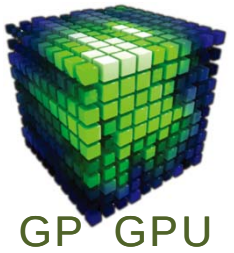
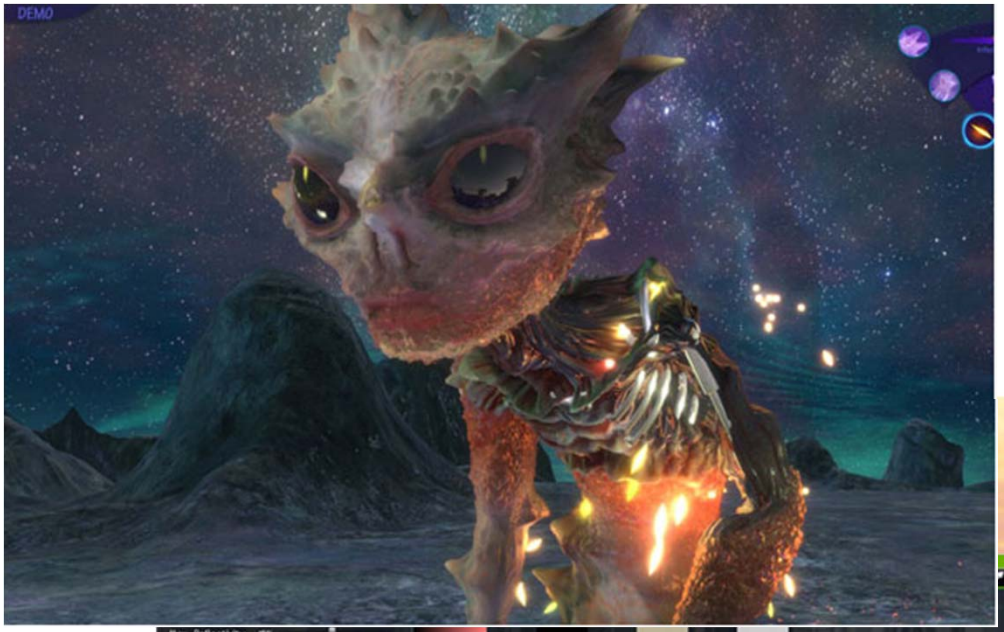


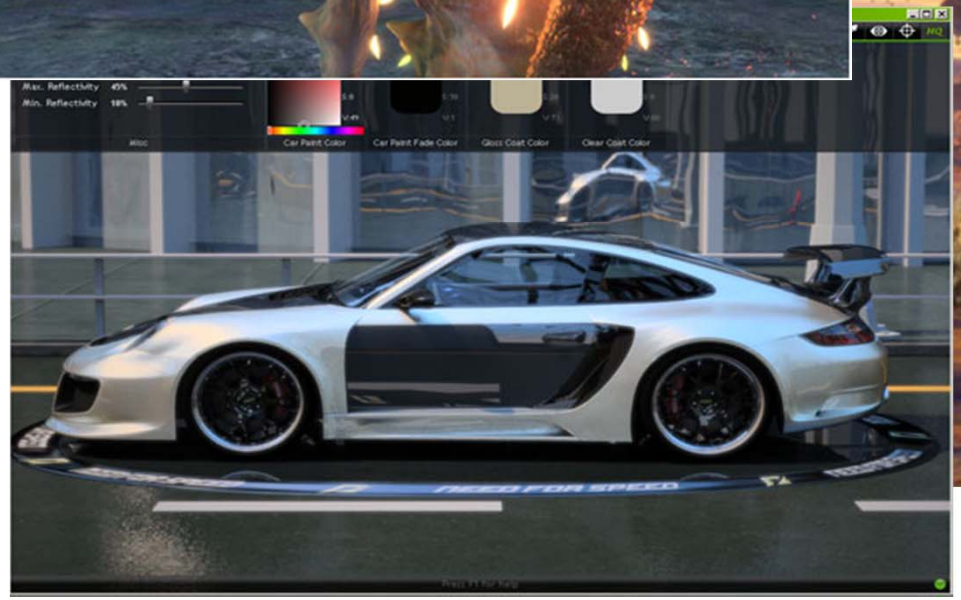
TSUBAME 2.0 の概要

東京工業大学・学術国際情報センター

青木　尊之

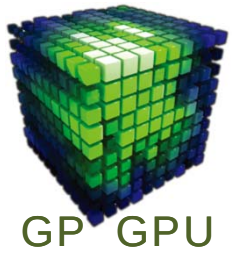
What is GPU ?





GP GPU

GP-GPU Computing



General-Purpose Graphics Processing Unit

- **High Performance over TFLOPS**

- **Major differences from Previous Accelerators**
  ClearSpeed, Grape, , ,

High Memory Bandwidth
suitable for CFD applications

Consumer Product
inexpensive

Software Development Environment
CUDA, Open CL

Supercomputer in the world

2010 November

Rank	Site	Computer/Year Vendor	Cores	R_{max}	R_{peak}	Power
1	National Supercomputing Center in Tianjin China	Tianhe-1A - NUDT YH Cluster, X5670 2.93Ghz 6C, NVIDIA GPU, FT-1000 8C / 2010 NUDT	186368	2566.00	4701.00	4040.00
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

TSUBAME 2.0

System (58 racks)

1442 nodes: 2952 CPU sockets,
4264 GPUs

Performance: 224.7 TFLOPS (CPU) ※ Turbo boost
2196 TFLOPS (GPU)
Total: **2420** TFLOPS

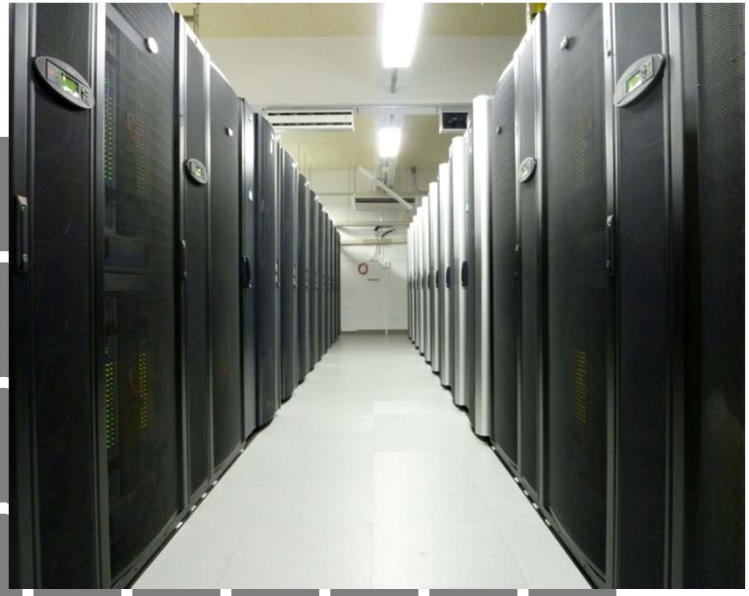
Memory: 103.9 TB

Rack (30 nodes)

Performance: 51.0 TFLOPS
Memory: 2.03 TB

Compute Node
(2 CPUs, 3 GPUs)

Performance: 1.7 TFLOPS
Memory: 58.0GB(CPU)
+9.7GB(GPU)



Details of Compute Node

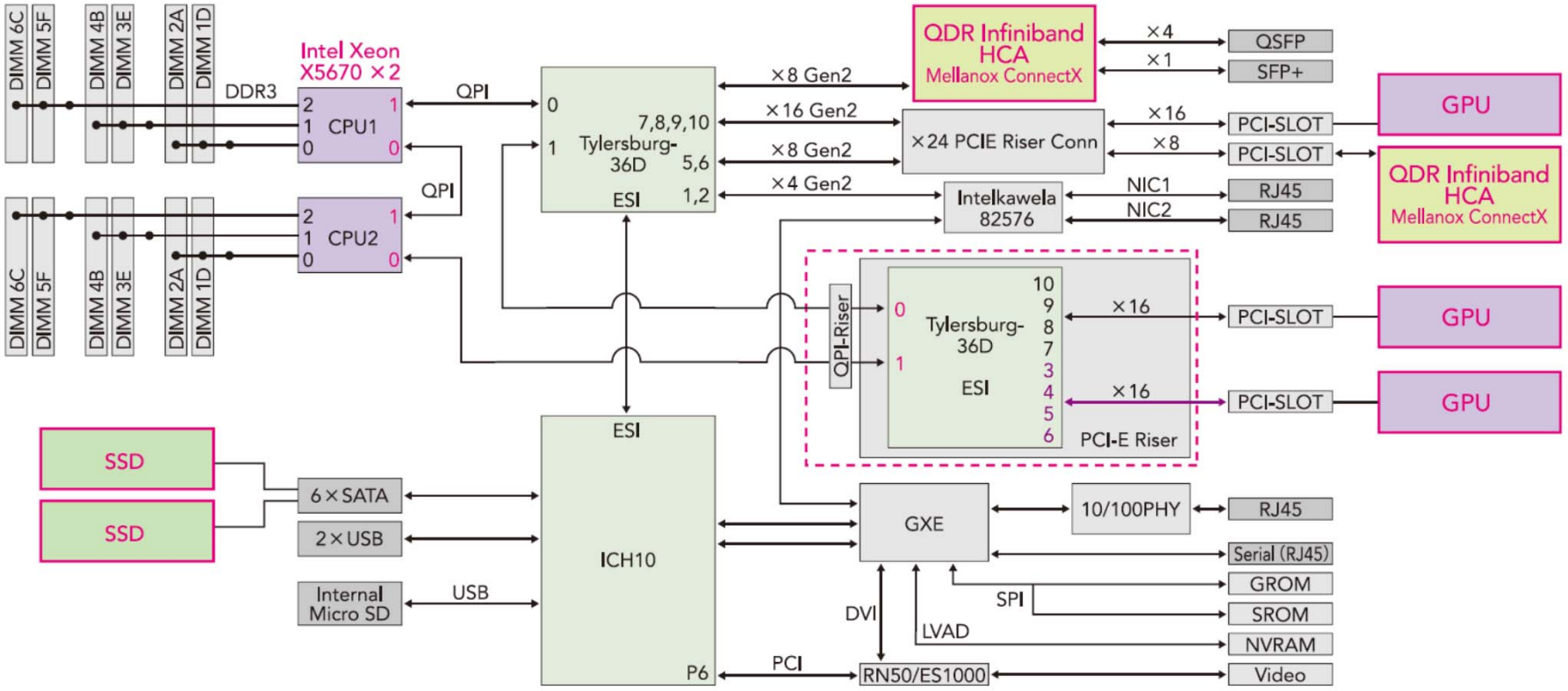
HP ProLiant SL390s

GPU : NVIDIA Tesla M2050 (Fermi Core) ×3 515GFLOPS VRAM 3GB/GPU

CPU : Intel Xeon X5670 2.93Ghz ×2
6 core/socket 76.7 GFLOPS (12cores/node) ※ Turbo boost: 3.196GHz

