TSUBAME2.0 2010年11月1日稼働開始

TSUBAME2.0: A GPU-centric Green 2.4 Petaflops Supercomputer

Tsubame 2.0: "Tiny" footprint, very power efficient
- Floorspace less than 200m² (2,100 ft²)
- Top-class power efficient machine on the Green 500

新規開発

Chip
(CPU ,GPU)

Compute Node
(2 CPUs,3 GPUs)

Node Chassis
(4 Compute Nodes)

Rack
(8 Node Chassis)

System
(42 Racks)
1408 GPU Compute Nodes,
34 Nehalem "Fat Memory" Nodes

CPU(Westmere EP)
76.8 GFLOPS

GPUs(Tesla M2050)
515 GFLOPS
3 GB

1.6 TFLOPS
55 GB/103 GB

6.7 TFLOPS
220 GB/412 GB

53.6 TFLOPS
1.7 TB/3.2 TB

2.4 PFLOPS
80 TB

Integrated by NEC Corporation


Tsubame2.0
2.4 Petaflops, 1408 nodes
~50 compute racks + 6 switch racks
Two Rooms, Total 160m²
1.4MW (Max, Linpack), 0.48MW (Idle)

2010年11月Top500, Green500において TSUBAME2.0上位入賞

- 省エネ性能958MFlops/W⇒世界2位!!
  - Greenest Production Supercomputer in the World賞獲得!!
- 演算性能1.192PFlops ⇒ 世界4位!!
  - 日本のスパコンで10位以内は4年ぶり，5位以内は5年ぶり

TOP500 SUPERCOMPUTER SITES

Rank	Site	Computer/Year Vendor				
1	National Supercomputing Center in Tianjin China	Tianhe-1A - NUDT TH MPP, X5670 2.93Ghz NVIDIA GPU, FT-100 IC / 2010 NUDT				
2	DOE/SC/Oak Ridge National Laboratory United States	Jaguar - Cray XT5-HE Opteron 6-core 2.6 GHz / 2009 Cray Inc.	224162	1759.00	2331.00	6950.60
3	National Supercomputing Centre in Shenzhen (NSCS) China	Nebulae - Dawning TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU / 2010 Dawning	120640	1271.00	2984.30	2580.00
4	GSIC Center, Tokyo Institute of Technology Japan	TSUBAME 2.0 - HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows / 2010 NEC/HP	73278	1192.00	2287.63	1398.61
5	DOE/SC/LBNL/NERSC United States	Hopper - Cray XE6 12-core 2.1 GHz / 2010 Cray Inc.	153408	1054.00	1288.63	2910.00
6	Commissariat a l'Energie Atomique (CEA) France	Tera-100 - Bull bullx super-node S6010/S6030 / 2010 Bull SA	138368	1050.00	1254.55	4590.00
		Roadrunner - BladeCenter QS22/LS21				

