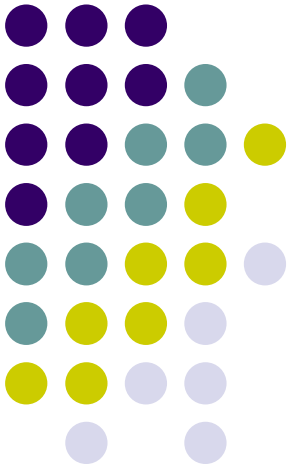
Practical Parallel Computing (実践的並列コンピューティング)

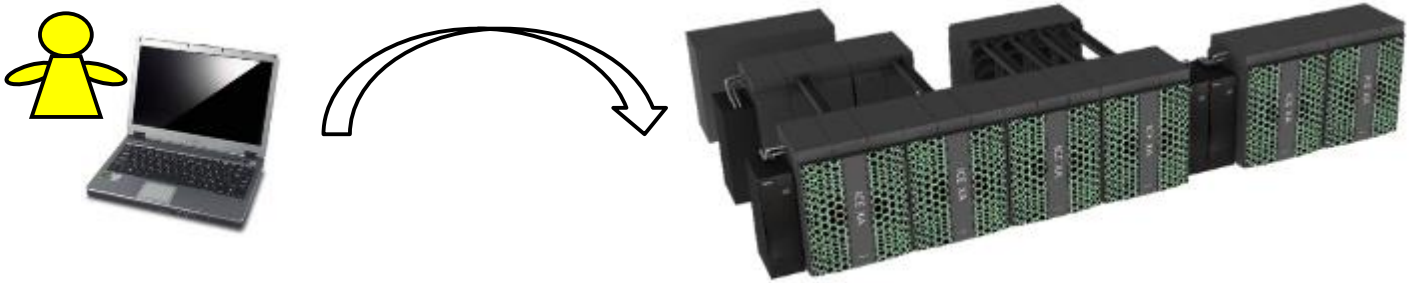
Introduction (1)
Apr 11, 2022



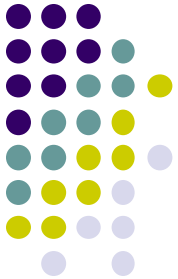
Toshio Endo
School of Computing & GSIC
endo@is.titech.ac.jp

Purpose of This Course

- To learn parallel computing practically
  - Lecture + Practice
  - We will use the TSUBAME supercomputer from your PC



Overview of This Course

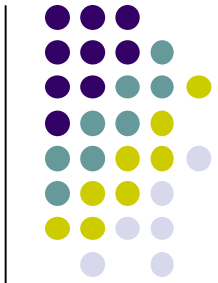


- Part 0: Introduction
  - 2 classes including today
- Part 1: OpenMP for shared memory programming
  - 4 classes
- Part 2: GPU programming
  - 4 classes, OpenACC and CUDA
- Part 3: MPI for distributed memory programming
  - 3 classes

In 2022, talk and presentations are done in Japanese

偶数年度は日本語です

Credits / 単位認定



Your score will be determined by the followings

- Each part (OpenMP, GPU, MPI) has homework.

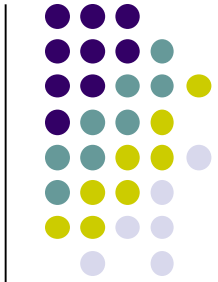
Reports submission for 2 parts is required

- The due date will be about 1.5 weeks after each part finished
- (You can submit more than 2)

下記により採点・単位認定

- OpenMP, GPU, MPIの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1.5週間後
  - (それ以上のレポート提出してもよい)

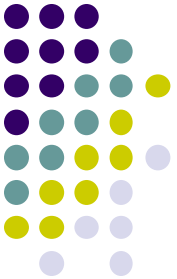
What We Do



We do

- Parallel programming by yourself
  - C language + something
  - CPU parallel programming and GPU parallel programming
- To evaluate speed-up of programs on TSUBAME supercomputer

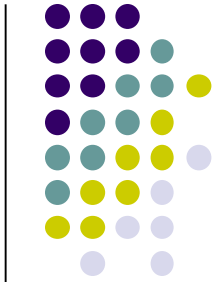
What We Don’t



We don’t

- To use Python
- To learn usage of machine learning framework
- To learn usage of parallel computation libraries
- To learn variety of parallel algorithms
  - Consider to attend “CSC.T526 High Performance Scientific Computing”
- To program network or client-server applications

Requirements (1)



- Knowledge of basic C language
  - Pointers, malloc/free
  - Relation between pointers and arrays

  - Knowledge of Pthread, Java threads is useful, but not required

Requirements (2)

- Knowledge of basic Linux commands
  - TSUBAME uses Linux OS
  - ls, cp, mkdir, gcc…
  - “make” command will help you
  - Optional: SSH public key authorization (SSH公開鍵認証)

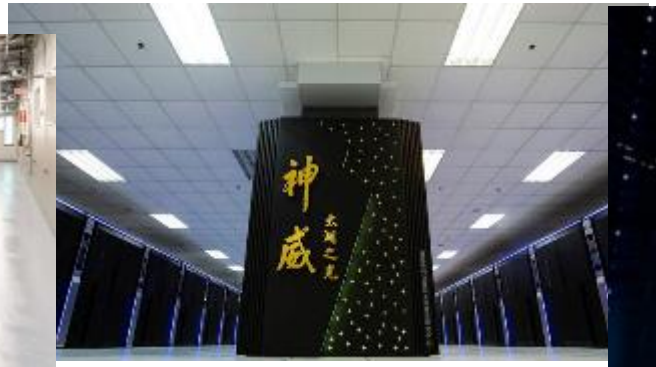
- A PC to log-in to TSUBAME3 and Internet connection
  - For the remote course, you will need to open both Zoom window and a terminal window on your PC ☹

