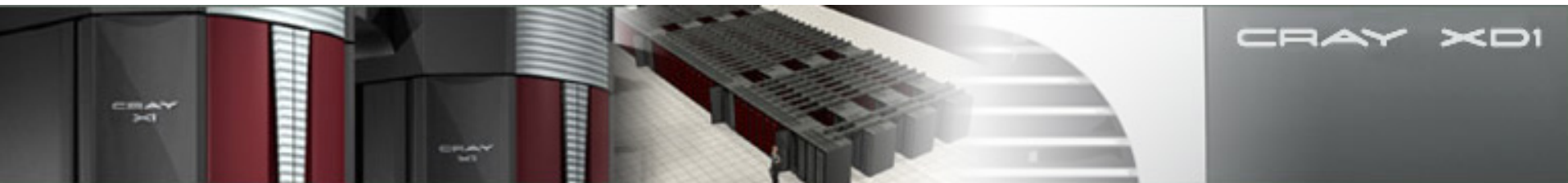
THE SUPERCOMPUTER COMPANY

The Cray Rainier System: Integrated Scalar/Vector Computing

Per Nyberg

11th ECMWF Workshop on HPC in Meteorology



Topics

- Current Product Overview
- Cray Technology Strengths
- Rainier System: Integrated Scalar/Vector Computing
  - Overview
  - Motivation
  - Benefits
- Cascade Project Overview

Cray's Family of Supercomputers

Cray X1

- 1 to 50+ TFLOPS
- 4 – 4,069 processors
- Vector processor for uncompromised sustained performance

Cray XT3

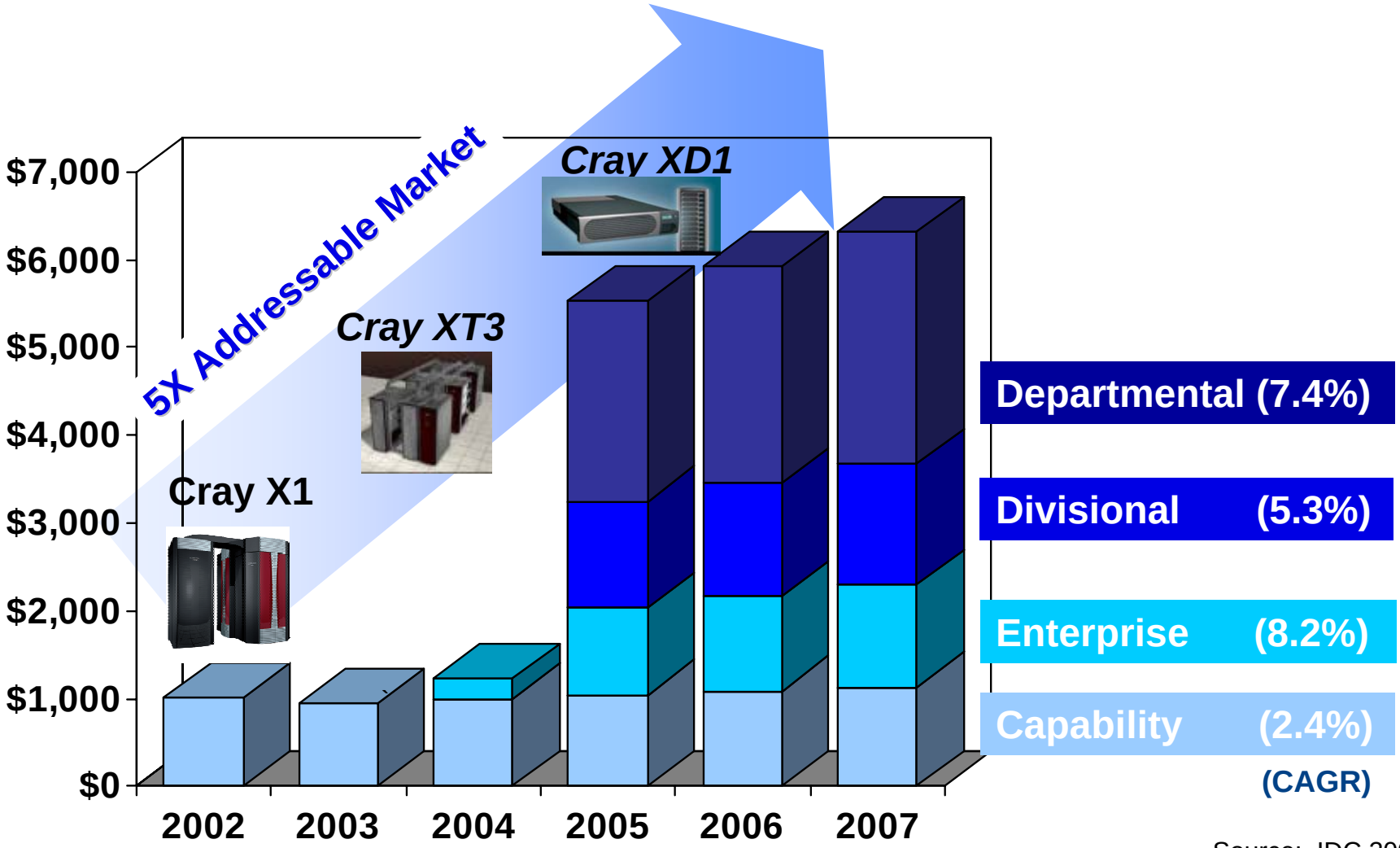
- 1 to 50+ TFLOPS
- 256 – 10,000+ processors
- Compute system for large-scale sustained performance

Cray XD1

- 48 GFLOPS - 2+ TFLOPS
- 12 – 288+ processors
- Entry/Mid range system optimized for sustained performance

Purpose-Built High Performance Computers

Growing the Addressable Market



Source: IDC 2003

Taking Success Formula to Broader HPC Market

Recent XD1 Announcements

Media:
Steve Conway
651/592-7441
sttico@aol.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. REPORTS FIRST EUROPEAN ORDER FOR CRAY XD1 SYSTEM, FROM HELMUT SCHMIDT UNIVERSITY

SEATTLE, September 2, 2004--Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that Helmut Schmidt University (HSU), Hamburg, Germany, is the first European organization to place an order for a Cray XD1™ supercomputer system. Financial terms were not disclosed.

