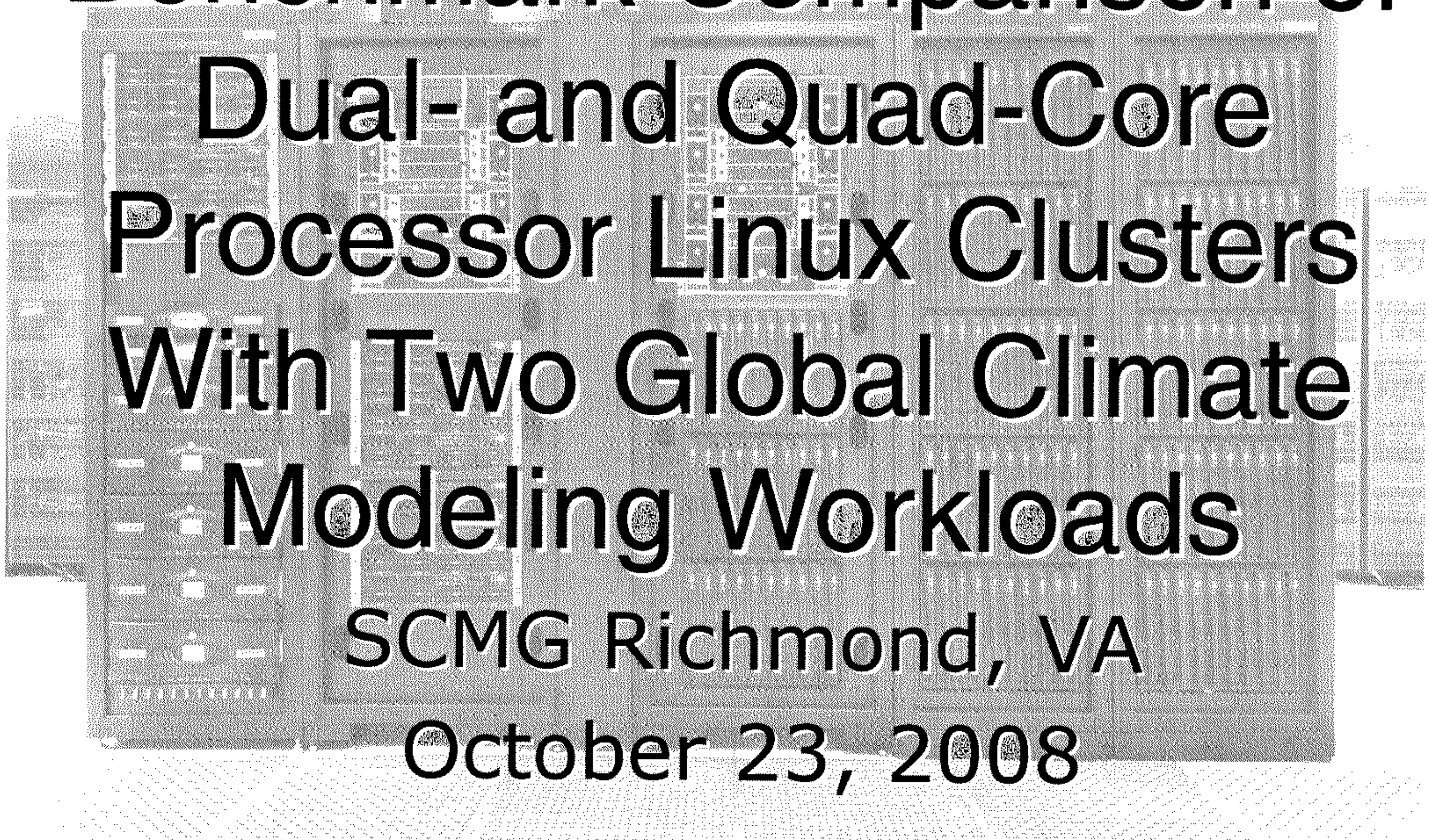
Benchmark Comparison of Dual- and Quad-Core Processor Linux Clusters With Two Global Climate Modeling Workloads

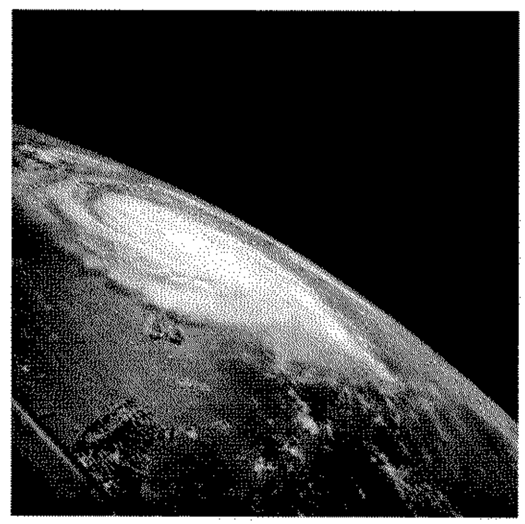


SCMG Richmond, VA

October 23, 2008

Presentation Overview

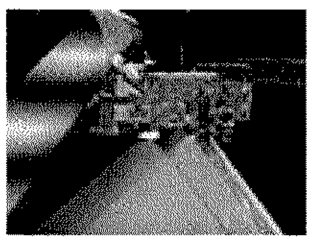
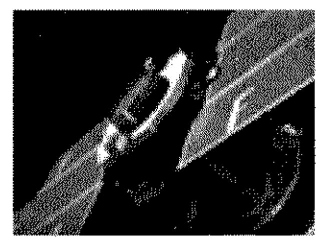
- Science Environment
- System Environment
- Workload
- Benchmark Results & Discussion
- Concluding Remarks

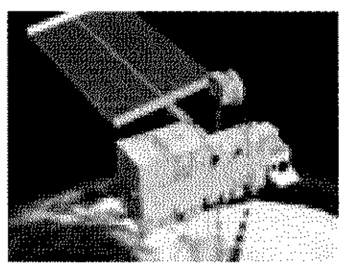


Science Environment

- NASA's High End Computing Program supports two supercomputer facilities.
- The NASA Center for Computational Sciences is located on the Goddard Space Flight Center in Greenbelt, MD.
- The Ames Research Center is located in Sunnyvale, CA.
- The primary processing platforms at both facilities, currently, are Linux clusters using multi-core Intel-architecture microprocessors.