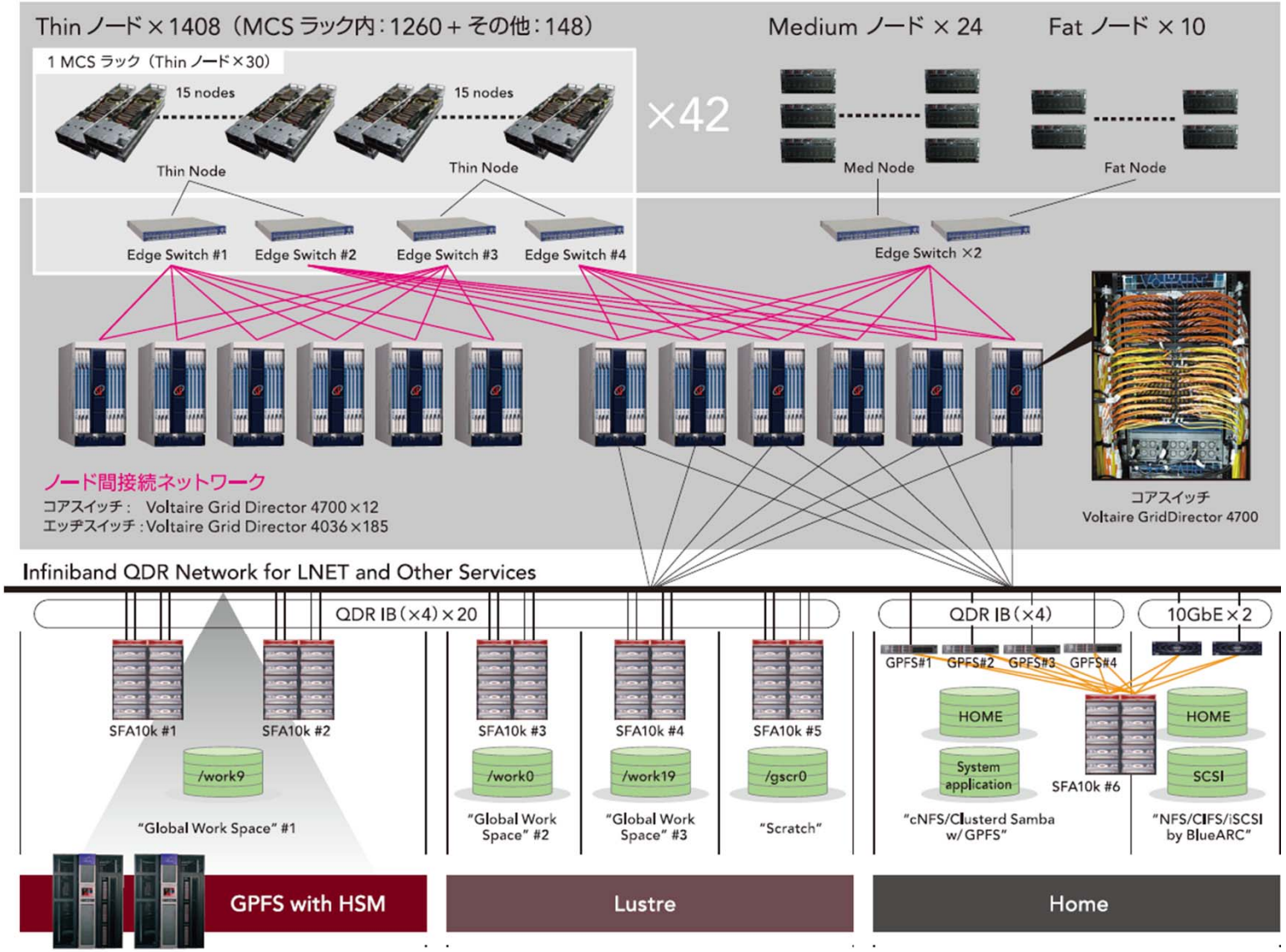
Memory : 58GB DDR3 1333MHz 一部 103GB

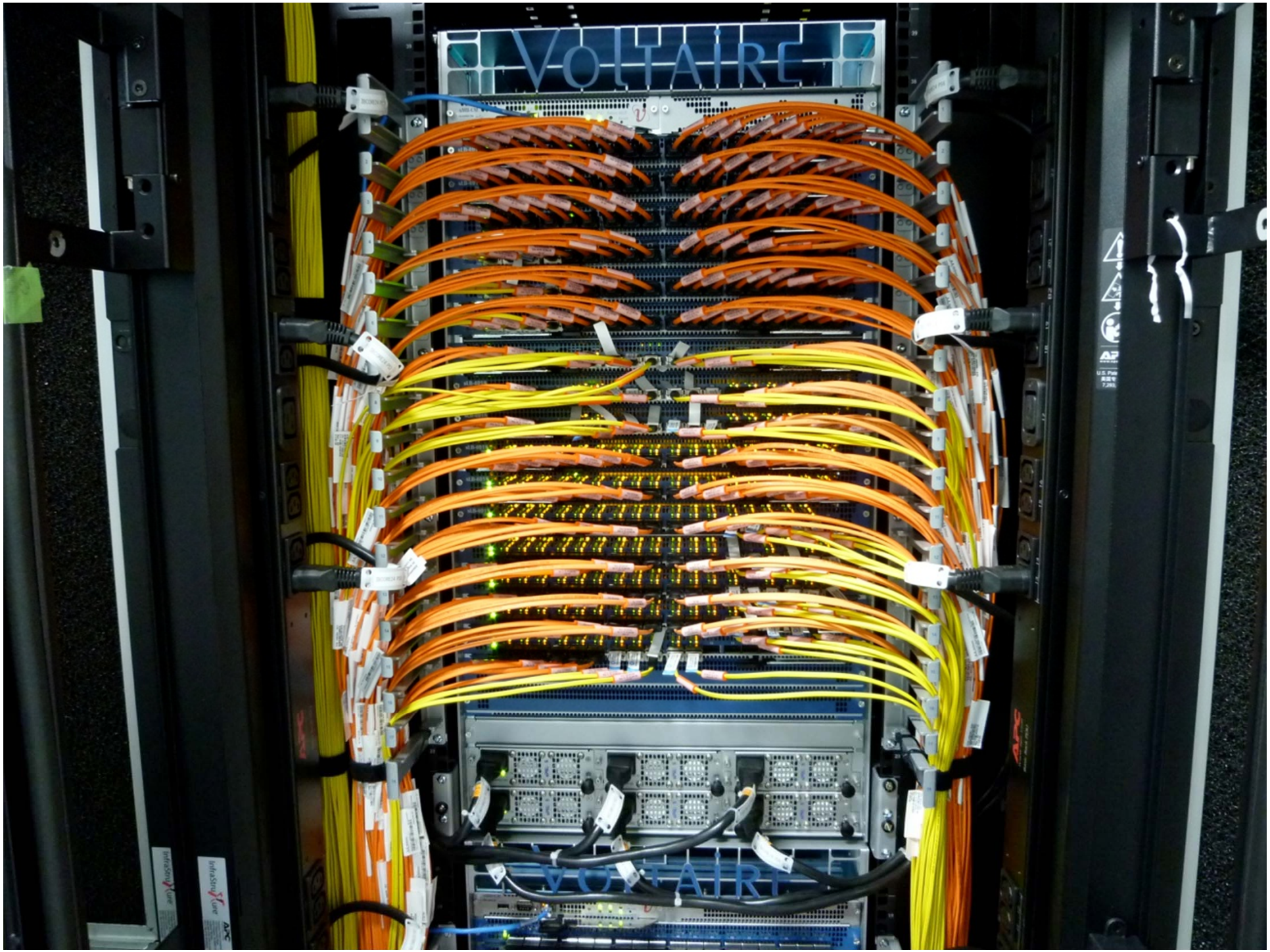
SSD : 60GB ×2 (120GB/node) 一部 120GB ×2 (240GB/node)



GPU M2050

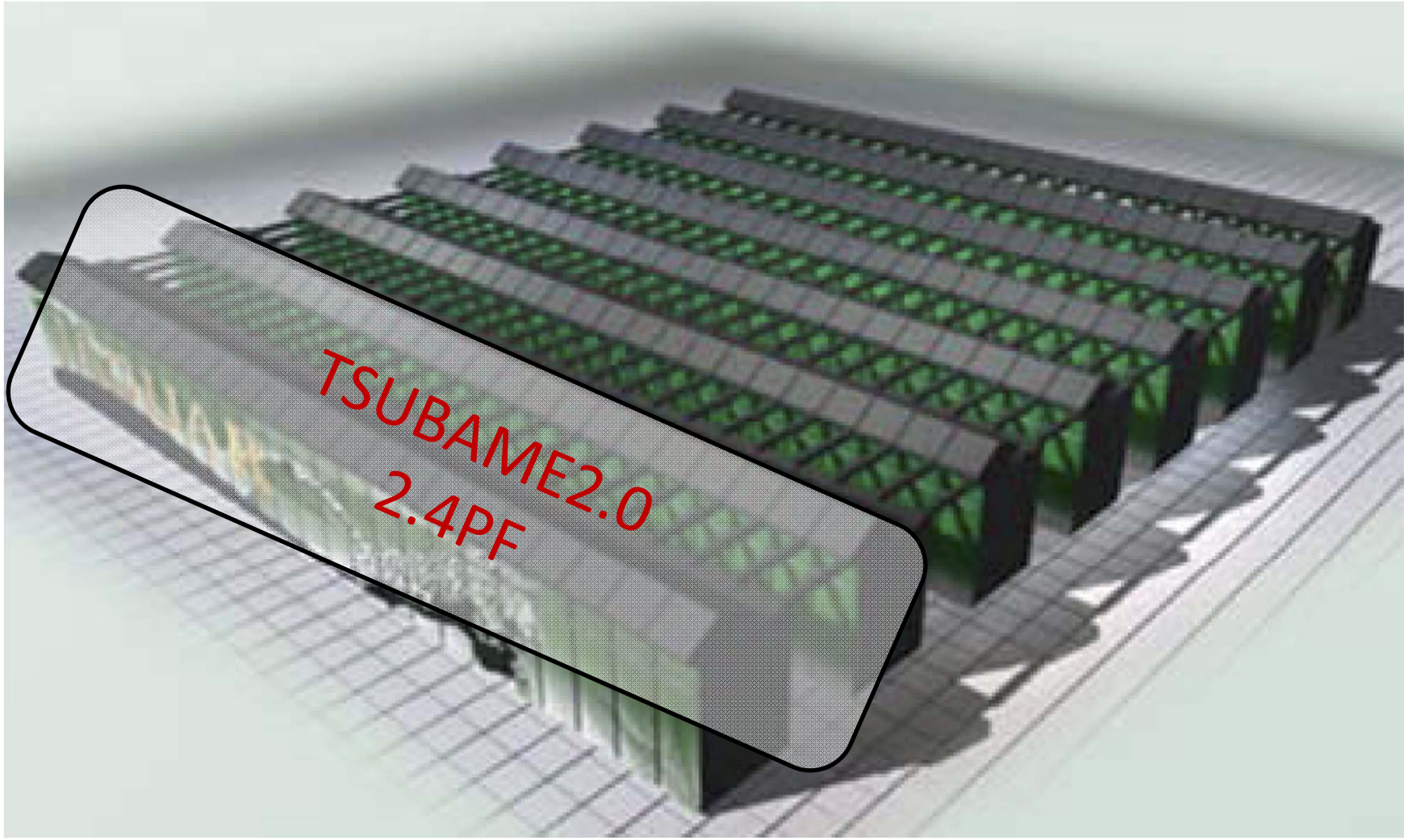
High-Speed Network and Reliable Storage System




