
NASA ADVANCED SUPERCOMPUTING • NASA ADVANCED SUPERCOMPUTING • NASA ADVANCED SUPERCOMPUTING
PROJECT COLUMBIA
SUPPORTING NASA'S MISSION DIRECTORATES
AERONAUTICS RESEARCH MISSION DIRECTORATE
SCIENCE MISSION DIRECTORATE
SPACE OPERATIONS MISSION DIRECTORATE
EXPLORATION SYSTEMS MISSION DIRECTORATE
intel
NASA
sgi
NASA ADVANCED SUPERCOMPUTING • NASA ADVANCED SUPERCOMPUTING • NASA ADVANCED SUPERCOMPUTING

"Project Columbia Development and Impact"

Salishan Conference April 19, 2005

Dr. Walter F. Brooks
NASA Ames Research Center
Moffett Field, CA USA

Outline

- Columbia Project Context
- Architecture and Development of the Columbia System
- Impact on Engineering
  - Return to Flight
  - Exploration
- Impact on Science
  - Space science
  - Earth Science

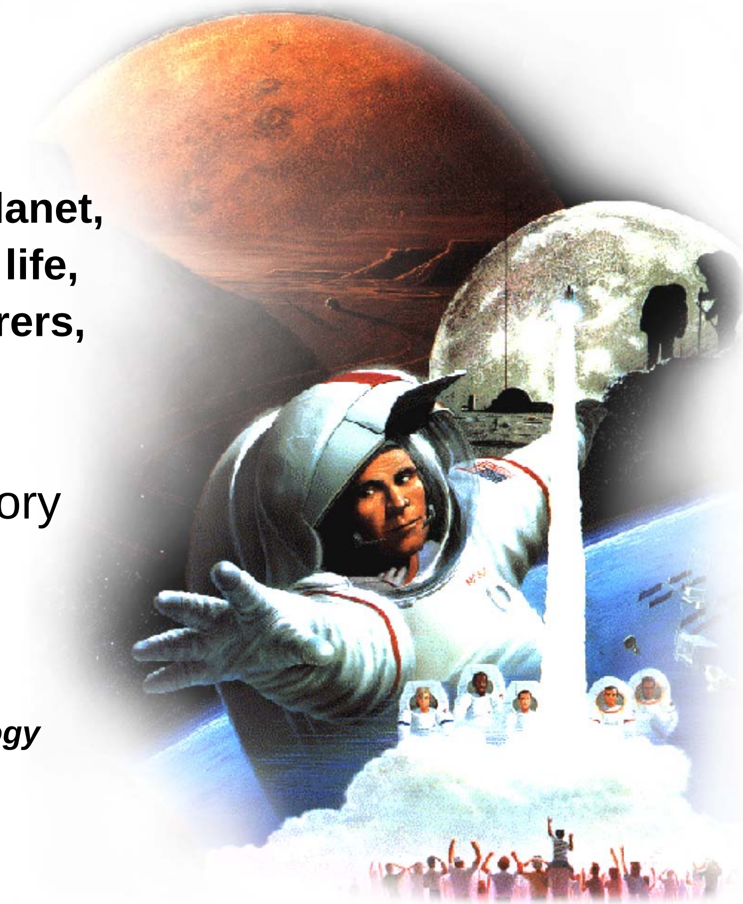
NASA Vision and Mission

- Vision:
  - **To improve life here,**
    To extend life to there,
    To find life beyond.

- Mission:
  - **To understand and protect our home planet,**
    To explore the universe and search for life,
    To inspire the next generation of explorers,
    … as only NASA can.

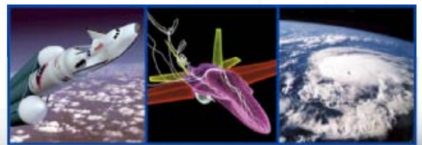
- President's Information Technology Advisory Committee quote:

  "Information Technology will be one of the key factors driving progress in the 21st century - it will transform the way we live, learn, work, and play.
  Advances in computing and communications technology will create a new infrastructure for business, scientific research, and social interaction."

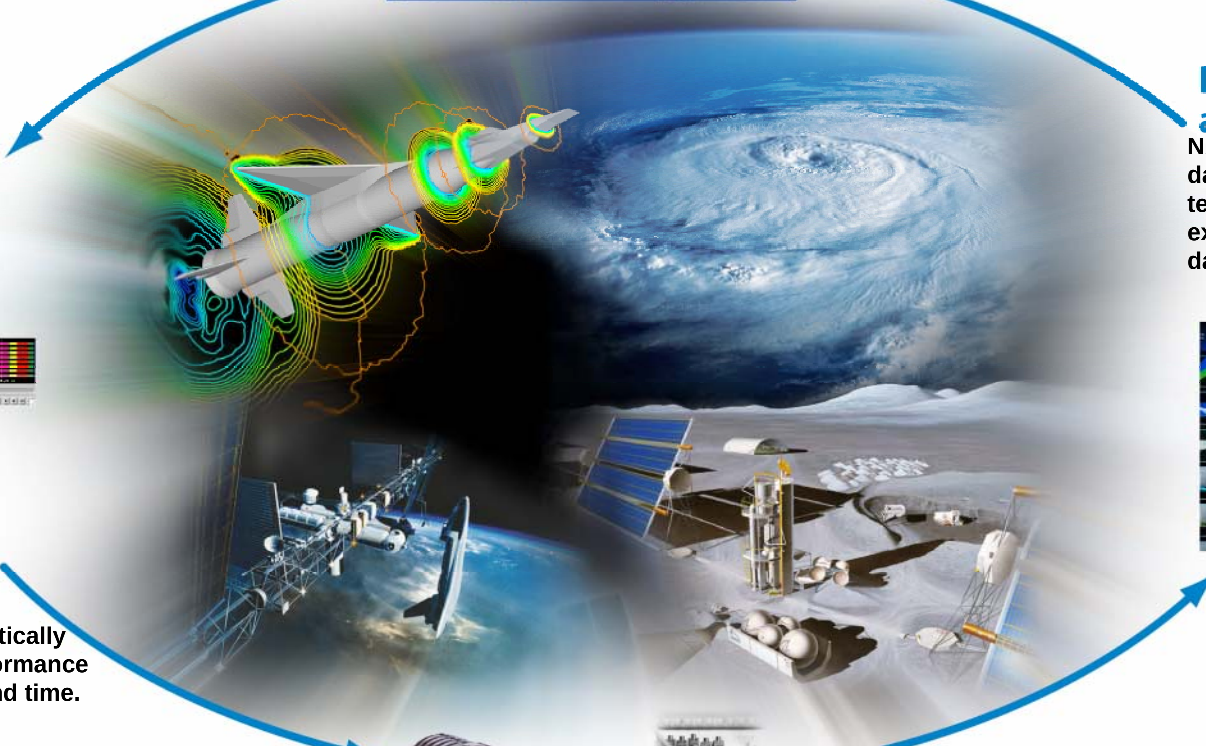


Integrated Support for High Performance Modeling and Simulation

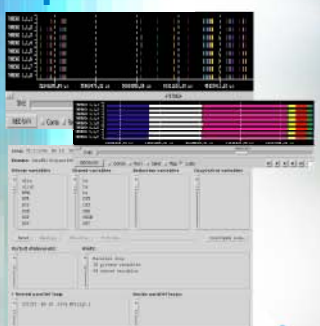
NASA Scientists and Engineers



Scientists and engineers set up computational problems, choosing effective codes and resources to solve NASA's complex mission problems.



Performance Optimization



NAS software experts exploit tools to parallelize and optimize codes, dramatically increasing simulation performance while decreasing turnaround time.

Data Analysis and Visualization

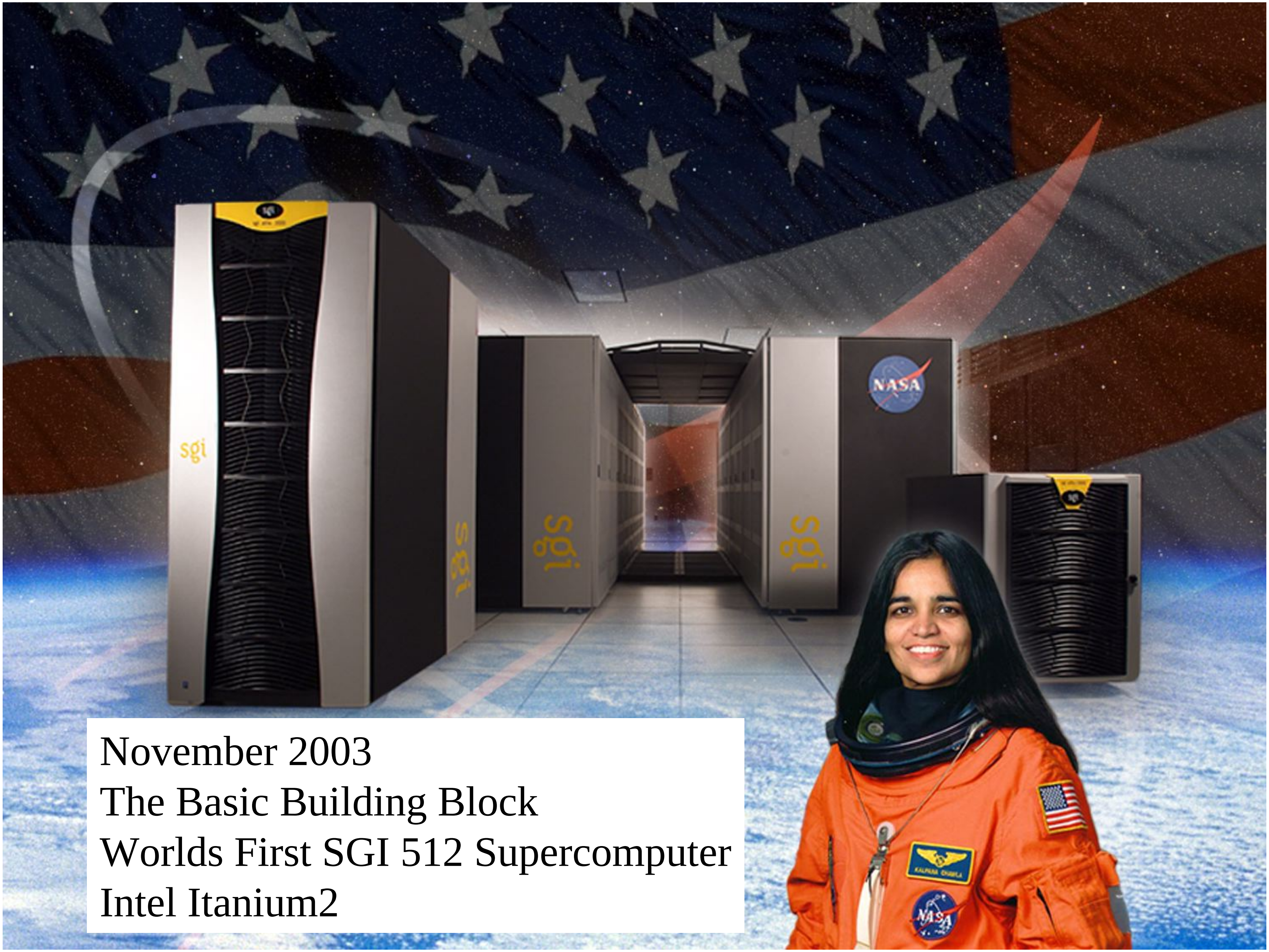
NAS experts apply advanced data analysis and visualization techniques to help scientists explore and understand large data sets.

Supercomputers, Storage and Networks

The NAS supercomputer environment (hardware, software, network and storage) is used to execute the optimized code to solve NASA's large computational problems.

60 TERAFLOPS in 120 DAYS

- Program
  - Beginning May 18,2004, obtain all of the necessary approvals and procure the system by June 18,2004
- Physical Plant
  - Make all of the necessary power and cooling changes to run Columbia
  - Reconfigure and retrofit decommissioned water cooling loops
- Production
  - SGI Build and deliver 19 Altix 512's in less then 4.5 months including first Altix 3700BX2
- Integration
  - Assemble, Test and 20x512p with GigE and Infiniband connectivity
- Continuous production
  - Continue NASA science and engineering in support of NASA Missions
- Provide a national capability
  - Build and utilize the 1st shared-memory 2048

