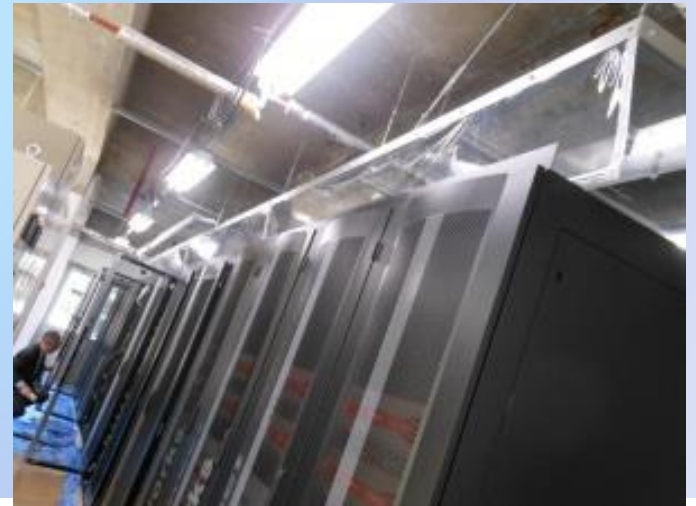


遺伝研スーパーコンピュータ (スパコン) の概要と利用申請

利用申請から約2週間後に
アカウントが郵送される





遺伝研スパコンの概要



遺伝研スーパーコンピュータには無料枠有り

■NGS解析にはスパコンが必要

NGS配列は規模大な為に、
スパコンでの解析が
デフォルト



スパコン利用には、課金
や「SSH公開鍵認証接
続」、「キューイングシ
ステムによるジョブ実
行」などの技術的ハード
ルが高い

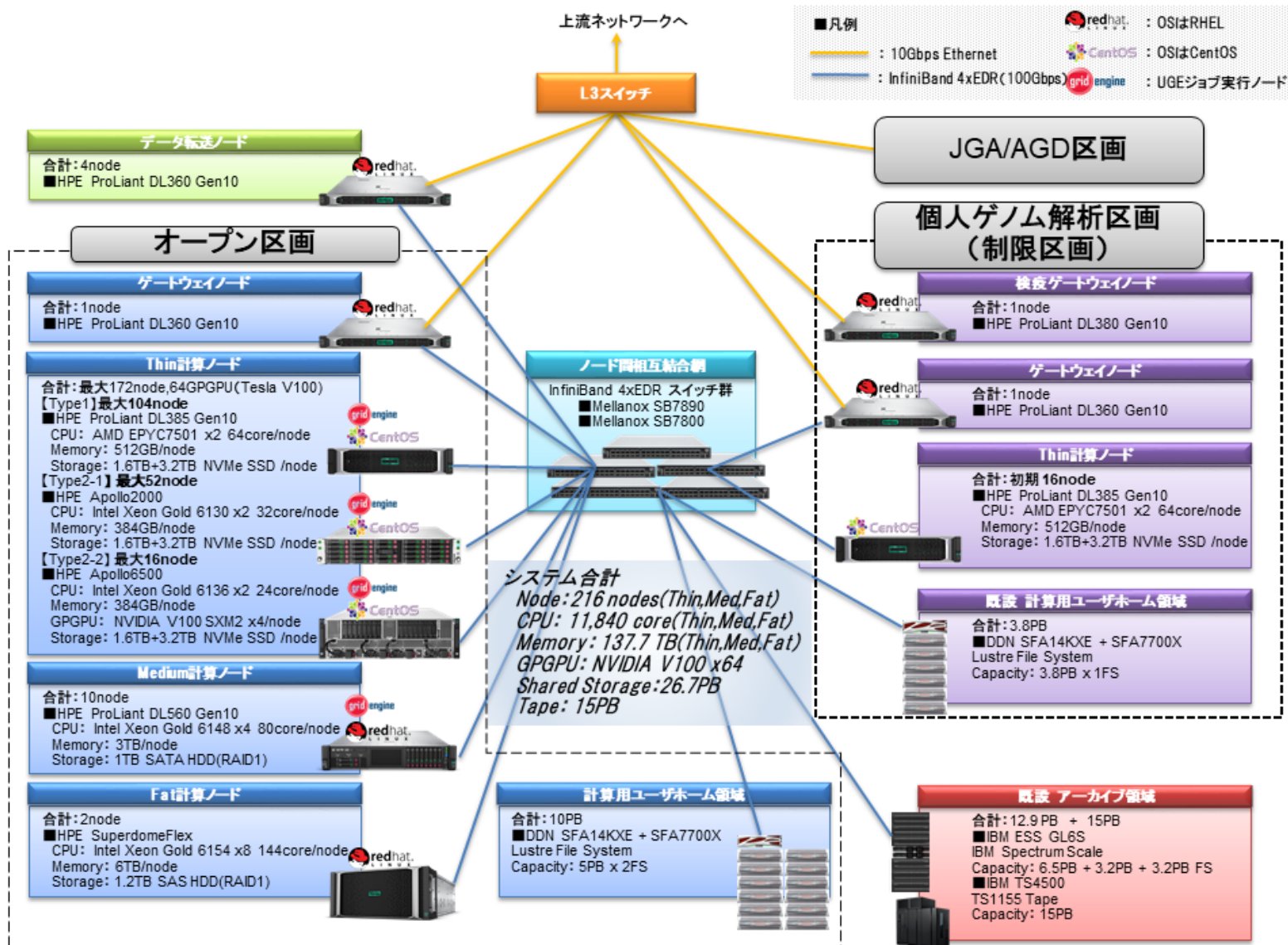


無課金の遺伝研スパコン
で初歩的技術が学べる

- * SSH接続（公開鍵設置）
- * ジョブ実行

一般解析環境	個人ゲノム解析環境 (個人ゲノム解析環境の利用はすべて課金 サービスとなります)
<ol style="list-style-type: none">一般解析環境の使い方（概要）一般解析環境へのログインの方法<ol style="list-style-type: none">Linux / Mac OSXの場合Windowsの場合インタラクティブノードを用いた対 話的作業の方法バッチ処理の実行方法 : Univa Grid Engine (UGE)の使い方ファイル転送の方法一般解析環境における課金サービス の利用方法<ol style="list-style-type: none">ストレージの大規模利用計算機の占有利用 (UGE Advanced Reservation)	<ol style="list-style-type: none">個人ゲノム解析環境の使い方（概 要）個人ゲノム解析環境へのログイン の方法<ol style="list-style-type: none">Mac OSXの場合Linuxの場合Windowsの場合対話的作業の方法バッチ処理の実行方法<ol style="list-style-type: none">Univa Grid Engine (UGE)の使 い方slurmの使い方ファイル転送の方法
<p>* 30TBまで無料 * 30TBを超えた場合10TBあ たり年間20万円</p> <p><u>年度末まで利用可能</u></p>	

遺伝研スパコンのシステム構成





遺伝研スパコン内に、主要な塩基配列データベースが設置されている。

