysis, weapons analysis, airborne contaminant modeling and biotechnology

Fill the Critical Technology and Capability Gap
Today (late 80’s HPC technology).....to.....Future (Quantum/Bio Computing)
Vision: Focus on the Lost Dimension of HPC – “User & System Efficiency and Productivity”

Parallel Vector Systems

1980’s Technology

Vector

Commodity HPCs

Tightly Coupled Parallel Systems

Moore’s Law
Double Raw Performance every 18 Months

New Goal: Double Value Every 18 Months

Fill the high-end computing technology and capability gap for critical national security missions
HPCS Technical Considerations

Architecture Types

<table>
<thead>
<tr>
<th>Custom Vector</th>
<th>Microprocessor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parallel Vector</td>
<td>Symmetric Multiprocessors Distributed Shared Memory</td>
</tr>
<tr>
<td>Scalable Vector</td>
<td>Massively Parallel Processors Commodity Clusters, Grids</td>
</tr>
<tr>
<td>Vector Supercomputer</td>
<td>Commodity HPC</td>
</tr>
</tbody>
</table>

HPCS Focus
Tailorable Balanced Solutions

Single Point Design Solutions are no longer Acceptable
HPCS Program Phases I - III

HPCS Capability or Products
Application Analysis Performance Assessment
Industry

HPCS

Requirements and Metrics
Concept Reviews
System Design Review
Technology Assessments
PDR
DDR
Research Prototypes & Pilot Systems

Phase I Industry Concept Study
Phase II R&D
Phase III Full Scale Development

Product Metrics, Benchmarks
Academia Research Platforms
Early Pilot Platforms

Early Software Tools

Industry Evolutionary Development Cycle

Fiscal Year
02 03 04 05 06 07 08 09 10

Reviews
Industry Procurements
Critical Program Milestones
HPCS Phase I Industry Teams

Industry:

Cray, Inc.  (Burton Smith)
Hewlett-Packard Company  (Kathy Wheeler)
International Business Machines Corporation  (Mootaz Elnozahy)
Silicon Graphics, Inc.  (Steve Miller)
Sun Microsystems, Inc.  (Jeff Rulifson)

Application Analysis/Performance Assessment Team:

MITRE  MIT Lincoln Laboratory
Phase II Milestones

- Award
  - Kick Off
    - System Design Review
  - Project Review
    - Technology Assessment Review
- Project Review
  - Preliminary Design Review
    - Readiness Review
      - Report Submittal
Application Analysis/Performance Assessment

Activity Flow

Inputs
- DDR&E & IHEC Mission Analysis

Mission Partners:
- DOD
- DOE
- NNSA
- NSA
- NRO

Application Analysis
- HPCS Applications
  1. Cryptanalysis
  2. Signal and Image Processing
  3. Operational Weather
  4. Nuclear Stockpile Stewardship
  5. Etc.
- Mission-Specific Roadmap
- Mission Work Flows

Benchmarks & Metrics
- Common Critical Kernels
- Compact Applications
- Productivity Ratio of Utility/Cost Metrics
  - Development time (cost)
  - Execution time (cost)
- Implicit Factors

Impacts
- Participants
- Mission Partners
- DARPA

Participants:
- Cray
- HP
- IBM
- SGI
- Sun

Mission Partners:
- DOD
- DOE
- NNSA
- NSA
- NRO
Application Focus Selection

**DDR&E Study**
- Operational weather and ocean forecasting
- Planning activities for dispersion of airborne/waterborne contaminants
- Cryptanalysis
- Intelligence, surveillance, reconnaissance
- Improved armor design
- Engineering design of large aircraft, ship and structures
- National missile defense
- Test and evaluation
- Weapon (warheads and penetrators)
- Survivability/stealth design

**IHEC Study**
- Comprehensive Aerospace Vehicle Design
- Signals Intelligence (Crypt)
- Signals Intelligence (Graph)
- Operational Weather/Ocean Forecasting
- Stealthy Ship Design
- Nuclear Weapons Stockpile Stewardship
- Signal and Image Processing
- Army Future Combat Systems
- Electromagnetic Weapons Development
- Geospatial Intelligence
- Threat Weapon Systems Characterization

**Bioscience**
Computational Biology: from Sequence to Systems

- Sequence Genome
- Assemble Genome
- Find the Genes
- Annotate the Genes
- Map Genes to Proteins
- Protein-Protein Interactions
- Pathways: Normal & Aberrant
- Protein Functions in Pathways
- Protein Structure
- Identify Drug Targets
- Cellular Response
- Tissue, Organ & Whole Body Response

