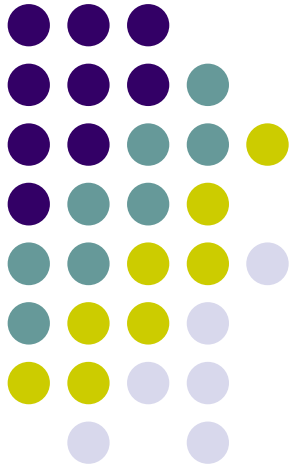
2017
Practical Parallel Computing
(実践的並列コンピューティング)
No. 1



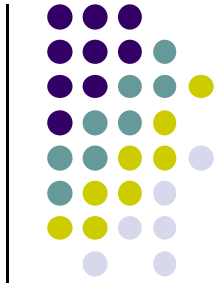
Overview of the course &
Basic usage of TSUBAME

Toshio Endo

School of Computing & GSIC

endo@is.titech.ac.jp

Purpose of This Course



- To learn parallel computing practically
  - Lecture + Practice
  - We use the TSUBAME supercomputer from this room
- Plan
  - Lecture about libraries/languages for parallel computing
    - OpenMP
    - MPI
    - CUDA (a programming environment for GPUs)

Materials, announcements will be found at OCW web page

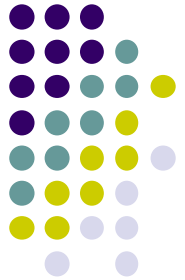
- http://ocw.titech.ac.jp

→ Search “practical parallel computing”

Required Knowledge

- This course uses C language as basis
  - Pointers, malloc/free
  - Relation between pointers and arrays
  - Knowledge of Pthread, Java threads is useful, but not required
- Basic Linux commands
  - ls, cp, mkdir…
  - Compilation, make would be helpful

Overview and Credits



- Part 1: OpenMP for shared memory parallel programming
- Part 2: MPI for distributed memory parallel programming
- Part 3: GPU programming

Your score will be determined by the followings

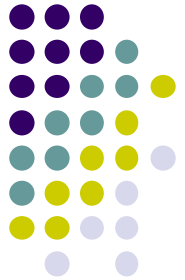
- There is homework for each part. Submission of reports for 2 parts is required
  - The due date will be about two weeks after each part finished
  - (You can submit reports more)
- Also attendances will be considered

講義の流れと単位認定

- Part 1: OpenMPによる共有メモリ並列プログラミング
- Part 2: MPIによる分散メモリ並列プログラミング
- Part 3: GPUプログラミング

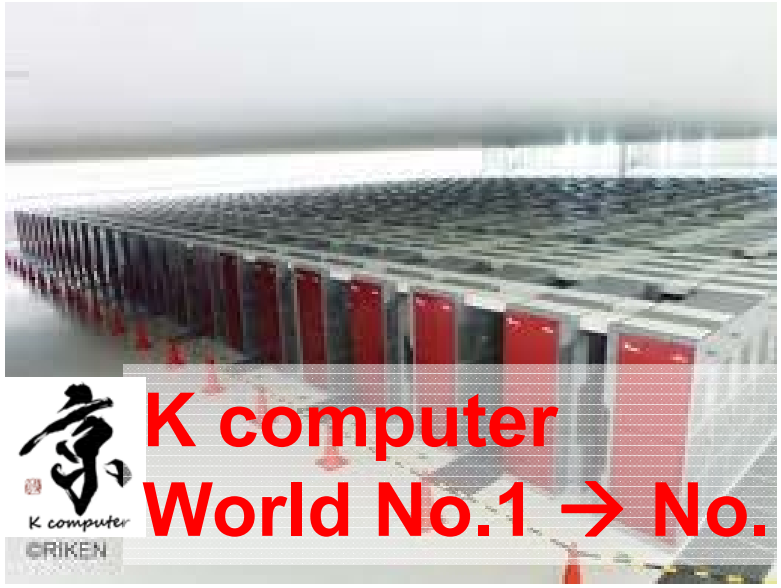
下記により採点・単位認定

- 各パートで課題を出す。2つのパートのレポート提出を必須とする
  - ∠切は、各パート終了の約2週間後
  - (それ以上のレポート提出してもよい)
- 出席点

