

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

Part 0: Introduction

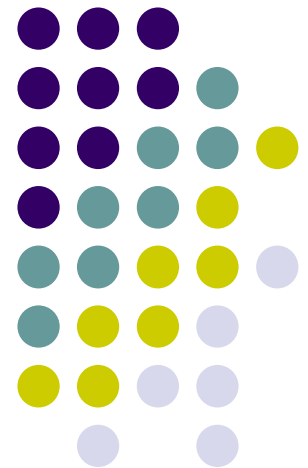
No 1: About Supercomputer

Apr 8, 2024

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
From this year, we use a brand-new TSUBAME4.0 !





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 2024, 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
 - 〆切は、各パート終了の約1.5週間後
 - (それ以上のレポートを提出してもよい)



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 TSUBAME4 and Internet connection
 - Such as “SciTokyo” or “TokyoTech” wifi



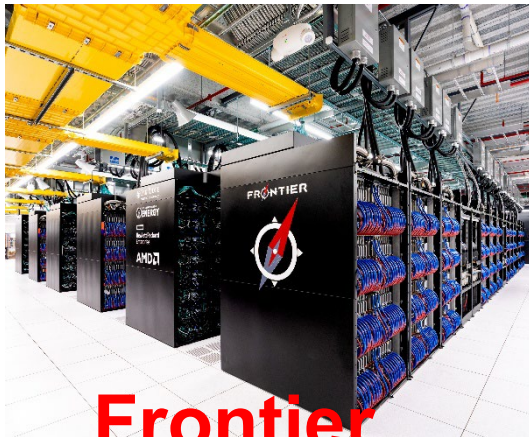
Supercomputers



Fugaku
World No. 4



Tokyo Tech
TSUBAME4.0



Frontier
World No.1



Taifulight

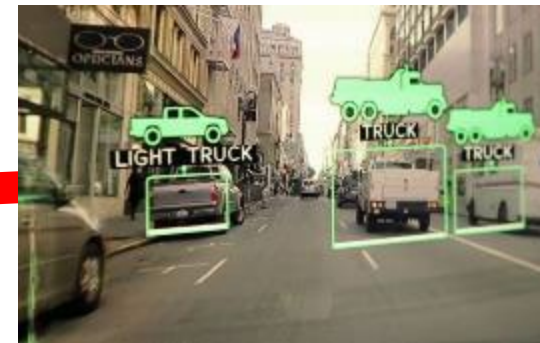
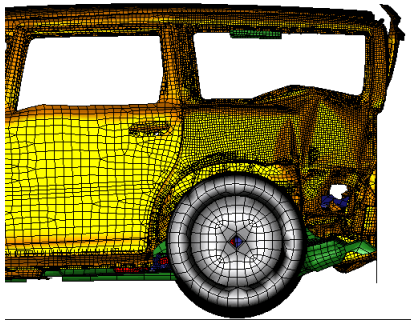
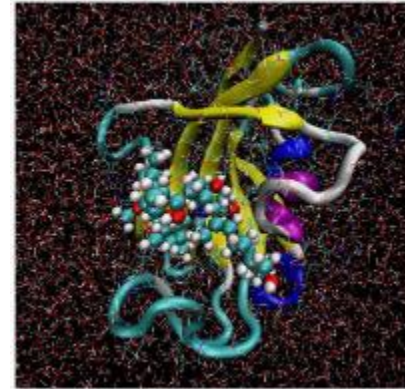
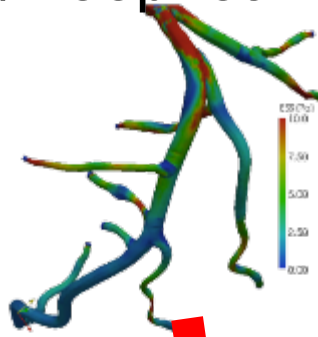


ABCI
AIST, Japan

What are Supercomputers (SCs) used for?



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



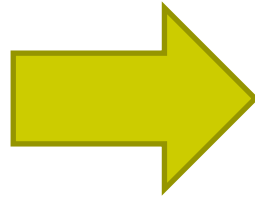
Difference with “Normal” Computers



- 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



~130,000x!

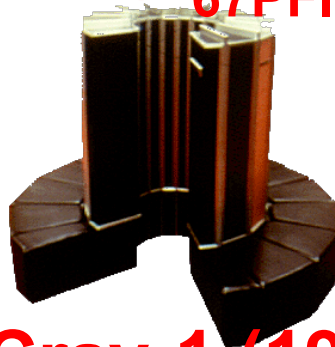


~500GFlops (5×10^{11} times per sec)

67PFlops (6.7×10^{16} times per sec)



PC in 1980



Cray-1 (1976)

160MFlops



Modern
Cell Phone

Why are Speed & Size Important?

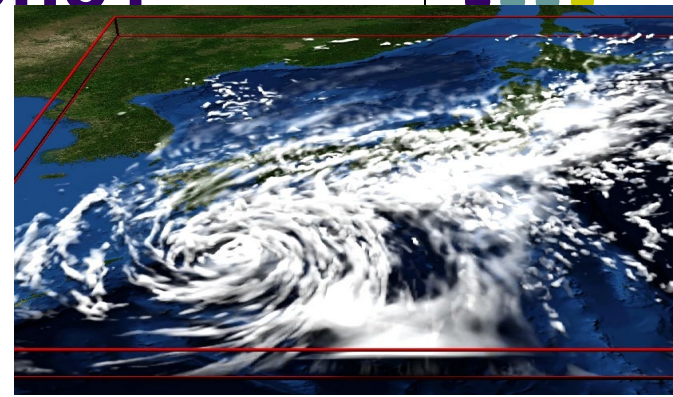


- For simulation & big-data analysis, **large number of computations** should be done speedily
 - ⇒ Want to obtain forecast of tomorrow weather by tomorrow (of course!)
 - ⇒ Want to develop and sell new medicine (than competitors)
- For simulation & big-data analysis, storing **large scale data** is 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