The Cray XD1 system is scheduled to be installed in the fourth quarter at the University of the Federal Armed Forces within HSU, where it will be used primarily to support advanced education and training.

Media:
Steve Conway
651/592-7441
sttico@aol.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. REPORTS ORDER FOR 144-PROCESSOR CRAY XD1 SYSTEM FROM ALABAMA SUPERCOMPUTER AUTHORITY

634-Gigaflop Opteron/Linux Supercomputer "Promises To Provide Substantial Improvements In Computational Efficiency" Over Clusters

SEATTLE, October 21, 2004--Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that it has received an order for a 12-chassis Cray XD1™ supercomputer system for the Alabama Supercomputer Authority (ASA). Financial terms were not disclosed.

Media:
Steve Conway
651/592-7441
sttico@aol.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

USDA Forest Service Selects Cray XD1 Supercomputer to Improve Wildfire Smoke Plume Path Prediction and Tracking

SEATTLE, WA -- October 4, 2004 -- Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that the United States Department of Agriculture Forest Service has selected the new Cray XD1™ supercomputer to help improve the Forest Service's ability to predict and track the paths of smoke plumes from wildfires. Financial details were not disclosed.

"Tracking a smoke plume as it moves downwind from a fire requires all the computational complexity of a weather model run over a nationwide domain. Tracking the evolving chemical composition of said plume produces a task so computationally intense that we assumed we would not be able to afford any computer capable of performing it," said Bryce Nordgren, a Physical Scientist with the Forest Service's Fire Science Lab. "Reviewing the test case results from Cray restored our hope that we would be able to perform a scientifically meaningful simulation on our budget."

Cray/Media:
Steve Conway
651/592-7441
sttico@aol.com

Hinditron:
Saurabh Sonawala
+91-22-2361-0015
saurabh@hinditron.com

Cray/Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

HINDITRON SECURES FIRST ORDER IN INDIA FOR NEW CRAY XD1 SUPERCOMPUTER, FROM SAHA INSTITUTE OF NUCLEAR PHYSICS

SEATTLE, September 30, 2004--Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that Hinditron, its representative in India, has secured the first order in that country for the new Cray XD1™ supercomputer product. A Cray XD1 system with 96 compute processors and more than 422 billion calculations per second (gigaflops) of peak performance will be installed in late 2004 at the Saha Institute of Nuclear Physics in Kolkata (Calcutta). Financial terms were not disclosed.

Recent XT3 Announcements

Media:
Steve Conway
651/592-7441
sttico@aol.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. BEGINS SHIPPING CRAY XT3 MASSIVELY PARALLEL SUPERCOMPUTER BASED ON SANDIA "RED STORM" DESIGN

First Shipment Is To Sandia National Laboratories; Other Initial Customers Include Oak Ridge National Laboratory, Pittsburgh Supercomputing Center

SEATTLE, October 25, 2004 –Global supercomputer leader Cray Inc. (Nasda
CRAY) today reported that it has begun shipping the Cray XT3™ supercompu
industry standard massively parallel processing (MPP) system that strongly ad
record-setting scalability and sustained application performance of the renown
T3D™ and Cray T3E™ systems. U.S. list pricing for the Cray XT3 supercom
begins at about $2 million.

Media:
Steve Conway
651/592-7441
sttico@aol.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. ANNOUNCES 'RED STORM' CONTRACT WITH PITTSBURGH SUPERCOMPUTING CENTER

SEATTLE, September 30, 2004 –Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that the Pittsburgh Supercomputing Center (PSC) has received a $9.7 million grant from the National Science Foundation that will be applied toward the contract for a Cray Red Storm-based supercomputer. The total value of the contract was not disclosed.

A Red Storm-based system with 2,000 AMD Opteron™ processors and 10 teraflops (trillions of calculations per second) of peak performance is scheduled to be installed at PSC by the end of 2004, where it will be made broadly available to the U.S. academic research community. The capability of the system can be greatly expanded in the future.

"The Red Storm system in Pittsburgh will enable researchers to explore the limits of high-performance computing and to demonstrate the potential of this architecture for a wide range of scientific applications," said Peter Freeman, head of NSF's Computer and Information Science and Engineering directorate. "The system will complement other systems already provided by NSF to the national community and will strengthen the growing high-end computing partnership between NSF and the Department of Energy."

Recent X1 Announcements

Media:
Lori Kaiser
206/701-2233
lkaiser@cray.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. SIGNS $43.2 MILLION SUPERCOMPUTER CONTRACT WITH KOREA METEOROLOGICAL ADMINISTRATION

SEATTLE, May 18, 2004 – Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that it has signed a five-year contract valued at $43.2 million to supply the next-generation supercomputer to the Korea Meteorological Administration (KMA), one of the world's premier weather forecasting and research centers.

Cray and KMA also will jointly establish an Earth System Research Center aimed at advancing atmospheric modeling in the East Asia Pacific region, the company said.

Media:
Lori Kaiser
206/701-2233
lkaiser@cray.com

Investors:
Victor Chynoweth
206/701-2280
victorc@cray.com

CRAY INC. RECEIVES ORDER FOR CRAY X1 SUPERCOMPUTER FROM GMRI FOR NASA AMES RESEARCH CENTER

SEATTLE, May 4, 2004 – Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that it has received an order for a Cray X1™ supercomputer system from GMRI. Under the contract terms, the Cray X1 system was recently installed at the NASA Ames Research Center facility in Moffett Field, California. Financial terms were not disclosed.

