NASA's Science and Engineering Applications in the Future

ZettaFLOPS Forum: Frontiers of Extreme Computing
October 26, 2005, Santa Cruz, California

Dr. Rupak Biswas

Chief (Acting), NASA Advanced Supercomputing (NAS) Division

NASA Ames Research Center

Moffett Field, California

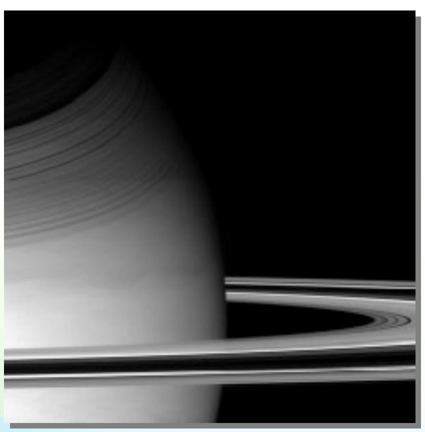
NASA's Mission Directorates

- ***Aeronautics Research Mission Directorate (ARMD):***
  - To pioneer the identification, development, verification, transfer, application, and commercialization of high-payoff aeronautics and space transportation technologies.

Artist concept of a vision for the National Air Transportation System in 2025, allowing airport and airspace capacity to be more responsive, adaptable and dynamic.

- ***Exploration Systems Mission Directorate (ESMD):***
  - To develop capabilities and supporting research and technology that enable sustained and affordable human and robotic exploration; includes the biological and physical research necessary to ensure the health and safety of crew during long duration space flight.

Artist concept of a future lunar exploration mission.



- ***Science Mission Directorate (SMD):***
  - To carry out the scientific exploration of the Earth, Moon, Mars, and beyond; charts the best route of discovery; and reaps the benefits of Earth and space exploration for society.

Sidelong view of Saturn's rings captured by Cassini spacecraft on Dec. 14, 2004.

NASA’s Mission Directorates (cont.)

International Space Station

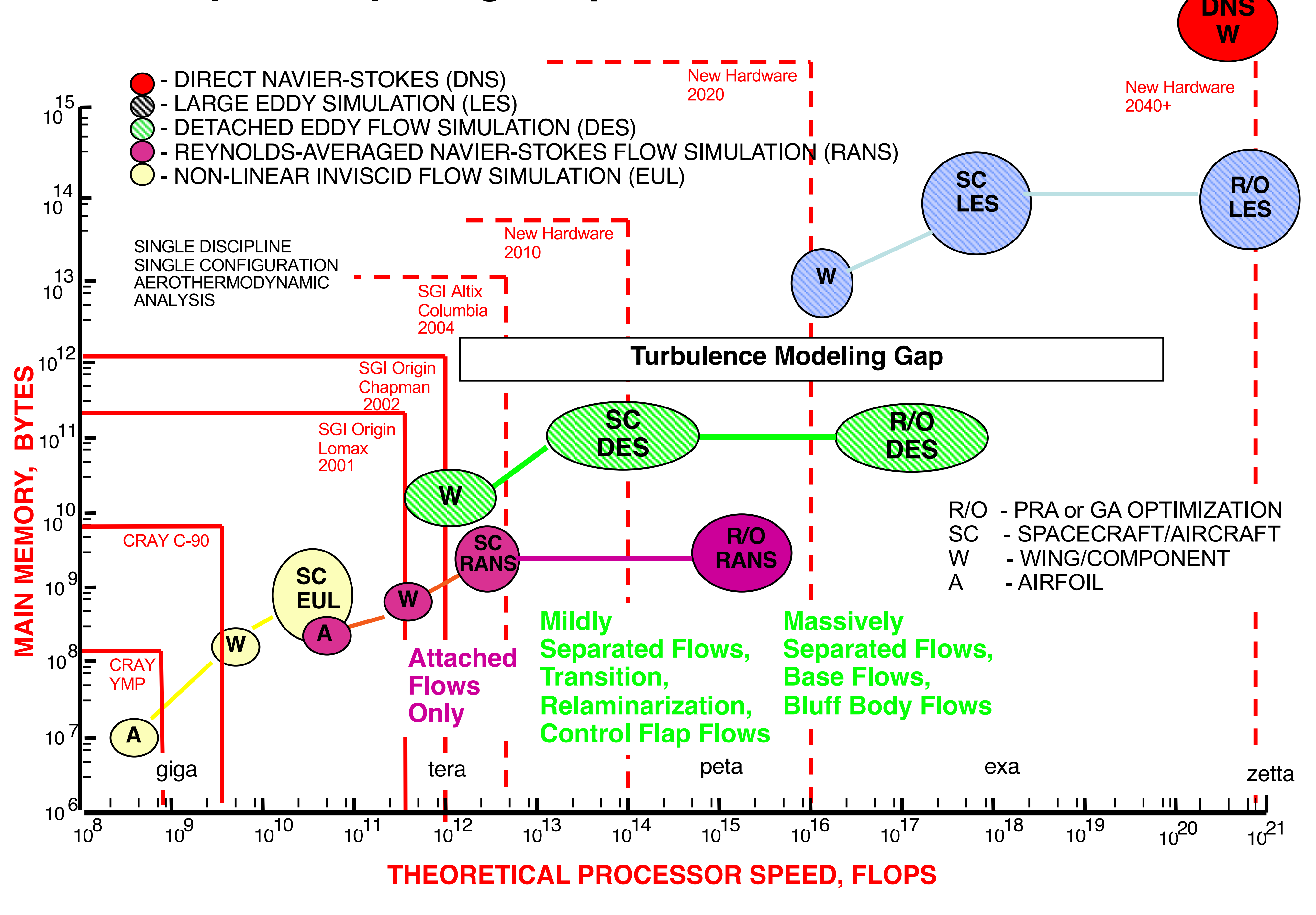
- ***Space Operations Mission Directorate (SOMD):***
  - To provide many critical enabling capabilities that make possible much of the science, research, and exploration achievements of the rest of NASA. It does this through the three themes of the International Space Station, the Space Shuttle Program, and Flight Support.