VOLTAIRE
VOLTAIRE

ORNL Jaguar vs Tsubame 2.0

Similar Peak Performance, 1/4 the Size and Power



東京工業大学
TOKYO INSTITUTE OF TECHNOLOGY
科学と技術で未来を創造する

Supercomputer in the world

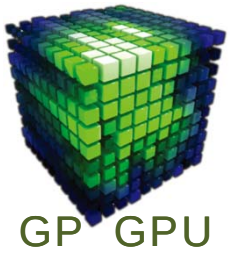
THE GREEN 500™

The Green500 list, November 2010

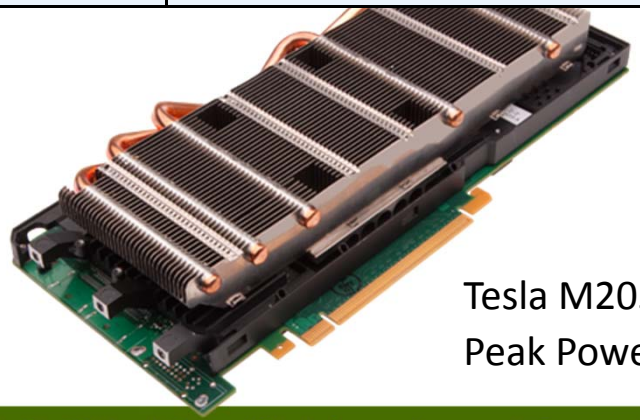
Green500 Rank	MFLOPS/W	Site*	Computer*	Total Power (kW)
1	1684.20	IBM Thomas J. Watson Research Center	NNSA/SC Blue Gene/Q Prototype	38.80
2+	1448.03	National Astronomical Observatory of Japan	GRAPE-DR accelerator Cluster, Infiniband	24.59
2	958.35	GSIC Center, Tokyo Institute of Technology	HP ProLiant SL390s G7 Xeon 6C X5670, Nvidia GPU, Linux/Windows	1243.80
3	933.06	NCSA	Hybrid Cluster Core i3 2.93Ghz Dual Core, NVIDIA C2050, Infiniband	36.00
4	828.67	RIKEN Advanced Institute for Computational Science	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	57.96
5	773.38	Universitaet Wuppertal	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54
5	773.38	Universitaet Regensburg	QPACE SFB TR Cluster, PowerXCell 8i, 3.2 GHz, 3D-Torus	57.54

TSUBAME2.0 PUE = 1.2 **(Power Usage Effectiveness)**

CPU/GPU Spec Sheet



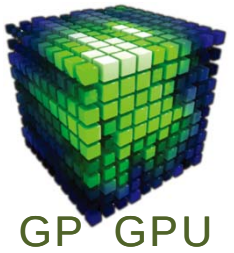
		Intel Xeon X5670	Tesla C2050 /M2050	GeForce GTX 580 Fermi
GPU	Peak Performance [GFlops]	**76.8*,153.6**	**515*,1030**	**197*,1576**
	Number of Processor	6	448	512
	Core Clock [GHz]	2.93	1.15	1.544
Memory	Bandwidth[GB/s]	**32.0**	**148.6**	**192.1**
	Memory Interface [bit]	64	384	384
	Memory Clock [GHz]	1.333 (DDR3)	1.50 (GDDR5)	2.00 (GDDR5)
B_{peak}/F_{peak}	Bandwidth/Performance	**0.416**	**0.289**	**0.974**

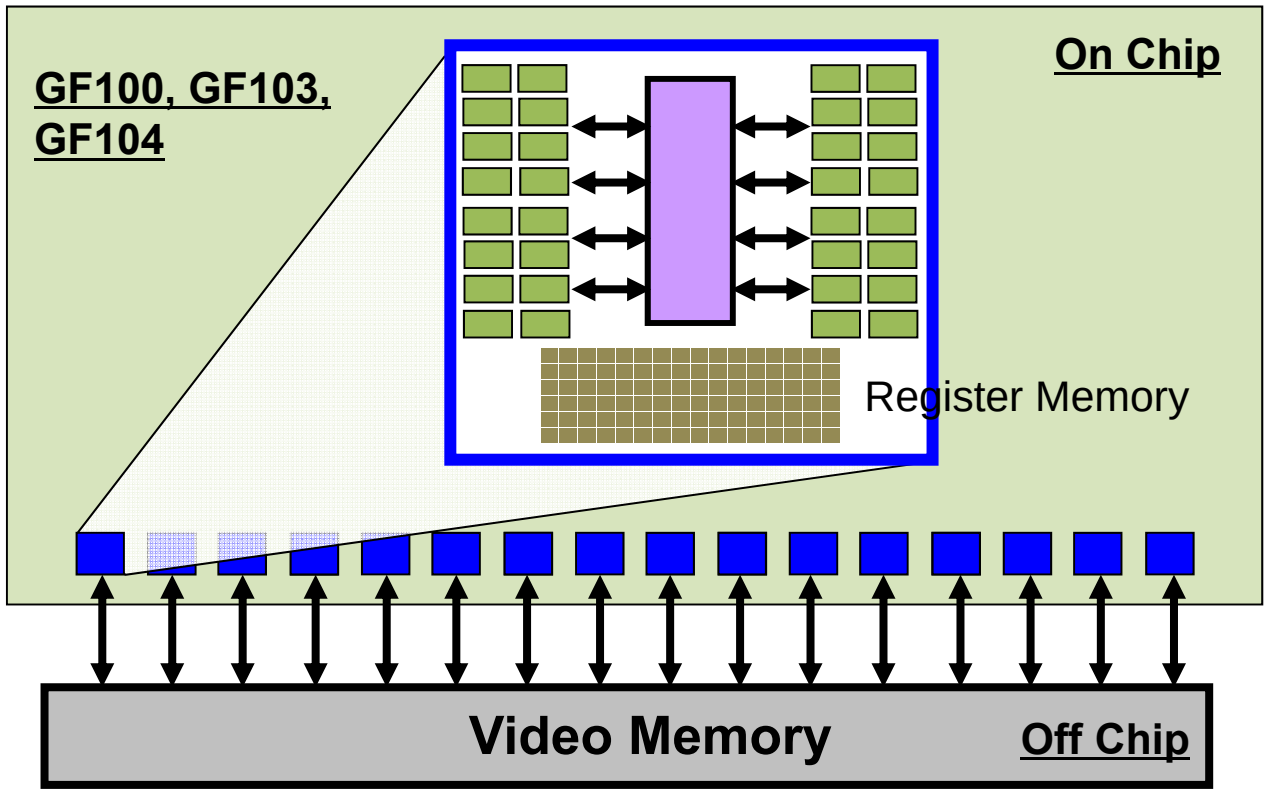


Tesla M2050
Peak Power : 225W

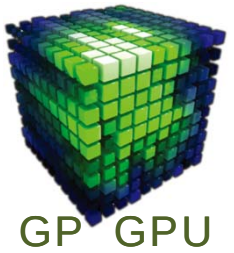
Peak Power : 244W

GPU Architecture



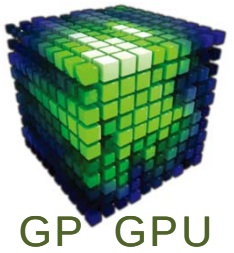


Global memory	~6GB (VRAM
Streaming Multiprocessor	~16 (C2050 (GF100): 14)
Shared memory + L1 Cache	64 Kbyte
Streaming Processor (CUDA core)	8～48 per SM, total 512



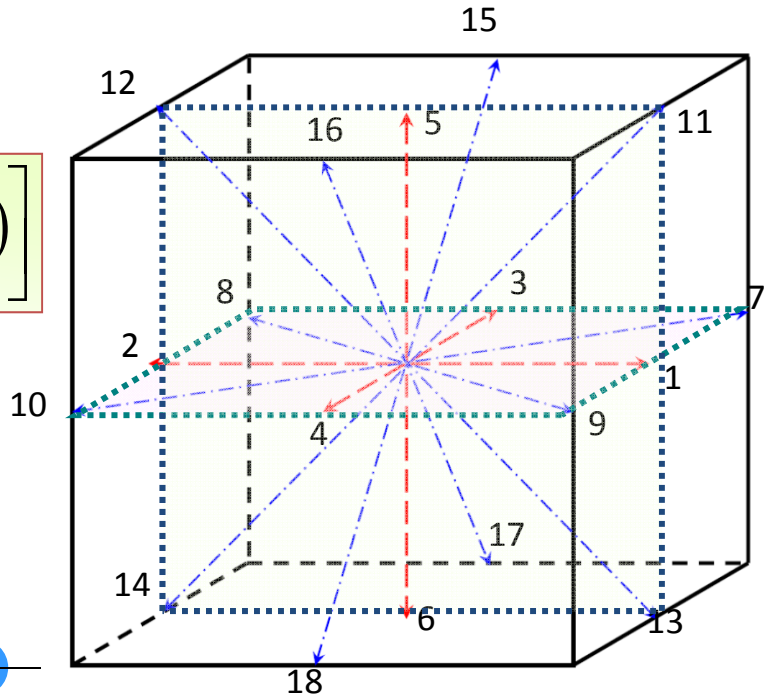
Showcase of CFD Applications

Lattice Boltzmann Method



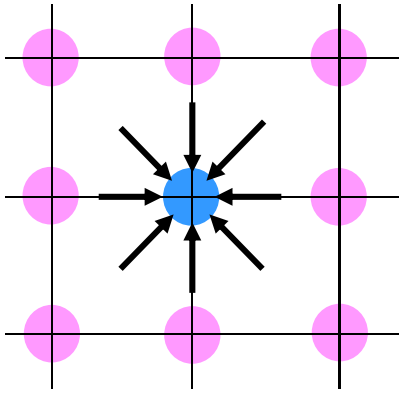
$$\frac{\partial f_i}{\partial t} + \mathbf{e}_i \cdot \nabla f_i = -\frac{1}{\lambda}\left(f_i - f_i^{eq}\right)$$

$$f_i^{eq} = \rho w_i \left[1 + \frac{3}{c^2}(\mathbf{e}_i \cdot \mathbf{u}) + \frac{9}{2c^4}(\mathbf{e}_i \cdot \mathbf{u})^2 - \frac{3}{2c^2}(\mathbf{u} \cdot \mathbf{u})\right]$$

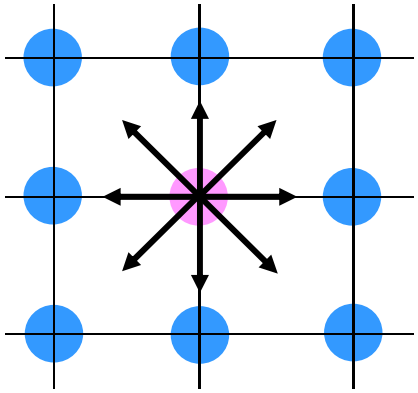


Strongly Memory Bound Problem:

Collision step:



Streaming step:



i is the value in the direction of *ith* discrete velocity
e_i is the discrete velocity set;
w_i is the weighting factor
c is the particle velocity
u is the macroscopic velocity

