High Productivity Computing Systems

Robert Graybill
DARPA/IPTO

March 2003

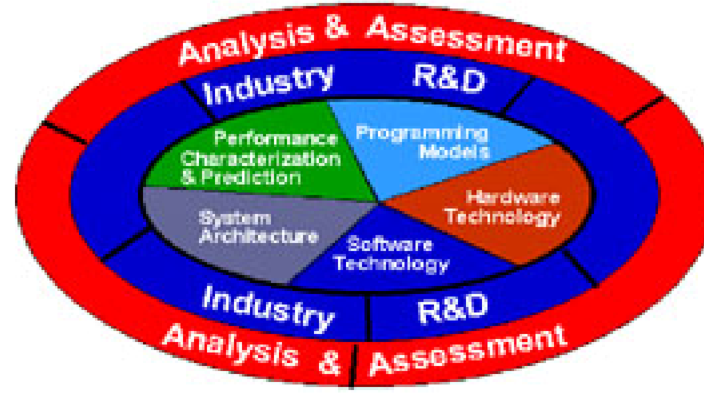
High Productivity Computing Systems

Goal:

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

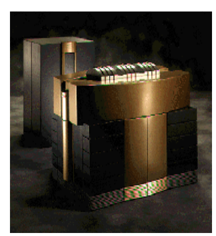
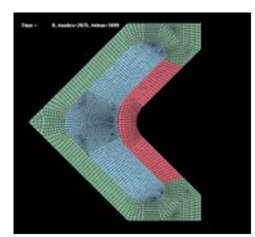
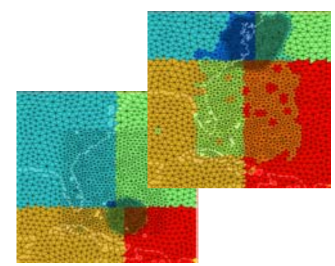
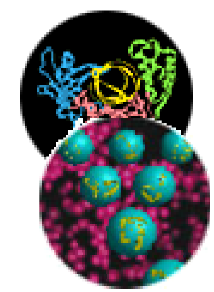
Impact:

- **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



HPCS Program Focus Areas



Applications:

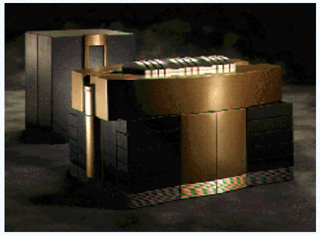
- Intelligence/surveillance, reconnaissance, cryptanalysis, weapons analysis, airborne contaminant modeling and biotechnology