TSUBAMEの自動ピークシフト&節電運用

TSUBAME Computing Services > TSUBAME 2.0 - MONITORING PORTAL > Power Monitoring

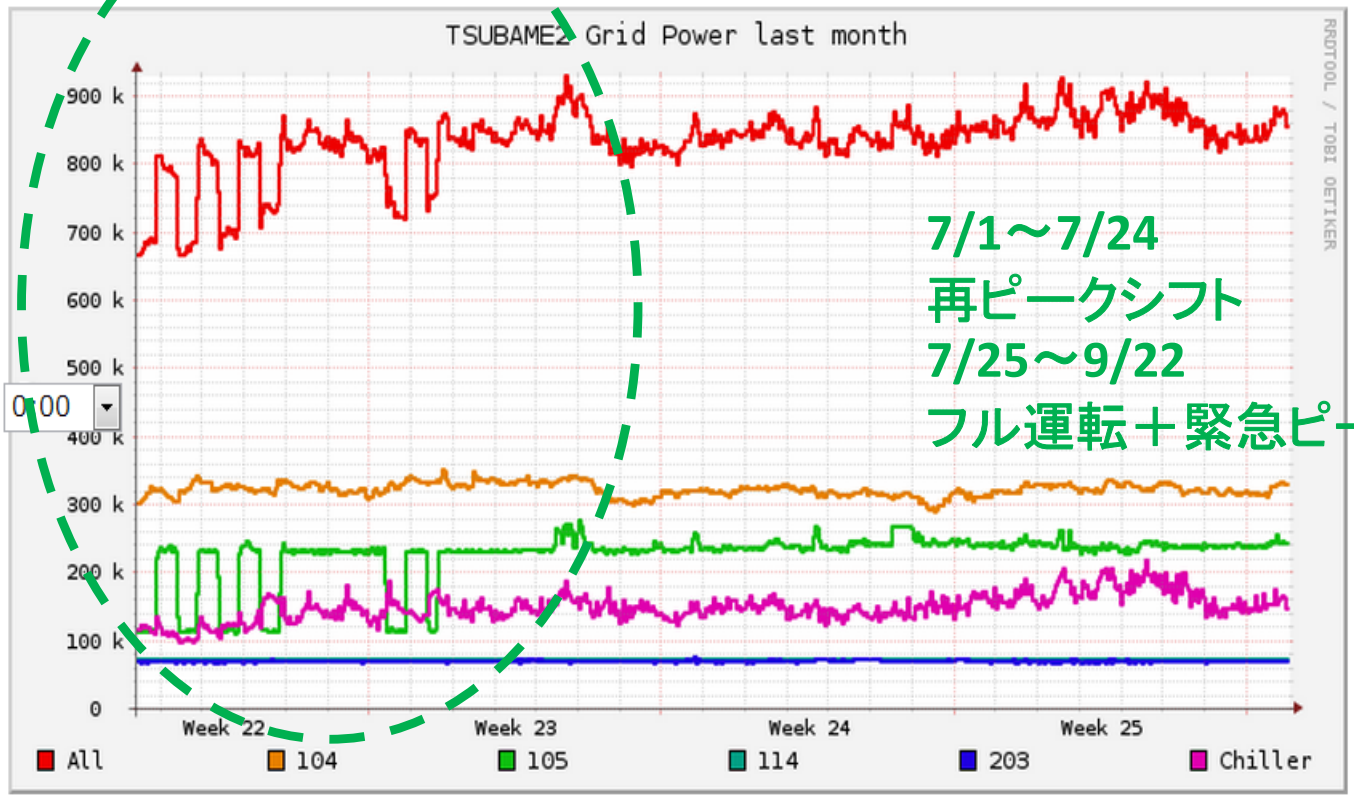
TSUBAME 2.0 Power Monitoring System

TSUBAME 2.0 All
Room 104
Room 105
Room 114
Room 203B
Chiller

Last
month

basedon
now
0:00
apply

TSUBAME 2.0 All Power Summary



7/1～7/24
再ピークシフト
7/25～9/22
フル運転＋緊急ピーク削減

NextGen Weather Prediction on Tsubame2 [Aoki SC10]

Collaboration: Japan Meteorological Agency

Meso-scale Atmosphere Model:

Cloud Resolving Non-hydrostatic model [Shimokawabe et. al. SC10 BSP Finalist]

ex. WRF(Weather Research and Forecast)

Typhoon ~ 1000km

1~ 10km
Tornado,
Down burst,
Heavy Rain

WSM5 (WRF Single Moment 5-tracer) Microphysics*

Represents condensation, precipitation and thermodynamic effects of latent heat release

1 % of lines of code, 25 % of elapsed time

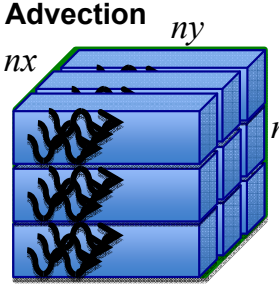
⇒ 20 x boost in microphysics (1.2 - 1.3 x overall improvement)

ASUCA : full GPU Implementation

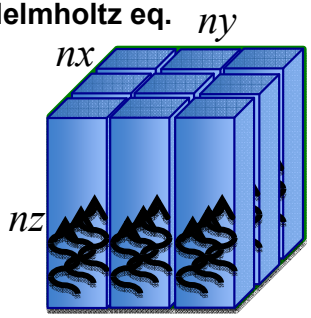
developed by Japan Meteorological Agency

