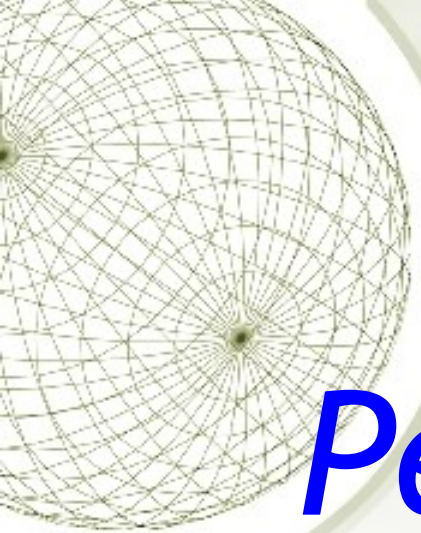


Solving a Blueice Performance Mystery

Wei Huang

Siddhartha Ghosh

Consulting Service Group (CSG)

High-End Services Section (HSS)

CISL/NCAR

Bluevista vs Blueice

--http://www.cisl.ucar.edu/docs/products/dsm.cpu.specs.html

	Bluevista: IBM Power5	Blueice: IBM Power5+
Clock Speed/time	1.9 GHz, 0.5262ns	1.9 GHz, 0.5262ns
Die Type	Single Core	Dual Core
Cache Size	L1 -- 96KB per core, 32-KB data cache, 64-KB instruction cache, 128-B cache line. 4-way associative L2 -- 1.92MB, 128-B line, 10-way set-associative. L3 -- 36MB, 12-way set-associative. 128-B cache lines	L1 -- 96KB per core, 32-KB data cache, 64-KB instruction cache, 128-B cache line. 4-way associative L2 -- 1.92MB, 128-B line, 10-way set-associative. L3 -- 36MB, 12-way set-associative. 128-B cache lines
Page Size	4KB	4 KB or 64 KB Settable by User
Procs per Node	8	16
Memory per Node	16 GB	32/64 GB
Total Batch Nodes	72	100

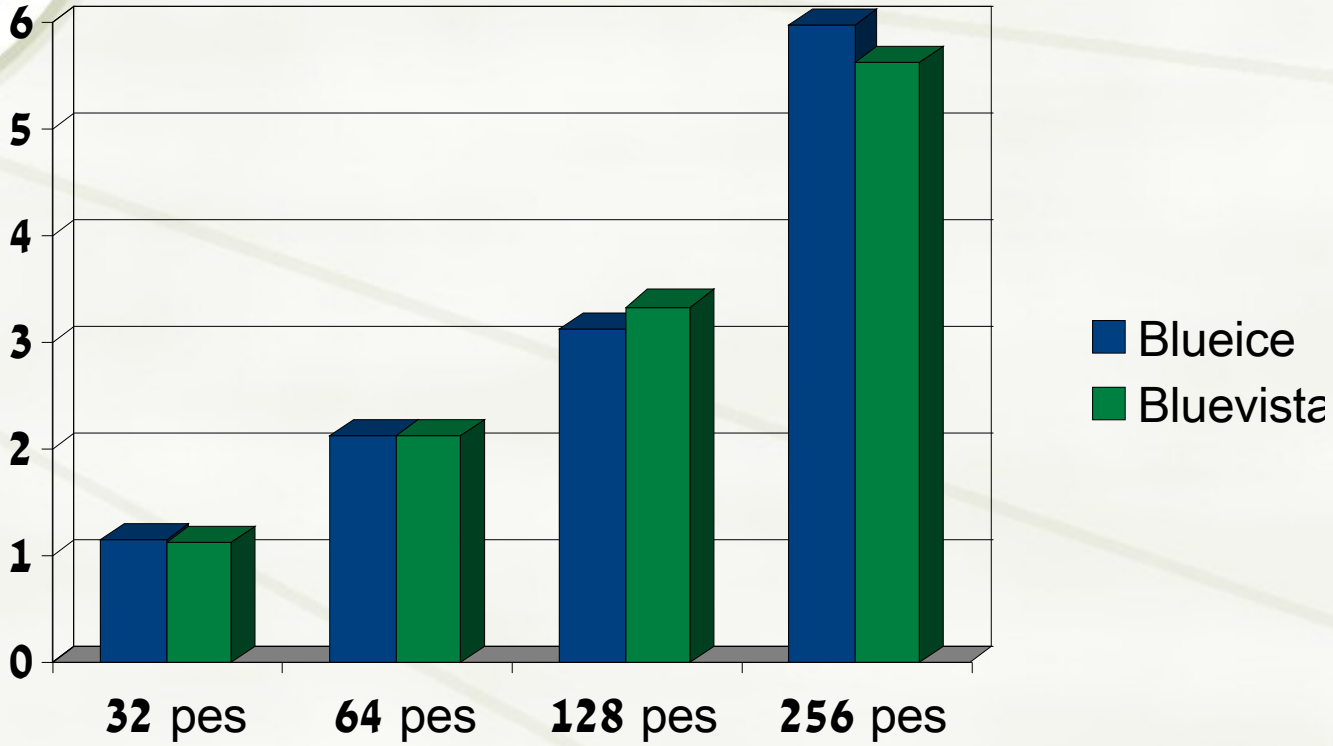
Model Performance: Bluevista vs. Blueice