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

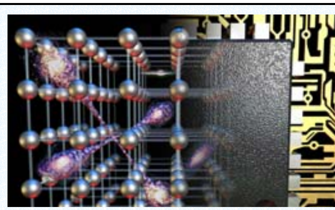


1980's Technology

Vector

Tightly Coupled Parallel Systems

Commodity HPCs



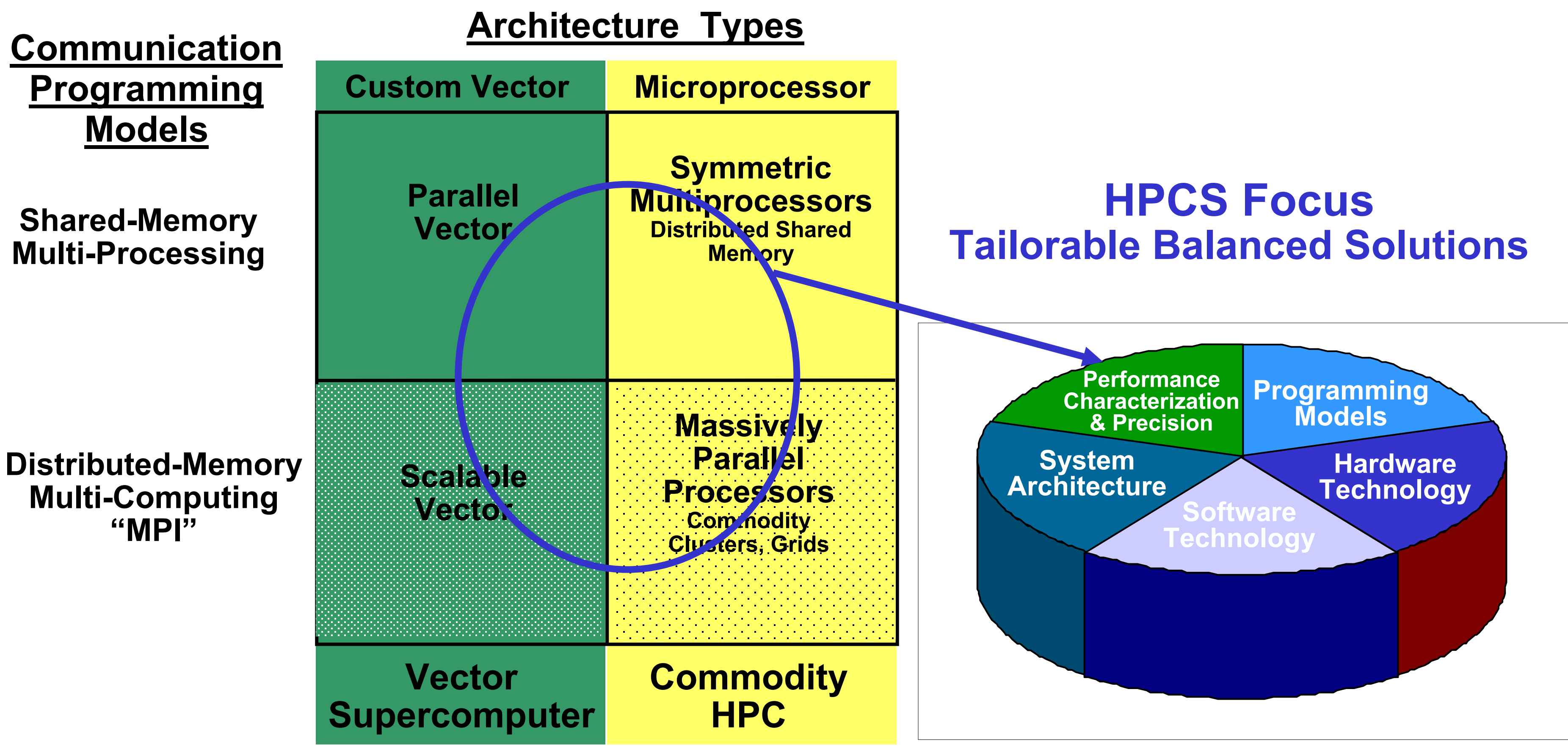
2010 High-End Computing Solutions

Moore's Law Double Raw Performance every 18 Months

New Goal: Double Value Every 18 Months