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

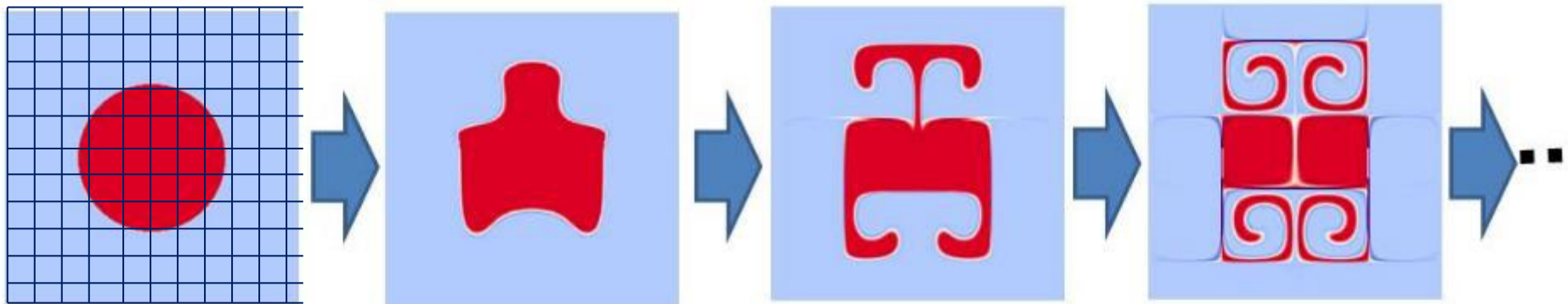


But no analytical solution for them, generally

By Takayuki Aoki

⇒ Instead, space and time are **discretized**

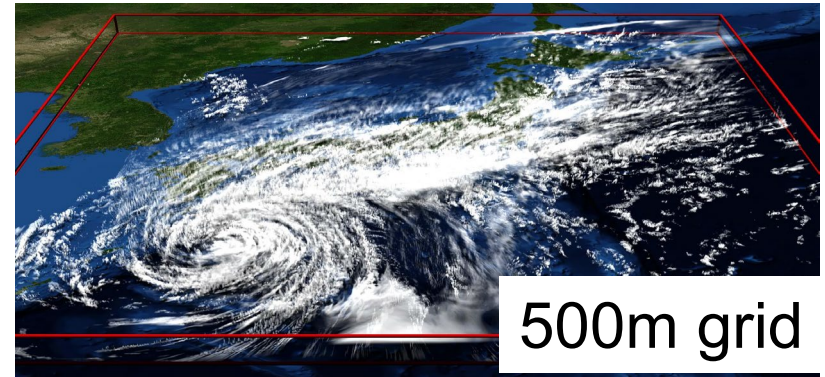
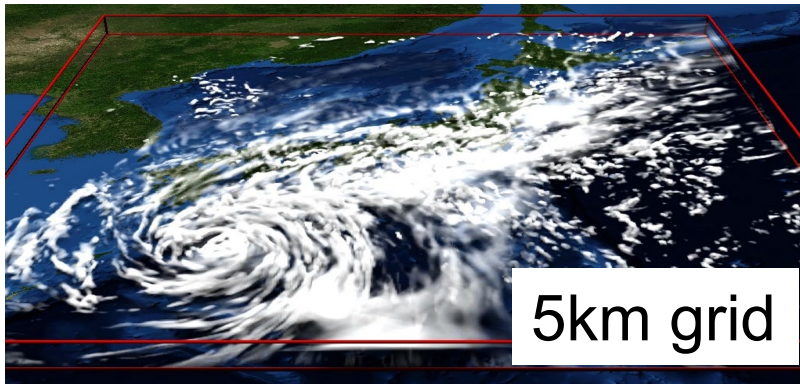
- The space is divided into small grids, expressed as an array
⇒ Each array element should be computed
- The time is divided into time steps
⇒ After a time step is computed, we go to next step, and so on



Why is Speed Important?



- Since we have to compute all points for every time step, computational complexity is
 - $O(\text{x-size} \times \text{y-size} \times \text{z-size} \times \text{time-steps})$



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

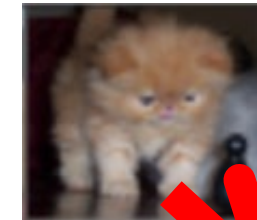
Inputs are lots of images with correct answers



AI, actually
numeric data

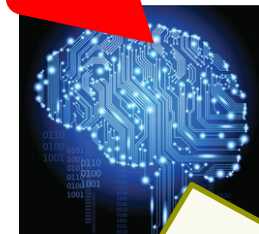
Inference

Done on cars/phones



What's
this?

Copy
data of AI

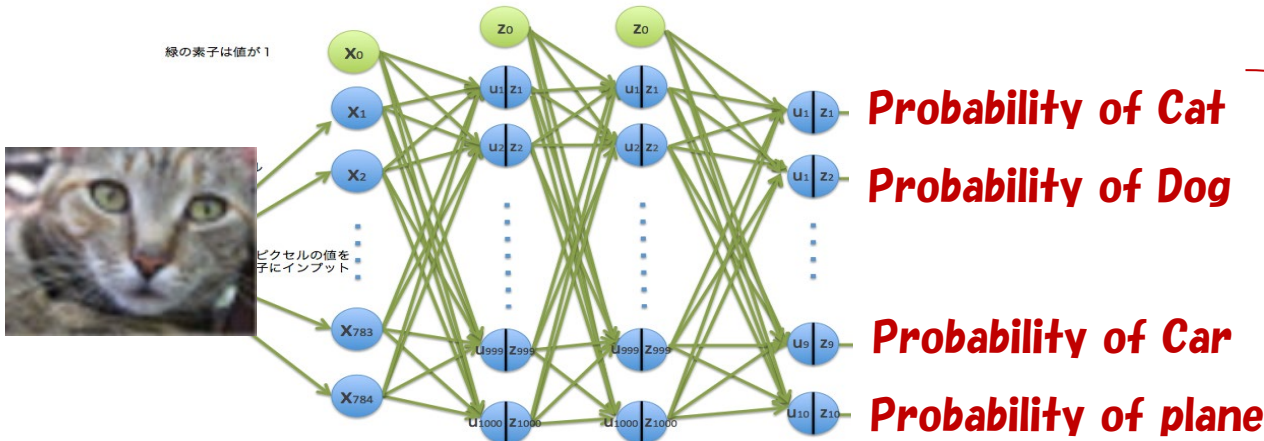


A cat (75%)
A dog (24%?)

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 (\neq correct answer)
 - By using difference of current answer and correct answer, it updates its parameter vectors



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

Vertices are neurons
Arrows are connections,
which have numeric data

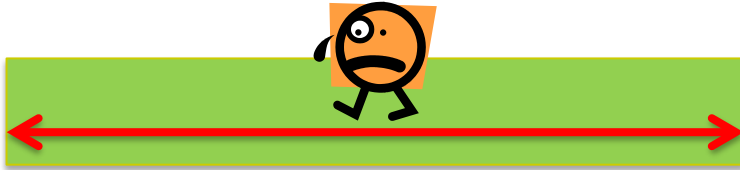
Complexity of (Number of images \times
Sizes of parameter vectors \times
Number of repetitions) !!



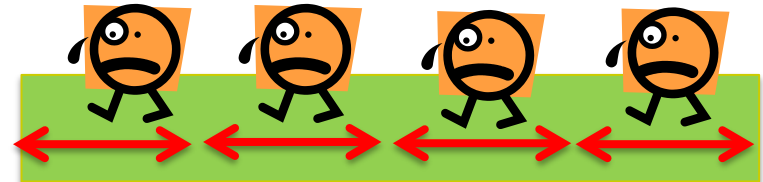
Why are SCs Fast?