Supercomputers

Fugaku
World No. 1

Tokyo Tech
TSUBAME3.0

Summit
World No.2

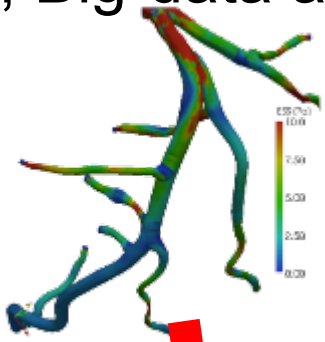


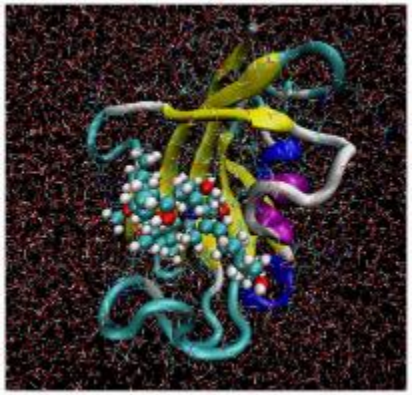
Taifulight

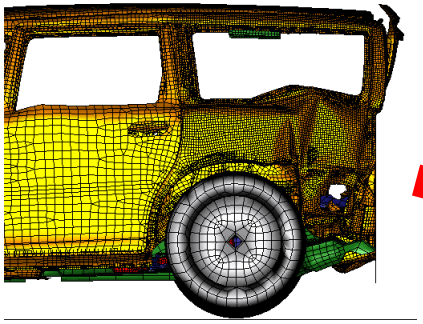
ABCI
AIST, Japan

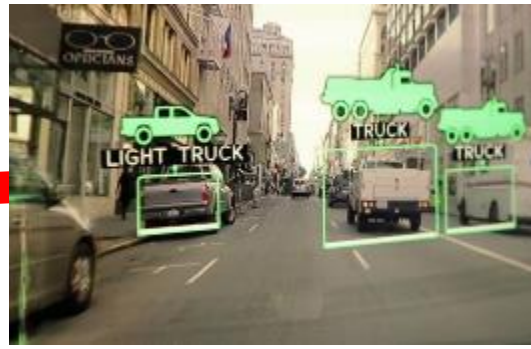
What are Supercomputers (SCs) used for?

- Simulations (Fluid dynamics, molecular dynamics, etc.)
- Machine learning, Big-data analysis





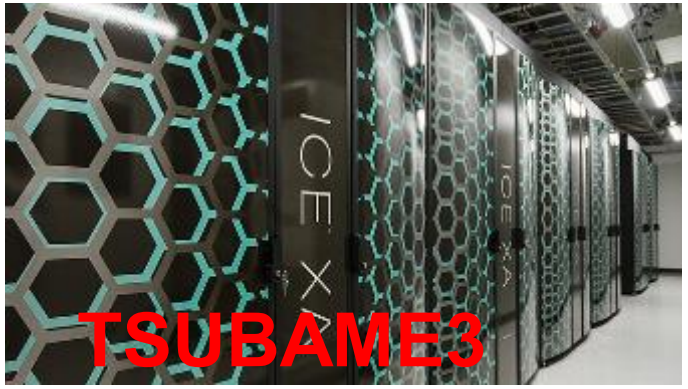




Difference with "Normal" Computers

- 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

~60,000x!



~200GFlops (2x10^{11} times per sec)

12PFlops (1.2x10^{16} times per sec)

<<

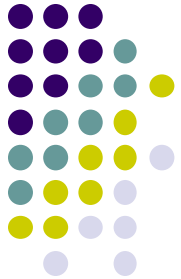
<<

PC in 1980

Cray-1 (1976)
160MFlops

Modern Cell Phone

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

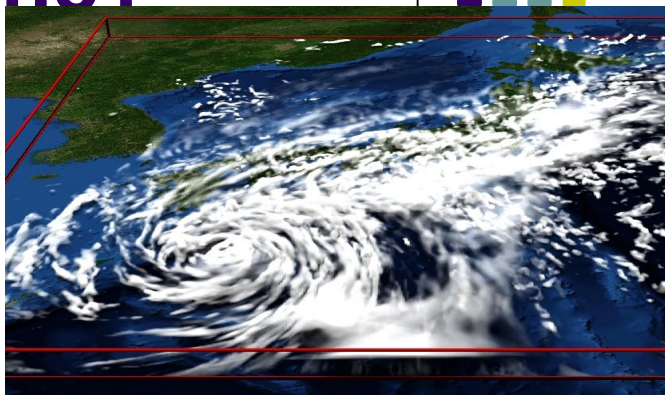
⇒Want to make discovery by comparing mass genome data

⇒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

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

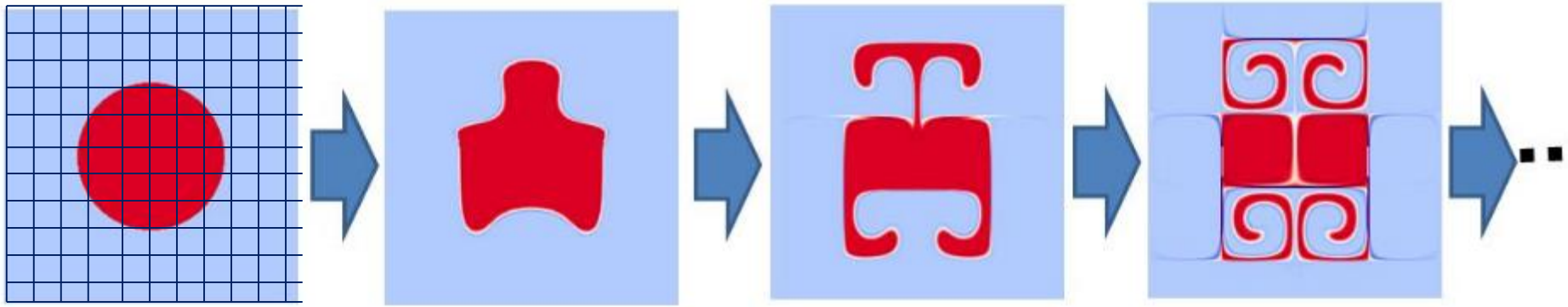


By Takayuki Aoki

But no analytical solution for them, generally

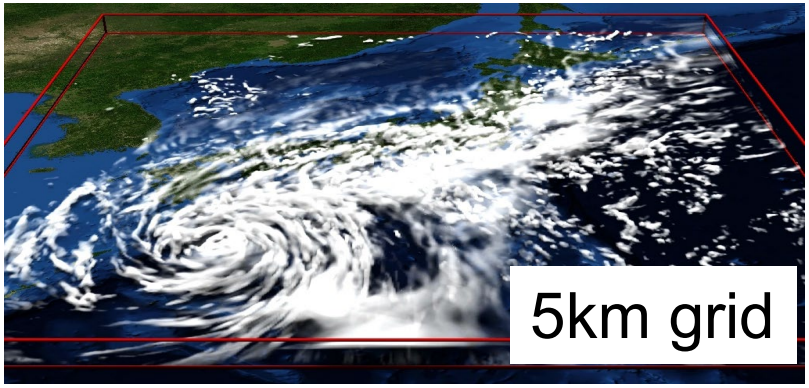
⇒ Instead, space and time are discretized