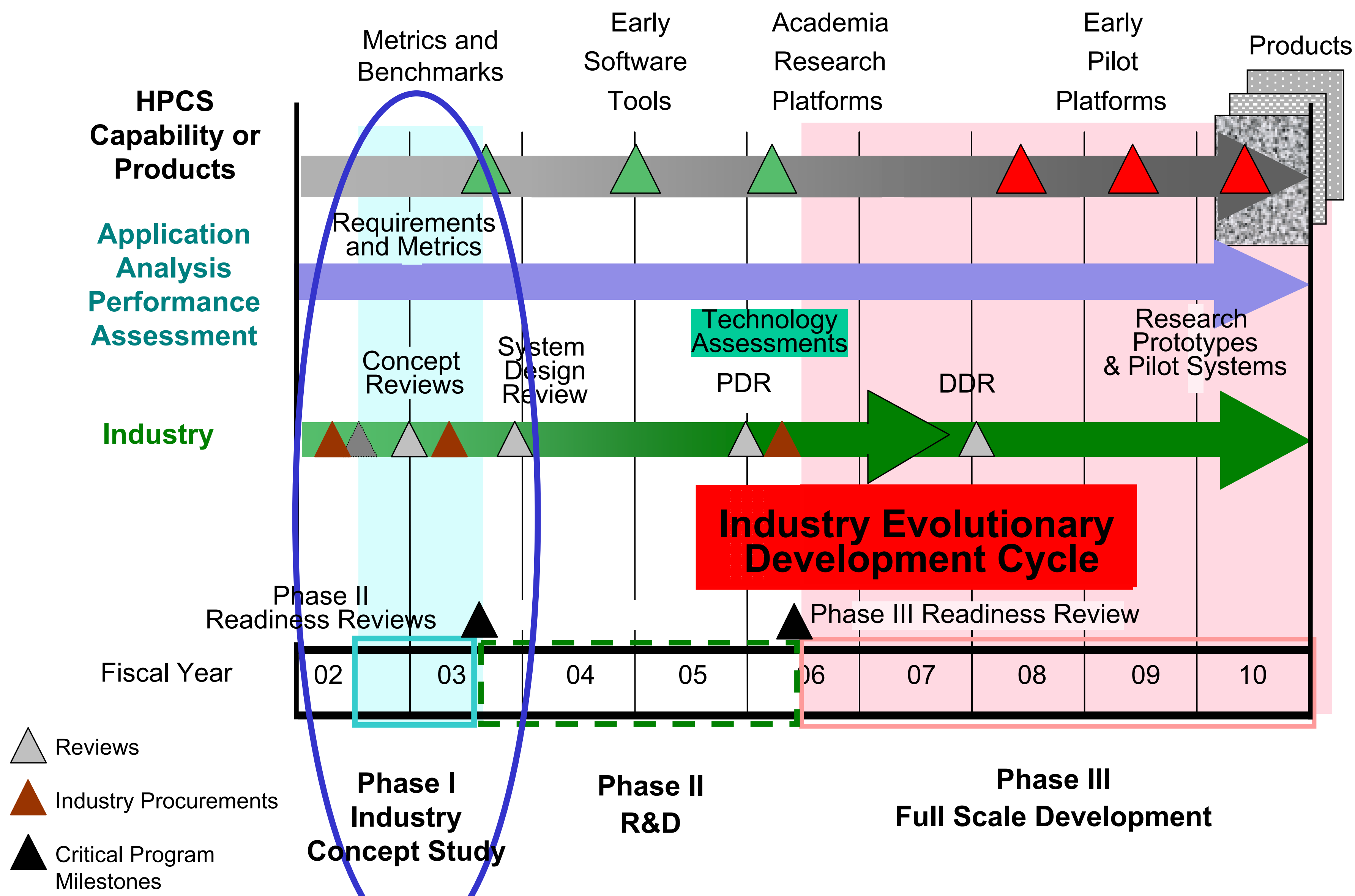
Fill the high-end computing technology and capability gap for critical national security missions

HPCS Technical Considerations



Single Point Design Solutions are no longer Acceptable

HPCS Program Phases I - III



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)