Science Environment - Goddard

- NASA Goddard is the world's largest organization of Earth scientists and engineers. Goddard designs, builds and operates approximately 60 spacecraft, including Earth-observing satellites, such as Aqua  , Cloudsat 

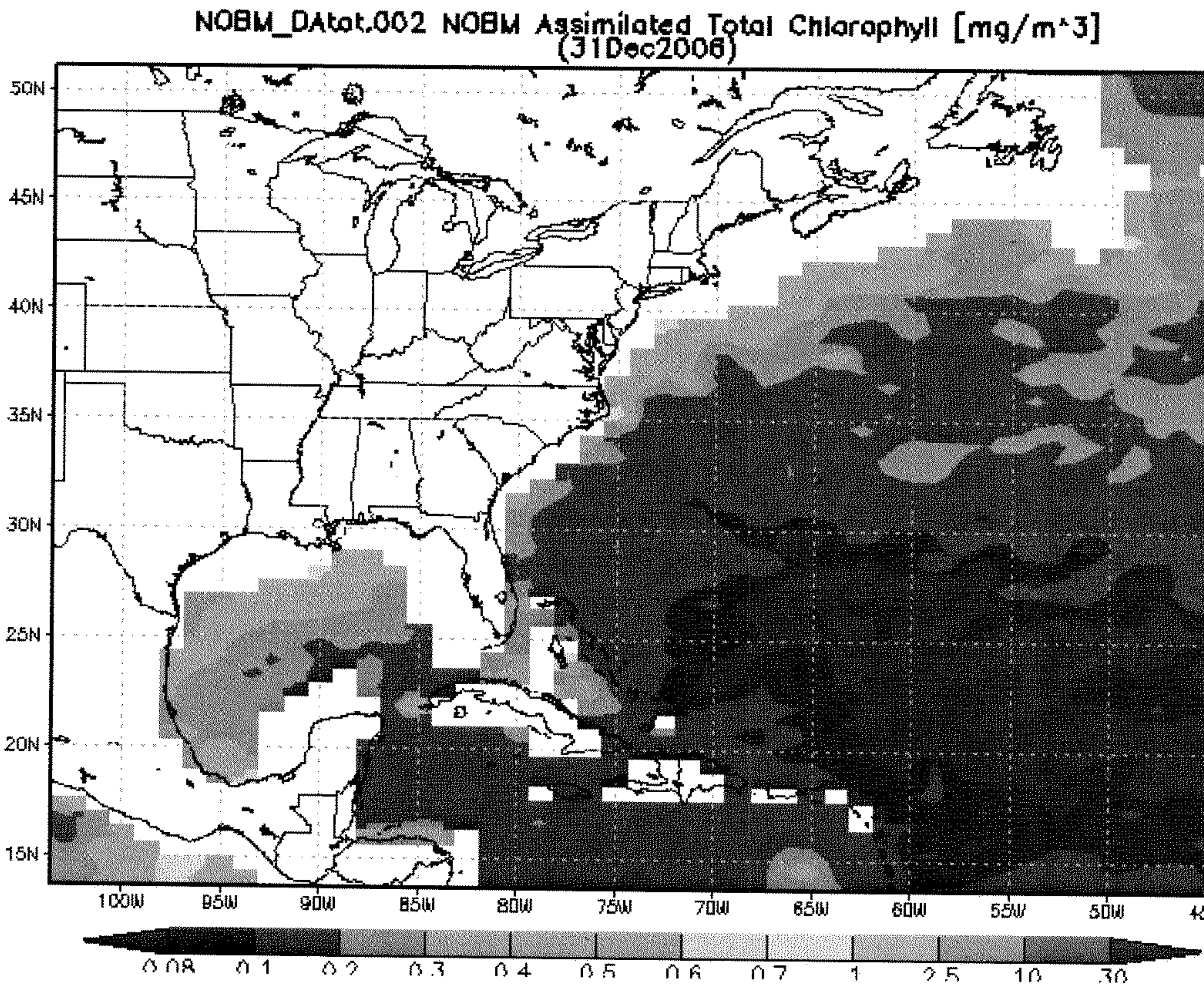
  and Terra 

Science Environment - Goddard

- ....and receives, stores, processes and distributes the data that their instruments transmit.
- Go to

  http://daac.gsfc.nasa.gov/techlab/giovanni/index.shtml

  For the Giovanni tool, which allows anyone to draw maps of selected Earth-observation data sets for user-selected time periods and areas of the globe.


NOBM_DAtot.002 NOBM Assimilated Total Chlorophyll [mg/m^3]
(31Dec2006)
50N
45N
40N
35N
30N
25N
20N
15N
100W
95W
90W
85W
80W
75W
70W
65W
60W
55W
50W
45W
0.08
0.1
0.2
0.3
0.4
0.5
0.6
0.7
1
2.5
10
30

Science Environment - Goddard

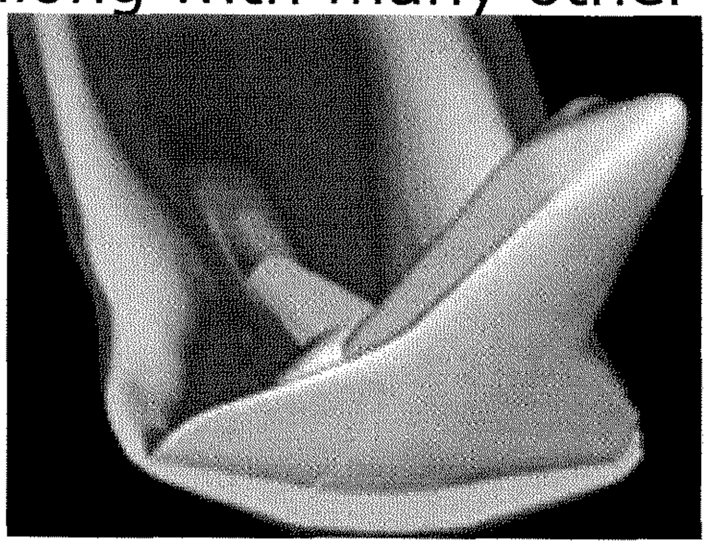
- Goddard supports climate- and weather-forecasting research and production. The NCCS' single largest user (in terms of processing hours) is the Global Modeling and Assimilation Office (GMAO).
- GEOS-5 is GMAO's production assimilation model and one source for the natural benchmarks ("Lat-Lon") in this presentation.

Science Environment – Goddard - GEOS-5

- The Goddard Earth Observing System Model, Version 5 (GEOS-5) is a system of models integrated using the Earth System Modeling Framework (ESMF). The GEOS-5 systems are being developed in the GMAO to support NASA's earth science research in data analysis, observing system modeling and design, climate and weather prediction, and basic research.