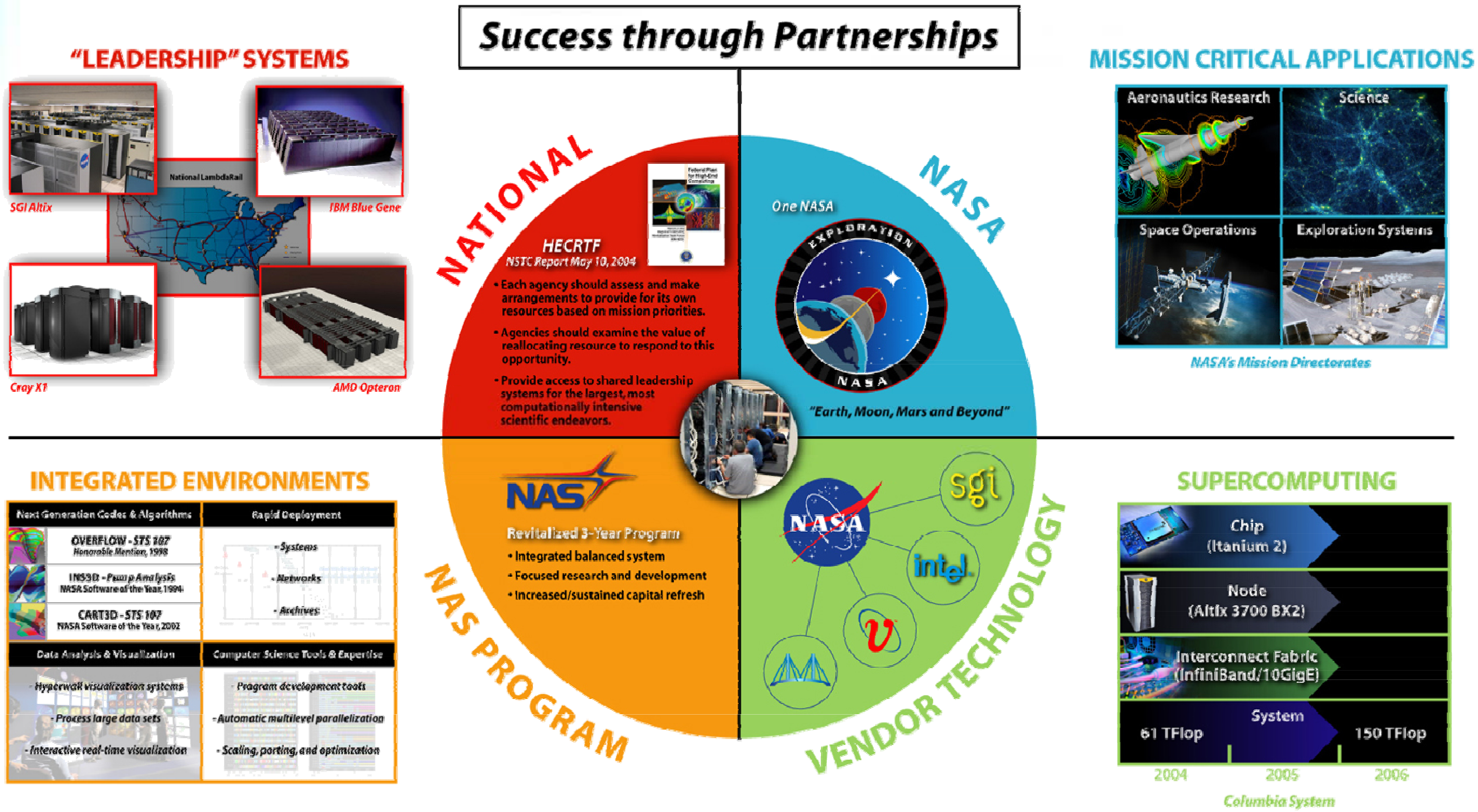


November 2003
The Basic Building Block
Worlds First SGI 512 Supercomputer
Intel Itanium2

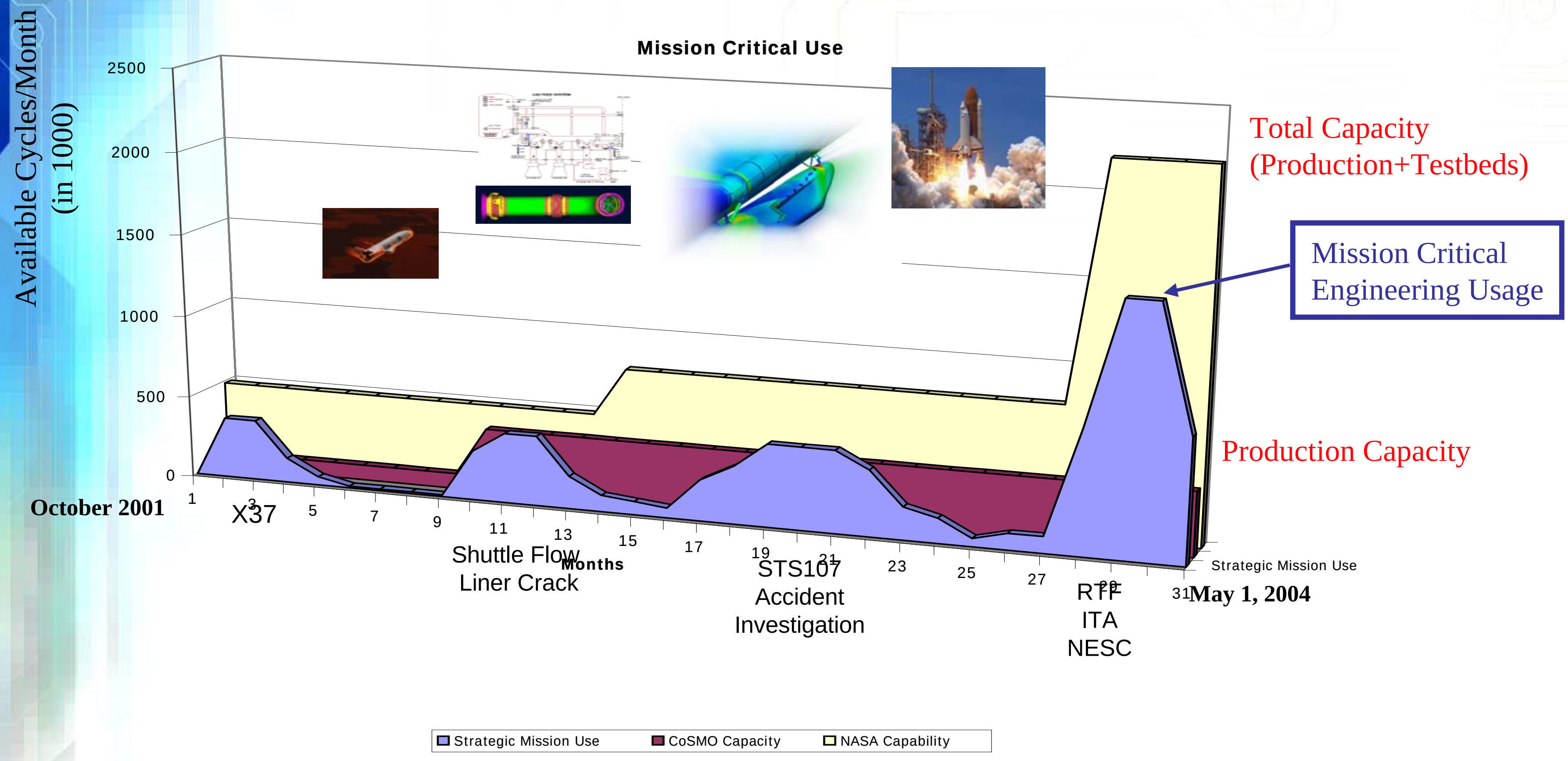
QuickTime™ and a
Animation decompressor
are needed to see this picture.

NASA Advanced Supercomputing Mission

An integrated modeling and simulation environment enabling NASA and its industry and academic partners to accelerate design cycle time, conduct extensive parameter studies of multiple mission scenarios, and increase safety during the entire life cycle of exploration missions, while satisfying the tight time constraints of fast-paced NASA exploration system design and acquisition cycles.



What is the demand on the current system based on the past and present users needs/pattern?



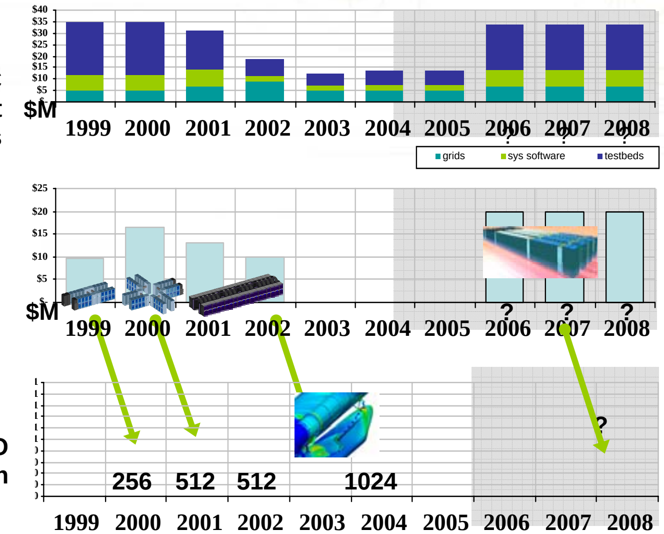
Over the last two and a half years, NASA's HEC requirements for mission critical engineering have been time-critical, cyclical, and growing

NASA’s Modeling and Simulation Capability 2004

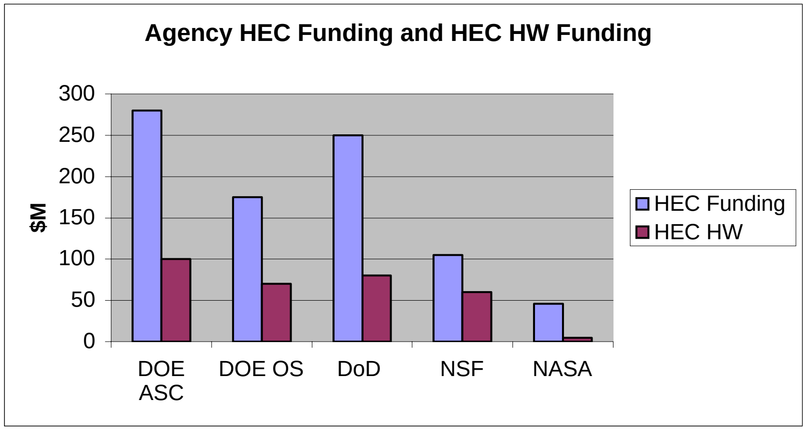
- NASA Status
  - Demise of HPCC program and other program shifts have resulted in major decrease in HEC investments in NASA since 2000
  - Without significant reinvestment in 04 and 05 “bridging” we will create a major gap in HEC computing
  - Need strategic investment in 05 to reinvigorate modeling and simulation in the agency
  - We still have a major capability based on smart investments and one time only buy Altix Kalpan

NASA HEC Budget Sources

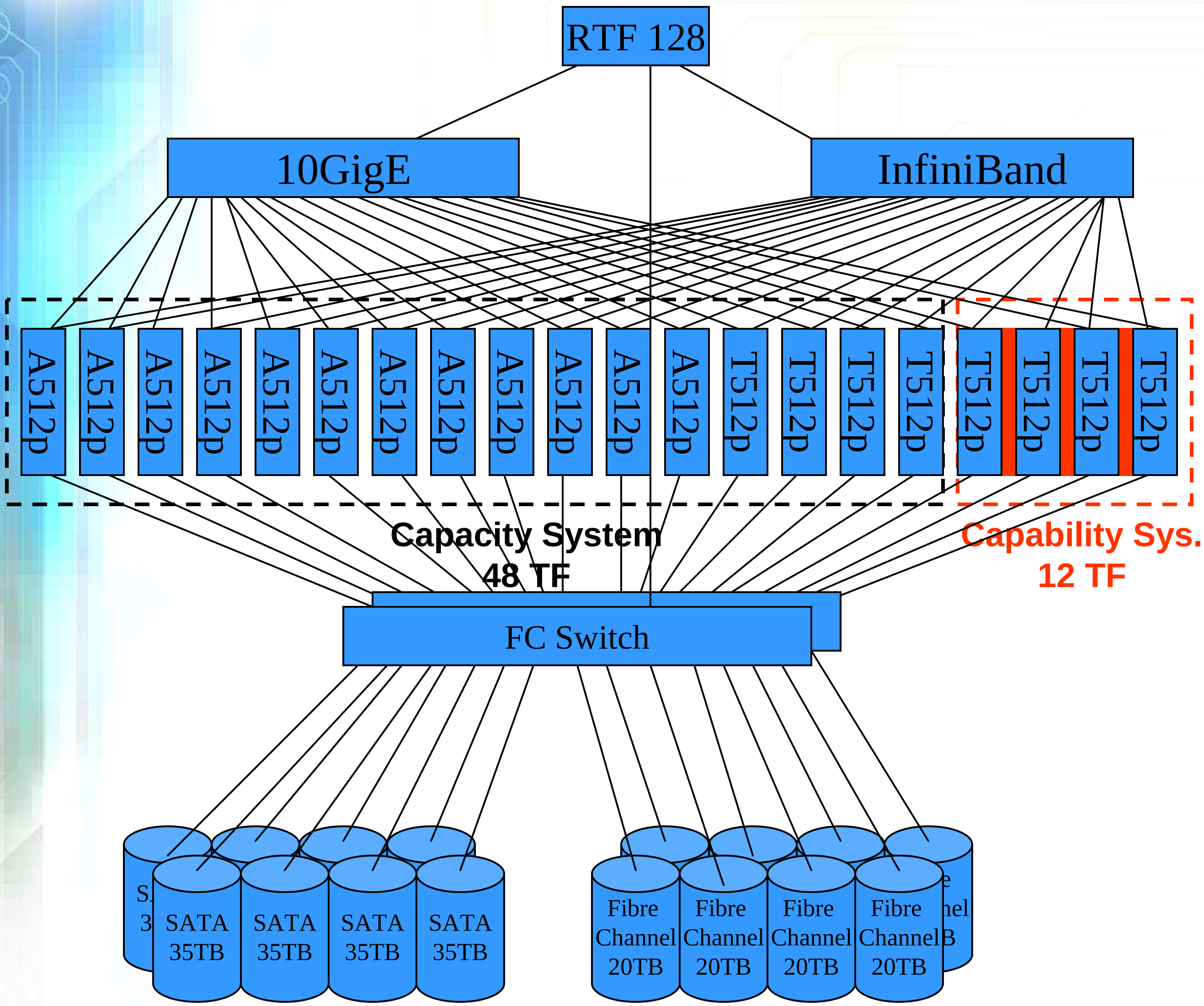
CoSMO Production



•Agency*	•FY03	•FY04	•FY05
•Agency 1	•3.24%	•4.09%	•4.23%
•Agency 2	•2.06%	•2.05%	•2.21%
•Agency 3	•1.72%	•1.72%	•Not Revealed
•Agency 4	•0.53%	•Not Revealed	•Not Revealed
•NASA	•0.21%	•0.21%	•0.19%
•Agency 5	•0.04%	•0.04%	•0.04%



Columbia Configuration



Front End
- 128p Altix 3700 (RTF)

Networking
- 10GigE Switch 32-port
-10GigE Cards (1 Per 512p)
- InfiniBand Switch (288port)
- InfiniBand Cards (6 per 512p)
- Altix 3700 2BX 2048 Numalink Kits