- ***NASA Engineering and Safety Center (NESC):***
  - The NESC is an independent organization, which was charted in the wake of the Space Shuttle Columbia accident to serve as an Agency-wide technical resource focused on engineering excellence. The objective of the NESC is to improve safety by performing in-depth independent engineering assessments, testing, and analysis to uncover technical vulnerabilities and to determine appropriate preventative and corrective actions for problems, trends or issues within NASA's programs, projects and institutions.

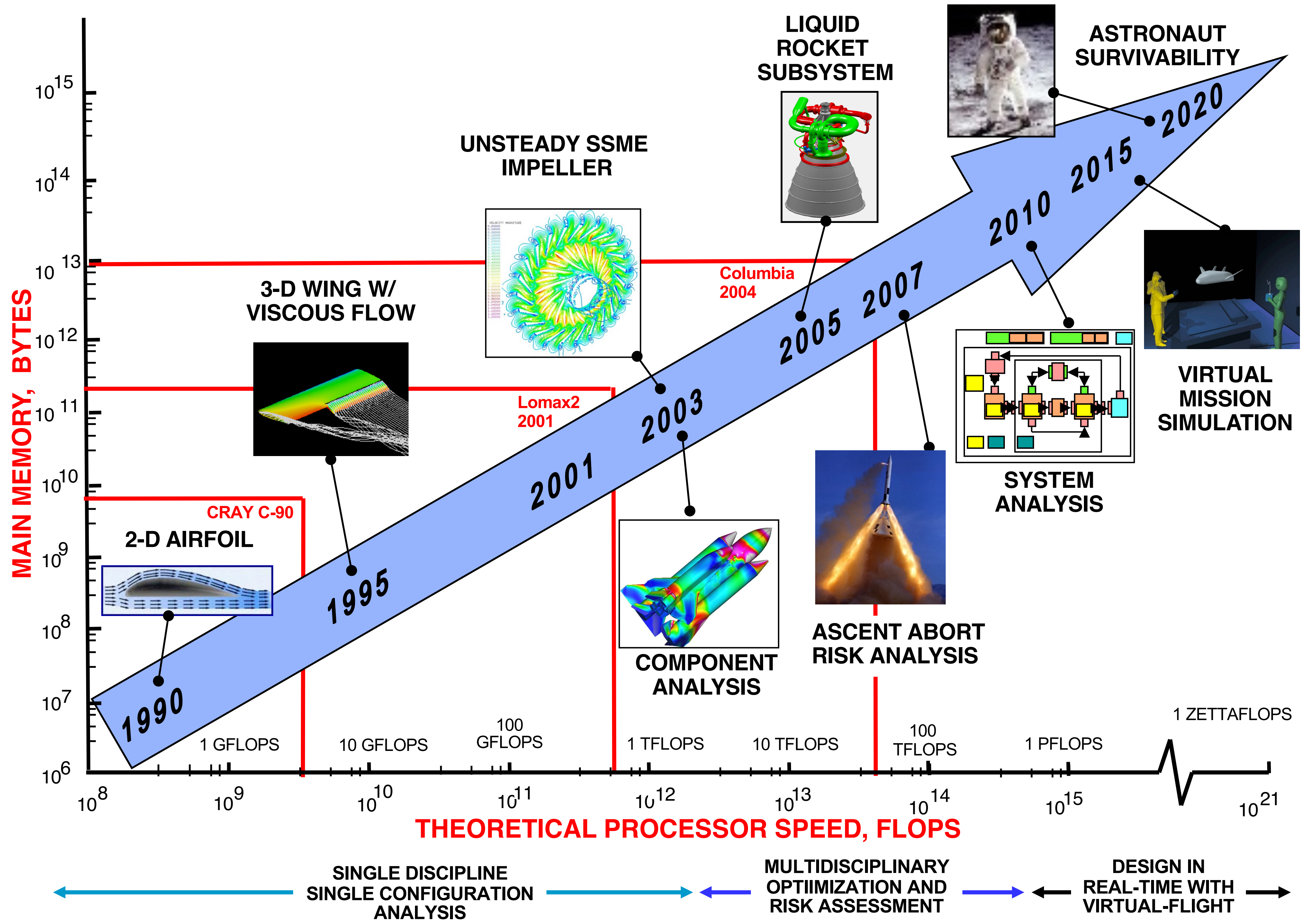
Integrated Safe Spacecraft Design: 2020 Goal

- **Vision**
  - Full simulation and optimization of multiple vehicle designs with safety analysis to enable automated identification and simulation of failures and effects against a suite of health management technologies for survivability analysis and cost trade-offs. Real-time generation of flight simulation enables pilot-in-the loop design.
- **Technology Advances**
  - Full, time-accurate, multi-disciplinary vehicle simulations with high-fidelity modeling of safety critical elements
  - Real-time data generation for piloted simulation
  - Integration of health management strategies into vehicle behavior models
- **Aerospace Technology Benefits**
  - Mission Safety - Supports order of magnitude improvement in mission safety from 2nd Gen RLV baseline
  - Mission Affordability - Supports development of cost-effective survivable systems through higher design certainty and lower requirement for safety margin
  - Development of advanced tools and processes for rapid, high-confidence design - Enables early evaluation and decision making within a virtual design process
  - Revolutionary solution for fundamentally new missions - Enables simulation and evaluation of self-repairing systems technologies

Supercomputing Requirements



Supercomputing Requirements: Mission Applications



Columbia: World Class Supercomputing

- Currently the world's third fastest supercomputer providing 62 Tflops peak and 52 Tflops Linpack sustained performance
- Conceived, designed, built, and deployed in just 120 days
- A 20-node constellation built on proven 512-processor nodes
- Largest SGI system in the world with over 10,000 Itanium 2 processors
- Provides the largest node size incorporating commodity parts (512) and the largest shared-memory environment (2048)
- 88% efficiency tops the scalar systems on the Top500 list
- Most importantly, having mission impact almost immediately

