Benefits of Scalable Servers with Global Addressable Memory for Crash Simulation

Authors:

Christian Tanasescu, Kevin Fox
SGI

Correspondence:

Christian Tanasescu
Director of Application Engineering
SGI

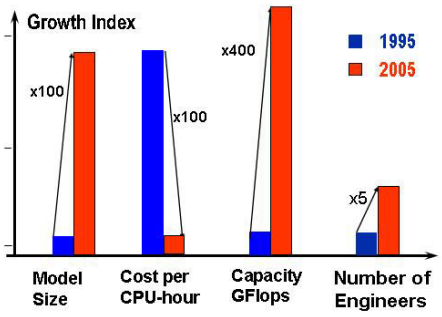
Email: christi@sgi.com

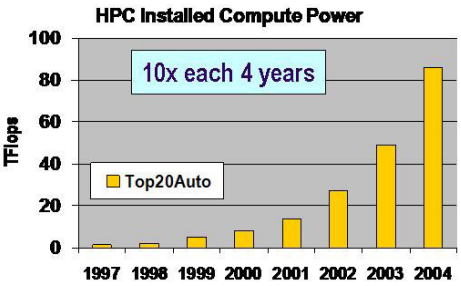
Topics

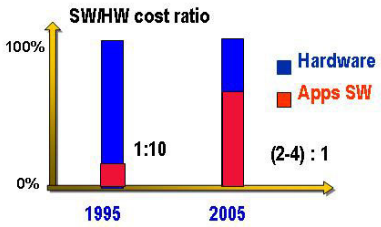
- Scope of CAE and its Growth Factors
- Overview of HPC for Crash Simulation
- LS-Dyna Performance Characterization
- Altix server family for LS-Dyna
- Simulation Environments

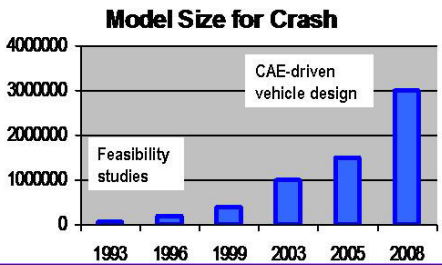
sgi

Growth Factors for CAE





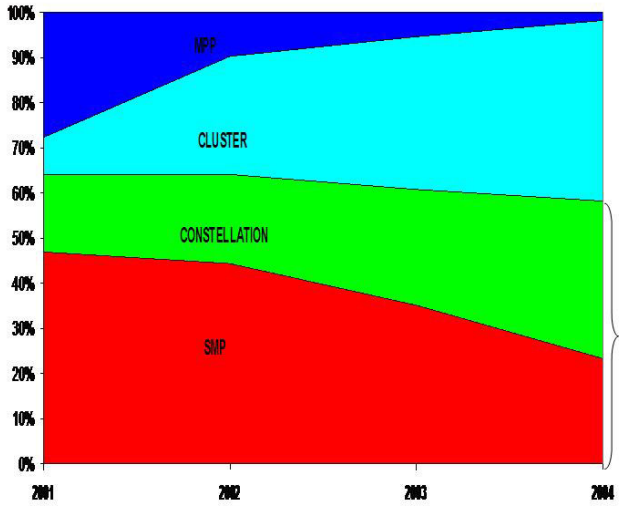




Source: Top20Auto, Ch.Tanasescu, SC'2003, www.top500.org http://www.top500.org/lists/2003/11/Top20Auto_Top500V2.pdf

sgi

System Architectures Market Shares for HPC in Automotive



CONSTELLATION is a cluster of SMP nodes, this means the SMP architecture accounts for 60% of all the installations.