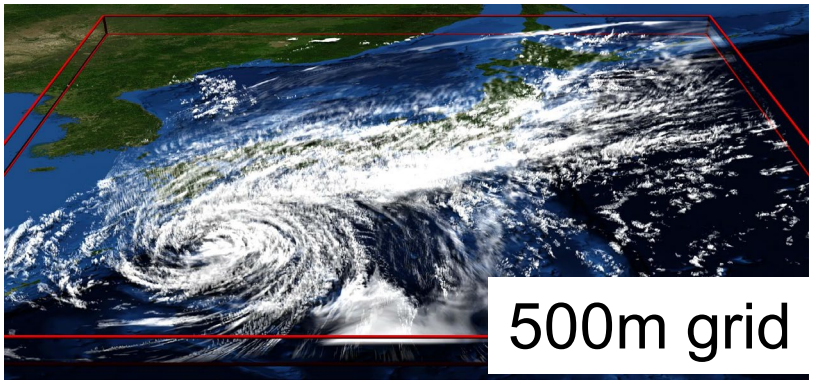
- The space is divided into small grids, expressed as an array
  ⇒ Each array element should be computed
- The time is divided into time steps
  ⇒ After a time step is computed, we go to next step, and so on



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 or smaller grids

Deep Learning (DL) needs SC

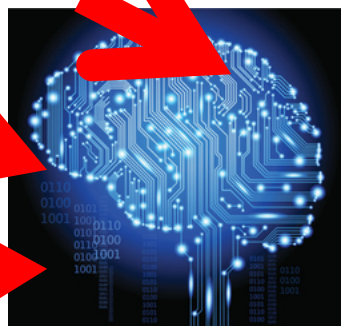
- Deep learning consists of "training phase" and "inference phase"
- Training requires supercomputers

In case of image recognition

Training

Done on supercomputers/cloud

Inputs are lots of images with correct answers

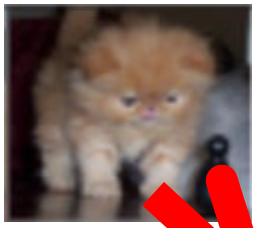


AI, actually numeric data

Copy data of AI

Inference

Done on cars/phones

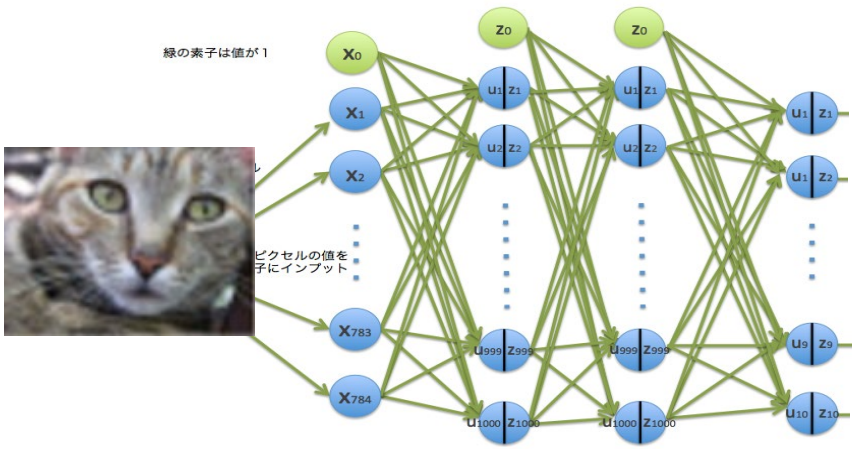


What's this?

A cat (75%)
A dog (24%?)

Training in DL Has Large Compute Complexity

- DL uses “network of neurons”
- (Roughly) “knowledge” is represented as connectivity of neurons
  - Sets of large parameter vectors, whose total sizes are >10^4, >10^8…
- Basic training method:
  - For an input image, it computes the current answer (≠ correct answer)
  - By using difference of current answer and correct answer, it updates its parameter vectors



Vertices are neurons
Arrows are connections,
which have numeric data

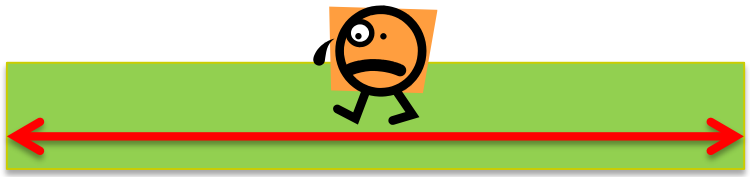
Compares current answer and correct answer, *“cat”*
→ It updates the knowledge (vectors) using the difference

Complexity of (Number of images × Sizes of parameter vectors × Number of repetitions) !!

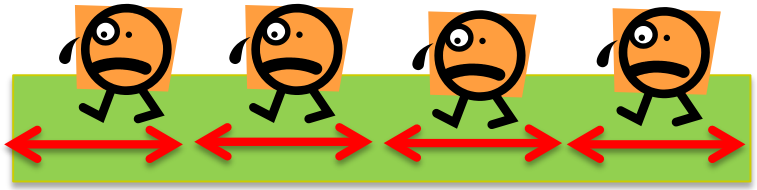
Why are SCs Fast?

- 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 worker is cultivating a large field

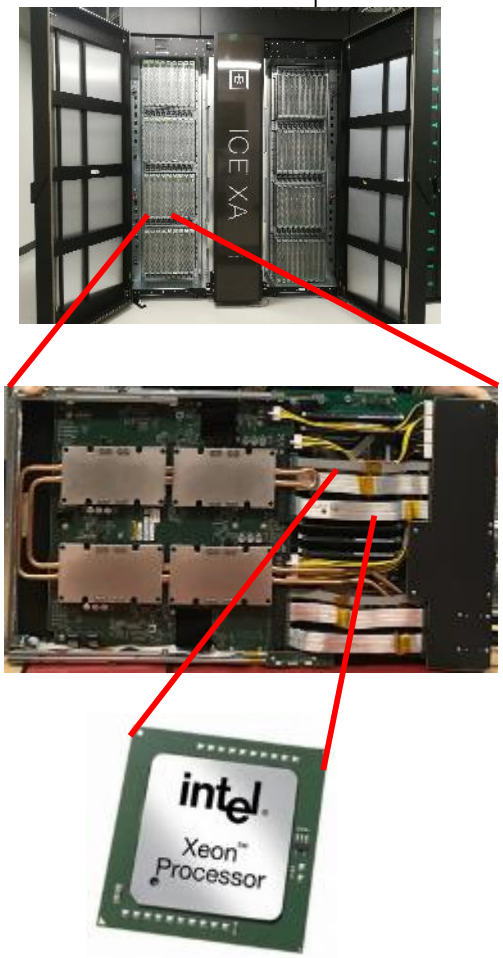


Multiple workers are working together → fast!



