SILICON GRAPHICS | The Source of Innovation and Discovery™

High-Performance Computing at SGI and the Status of Climate and Weather Codes on the SGI Altix

Gerardo Cisneros, Ph.D.

Scientist

sgi®

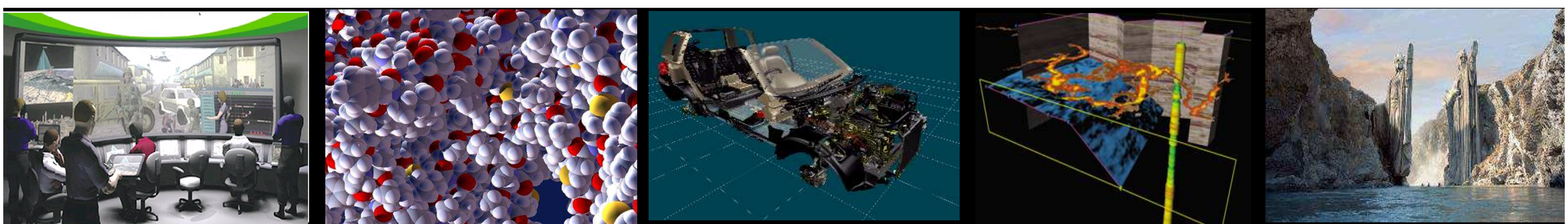
Overview

- Company focus
- SGI Altix: present and future
- Performance of NWS codes
- Conclusions

Silicon Graphics

Providing the Industry's Highest-Performing Compute, Storage and Visualization Products

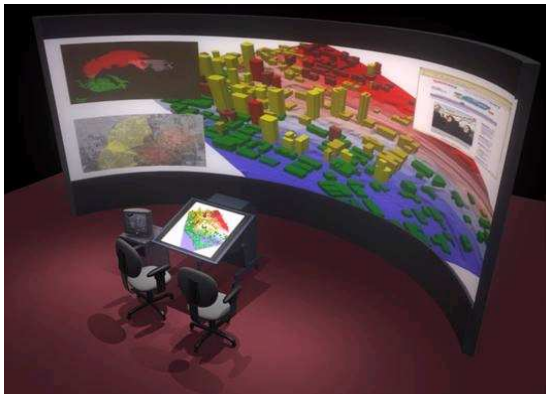
•Exclusively focused on the technical computing market
•Technology is designed to enable the most significant scientific and creative breakthroughs of the 21st century
•Products and services are mission critical to government and defense, science and research, manufacturing, energy and media industries



Images courtesy of SCI Institute, University of Utah, Navy Rehearsal TOPSCENE Program, Magic Earth LLC, and WETA Digital

Strategic Focus Areas

High Performance Computing	Storage	Advanced Visualization



High-performance NUMAflex™ architecture delivers unprecedented flexibility and performance.

CXFS™ shared file system allows transparent, heterogeneous file access everywhere, without copying data.

Onyx4™ allows users to combine multiple industry-standard graphics cards in a high-bandwidth, low-latency architecture, for cost-effective high performance visualization.

Architecture Designed for HPC; Choice of Deployments

NUMAflex™ Global Shared-Memory Architecture

Balanced, scalable performance
Operating environment optimized for HPC
Low-latency memory access
Easily Deployable

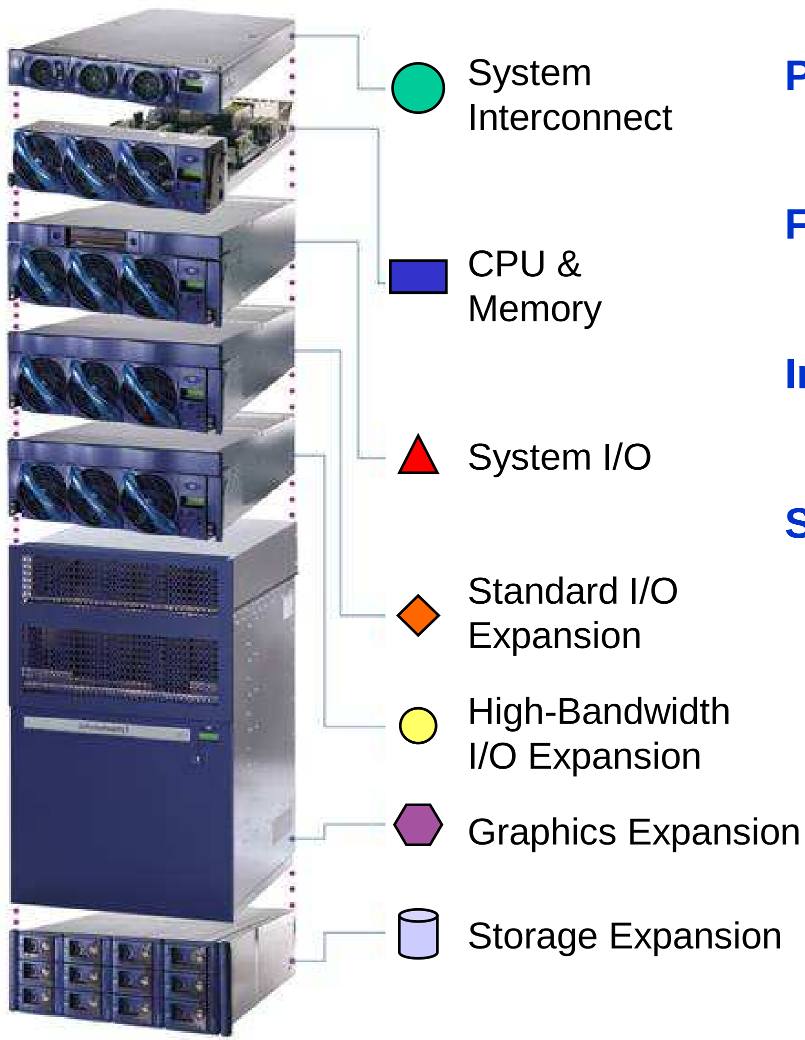
MIPS® and IRIX®

SGI® Origin® Family

Intel® Itanium® 2 and Linux®

SGI® Altix® Family

Modular SGI® NUMAflex™ Architecture

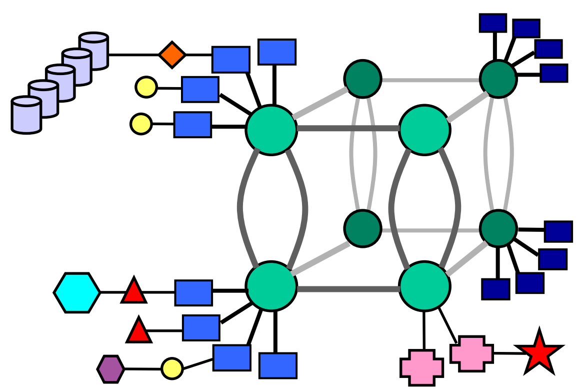


Performance: High-bandwidth interconnect with very low latency

Flexibility: Tailored configurations for different dimensions of scalability

Investment protection: Add new technologies as they evolve

Scalability: No central bus or switch; just modules and NUMAlink™ cables



SGI Family of Scalable Linux® Solutions

SGI® Altix® 350
Servers and Clusters

SGI® Altix® 3000
Servers and Superclusters

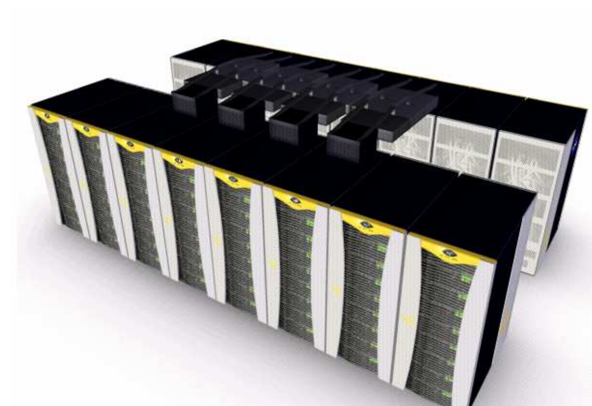
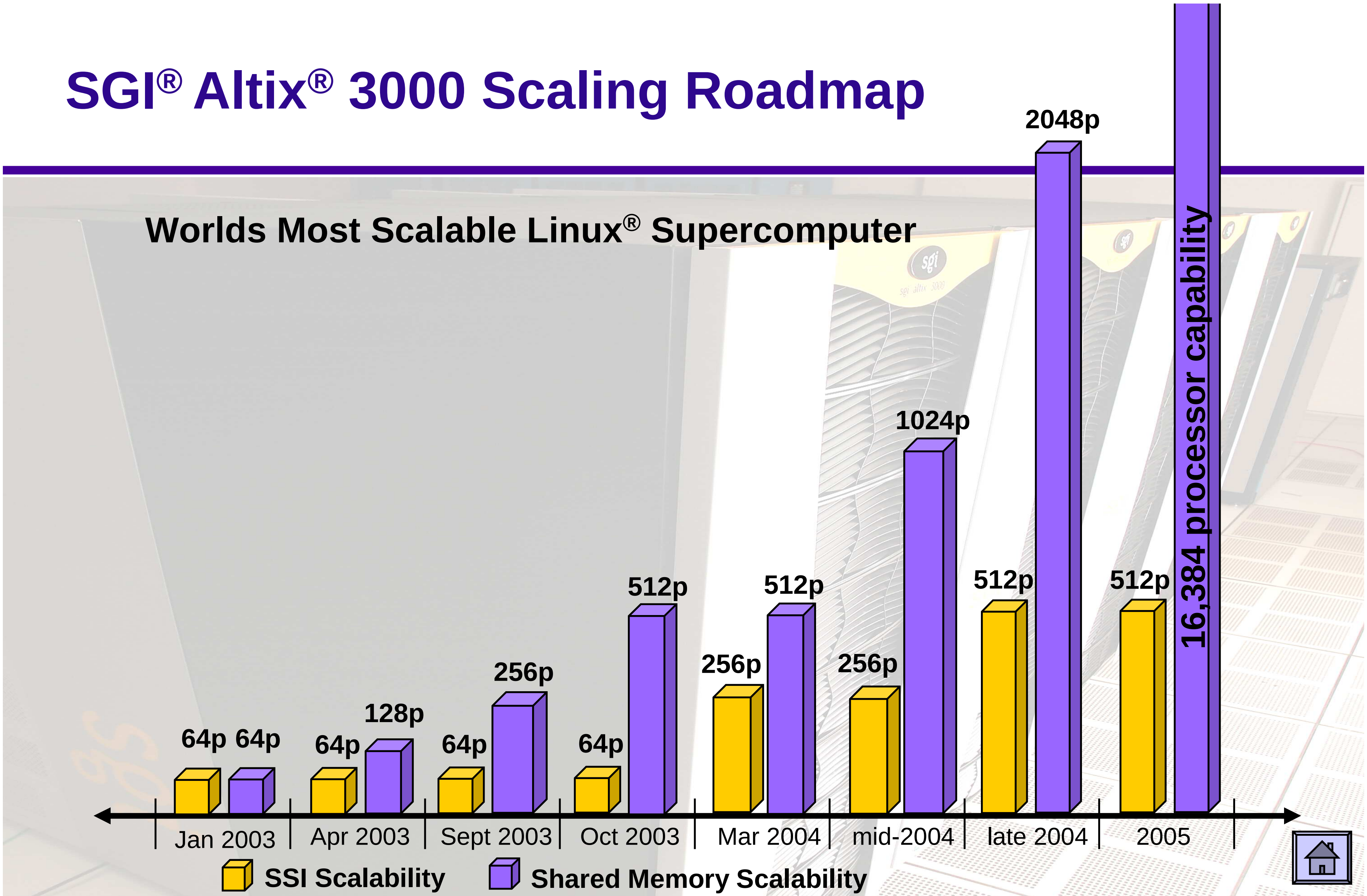