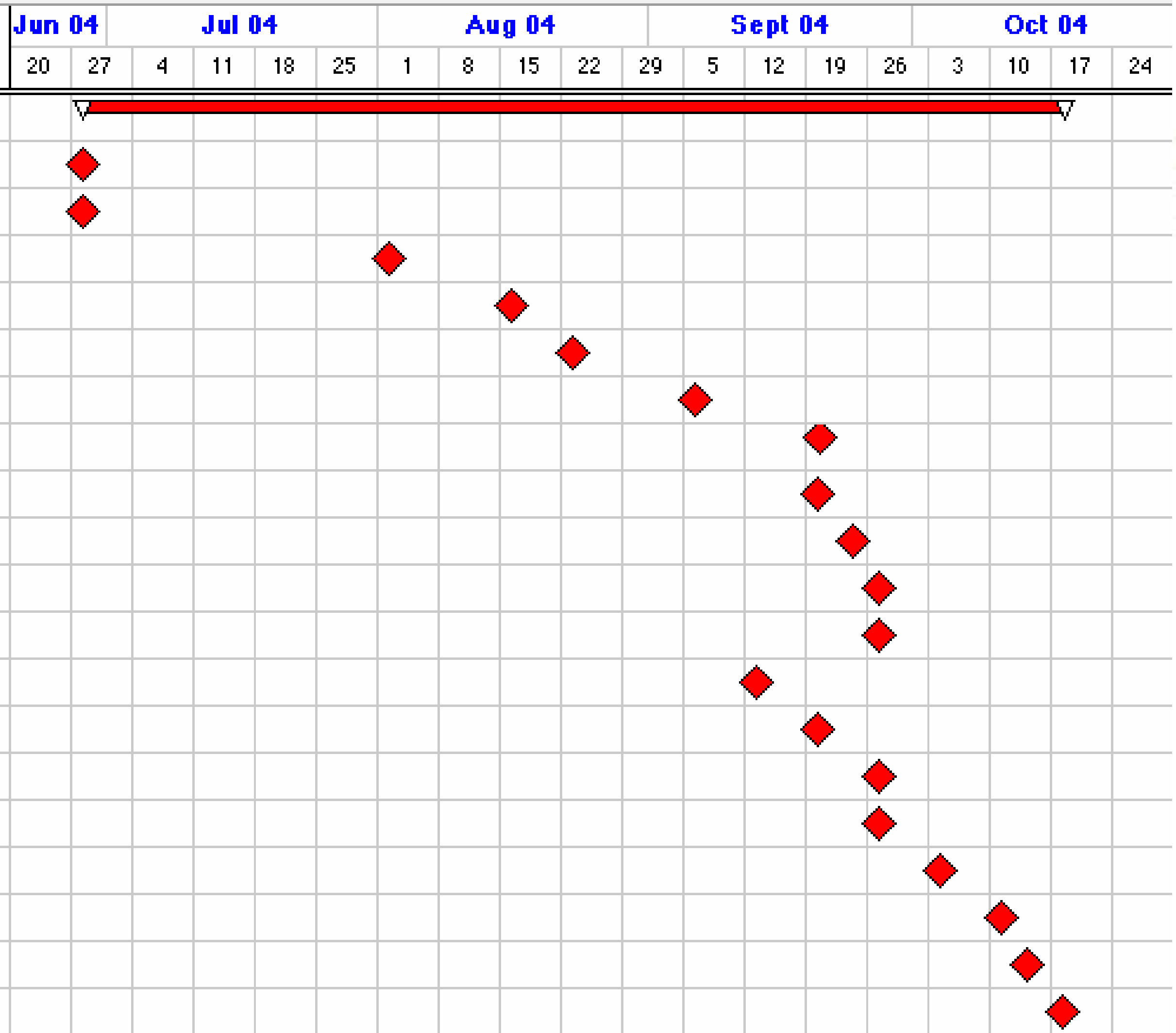
Compute Node (Single Sys Image)
- Altix 3700 (A) 12x512p
- Altix 3700 BX2 (T) 8x512p

Storage Area Network
-Brocade Switch 2x128port

Storage (440 TB)
-FC RAID 8x20 TB (8 Racks)
-SATARAID 8x35TB (8 Racks)

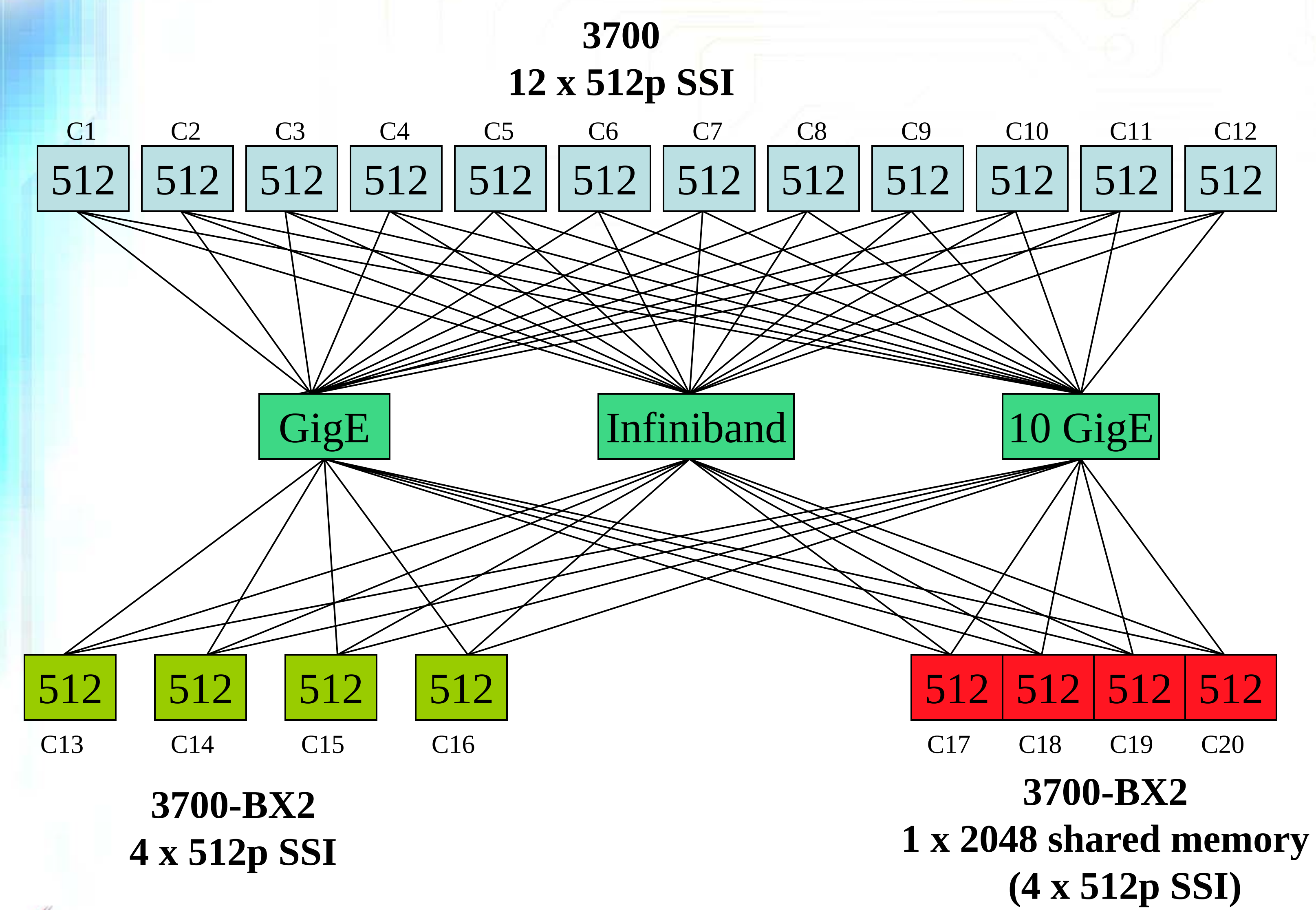
Delivery Schedule

Activity Name	Start Date	Finish Date
System Arrival	6/28/04	10/18/04
3700-2	6/28/04	
3700-3	6/28/04	
3700-4	8/2/04	
3700-5	8/16/04	
3700-6	8/23/04	
3700-7	9/6/04	
3700-8	9/20/04	
3700-9	9/20/04	
3700-10	9/24/04	
3700-11	9/27/04	
3700-12	9/27/04	
3700BX2-1	9/13/04	
3700BX2-2	9/20/04	
3700BX2-3	9/27/04	
3700BX2-4	9/27/04	
3700BX2-5	10/4/04	
3700BX2-6	10/11/04	
3700BX2-7	10/14/04	
3700BX2-8	10/18/04	