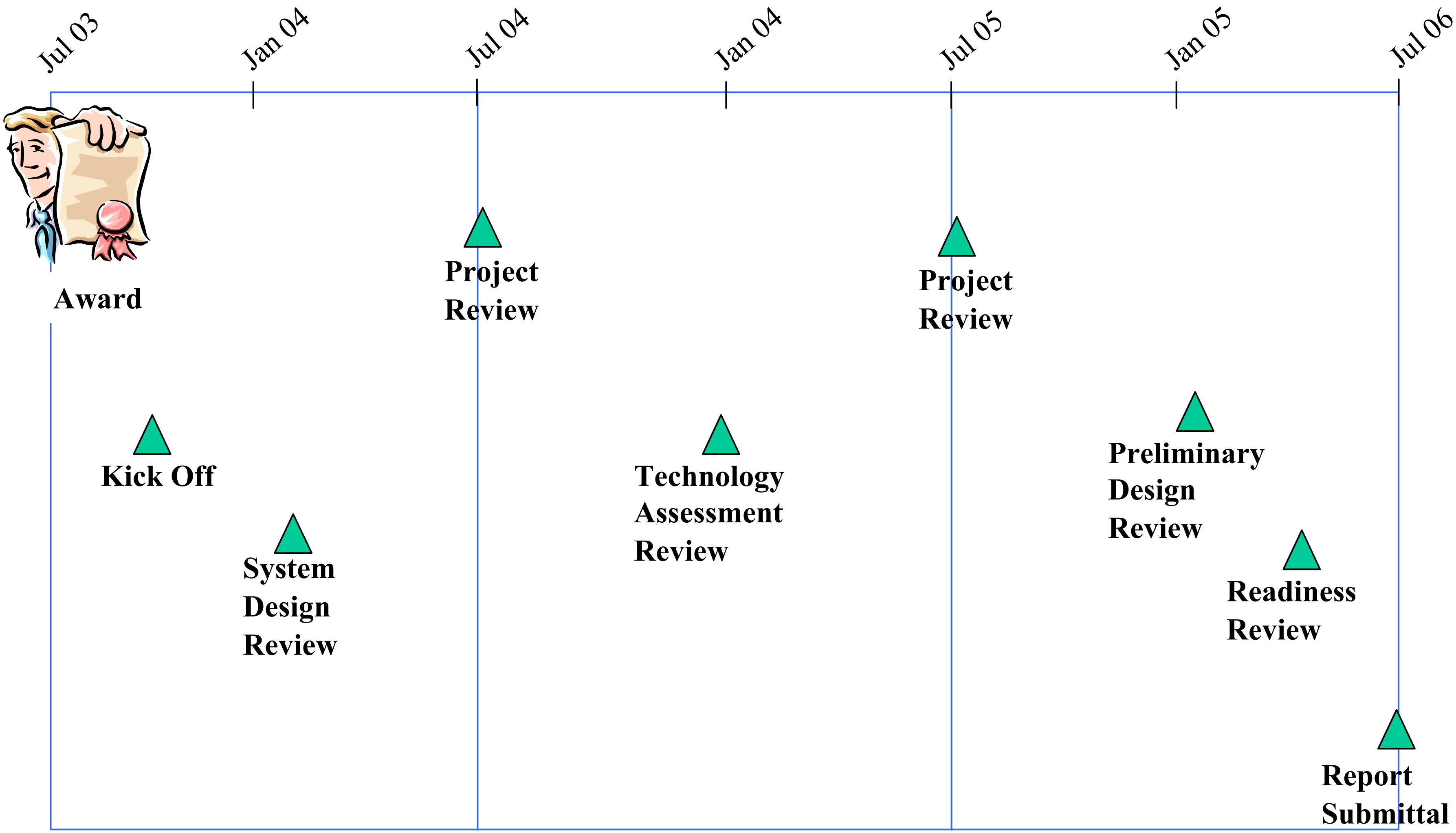
Sun Microsystems, Inc. (Jeff Rulifson)

Application Analysis/Performance Assessment Team:

MITRE

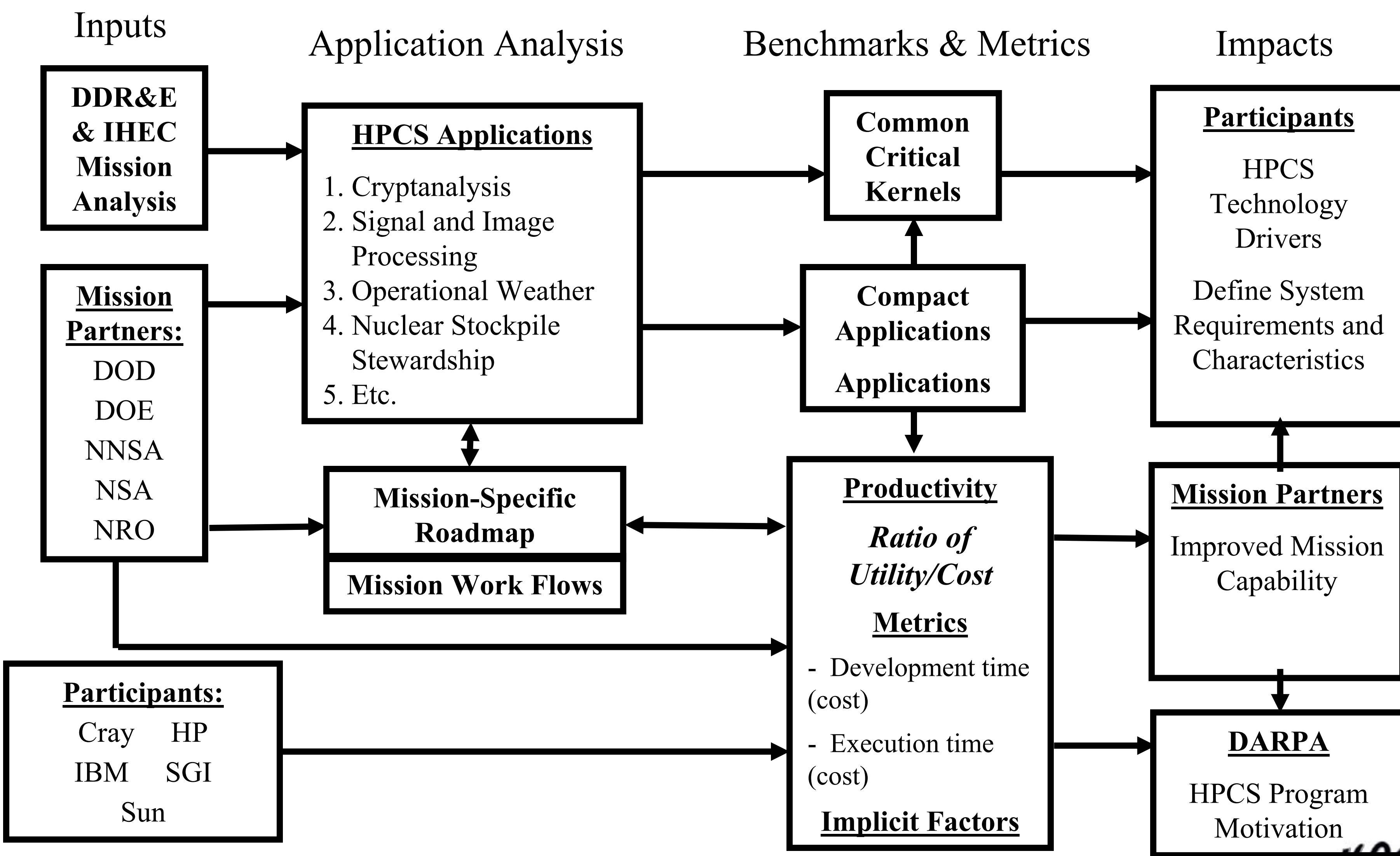
MIT Lincoln Laboratory

Phase II Milestones



Application Analysis/ Performance Assessment

Activity Flow



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

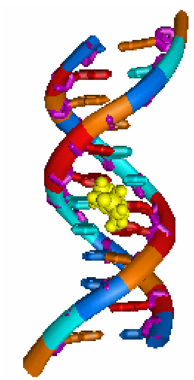
- **Bioscience**

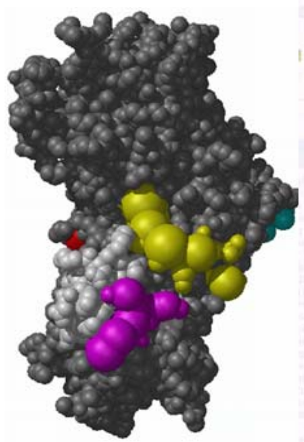
IHEC Study

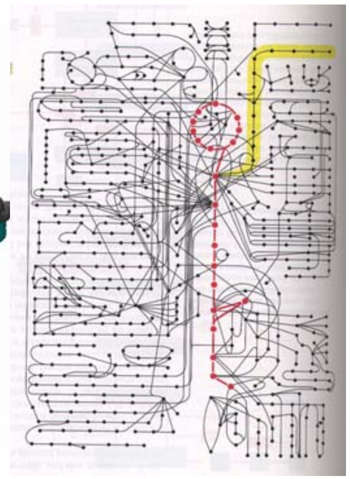
- Comprehensive Aerospace Vehicle Design
- Signals Intelligence (Crypt)
- Signals Intelligence (Graph)
- Operational Weather/Ocean Forecasting
- Stealthy Ship Design
- Nuclear Weapons Stockpile Stewardship
- Signal and Image Processing
- Army Future Combat Systems
- Electromagnetic Weapons Development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