- Models
  - CAM-WACCM
  - POP
  - HD3D
  - WRF

MODEL PROCS	BV	BL		FAST	100*BL/BV	DIFF %
cam_waccm 256	5.61	5.96		BL	106.23	-6.23
cam_waccm 128	3.34	3.12		BV	93.41	6.58
cam_waccm 64	2.13	2.13		SAME	100	0
cam_waccm 32	1.13	1.15		BL	101.76	-1.76
POP 128	22.29	21.67		BL	97.21	-2.78
POP 64	34.96	36.24		BV	103.66	3.66
POP 48	42.2	43.76		BV	103.69	3.69
POP 32	56.8	65.34		BV	115.03	15.03
POP 24	71.83	79.96		BV	111.31	11.31
POP 16	103.96	112.75		BV	108.45	8.45
POP 8	197.11	231.32		BV	117.35	17.35
hd3D 128	0.1857	0.2188		BV	117.82	17.82
hd3D 64	0.3046	0.3961		BV	130.03	30.03
hd3D 32	0.5261	0.5655		BV	107.48	7.48
hd3D 16	0.9408	0.9459		BV	100.54	0.54
hd3D 8	1.7917	1.7028		BL	95.03	-4.96
WRF 256	0.158	0.1466		BL	92.78	-7.21
WRF 128	0.2316	0.239		BV	103.19	3.19
WRF 64	0.3849	0.3934		BV	102.20	2.20
WRF 32	0.6736	0.7274		BV	107.98	7.98
WRF 16	1.3584	1.407		BV	103.57	3.57
WRF 8	2.6308	2.5783		BL	98.00	-1.99
WRF 4	4.9209	4.8376		BL	98.30	-1.69
WRF 2	9.8538	9.5474		BL	96.89	-3.10
WRF 1	19.8319	17.6361		BL	88.92	-11.07