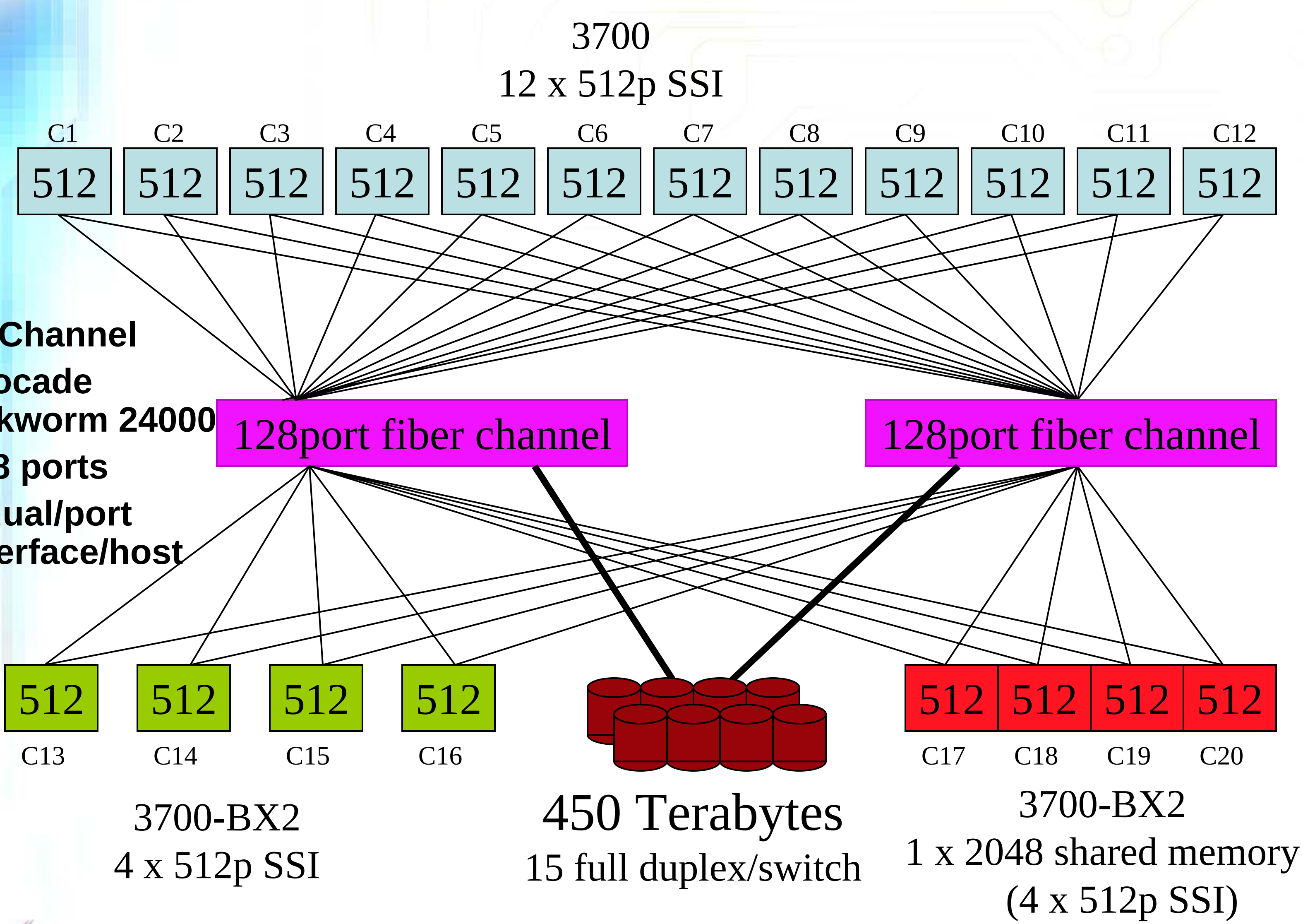
Project Columbia Dedicated October 2004

Configuration Details - Compute Nodes

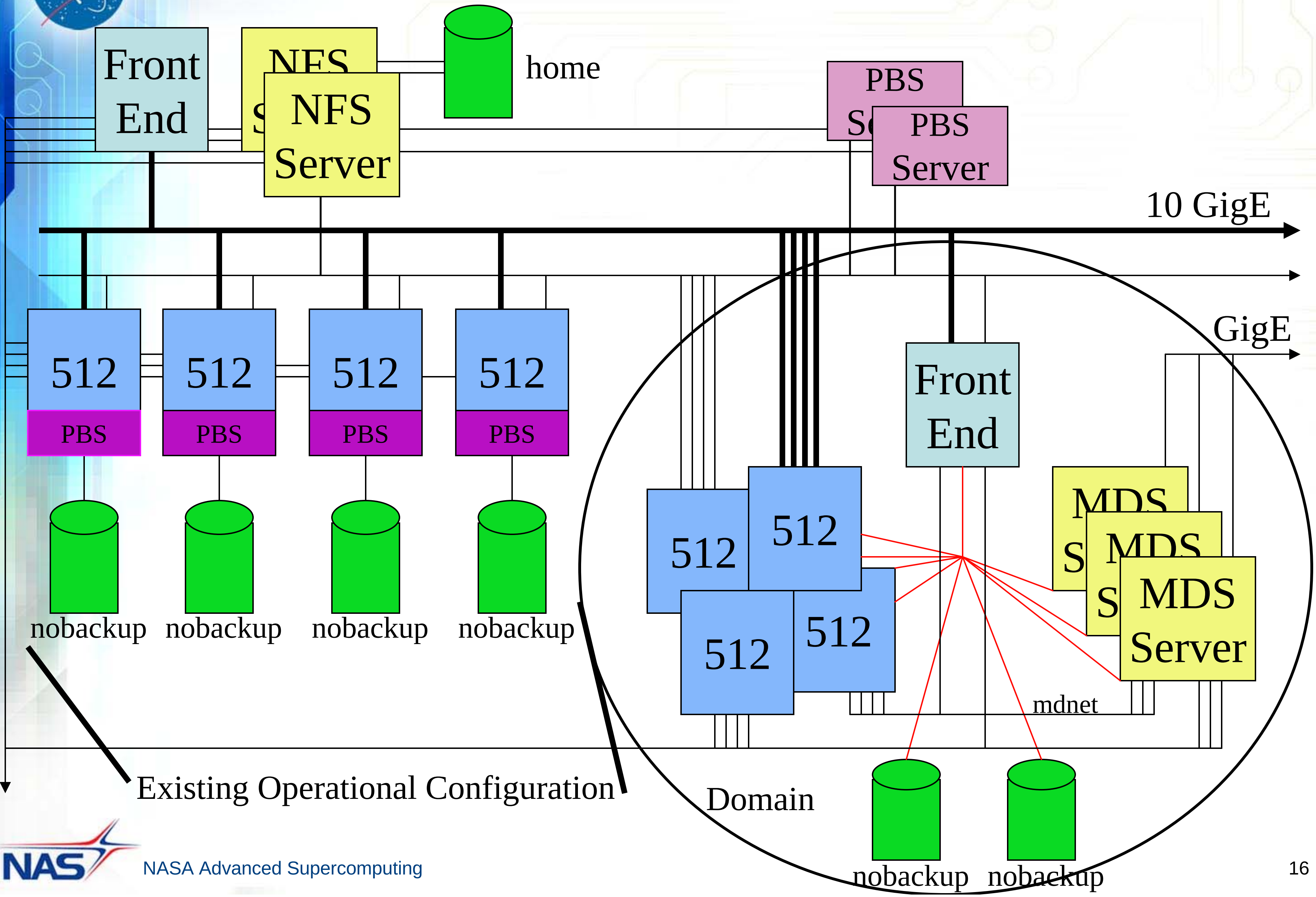


Configuration Details - SAN Fabric

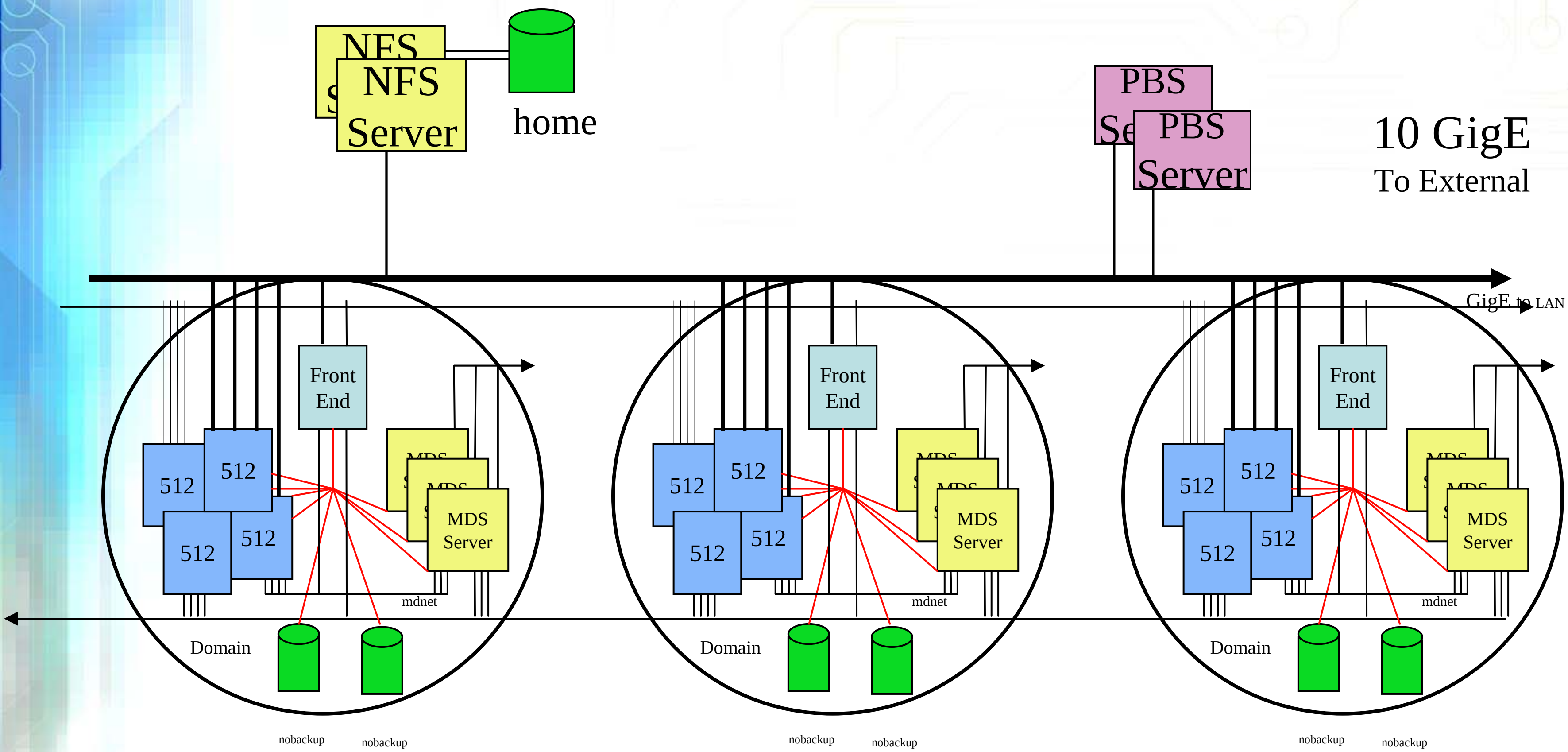
Fiber Channel
- **Brocade Silkworm 24000**
- **128 ports**
- **4 dual/port interface/host**



Architecture Target - Phase II



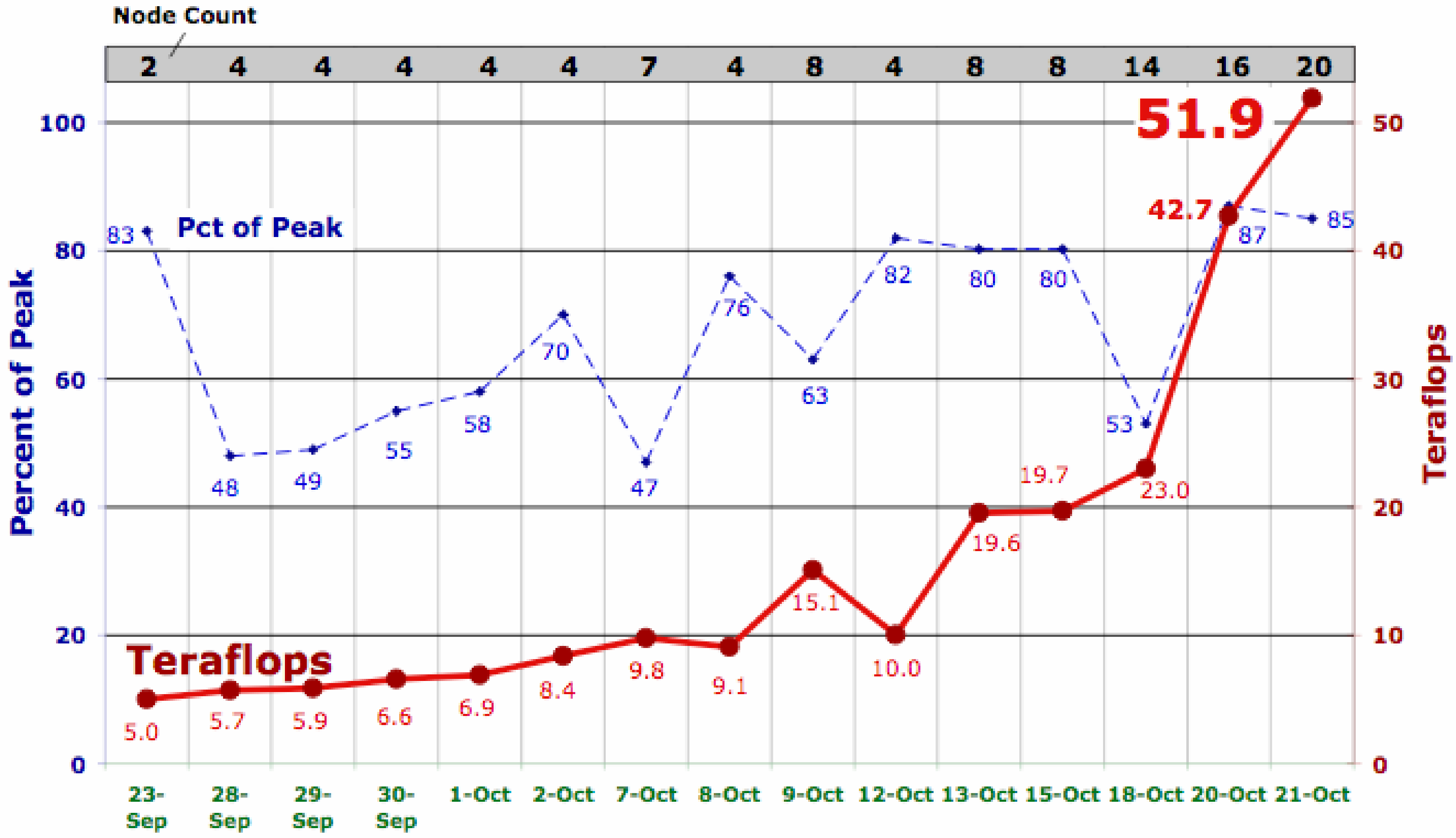
File SystemTarget - Phase III



Domain counts and sizes will change based on experience with cxfs and mission requirements. Most likely config is 2-4 domains

(Job scheduling is independent of filesystem domain – i.e. span domain)