- 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



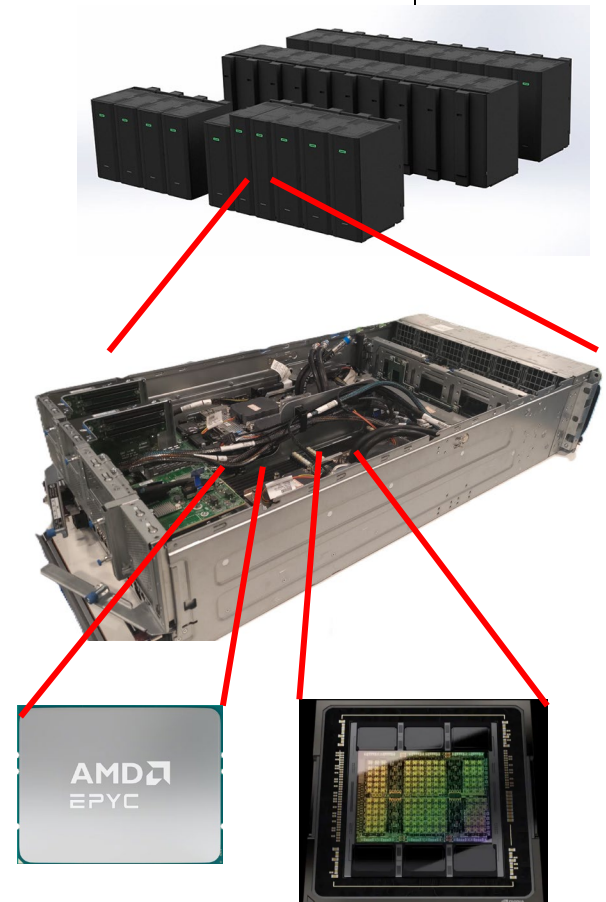
Multiple workers are working together → **fast!**



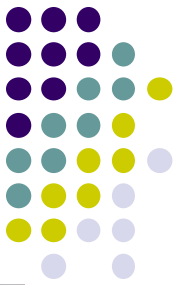
SC Structure is Hierarchical



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



Structure of TSUBAME4 and Fugaku



TSUBAME4



Fugaku

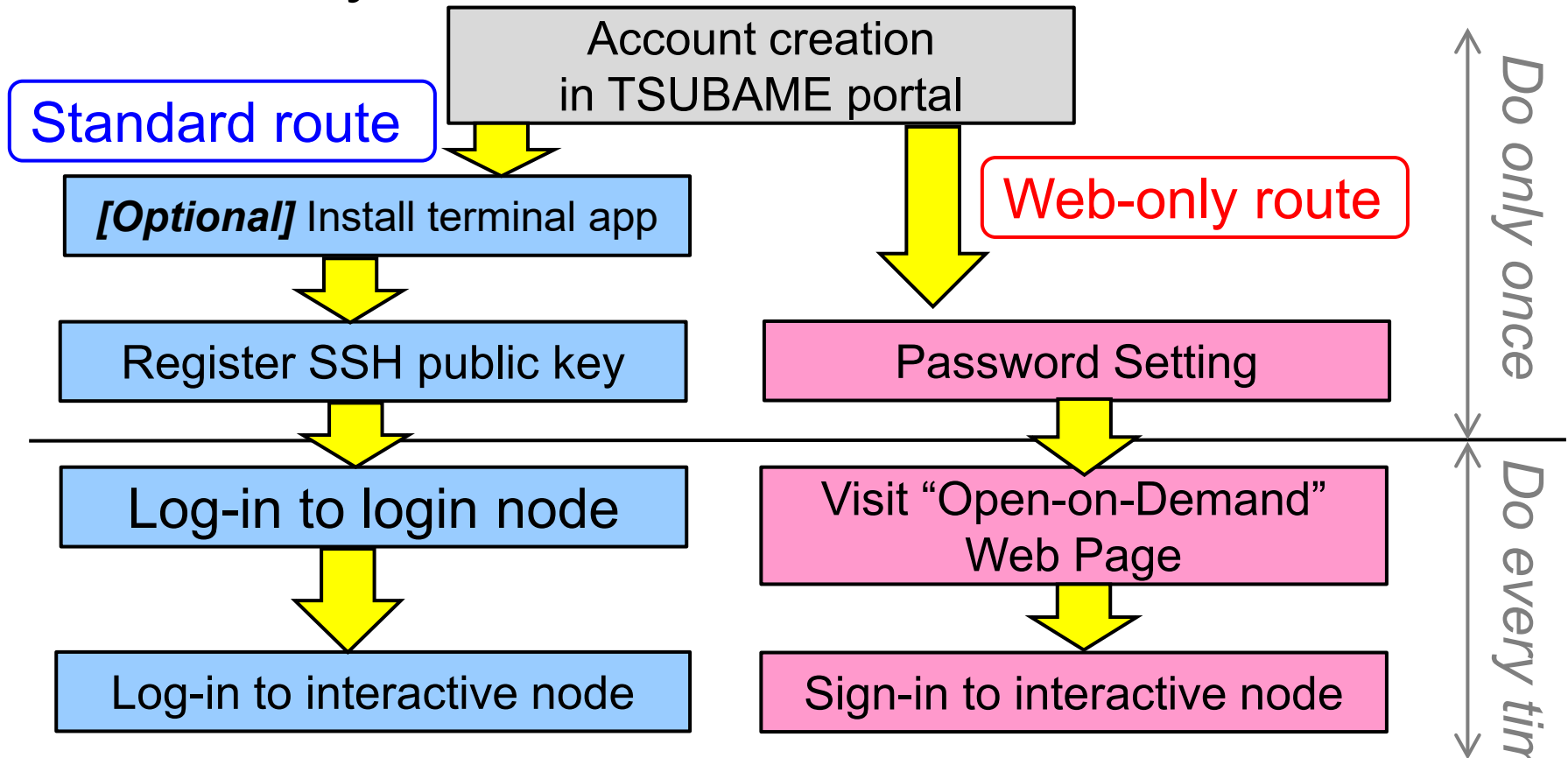
System	240 nodes	66.8PFlops	160,000 nodes	540PFlops
Node	2 CPUs + 4 GPUs	2 x 5.45TFlops + 4 x 66.9TFlops = 278.5TFlops	1 CPU	3.38TFlops
Processor	CPU: 96 cores GPU: 132 SMXs	CPU: 5.45TFlops GPU: 66.9TFlops	48 cores	3.38TFlops
Core	CPU core: 3.55GHz x 16 = 56.8GFlops GPU SMX: 507 Gflops (Tensor core)		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.t4.gsic.titech.ac.jp>
 - Manuals are here

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

➔ Log-in and then click
“TSUBAME portal”

- Open-on-Demand page
 - <https://ood.t4.gsic.titech.ac.jp/>
 - For **Web-only** route



TSUBAME計算サービス
TSUBAME4.0

メニュー

トップ

マニュアル・FAQ

お問い合わせ

TSUBAME4ポータル

▼ アカウント取得方法

アクセスコード所持者の課金優遇

外部資金による学外者の利用

科目等履修生・特別聴講学生

東工大以外からの利用(共同利用)

▼ 利用規約・利用料など

規則・利用細則

利用料の概略

定額制

(学内向け) 課金請求時期

講義用の無償利用

JST創発的研究支援事業に基づく割引

謝辞記載のお願い

▼ 運用情報

お知らせ一覧

各種制限値

現在のTSUBAME4の運用状況 / Current TSUBAME4 Operational Status

2024-04-01 15:00 テスト運用期間中(4/22まで) / Test Operation until April 22th.