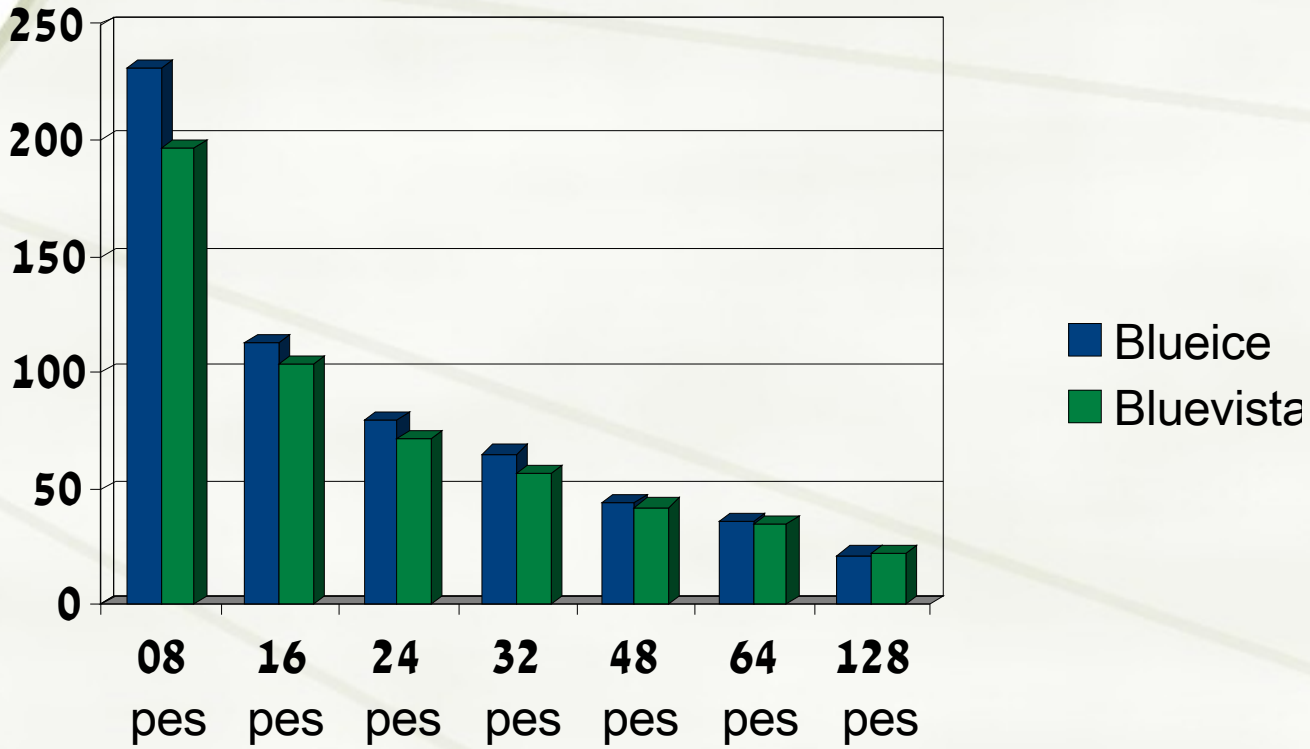
CAM-WACCM (simu-year/day)

-- CAM: Community Atmosphere Model, http://www.ccsm.ucar.edu/models/atm-cam/
-- WACCM: Whole-Atmosphere Community Climate Model, http://waccm.acd.ucar.edu/