Image courtesy: NASA Ames

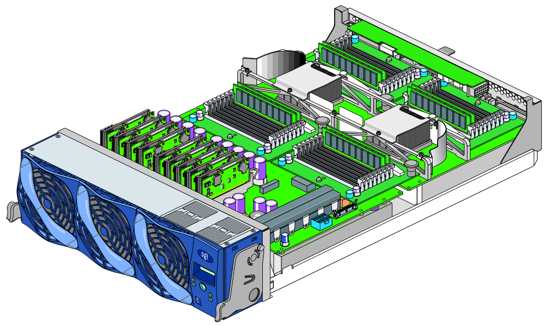
Mid-range **Departmental** **Capability**

SGI® Altix® 3000 Scaling Roadmap

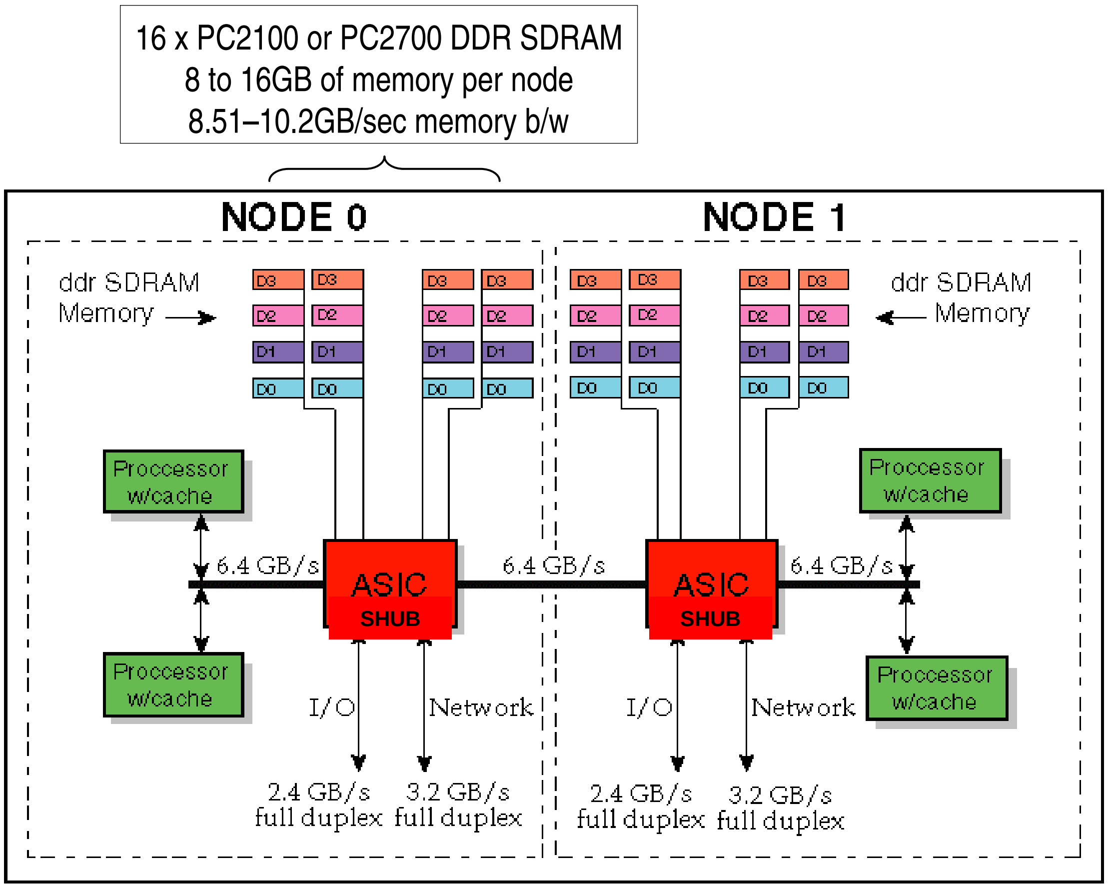


This slide contains forward-looking statements. The results and forecasts as stated may vary. Other risks and uncertainties relating to this slide may be found in the "Safe-Harbor" statement at the beginning of this presentation.

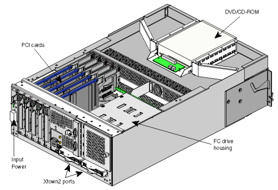
SGI® Altix® 3000 C-Brick Detail



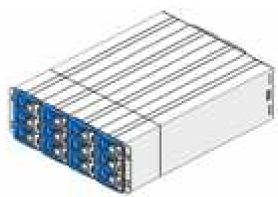
- **4x Intel® Itanium® 2 processors**
- **2 processor per 6.4GB/sec frontside bus**
- **4–64GB memory C-brick**
- **SHUB memory controller 8.51–10.2GB/sec memory bandwidth**
  (varies with memory speed 133 MHz vs. 166 MHz)
- **6.4GB/sec aggregate interconnect bandwidth**
- **4.8GB/sec aggregate I/O bandwidth**