重要なお知らせ

2024-03-07

TSUBAME4.0 テスト運用期間について

2023-11-09

2024年度のTSUBAMEの運用予定について

お知らせ

2024-04-01

TSUBAME4.0の運用を開始しました

ログインノード

2024-04-04 9:00 利用できます。

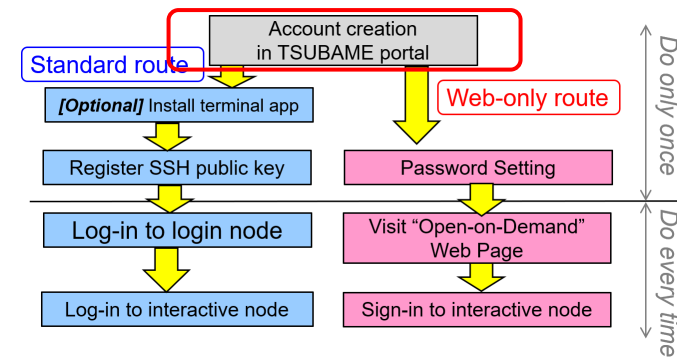


The screenshot shows the Tokyo Tech Portal homepage. At the top, there's a navigation bar with 'Tokyo Tech Portal' and '東京工業大学' (The University of Tokyo). Below the navigation bar, there's a section for '東京工業大学情報基盤利用承諾' (The University of Tokyo Information Infrastructure Usage Consent). The main content area is divided into several sections, including '重要なお知らせ' (Important Notice) and 'お知らせ' (Notice). The '重要なお知らせ' section contains a notice about the TSUBAME4.0 test operation period, dated 2024-04-01. The 'お知らせ' section contains a notice about the start of TSUBAME4.0 operation, dated 2024-04-01. At the bottom of the page, there's a section for 'TSUBAME4.0ポータル' (TSUBAME4.0 Portal) with a link to 'ログイン' (Login).

Standard route

Web-only route

Account Creation in TSUBAME Portal



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



Click **“TSUBAME4 portal”**

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

➔ You will get an account.

- An account name looks like:
ux02345 (not the student ID)

- For details, visit <https://www.t4.gsic.titech.ac.jp/> and click “Getting Accounts” / “アカウント取得方法”

Web-only route

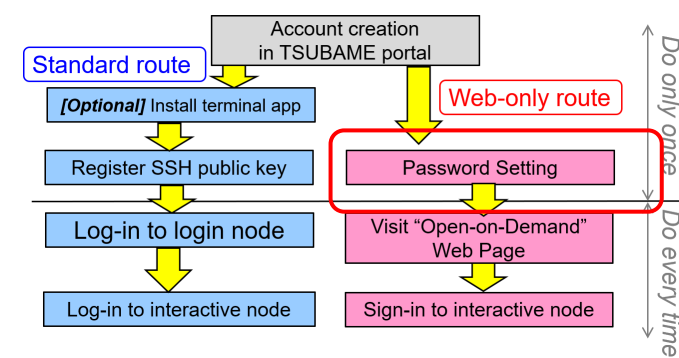
Password Setting

Go to “TSUBAME4 portal” again

- Visit “Tokyo Tech portal”
- Log-in
- Click “TSUBAME4 portal”

Your account name

Set a new password
→ used in next page



TSUBAMEポータルページ

Logout Language

所属グループ:

[TSUBAME利用状況]

ジョブ一覧

予約ノード一覧

[利用者情報]

利用者情報表示

SSH公開鍵登録

パスワード設定

利用規約一覧

[課金管理]

支払コード管理

予算コード承認依頼 0

[グループ]

所属グループ管理

TSUBAMEパスワードの設定

新パスワード

新パスワード再入力

登録

TSUBAMEポータルページ

Logout Language

利用者権限:

アカウント: ua02023

現在の状態: 利用中

所属グループ:

[TSUBAME利用状況]

ジョブ一覧

予約ノード一覧

[利用者情報]

利用者情報表示

SSH公開鍵登録

パスワード設定

利用規約一覧

[課金管理]

支払コード管理

予算コード承認依頼 0

[グループ]

所属グループ管理

お知らせ

[TSUBAME4.0はテスト運用期間中です](#)

テスト運用期間中につき、無予告でサービス停止・縮退等が発生する可能性があります。

[お問い合わせフォーム](#)

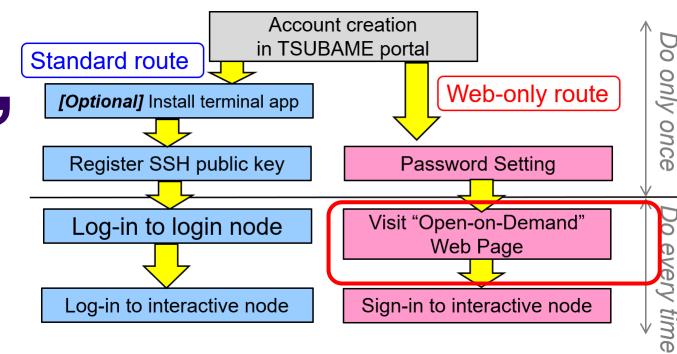
不具合報告・改修予定・Q&Aなど

- 2024-04-05 14:30 アプリの利用権購入]権限を付与してもアプリケーション利用の設定で「権限がないため表示できません。」と表示されます。 → 現在確認中です
- 2024-04-05 9:40 TSUBAME3.0のグループ情報を4.0に引き継ぎますか? → 出来ません
- 2024-04-04 17:15 諸事情によりTSUBAMEポータル内のお問い合わせフォームを停止しました。お問い合わせは <https://www.t4.gsic.titech.ac.jp/contact-t4> からお願いします。
- 2024-04-04 17:00 支払いコードの承認はどのくらいかかる? → 申請が集中している。不備がなければ遅くとも2営業日以内には
- 2024-04-04 利用分野の選択肢が少ない → 増やしました
- 2024-04-04 ログイン名が覚えにくいに変更できるか? → 出来ません
- 2024-04-04 一度に100口までしか購入できない → 上限値を大きくします