Systems: SGI Altix 3700 and 3700-BX2
Processors: 10,240 Intel Itanium 2
Global Shared Memory: 20 Terabytes

Front-End: SGI Altix 3700 (64 proc.)
Online Storage: 440 Terabytes RAID
Offline Storage: 6 Petabytes STK Silo

Internal Networks:
Internode Comm: Infiniband
Hi-Speed: 10 Gigabit Ethernet

Exploration Systems: Space Flight Applications

- **In computational fluid dynamics:**
  - Real time, high-fidelity simulation for digital flight will be possible.
  - With today's technology and computing capabilities, we focus on high-fidelity simulation of a certain phenomena on a specific section of the vehicle. Some examples are propulsion, external body dynamics with six degree of freedom (debris transport analysis), re-entry, fluid/structure interaction, etc.
  - In future, these simulations have to be very fast and integrated at the system level so that complete flight can be simulated in real time.

QuickTime™ and a
Animation decompressor
are needed to see this picture.

Return to Flight: Six-degree-of-freedom CFD analyses to determine the impact conditions and locations, using the aerodynamic characteristics of potential debris.

QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

Flowliner: Instantaneous snapshot from time-accurate fuel flowliner analysis using 66 million grid points with 262 overlapped zones.

POC: Cetin Kiris, Mike Aftosmis, Stuart Rogers, NASA Ames Research Center, CA

Exploration Systems: Digital Astronaut

Human Brain Circulatory System under Altered Gravity

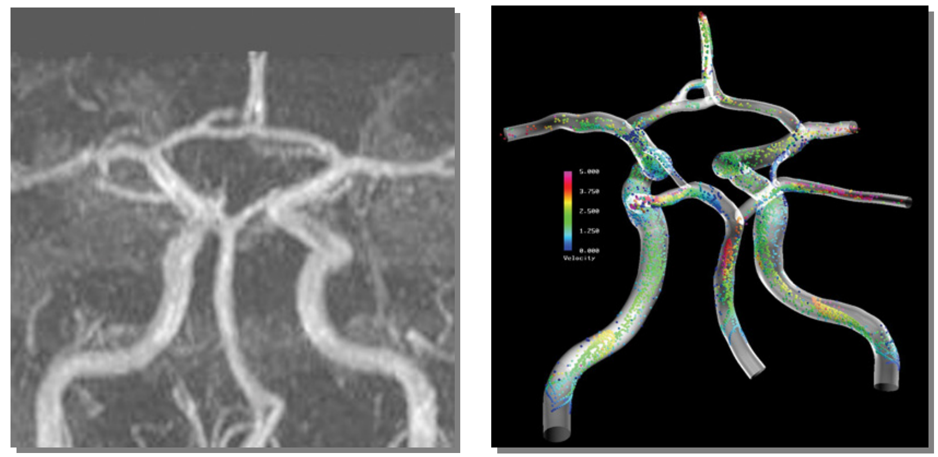
- For astronauts, blood circulation and body fluid distribution undergo significant adaptation both during and after long-duration space flights.
- To assess the impact of changing gravitational forces on human space flight, it is essential to quantify the blood flow characteristics in the brain under varying gravity conditions.
- Currently, NASA is working on blood flow simulations in the arterial system of an astronaut.

With increased computational capabilities, we will be able to:

- – Extend the simulations from just the arterial system to the entire body; then, extend this capability to couple with other systems such as the respiratory system
- – Construct a bridge between macroscopic and microscopic (molecular) scal; then, extend studies from the capillary level to the cell level

This will enable us to predict astronauts' performance during long space flights.

POC: Cetin Kiris, NASA Ames Research Center, CA



Human-specific geometry of the cerebral arterial tree reconstructed from magnetic resonance images are used in conjunction with supercomputing technology to establish large-scale continuum fluid simulations.

MICROGRAVITY CIRCULATORY SYSTEM

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

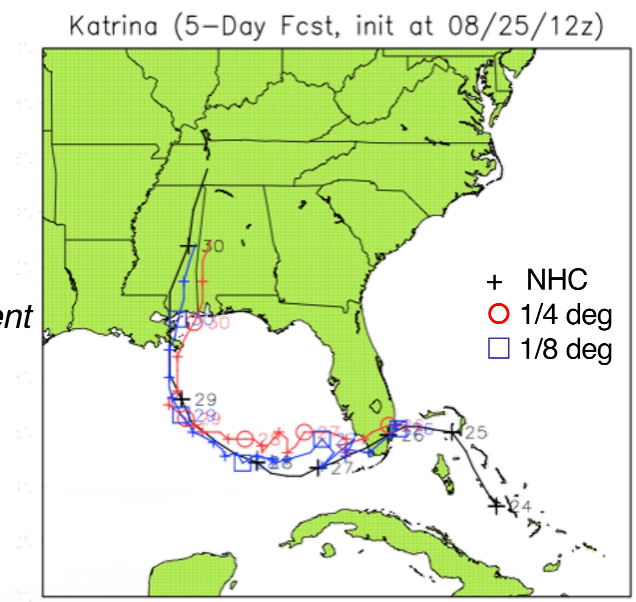
RADIATION SHIELDS

Earth Science: Finite-Volume General Circulation Model (fvGCM)

- Even with unlimited computing resources, there will be a hard limit on how far we can go in resolution beyond which we cannot possibly model without also modeling society, biology (such as whale movements), etc. We will also need to model human behavior, if the resolution is of the order of 1 meter.
- The ultimate useful min(dx, dy, dz), in a global model, would be about 10 meters. In that case, it would be an increase in computing power that is

  ~ (10km/10m)**4 = (1.E3) ** 4 = 1.E12 times more than what *Columbia* currently provides!