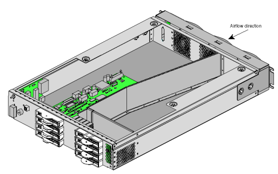
Other Altix Bricks



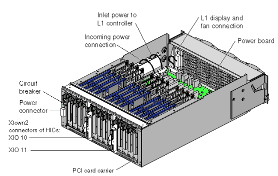
IX-brick
Base I/O module



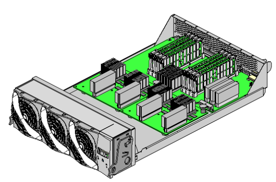
D-brick2
Disk expansion



R-brick2
Router interconnect

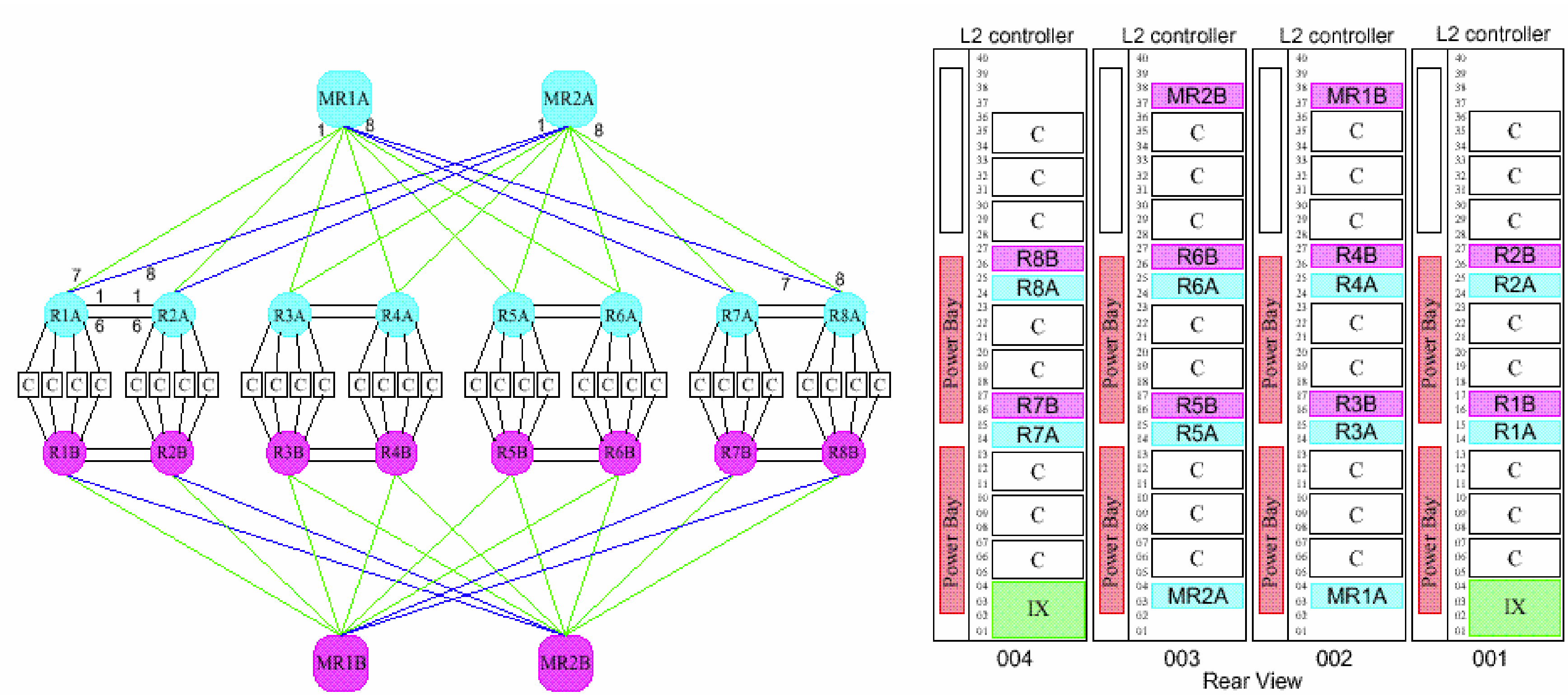


PX-brick
PCI-X expansion



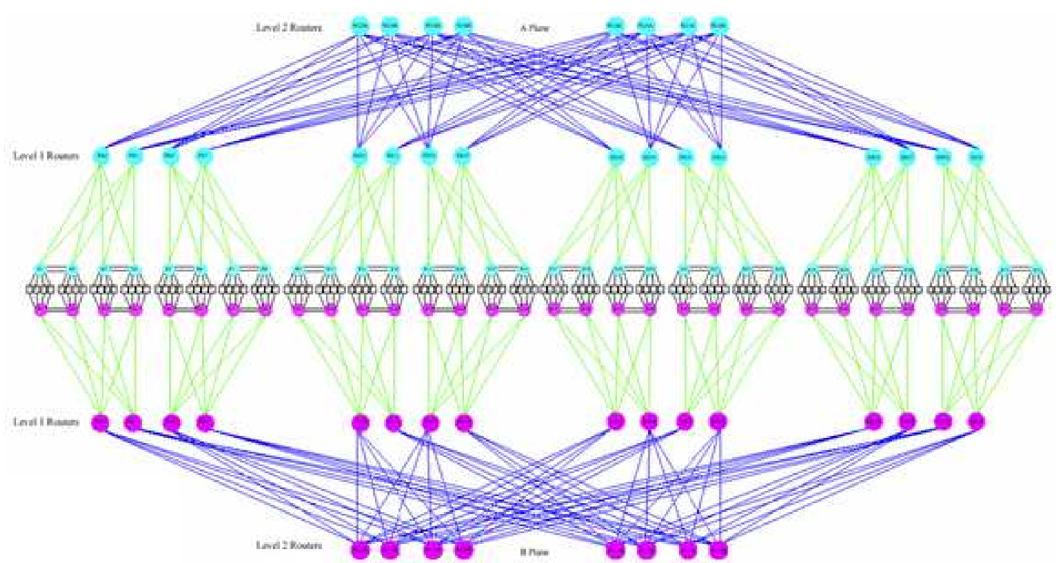
M-brick
Memory expansion

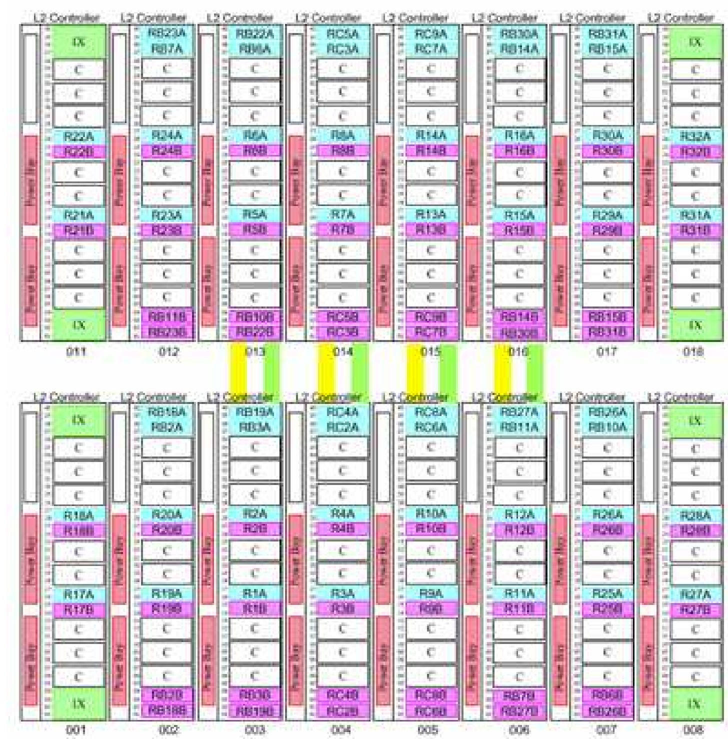
Altix 3700 — 128 processors



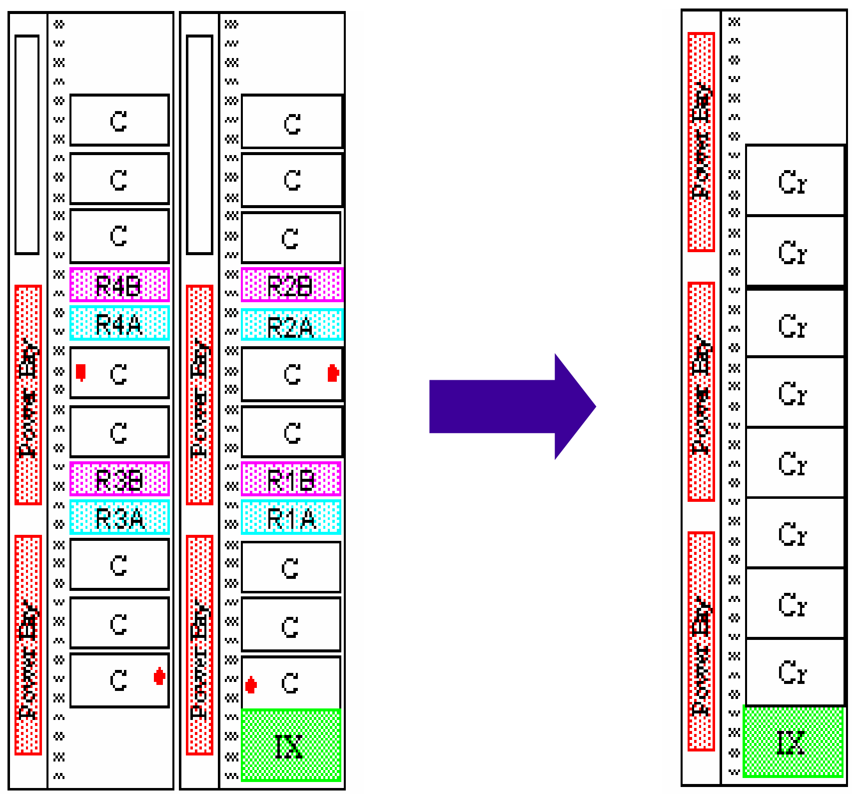
sgi

Altix 3700 — 512 processors





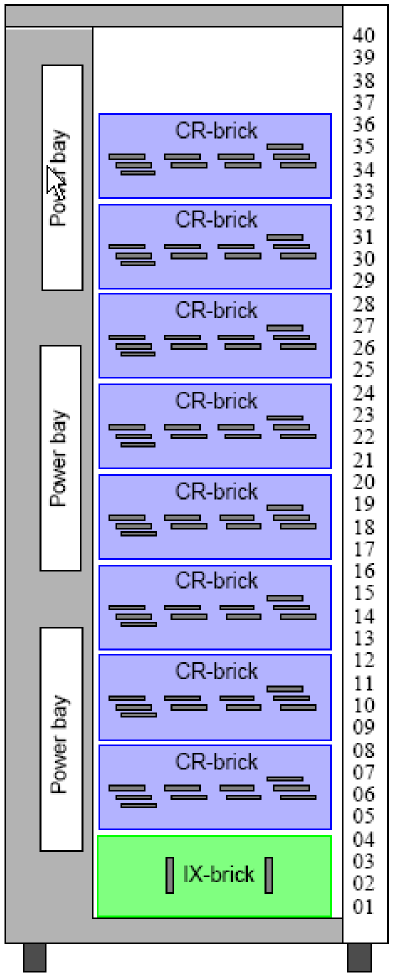
Next generation Altix



Altix 3700 **"Tornado"**

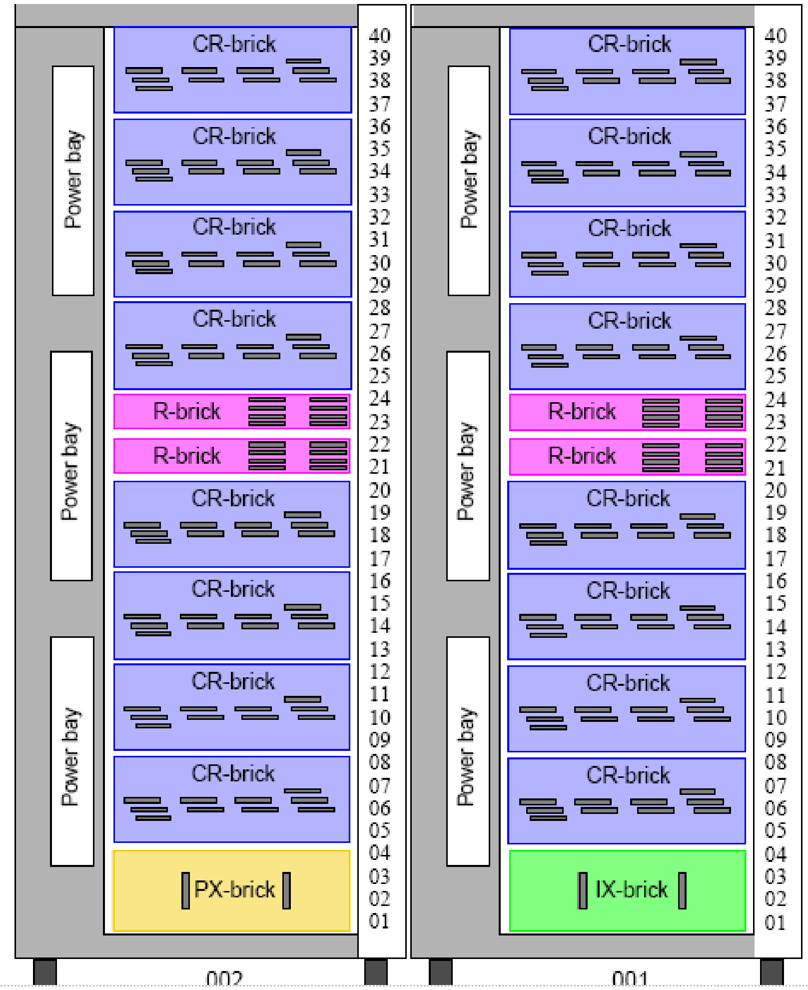
(8) Fewer External Routers
(1) Less Power Bay
(1) Less Rack
(20) Fewer Cables

Next generation Altix



Dual-plane
64-processor

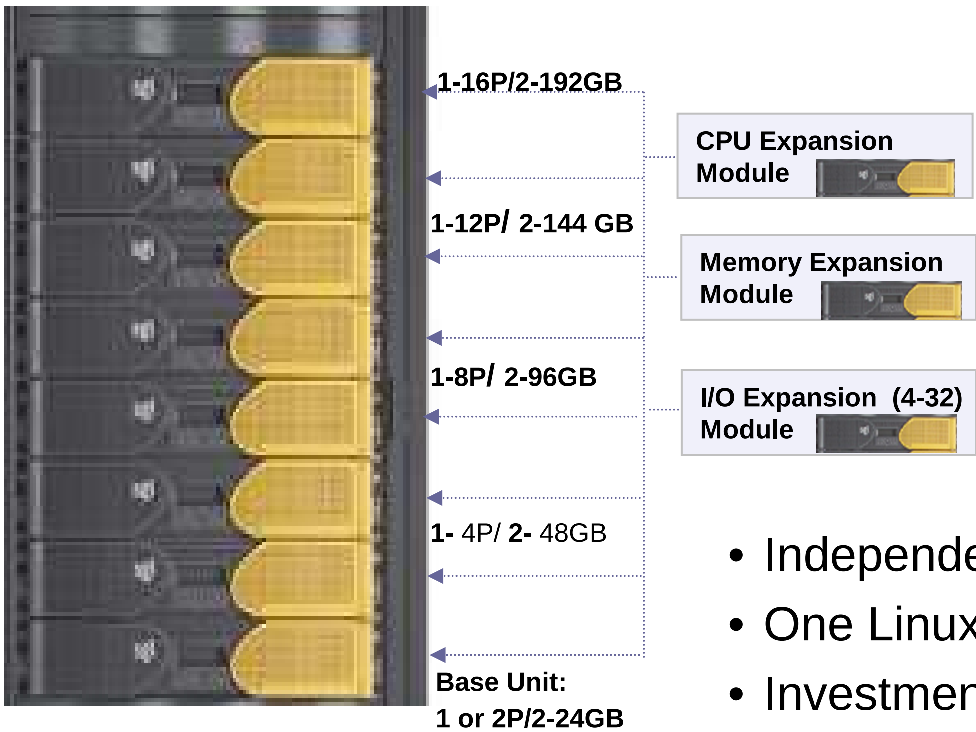
Note:
Additional D-brick racks not shown.