Science Environment - Ames

- NASA Ames expertise includes aerodynamics and other disciplines. Computer simulation has replaced physical wind tunnels to study the behavior, e.g., of the Space Shuttle on re-entry into the atmosphere. NASA uses Ames' systems to predict the flight characteristics and stability of the shuttle along with many other computational tasks.



Systems Environment

- The primary platforms at both Goddard and Ames include cluster supercomputers running the Linux operating system and using Intel X86 architecture microprocessors.

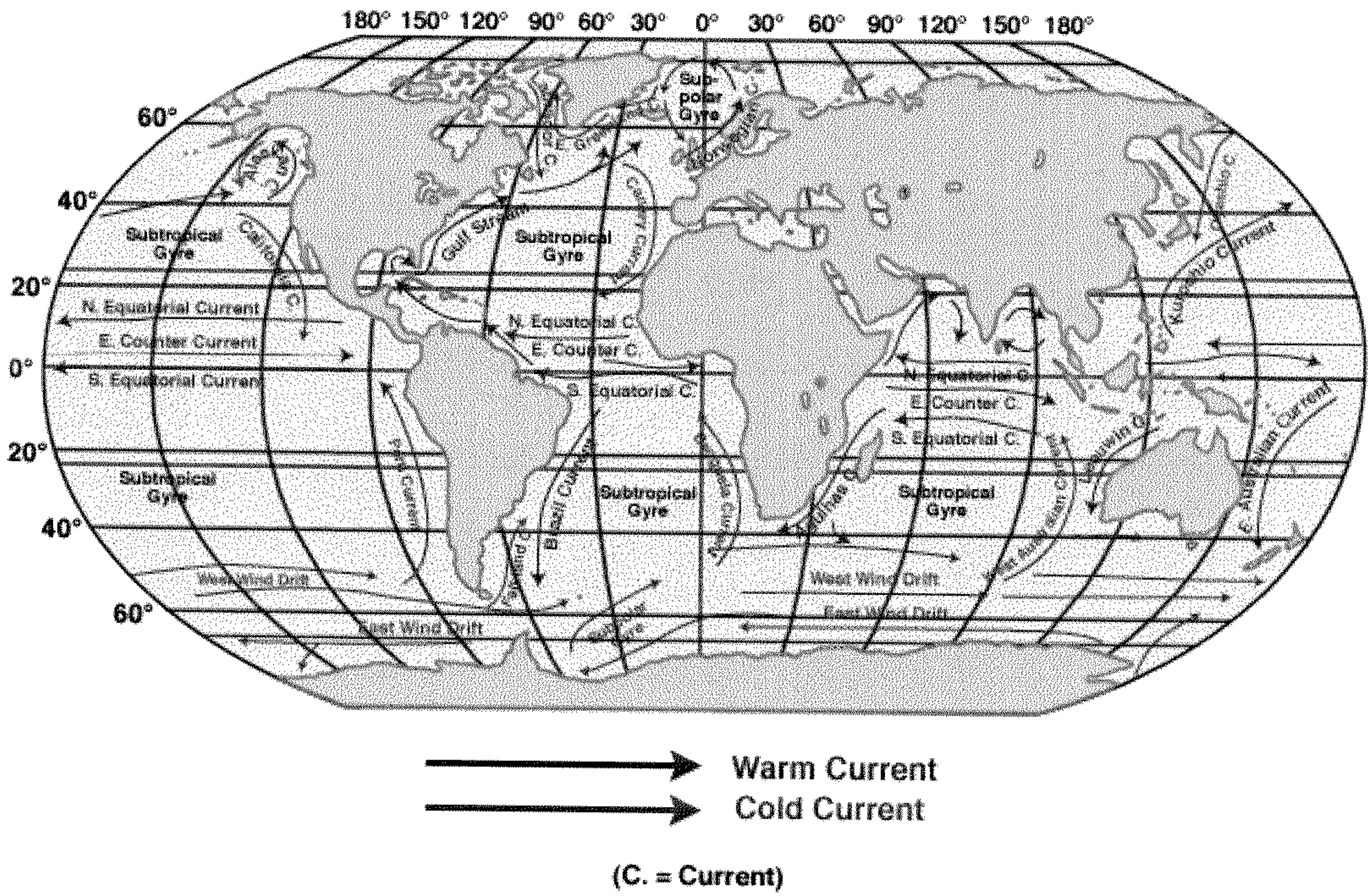
Systems Environment

Site	Goddard	Goddard	Goddard	Ames
System	Discover - Base	Discover - SCU 1&2	Discover - SCU 3&4	RTJones
CPU	Intel 5060 (Dempsey)	Intel 5150 (Woodcrest)	Intel 5420 (Harpertown)	Intel 5355 (Clovertown)
Clock - GHz	3.2	2.66	2.5	2.66
Release Date	May 06	June 06	Nov 07	Nov 06
MB L2 Cache/Core	2	2	3	4
Flops/Clock	2	2	4	4
Cores/Socket	Dual	Dual	Quad	Quad
Nodes/System	128	512	512	512
Total Cores	512	2048	4096	4096
Peak TF Calc	3.278	10.8954	40.96	43.5
GB Memory/Core	0.6	0.6	2	1
Front Side Bus MHz	1066	1066	1333	1333
Switch	Infiniband	Infiniband	Infiniband	Infiniband
OS	SUSE Linux	SUSE Linux	SUSE Linux	SUSE Linux
Scheduler	PBS	PBS	PBS	PBS
MPI	Scali-MPI	Scali-MPI	Open MPI 1.2.5	MVAPCH or MPT
Compiler	Intel Fortran 10.1.013	Intel Fortran 10.1.013	Intel Fortran 10.1.013	Intel Fortran 10.1.013
Manufacturer	LNXI	LNXI	IBM	SGI