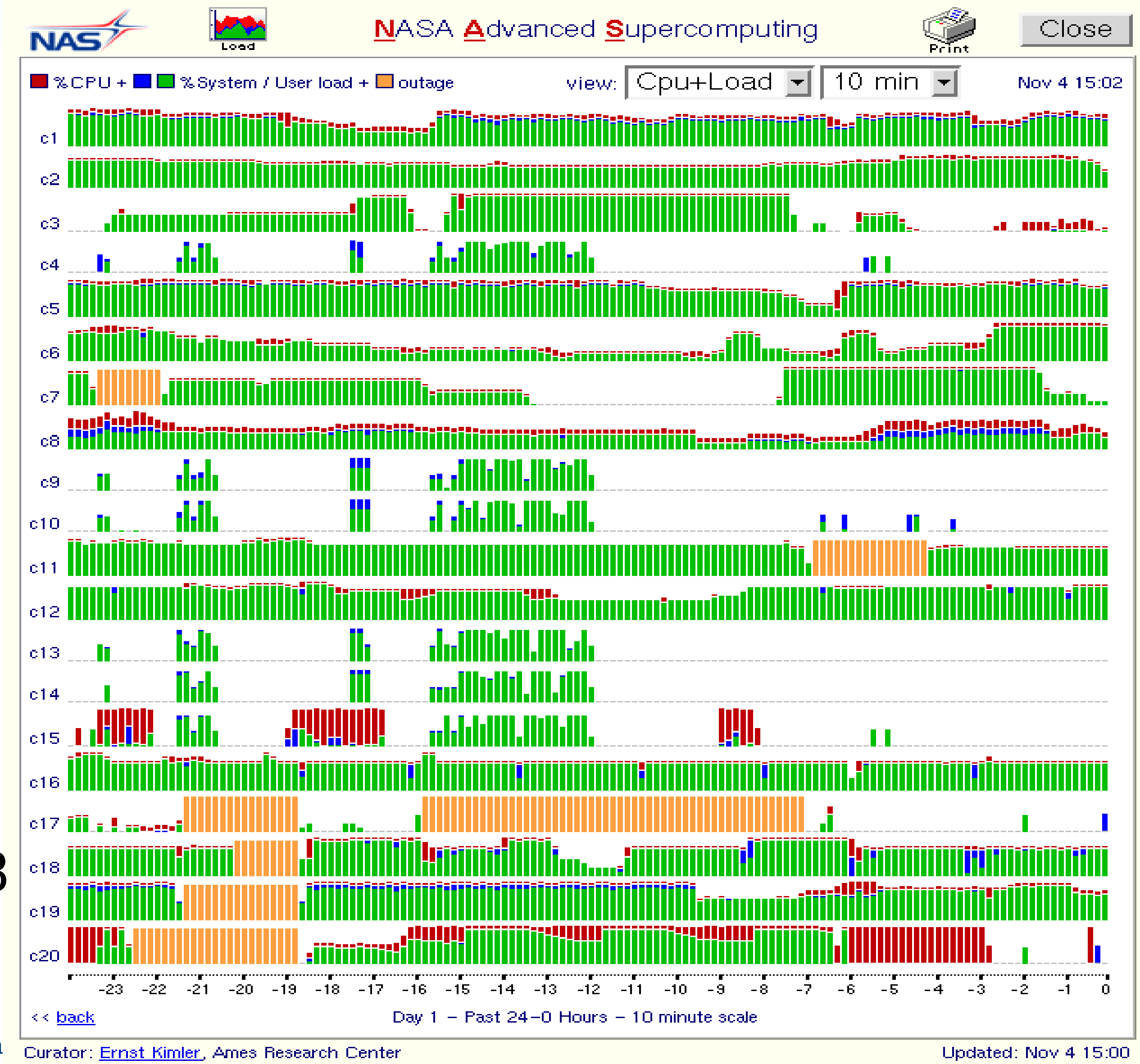
Linpack Results



High Utilization -600 Users-Major Impact

20 Nodes

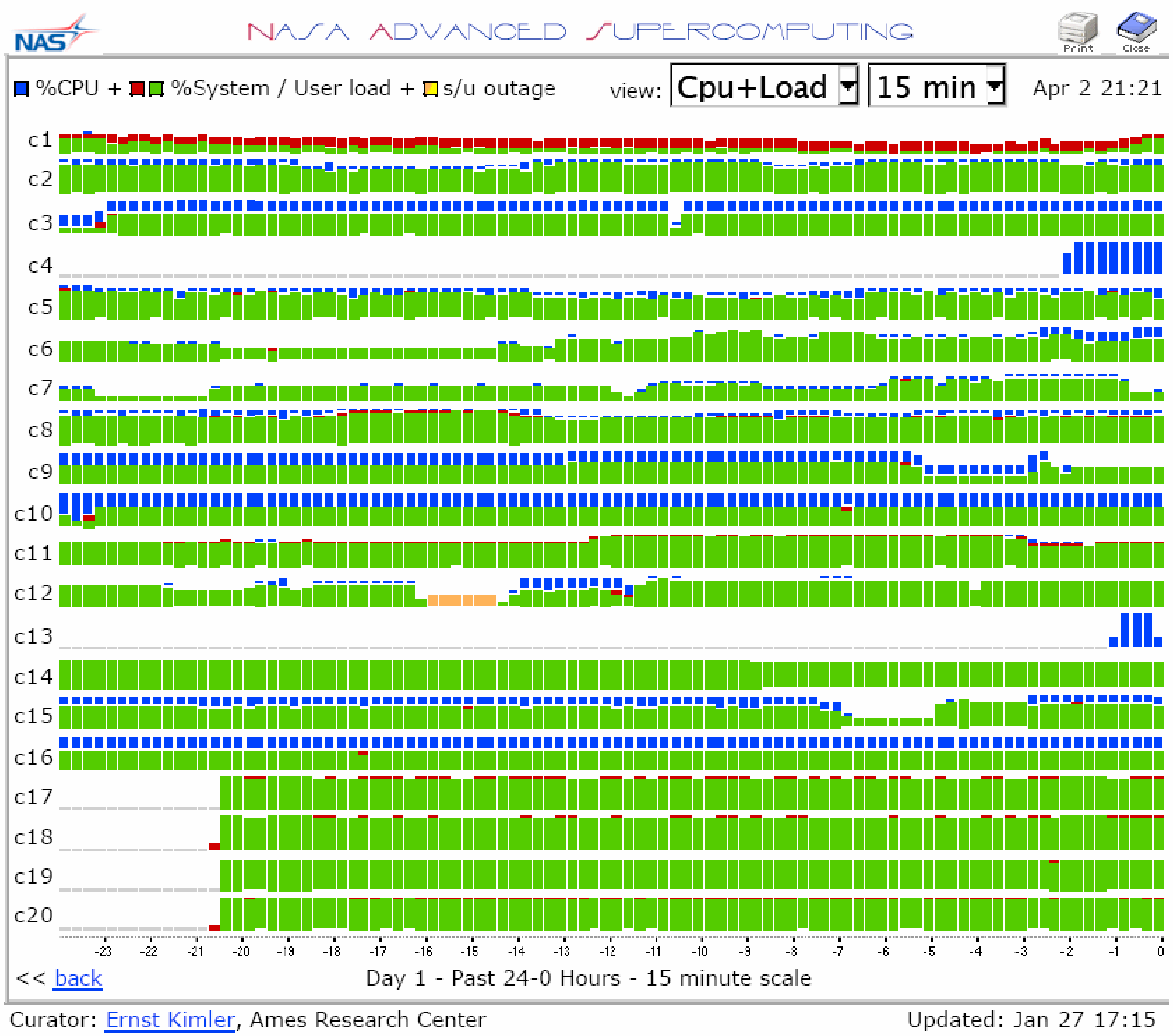
2048



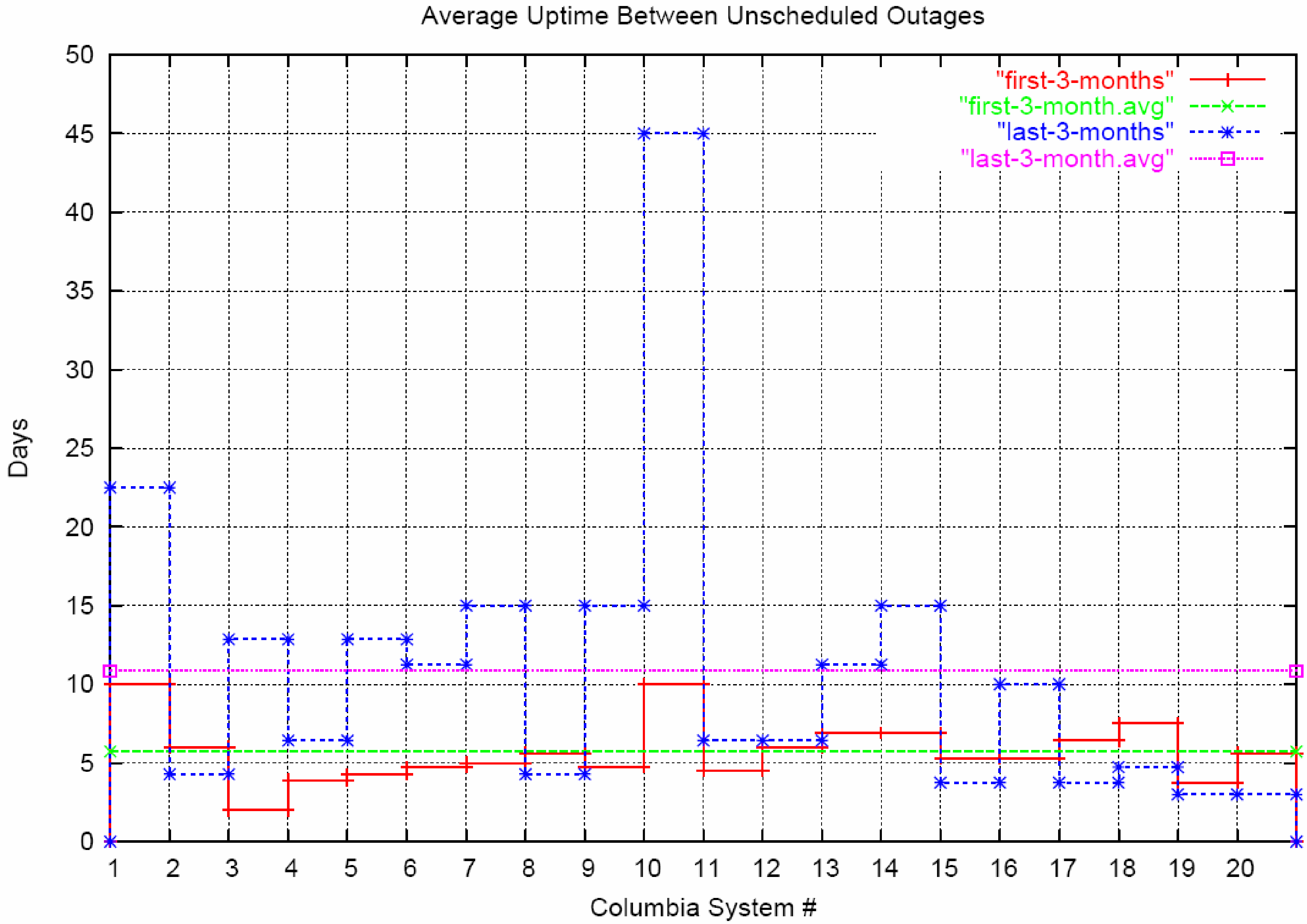
Fastest Production Supercomputer

20 Nodes

2048

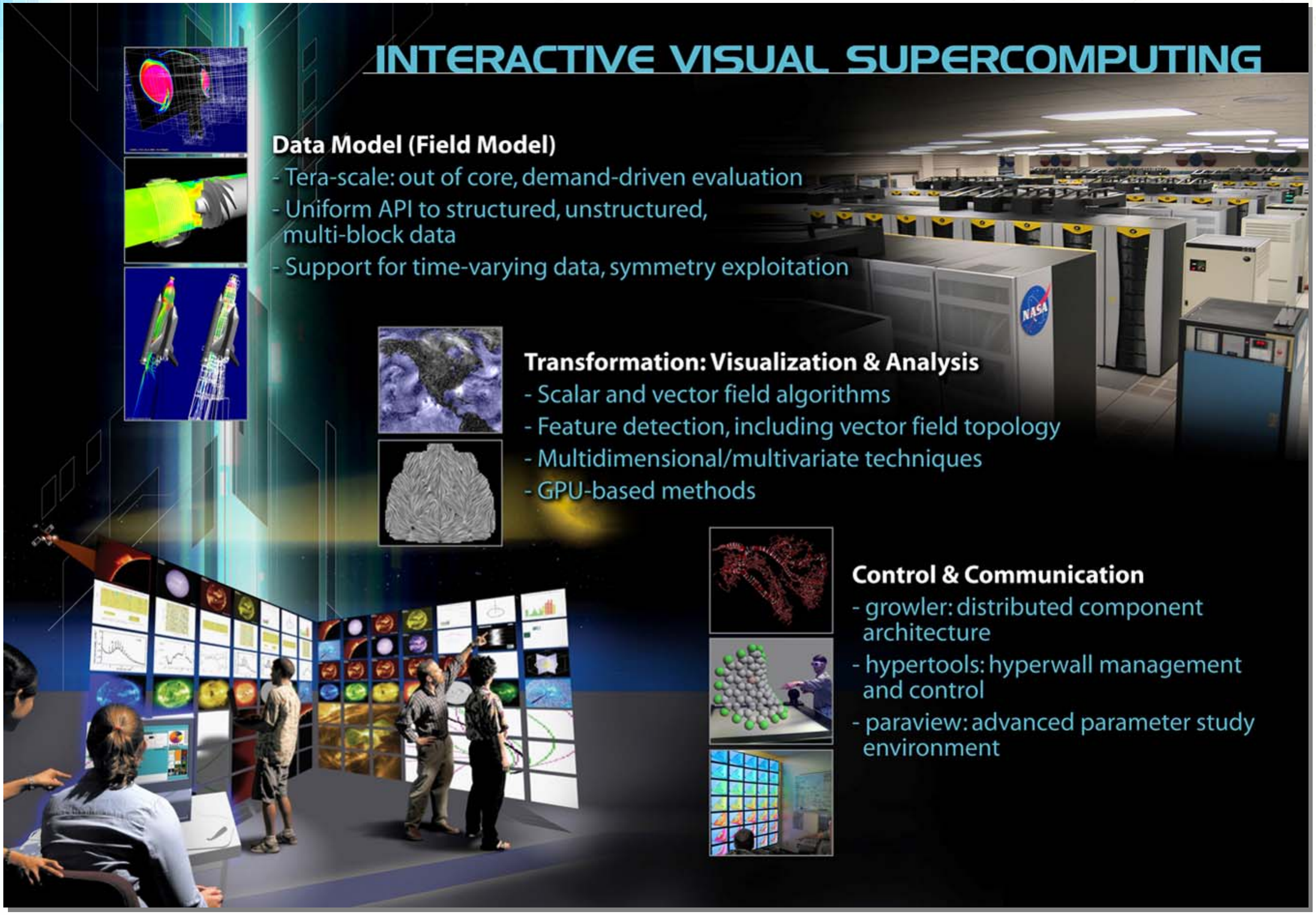


Average uptime between Scheduled Outages



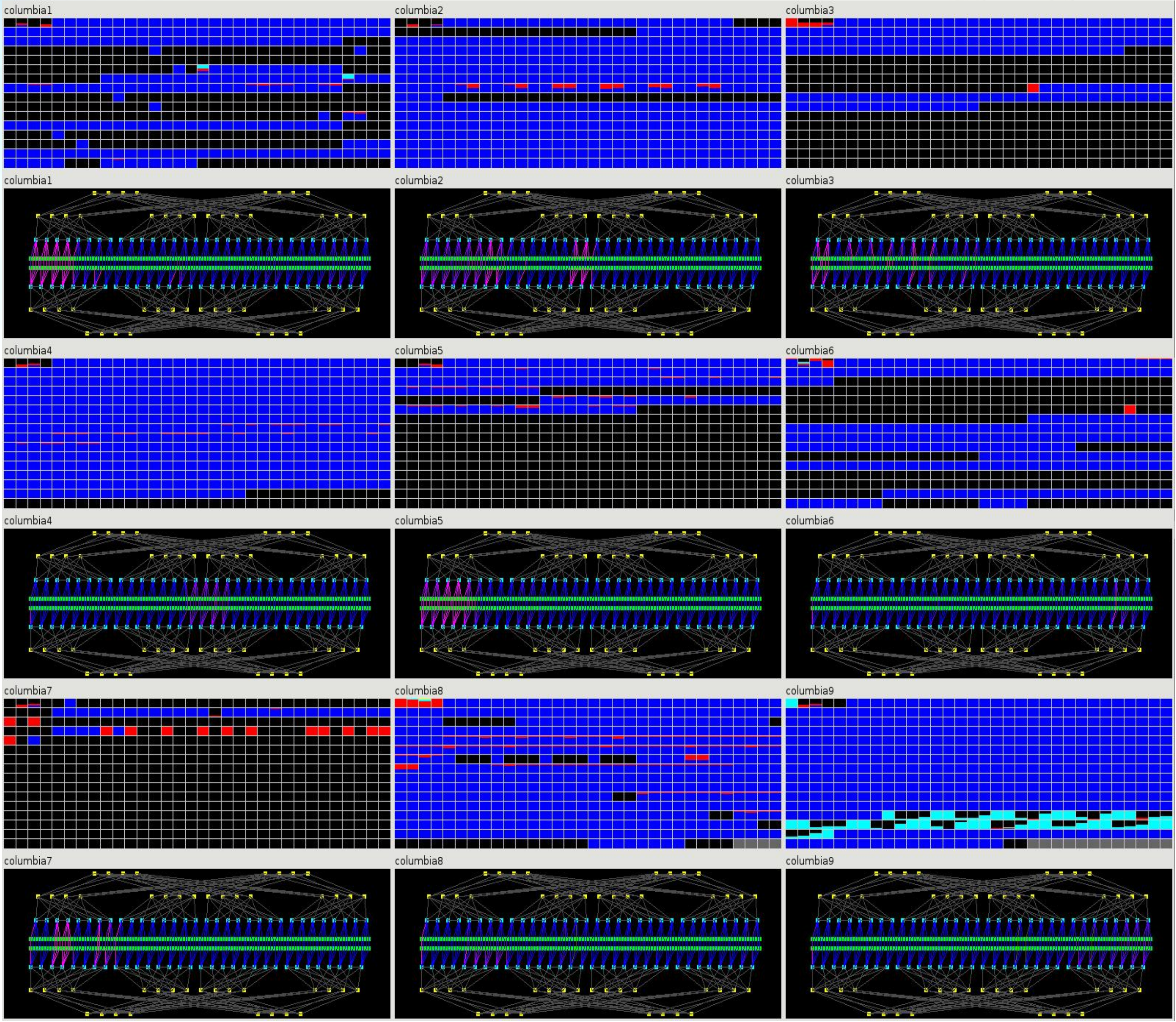
Networking: NASA Research and Engineering Network (NREN) - Planned Upgrade