Katrina:
Very promising and comparable track predictions at different resolutions from a 5-day forecast (1/8 degree fvGCM)



QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

Higher Resolution Hurricane Track Prediction fvGCM Code Simulations - Hurricane Francis 09/04 (Total Precipitable Water - Resolution: 1/12th of a degree)

POC: Bowen Shen, NASA Goddard Space Flight Center

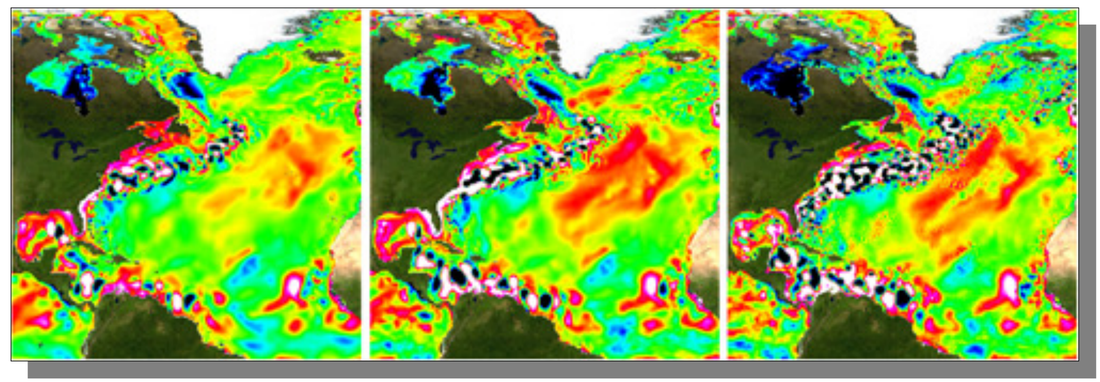
Earth Science: Estimating the Circulation and Climate of the Ocean (ECCO)

Two CPU-intensive problems that ECCO consortium is working on but are unlikely to be solved in a definitive way during the next 25 years.

- First problem is convergence of numerical ocean model solutions as resolution is increased. By some estimates, the ocean is a turbulent fluid with upwards of 10^{24} degrees of freedom at each instant of time. To date, the largest computation that ECCO has conducted on Columbia is an ocean simulation with approximately 10^9 degrees of freedom at each time step. Taking into account shorter time steps that are needed to simulate smaller volumes of water, maybe we will not have a definitive answer to the question of convergence until available computational power is increased by a factor of 10^{20}.

- Second problem is ocean state estimation. Assuming 1-s time steps, an exhaustive search of all possible solutions for above ocean model for 1000 years (the overturning time scale of the oceans) would require approximately 10^{60} increase in computer FLOPS relative to *Columbia*.

- Add to above model, atmosphere, land, and ice processes, and clearly, there is a very long way to go before earth scientists will be fully satisfied with computing capability.



To improve specification of error statistics and parameterization of small-scale processes in ECCO and to investigate solution convergence, a series of full-depth, global-ocean, and sea-ice simulations at increasingly higher resolution (1/4, 1/8, and now 1/16-deg) are being carried out on the 2048-CPU partition of Columbia. The figure shows one-month sea-surface height difference in the Gulf Stream region from these three integrations (left panel: 1/4 deg; middle panel: 1/8 deg; and right panel: 1/16 deg). Color scale is -0.125 m to 0.125 m.

POC: Dimitris Menemenlis, Jet Propulsion Lab, California Institute of Technology, Pasadena, CA

Space Science: Stellar Models and Supernovae

The influence of computers in the next 25 years will be much greater than the huge impact they have had in the last 25.

- In astronomy, large ground-based telescopes will use adaptive optics and other computer-assisted data enhancement techniques to do observations from the ground that presently can only be done from space.
- With a 1000-fold increase in present computer power, models will start from a given presupernova model (mass, angular momentum, distribution, etc) and determine the explosion - including gamma-ray bursts as a subset, as well as the properties of a neutron star, pulsar, magnetar, or black hole that is produced, the nucleosynthesis, and the appearance of the supernova remnant. This includes a detailed description of the neutron star magnetic field inside and out.

- Within 10 years, snapshots of presupernova evolution studied in 3D with magnetic fields will give a much better understanding of the transport of angular momentum, convection, convective overshoot, etc so that the presupernova model has a good physical basis.
- Nucleosynthesis will be calculated in all stellar models and supernovae with unprecedented accuracy. Improvements in cross sections will also occur in laboratory and computational nuclear physics. The models will be able to describe the chemical evolution of galaxies of all types, not just the Milky Way.

QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

POC: Stan Woosley, University of California, Santa Cruz

Space Science: Stellar Models and Supernovae

- Shown here is an animation of a reactive rising bubble in conditions appropriate for Type Ia supernova. The standard picture of an SNe Ia is that it begins as one or more hotspots near the center of a carbon/oxygen white dwarf star. These hotspots quickly burn the carbon fuel to nickel, via thermonuclear fusion reactions, and a flame is formed. The hot ash is less dense than the surrounding fuel, so the bubble of ash will buoyantly rise, while the flame continues to burn outward.
- In these simulations, we were interested in understanding the role of the turbulence that develops on the sides of the bubble. In particular, can these turbulent eddies cause the bubble to shed some sparks of hot partially burned fuel or ash, which would then ignite the star in other regions.
- These calculations are very computationally demanding, requiring 100s of millions of zones to accurately capture the flame structure and the developing turbulence. With zettaflop capability, we could certainly capture this transition to turbulence and gain a detailed understanding of the evolution of these bubbles.