Dual-plane
128-processor

Note:
Additional racks with D-bricks not shown.

The Altix® 350: “Expand on Demand” Growth Path



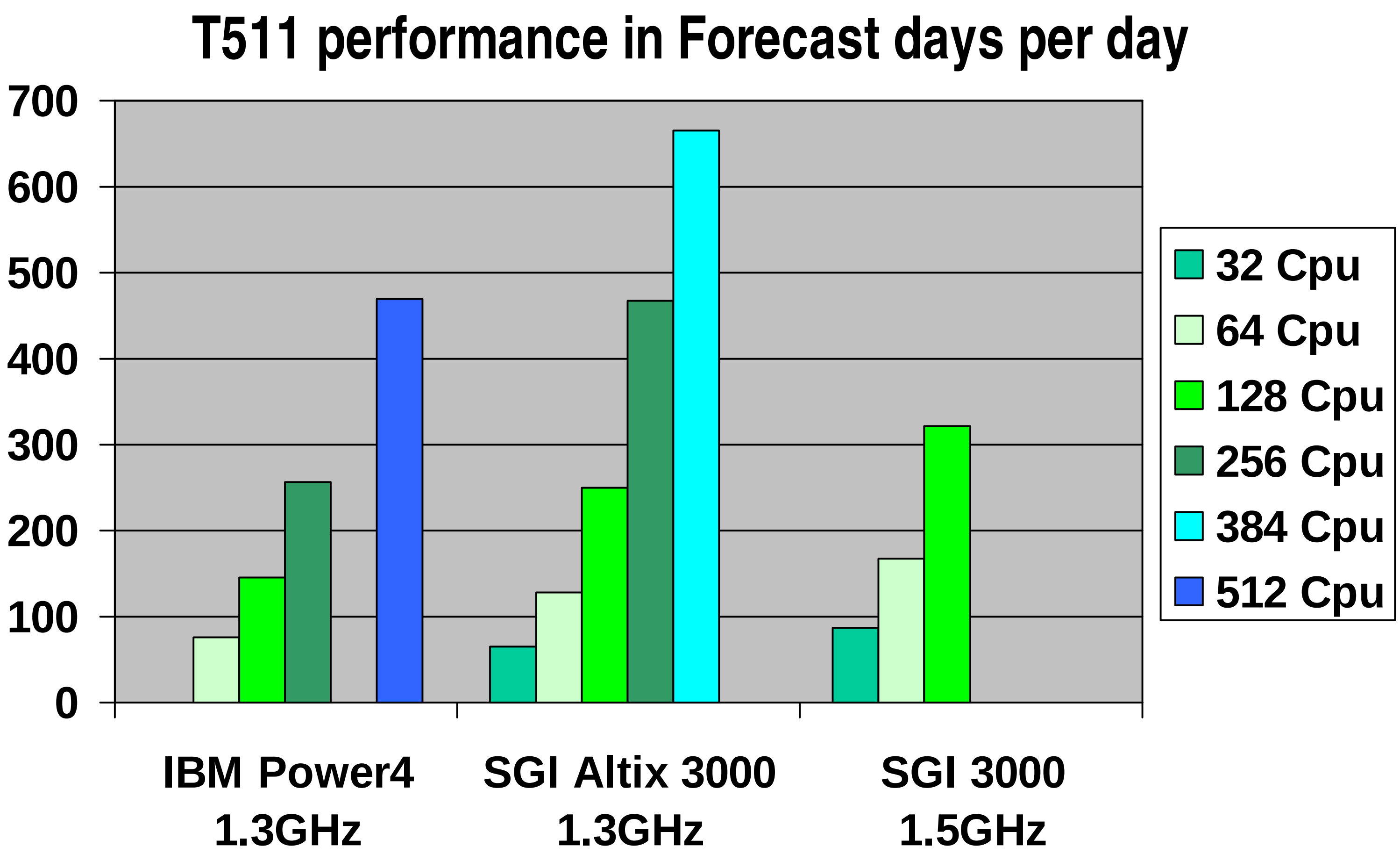
Right-size systems for the ultimate price/performance

- Independently scale CPU, memory, I/O
- One Linux® instance to manage
- Investment protection & leverage current assets
- Allocate budget and resources to ongoing needs

Customers with Altix systems running climate and weather applications

- NASA – 10240p – ECCO, CAM, CCSM, etc.
- NCSA – 1024p – WRF
- GFDL – 608p (2x256, 1x96) – MOM4, CM2.1
- NRL – 384p (1x128, 1x256) – various
- ORNL – 256p – CCSM, CAM, POP
- LANL – 256p – POP, HYCOM
- BAMS – 20p – MM5 and MAQSIP
- CMMACS – 12p Altix350 – MOM4
- APAT – 8p Altix350 – MM5
- Romanian Nat'l Met Admin – 2p Altix350 – HRM