Visualization: Interactive Visual Supercomputing



Hyperwall-1: parameter studies

System and Application Performance Monitoring



Impact of Columbia After 150 days

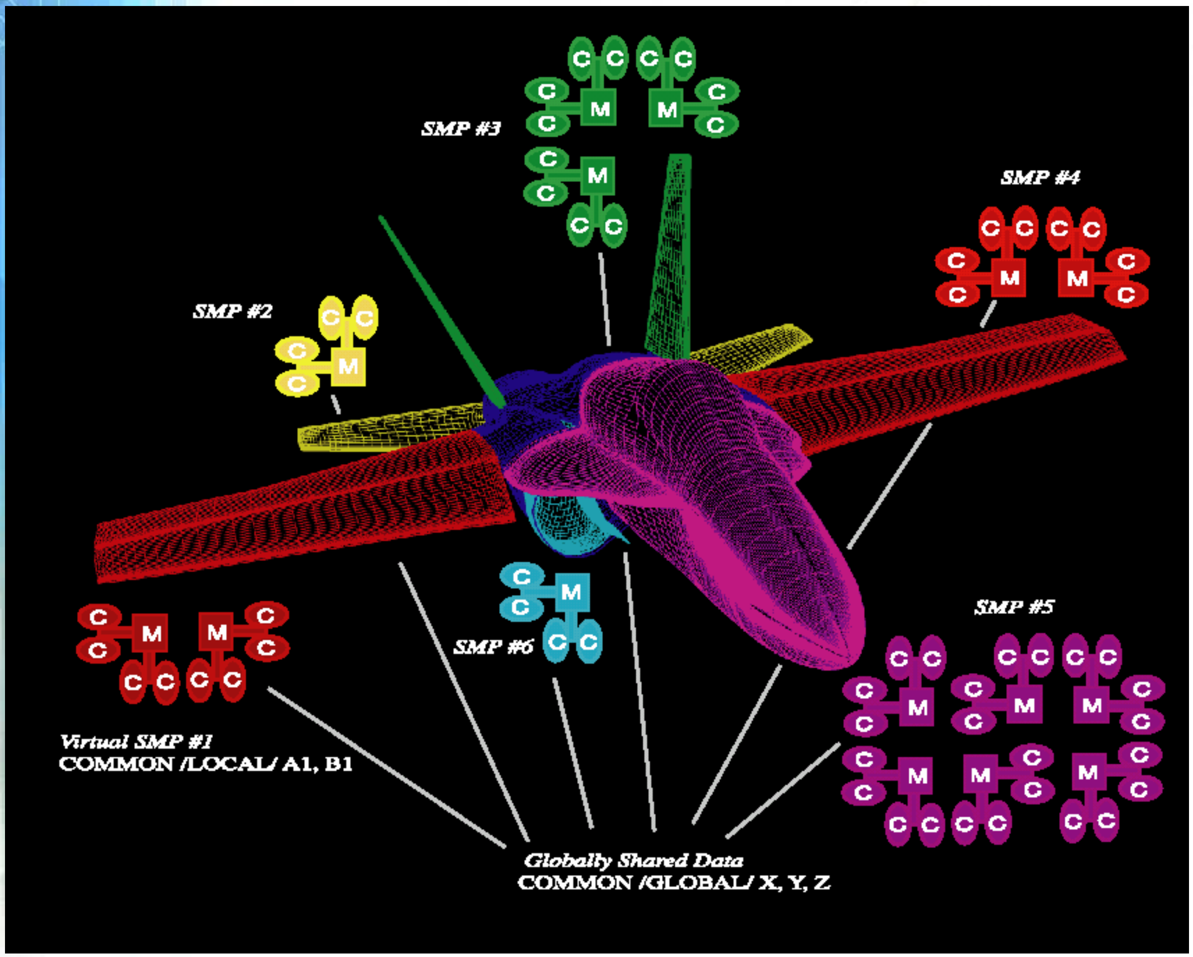
- Engineering
  - Emergency Response
  - Return To Flight
  - NASA Engineering and Safety Center (NESC)
  - Digital Astronaut
  - Genomics/Nanotechnology

Classes of Computation

- Breakthrough Investigations
  - Hurricane Forcast, Ocean Modeling
- Baseline Computational Workload
  - Existing Engineering/Science Workloads
- Emerging Workloads (near term)
  - Return to Flight
  - NASA Engineering Safety Center
- Emergency Response
  - Periodic requirement for mission critical analysis work
  - STS107, STS fuel line, X37 heating

Why Single System Image?

Mapping problems onto a Shared Memory System is conceptually simple



– Shared Memory Enhances Development
  - *Less code*
  - *Less development time*
  - *Less debug time*
  - *More algorithmic flexibility*
  - *Ability to retain existing features/algorithms*

– All memory references utilize microprocessor optimizations
  - caching
  - out of order execution
  - pre-fetch

– Higher levels of parallel efficiency

– Fast Local Filesystems

Interconnect latency <1us worst case with a TLB miss

CART3D OpenMP/MPI Scaling

(results for 4.6 and 6.4 million cell shuttle configuration)

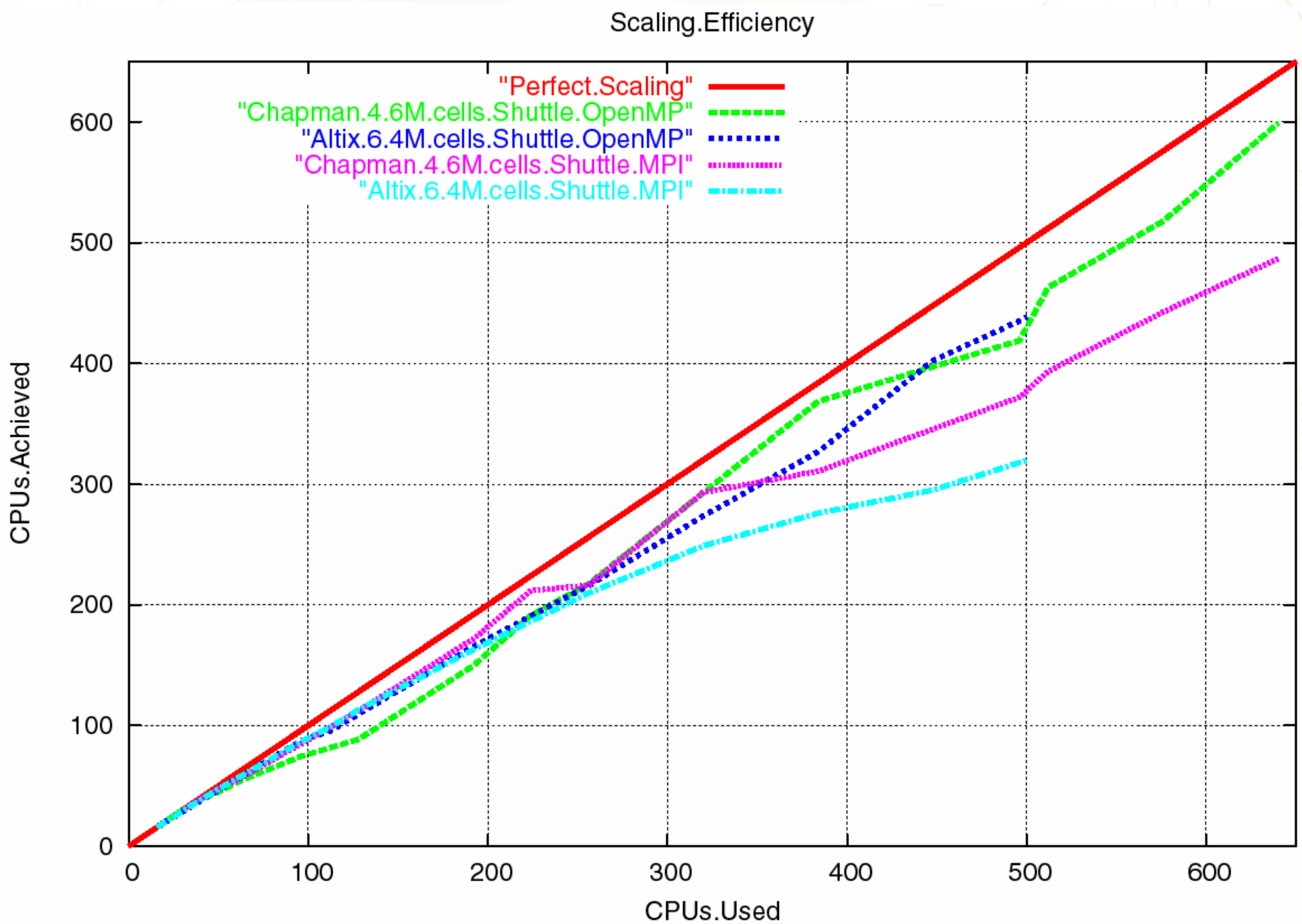
Scaling Efficiency

Altix-512p SSI (6.4 million cells)

	NCPUS	Efficiency
OpenMP	500	88 %
MPI	500	64 %

Origin-1024p SSI (4.6 million cells)

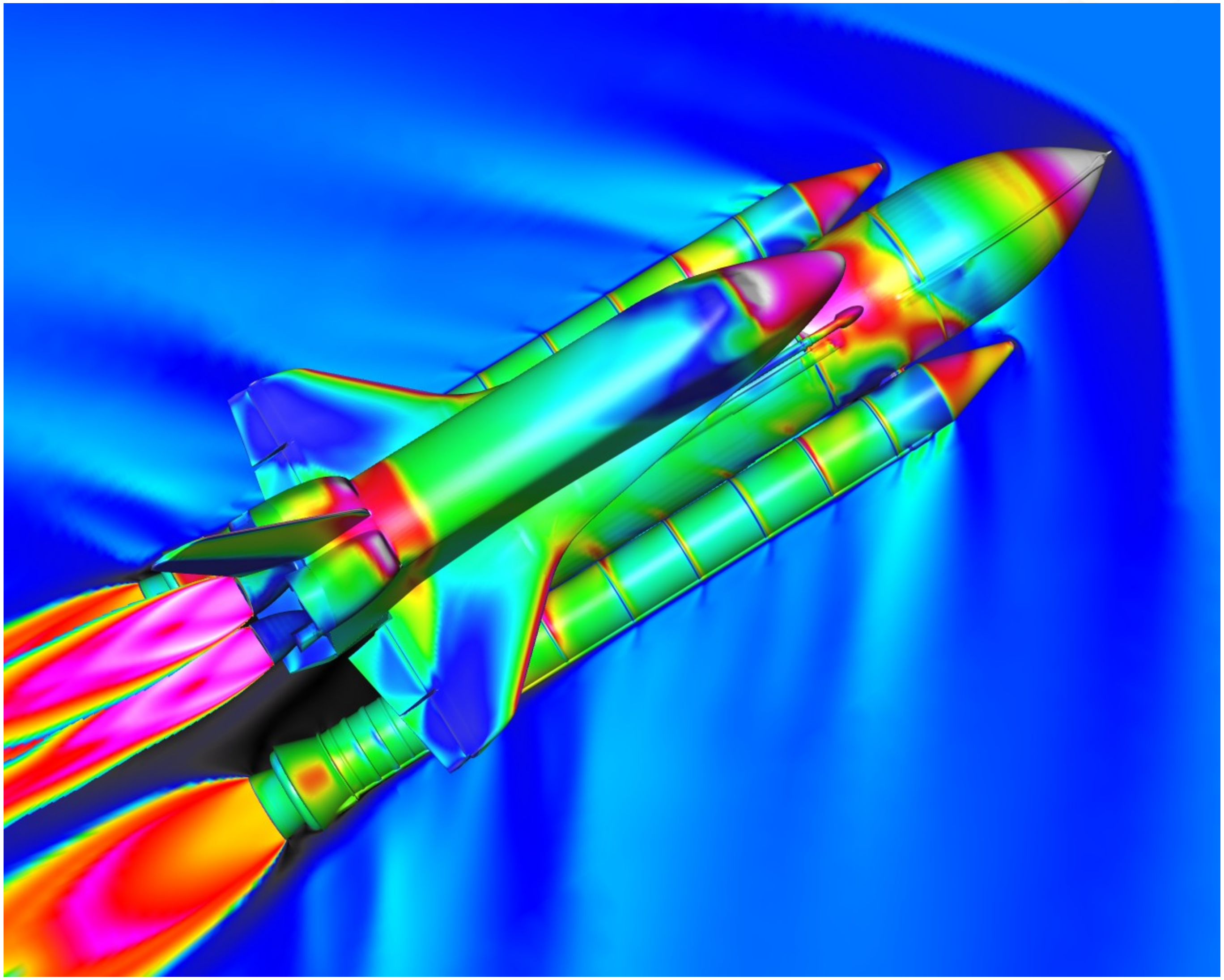
	NCPUS	Efficiency
OpenMP	640	94 %
MPI	640	76 %



OpenMP shows best scaling efficiency of 88% on 500 CPUs for Altix System.