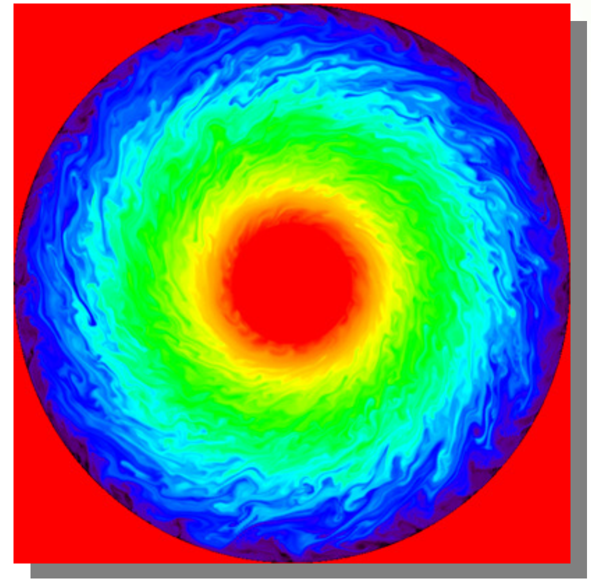
QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

POC: Mike Zingale, Stan Woosley, University of California, Santa Cruz; John Bell, Marc Day, and Charles Rendleman at Lawrence Berkeley National Laboratory.

Space Science: Simulating Convection and Magnetic Field Generation in the Interiors of Planets and Stars

Our goals and dreams expand much faster than computer power…

- With four or five times the computing resources than currently available today, it would be possible to simulate the interior dynamics of stars and planets as strongly turbulent convection in 3D, as can only now be done in 2D. Comparisons of 2D laminar and turbulent simulations clearly show fundamental differences. This suggests that our current 3D simulations, which are at best weakly turbulent, may be still far from realistic. Simulating strong turbulent convective dynamos requires much greater spatial and temporal resolution.
- So, it's not that our solutions would be just a little more accurate, if we had more computational resources; they would likely be fundamentally different and lead to new discoveries and predictions.



Snapshot of the entropy from one of our simulations of turbulent convection in a rapidly rotating disk or equatorial plane of a star or giant planet

- Although the current solutions do resemble observations to first order and our understanding of these processes continues to improve, we cannot include all the spatial and temporal scales that are part of the actual turbulent mechanisms. The situation has improved significantly over the past two decades and no doubt will continue to improve over the next two decades. Hopefully by then, it will be clear that we will be simulating all the important scales.
- We would also like to include the more detailed physics, chemistry and radiative transfer in our 3D time-dependent models that currently only 1D (spherically-symmetric) evolution models can include.
- We would like to simulate every major body in the solar system simultaneously with all the interactions among them included, while simulating their internal dynamics. The computational resources needed to do this would be difficult to estimate - but there will never be a time when those working on state-of-the-art problems will feel they have enough resources.

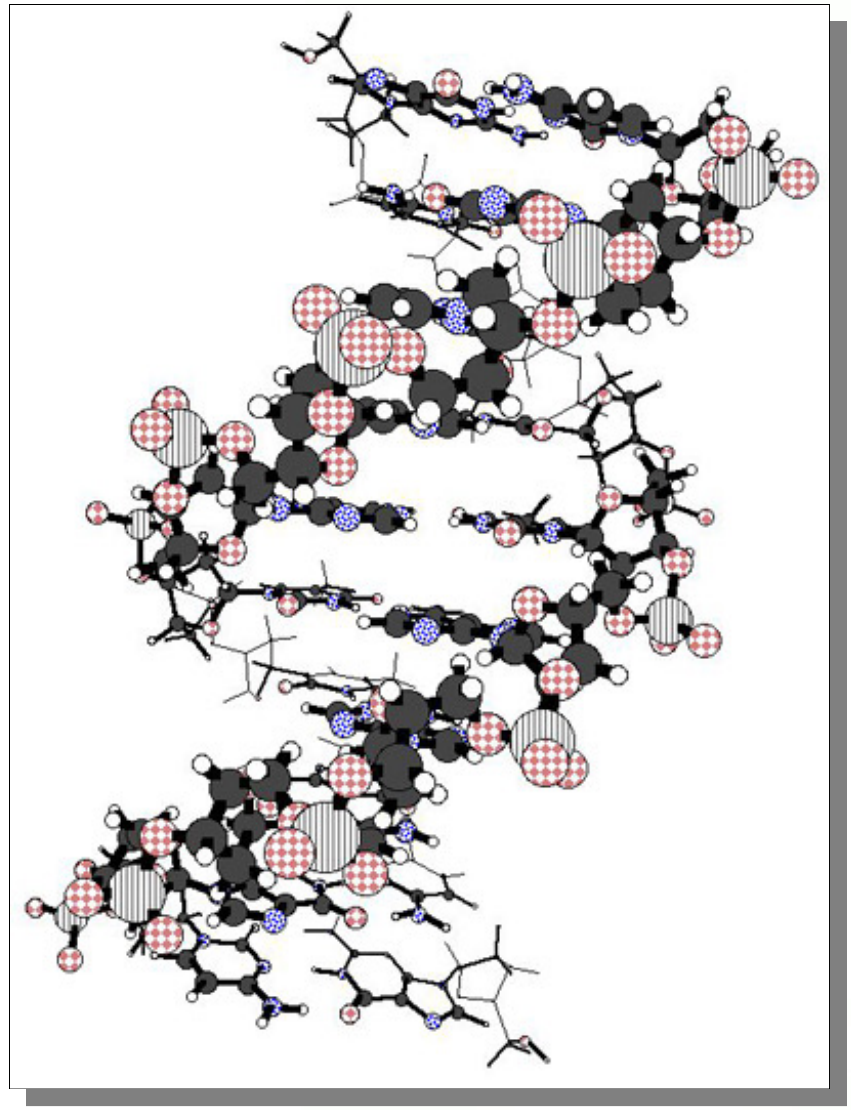
POC: Gary Glatzmaier, Earth Science Dept., University of California, Santa Cruz

Backup Slides

Computational Chemistry

Computational chemists are currently interested in two areas, radiation biology and computational material science.