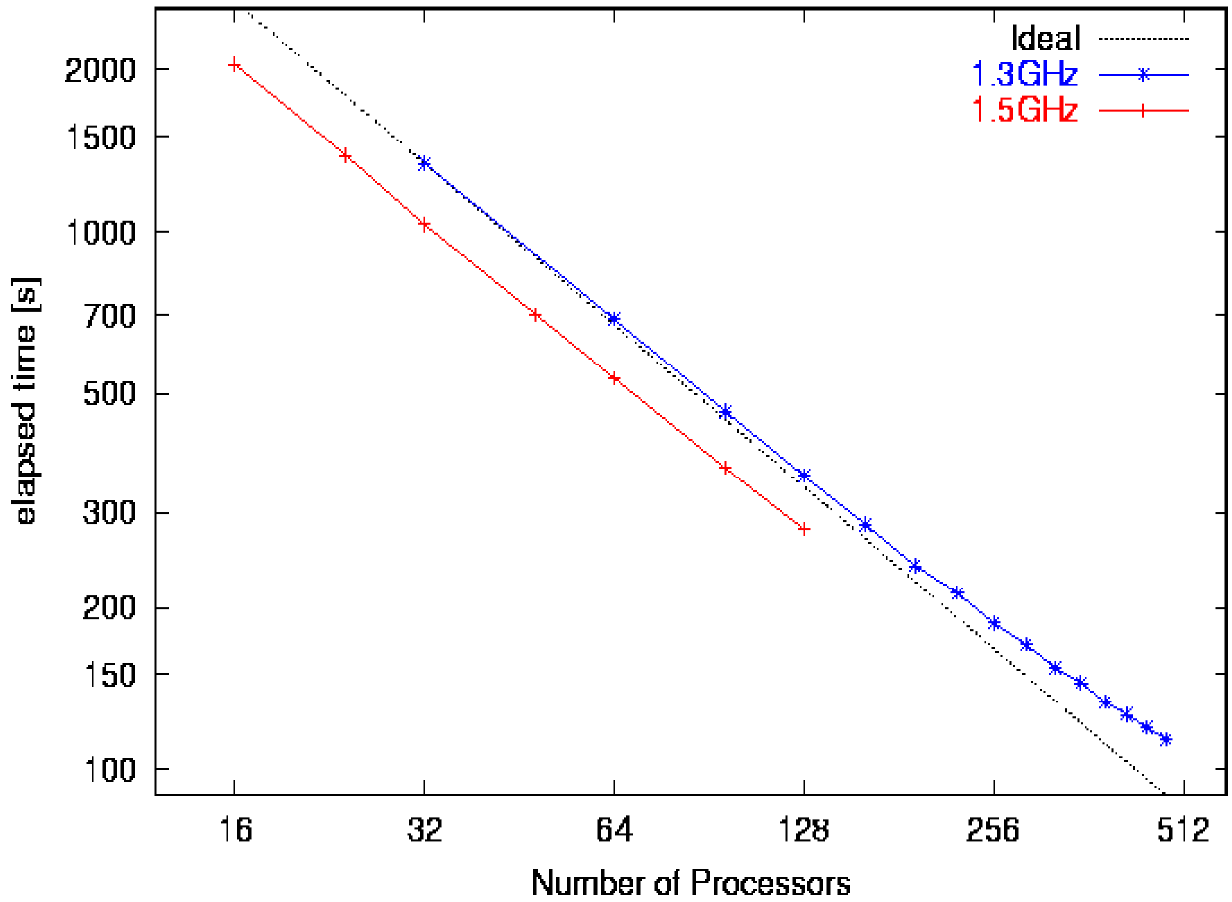
IFS - Performance on SGI Altix 3000



Altix 3000 @1.5GHz is 2.22 x faster IBM Power4 @1.3GHz

Source: Roland Richter, SGI, May 2004

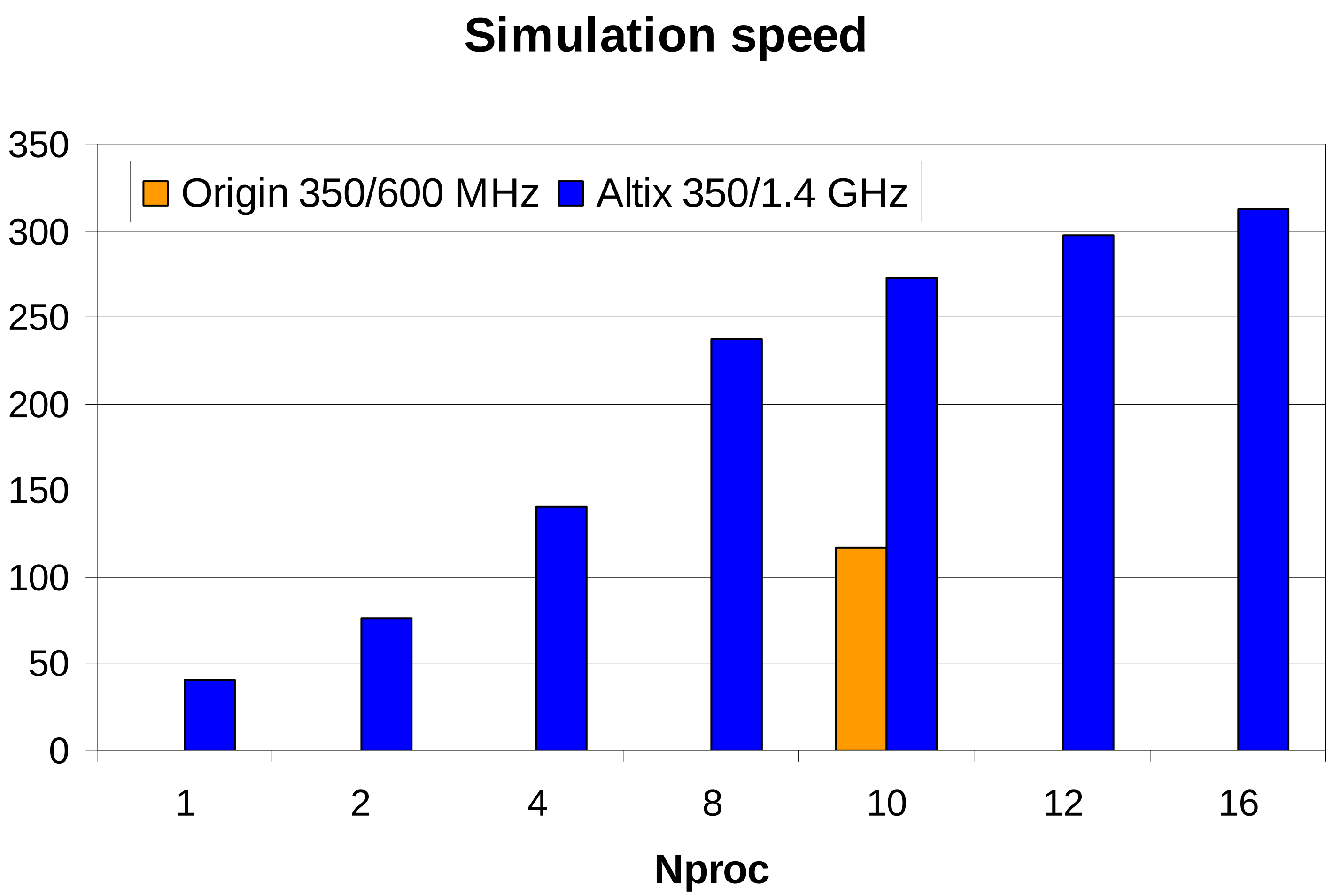
IFS - Scalability on SGI Altix 3000



Itanium2 @1.5GHz is 1.3x faster than Itanium @1.3GHz because of the larger cache and higher clock rate

Source: Roland Richter, SGI, May 2004

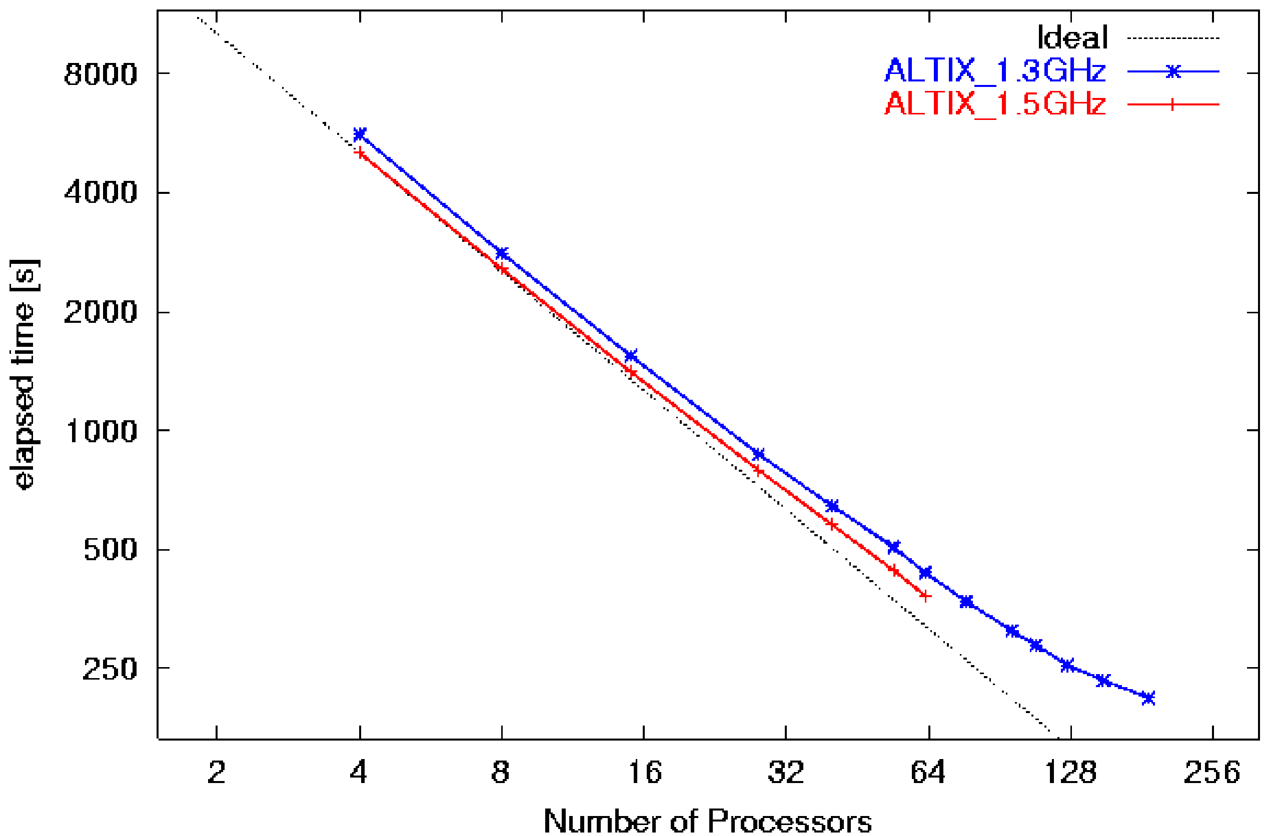
HRM Performance on the Altix 350



78h forecast in 2340 time steps over a 181x217 horizontal grid with 26 vertical levels

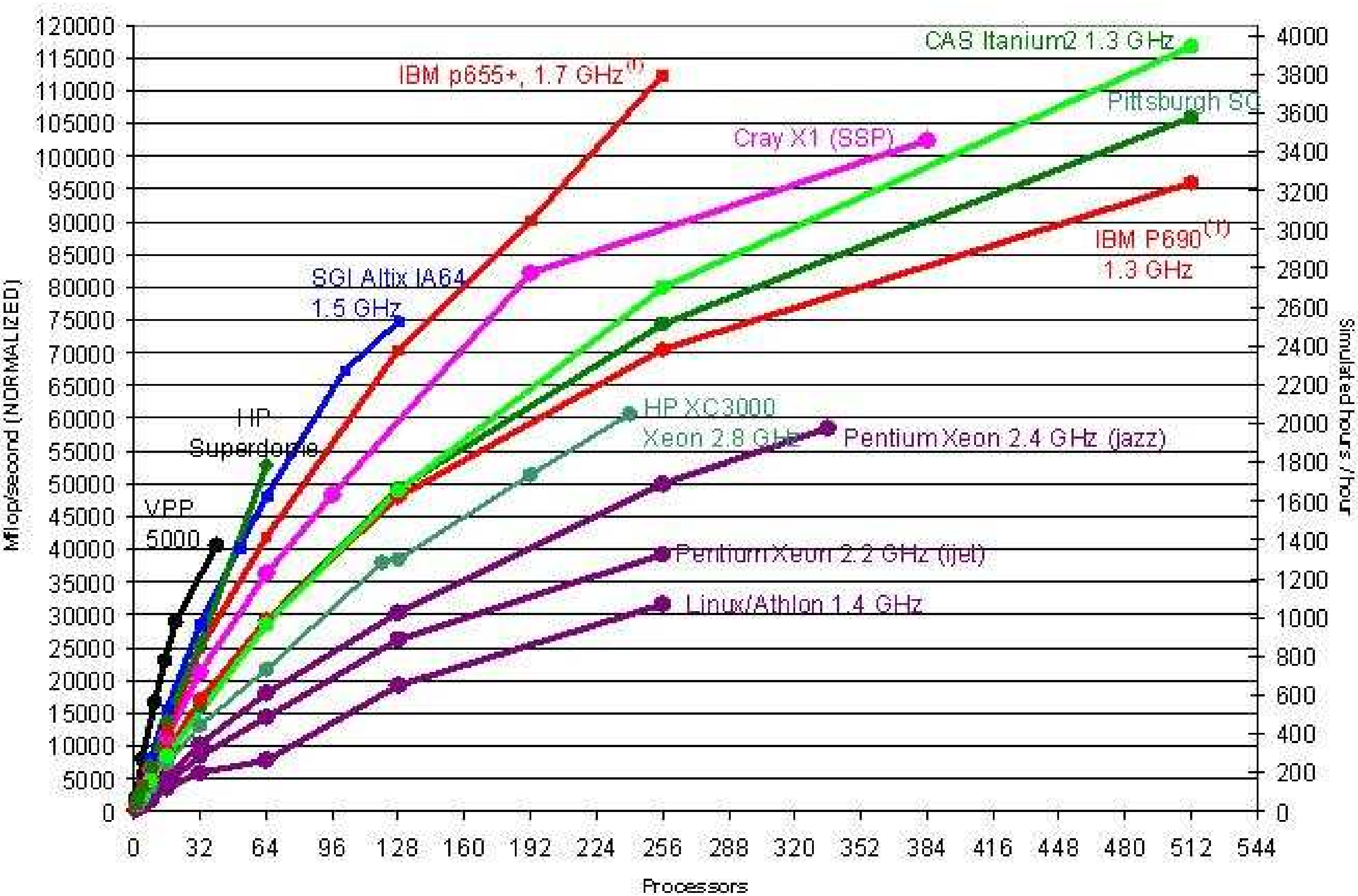
LM-RAPS 2.1
(Optimization using shared Memory)