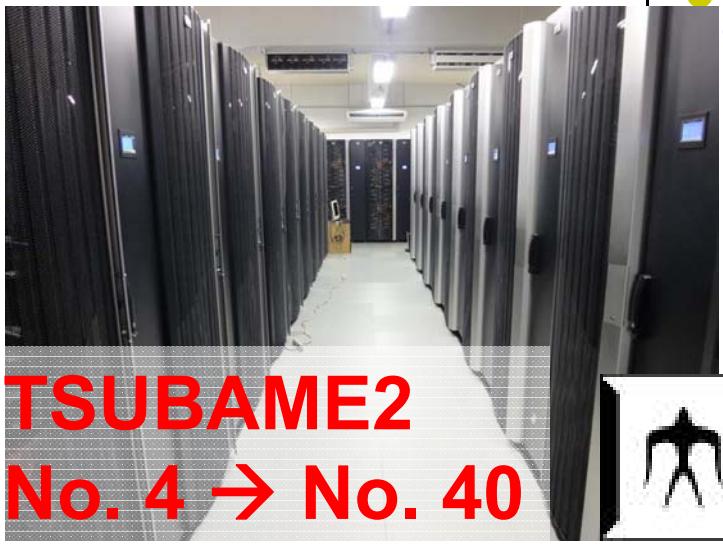


What are supercomputers?

Variety of Supercomputers (SC)



K computer
World No.1 → No. 7



TSUBAME2
No. 4 → No. 40

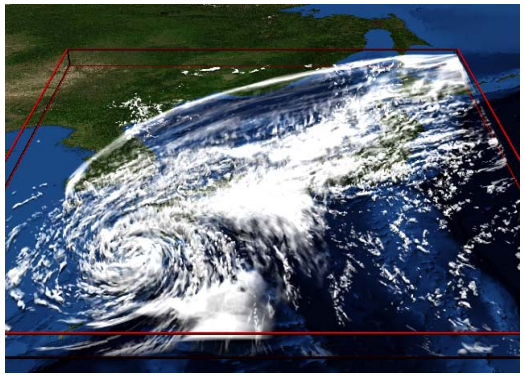
Oakforest-PACS

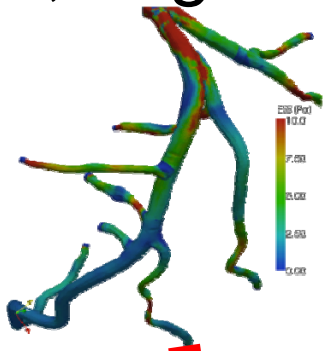
Jaguar

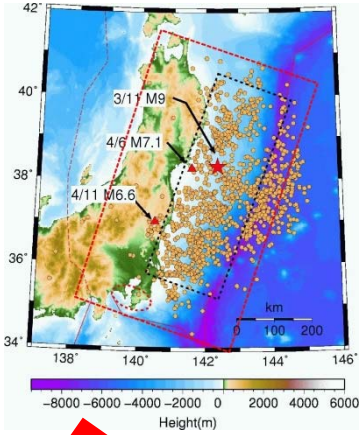
Taifulight
World No.1

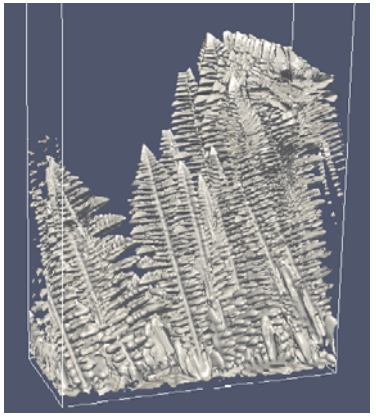
What are Supercomputers used for?