Web-only route

Visit “Open-on-Demand” Web Page

- Visit <https://ood.t4.gsic.titech.ac.jp/>



ONDEMAND LOGIN PAGE

Sign in to your account

Username

Password

Sign In

Type your account name

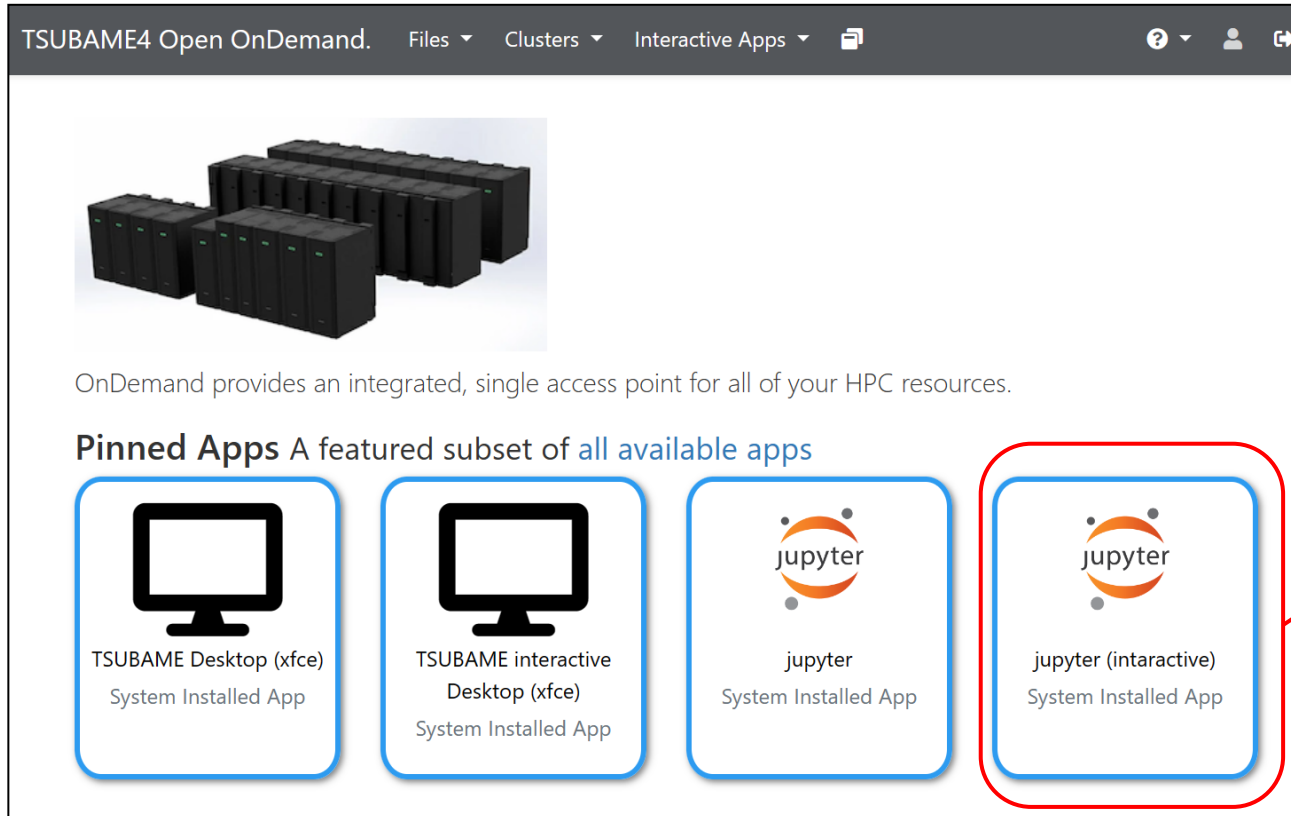
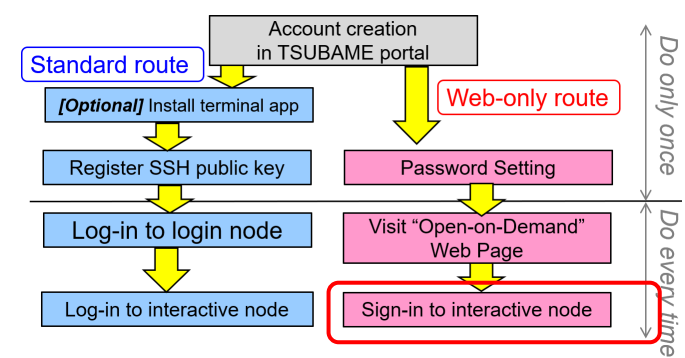
Type your password

Two-factor authentication is used:

- You will receive an e-mail with “Access code”
➔ Type it

Web-only route

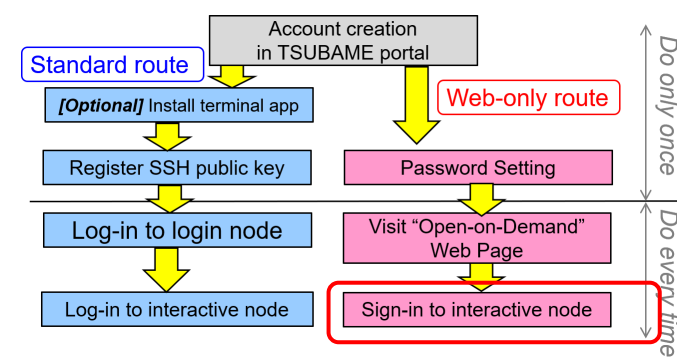
Start “Web Service” (1)



Select
jupyter (interactive)

Web-only route

Start “Web Service” (2)



TSUBAME4 Open OnDemand. Files ▾ Clusters ▾ Interactive Apps ▾

Home / My Interactive Sessions / jupyter (interactive)

Interactive Apps

Desktops

- TSUBAME Desktop (xfce)
- TSUBAME interactive Desktop (xfce)

Servers

- jupyter
- jupyter (interactive)**

jupyter (interactive)

This app will launch a Jupyter Notebook server on one or more nodes.

TSUBAME group

Free

Maximum run time(hh:mm:ss)

2:00:00

Specify the job execution time in the format hh:mm:ss.

Launch

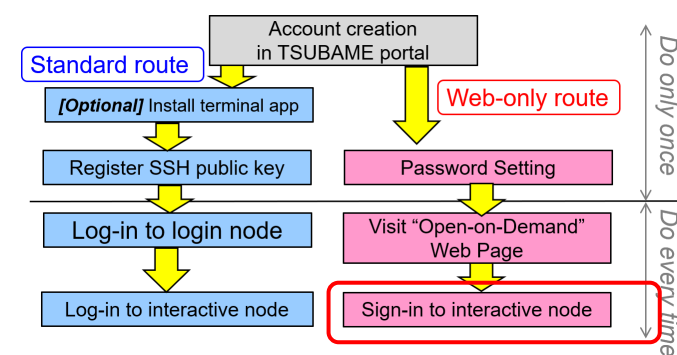
* The jupyter (interactive) session data for this session can be accessed under the [data root directory](#).