Scalability on Altix using Intel Compiler and SGI's MPT library



Benchmark case used by Emy (Greece) in 2003
Grid 363x263x45

MM5 3.6.3 - Standard benchmark, 2004

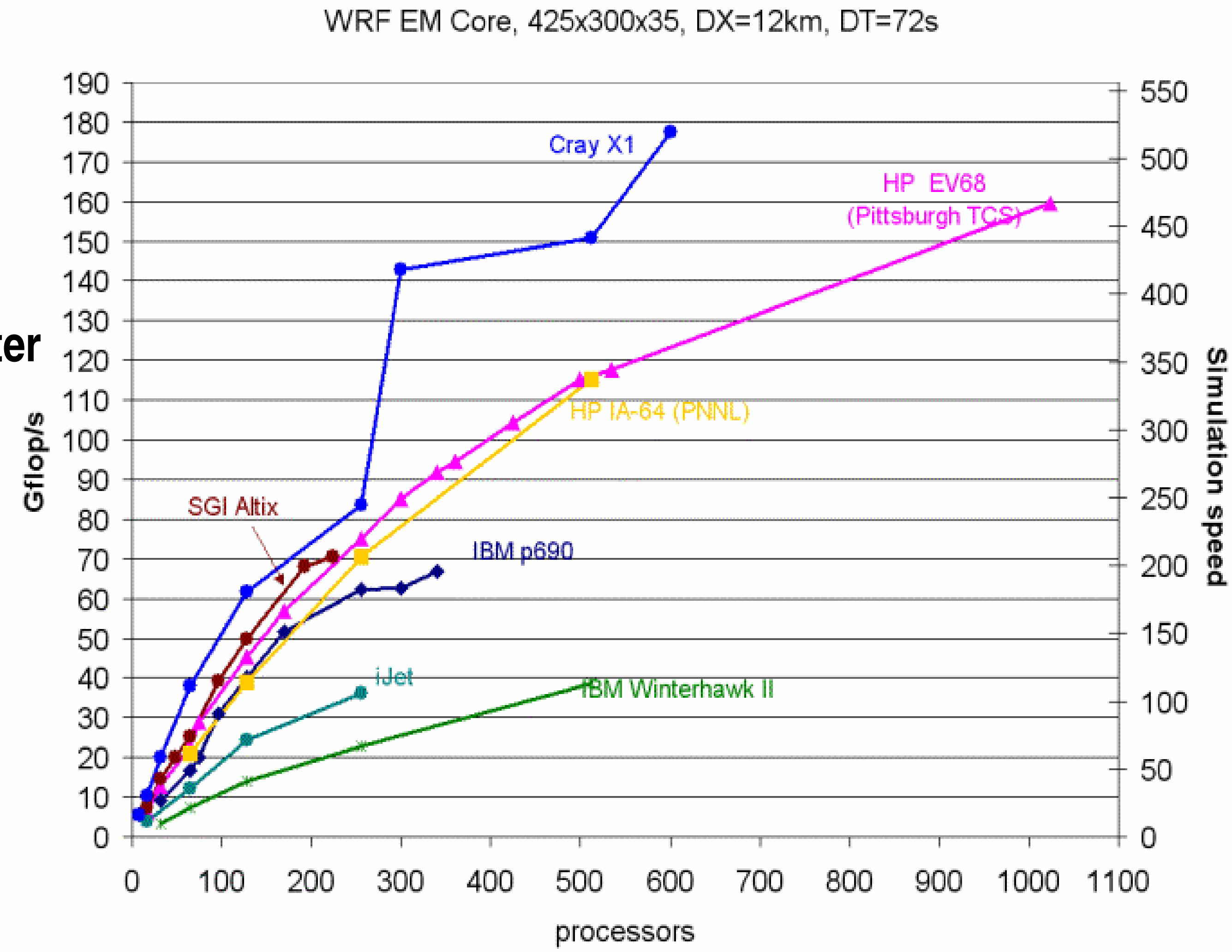


Altix version was compiled with the Intel 8.1.007 (beta) ifort Fortran compiler and the Intel 8.1.010 (beta) icc C compiler, and linked with SGI's MPI from MPT 1.10.

Source: http://www.mmm.ucar.edu/mm5/mpp/helpdesk/20040304a.htr

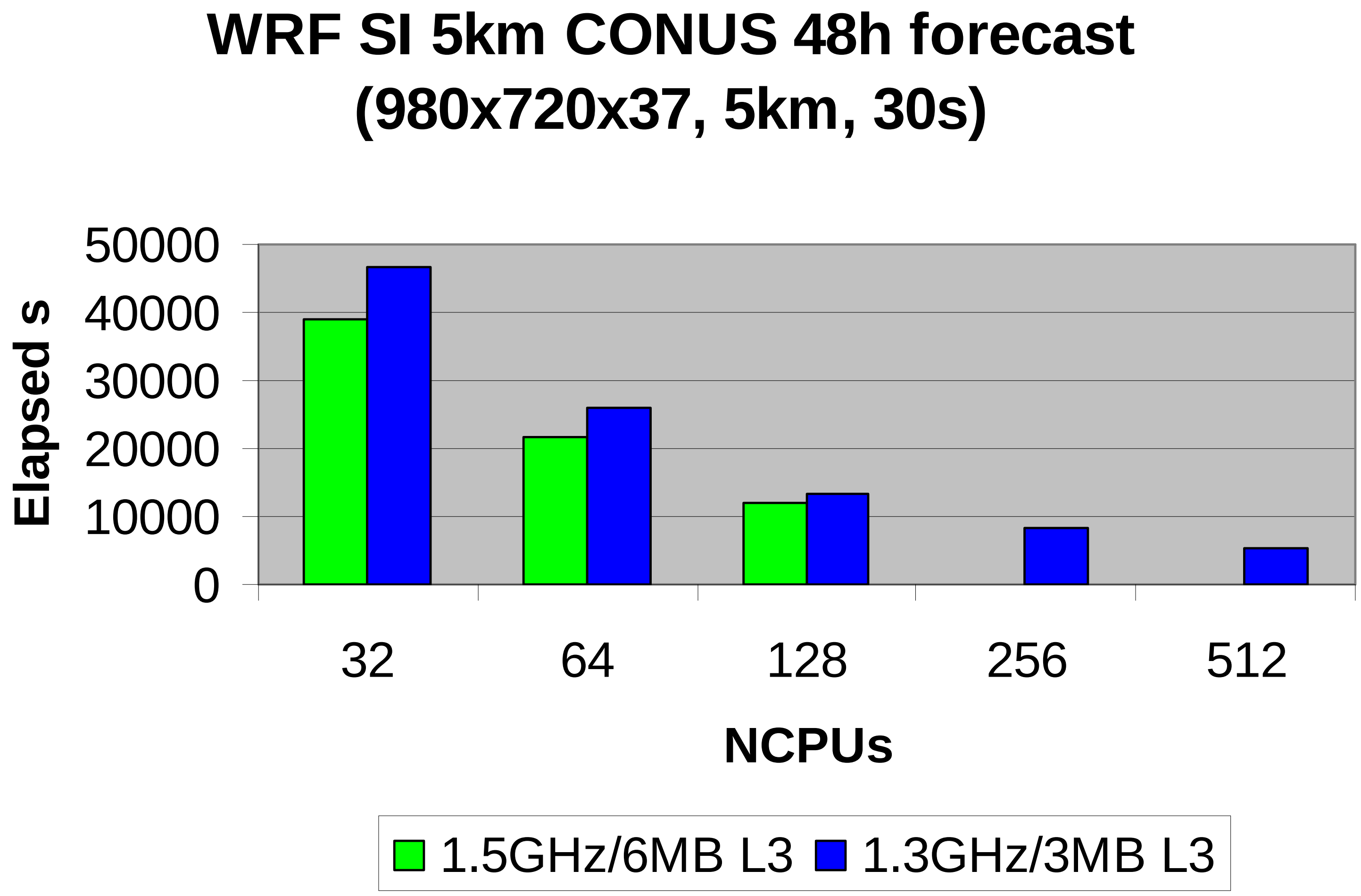
WRF 1.3 - Scalability & Performance on Altix 3000

Altix 3000 is 1.28x faster than HP cluster at the same clock frequency.