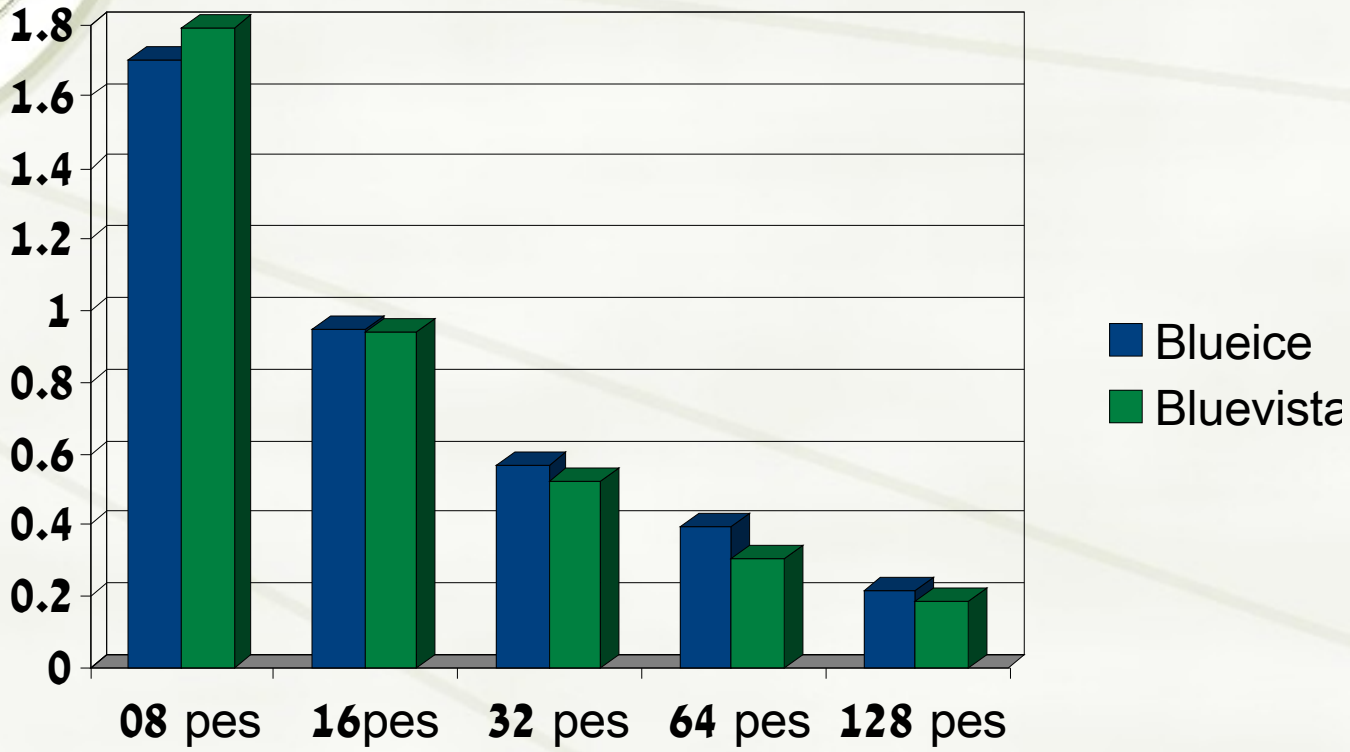
POP (seconds used for step 11)