Select “free”

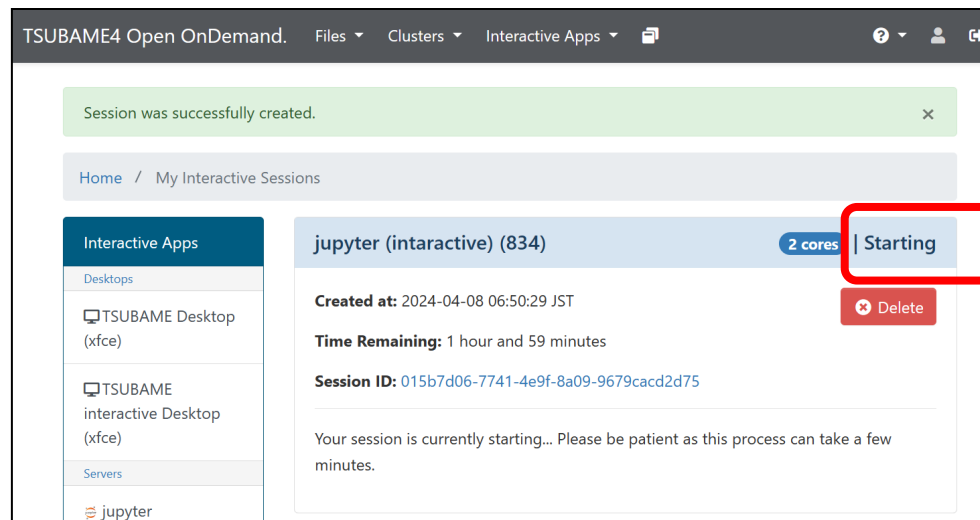
Type maximum time, such as “2:00:00”

Web-only route

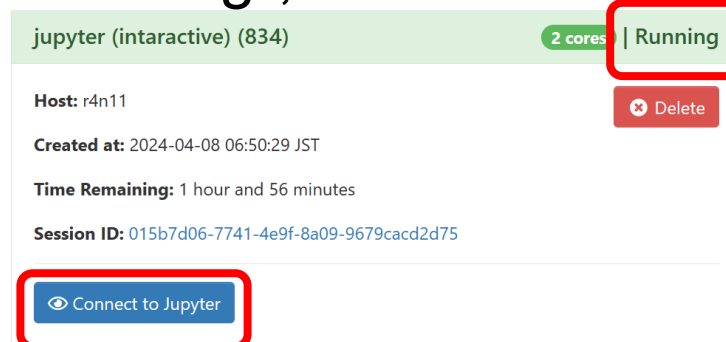
Start “Web Service” (3)



- Please wait until your session becomes “Running”



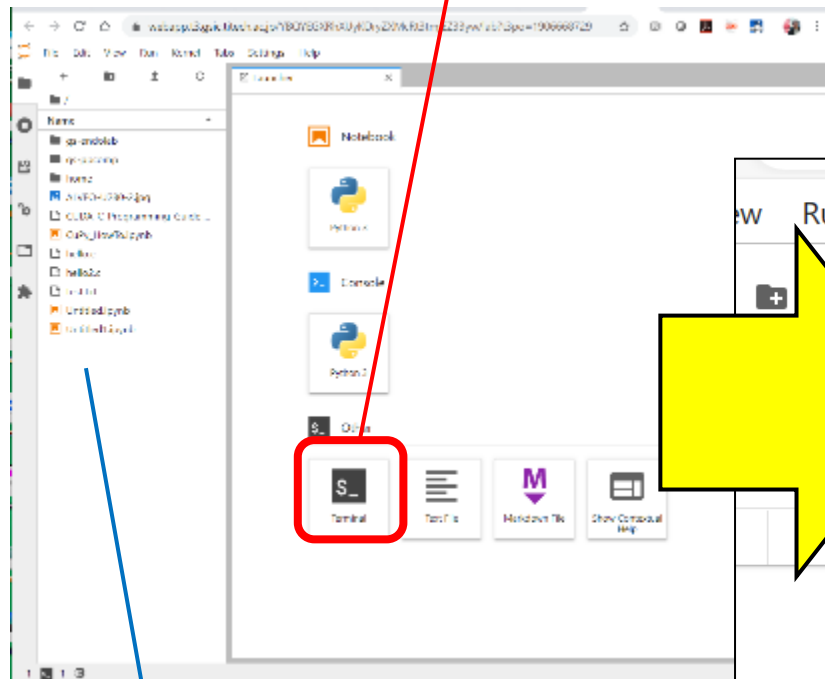
- If you see “Running”, click “Connect to Jupyter”



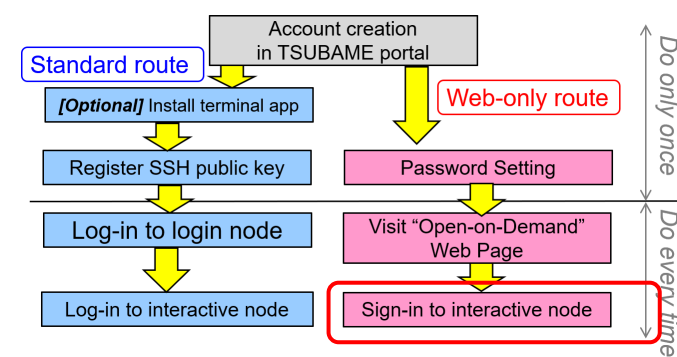
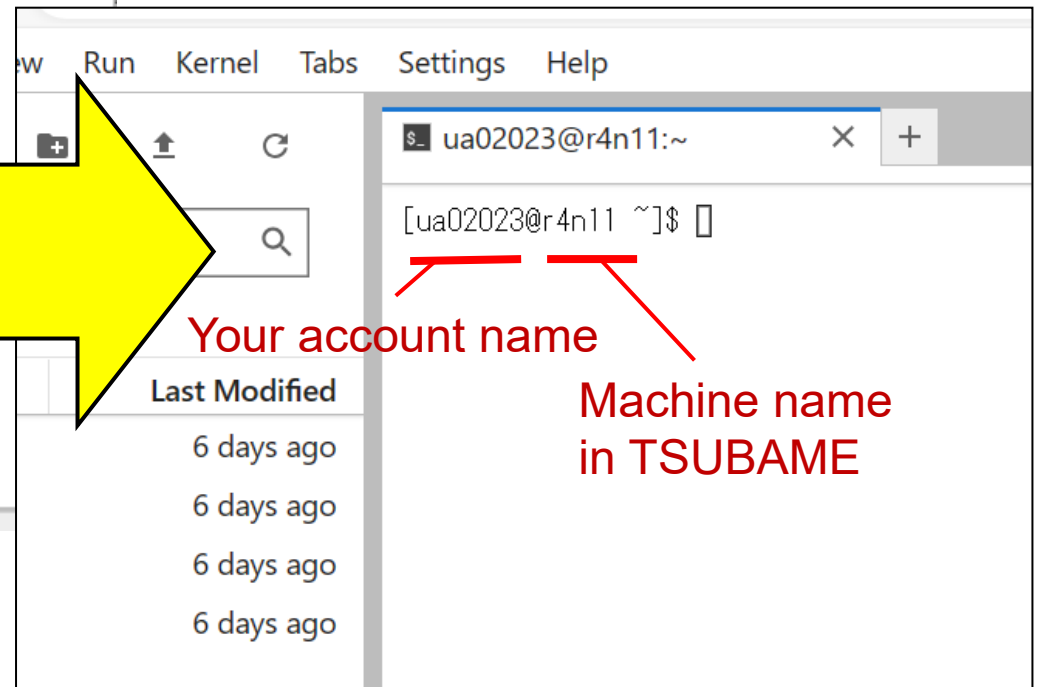
Web-only route