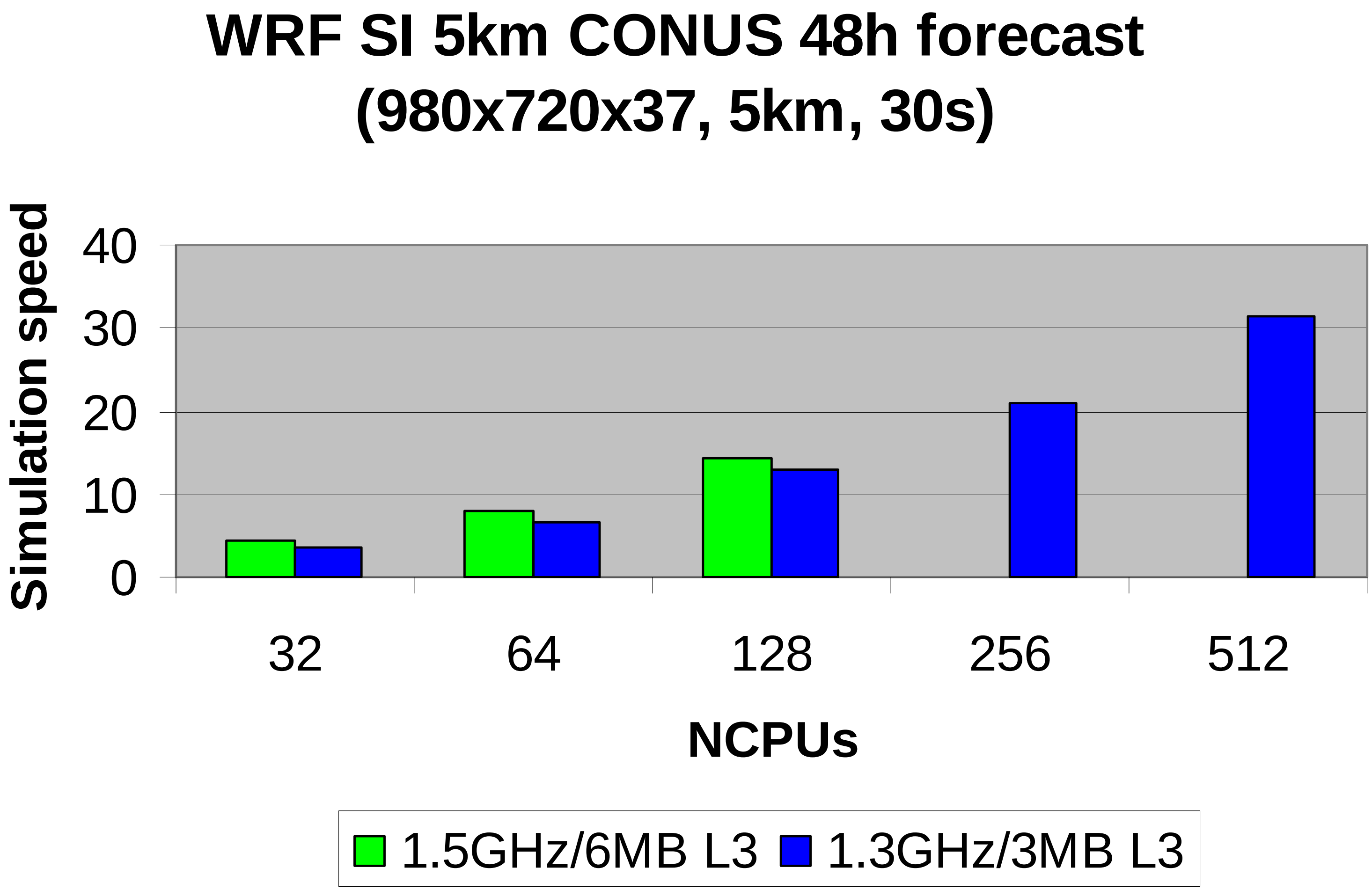


Source: http://www.mmm.ucar.edu/wrf/bench and Gerardo Cisneros, SGI

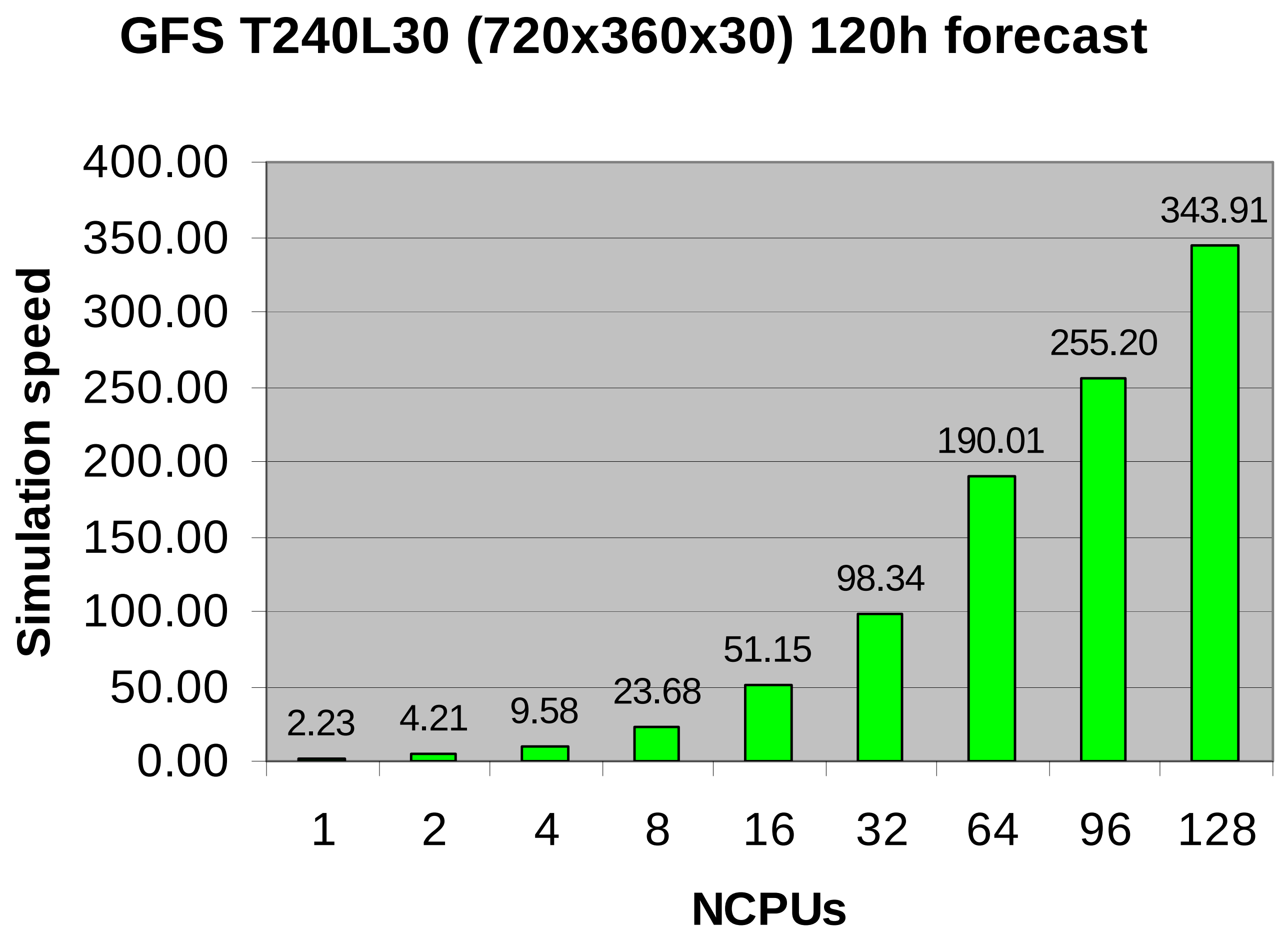
WRF 2.0.2 on a Large Problem



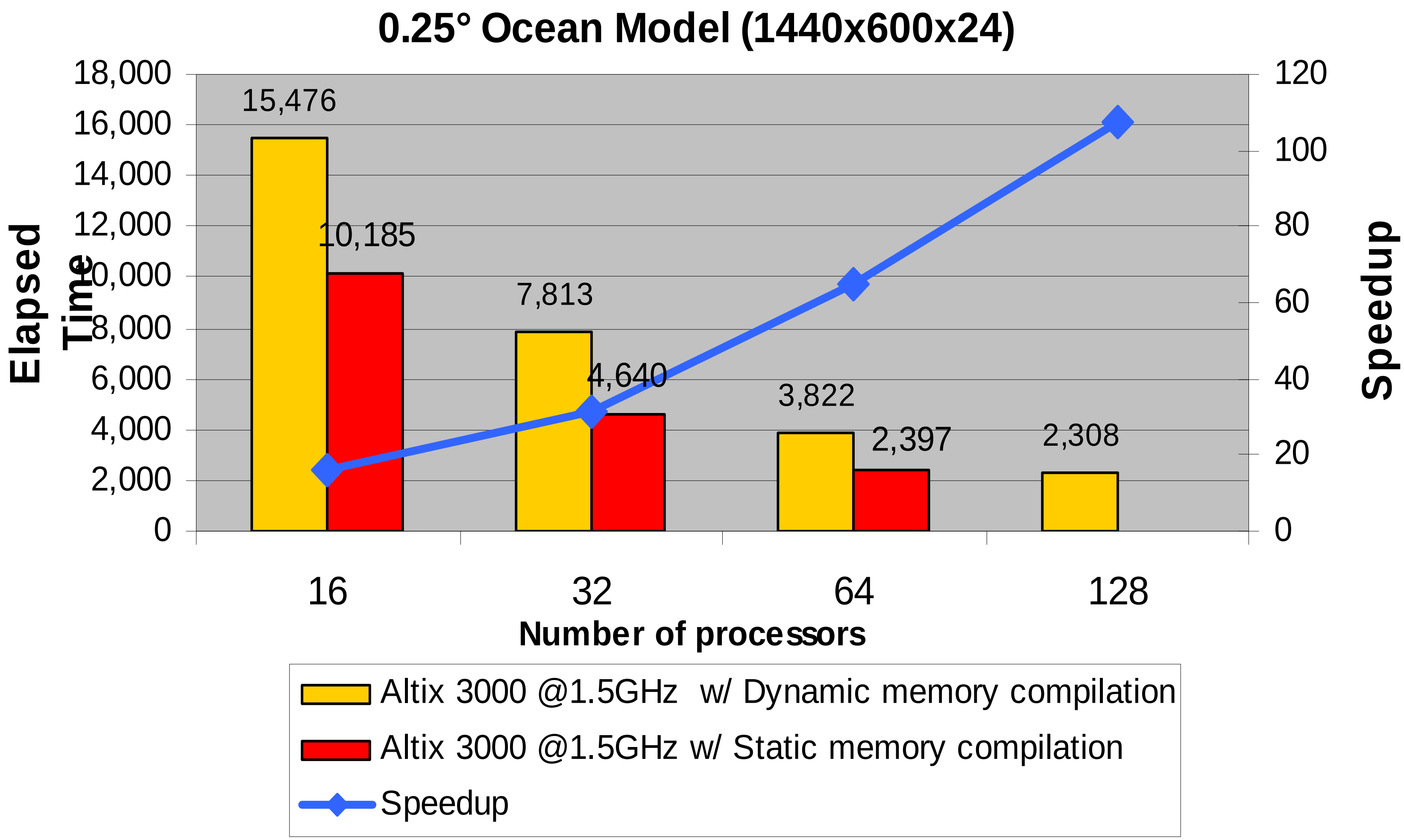
WRF 2.0.2 on a Large Problem



GFS Performance on the Altix 3700 (1.5GHz)