-- POP: Parallel Ocean Program, http://climate.lanl.gov/Models/POP/index.shtml



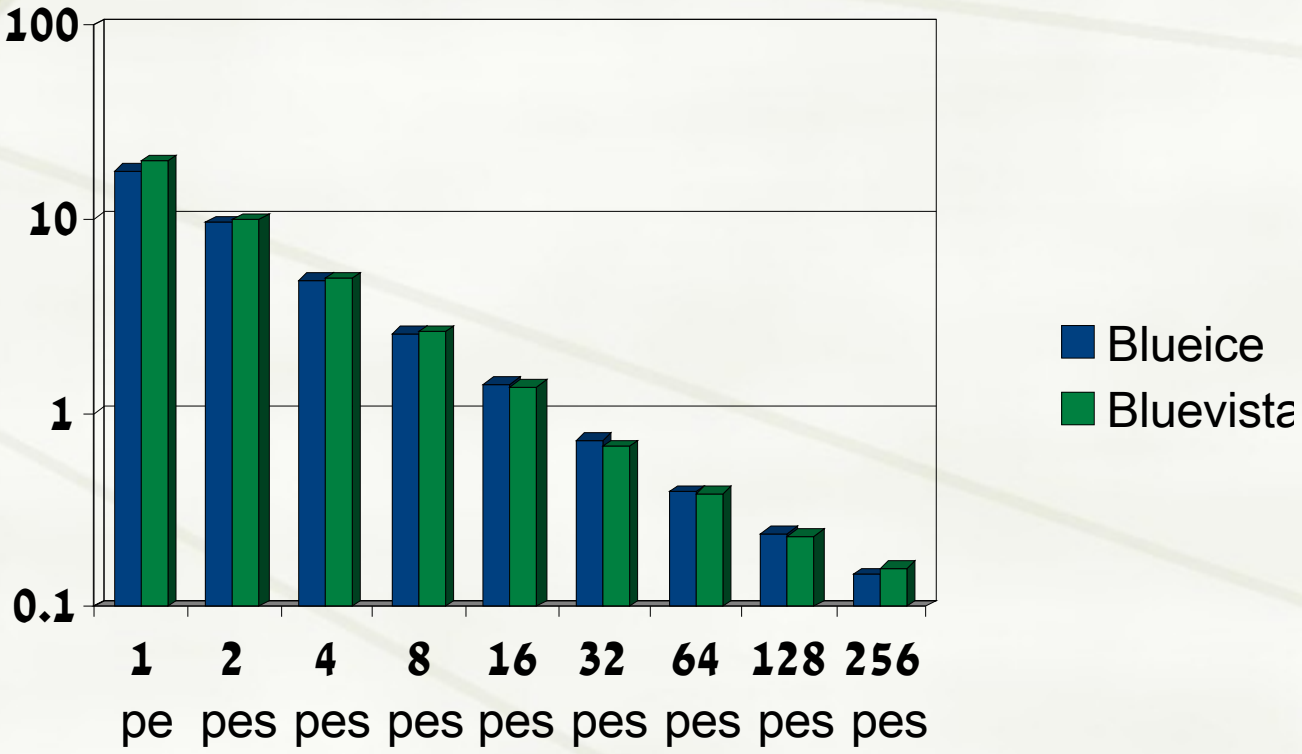
HD3D (seconds/step)

-- HD3D: pseudospectral three-dimensional periodic hydrodynamic/magnetohydrodynamic/ Hall-MHD turbulence model.



WRF (seconds/step)

-- WRF: Weather Research & Forecasting Model, http://www.wrf-model.org/index.php



Model Performance *(continue)*

- Overall
  - Blueice is about 4% slower than bluevista
- CAM-WACCM
  - Almost no difference
- POP
  - Blueice is about 10% slower
- HD3D
  - Blueice can be as much as 30% slower
- WRF
  - Blueice is about 3% slower (did not count small processors)

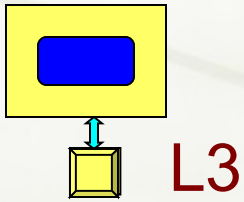
Model Performance *(continue)*

- POP 16 (physical) Processors
  - Bluevista: 103.96sec 2-8 way
  - Blueice: 112.75sec 16 way
  - Difference: 8.45%

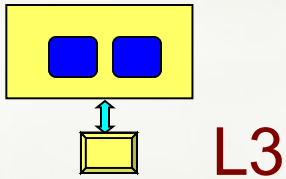
- What Causes the Difference?
  - We compiled the same way
  - We run the same way