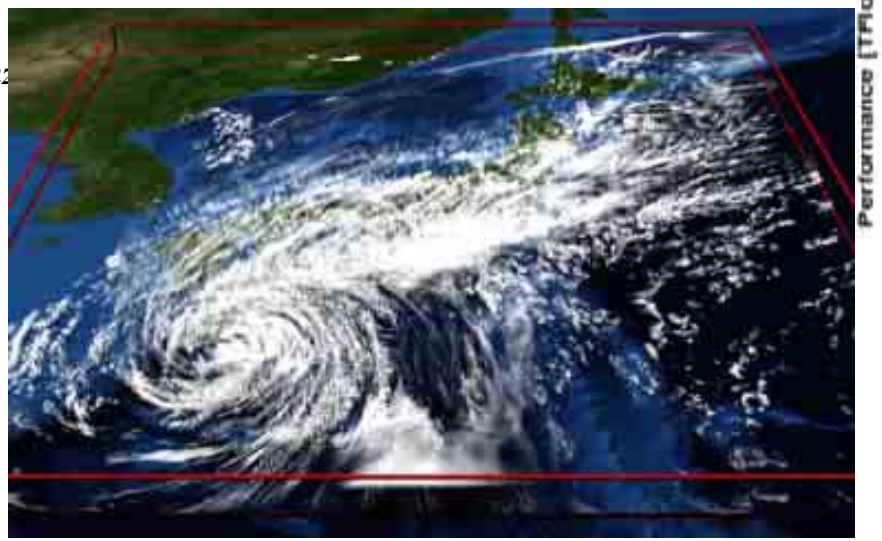
TSUBAME 2.0 : 145 Tflops
@4000 GPUs World Record !!!