SC Structure is Hierarchical

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



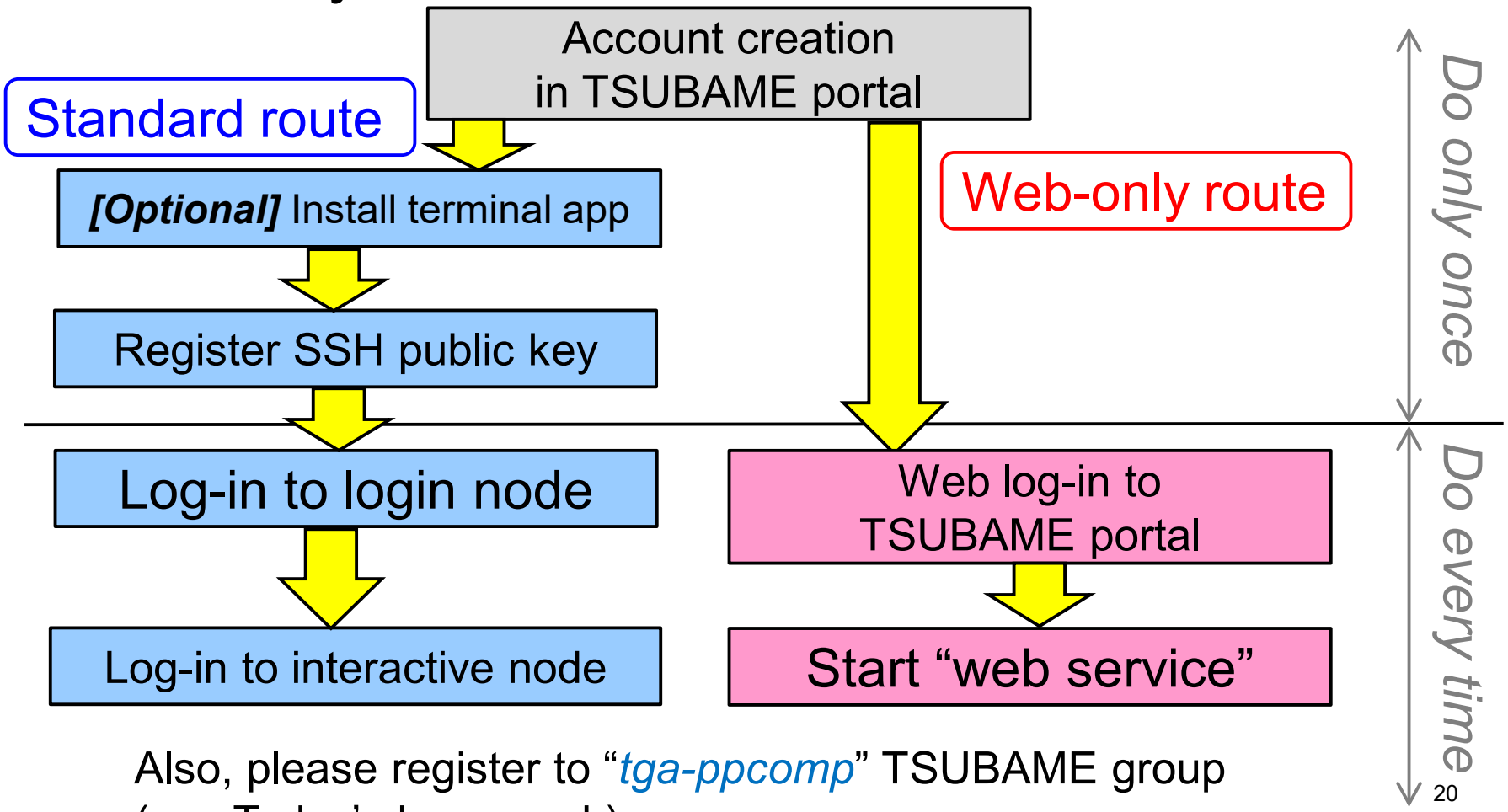
Structure of TSUBAME3 and Fugaku

System	540 nodes	12.15PFlops	160,000 nodes	540PFlops
Node	2 CPUs + 4 GPUs	2 x 425GFlops + 4 x 5300GFlops = 22.05TFlops	1 CPU	3.38TFlops
Processor	CPU: 14 cores GPU: 56 SMXs	CPU: 425GFlops GPU: 5300GFlops	48 cores	3.38TFlops
Core	CPU core: 1.9GHz x 16 = 30.4GFlops GPU SMX: 1.48GHz x 64 = 94.6GFlops		2.2GHz x 32 = 70.4GFlops	

Here "Flops" shows speed in "double precision (FP64)"

Start to Use TSUBAME

- Two ways to use TSUBAME in this class



Also, please register to "*tga-ppcomp*" TSUBAME group
(see Today's homework)

Important URLs about TSUBAME

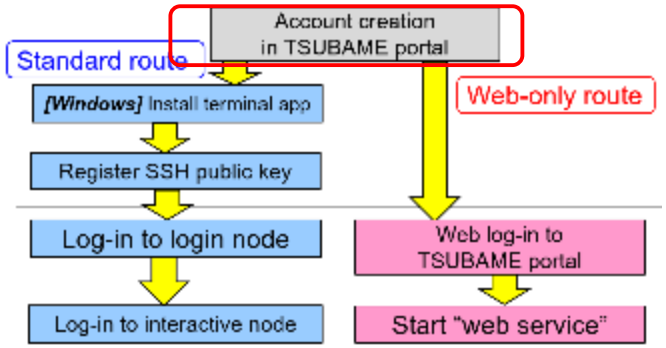
- TSUBAME official site
  - https://www.t3.gsic.titech.ac.jp
  - Manuals are here

- Tokyo Tech portal
  - https://portal.titech.ac.jp/

➔ Log-in and then click “TSUBAME portal”

Standard route | Web-only route

Account Creation in TSUBAME Portal



- Visit TokyoTech Portal / 東工大ポータル and log-in

Click “TSUBAME portal”

➔ If you are new to TSUBAME, you will be taken to account creation pages

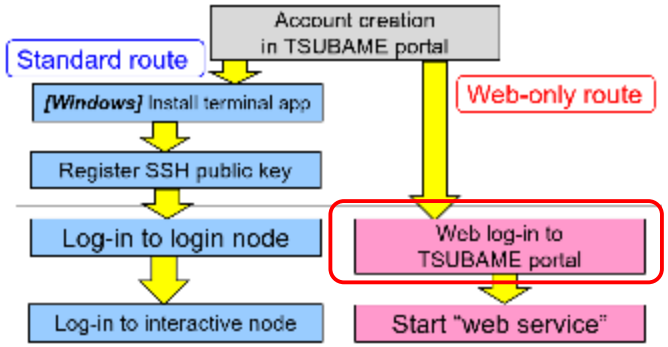
➔ You will get an account.