<https://sc2.ddbj.nig.ac.jp/>



大学共同利用機関法人 情報・システム研究機構 国立遺伝学研究所
スーパーコンピュータシステム
SuperComputer Facilities of National Institute of Genetics

サイトポリシー
アップ

検索...

検索

現在地: ホーム ▶ 利用可能DB

Language/言語

JA EN

ホーム

このサイトへのログイン

ユーザ名

ユーザ名

パスワード

パスワード

ログイン

システム構成

ハードウェア構成

ソフトウェア構成

利用可能DB

システム使用方法

基本的利用方法

その他UGE利用方法

スーパーコンピュータシステム 利用可能DB

スーパーコンピュータシステムでは、各計算ノード、各ログインノードから各種バイオ系DBが利用可能です。

1.DDBJ,NCBI,EBI等の公共DBを利用したい場合

スーパーコンピュータシステムにて利用可能なDBおよびパスは[利用可能DB一覧](#)をご覧ください。

2.DRAを含むその他のDDBJ DBを利用したい場合

上記利用可能DB以外のDDBJ DBについては[下記方法](#)にてデータをコピーしてご利用下さい。

Singularityコンテナ内からDBを参照したい場合

[SingularityコンテナからのDB利用方法](#)をご覧ください。

利用可能DB一覧

Available Databases

DB名	パス (/usr/local/seq/)	設置されているファイルの詳細	更新頻度
DDBJ- unified-all	-	-	毎日
	fasta/	ddbj-unified-all/ ftp://ftp.ddbj.nig.ac.jp/ddbj_database/ddbjnew/unified-all/fasta/以下を解凍したFASTA形式ファイル	
	blast/	ddbj-unified-all/ ftp://ftp.ddbj.nig.ac.jp/ddbj_database/ddbjnew/unified-all/blastdb/以下を解凍したBLASTデータベース	
DDBJ- unified-	-	-	毎日
	fasta/	ddbj-unified-new/ ftp://ftp.ddbj.nig.ac.jp/ddbj_database/ddbjnew/unified-new/fasta/以下を解凍したFASTA形式ファイル	