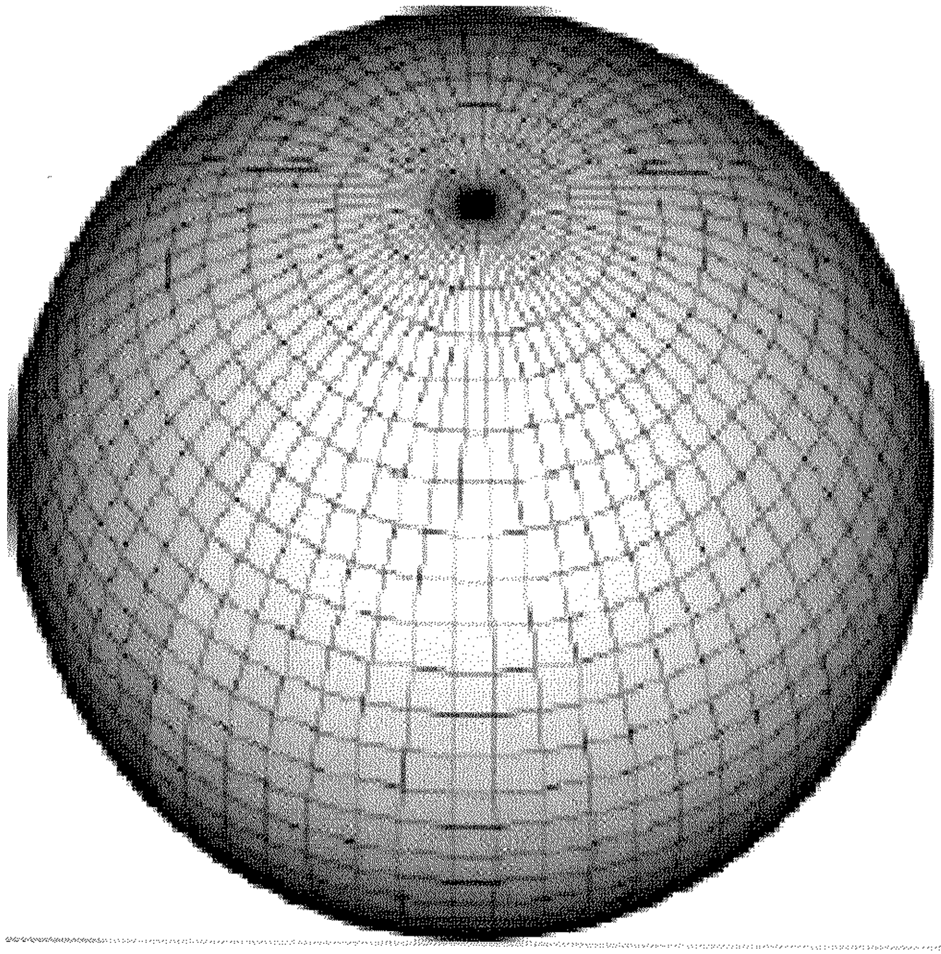
Workload

- Lat-Lon vs. Cubed Sphere
- Lat-Lon. The Earth's atmosphere is mapped into a three-dimensional grid. Each cell is ½ degree east-to-west, ½ degree north-to-south, and 1/72 of the distance from sea level to the top of the atmosphere.
- The time step simulated is a function of the cell size. With smaller cells, weather phenomena such as wind carry over cell boundaries more quickly, so shorter time steps are needed with smaller cell sizes. This is called the Courant condition.

Lat/Lon Division



Lat/Lon Division



Workload Lat/Lon

- Because the atmosphere is so shallow compared to its lateral dimensions (north-south and east-west), usually the number of levels does not change as cell sizes shrink.
- The Lat-Lon mapping has been in common use by weather and climate modelers for many years, with decreasing cell sizes and an increase in the number of levels over time as more powerful systems become available to process them. In general, smaller cells produce more accurate predictions.

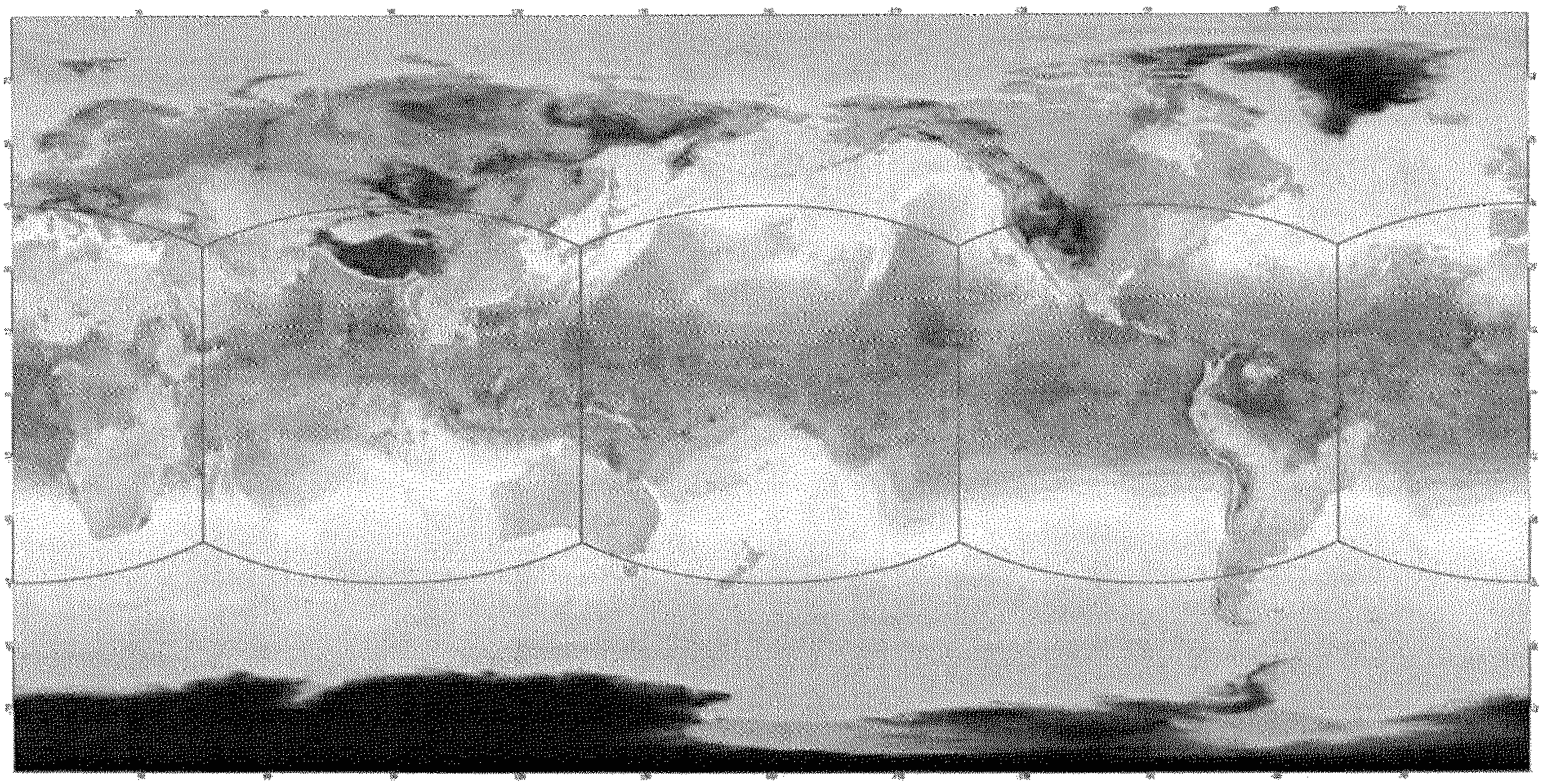
Workload – Lat/Lon

- Reducing cell width laterally by one half requires about a one order of magnitude increase in processing power – 2x (east-west), 2x (north-south), 2x (time step). Vertical levels are unchanged.
- One problem with lat-lon mappings occurs at the poles. Cells become small close to the north and south poles as longitudinal lines (cell boundaries) converge. Due to the Courant condition (smaller cells need shorter time steps), the poles require special treatment and ultimately limit the utility of Lat-Lon models.