- Account name is same as the student ID (like 22M12345) for Tokyo Tech students

- For details, visit https://www.t3.gsic.titech.ac.jp/ and click “Getting Accounts” / “アカウント取得方法”

Web-only route

Log-in to TSUBAME Portal



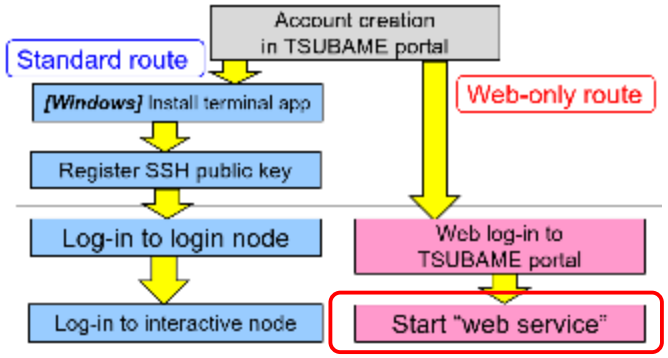
- Visit “Tokyo Tech portal”
- Log-in
- Click “TSUBAME portal”

Your account name

Will be used in “standard route”

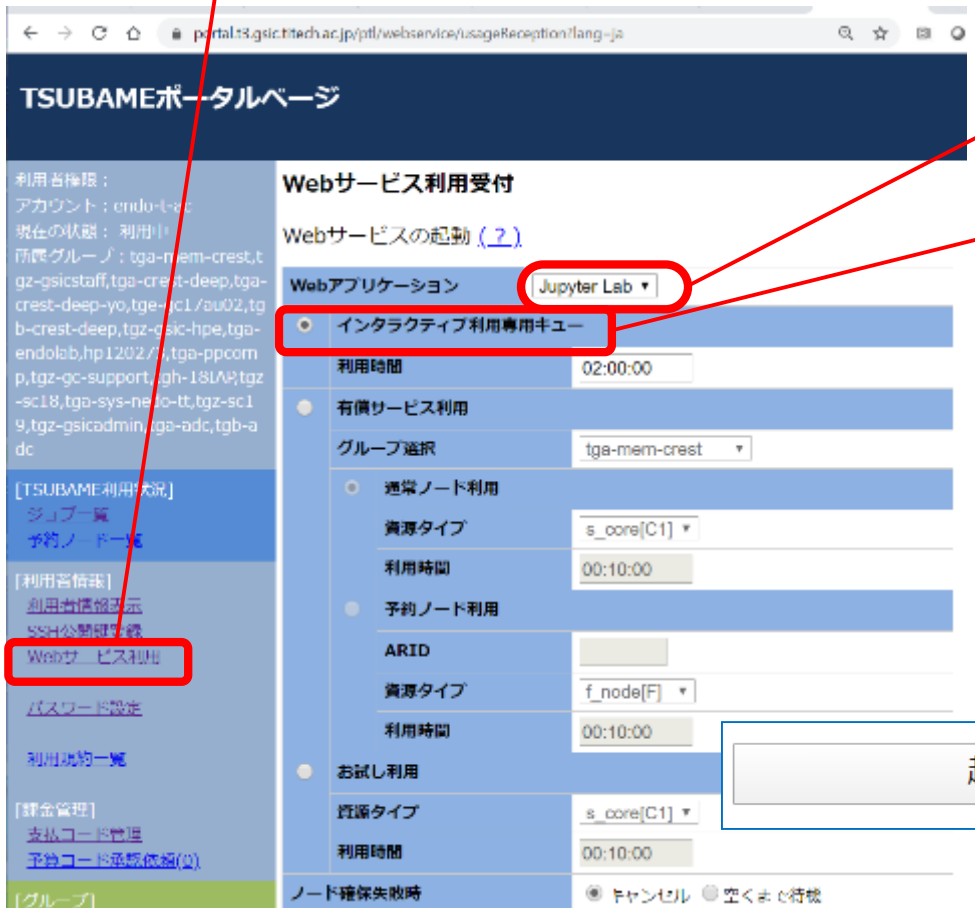
Used in the next page

Web-only route



Start “Web Service” (1)

- In TSUBAME portal, click Use Web service / Webサービス利用

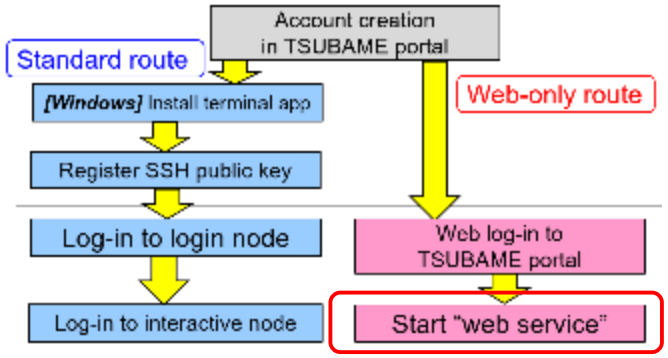


Choose
Jupyter Lab

Choose
Queue for interactive use only/
インタラクティブ専用キュー

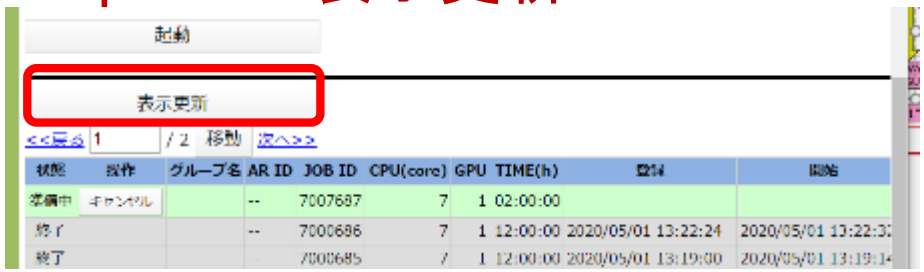
And click
Start-up / 起動 button

Web-only route

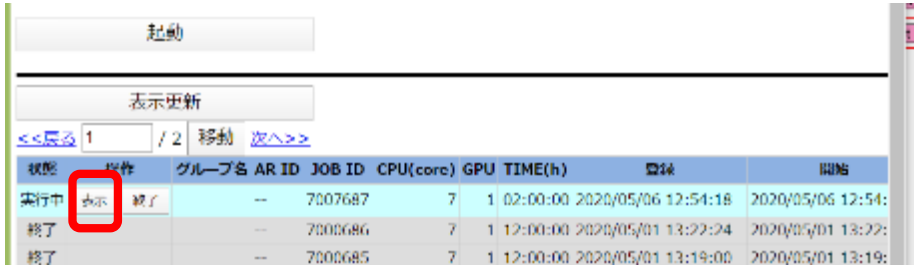


Start “Web Service” (2)

- When you see a green row, please wait and push update / 表示更新 button



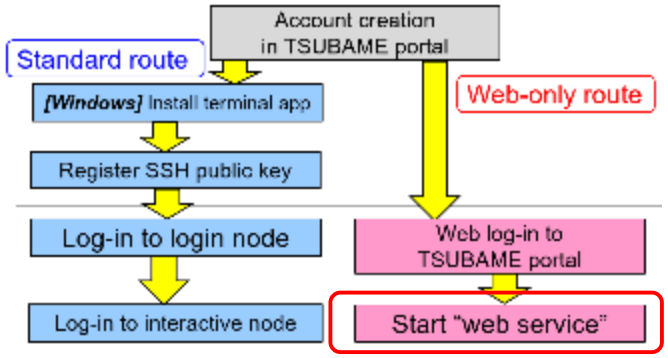
Wait for 1-2 minutes and “update”