ORNL National Leadership Computing Facility

CRAY INC. WILL TEAM WITH ORNL TO PROVIDE 100-TERAFLOP DEPARTMENT OF ENERGY LEADERSHIP-CLASS SUPERCOMPUTER

World's Most Powerful Computer Could Grow To 250 Teraflops In 2007

SEATTLE, May 12, 2004 – Global supercomputer leader Cray Inc. (Nasdaq NM: CRAY) today announced that it will collaborate with the Department of Energy's Oak Ridge National Laboratory (ORNL) to provide the world's most powerful supercomputer. The leadership-class system will be used for advanced scientific research.

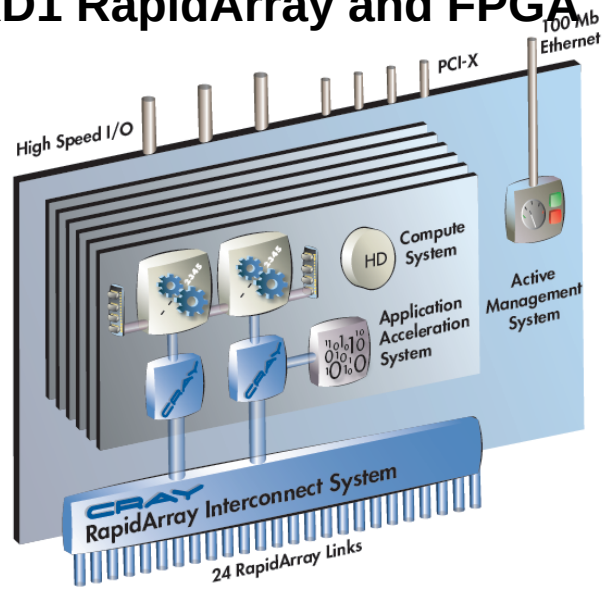
A 100-teraflop (trillions of calculations per second) Cray system at Oak Ridge is planned for 2006, with the potential to grow to 250 teraflops in 2007. Near-term plans call for increasing the capacity of the current Cray X1™ supercomputer at ORNL to 20 teraflops in 2004, with a 20-teraflop Red Storm-based system from Cray added in 2005. The systems will be housed in ORNL's new National Leadership Computing Facility in Oak Ridge, Tennessee.

Cray officials placed the contract value at about $25 million for 2004 and said it could be worth more than $125 million to the company over several years, subject to future federal funding.

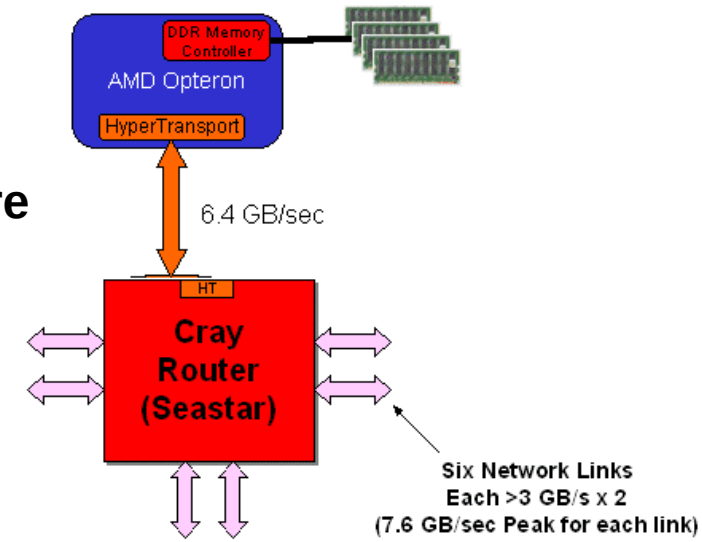
The leadership-class computing capability is part of the DOE Office of Science's 20-year plan to provide facilities "needed to extend the frontiers of science, to pursue opportunities of enormous importance, and to maintain U.S. science primacy in the world," according to Raymond L. Orbach, director of DOE's Office of Science. "Investment in these facilities will yield extraordinary scientific breakthroughs – and vital societal and economic benefits."