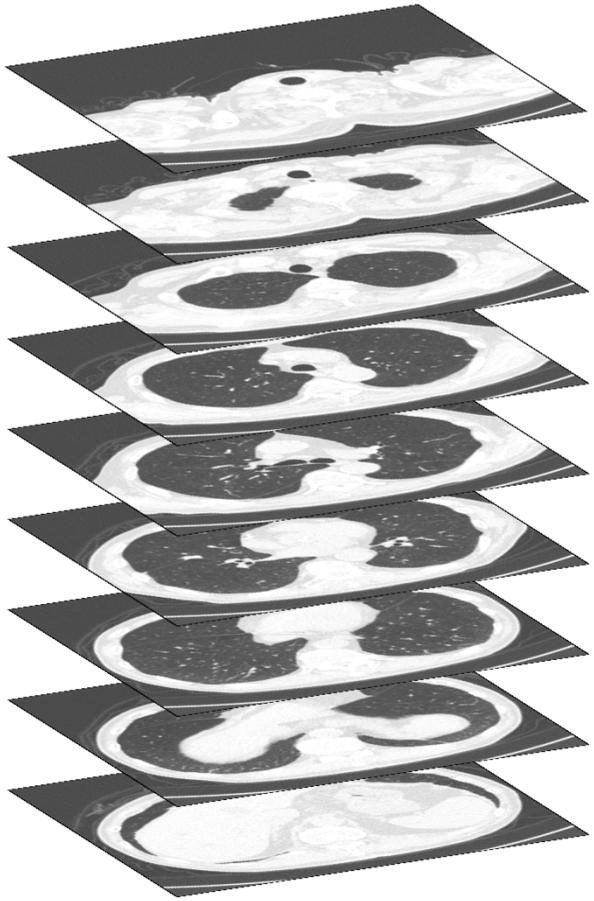
Pulmonary Airflow Study

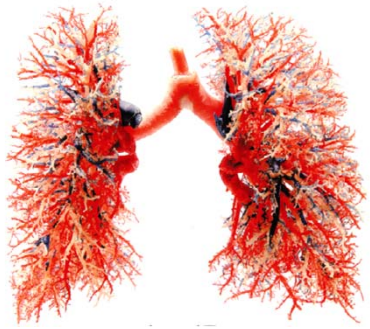
Collaboration with Tohoku University

GP GPU

「人体の不思議展」より

X-Ray CT images
512×512×512

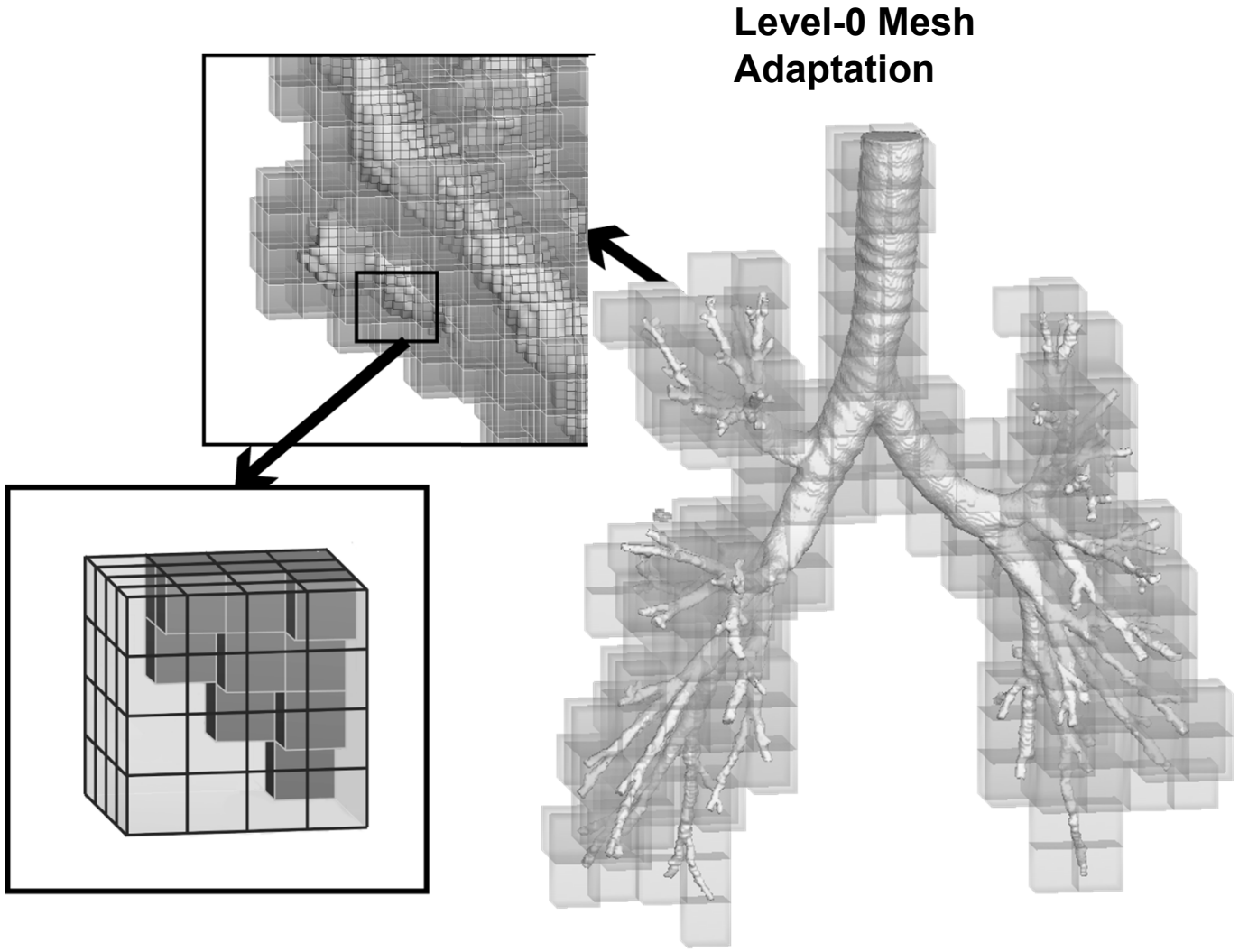




Airway structure Extraction

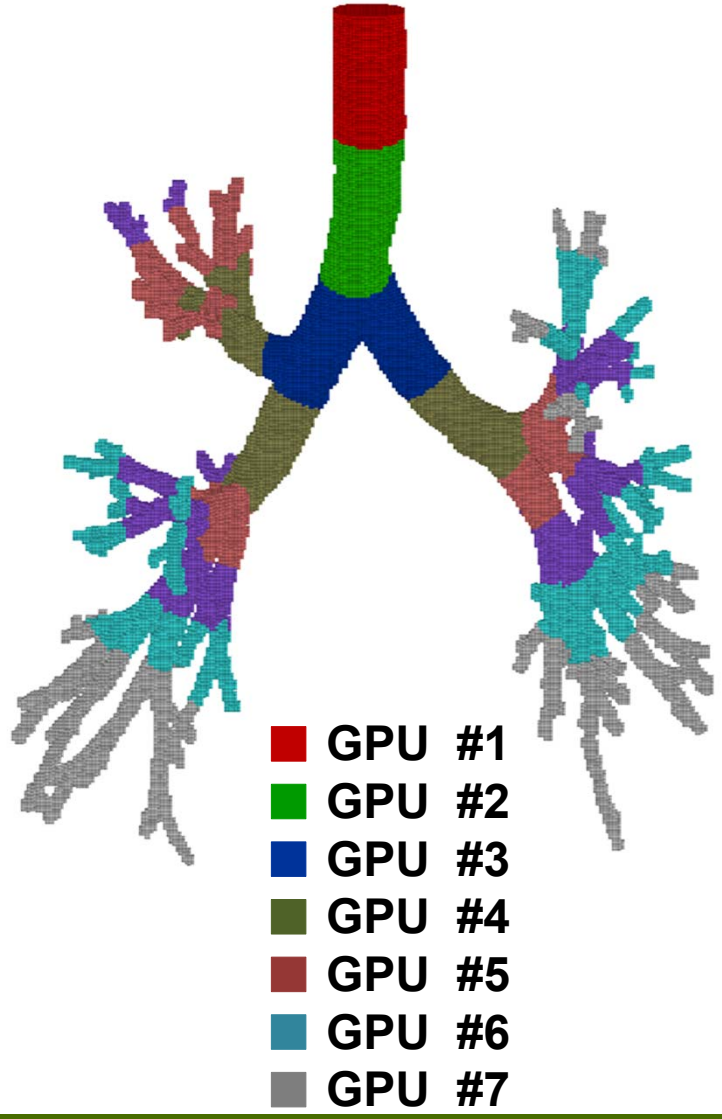
Pulmonary Airflow Study

Collaboration with Tohoku University

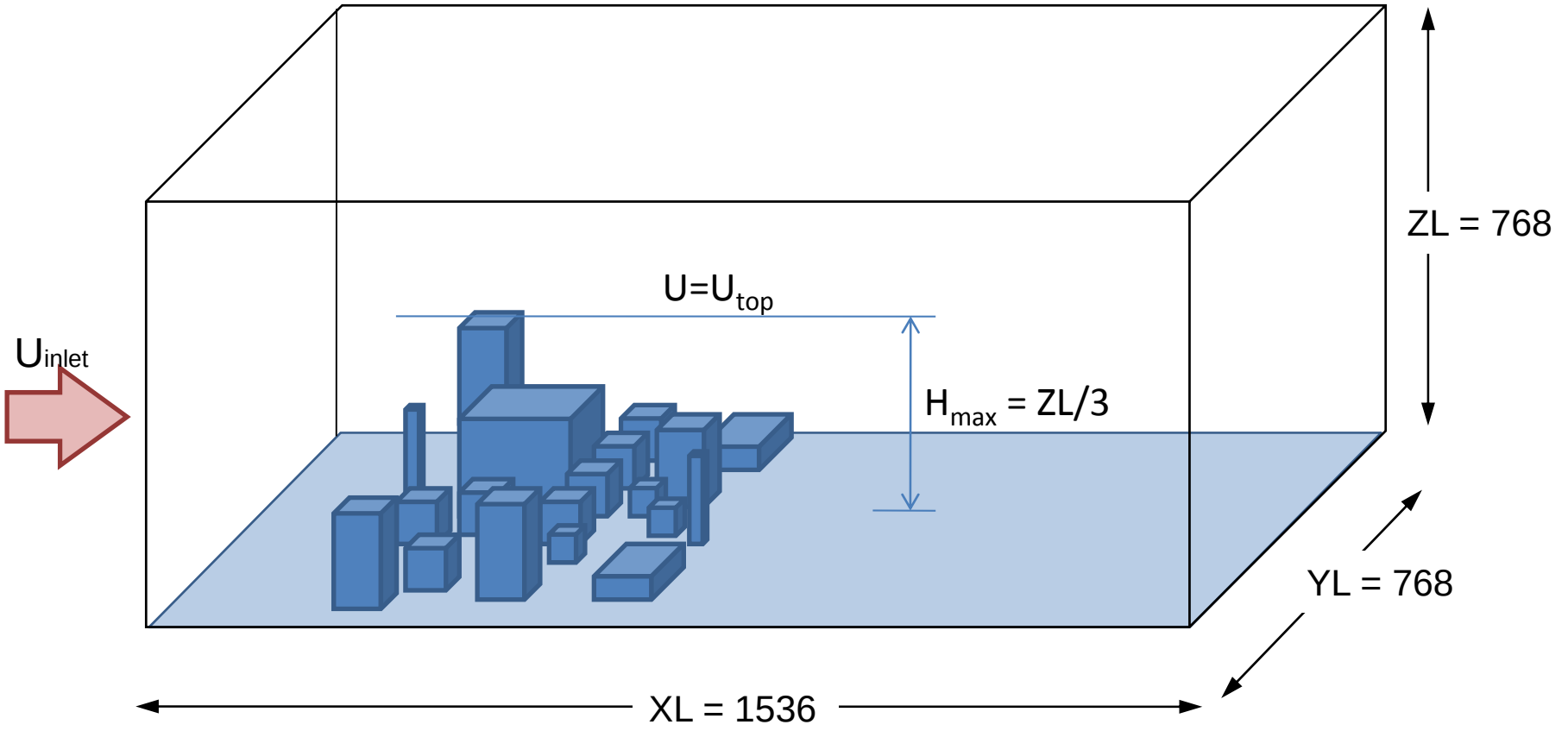


Pulmonary Airflow Study

Collaboration with Tohoku University



Building Model



$U_{inlet}(z)=U_{top}(z/ H_{max})^{0.2}$

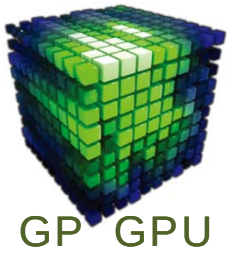
$Re= U_{top} H_{max}/\nu$

BC: $U_{inlet}(z)$ for inlet and v=w=0

BC: u,v,w,ρ flux=0 for sides and top and outlet

BC: bounce back for bottom and building surface

Real-time TSUNAMI Simulation



ADPC : Asian Disaster Preparedness Center

Early Warning System:

Data Base → Real-time CFD

high accuracy

Shallow-Water Eq.