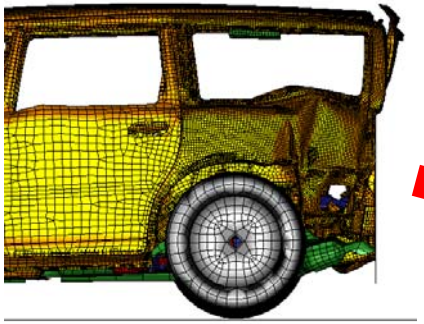
Simulations and Big-data analysis are important for area of science, engineering, social…

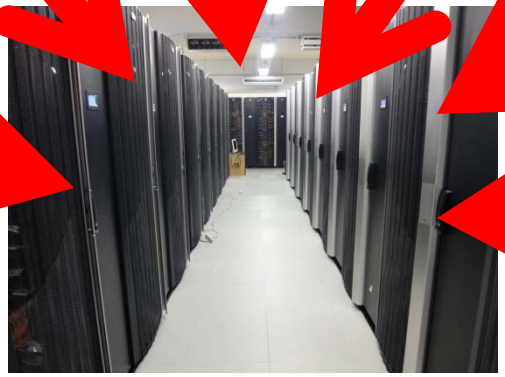












Difference with "Normal" Computers

- SCs are computers that support much faster and much larger computation than normal computers
  - Speeds are often compared in "Flops": The number of possible add/subtract/multiplication operations per second

~50,000x!

~100GFlops (10^{11} times per sec)

5.7PFlops (5.7×10^{15} times per sec)

<<

<<

PC in 1980

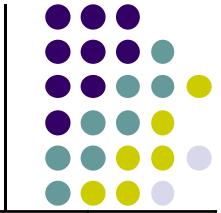
Cray-1 (1976)
160MFlops

Modern Cell Phone

Development of Supercomputers (1)

(from www.top500.org)

Linpack benchmark speed (Gflops)