A Wealth of Technology

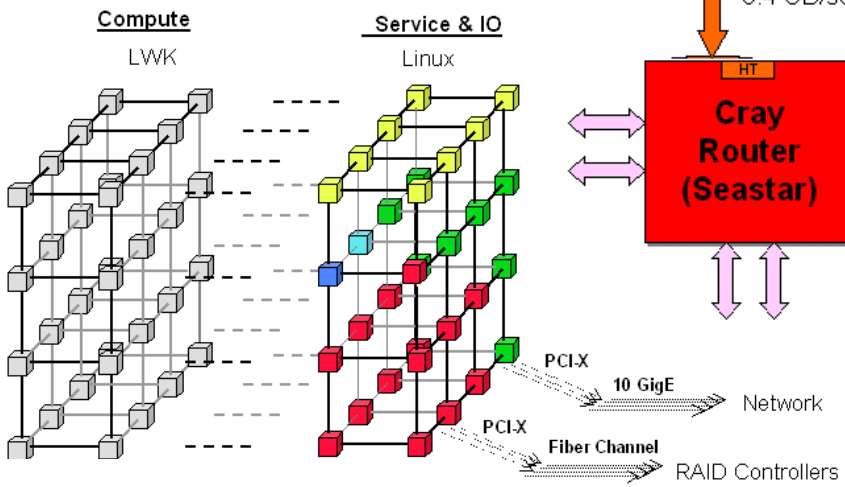
XD1 RapidArray and FPGA



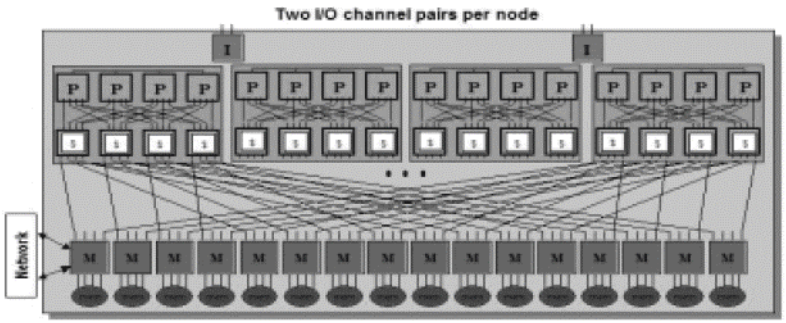
XT3 Compute PE



Red Storm Architecture

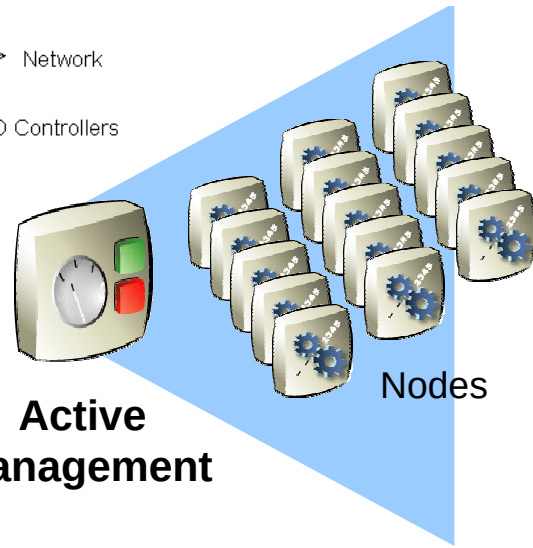


X1 Vector Node and Global Address Space



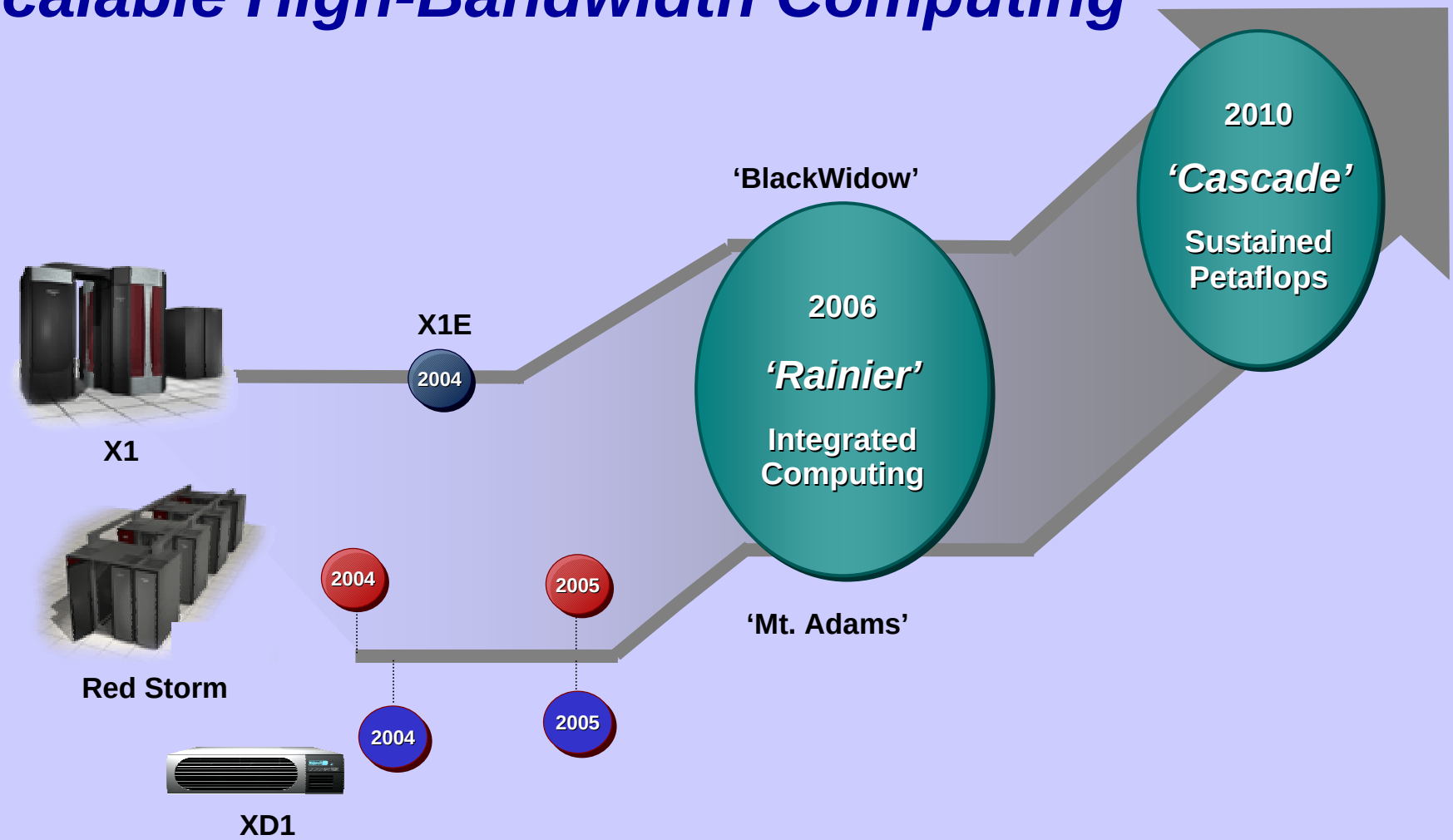
Inter-node network
two ports per M-chip
1.6 GB/s peak both directions per port
⇒ Node bandwidth is 1.6 GB/s x 2 directions
x 2 ports x 16 M-chips = 102.4 GB/s

Local node memory
peak BW=16 slices x 12.8 GB/s/slice = 204.8 GB/s
capacity = 16, 32GB