Model Performance *(continue)*

- It Must Be The Core



Single Core

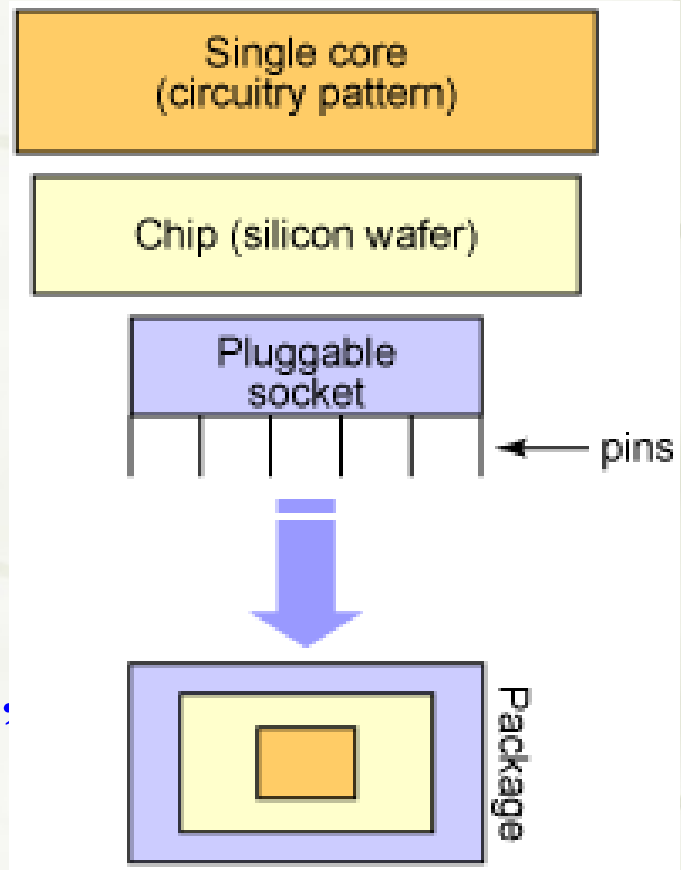


Dual Core

- Can we verify it is due to core difference?

What is "Core"

- What is a Core
  - A core is the circuitry that executes computer commands
- What is Chip
  - Chip, is a silicon wafer that a core resides on
- Bluevista is "Single Core"



What is "Dual Core"

- What is "Dual Core"
  - Dual-core refers to a chip design and fabrication capability that results in two processor cores per physical chip
- Blueice is "Dual Core"

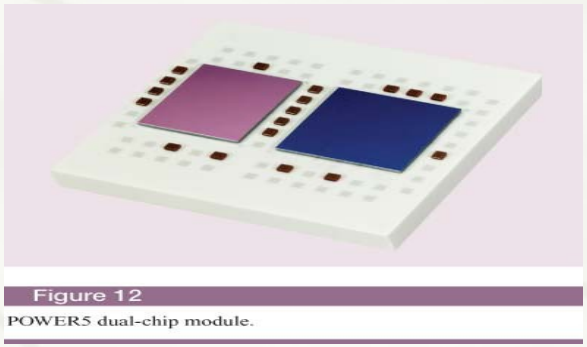
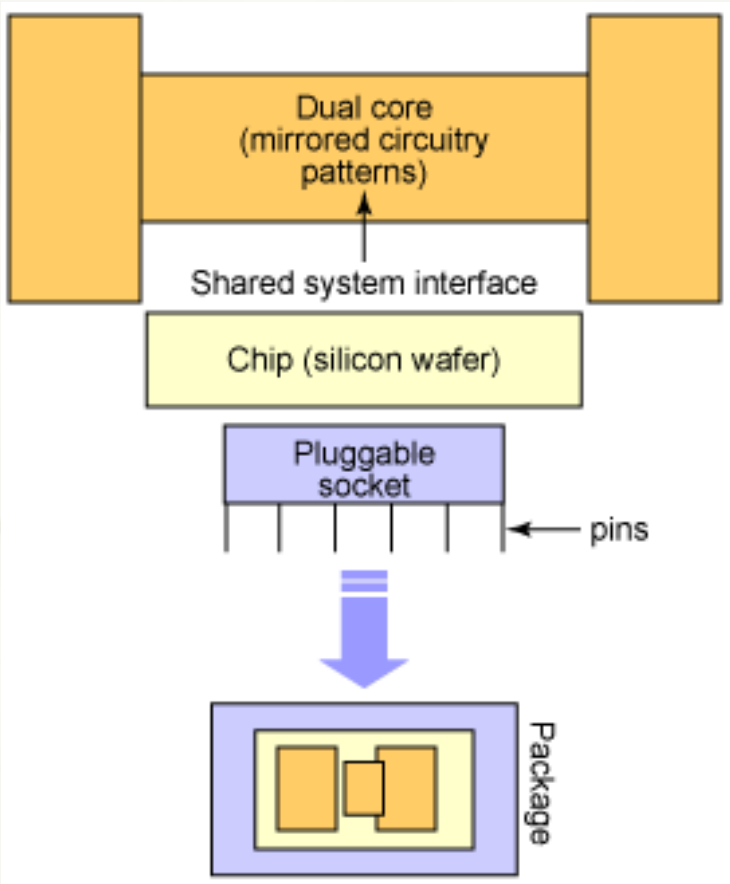