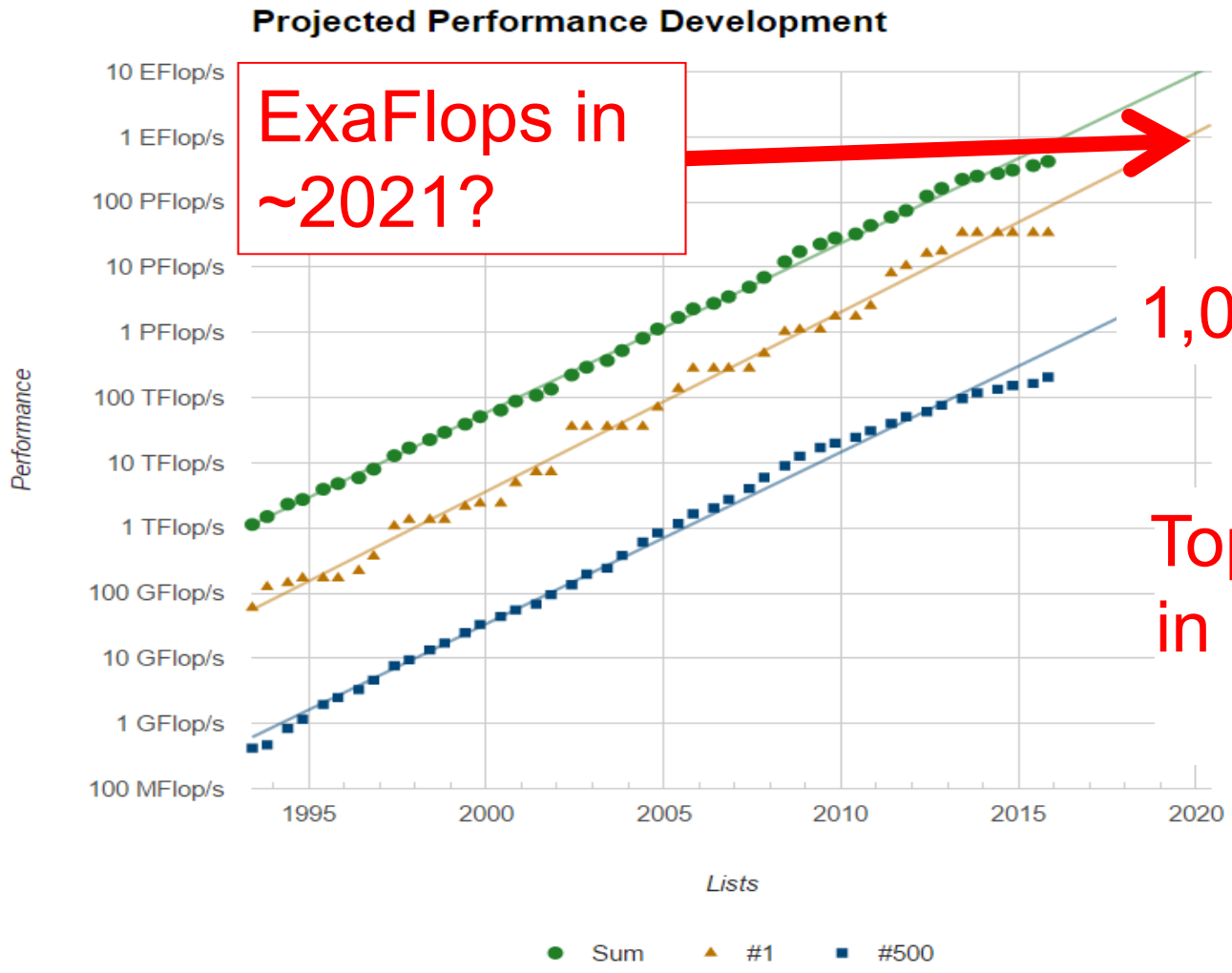
60GFlops

500,000x Faster in 20 years

33.9PFlops

Jun-93	CM-5/1024	60	LANL	US
Nov-93	Numerical Wind T		AL	Japan
Jun-94	XP/S140		NL	US
Nov-94	Numerical Wind Tunnel		IAL	Japan
Jun-95	↓	1		Japan
Nov-95	↓	170		Japan
Jun-96	SR2201/1024			
Nov-96	CP-PACS			
Jun-97	ASCI I			
Nov-97	↓	1338	↓	
Jun-98	↓	1338	↓	
Nov-98	↓	1338	↓	U
Jun-99	↓	2121	↓	US
Nov-99	↓	2379	↓	US
Jun-00	↓	2379	↓	US
Nov-00	ASCI White	4938	LLNL	US
Jun-01	↓	7226	↓	US
Nov-01	↓	7226	↓	US
Jun-02	Earth-Simulator	35860	ES Center	Japan
Nov-02	↓	35860	↓	Japan
Jun-03	↓	35860	↓	Japan
Nov-03	↓	35860	↓	Japan
Jun-04	↓	35860	↓	Japan
Nov-04	BlueGene/L beta	70720	IBM/DOE	US
Jun-05	BlueGene/L	136800	DOE/NNSA/LLNL	US
Nov-05	↓	280600	↓	US
Jun-06	↓	280600	↓	US
			↓	US
			↓	US
			↓	US
Jun-08	RoadRunner	1026000	↓	US
Nov-08	↓	1105000	↓	US
Jun-09	↓	1105000	↓	US
ov-09	Jaguar	1759000	ORNL	US
10	↓	1759000	↓	US
	Tianhe-1A	2566000	NSC	China
	K computer	8162000	RIKEN AICS	Japan
N	↓	10510000	↓	Japan
J	Sequoia	16324000	DOE/NNSA/LLNL	US
Nov-12			DOE/SC/ORNL	US
Jun-13	Tianhe-2	33863000	NSC	China
Nov-13	↓	33863000	↓	China

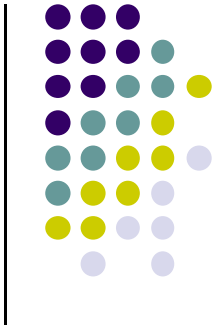
Development of Supercomputers (2)
(from www.top500.org)



1,000x in 10 years

Top Supercomputer in 1993 (60GFlops) ≒ PC in 2011

Why are Speed & Size Important?



- For simulation & big-data analysis, large number of computations should be done speedily

⇒Want to obtain forecast of tomorrow weather by tomorrow (of course!)

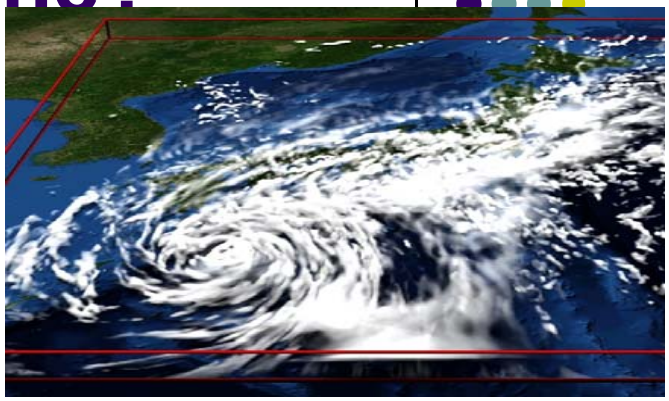
⇒Want to develop and sell new medicine (than competitors)

- For simulation & big-data analysis, storing large scale data is needed

⇒Want to make discovery by comparing mass genome data

⇒Want to visualize motion of molecules for every time step

How is Weather Forecast done?