Conservative Form:

$$\frac{\partial h}{\partial t}+\frac{\partial hu}{\partial x}+\frac{\partial hv}{\partial y}=0$$

Assuming

hydrostatic balance

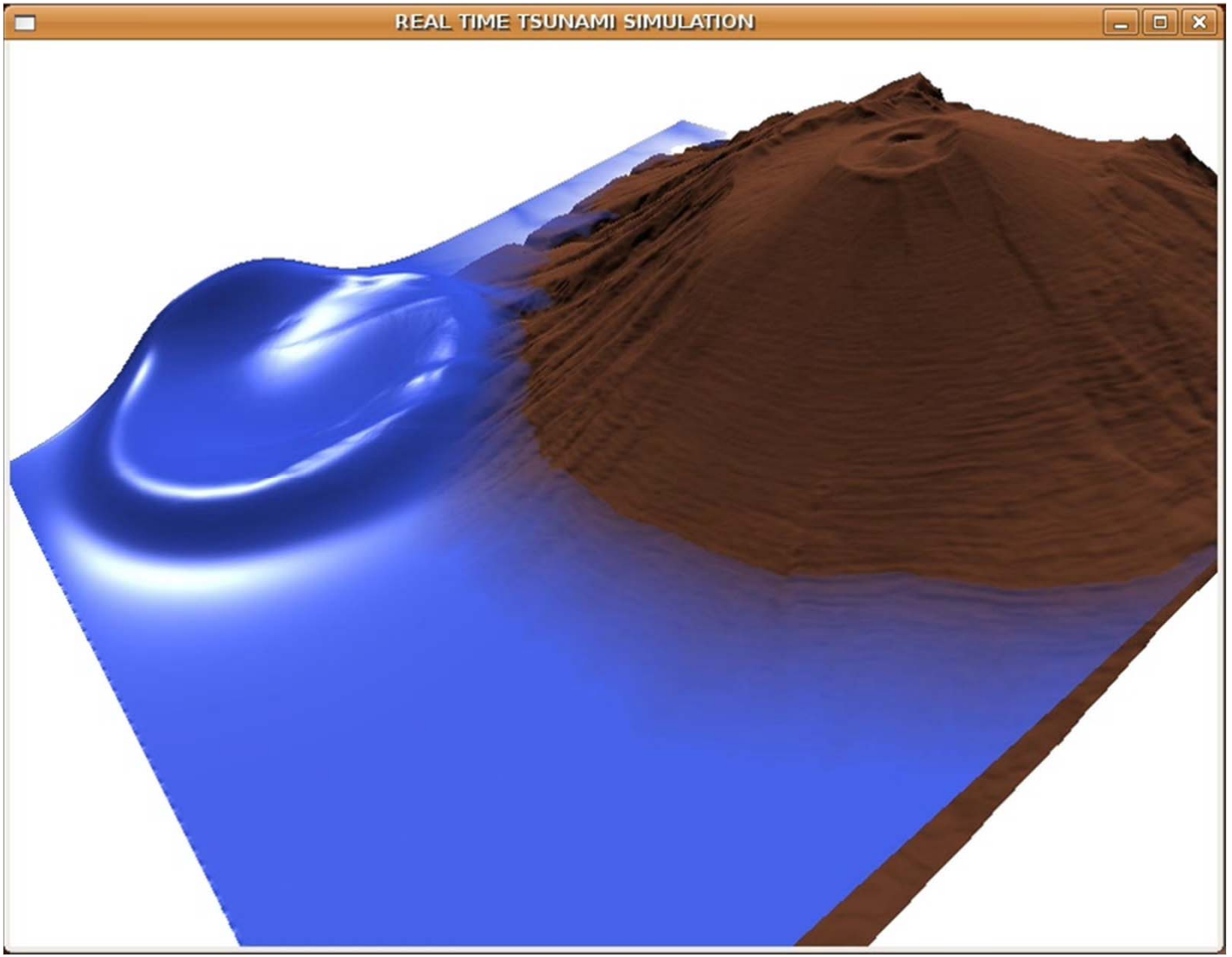
in the vertical direction,

$$\frac{\partial hu}{\partial t}+\frac{\partial}{\partial x}\left(hu^2+\frac{1}{2}gh^2\right)+\frac{\partial huv}{\partial y}=-gh\frac{\partial z}{\partial x}$$

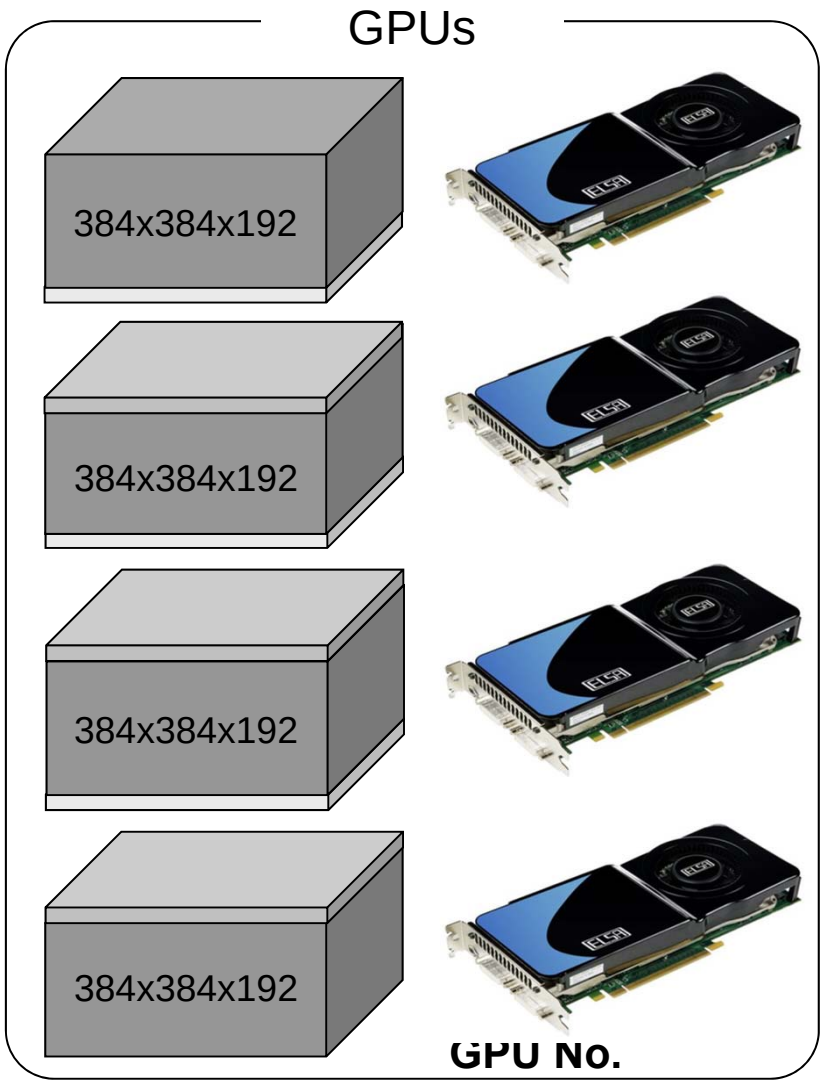
3D → 2D equation

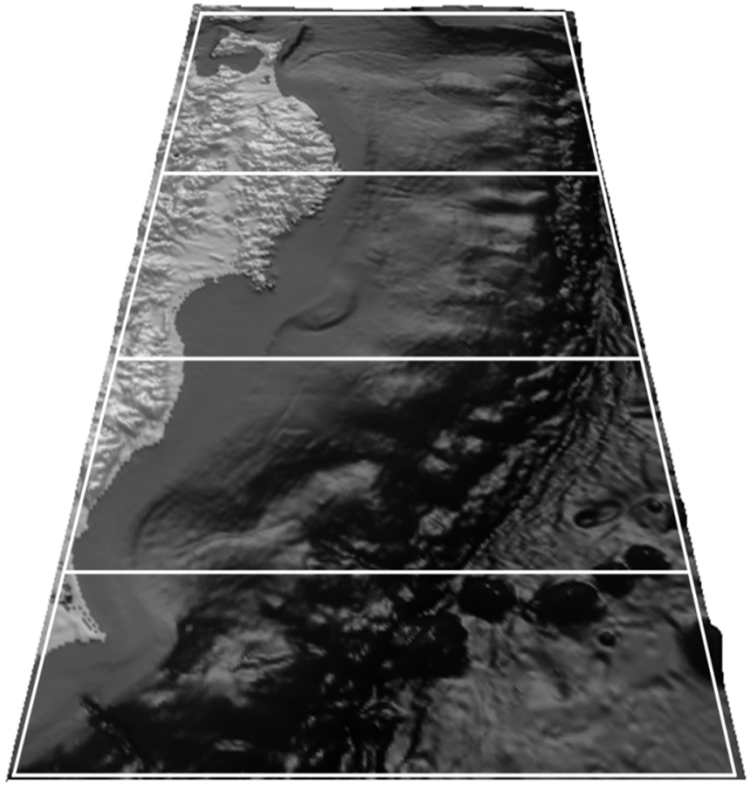
$$\frac{\partial hv}{\partial t}+\frac{\partial huv}{\partial x}+\frac{\partial}{\partial y}\left(hv^2+\frac{1}{2}gh^2\right)=-gh\frac{\partial z}{\partial y}$$