Cray's Vision…

Scalable High-Bandwidth Computing

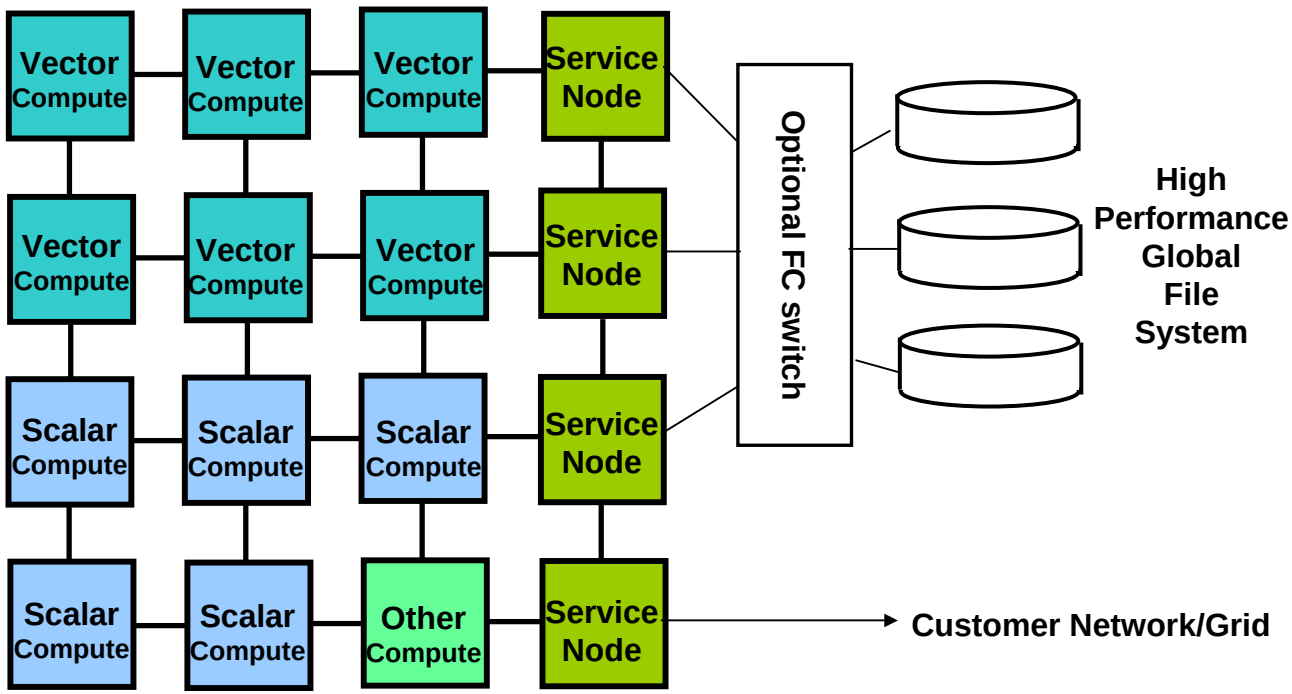


Rainier Integrated Computing

- The Concept:
  - Single system:
    - Common infrastructure and high performance network.
    - Common global address space.
    - Common OS, storage and administration.
  - Variety of compute nodes:
    - Follow on nodes for vector line (X1/X1E) and Opteron (XT3/XD1) lines.
  - Opteron based login, service and I/O nodes.
- First customer shipment in 2006.

Integrated Product Concept

- Single hardware infrastructure (cabinet, power, cooling, etc.)
- Common high speed network
- Heterogeneous compute capabilities
- Service nodes based on COTS processors
- Global address space across machine
- Linux-based OS



Motivation

- Different algorithms are appropriate for different architectures.
- Different requirements for:
  - flops vs. memory bandwidth
  - local memory size
  - mixed MPI/SMP vs. pure MPI
  - granularity of computation and communication
  - regular vs. irregular memory accesses and communication
  - network bandwidth
  - global vs. nearest neighbor communication
  - ability to tune application
  - capacity vs. capability
- Benchmark suites are often split as to best platform.

⇒ One size does not fit all.

⇒ Design a single system with heterogeneous computing capabilities.

Benefits of Integrated Computing

- Customer:
  - **Single solution for diverse workloads**
  - **Maintain and foster architectural diversity**
  - **Reduced administration and training costs**
  - **Single, unified user interface and environment**
  - **Better login, pre/post processing environment for vector machines**
  - **More configuration and upgrade flexibility**
  - **Improved performance by matching processor to the job**
- Cray:
  - **Better focus**
  - **Reduced development costs through commonality**
  - **Reduced manufacturing costs through increased volumes**
  - **Able to support specialized computing (vectors, MTA, FPGAs)**

Fit to Earth Science Requirements

- **Rainier architecture offers strong fit for diverse, complex and evolving workload:**
  - **Heterogeneity ideal for coupled modeling.**
  - **Capability features well suited to:**
    - **Production workload**
    - **Advanced research requiring high-resolution and complexity (eg: high-resolution, sub-scale processes, atmospheric chemistry).**
  - **Ability to explore alternative processor architectures (MT, FPGA).**
  - **Architecture flexibility at upgrade points; better leveraging investment.**