Please be patient…

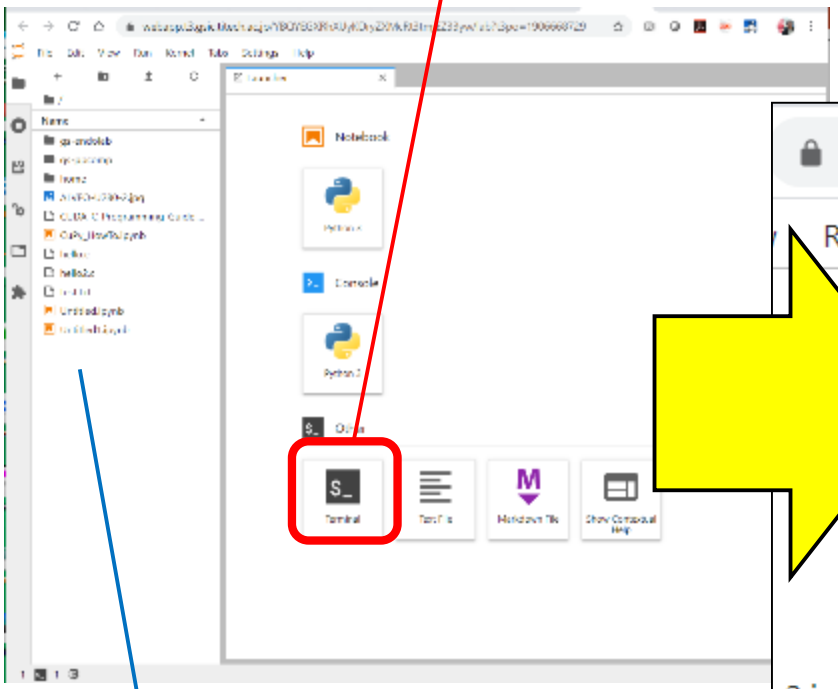
- If you see a blue row, push Display /表示 button

Web-only route

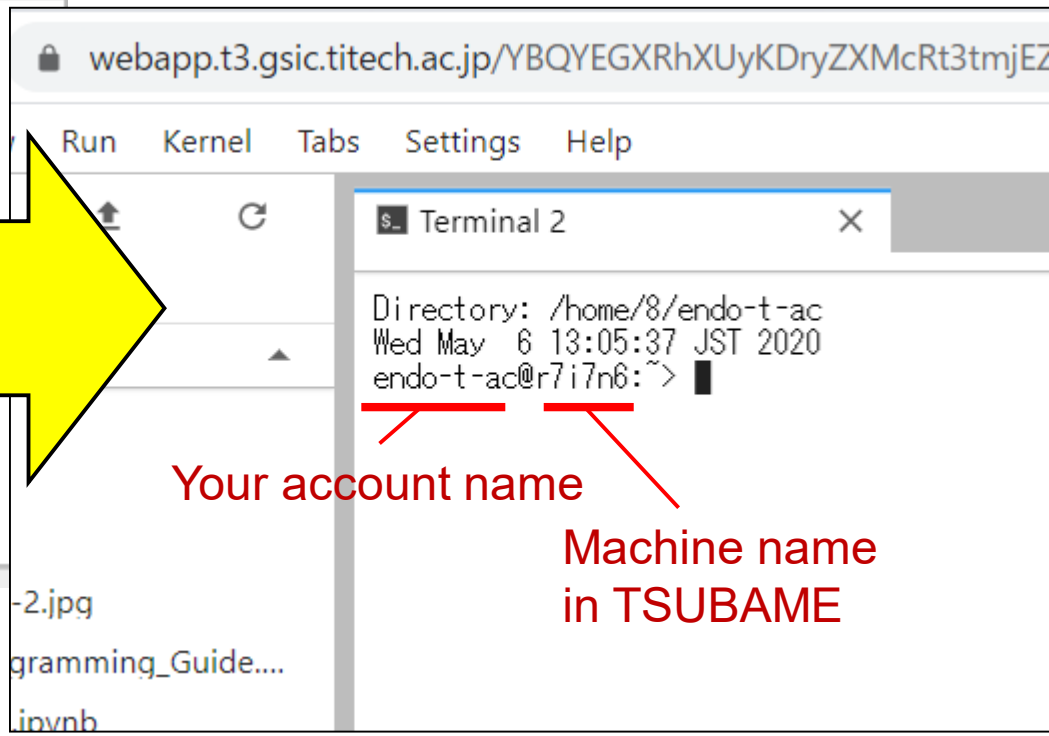


Start “Web Service” (3)

- You will see a “Jupyter” screen
- Click the Terminal icon



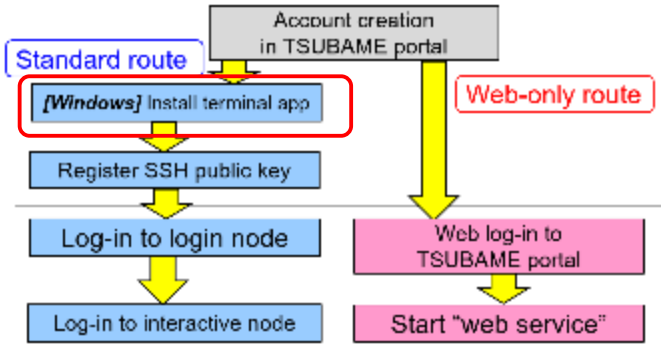
~/t3workspace/ directory
in TSUBAME



Your account name

Machine name
in TSUBAME

Standard route



[Optional]
Install a Terminal Application

(In standard route) Your PC must have a terminal application that supports SSH protocol

On Mac

- Start terminal / ターミナル app → use ssh command

On Windows…

- Recently, command prompt / コマンドプロンプト is ok
  - On Windows10 after 2018, ssh command has been installed