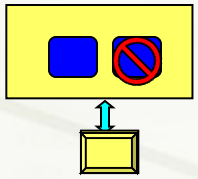
Figure 12
POWER5 dual-chip module.



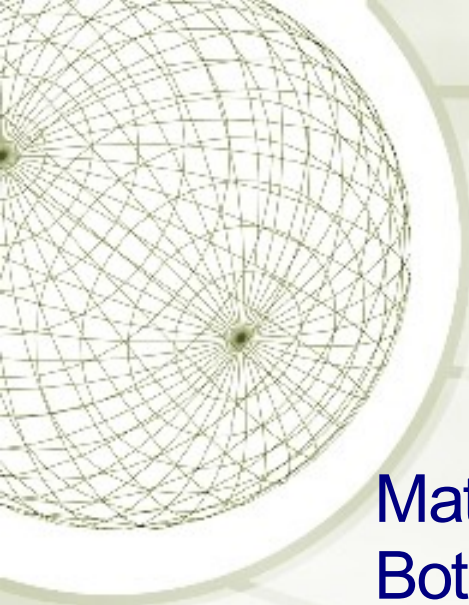
This figure and the one in last slide are from IBM techarticle

Model Performance *(continue)*

- Use one core on blueice chip



- The new results
  - Blueice: 99.86sec 2 nodes (with medium page size)
  - Bluevista: 103.96sec 2 nodes



Matrix Addition

Matrix size: 6000*6000, 200 Iteration
Both use 16 processors

Node Number		1	2
Naïve Code	Time (Sec)	84.91	50.01
	Efficiency	1.12%	1.89%
Use IBM Lib	Time (Sec)	86.18	47.73
	Efficiency	1.10%	1.98%

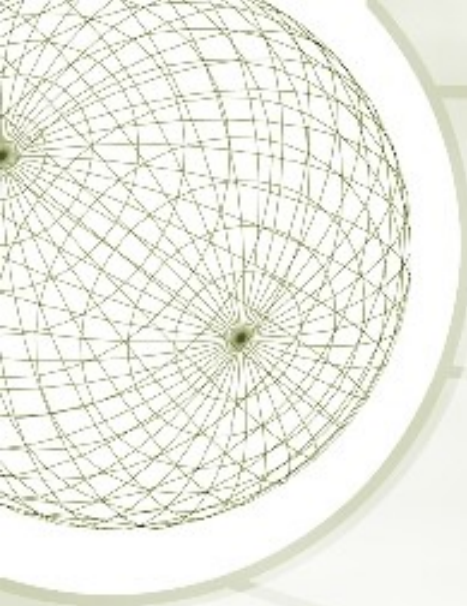
Matrix Multiplication

Matrix size: 1500*1500, 100 Iteration
Both use 16 processors

Node Number		1	2
Naïve Code	Time (Sec)	156.81	128.70
	Efficiency	56.00%	69.01%
Use IBM Lib	Time (Sec)	95.04	92.76
	Efficiency	93.45%	95.75%

Conclusion

- We can get similar performance on blueice as on bluevista
  - Under-subscribe blueice can beat bluevista performance.
- The performance difference on blueice and bluevista is mainly due to L2 cache-miss.



Questions?

Thank You!