Increased Modeling Complexity

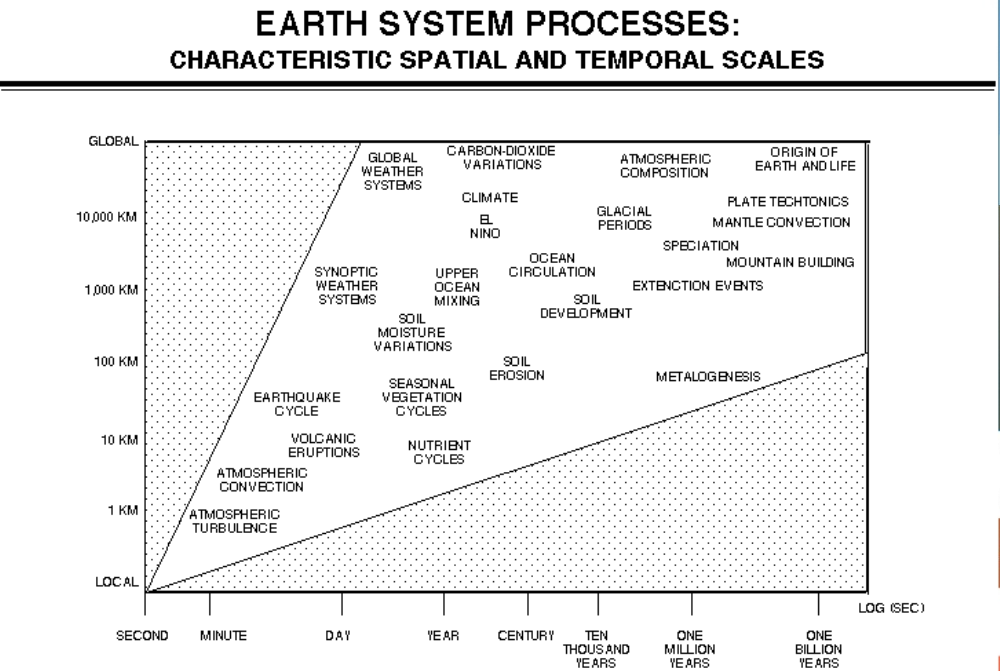
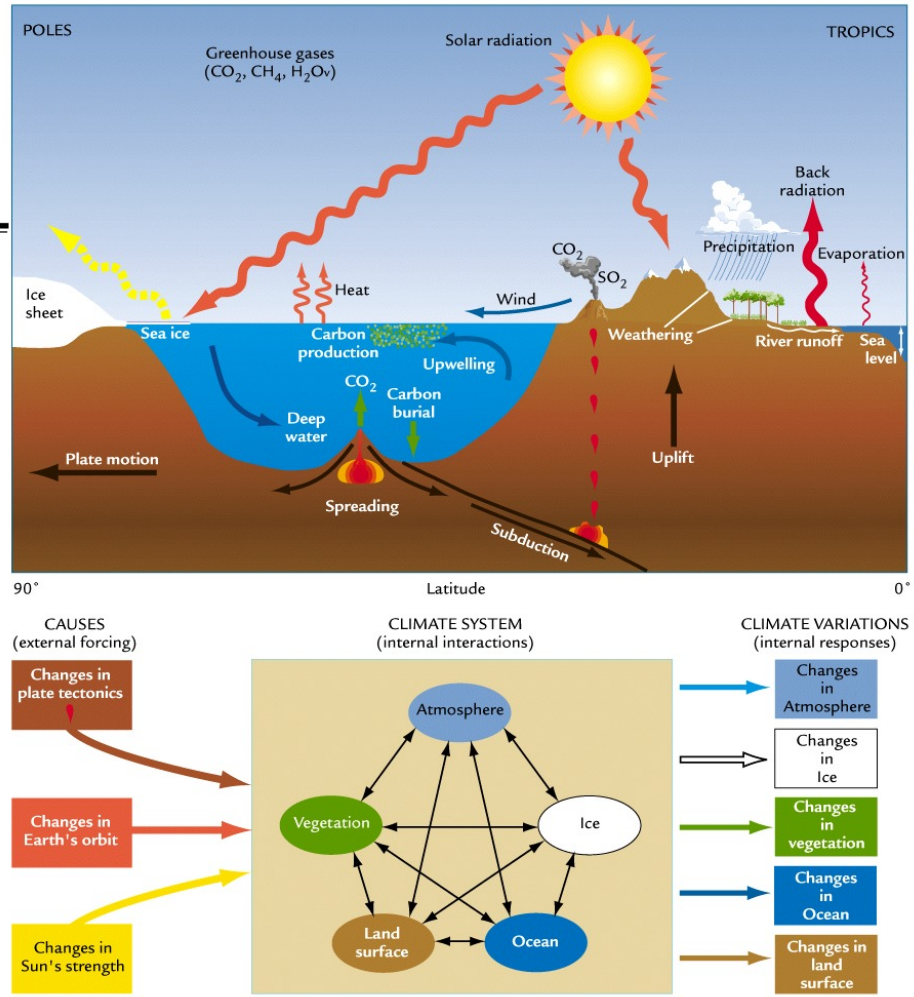
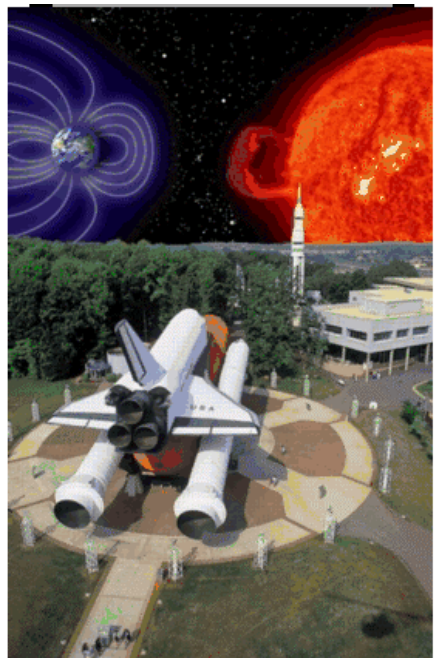