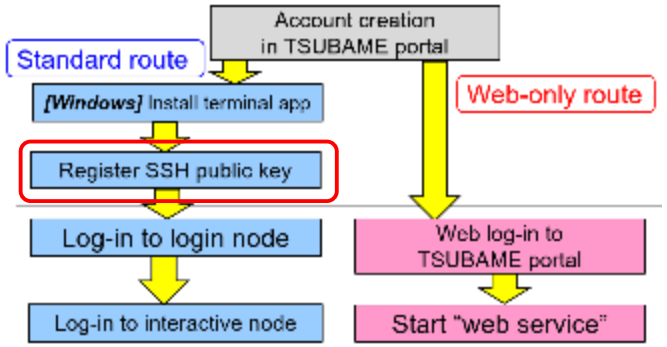
Also you can install other applications

- MobaXterm on Windows
- Putty on Windows/Mac
- iTerm on Mac …

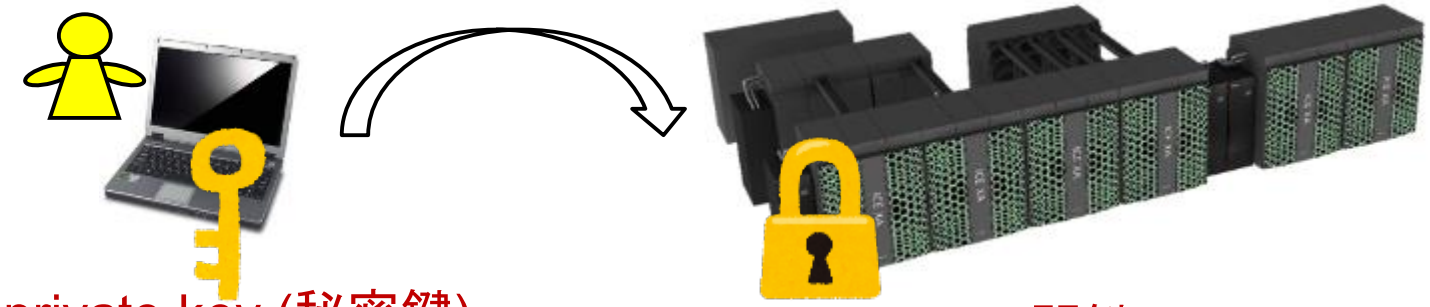
Or try google
“windows ssh client”
“mac ssh client”

Standard route

Register an SSH Public Key (1)



- To log-in to TSUBAME,
  - A password is NOT used
  - instead, public key (公開鍵) method is used



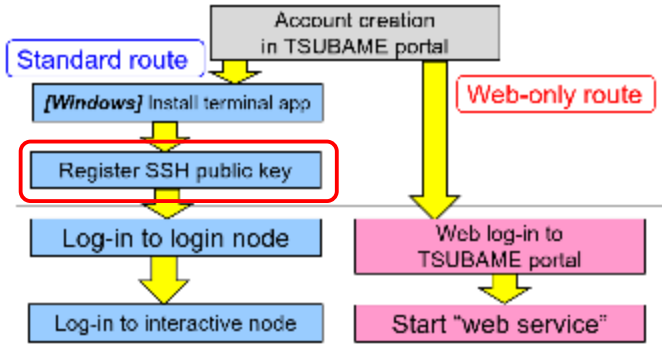
private key (秘密鍵)

public key (公開鍵)

Actually, they are two text files

Standard route

Register an SSH Public Key (2)



Please create key-pair on your PC by yourself

- On Mac terminal or Windows command prompt
  - Use ssh-keygen command
  - Setting "passphrase" is strongly recommend
  - → Two files are created

  In default,

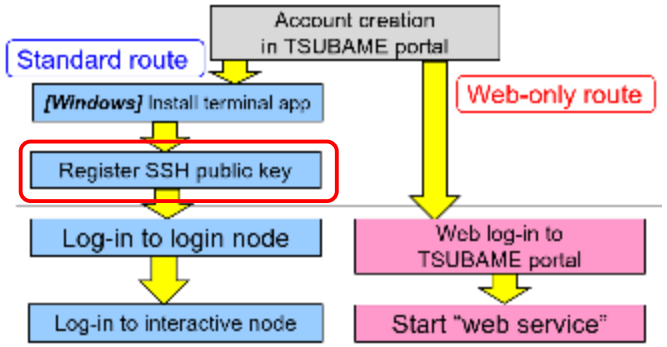
  private key → .ssh/id_rsa (Mac) .ssh¥id_rsa(Win)

  public key → .ssh/id_rsa.pub (Mac) .ssh¥id_rsa.pub (Win)

- If you are using another terminal application, try Google
  - Such as "mobaXterm public key"
  - If you are asked the key format, choose "OpenSSH" format, NOT "ssh.com format"

Standard route

Register an SSH Public Key (3)



Register your public key on TSUBAME portal

1. Click Register SSH public key / SSH公開鍵登録

2. Upload your public key

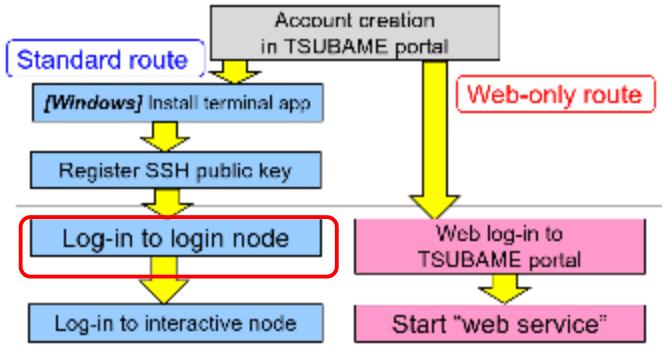
- Don't upload the private key!!