Source: Top20Auto 2004, Tanasescu

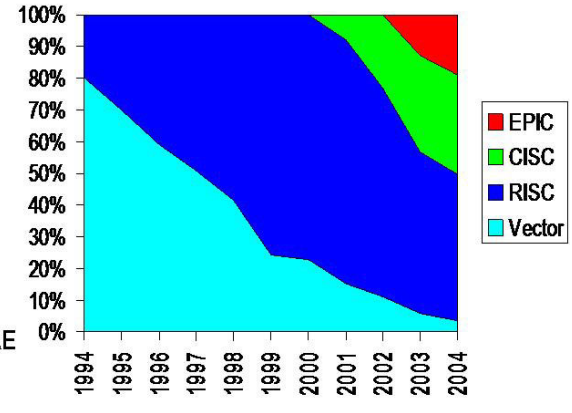
sgi

Processor Architecture Trends
HPC Trends for CAE – Study for Automotive

-Competing microprocessor architectures at instruction set level:

EPIC	Itanium2
RISC	Power, SPARC, HP, MIPS
CISC	x86, x86-64 (Intel, AMD)
Vector	Cray, NEC

-Trend to platform bifurcation in CAE
high-end (EPIC, RISC)
low-end (CISC)

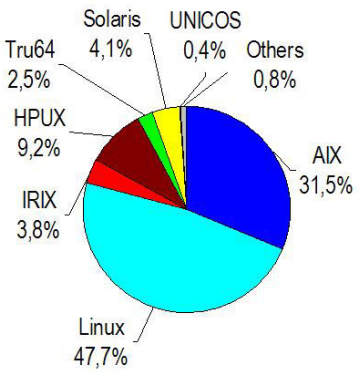


Source: Top20Auto – 2004 Tanasescu

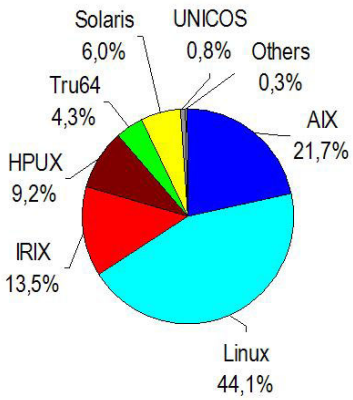
sgi

Operating Systems in HPC Automotive- 2004

Performance (GFlops)



Capacity (Processors)

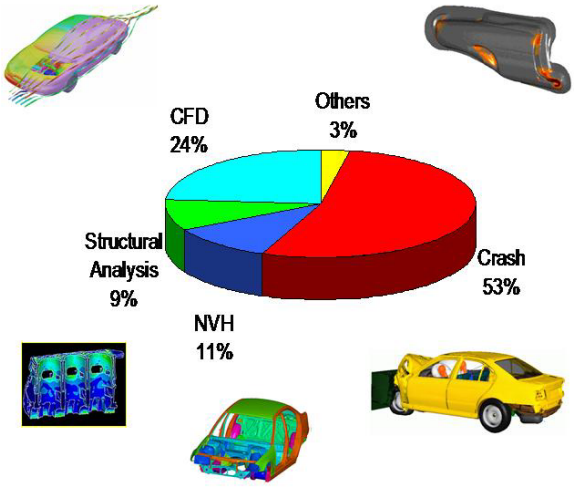


LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

Source: Top20Auto 2004, Tanasescu

sgi

CAE Application Segments in Automotive 2004



-Crash is the application segment #1

-The fastest growing applicat segments are Crash and CF

-NVH is the third application segment. However, it's the m demanding in terms of memo and IO bandwidth.

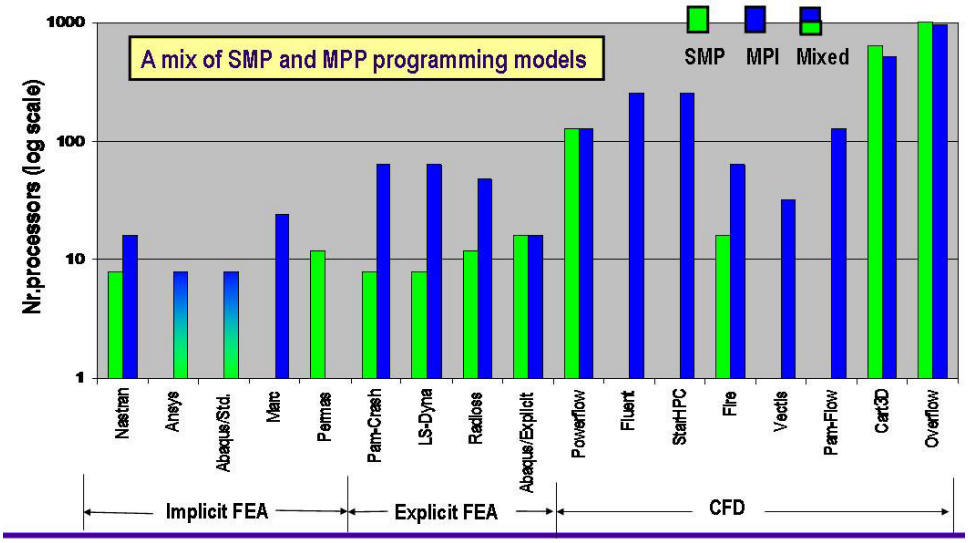
CFD - Computational Fluid Dynamics
NHV - Noise, Vibration, Harshness
Others -Stamping, Casting, Forging, CEM

LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

Source: Top20Auto 2004, Tanasescu

sgi

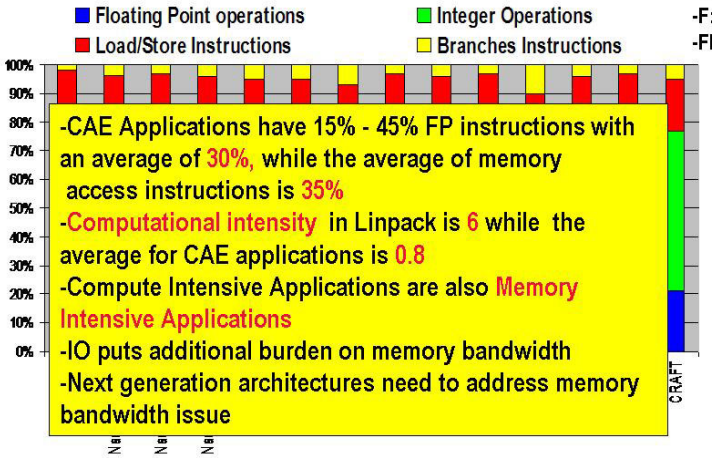
Key CAE Applications
Scalability and Implementation



sgi

Application Performance Dimensions

In depth understand real of application profiles



-CAE Applications have 15% - 45% FP instructions with an average of 30%, while the average of memory access instructions is 35%
-Computational intensity in Linpack is 6 while the average for CAE applications is 0.8
-Compute Intensive Applications are also Memory Intensive Applications
-IO puts additional burden on memory bandwidth
-Next generation architectures need to address memory bandwidth issue

Processor
-Flop rate
-Integer rate
-F:B rate
-Float/integer

Interconnect
-BW from Node
-Latency
-Bisection BW

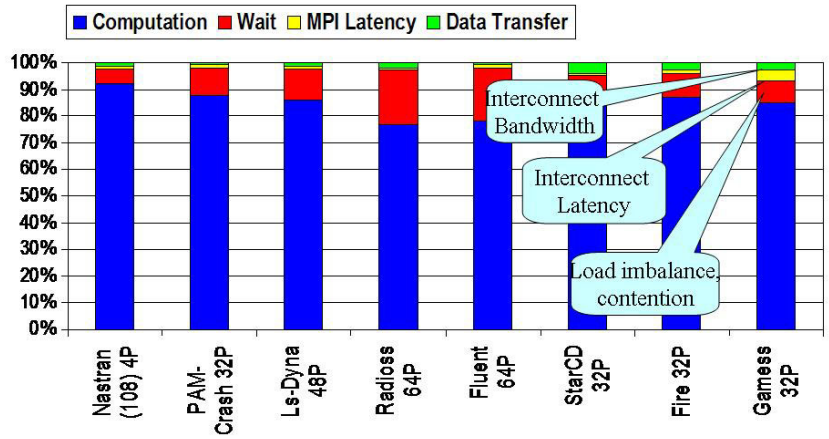
Apps Performance

Memory
-BW
-Latency
-Size

IO
-BW
-Size

sgi

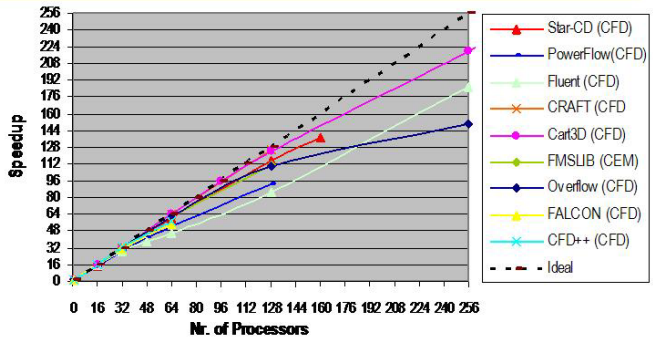
Communication vs. Computation Ratio in Key CAE Applications - measured with BandeLa