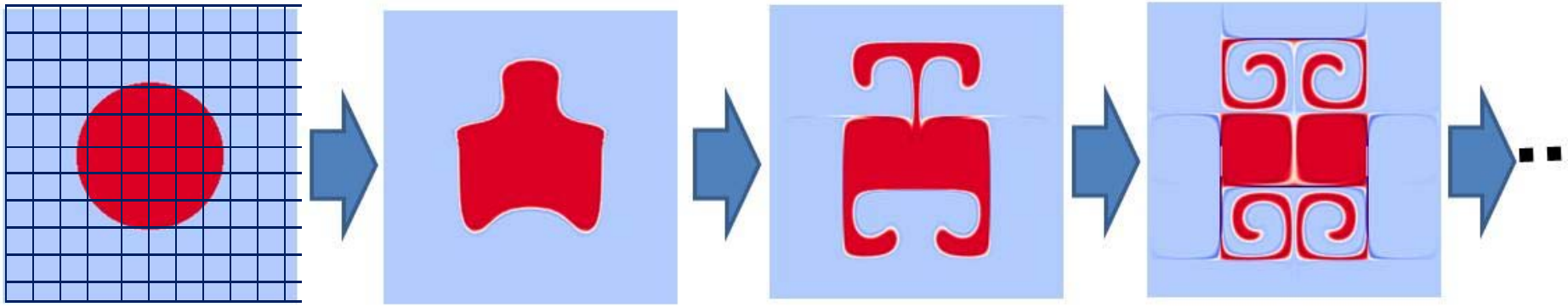
Motions of air, clouds, water are expressed by differential equations

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} = -\frac{1}{\rho}\nabla p + \nu \Delta \boldsymbol{v} - g\hat{\boldsymbol{z}}$$

But no analytical solution for them, generally

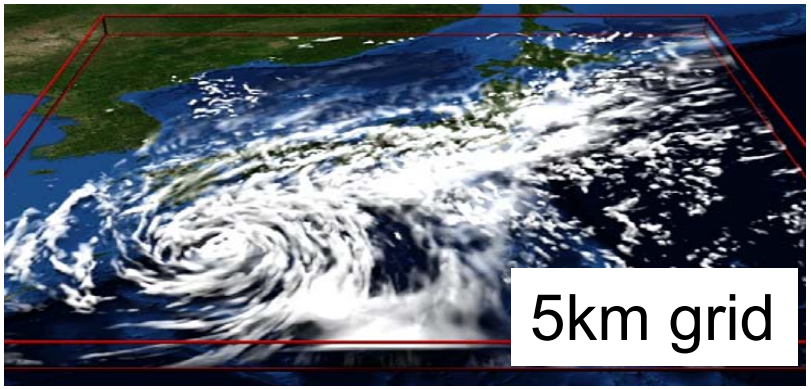
⇒ Instead, space and time are discretized

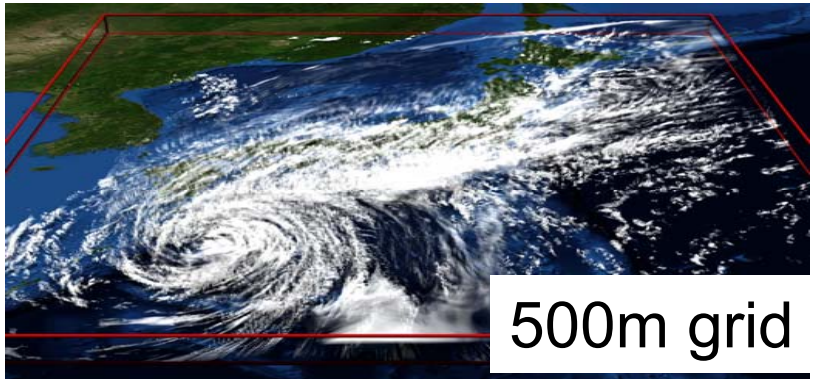
- The space is divided into small grids, expressed as an array
  ⇒ Each array element should be computed
- The time is divided into time steps
  ⇒ After a time step is computed, we go to next step, and so on



Why is Speed Important?