- Simulation of Radiation Damage to DNA:
  - Double or triple the computing power allows us to study damages to the Watson-Crick base pair quantum mechanically. Currently, we can only apply quantum mechanics to individual bases. It will also allow us to study the role of water and protein in more detail.
  - Unlimited computing facility will allow us to follow the radiation damage from initial hit by the space radiation, subsequent chemical reactions that occur in the cell leading to the biological response. At present these studies are piecemeal.
- Computational Material Science:
  - In a multi-scale modeling of materials, double or triple the computing power allows us to extend both the size of the quantal region as well as the molecular dynamics region. This is important to simulate the energetic reactions such as pyrolysis of TPS during a high-speed vehicle entry into the atmosphere.



Multi-scale modeling of materials and bioscience - 10-base pair DNA

POC: Winifred Huo, NASA Ames Research Center

ZettaFLOP Visualization and Data Analysis

With zettaFLOP capabilities, we would be able to achieve:

- Visualization of zettabyte datasets
- High-quality ray traced volume rendering with realistic shading models (true shadows, accurate material reflectance & absorption)
- Interactive radiosity calculations
- Interactive 3D LIC (line integral convolution - "van Gogh" technique)
- Interactive feature exploration and detection, using sophisticated kernel methods, non-linear fitting, etc.
- Interactive "causality exploration", using high-order Bayesian conditional probability networks
- Natural language interfaces to visualization applications
- Simulations would be the vis-techniques (i.e. there would be no separation between the computation/ analysis/visualization stages (true "interactive visual supercomputing")
- Sensory devices could provide extremely good immersion, using feedback even of saccadic eye movements
- Neural network-based "cognitive prosthetics" could assist data analysis and exploration, using, e.g., map seeking circuits, adaptive resonance, probability collectives and other information theoretic techniques.

POC: Chris Henze, NASA Ames Research Center

Artist concept of a visualization tool - a double hyperwall

Integrated Safe Spacecraft Design: 2010 Goal

- **Vision**
  - Single vehicle design integrating full, high fidelity multi-disciplinary analyses with FMEA. Enables perturbation of the simulation to introduce failures and re-fly through mission profiles to determine survivability.

- **Technology Advances**
  - Full 3-D multidisciplinary simulations

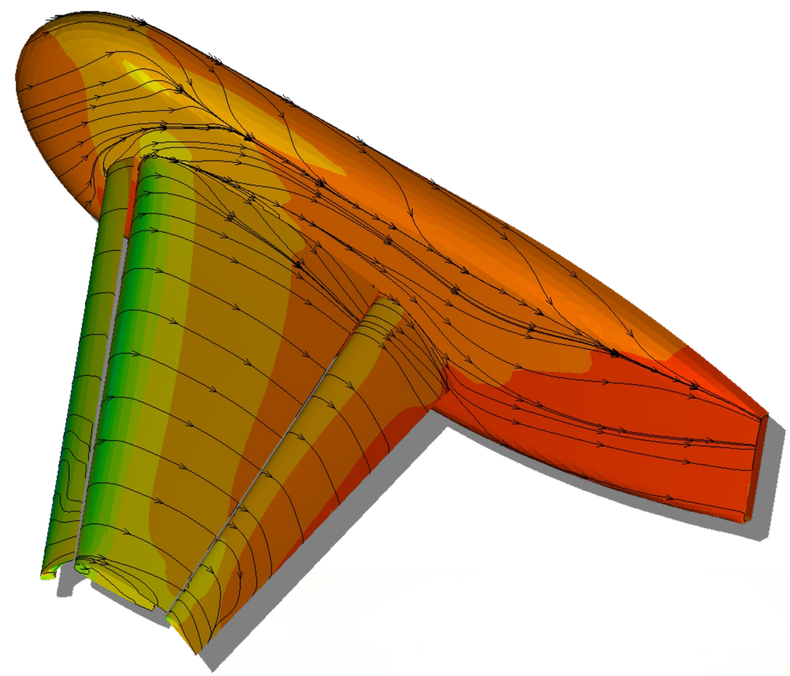
- **Benefits**
  - Mission Safety - Supports 2nd Generation RLV goals of 1:10,000 risk of crew loss
  - Develop revolutionary technologies to enable new aerospace capabilities - Enables an order of magnitude safer human space flight missions.

Aeronautics Research: High-Lift Aerodynamics

- The grid requirements for an accurate computation of high-lift aerodynamics is staggering. For the simple geometry in the figure below, systematic refinement of the grid resulted in 46 million cells before a reasonable level of CLmax agreement was achieved. With the combination of *Columbia* run time and queue structure, it took 135 days of round-the-clock submittals to get **one** 13 point lift polar.

- A colleague, Dr. Shahyar Pirzadeh, is presently trying to apply these guidelines to a Boeing 777 in high-lift configuration. He is presently up to 108 million cells and is getting some results indicating that this may not be adequate. These calculations are taking weeks and weeks on 360 processors.

- Therefore, if we could do what we would like to do with unlimited computational capacity, we would like to perform these computations in a few days or less.