Block Division for Advection

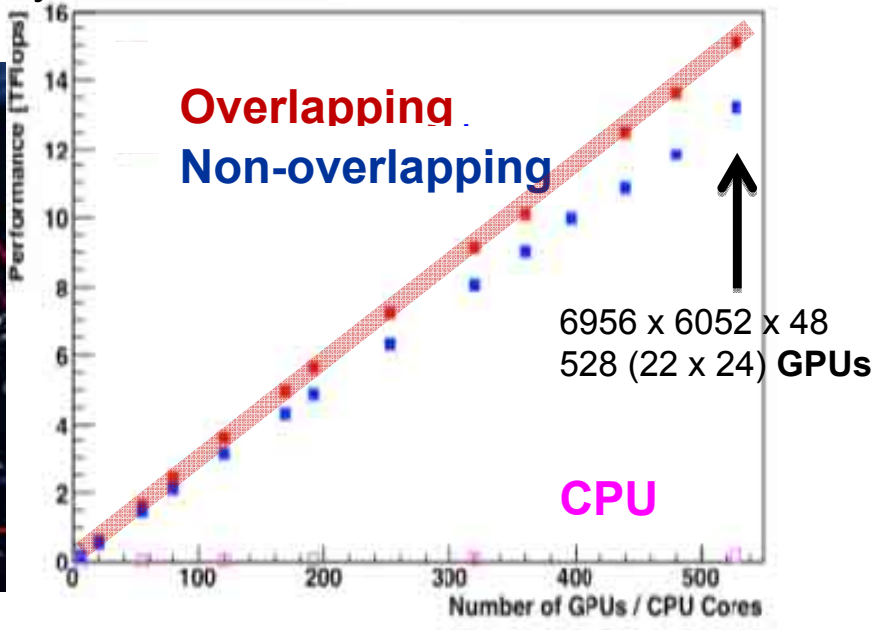


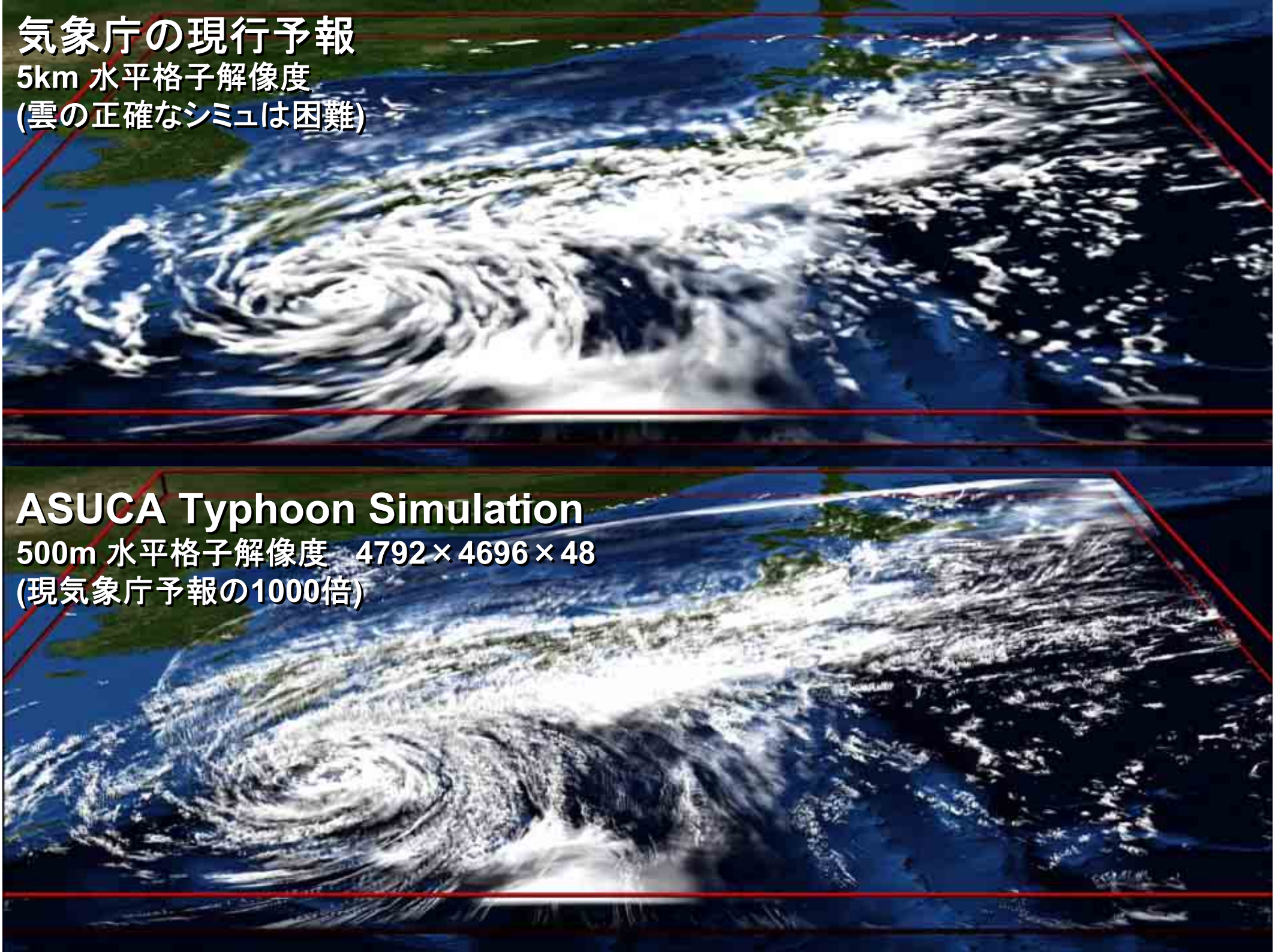
for 1-D Helmholtz eq.





Typhoon




気象庁の現行予報
5km 水平格子解像度
(雲の正確なシミュは困難)
ASUCA Typhoon Simulation
500m 水平格子解像度　4792×4696×48
(現気象庁予報の1000倍)

Example Grand Challenge, Petascale Applications on TSUBAME2.0 early 2011 (~10 apps)

PetaFLOPS Phase-Field Simulation (Aoki)

Metal dendritic solidification is simulated

Phase-field Model:

Time integration of Phase-field

$$\frac{\partial \phi}{\partial t} = M_\phi \left[\nabla \cdot (a^2 \nabla \phi) + \frac{\partial}{\partial x}\left(a \frac{\partial a}{\partial \phi_x} |\nabla \phi|^2\right) + \frac{\partial}{\partial y}\left(a \frac{\partial a}{\partial \phi_y} |\nabla \phi|^2\right) + \frac{\partial}{\partial z}\left(a \frac{\partial a}{\partial \phi_z} |\nabla \phi|^2\right) - \Delta S \Delta T \frac{dp(\phi)}{d\phi} - W \frac{dq(\phi)}{d\phi} \right]$$

M_ϕ Mobility, a Interface anisotropy, $\frac{\Delta S}{\Delta T}$ Entropy of fusion / Undercooling

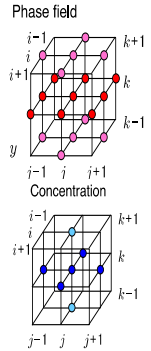
Time integration of solute concentration