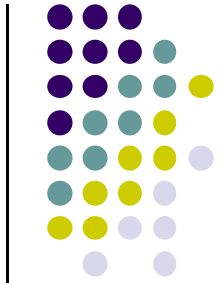
- Since we have to compute all points for every time step, computational complexity is
  - O(x-size × y-size × z-size × time-steps)





For better prediction, we need to make grid finer (arrays larger)
If resolution is 10x higher, we need 10000x computations! (10x10x10x10)
→ In future, we are going to 50m grid…

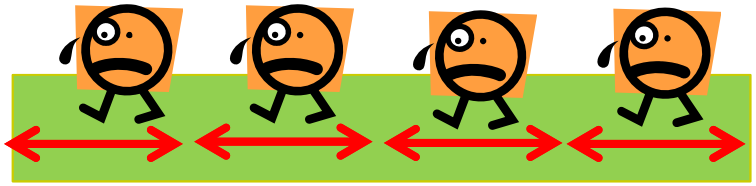
Why are SCs Fast?



- Do SCs have 10THz CPUs? → No!!
- Basic idea : If multiple workers work cooperatively and simultaneously, they can do great tasks than a single worker ⇒Parallel execution

A work is cultivating a large field

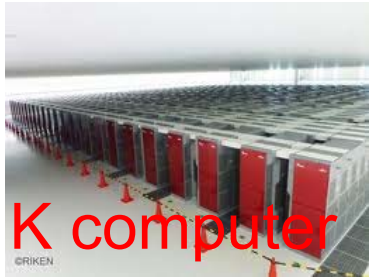
Multiple workers are working together → fast!



Hierarchical SC Structure

- System = Many nodes (=computers) + External storage
  - Parts are connected by Network
- Node = Several processors (CPU etc.) + Memory + Local storage
  - Parts are connected by PCI-e, QPI, etc.
- Processor = Several cores + Cache

Structure of TSUBAME2 and K



System	1408 nodes	5.73PFlops	88000 nodes	11.26PFlops
Node	2 CPUs + 3 GPUs	140GFlops + 3930GFlops = 4070GFlops	1 CPU	128GFlops
Processor	CPU: 6 cores GPU: 14 SMXs	CPU: 70GFlops GPU: 1310GFlops	8 cores	128GFlops
Core	CPU core: 2.93GHz x 4 = 11.7GFlops GPU SMX: 0.73GHz x 128 = 93.4GFlops		2GHz x 8 = 16GFlops	

GPUs are focused in Part 3
In this course

- “TSUBAME2.5 Guidance” is explained here

TSUBAME Group in This Course

- Students of this course will become members of "t2g-ppcomp" TSUBAME group
  - Use t2sub command with "-W grouplist=t2g-ppcomp"

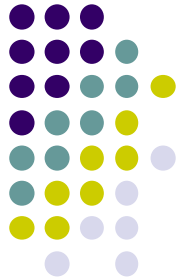
Please do the following by Apr 20 (earlier is better)

- Please make your account on TSUBAME2
  - Tokyo Tech Portal (portal.titech.ac.jp) → TSUBAME2 Portal
- Please send an e-mail to ppcomp@el.gsic.titech.ac.jp

Subject: TSUBAME2 ppcomp account
To: ppcomp@el.gsic.titech.ac.jp

Lab name/Department name
School year
Name
Your TSUBAME account name

Note in This Course



Since each student logins to TSUBAME from this room,

- We recommend you to bring your laptop PC
  - If it is difficult, please practice in home
- Please install SSH terminal software
  - For Windows PC, Putty or Teraterm+SSH will work
- Make sure use can connect to Wifi network
  - We recommend “TokyoTech” Wifi network

Contact/ Information on the Course

- About this course
  - ppcomp@el.gsic.titech.ac.jp (Endo)
- Tokyo Tech OCW
  - http://ocw.titech.ac.jp

→ Search with “Practical Parallel Computing”School of 「実践的並列コンピューティング」で検索