Start “Web Service” (4)

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



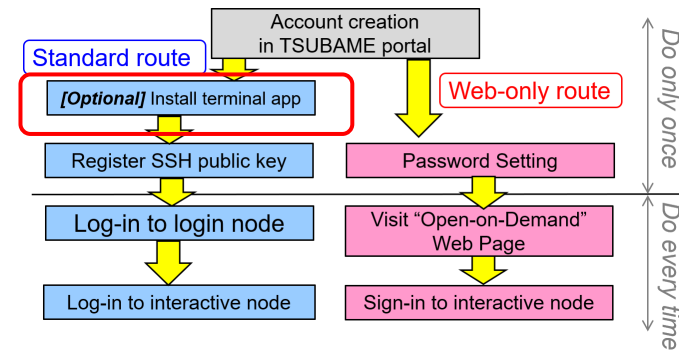
~/ (your home) directory
in TSUBAME



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

Also you can install other applications

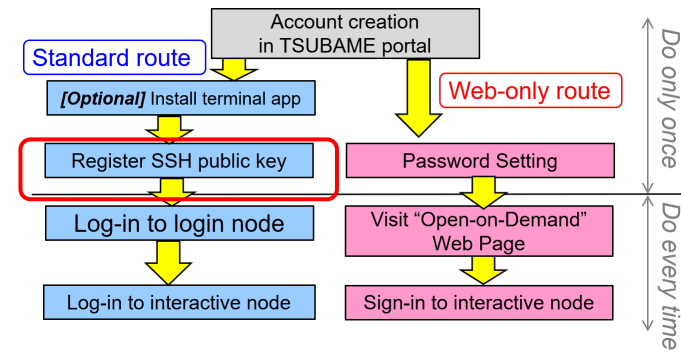
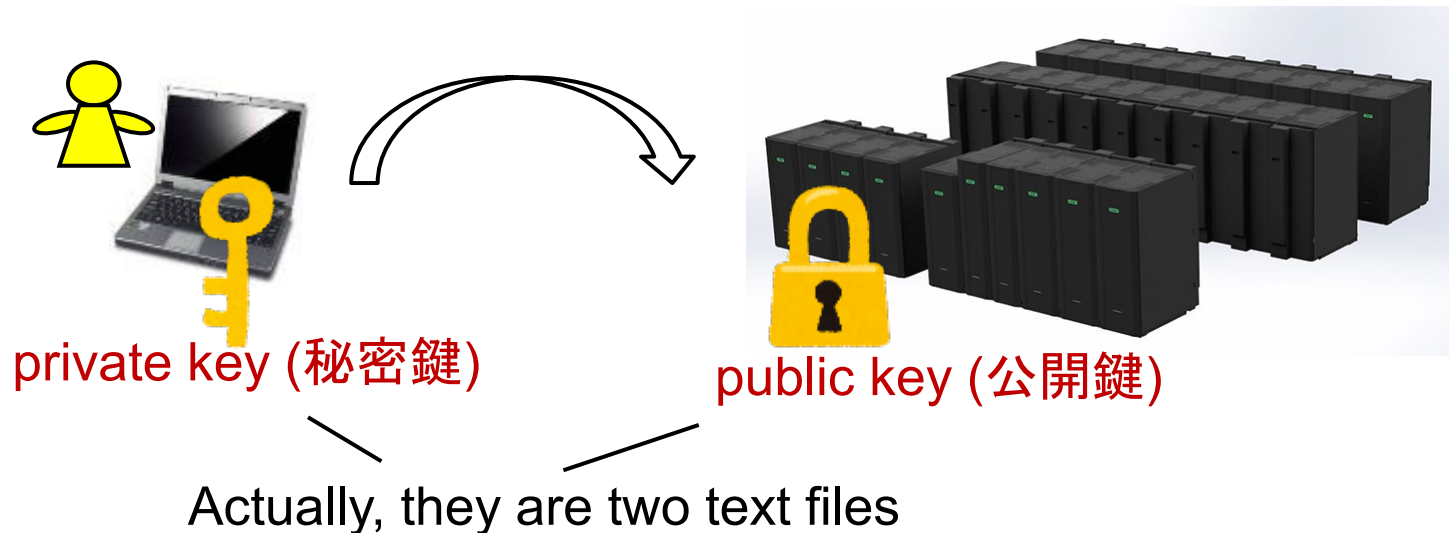
- MobaXterm on Windows**
- Putty on Windows/Mac**
- iTerm on Mac ...**

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

Standard route

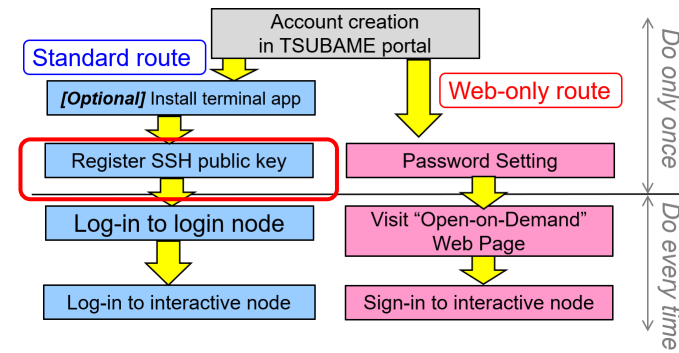
Register an SSH Public Key (1)

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



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

For example,

private key → `.ssh/id_ecdsa` (Mac) `.ssh\id_ecdsa` (Win)
public key → `.ssh/id_ecdsa.pub` (Mac) `.ssh\id_ecdsa.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

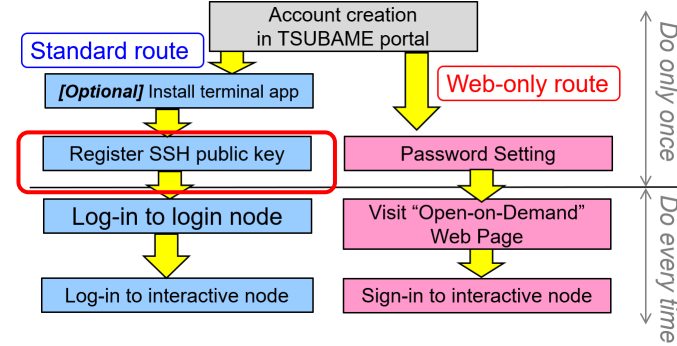
1. Click **Register SSH public key / SSH公開鍵登録**
2. Upload your **public key**
 - Don't upload the private key!!