遺伝研スパコンの主要Bioinformaticsツールは全て、「Singularity」コンテナで提供されている

* Singularity = 共用計算機システム上での高性能な計算環境(HPC)に適した、Linuxコンテナソフトウェア

+ 遺伝研スパコンだけでなく、産総研AIスパコン (ABCI) でも提供

* 2019年2月時点で約9000程度のツール利用可

	Docker	仮想マシン	Singularity
サービス			
動作	○	○	○
結果	<ul style="list-style-type: none">・ 一般ユーザーにDockerを使わせる設定をする とセキュリティなどの問題も対処が必要。・ GUIを表示させるための設定も手間がかかる。	<ul style="list-style-type: none">・ 仮想マシンをその都度起動されるのは面倒。・ 常時起動させるとCPU、メモリなどのリソース を無駄にしまうためあまりよくない。	<ul style="list-style-type: none">・ 一般ユーザーにSingularityを使わせる ために特別な設定は必要なし！・ グラフィカルなソフトでも表示させるための 設定を追加する必要なし！
判定	△	△	◎



遺伝研スパコンの利用申請



遺伝研スパコンの新規利用申請

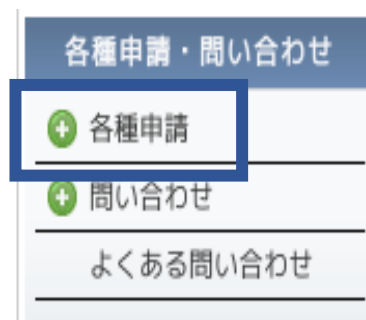
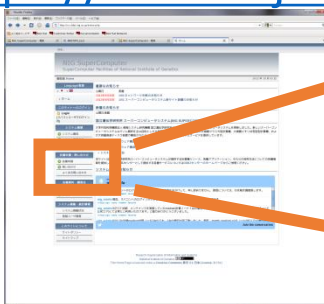
■新規ユーザ登録申請

<https://sc2.ddbj.nig.ac.jp/index.php/ja-new-application>

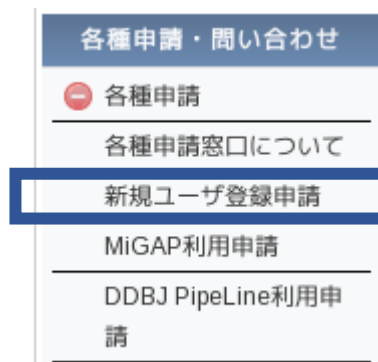
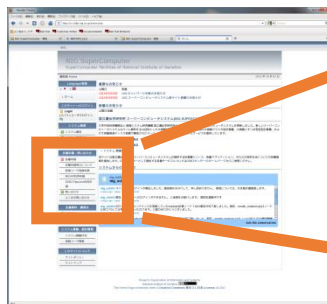
→登録画面が開く

or

<https://sc2.ddbj.nig.ac.jp/>



of supercomputer



on the page for the use of the supercomputer



新規ユーザ登録申請の記入①

(アカウント名・名前・所属等の**必須分**を記入してください)

■ 遺伝研スパコンの利用申請ページ

大学共同利用機関法人 情報・システム研究機構 国立遺伝学研究所
スーパーコンピュータシステム
SuperComputer Facilities of National Institute of Genetics