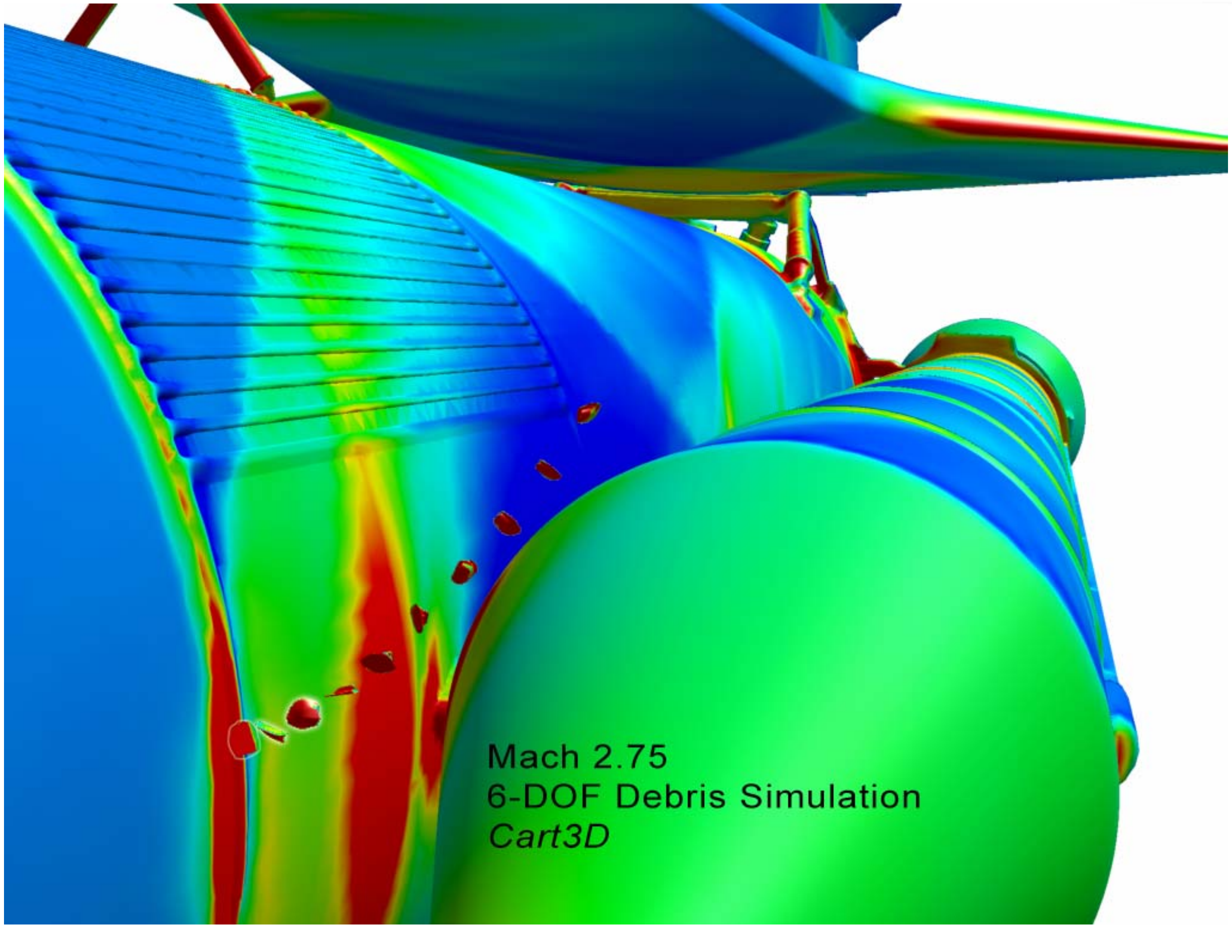
Source: mike aftosmis@nas.nasa.gov

Shuttle Return to Flight



Example: Current RTF Support Activity

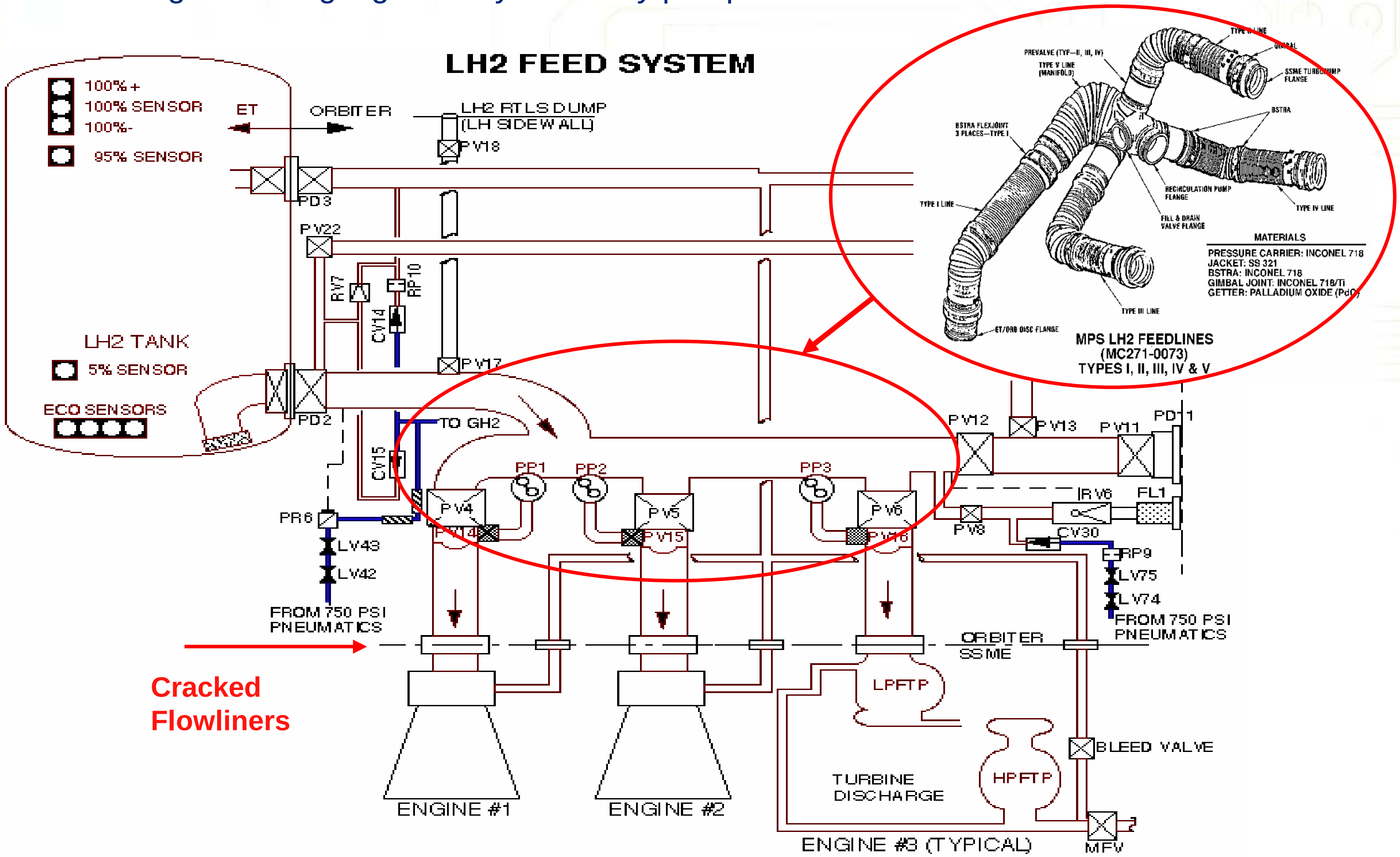
- Full-up moving body 6-DOF simulation for a release of .023 lb of foam from the intertank region near the SRB thrust panel. The altitude is at 80,000 ft, Mach 2.75. The foam released is about 5in in diameter. These simulations are being used to check the validity of the models used in the lower-fidelity debris simulations.
- Simulation will be on the order of 1-200 of these over the next couple of weeks.



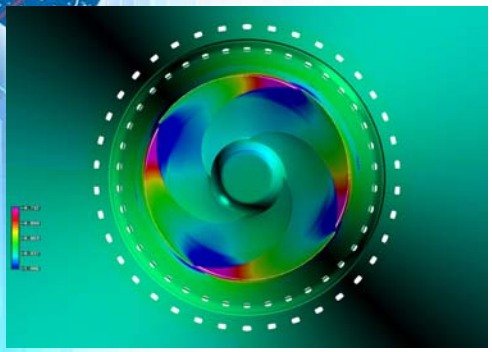
Vehicle-Propulsion Integration

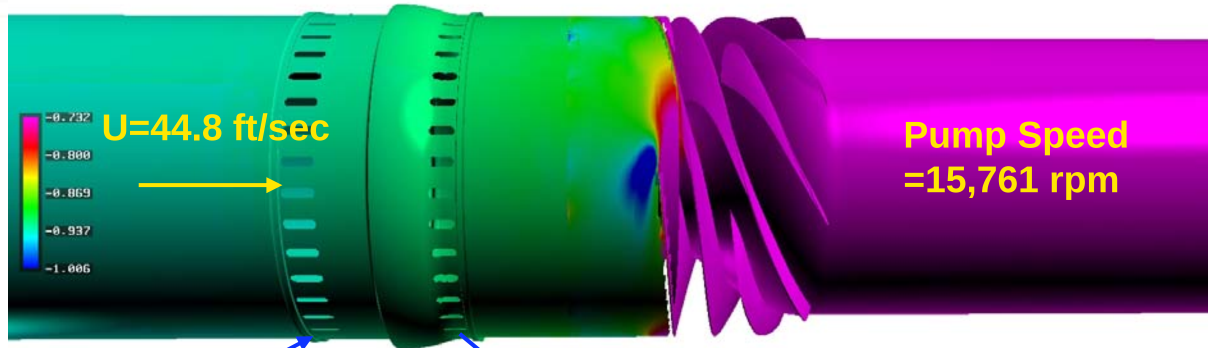
High-fidelity Analysis of LH2 Flowliner (NESC ITA)

Structural load is investigated using high-fidelity unsteady pump simulation tools.



High-Fidelity Unsteady Simulation of SSME LH2 Flowliner

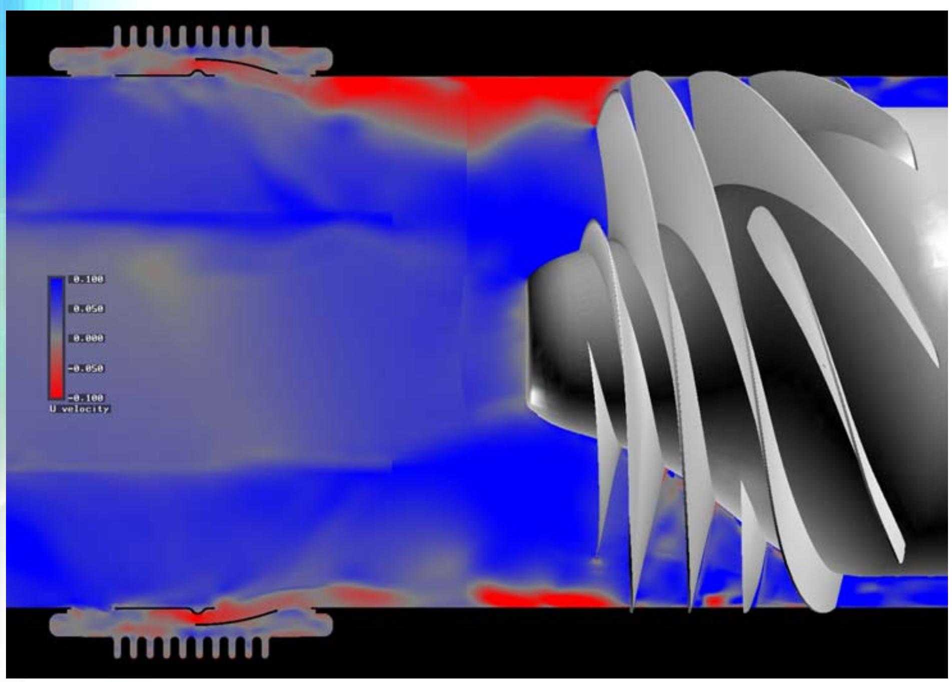


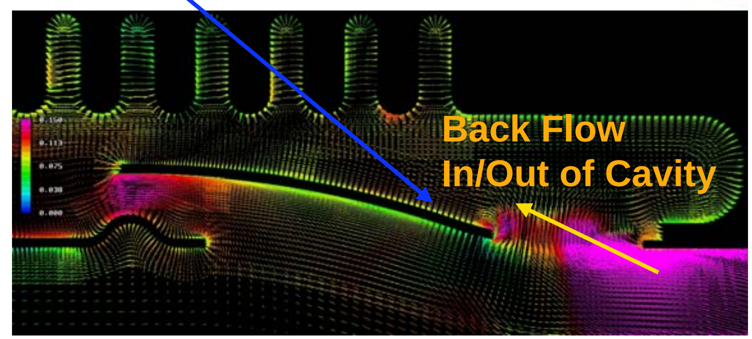


Strong backflow causing high-frequency pressure oscillations

UPSTEREAM LINER

DOWNSTREAM LINER





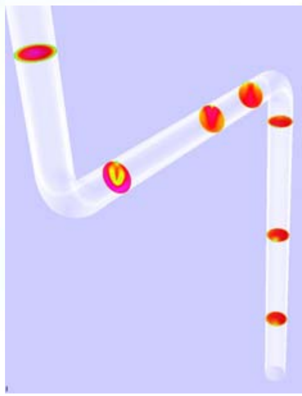
Damaging frequency on flowliner due to LH2 pump back flow has been quantified in developing flight rationale for the flowliner.

Columbia Impact at 120 days

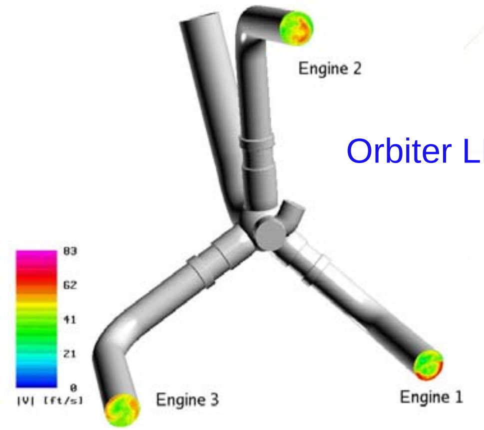
QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

High-Fidelity Simulation of SSME LH2 Flowliner

For Propulsion Vehicle Integration and Test Support



A1 Test Stand



Orbiter LH2 Feed System

FEB 2004 . . . **JUNE 2004** . . **NOV 2004**

NESC ITA Kick-Off →

1. Straight Pipe + Inducer
- 24 Million points
- 3.3 speed-up from Origins to Columbia
- 4 days/rotation (24 CPU)
- Validated w/ exp (14 rotation)

→

2. Model 1 + Flowliner
- 66 Million points
- GTA Test article
- 2-3 days/rotation (128 CPUs)
- Root cause investigated (16 rotations)

3. Engine I line +Flowliner + Inducer
- 78 Million points
- Engine I flight config.
- 5 days per rotation with 128 CPUs. Scalable to 512 CPUs.