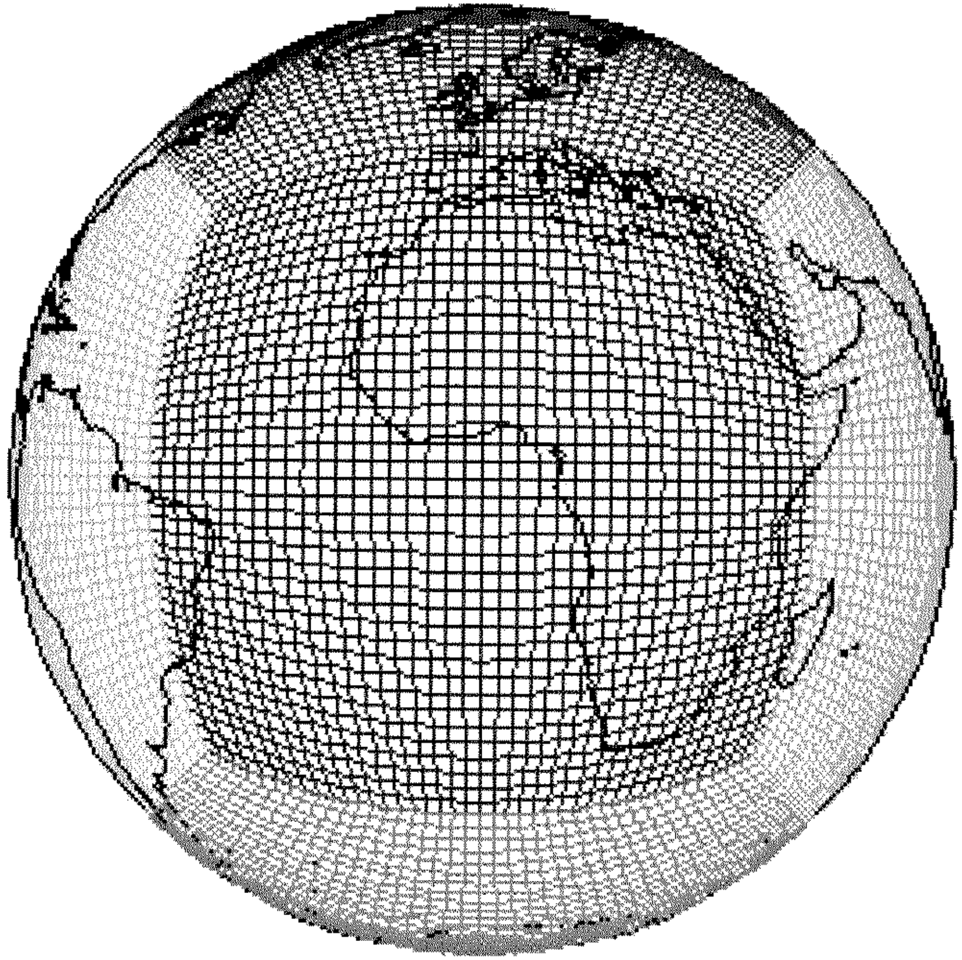
Workload – Cubed Sphere

- Cubed Sphere. The cubed sphere maps the Earth's (nearly) spherical surface and atmosphere onto a cube. Imagine a point source (a small light bulb) shining through the Earth's spherical surface and projected onto a cube that completely encloses the sphere. This projection, although not familiar to elementary school students, avoids the problem of converging cell boundaries and the Courant condition. Cell dimensions east to west don't shrink to nothing at the poles.