SCREEN Capture



Large-scale Real-time Tsunami Simulator



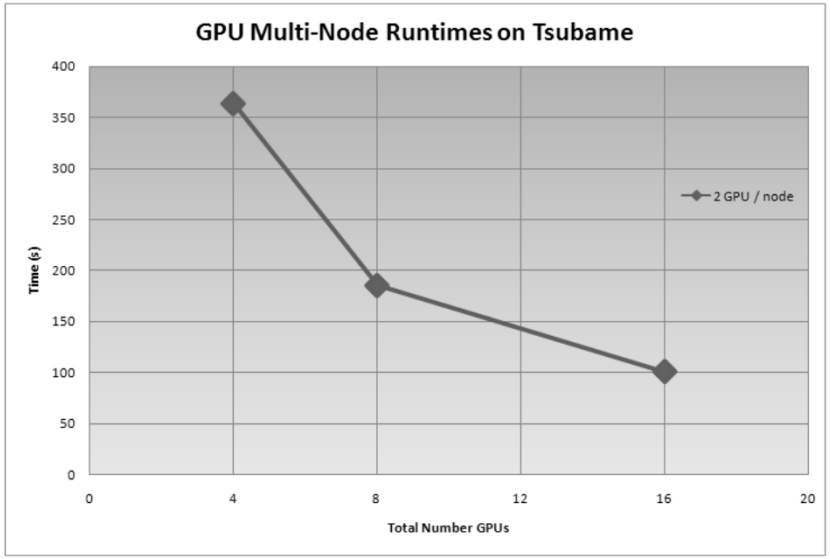


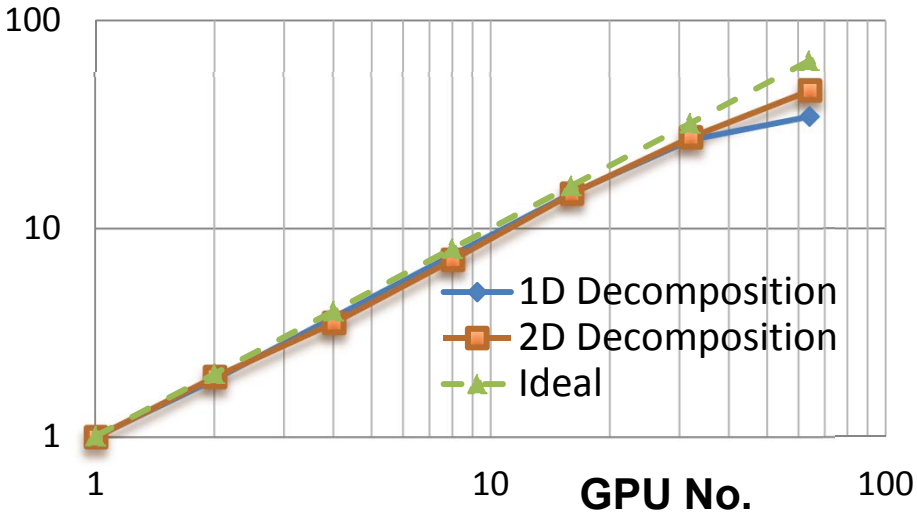
8 GPU 400km×800km
（100m mesh）

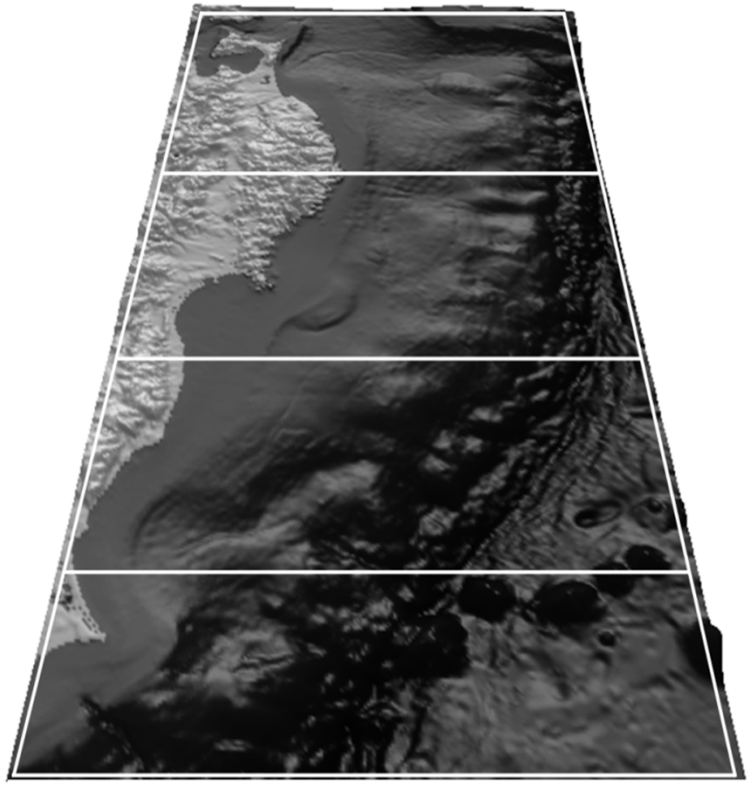
within 3 min

Large-scale Real-time Tsunami Simulator

GP GPU



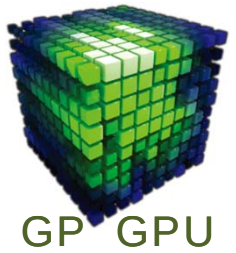




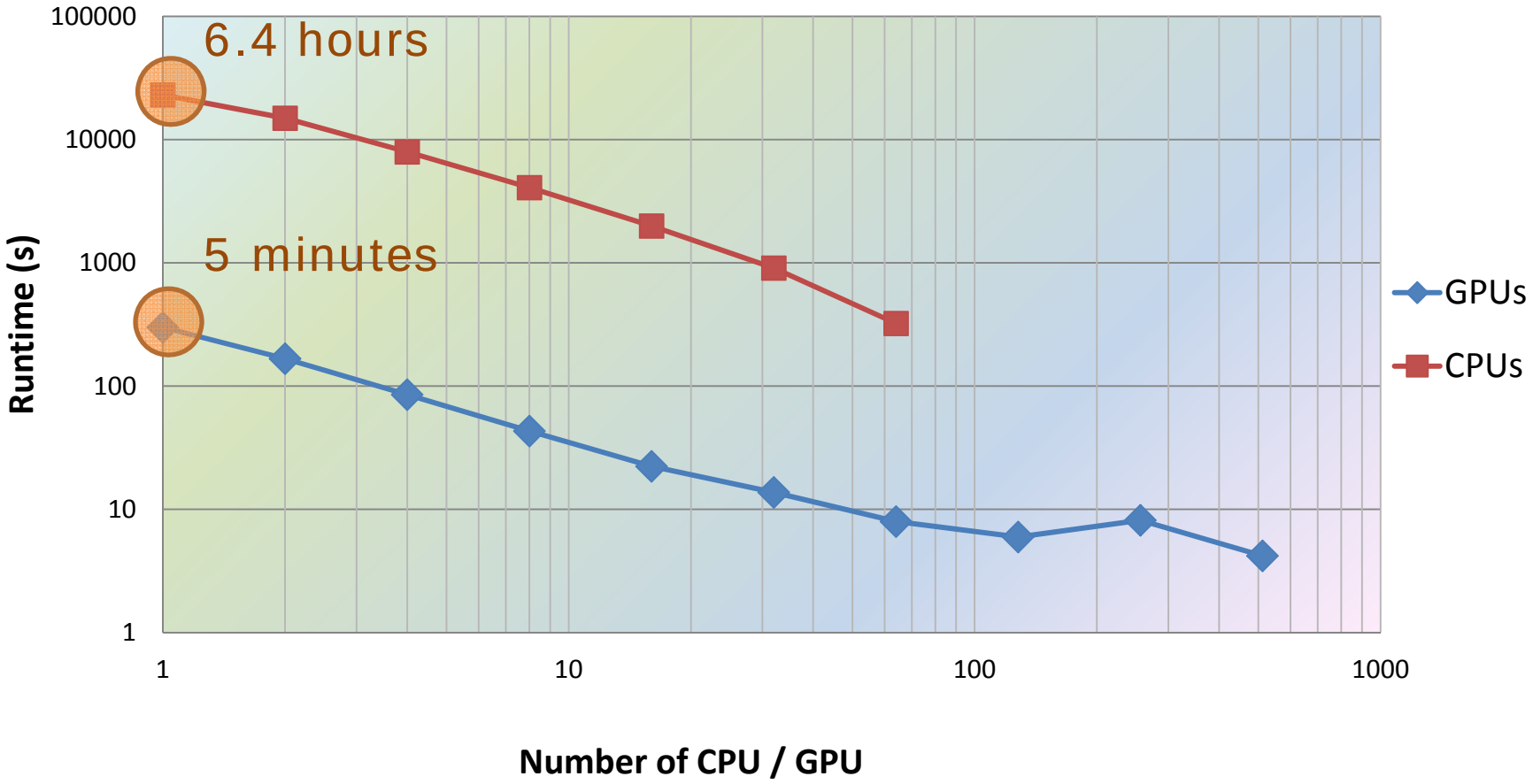
8 GPU 400km×800km
（100m mesh）

within 3 min

Performance on Multi-GPU



Tsubame 2.0



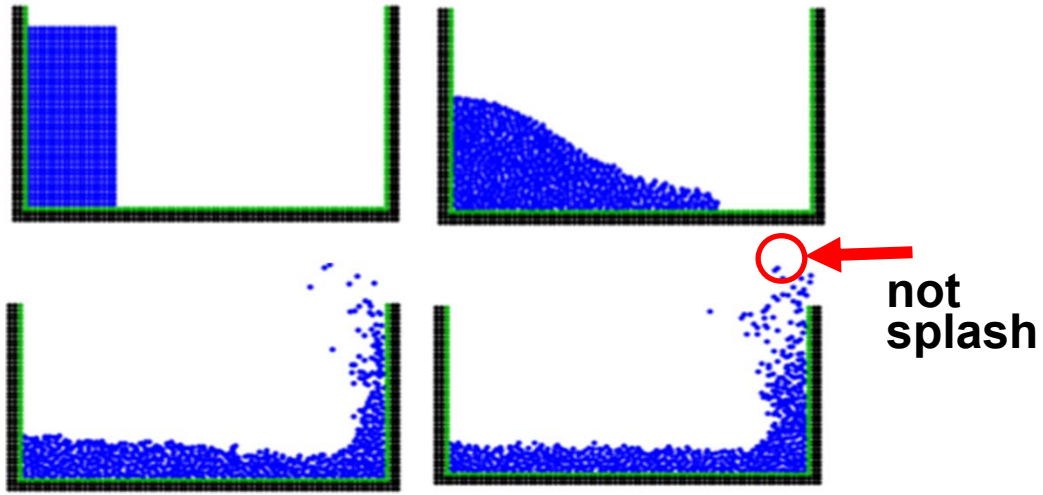
Two-Phase Flow Simulation

Particle Method
ex. SPH

Low accuracy
< $10^{6\text{-}7}$ particles

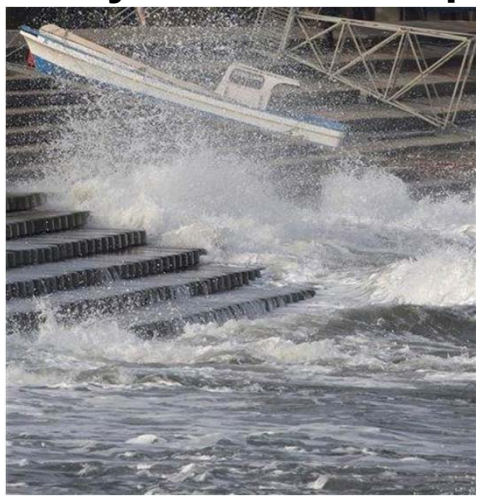
Mesh Method (Surface Capture)

- Navier-Stokes solver : Fractional Step
- Time integration : 3rd TVD Runge-Kutta
- Advection term : 5th WENO
- Diffusion term : 4th FD
- Poisson : MG-BiCGstab
- Surface tension : CSF model
- Surface capture : CLSVOF(THINC + Level-Set)

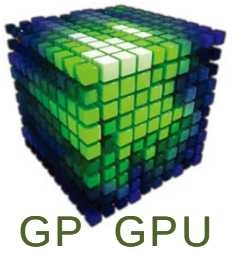


Numerical noise and unphysical oscillation

High accuracy > $10^{8\text{-}9}$ mesh points



Sparse Matrix Solver

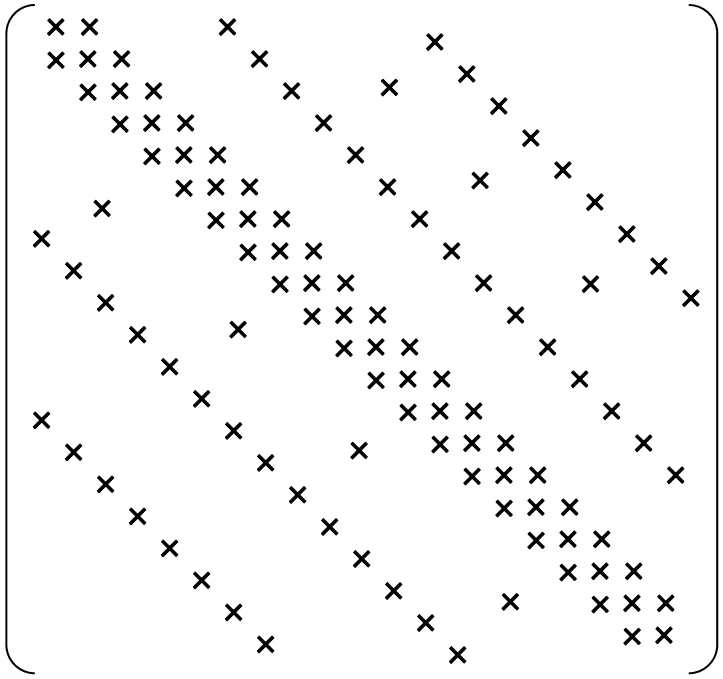


$$Ax = b \quad \textit{for} \quad \nabla \cdot \left(\frac{1}{\rho} \nabla p \right) = \frac{\nabla \cdot \boldsymbol{u}}{\Delta t}$$

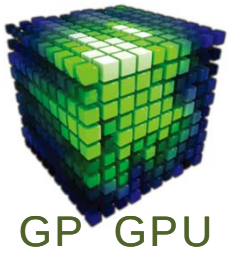
Krylov sub-space methods:
CG, BiCGStab, GMRes, , ,