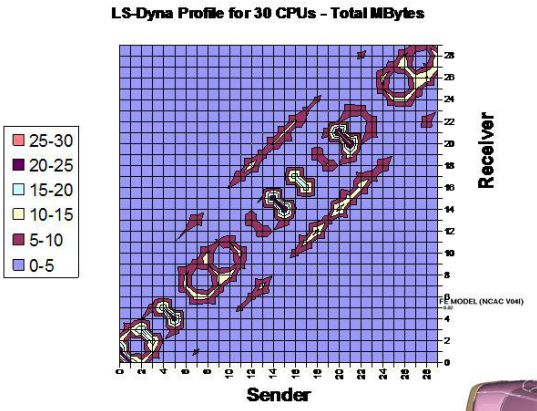
MPI In Clusters vs Global Addressable Memory

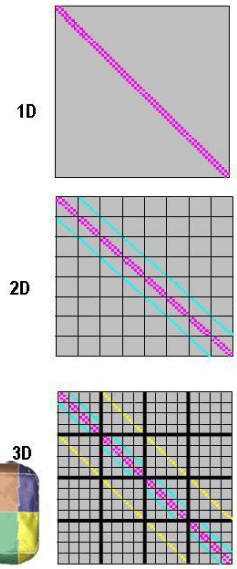
• Gigabit Ethernet	40.0 us	(Low-cost Clusters)
• Myrinet	6.3 us	(Mid-range Clusters)
•Infiniband (Voltaire)	5.0us	(High-end Clusters)
• Quadrics	2.9 us	(High-end Clusters)
• SGI Altix 3000 @1.5GHz	0.98 us	(Supercluster]

(HPC Challenge Benchmark)



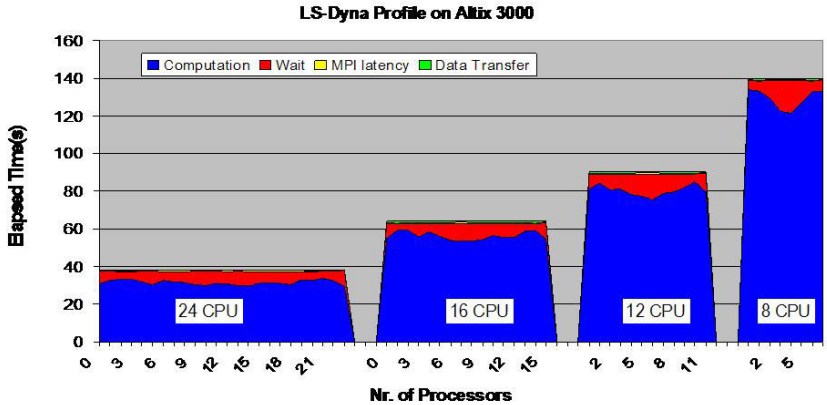
LD-Dyna mpp970
Communication Matrix





-2D Decomposition
-90% or messages shorter than 3kbyte , Latency is important for scalability

LS-Dyna mpp970
Communication vs. Computation



- This input case shows less sensitivity to latency and bandwidth
- Load balance and contention on the interconnect principal factors for scalability
- Load imbalance improves on higher CPU count
- Communication costs increase from 8% on 8P to 16% on 64P

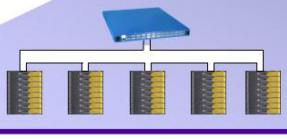
SGI Altix and Production CAE Applications

Features:
- Standard Linux kernel
- Red Hat ES or SuSe LES
- Intel Itanium 2 roadmap
- SGI NUMAFlex and ProPack