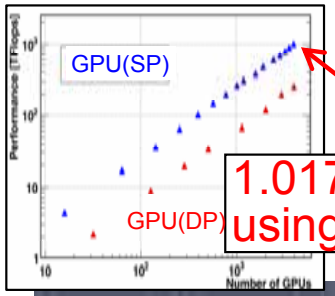
$$\frac{\partial c}{\partial t} = \nabla \cdot [D_S \phi \nabla c_S + D_L (1-\phi) \nabla c_L]$$

$$c_S = \frac{kc}{1-\phi+k\phi}, \; c_L = \frac{c}{1-\phi+k\phi}, \; k = c_S / c_L$$

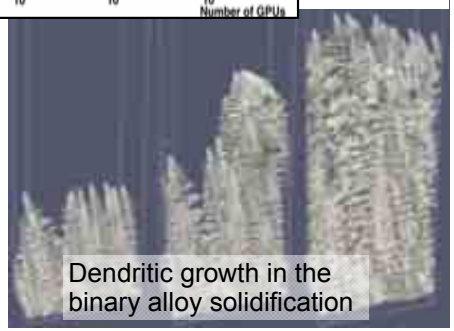
D_S D_L Diffusion coeff. in solid and liquid

Access pattern





1.017 PFlops in SP using 4,000 GPUs



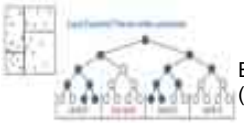
(Mesh:768×1632×3264)

SC11 Gordon Bell Finalist + Tech Paper, fastest Stencil ever

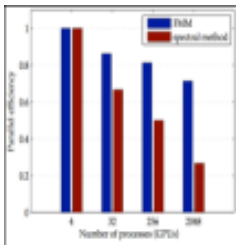
Turbulance Simulation using FMM (Yasuoka)

Vortex method with fast multipole method (FMM) is used

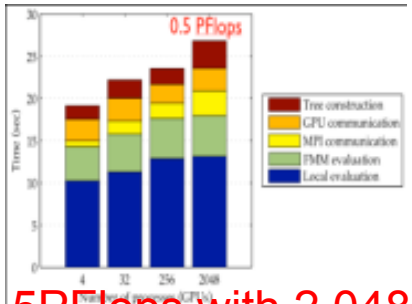
Q criteria in isotropic turbulence



Efficiency in weak scaling (4×10^6 particles per proc)



Breakdown of exec time



0.5PFlops with 2,048GPUs

Fastest FMM ever? Can hit 1PF w/ 4000 GPUs

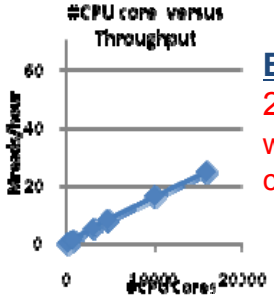
BLASTX for Millions of DNAseq (Akiyama)

Metagenome Analysis for Bacteria in Soil

Data: 224million DNA reads(75b) /set

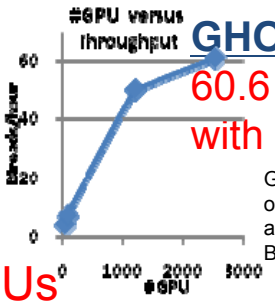
Pre-filtering: reduces to 71million reads

BLASTX: 71million DNA vs. nr-aa DB (4.2G)



BLASTX

24.4 Million/hour with 16008 CPU cores



GHOSTM

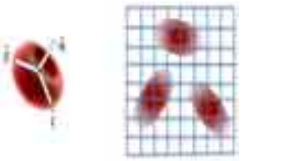
60.6 Million/hour with 2,520 GPUs

GHOSTM is our original CUDA app almost compatible to BLASTX.

Can cope next gen Giga Sequencers

Multiphysics Biofluidics Simulation (Bernaschi)

Simulation of blood flows that accounts for from red blood cells to endothelial stress

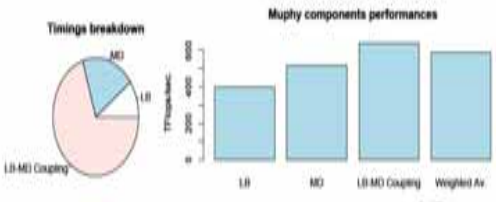


Red blood cells as ellipsoidal particles

Multiphyics simulation with *MUPHY* software

Fluid: blood plasma → Lattice Boltzmann(LB)

Body: RBC → Extended MD



0.6PFlops with 4,000GPUs

450M RBCs are simulated

SC11 Gordon Bell Finalist Fastest LBM ever? x3 faster than 1PF BG/P