Pre-conditioner:
Incomplete Cholesky,
ILU, MG, AMG,
Block Diagonal Jacobi

Non-zero Packing:
CRS → ELL, JDL



BiCGStab + MG Pre-conditioner



Collaboration with
Mizuho Information & Research Institute

Set $k = 0 \quad r_0 = p_0 = M^{-1}(b - Ax_0)$

for $k = 0;\ k < N;\ k{+}{+};$

$$\alpha_k = \frac{(r_0, r_k)}{(r_0, M^{-1}Ap_k)} \qquad q_k = r_k - \alpha_k M^{-1}Ap_k \qquad \omega_k = \frac{(q_k, M^{-1}Aq_k)}{(M^{-1}Aq_k, M^{-1}Aq_k)}$$

$$x_{k+1} = x_k + \alpha_k p_k + \omega_k q_k$$

$$r_{k+1} = q_k - \omega_k M^{-1}Aq_k$$

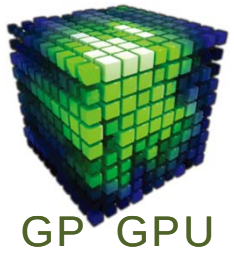
if $(r_{k+1}, r_{k+1}) < \varepsilon^2 (b, b)$ exit;

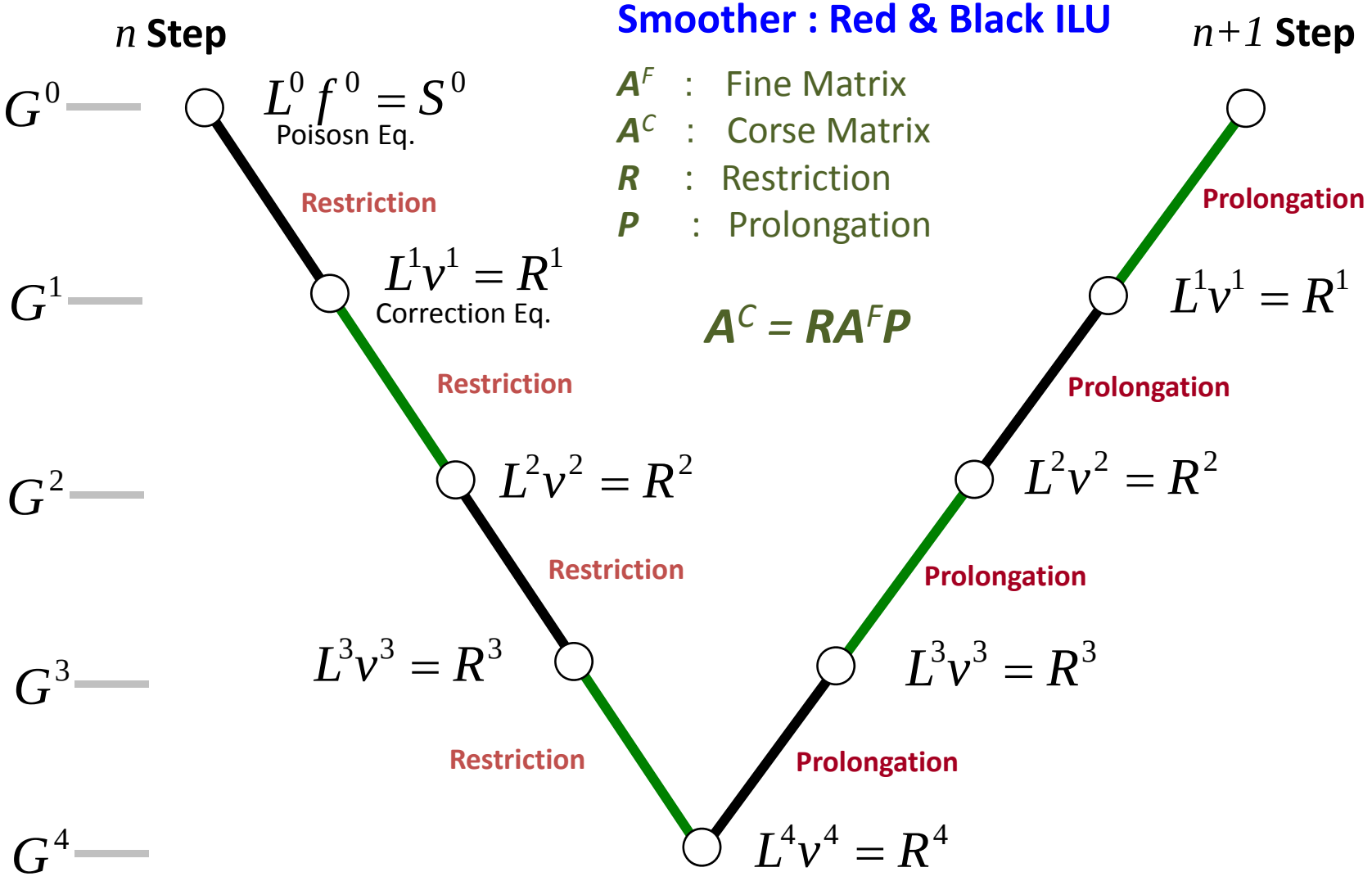
$$\beta_k = \frac{(r_0, r_{k+1})}{\omega_k (r_0, M^{-1}Ap_k)}$$