Workload – Cubed Sphere

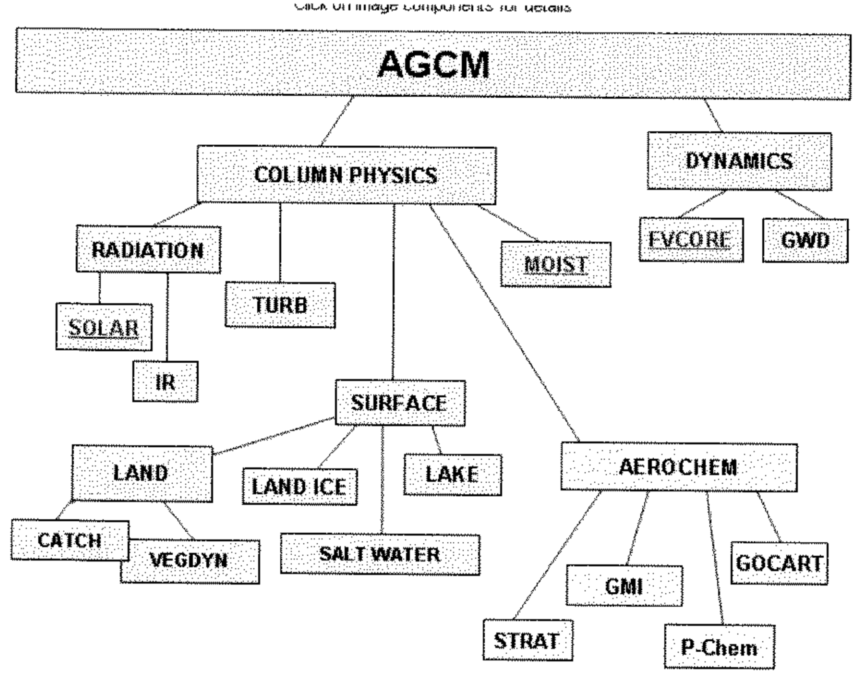


Workoad – Cubed Sphere

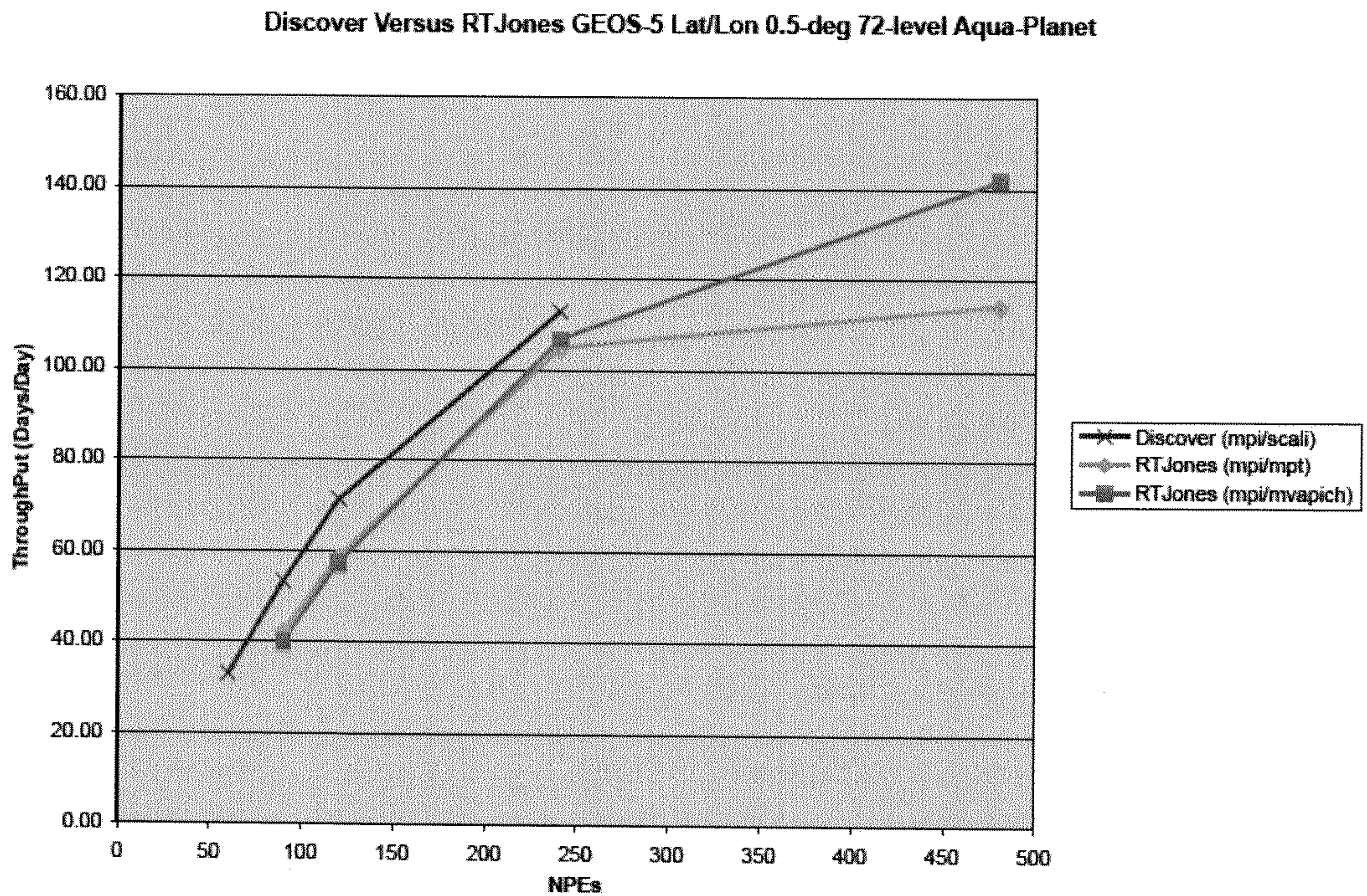


Workload

- Dynamics (e.g., wind, air pressure) – about half of processing time. Generally more parallel. Runs more often.
- Physics (e.g., heat, humidity & precipitation, topography, turbulence, chemistry)