Figure 2. Earth System Processes Characteristic Scales



Rudiman, 2001

Increased Multi-Disciplinary Coupling



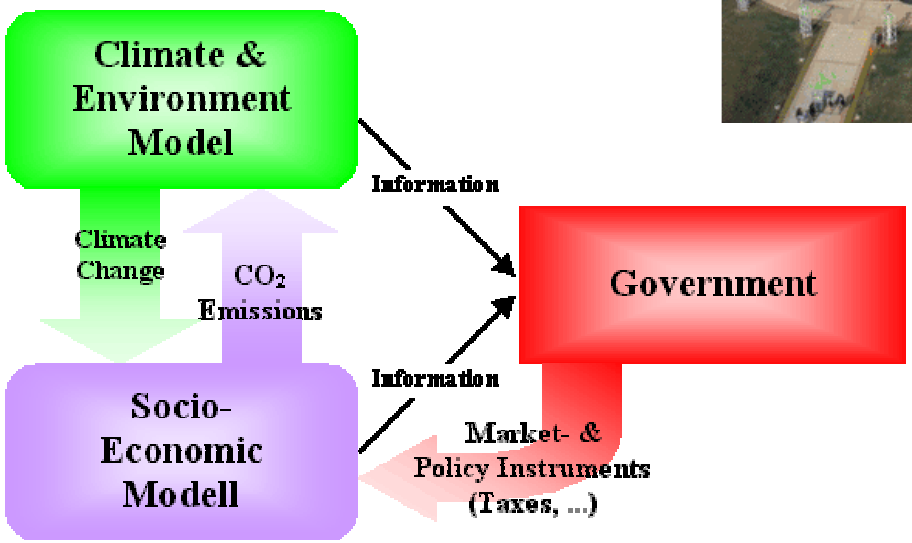
Huntsville 2004 Workshop

CHALLENGES TO MODELING THE SUN-EARTH SYSTEM

A discussion-oriented forum to advance the modeling of the coupled Sun-Earth system.

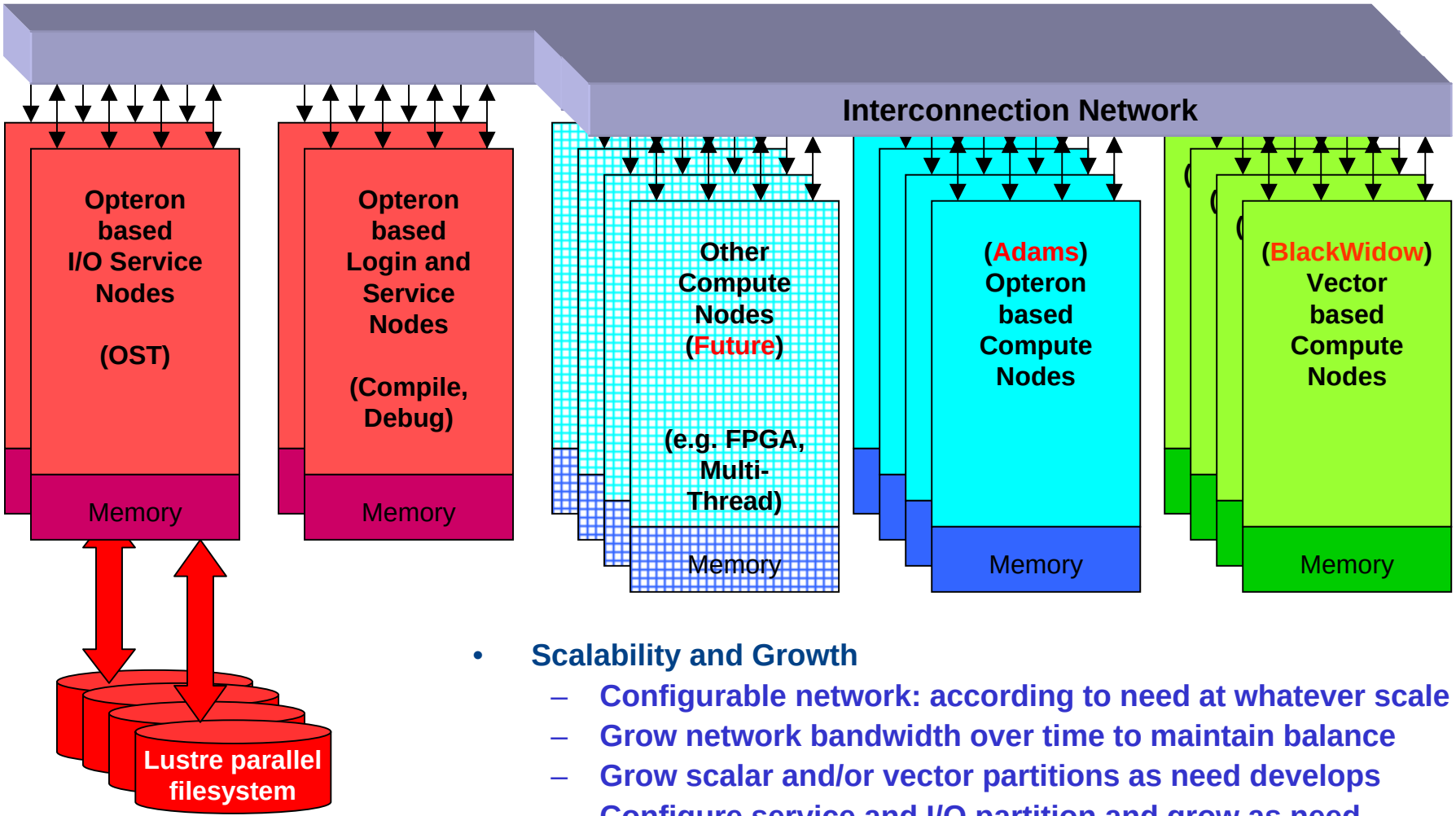
18-22 OCTOBER 2004
HILTON HUNTSVILLE HOTEL
HUNTSVILLE, ALABAMA

Sponsoring Institutions



MPI-M SDEM Structural Dynamical Economic Model

Scalable, Common Infrastructure



- **Scalability and Growth**
  - **Configurable network: according to need at whatever scale**
  - **Grow network bandwidth over time to maintain balance**
  - **Grow scalar and/or vector partitions as need develops**
  - **Configure service and I/O partition and grow as need develops**