id_rsa　　id_rsa.pub

NOTE: The key is automatically added to ~/.ssh/authorized_keys on TSUBAME

Standard route

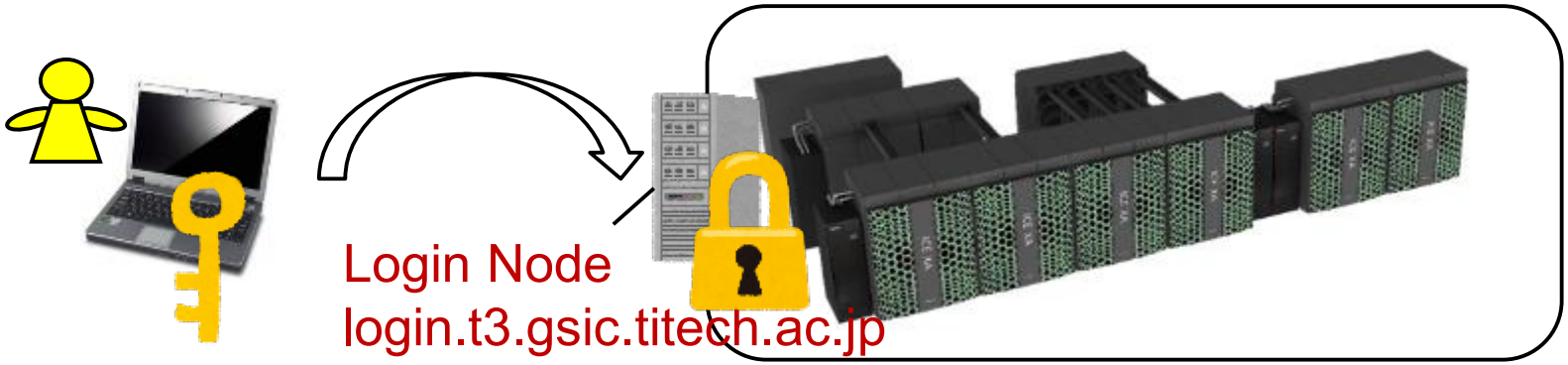
Log-in to TSUBAME Login Node



- On Mac terminal or Windows command prompt

  ssh -i *[private key] [account]*@login.t3.gsic.titech.ac.jp

  ex) ssh -i .ssh/id_rsa 20M12345@login.t3.gsic.titech.ac.jp



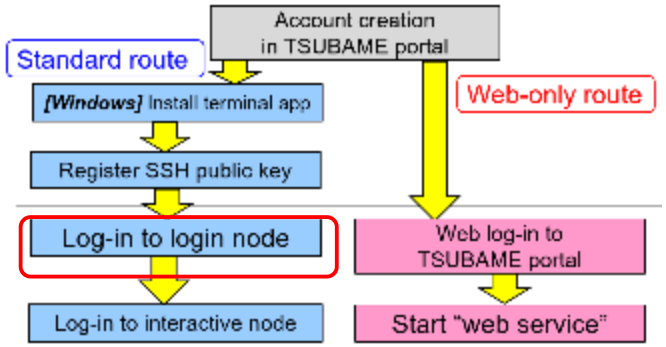

```
Last login: Wed May  6 13:30:17 2020 from 160.13.82.172
------------------------------------------------------------------------
Last modified: Apr 6 17:00:00 JST 2020

 *** Do not run programs with high load average such as ISV  ***
 *** on login0 and 1.                                        ***

    (The current TSUBAME 3.0 operational status)
    http://www.t3.gsic.titech.ac.jp/
------------------------------------------------------------------------
endo-t-ac@login1:~>
```

← If successful, you will see

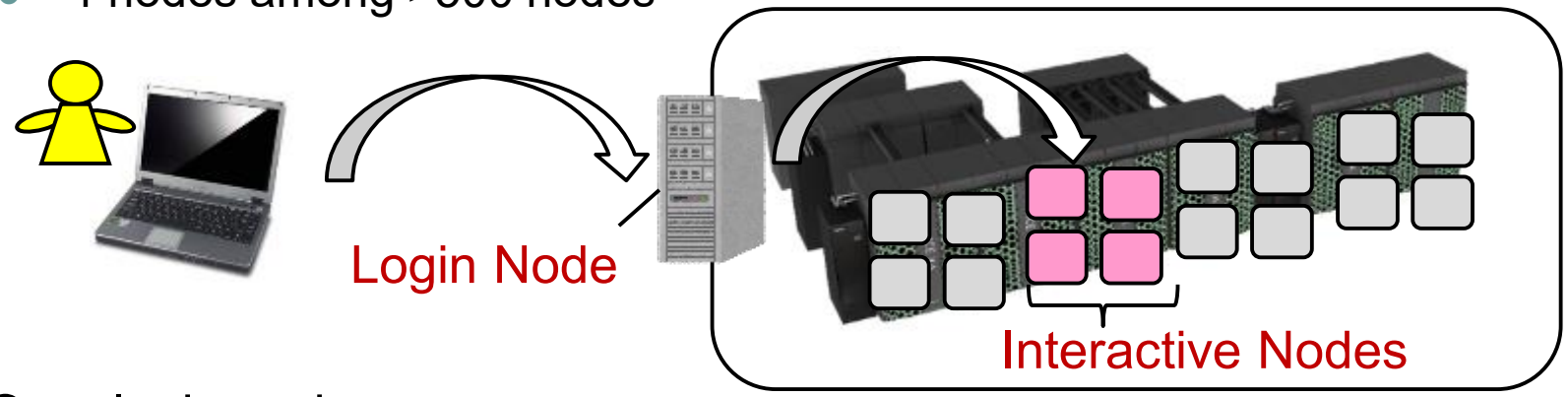
On other terminal applications, try Google

Standard route

Log-in to Interactive Node



- Log-in nodes are only entrance of TSUBAME and not powerful, no GPU on them
- In this lecture, we mainly use “interactive nodes”
  - 4 nodes among >500 nodes



- On a login node,

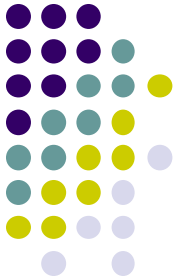
  iqrsh -l h_rt=2:00:00

NOTE: The command is changed in Apr 2021

If successful, you will see ➔

```
endo-t-ac@login1:~> iqrsh -l h_rt=2:00:00
Directory: /home/8/endo-t-ac
Mon Apr 11 07:14:16 JST 2022
endo-t-ac@r7i7n8:~>
```

Notes in Using TSUBAME



- Use your account only by yourself

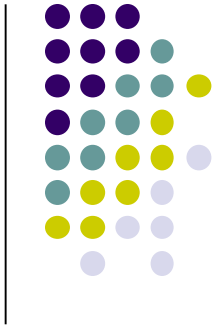
Standard route

- Don’t share private or public keys with other people
- Login nodes are shared by many TSUBAME users. Avoid running CPU/GPU heavy jobs there

[Interactive nodes] Web-only route Standard route

- Each user can use only one session
- 7 CPU cores + 1 GPU (= ¼ node) are assigned
- CPU cores, GPU may be shared by several users → you may suffer from slow down

Today's Homework (1)



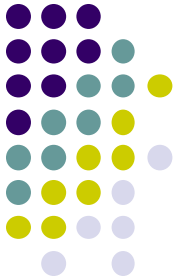
1. (If you are new to TSUBAME) please make your account on TSUBAME by 13:00, Apr 13

   (まだ作ったことがなければ) TSUBAMEアカウントを、4月13日13:00までに作成してください

2. You will receive an invitation e-mail to tga-ppcomp TSUBAME group. Please read it and accept the invitation.

   tga-ppcomp TSUBAMEグループへの招待e-mailが届くはずです。指示に従って招待を受けてください。

Today's Homework (2)



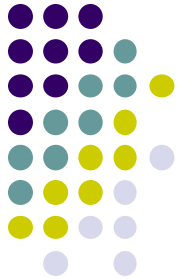
- If you cannot make an account by due time, please send an e-mail after account creation
- もし期限までにアカウントを作成できなければ、作成してから下記のようなe-mailを送ってください

To: ppcomp@el.gsic.titech.ac.jp
Subject: TSUBAME3 ppcomp account

Department name:
School year:
Name:
TSUBAME account name:

Then we will invite you to the TSUBAME group
その後、TSUBAMEグループへの招待を送ります

Contact/ Information on the Course



- About this course
  - e-mail ppcomp@el.gsic.titech.ac.jp or endo@is.titech.ac.jp
- T2SCHOLA
  - https://t2schola.titech.ac.jp/
- Slides
  - https://github.com/toshioendo/ppcomp

- TSUBAME official site
  - https://www.t3.gsic.titech.ac.jp manuals are here
- Tokyo Tech portal
  - https://portal.titech.ac.jp/

➔ Log-in and then click “TSUBAME portal”