MOM4 - Performance on SGI Altix 3700

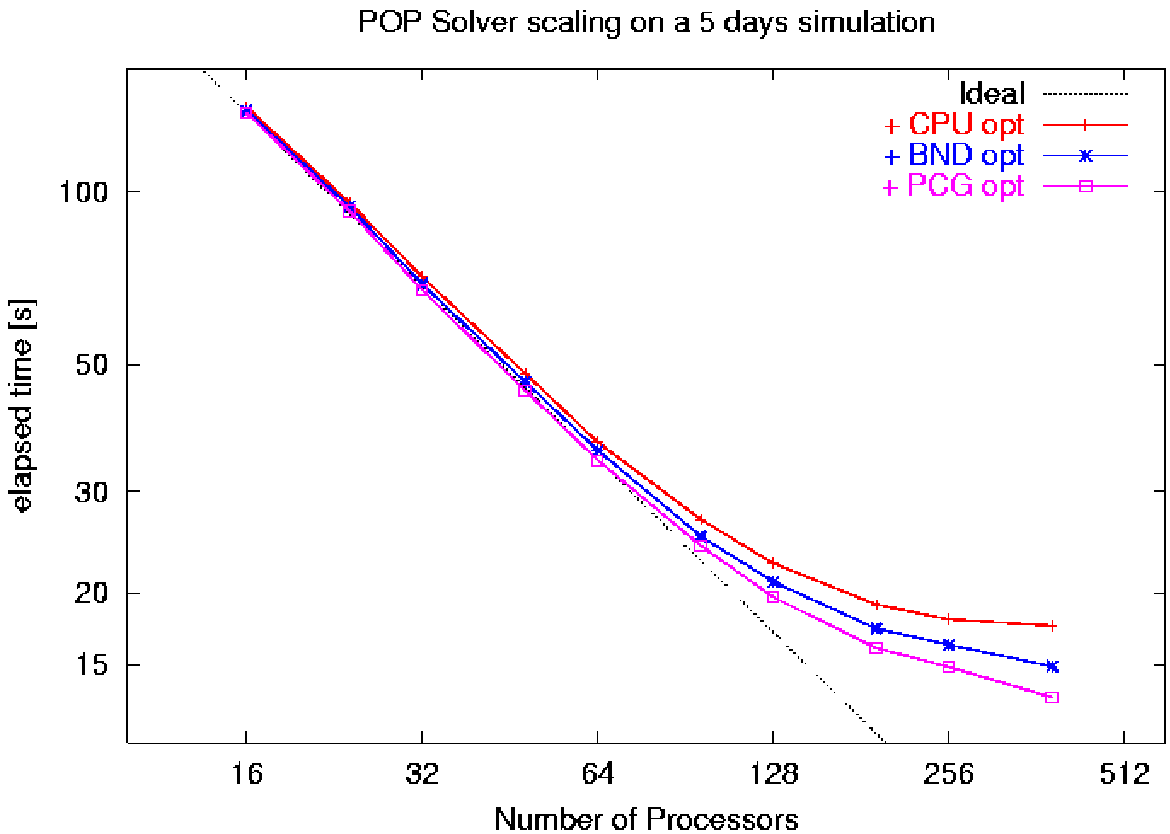


Source: Gerardo Cisneros, SGI, May 2004

POP 1.4.3

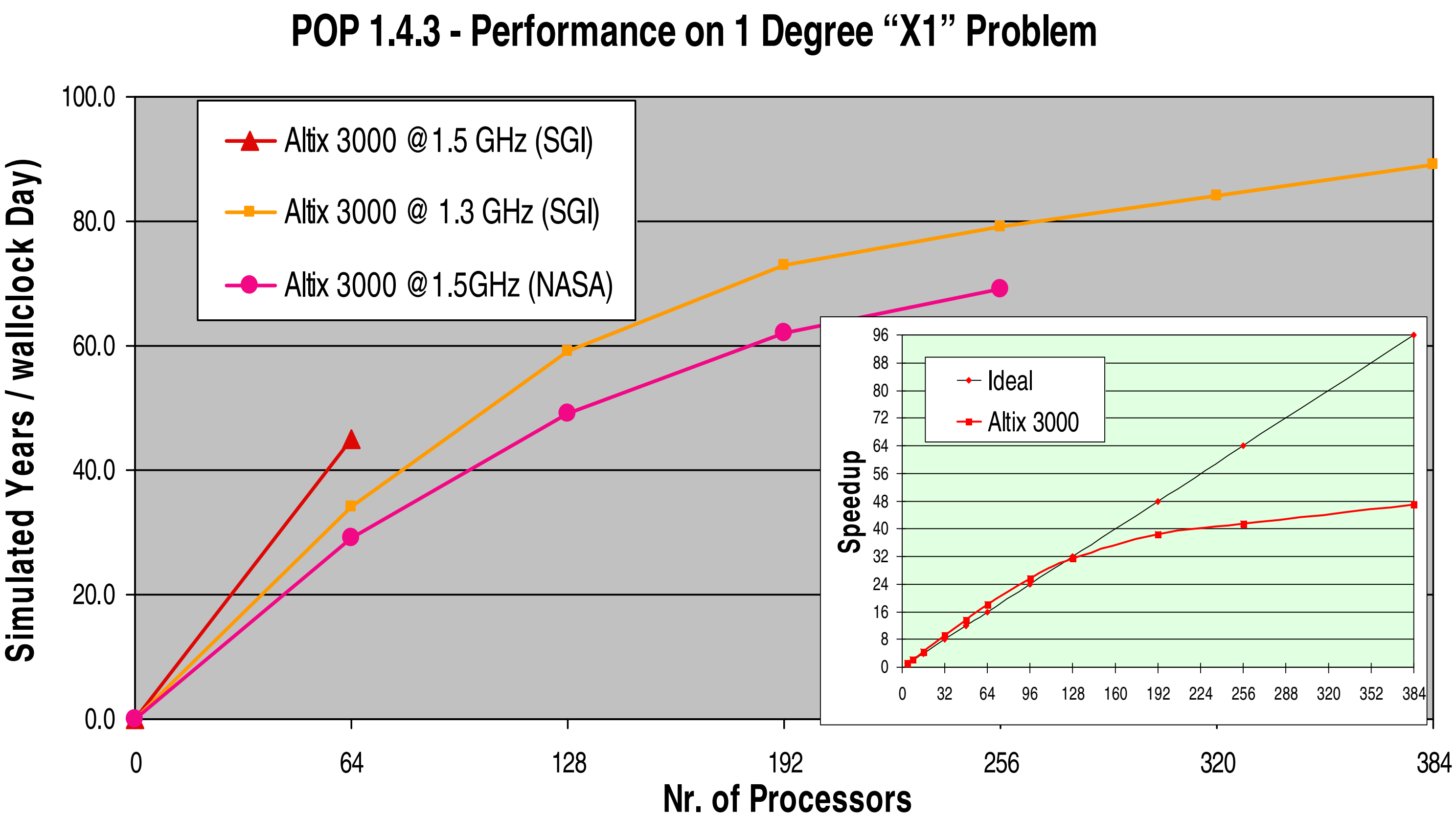
(Optimization by reducing the # of synchronizations)

Scalability on ALTIX using Intel Compiler and SGI's MPT library



Increasing Message length and reducing synchronization frequency improved the performance by 20% at 256 CPU

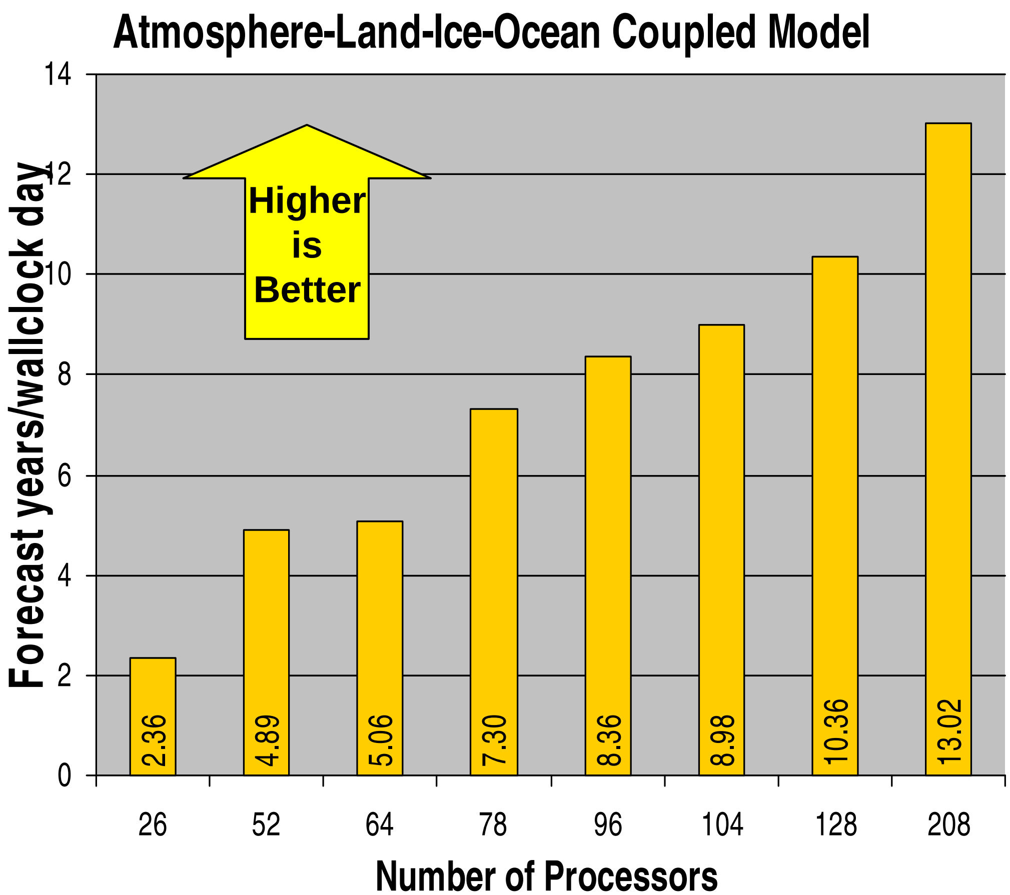
POP 1.4.3 - Performance on SGI Altix 3000 SSI

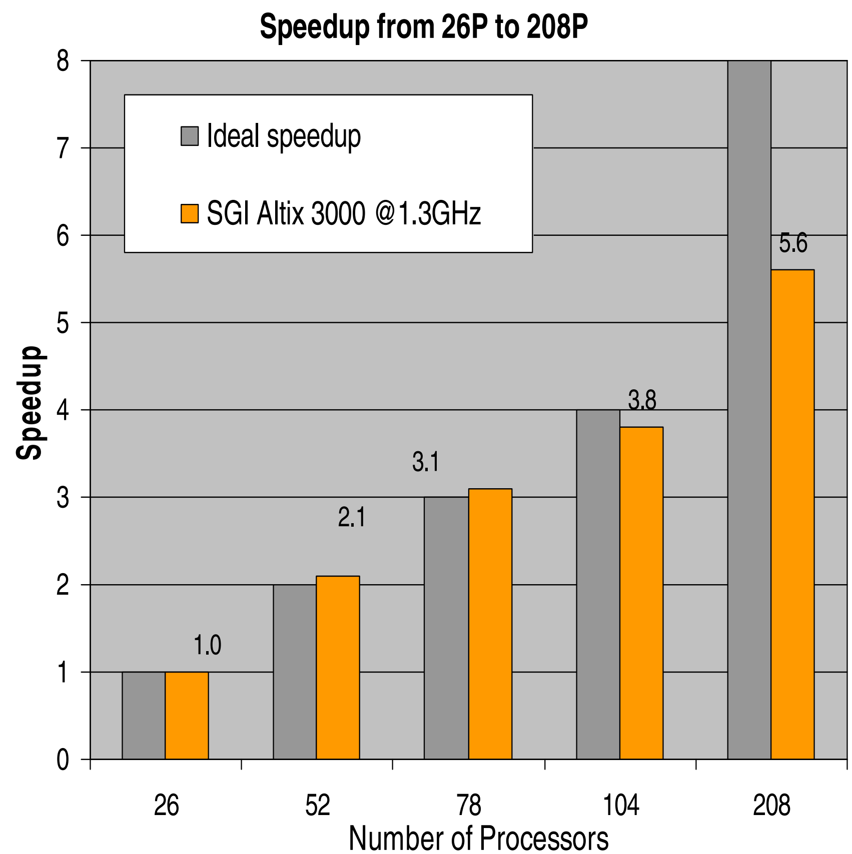


Configuration:
Altix 3000 @1.3GHz, 512P SSI,
Intel compiler 7.1.035 MPT 1.9

Source: Altix 3000 (SGI) - version optimized by Roland Richter using MPT
Altix 3000 (NASA) - Jim Taft's version using MLP

CCSM 3.0 (beta08) Performance on SGI Altix 3000





Run on 512P Altix 3000 @1.3GHz with efc 7.1.035 compiler

Source: Roland Richter, SGI

sgi

Conclusions

The Altix is proving to be an excellent system for weather and climate codes

- Shared memory allows very large models to run
- The compilers are getting better
- The libraries are excellent and still improving
- SGI's software layers on top of Linux help jobs attain their best performance
- SGI engineers have the know-how to make weather and climate codes run well on its customers' SGI systems

sgi®