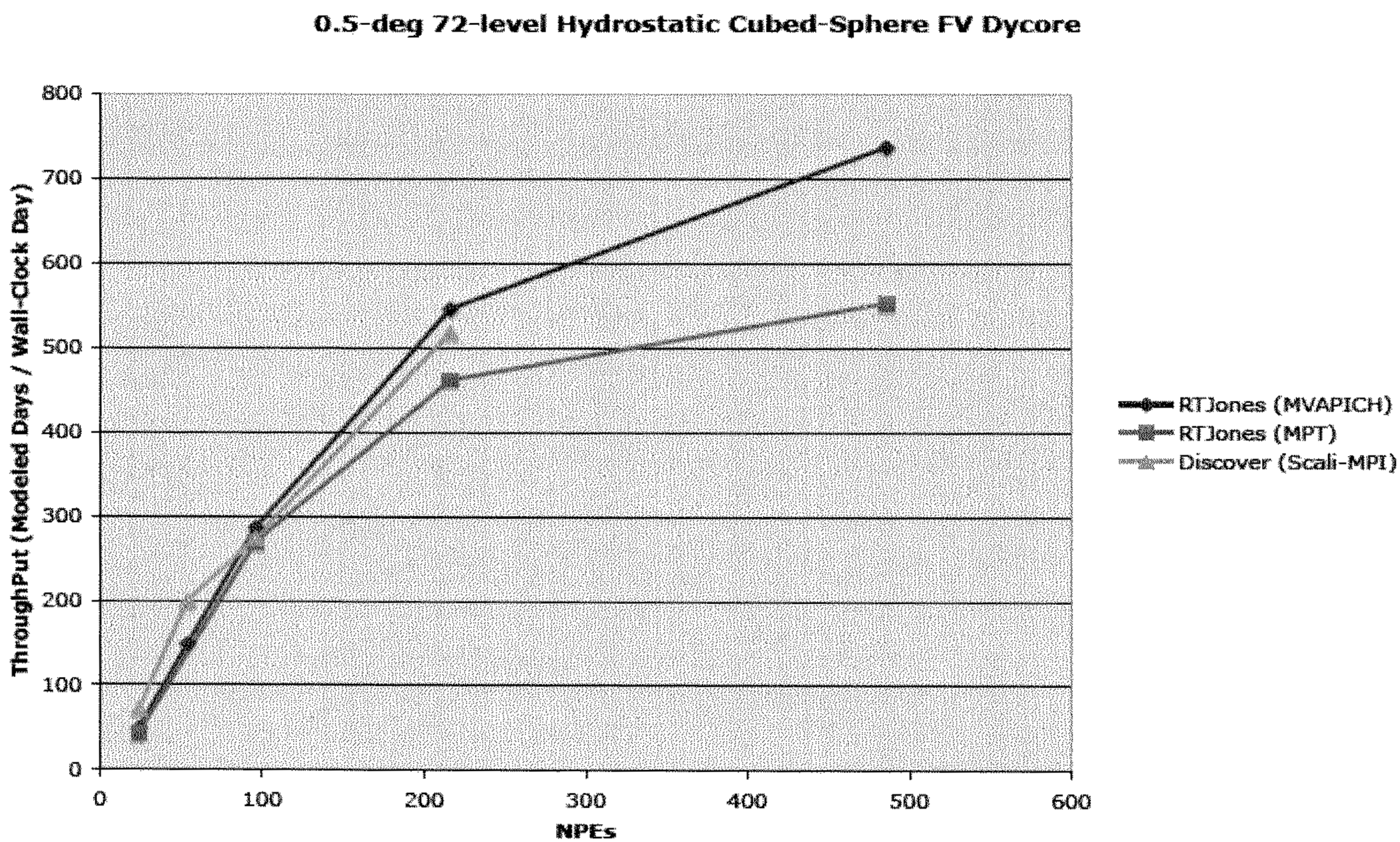
Benchmark Results – GEOS-5



Benchmark Results – GEOS-5

- Discover vs. RTJones. Two MPI versions on RTJones.
- GEOS-5 workload – dynamics and physics. Cells .5 degrees and 72 vertical atmospheric levels.
- Horizontal Axis – number of processor elements – the same for all charts
- Vertical Axis – Simulated days per wall-clock day. GEOS-5 is production system, runs every 6 hours and is used by weather forecasters and others.

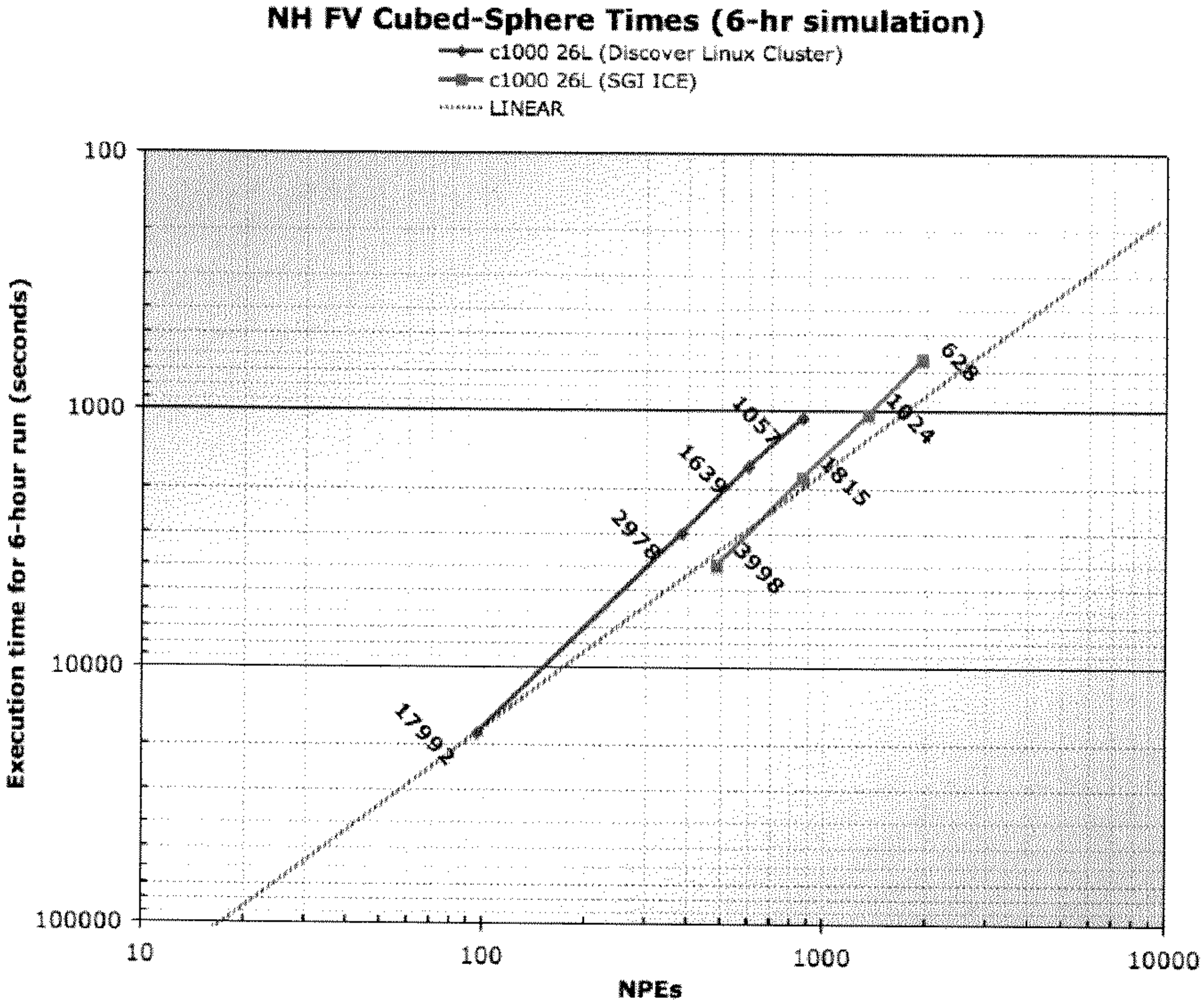
Benchmark Results – Cubed Sphere



Benchmark Results – Cubed Sphere

- Cubed Sphere model, same resolution as the Lat-Lon model
- Dynamics only
- Vertical Axis - days per day - but larger range