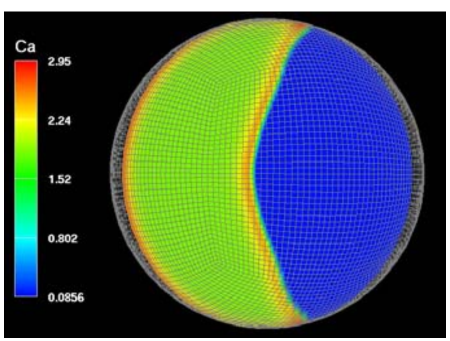
Biomedical Computing Requirements

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

TeraOps

Trivially Parallel

1

10

100

1000

Peta-Scale Computing

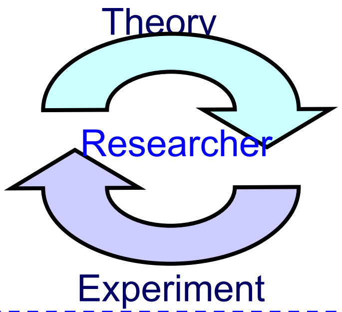
Slide provided by IDC

Biomedical Application and Kernels

	Kernels	Application	Source	Today
BioCatalysis	Ab Initio Quantum Chemistry	GAMESS	DoD HPCMP TI-03	TeraOp/s sustained
	Quantum Chemistry	GAUSSIAN	www.gaussian.com/	TeraOp/s sustained
	Quantum Mechanics	NWChem	PNNL	TeraOp/s sustained
Quantum and MM	Macromolecular Dynamics	CHARM	http://yuri.harvard.edu/	10 TeraOp/s sustained
	Energy Minimization			
	MonteCarlo Simulation			
	Molecular Mechanical Field Force	AMBER	http://www.amber.ucsf.edu/	10 TeraOp/s sustained
m-Array 8000 Genes	Clustering	CLUSTALW	http://bimas.dcrt.nih.gov/sw.html	200 GigaOps/s sustained
Multiple Alignment Phylogenetics	Pattern Matching	NONMEM	http://www.globomaxservice.com/ products/nonmem.html	100 GigaOps/s sustained
	Pattern Matching	PHYLIP	http://evolution.genetics.washington.edu/ phylip.html	100 GigaOps/s sustained
	Pattern Matching	FASTme	http://www.ncbi.nlm.nih.gov/ CBBresearch/Desper/FastME.html	100 GigaOps/s sustained
Whole Genome Analysis	Sequence Comparison	Needleman-Wunsch	http://www.med.nyu.edu/ rcr/rcr/course/sim-sw.html	100 TeraOps/s sustained
	Sequence Comparison	FASTA	http://www.ebi.ac.uk/fasta33/	100 TeraOps/s sustained
	Sequence Comparison	HMMR	http://hmmer.wustl.edu/	100 TeraOps/s sustained
	Sequence Comparison	GENSCAN	http://genes.mit.edu/GENSCANinfo.html	100 TeraOps/s sustained
Systems Biology	Functional Genomics		http://genomics.lbl.gov/~aparkin/ Group/Codebase.html	
	Biological Pathway Analysis			
	Complex Systems Simulation and Analysis		http://ecell.sourceforge.net/	
	Partial Differential Equation Solver		http://www.nrcam.uchc.edu/	
	Ordinary Differential Equation Solver			
Digital Imaging	Marching Cubes		Paper & Pencil for Kernels	
	Triangle Reduction		Paper & Pencil for Kernels	
	Triangle Smoothing		Paper & Pencil for Kernels	
	Noise Reduction		Paper & Pencil for Kernels	
	Artifact Removal		Paper & Pencil for Kernels	