Compute Nodes

- Adams (Opteron scalar nodes)
  - Excellent scalar performance
  - Very low memory latency
  - Many applications available (Linux + x86-64)
  - Potential for both uniprocessor and SMP nodes, single and dual cores
  - But, requires high degree of cache effectiveness
- BlackWidow (Cray vector nodes)
  - Fast, high bandwidth processors
  - 4-way vector SMP nodes
    - Large local memory
    - Supports hierarchical parallelism
  - Latency tolerance for global and irregular references
  - But, requires vectorizable code
- Other future planned compute capabilities
  - FPGA: direct hardware execution of kernels
  - MTA: highly threaded access to global memory

Reliability and Scaling Features

- Fault detection, diagnoses and recovery
  - Enhanced diagnostic error reporting
  - Memory retries for transmission-induced multi-bit errors
  - Timeouts and self-cleansing datapaths (no cascading errors)
- Hardware firewalls for fault containment
  - Secure, hierarchical boundaries between kernel groups
  - Protects the rest of the system even if a kernel is corrupted
- Graceful network degradation
  - Auto-degrade rails rather than lose a whole link
  - Hot swappable boards and reconfigurable routing tables
- Full node translation tables (NTTs)
  - Allows scheduling of parallel jobs across an arbitrary collection of processors, with efficient, scalable address translation

  $\Rightarrow$ Much higher system utilization under heavy workloads

Cascade Project

Cray Inc.

Stanford

Caltech/JPL

Notre Dame

Jet Propulsion Laboratory
California Institute of Technology

UNIVERSITY OF NOTRE DAME

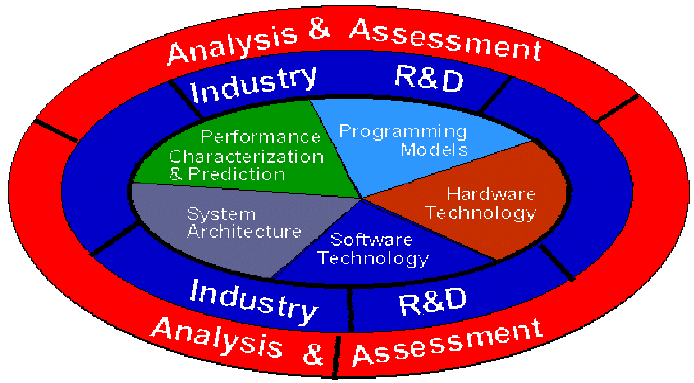
High Productivity Computing Systems

Goals:

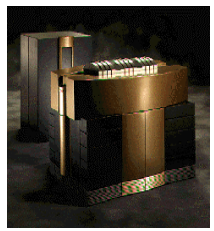
- **Provide a new generation of economically viable high productivity computing systems for the national security and industrial user community (2007 – 2010)**

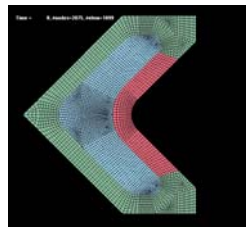
Impact:

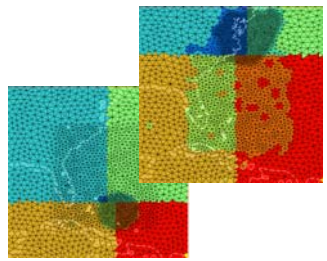
- **Performance** (efficiency): critical national security applications by a factor of 10X to 40X
- **Productivity** (time-to-solution)
- **Portability** (transparency): insulate research and operational application software from system
- **Robustness** (reliability): apply all known techniques to **protect against outside attacks**, hardware faults, & programming errors

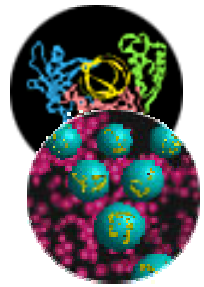


HPCS Program Focus Areas









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)

HPCS Phases

CRAY | SGI | Sun | HP | IBM

1 Year

2H 2002 – 1H 2003

$3M/year

Phase I: Concept Development

- Forecast available technology
- Propose HPCS hw/sw concepts
- Explore productivity metrics
- Develop research plan for Phase II

CRAY | Sun | IBM

3 Years

2H 2003 – 1H 2006

$17M/year

Phase II: Concept Validation

- Focused R&D
- Hardware and software prototyping
- Experimentation and simulation
- Risk assessment and mitigation

? | ?

4 Years

2H 2006 – 2010

$?/year

Phase III: Full Scale Product Development

- Commercially available system by 2010
- Outreach and cooperation in software and applications areas

The HPCS program lets us explore technologies we otherwise couldn't.
A three year head start on typical development cycle.

Cray's Approach to HPCS

- High system efficiency at scale
  - Bandwidth is the most critical and expensive part of scalability
  - Enable *very* high (but configurable) global bandwidth
  - Design processor and system to use this bandwidth *wisely*
  - Reduce bandwidth demand architecturally
- High human productivity and portability
  - Support legacy and emerging languages
  - Provide strong compiler, tools and runtime support
  - Support a mixed UMA/NUMA programming model
  - Develop higher-level programming language and tools
- System robustness
  - Provide excellent fault detection and diagnosis
  - Implement automatic reconfiguration and fault containment
  - Make all resources virtualized and dynamically reconfigurable

Summary

- Cray offers a unique range of science-driven technologies:
  - XD1, XT3, X1/X1E
- Rainier architecture offers strong fit for diverse, complex and evolving earth sciences workload.
- Cray continues to support sustained innovation to meet the needs of the scientific community:
  - Rainier: Integrated computing capability in 2006
  - Cascade: Aggressive research program for 2010


What
Do
You
Need
To
Know ?
CRAY