Shared Memory Server

SGI® Altix™ 3700 BX2
- SMP 4-512 cpu
- SUPERCLUSTER up to 2048P via NUMAlink
- CLUSTER via IB
- Global addressable memory up to 32GB
- Centralized HPC server for job mix of SMP and DMP/MPI

SGI® Altix™ 1350
- SMP up to 16/32 cpu
- CLUSTER via IB, GigE, Myrinet
- 2 to 192 GB per node



Large Node Cluster

LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

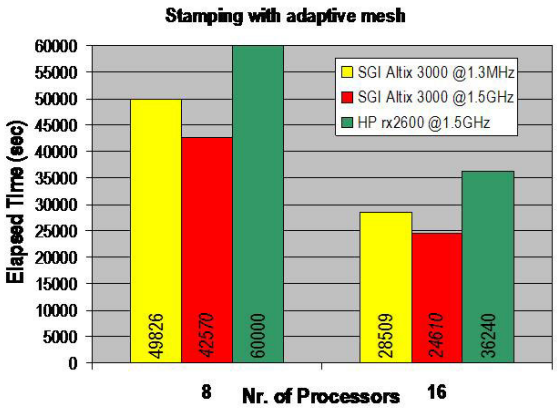
sgi

LS-Dyna mpp970.r3858 Metal Stamping

Automotive stamping job with adaptive mesh (30 times)

Total no. of Nodal points = 398602
Total no. of Shell elements = 431870
Total no. of Beam elements = 205
Total number of states = 2
Total no. of active parts = 10
Simulation Time: 21msec
No. of Cycle: 74928

8x faster better perforamance then Opteron cluster on 32p



Altix 3000 @1.5GHz is 1.47x faster than an Itanium2 1.5GHz cluster with Gigabit Ethernet. Stamping with adaptive mesh requires more memory BW than crash.

LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

Source: Nick Meng, SGI

sgi

Customer Example

- Requirement
  - High performance
  - To enable combined Explicit & Implicit solutions for same model to be run on same system
  - Explicit
    - 4.2M nodes
    - 3.7M elements
    - Run up to 128p
    - 100+ GB shared memory
  - Implicit
    - 940,000 nodes (2.8M DOFs)
    - 830,000 elements
    - 40+ GB shared memory (in–core for optimum performance)

Customer Example

System

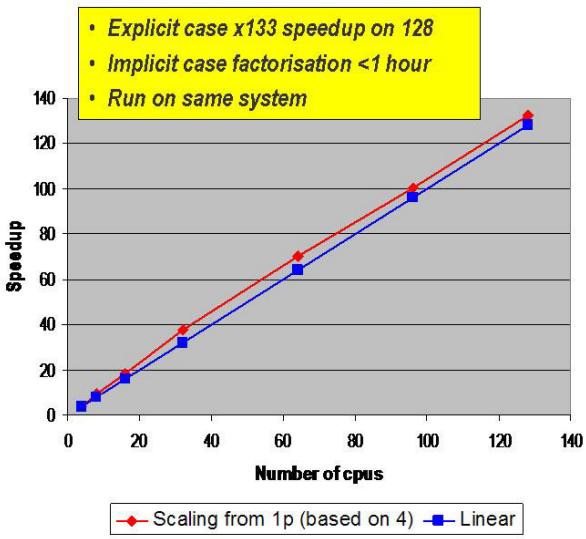
SGI Altix3700Bx2 128 Itanium2 processors, 512 GB memory

Linux (SGI ProPack 2.4)

MPI (SGI MPT 1.8)

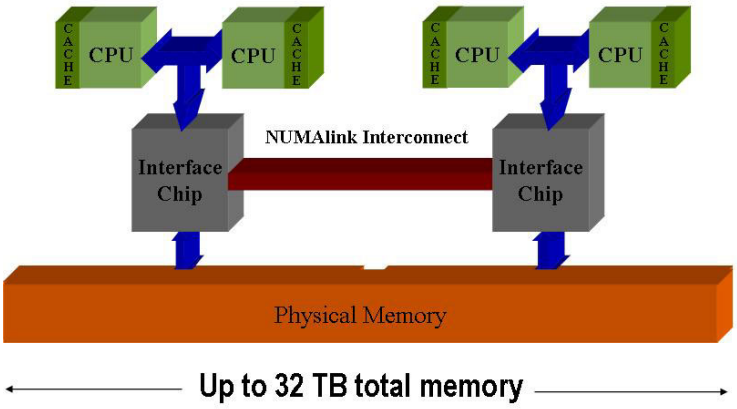
Application

LS-Dyna v970.5434 (double precision)



LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI
sgi

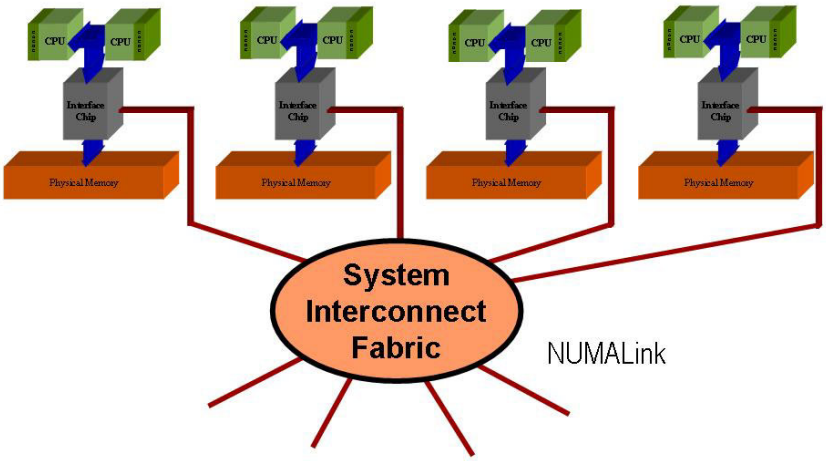
SGI Scalable ccNUMA Architecture
Basic Node Interconnect



LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

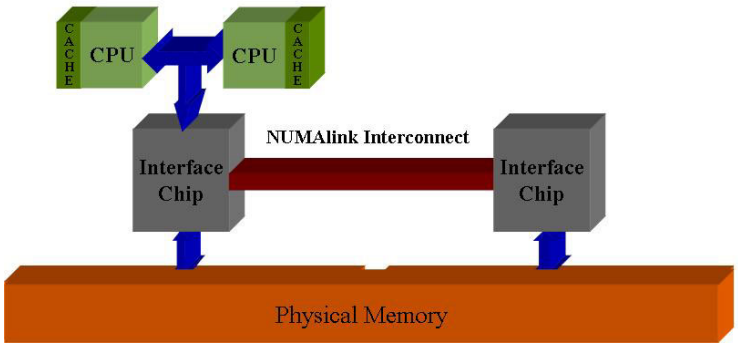
SGI Scalable ccNUMA Architecture
Scaling Communication Bandwidth



LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

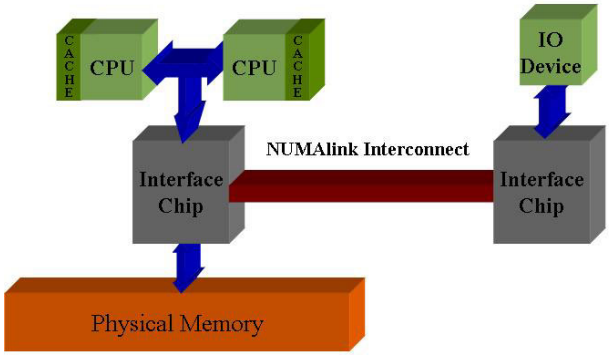
SGI Scalable ccNUMA Architecture
Scaling memory



LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

SGI Scalable ccNUMA Architecture
Scaling IO

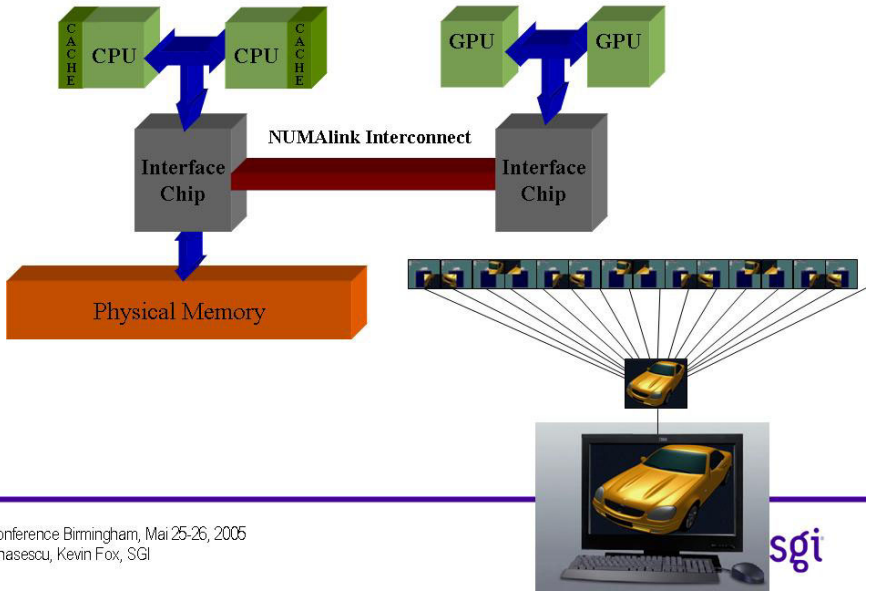


LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

SGI Scalable ccNUMA Architecture

Scaling Visualization



LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

Total Value Proposition

SGI Addresses Scientific, Engineering, and Creative Challenges

SGI customers can access, process and visualize massive amounts of data; speeding design cycles and delivering a greater return on investments

Access

Process

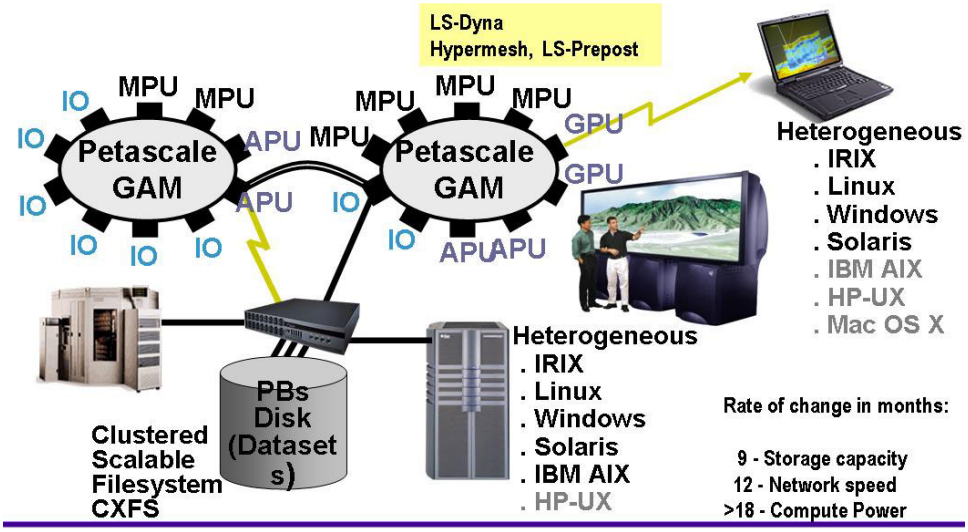
Visualize

SGI's Mission: Give the Scientist more time to solve the Science Questions, by removing the Computer Science Questions and addressing High Productivity Computing

LSTC UG Conference Birmingham, Mai 25-26, 2005
Christian Tanasescu, Kevin Fox, SGI

sgi

Strategy for Large Data Simulation Environments
Compute and Visualization Server



Conclusions

- **Performance in LS-Dyna depends on**
  - CPU power
  - Interconnect latency
  - Interconnect ability to sustain lot of messages in flight
- **SGI Altix 3700 & 350 server family is the first cluster with global shared memory across multiple nodes**
  - SSI -> 512P
  - Supercluster with Global memory -> 2048 using NUMAlink
  - Cluster -> 32P Nodes and commodity interconnect
- **SGI targeting to improve productivity for entire PLM solution**
  - Server -> Latency and Bandwidth, Petascale memory Independent scaling (CPU, IO, memory)
  - Storage -> Heterogeneous storage
  - Visualization -> Interactivity and Collaboration
- **Enabling Simulation Grid**
  - Collaborative Design and Engineering
  - Ensemble simulation (multi-scale, multi-physics simulation,MDO