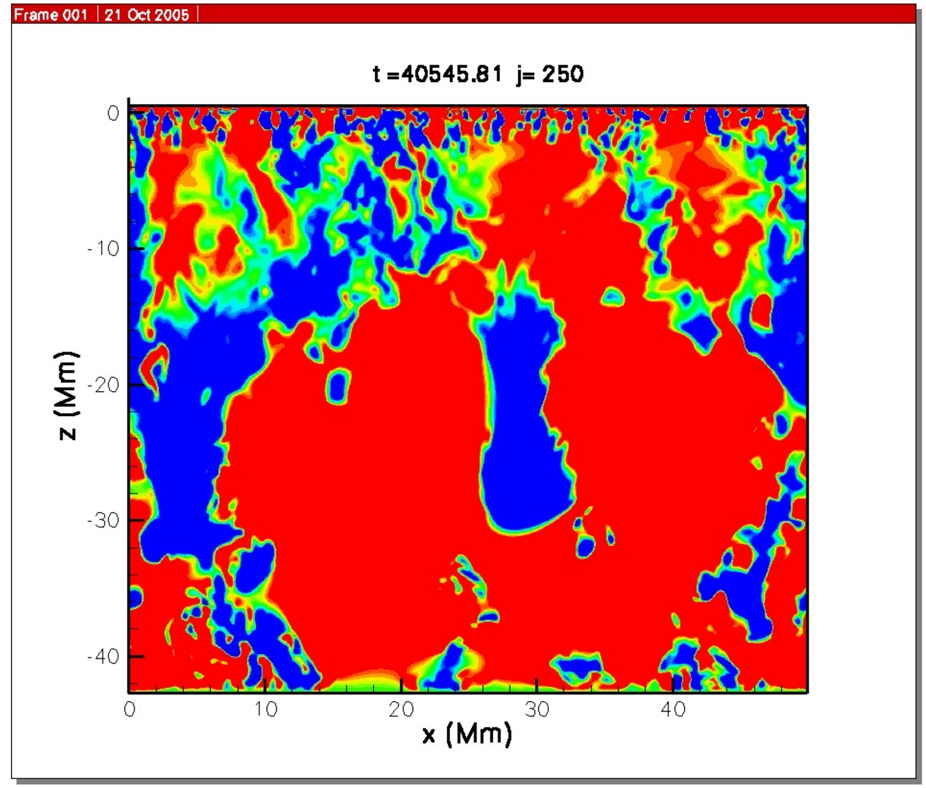


Trapezoidal wing high-lift geometry and typical lift-polar

POC: Neal Frink, NASA Langley Research Center, Virginia;
Mark S. Chaffin, Cessna Aircraft Company

Space Science: Solar Simulations in the Zettaflop Era

- Solar convection zone simulations could be expanded to include multiple super-granules with a 2-4x increase in computer power. This would allow a highly credible analysis of the physics of large-scale photospheric phenomena.
- Another 2-4x would allow simulation of the largest photospheric scales, the giant cells.
- Zettaflop performance would allow a simulation of the full convection zone, from 70% of the solar radius out into the atmosphere, at a horizontal resolution sufficient to resolve granules. This would include all important scales of motion and so give a complete picture of internal solar dynamics. A very thorough understanding of solar activity and space weather generation would then follow.



Current solar convection zone simulations are limited to boxes of approximately 10% of the solar radius on a side. These require roughly 200,000 processor hours on Columbia.

POC: Alan Wray, NASA Ames Research Center

So Where Are We?

- **The Science**
  - Production CFD codes executing 100x C90 numbers of just a few years ago.
  - Throughput 100x (or more) above that of a few years ago.
  - Earth/Space Science codes executing 2-4x faster than last year's best efforts, 100x throughput over last year's efforts.

- **The Systems (1997 - present)**
  - New expanded shared memory architectures: First 256, 512, and 1024 CPU Origin systems. First 256p, 512p Altix SSI systems.
  - First 2048p NUMAlinked 512p Altix cluster.

- **The Future?**
  - Expanded Altix SSI to 4096?
  - Expanded Altix NUMAlinked clusters to16Kp?
  - Serious upgrades to CPUs

Conclusion: Advanced Development Concepts

- Several orders of magnitude increase in effective computational power needed to radically extend the range of design options to be explored or radically shorten the design cycle
- Computer technology of massively parallel processing combined with single processor speed increases will support the above
- Computing methods and new architectures are needed to match over a spectrum of applications
- New paradigms are needed to harness a very large number of processors
- Need to provide advanced development tools, processes and products to increase design confidence, and reduce the design cycle time for aircraft and space vehicles by 50% in 10 years and 75% in 25 years
- Currently, answers to “what if” questions require hours, days, even months. To support designer’s train of thought, these answers should be coming in seconds
- Progress in computer technology will be achieved by two ingredients: faster processors, and more of them - yet needs to maintain a single virtual computer appearance to the user

POC: Jaroslaw Sobieski (LaRC), *Ultrafast Computing Team Report, Feb. 1999*

Consequences of Architecture Diversity

In the old days, single processor speed increases made our codes run faster; simple and easy.

- **Now**, there are a multitude of processors and memory architectures available, in a single or virtual computer. It is unlikely that smart operating systems will completely mask the architectural diversity
  - New task: tailor solution to architecture
  - New opportunity: specify architecture that suits a class of applications
- We need many processors, do we know how to use them?
  - Current experience shows diminishing returns setting in when the number of processors in 100's is reached
- **Why: Types of Parallelism**
  - Coarse-grained: replicated code, different inputs (problem-dependent)
  - Coarse-grained: partitioned domain (diminishing returns)
  - Fine-grained: existing code rearranged (machine-dependent, almost useless)
  - Fine-grained: existing solution algorithm recoded (machine-dependent, limited usefulness)
  - Radical, new paradigms to be invented
- **New paradigms are needed to exploit more than 100's processors**

POC: Jaroslaw Sobieski (LaRC), *Ultrafast Computing Team Report, Feb. 1999*

How to get engineering computing to ride the wave of the future in computer technology