$$p_{k+1} = r_{k+1} + \beta_k (p_k - \omega_k M^{-1}Ap_k)$$

loop end

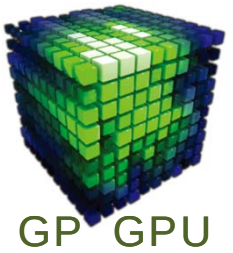
MG V-Cycle



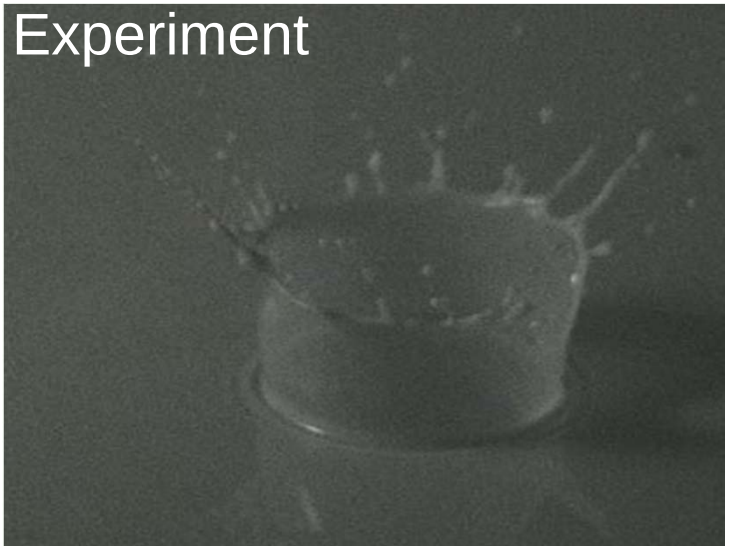


Milk Crown Formation

4.0 m/sec impact speed

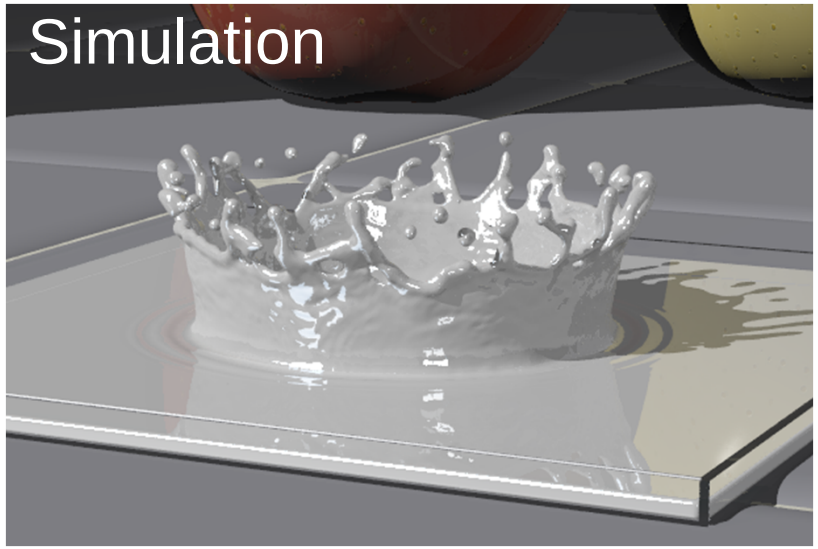


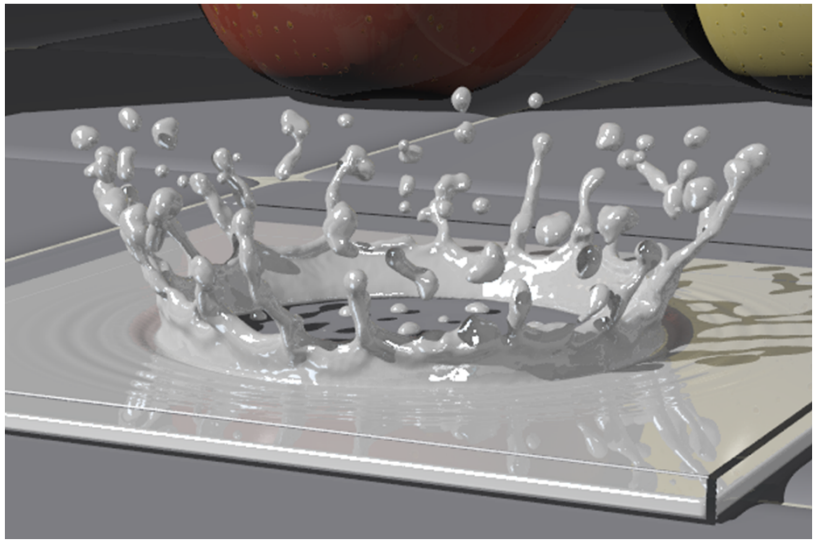
Experiment



Gunji, ishii, Saito, Sakai

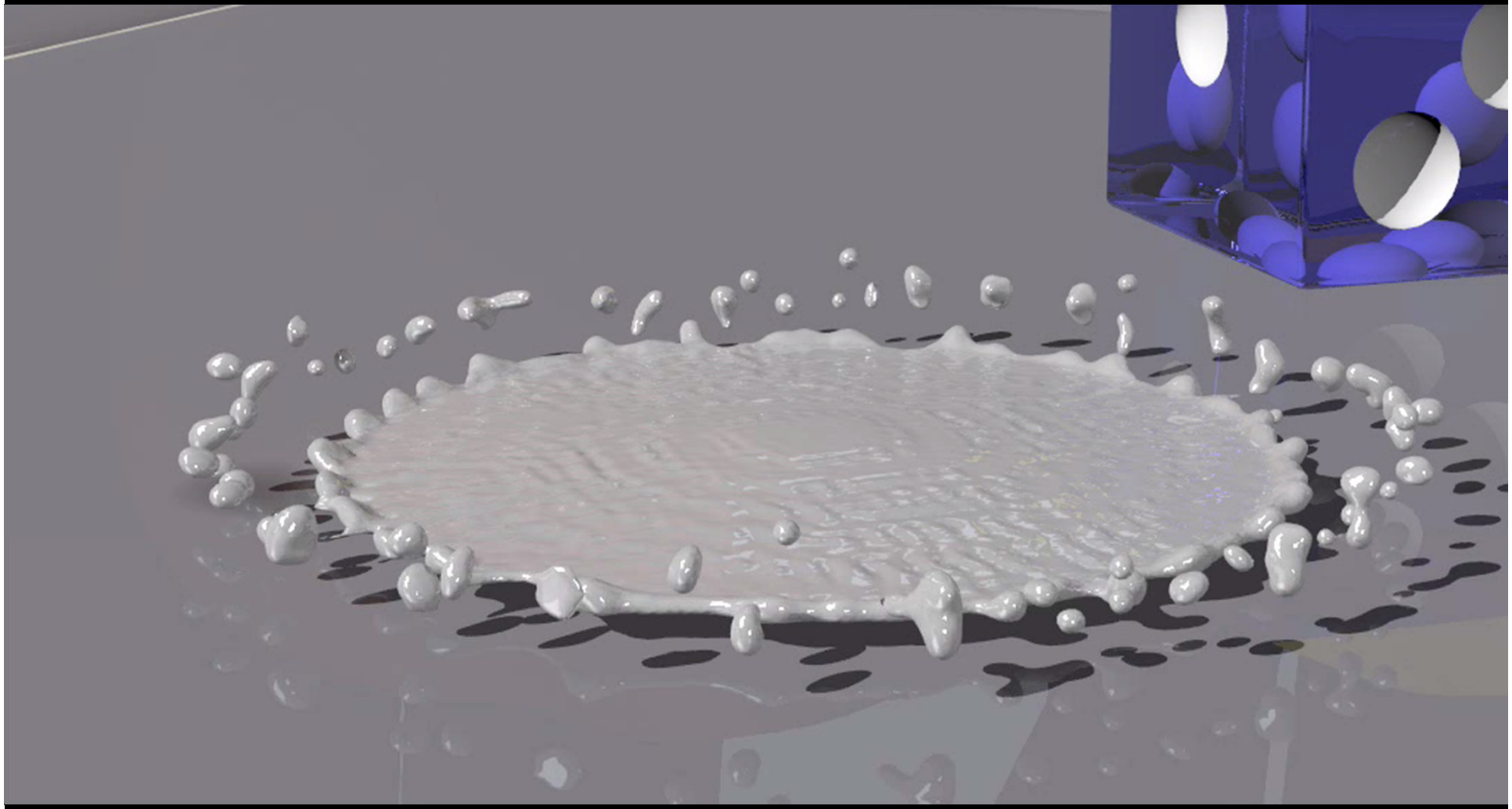
Simulation

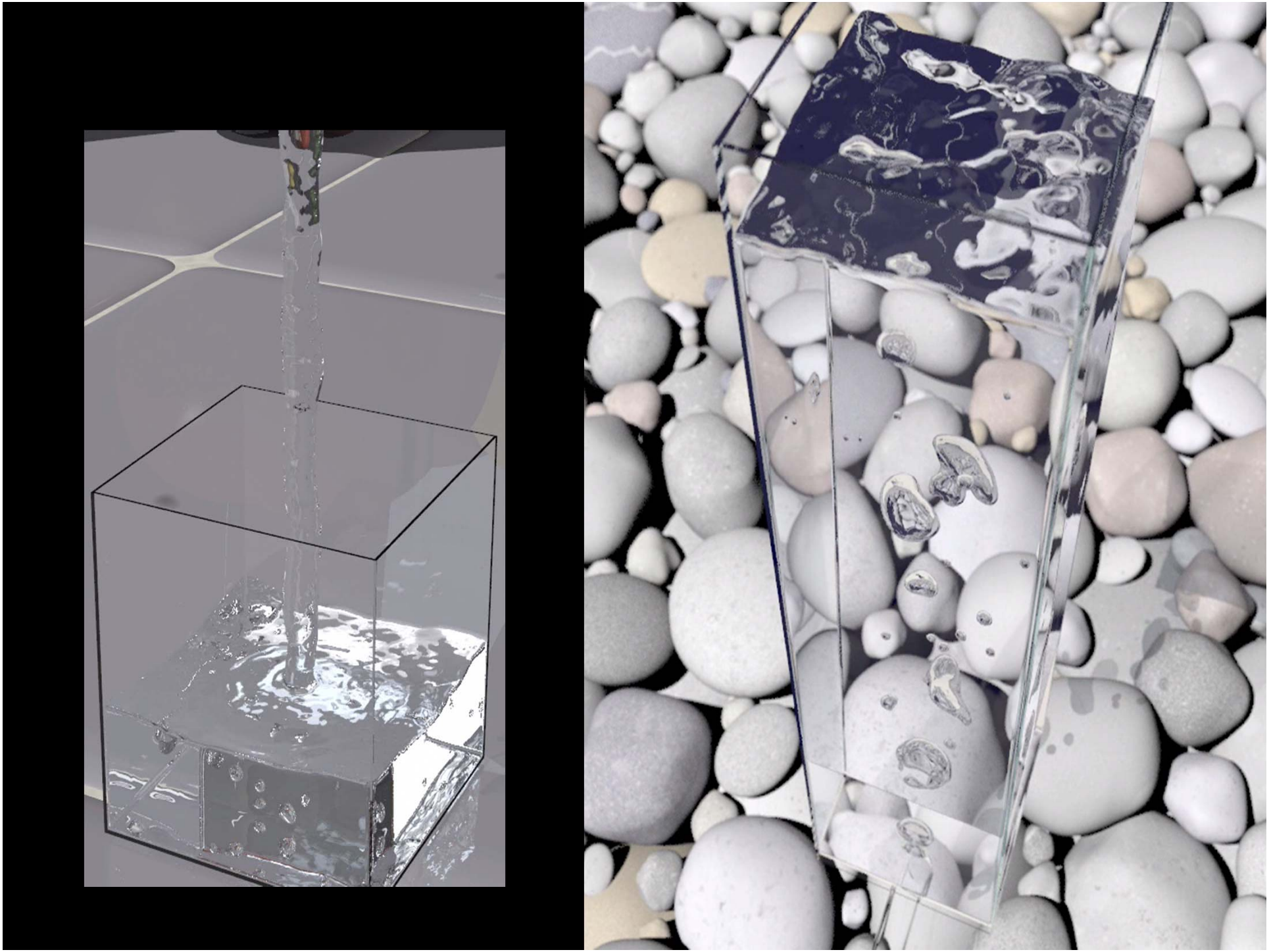




Milk Crown

Drop on dry floor





Initial stages of dam-break flow

P.K.Stanby, A.Chegini and T.C.D.Barnes (1998)

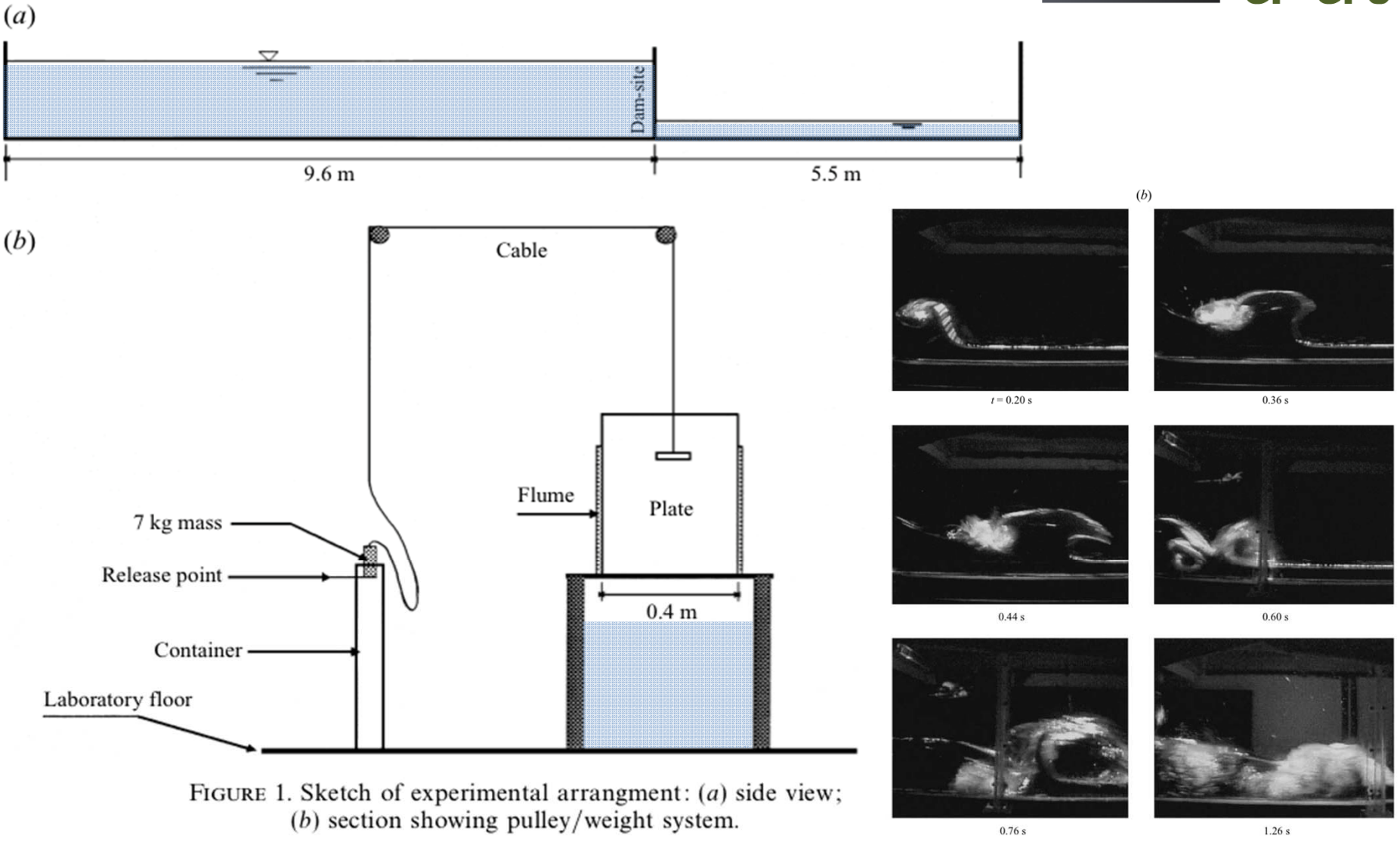
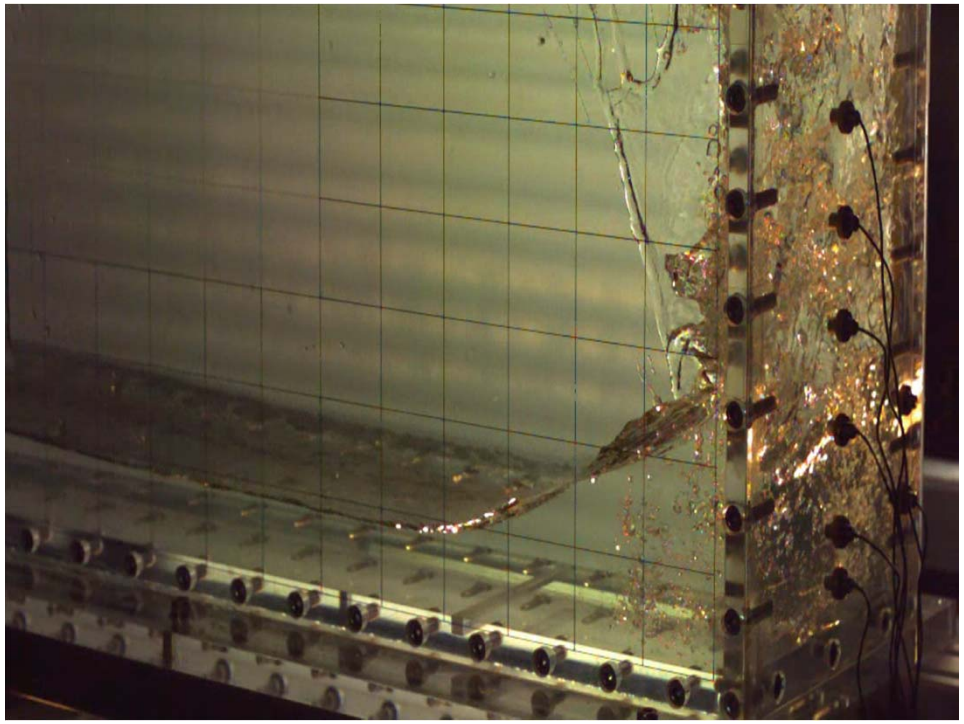


FIGURE 1. Sketch of experimental arrangment: (*a*) side view; (*b*) section showing pulley/weight system.

Experiment

Simulation



Collaboration: Prof. Hu and Dr. Sueyoshi, RIAM, Kyusyu University

MULTI-GPU Performance

（**TSUBAME 1.2**）