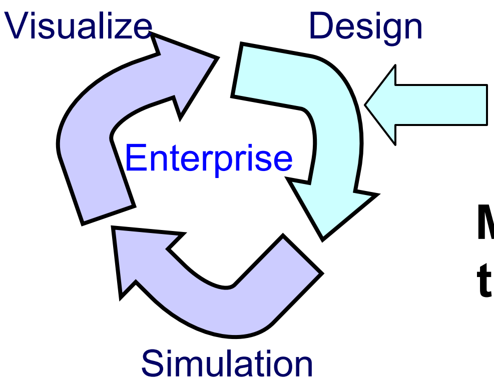
HPCS Mission Work Flows

Overall Cycle



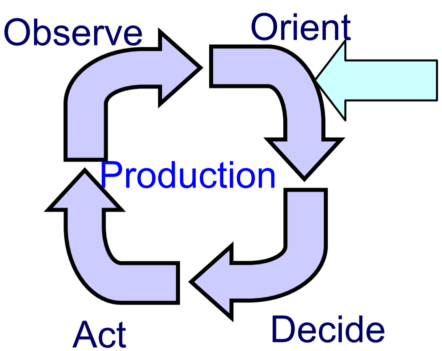
Days to hours

Development
Execution



Port Legacy Software

Months to days

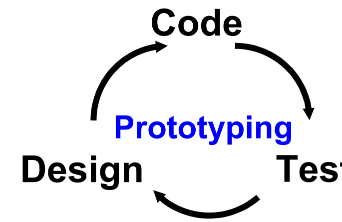


Initial Product Development

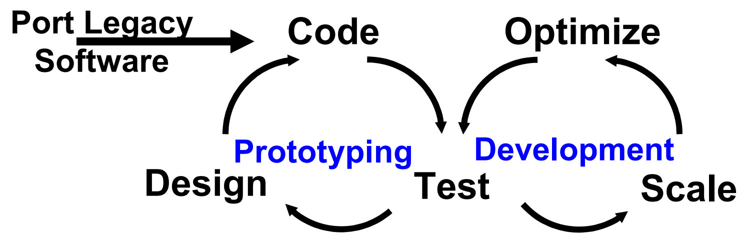
Hours to Minutes
(Response Time)

Development Cycle

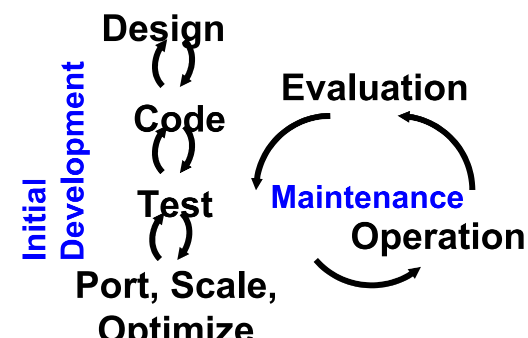
Hours to minutes



Months to days



Years to months

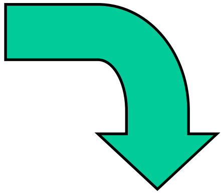


HPCS Productivity Factors: Performance, Programmability, Portability, and Robustness are very closely coupled with each work flow

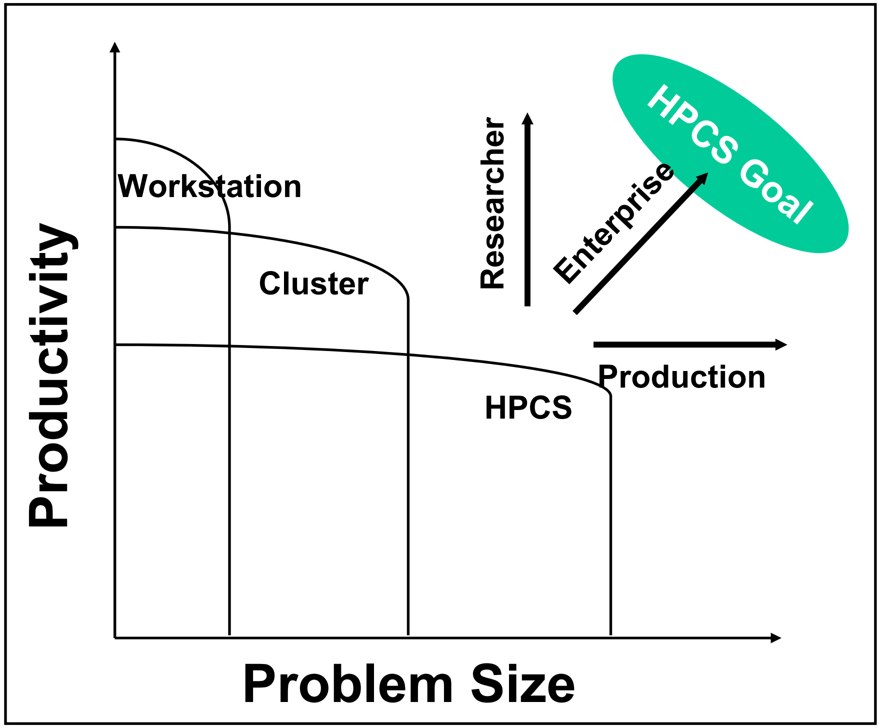
Workflow Priorities & Goals

	Implicit Productivity Factors			
Workflow	**Perf.**	**Prog.**	**Port.**	**Robust.**
Researcher		High		
Enterprise	High	High	High	High
Production	High			High