- The engineering computing market is small relative to that in business and entertainment. Therefore, it constitutes a niche where the Government seed money might make a real difference.
- In the interdisciplinary arena, one should continue to
  - monitor, understand the new computer hardware and software technologies and architectures
  - develop an understanding of the capabilities that are likely to be delivered by the commercial development regardless of the Government actions
  - Influence development of the new computer hardware and software technologies and architectures
  - Develop understanding of the match between various types of engineering computing jobs and various computer architectures, and the match frequencies
  - Formulate the need for new developments at the integrating framework level and at the disciplinary leveln particular discipline
  - Formulate standards and requirements as needed by the tool integration, MDO environment, and the new architectures
  - Develop methods for effective utilization of the system analysis and MDO for various classes of the new architectures, taking into consideration the computing load balancing among the processors
  - Recommend long term investment strategy based on the above information
  - Foster and coordinate disciplinary developments and application projects
  - Facilitate education and training 2)
- In each disciplinary domain, one will need to
  - Commit to gearing-up to the exploitation of new computer architectures in hardware and software.
  - Reexamine and restructure the disciplinary algorithms, and to develop new paradigms where needed, accounting fully for MDO
  - formulate local disciplinary standards and requirements compatible with the ones established in the interdisciplinary arena
  - develop and validate the restructured algorithms and the new paradigms, implementing the standards and requirements

POC: Jaroslaw Sobieski (LaRC), *Ultrafast Computing Team Report, Feb. 1999*

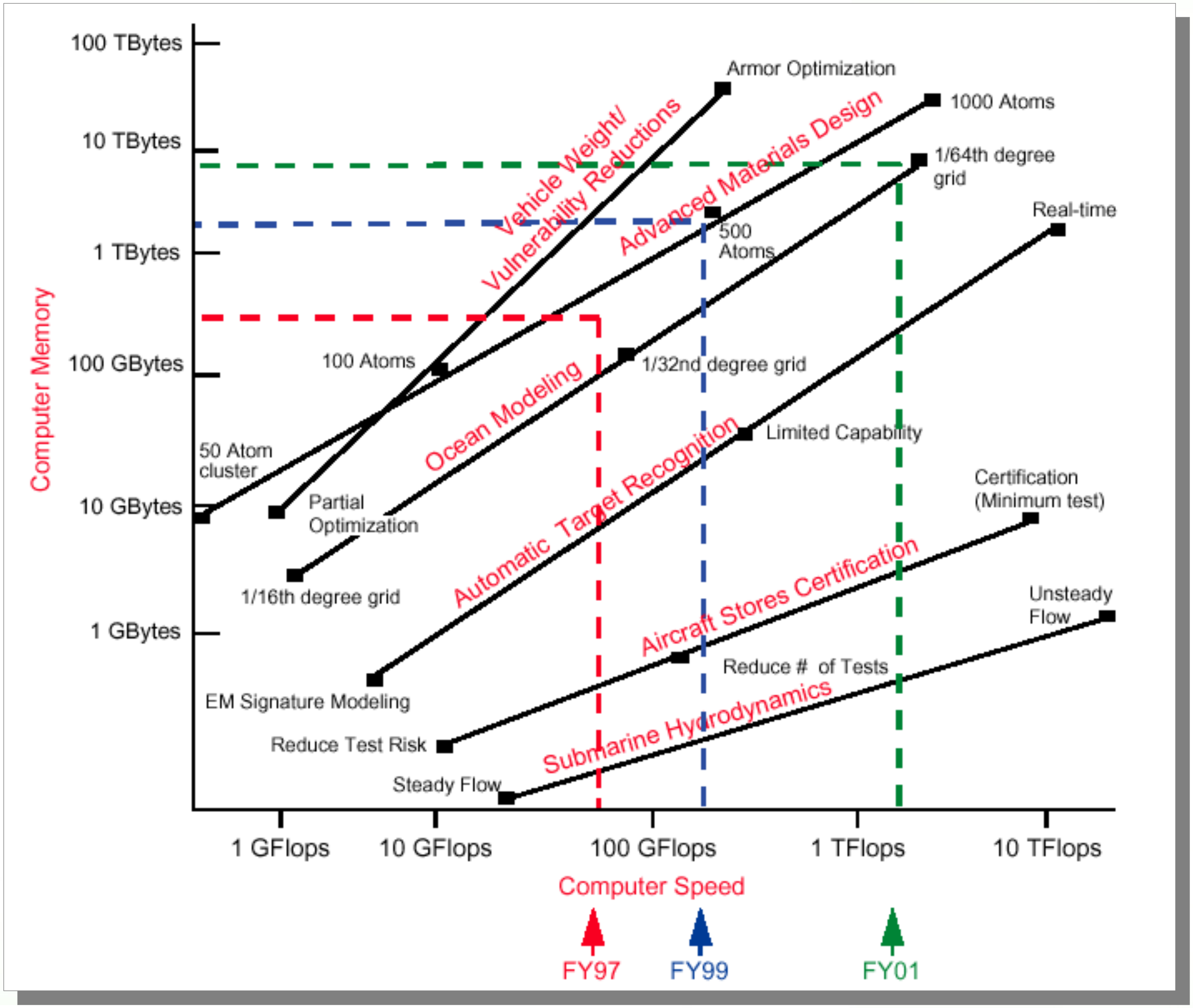
“Compute as Fast as the Engineers can Think!”

- The charter for the *Ultrafast Computing Team Report* (Feb. 1999) was to examine impact of new computer architectures on computing in the engineering design process because:
  – The aerospace vehicle design process is too long; not computing fast enough is a major culprit
  – Computer technology offers new opportunities in massively heterogeneous and concurrent processing that should be exploited.

- **Examining two user scenarios: RLV and HSCT, it was determined that:**
  – Major computing tasks need to be reduced from hours to seconds
  – Effective computing speed need to increase by several orders of magnitude to achieve that
  – Computer technology of massively parallel processing must combine with new methods to achieve that goal
  – There is usually one week for the partnership to determine which proposed configuration to pursue.
  – The objective is to maximize the return on investment over the life of the vehicle, including the assumptions of 10 years and 36 launches per year.

POC: Jaroslaw Sobieski (LaRC), *Ultrafast Computing Team Report, Feb. 1999*

Changing Engineering Paradigms: Moving from Capability to Capacity Systems



POC: Jeffrey Mohr, Computer Sciences Corp., 1999