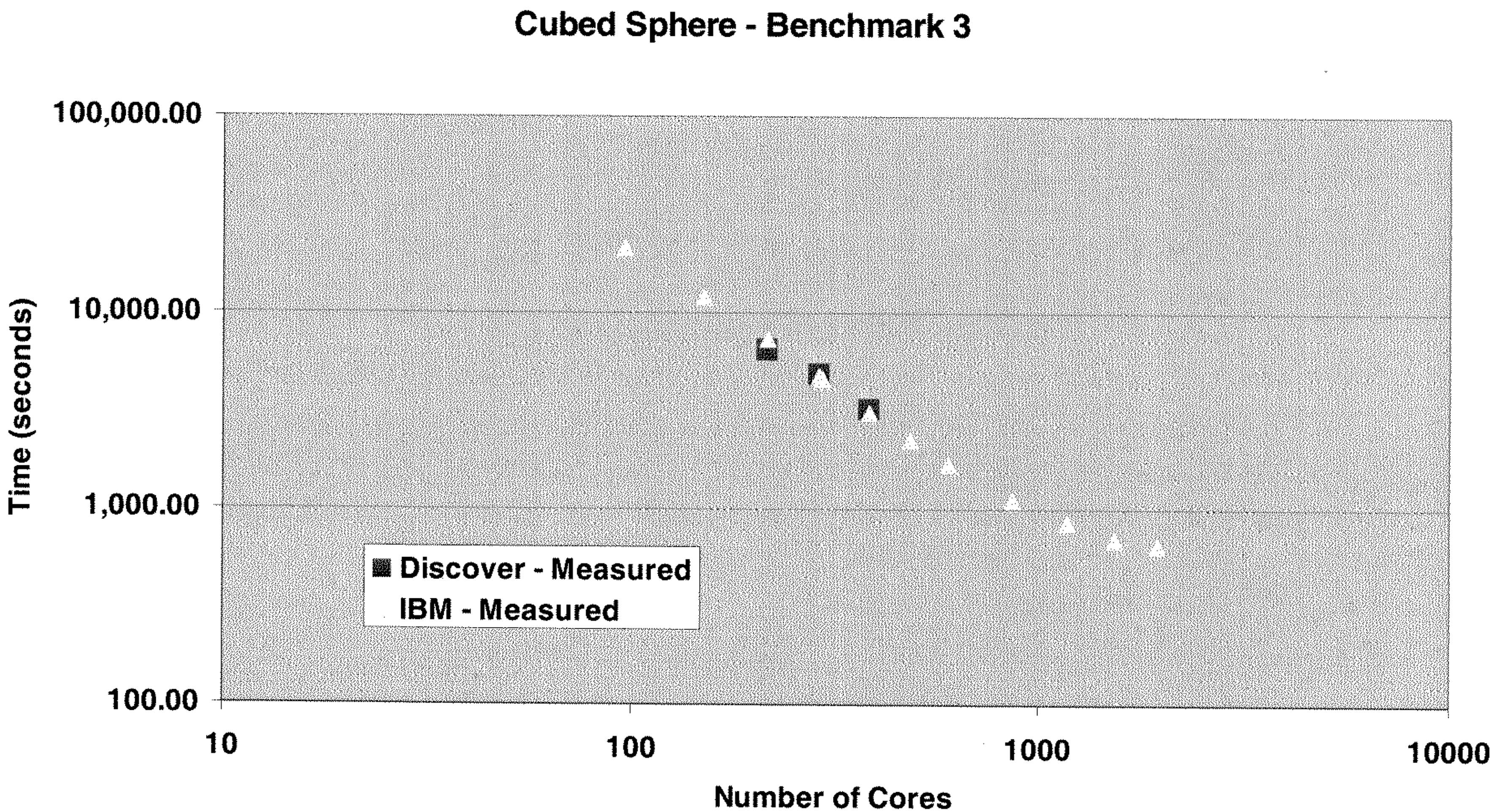
Benchmark Results – Cubed Sphere



Benchmark Results – Cubed Sphere

- Cubed-Sphere model, different resolution
- Horizontal and Vertical Axes have logarithmic scales
- Vertical Axis is execution time
- Fixed Workload
- Results plotted against linear speedup

Benchmark Results – Discover – LNXI & IBM



Benchmark Results – Discover – LNXI & IBM

- Compares performance for two incremental upgrades to the Discover system.
- Two processor models – "Woodcrest" (dual core) and "Harpertown" (quad core)
- Log/log scale was with previous chart
- Fixed workload and vertical axis is execution time

Discussion

- Processors
- Cache
- Message Passing
- Memory
- I/O

Concluding Remarks

- Use of benchmark tests in computer acquisitions - particularly natural benchmarks and also custom synthetic benchmarks – by U.S. Government agencies has declined significantly since the early 90s.
- This is mostly due to two factors. One, continuing improvements in processor price/performance compared to the significant expense of benchmark development and execution.
- Two, procurement reform in the middle 90s that reduced the incidence of disputes with vendors over acquisition processes and competition.

Concluding Remarks

- It's interesting to note that the current high performance computing marketplace in some ways resembles the market for IBM -compatible mainframe computers in the 80s and early 90s.
- Intel x86 architecture dominates – which means the same instruction set across vendors.
- The same operating system – Linux.
- The same compilers, libraries and much of the other system software.
- Processor workload dominates, allowing I/O subsystem impacts to be ignored.

Concluding Remarks

- NCCS' specific circumstances made a natural benchmark a good fit for the most recent acquisition. (Note: NCCS also used standard synthetic benchmarks, but did not weight them as heavily.)
- Processor workload dominates.
- A key user/workload – GMAO.
- Batch mode packaging and execution is cheaper and easier for all parties than, say, an interactive benchmark using remote terminal emulation.

Concluding Remarks

- Large latent demand – user scientists can profitably employ ever larger/faster systems to get better science results. E.g., finer grids with smaller cells and shorter time steps yield better science results but require faster computers.
- Workload lends itself to parallel processing.
- Benchmark results also support system contract administration – installed hardware has to meet proposed numbers or the vendor must fix.

Concluding Remarks

- A common criticism of kernel benchmarks is too much reference locality. These results show that a realistic memory footprint can be crucial to discriminating system performance.
- If clock speed increases are no longer feasible, increasing parallelism is crucial to system performance increases. But increasing hardware parallelism has consistently outpaced software's ability to exploit it.

Acknowledgements

- Dr. Daniel Duffy – Computer Sciences Corporation – NCCS Benchmark Results and Analysis
- Drs. William Putnam and Thomas Clune – NASA – Ames and NCCS Benchmark Results and Analysis

Bibliography

- William Putnam and others. "The Finite Volume Dynamical Core on the Cubed Sphere"
- Glassbrook and McGalliard, "Performance Management at an Earth Science Supercomputer Center " CMG 2003.
- Spear and McGalliard, "A Queue Simulation Tool for a High Performance Scientific Computing Center" CMG 2007.

Webliography

- http://www.nccs.nasa.gov/ (NCCS)
- http://www.nccs.nasa.gov/systems.html (Discover)
- http://www.nasa.gov/centers/ames/home/index.html (Ames)
- http://www.nas.nasa.gov/Resources/Systems/rtjones.html (RTJones)
- http://gmao.gsfc.nasa.gov/index.php (GMAO)
- http://gmao.gsfc.nasa.gov/systems/geos5/ (GEOS-5)