Mission Needs

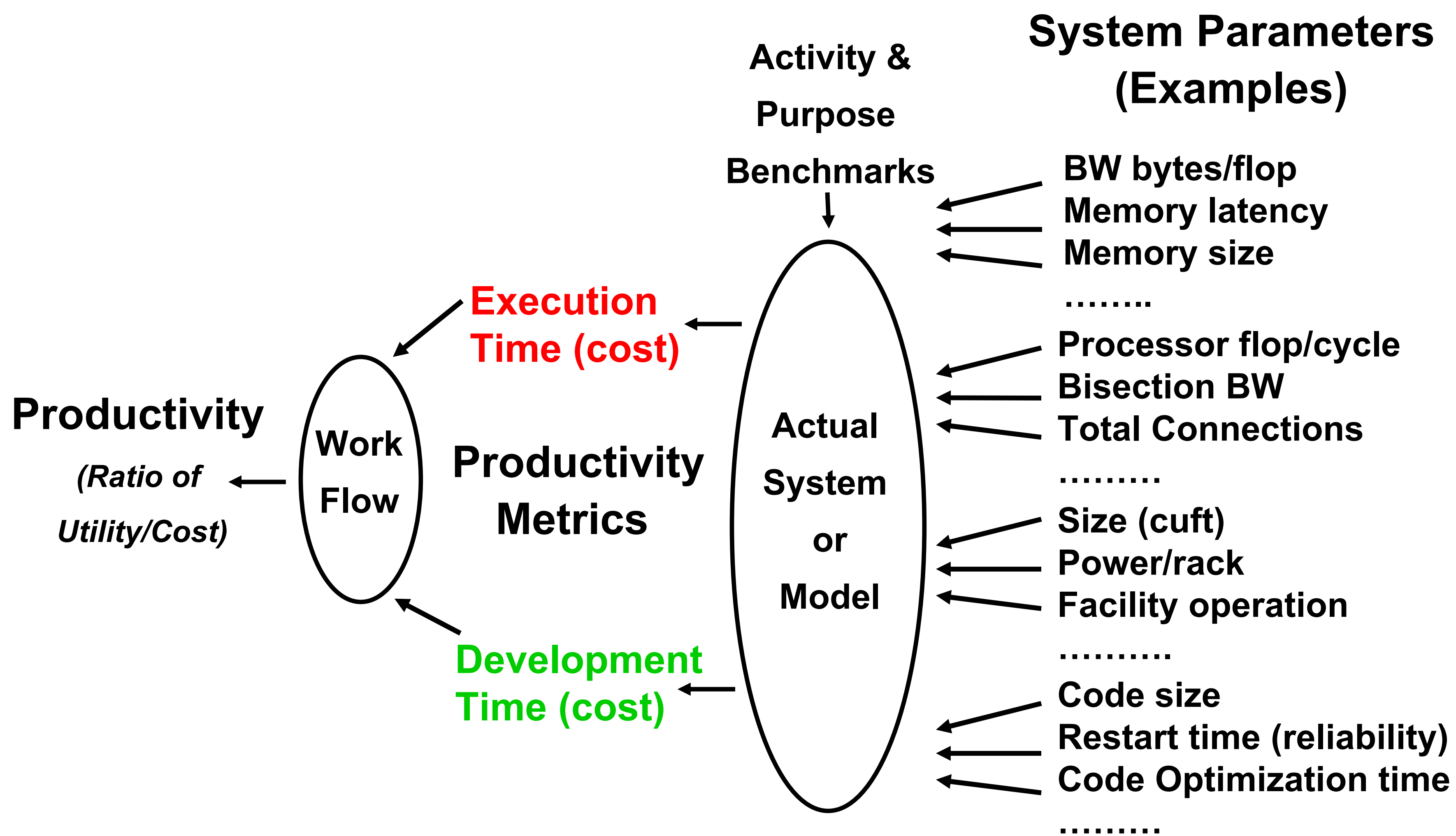


System Requirements



- 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

HPCS Productivity Framework



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

Benchmark Relationships

	Fixed Size	Scalable
Activity Based (Well Suited for Execution Measurement)	LINPACK NAS Parallel SPEC *HPCS Suite*	LINPEAK NAS Parallel Streams, GUPS
Purpose Based (Ideal for Development Measurement)	NSA Suites Some RFP Suites	HPCS Suite (Planned)

HPCS Focus

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