→

4. A1 Test vs Orbiter Fuel line System
- 7 and 17 Million points
- Test article was compared against flight conditions.
- 1/2 and 1 days per case with 128 CPUs

Goal: 3 Engines

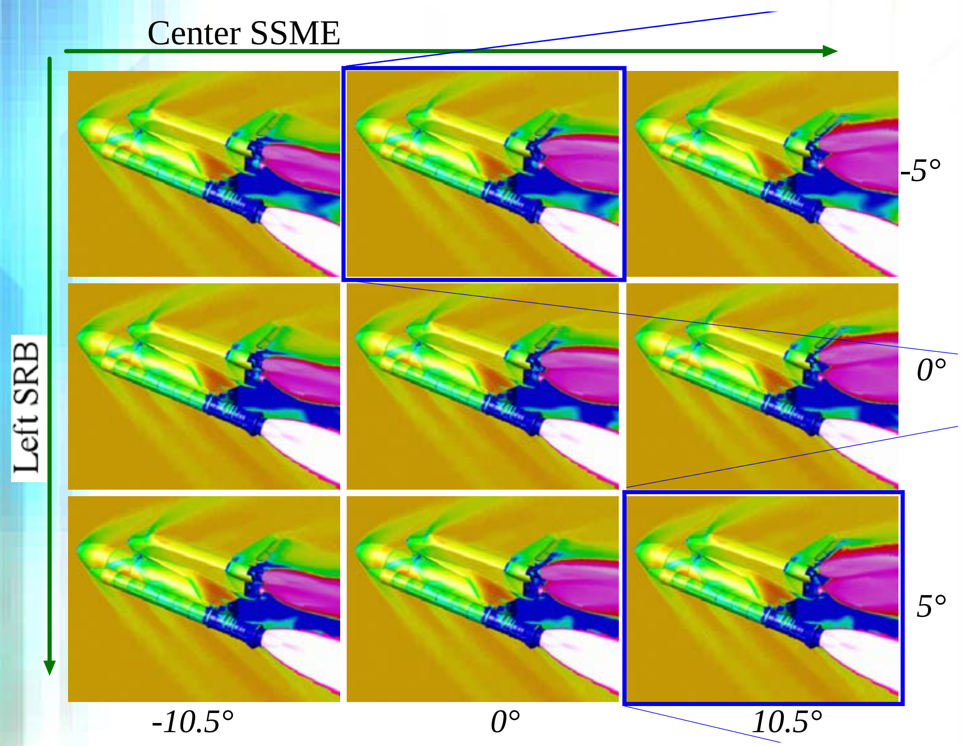
- **200 Million points**
- **100+ rotations for convergence**
- **1 rotation/day (512 CPUs)**
- **5 days (20x512 CPUs)**

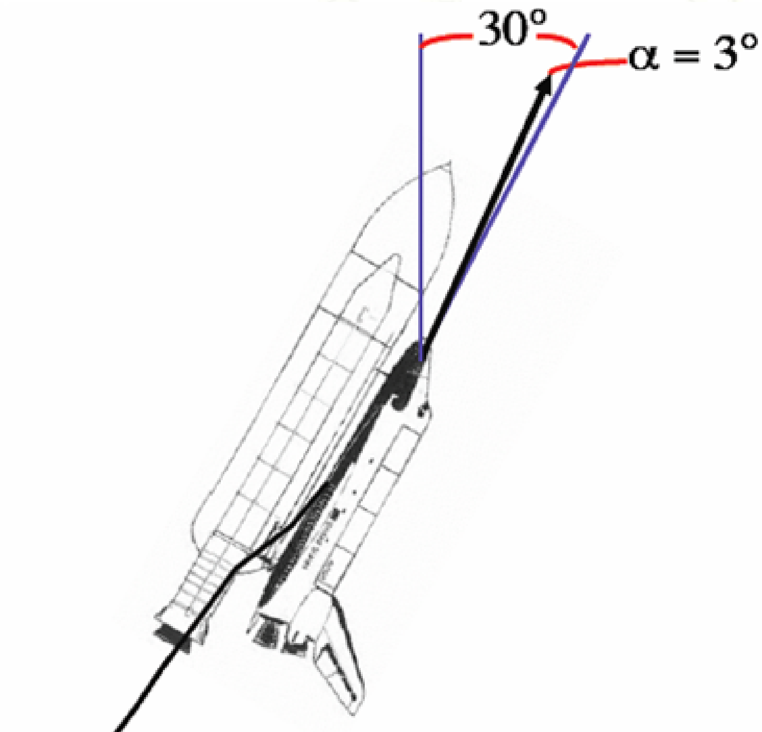
⇒ Flight Rationale

Database Generation and Trajectory Analysis (Cart3D)

(Digital Flight)

Configuration-space



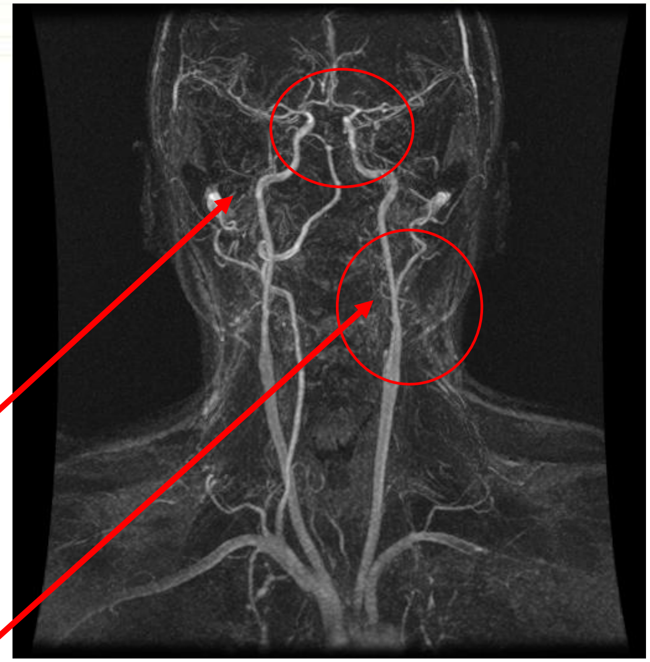


- Trajectory start
  - 69 sec. MET, full throttle,
  - Alt.=47kft
  - Mach=1.75
- Fly 27.5 sec through database
- Seek optimal controller gains holding constant AoA, max acceleration

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

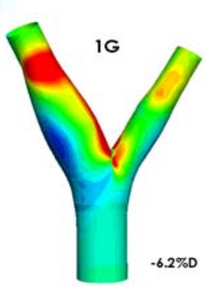
Digital Astronaut System: Near-term Performance Model

- Why Heart-Brain circulation model first
  - Black-out:
    +8G causes unconsciousness
  - Red-out:
    -3G makes retina engorged with blood



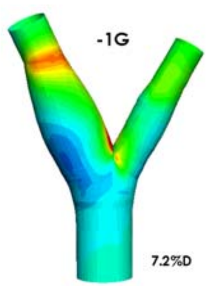
Circle of Willis
(Velocity Magnitude)

Carotid Bifurcation (Wall shear)



Standing

ion here



Hand-Standing

Impact of Columbia after 150 days

- Cosmology
  - Dark Matter halos
  - Supernova Detonation
  - Colliding black hole galaxies
- Solar Modeling
- Climate and Weather

N-Body Simulation of Galaxy Dark Matter Halos

- Dark matter halos are the seed for galaxies
- Current simulation is a multimass with 20 million particles in a Periodic 120Mpc box
- Completed in under a week using hybrid open MP/MPI code on 256processors
- Multi Mass resolution used to resolve while preserving large scale gravitational effects
  - Cubic grid Mesh adjusted at each time step based on evolution of particle distribution

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

Climate and Weather Modeling using fvGCM

The finite-volume GCM (fvGCM) is a next generation modeling system based on a state-of-the-art finite-volume dynamical core and the ***community built*** physical parameterizations and land surface model.

Timeline of Development and Evaluation of .25 deg fvGCM

- Accounts created on Altix and codes ported late March
- **Codes optimized, 3~4 weeks**, and initial climate spin-up runs started (Apr)
- Started transferring ICs and BCs from NCCS to NAS in late June
- The first .25 deg simulations (for September 2002 an 2003 were run (July)
- The first real-time hurricane simulation was run on August 10
- **Near real-time NWP started on August 11**.

Current Status

- Tuning runs for the .25 version of the fvGCM are underway
- Evaluations of the .25 deg fvGCM forecasts have begun, which are indicative of significant improvements in the representation of Tropical convection and storms.
- An extensive evaluation in collaboration with NOAA is planned and should be completed before the next hurricane season.

Courtesy of Robert Atlas, Chief Meteorologist and the fvGCM model development group

NASA fvGCM Hurricane Ivan Track

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

Courtesy of Robert Atlas, Chief Meteorologist and the fvGCM model development group

Landfall time
9/26/04Z

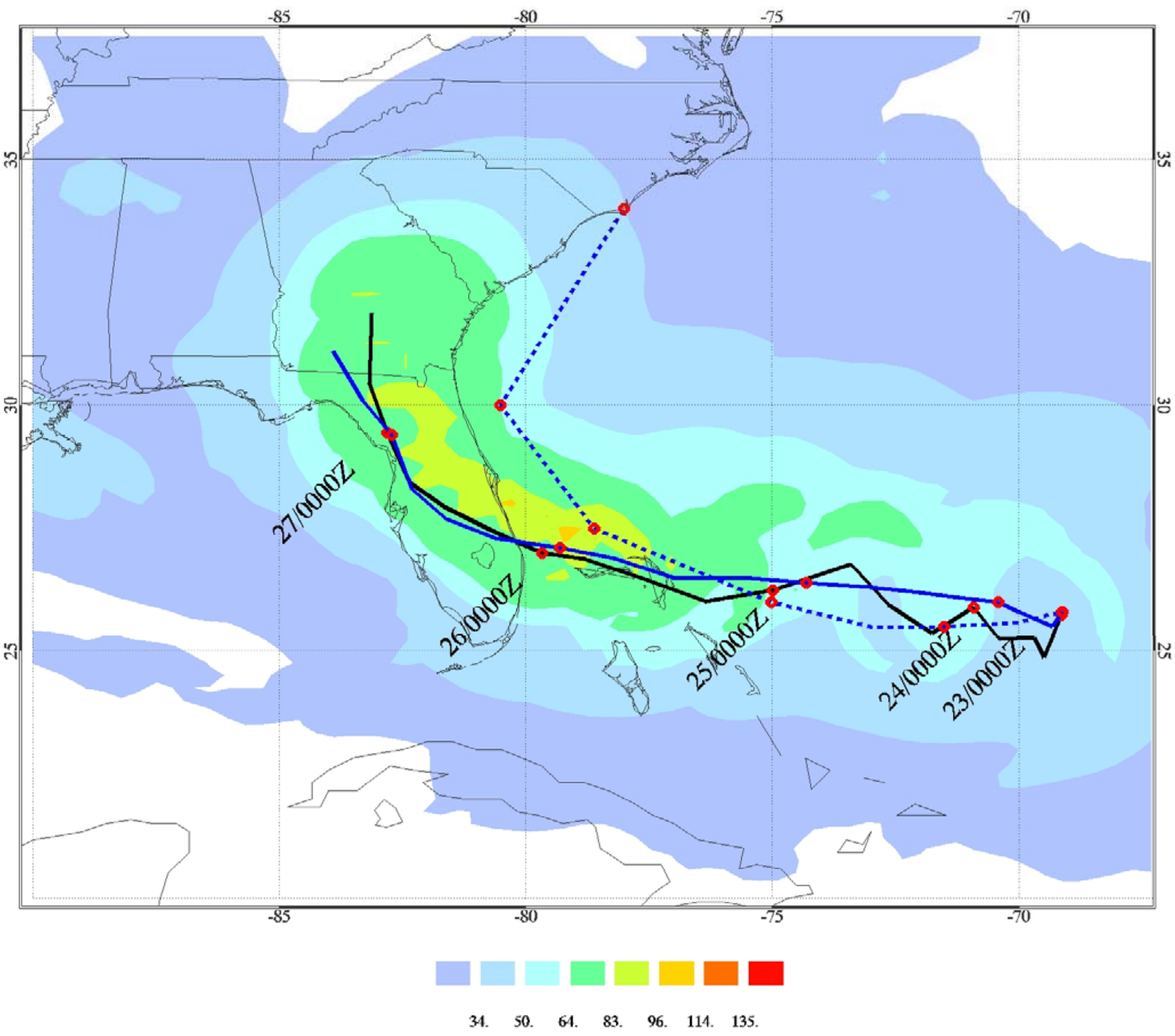
Landfall time error
NHC: 43hrs
FV: -2hrs

Landfall position error
NHC: 886 km.
FV: 2 km.

NASA fvGCM Hurricane Jeanne Forecast Track [Black] and NHC Observed [Blue] and NHC Forecast [Dashed]

Maximum Sustained Surface Wind Speed [knots]

Initialized 2004 SEP 23 00Z



Courtesy of Robert Atlas, Chief Meteorologist and the fvGCM model development group

Hurricane Ivan

Color Indicates Altitude

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

Value of "Productivity"

(do you see value in shared memory systems?)

- Full Cost of Implementation
  - Design/Develop/Debug/Maintenance
- Time Sensitive Value
- Opportunity Cost
  - What aren't you doing because you are too busy developing parallel code?

Shared Memory Enables

- Flexibility in approach
  - OpenMP/Multi-Level Parallel (MLP)/Shmem/MPI/Other
- Scalability/Performance
  - 21% of peak on Itanium2 (1.3 Gigaflops on 3-D CFD code – Cart3D)
  - 80% of 500 processors w/ OpenMP (Cart3d – POC Aftosmis@nas.nasa.gov)
- Efficient access to data
  - Local high performance file systems
  - High sustained performance on entire problem
- Deployment
  - Quick and Straight Forward

Project Columbia

120 Days to Build it

and

180 Days

of

Science and Engineering