Slide provided by IDC
<table>
<thead>
<tr>
<th>Kernels</th>
<th>Application</th>
<th>Source</th>
<th>Today</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioCatalysis</td>
<td>Ab Initio Quantum Chemistry</td>
<td>GAMESS</td>
<td>DoD HPCMP TI-03</td>
</tr>
<tr>
<td></td>
<td>Quantum Chemistry</td>
<td>GAUSSIAN</td>
<td><a href="http://www.gaussian.com/">www.gaussian.com/</a></td>
</tr>
<tr>
<td></td>
<td>Quantum Mechanics</td>
<td>NWChem</td>
<td>PNNL</td>
</tr>
<tr>
<td>Quantum and MM</td>
<td>Macromolecular Dynamics</td>
<td>CHARM</td>
<td><a href="http://yuri.harvard.edu/">http://yuri.harvard.edu/</a></td>
</tr>
<tr>
<td></td>
<td>Energy Minimization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MonteCarlo Simulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Molecular Mechanical Field Force</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m-Array 8000 Genes</td>
<td>Clustering</td>
<td>CLUSTALW</td>
<td><a href="http://bimas.dct.nih.gov/sw.html">http://bimas.dct.nih.gov/sw.html</a></td>
</tr>
<tr>
<td></td>
<td>Whole Genome Analysis</td>
<td>Sequence Comparison</td>
<td><a href="http://www.med.nyu.edu/rcr/rcrcourse/sim-sw.html">http://www.med.nyu.edu/rcr/rcrcourse/sim-sw.html</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Needleman-Wunsch</td>
<td>100 TeraOps/s sustained</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FASTA</td>
<td>100 TeraOps/s sustained</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HMMR</td>
<td>100 TeraOps/s sustained</td>
</tr>
<tr>
<td>Whole Genome Analysis</td>
<td>Sequence Comparison</td>
<td>GENSCAN</td>
<td><a href="http://genes.mit.edu/GENSCANinfo.html">http://genes.mit.edu/GENSCANinfo.html</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100 TeraOps/s sustained</td>
</tr>
<tr>
<td></td>
<td>Biological Pathway Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Complex Systems Simulation and Analysis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Partial Differential Equation Solver</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ordinary Differential Equation Solver</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Digital Imaging</td>
<td>Marching Cubes</td>
<td>Paper &amp; Pencil for Kernels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Triangle Reduction</td>
<td>Paper &amp; Pencil for Kernels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Triangle Smoothing</td>
<td>Paper &amp; Pencil for Kernels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Noise Reduction</td>
<td>Paper &amp; Pencil for Kernels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Artifact Removal</td>
<td>Paper &amp; Pencil for Kernels</td>
</tr>
</tbody>
</table>
HPCS Mission Work Flows

Overall Cycle

Days to hours

Orient
Act
Decide
Production

Observe

Visualize

Design
Enterprise
Simulation

Experiment
Researcher

Theory

Development
Execution

HPCS Productivity Factors: Performance, Programmability, Portability, and Robustness are very closely coupled with each work flow

Development Cycle

Hours to minutes

Code
Prototyping
Design
Test

Months to days

Optimize

Port Legacy Software
Code

Days to hours

Span

Years to months

Evaluation

Maintenance
Operation

Design
Test
Port, Scale, Optimize

Initial Development

Initial Product Development

HPCS Productivity Factors: Performance, Programmability, Portability, and Robustness are very closely coupled with each work flow
Workflow Priorities & Goals

Implicit Productivity Factors

<table>
<thead>
<tr>
<th>Workflow</th>
<th>Perf.</th>
<th>Prog.</th>
<th>Port.</th>
<th>Robust.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Researcher</td>
<td>High</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Production</td>
<td>High</td>
<td></td>
<td></td>
<td>High</td>
</tr>
</tbody>
</table>

- Workflows define scope of customer priorities
- Activity and Purpose benchmarks will be used to measure Productivity
- HPCS Goal is to add value to each workflow
  - Increase productivity while increasing problem size

Mission Needs

System Requirements

HPCS Goal
HPCS Productivity Framework

Productivity

(Ratio of Utility/Cost)

Work Flow

Activity & Purpose

Benchmarks

Development Time (cost)

Execution Time (cost)

Productivity Metrics

Actual System or Model

System Parameters (Examples)

- BW bytes/flop
- Memory latency
- Memory size
- Processor flop/cycle
- Bisection BW
- Total Connections
- Size (cuft)
- Power/rack
- Facility operation
- Code size
- Restart time (reliability)
- Code Optimization time

Implicit HPCS Productivity Factors:
Performance, Programmability, Portability, and Robustness

RBG 10/9/2002
Benchmark Relationships

Fixed Size

Activity Based
(Well Suited for Execution Measurement)

LINPACK
NAS Parallel
SPEC
HPCS Suite

Purpose Based
(Ideal for Development Measurement)

NSA Suites
Some RFP Suites

Scalable

LINPEAK
NAS Parallel Streams, GUPS
HPCS Suite (Planned)

Phase I – Scope Benchmarks
Phase II – Activity and Purpose Benchmarks
HPC Community Reactions

- DoD User Community
  - Active participation in reviews
  - Providing challenge problems
  - Linking with internal efforts
  - Providing funding synergism
- Industry
  - Finally an opportunity to develop a non evolutionary vision
  - Active program support (technical, personnel, vision)
  - Direct impact to future product roadmaps
- University
  - Active support for Phase 1 (2X growth from proposals)
- Extended Community
  - HPCS strategy embedded in Congressional IHEC Report

Productivity a new HPC Sub-discipline