id_ecdsa



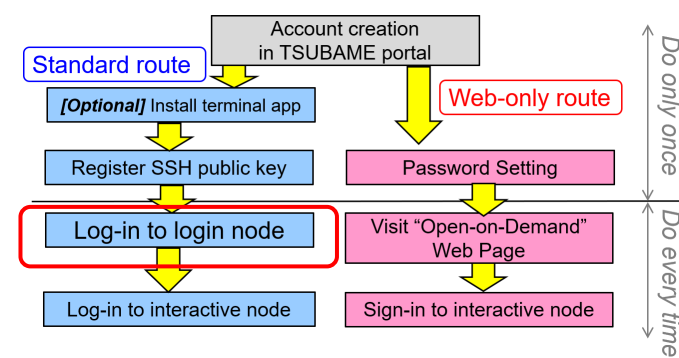
id_ecdsa.pub



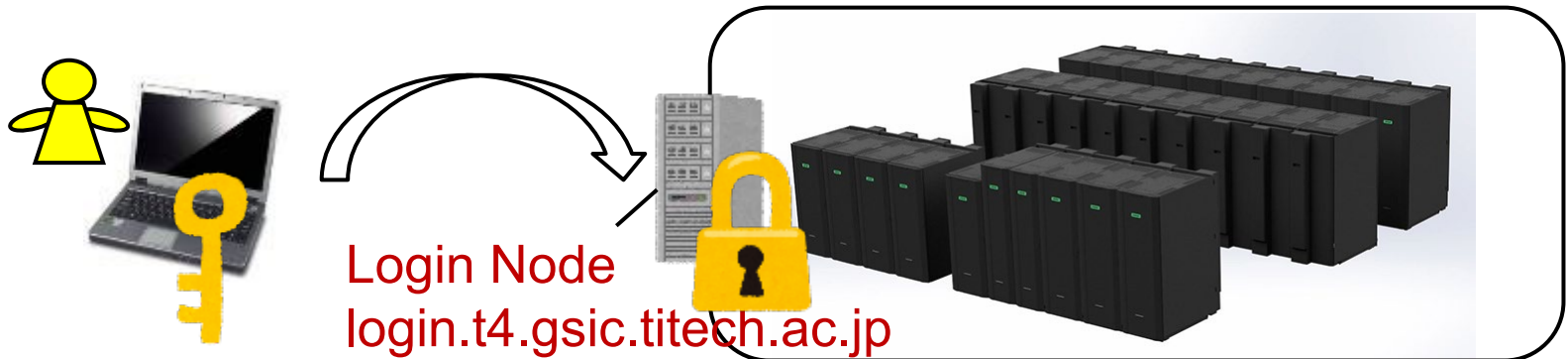
NOTE: The key is automatically added to
~/.ssh/authorized_keys on TSUBAME

Standard route

Log-in to TSUBAME Login Node



- On Mac terminal or Windows command prompt
`ssh -i [private key] [account]@login.t4.gsic.titech.ac.jp`
ex) `ssh -i .ssh/id_rsa ux02345@login.t4.gsic.titech.ac.jp`



```
Last modified: Mon Apr 1 10:00:09 JST 2024
```

```
(The current TSUBAME 4.0 operational status)
https://www.t4.gsic.titech.ac.jp/ X(Twitter):@Titech_TSUBAME
```

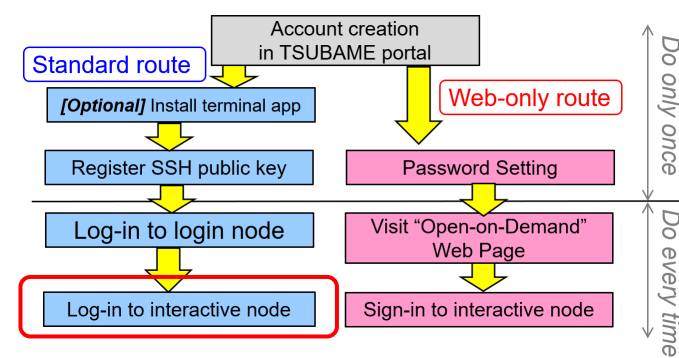
```
Last login: Wed Apr 3 18:41:29 2024 from 131.112.132.225
[ua02023@login1 ~]$
```

← If successful,
you will see

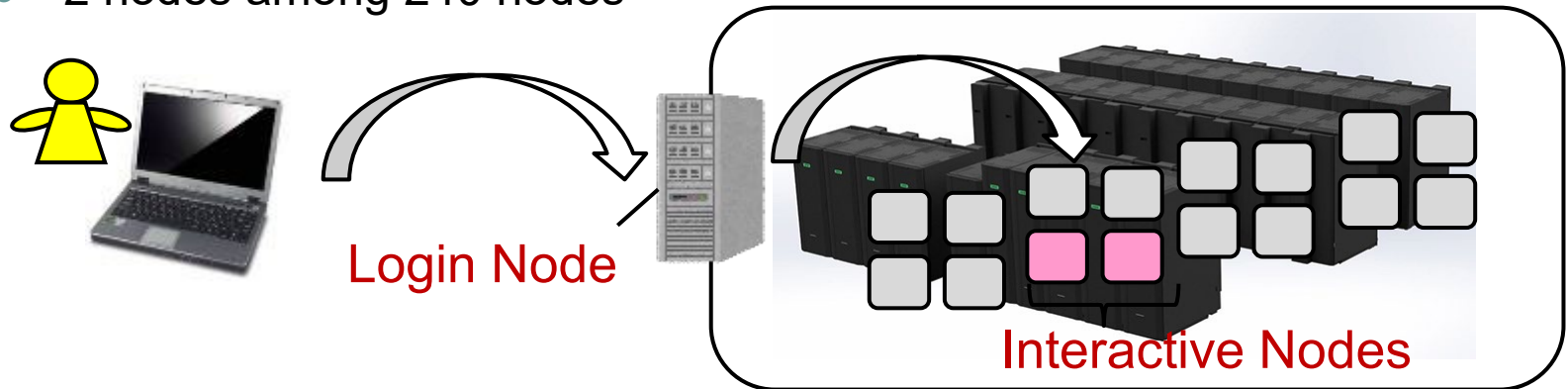
On other terminal
applications, try Google

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”
 - 2 nodes among 240 nodes



- On a login node,
`iqrsh -l h_rt=2:00:00`

If successful, you
will see →

```
[ua02023@login1 ~]$ iqrsh -l h_rt=2:00:00
[ua02023@r4n11 ~]$
```



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
- (24 CPU cores + 0.5 GPU) (= 1/8 node) are assigned
- CPU cores, GPU may be shared by several users → you may suffer from slow down



Today's Homework (1)

Please do 1&2 **by 17:00, Aug 10**. We accept the mail later.

1と2を**4/10 17:00まで**に行ってください。それより後でも受け付けます

1. (If you are new to TSUBAME) please make your account on TSUBAME
(まだ作ったことがなければ) TSUBAMEアカウントを作成してください
2. Please send an e-mail after account creation
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: **TSUBAME4 ppcomp account**

Department name:

Student ID:

Name:

TSUBAME account name:



Today's Homework (2)

3. 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。

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.t4.gsic.titech.ac.jp> manuals are here
 - Tokyo Tech portal
 - <https://portal.titech.ac.jp/>
- ➔ Log-in and then click “TSUBAME4 portal”