サイトポリシー サイトマップ

検索... 検索

現在地: ホーム > 新規ユーザ登録申請 2018 年 07 月 04

Language/言語
日本 英国

ホーム

このサイトへのログイン

ユーザ名
ユーザ名
パスワード
パスワード
ログイン

システム構成
ハードウェア構成
ソフトウェア構成
プログラミング環境
利用可能バイオツール
利用可能OSS
利用可能DB

システム使用方法
基本的利用方法

Page 1 of 2

新規ユーザ登録申請

アカウント申請に際しては

- 国立遺伝学研究所 スーパーコンピュータシステムの利用規程について
- ログインユーザーアカウント発行基準

をご一読ください。

※利用目的 = データサイエンスの講義利用のため

※チェックを入れる

記入例

希望アカウント名 (必須) アカウント名は15文字以内の英数字の半角小文字で指定してください。記号は-(ハイフン)のみ使用可能です。

利用目的 (必須) http://やhttps://から始まるURL情報を記入しないで下さい。

倫理指針 (必須) ☒ 「ヒトゲノム・遺伝子解析研究に関する倫理指針」上、問題となるような解析はいたしません。

申請者の名前 (必須) 姓 名 利用者の姓と名を記入してください。

申請者の名前(ローマ字) (必須) 姓 名 利用者の姓と名(ローマ字)を記入してください。



利用申請の記入②（責任者の欄の記入）

→年度末まで利用可。問い合わせから責任者の変更申請をすれば、
次年度以降もアカウントは継続して使えます

■責任者欄（必須分）に講師情報を入力

責任者情報

所属、利用目的は公表されることになります。また、利用結果又は過程の報告書についても原則公開されることになります。
責任者の情報は、これらが判断できる方の情報をお願いします。

記入例

☐ 利用者情報をコピーする

研究責任者の名前 (必須)	姓 神沼	名 英里	責任者の姓と名を記入してください。
研究責任者の名前(ローマ字) (必須)	姓 Kaminuma	名 Eli	責任者の姓と名(ローマ字)を記入してください。
所属機関 (必須)	東京医科歯科大学		責任者の所属機関を記入してください。
研究分野 (必須)	バイオインフォマティクス 文部科学省 平成27年度科学研究費補助金 系・分野・分科・細目表 http://www.jsps.go.jp/j-grantsinaid/03_keikaku/data/h27/h27_koubo_08.pdf		
職名 (必須)	特任講師		責任者の職名を記入してください。
メールアドレス (必須)	ekmds@tmd.ac.jp		責任者のメールアドレスを記入してください。
郵便番号 (必須)	113-8510		責任者の所属機関の所在地の郵便番号を記入してください。
住所 (必須)	東京都文京区湯島1-5-45		責任者の所属機関の所在地の住所を記入してください。
電話番号 (必須)	03-5803-4176		責任者の所属機関の電話番号を記入してください。(携帯番号不可)
内線番号	内線があれば内線番号を記入してください。		
国籍 (必須)	日本		責任者の国籍を記入してください。



遺伝研スパコンへの SSH公開鍵設置とログイン



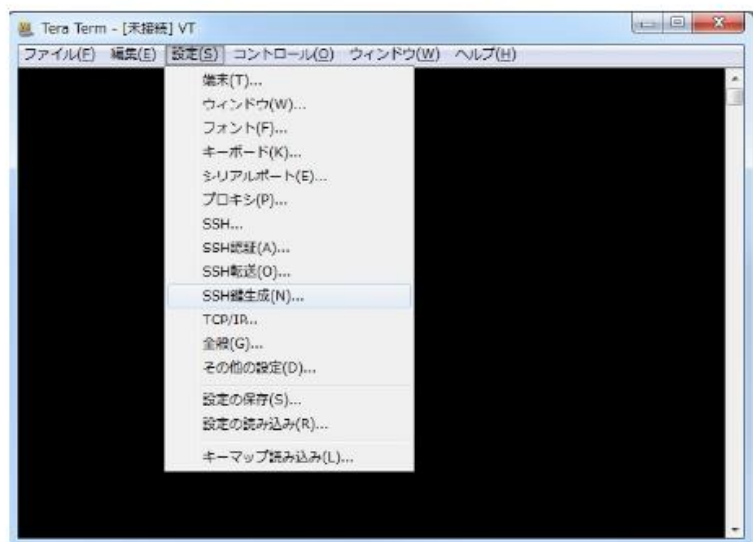
TeraTermでSSH公開鍵生成

<https://sc2.ddbj.nig.ac.jp/index.php/2014-09-17-05-42-33>

Windows (TeraTerm) topへ移動

(1) ユーザ端末におけるペア鍵の生成

TeraTerm で鍵を生成するには、TeraTermを起動します。



、【設定】－【SSH鍵生成】にて鍵の生成を行います。

- ①TeraTerm起動
- ②設定> SSH鍵生成
- ③種類: RSA, ビット数:2048,を選択
- ④パスフレーズ入力

絶対忘れないフレーズにする

- ⑤公開鍵の保存 場所はデスクトップ
- ⑥秘密鍵の保存 場所はデスクトップ



鍵の種類をRSA, ビット数(B)を2048として【生成】ボタンを押下して、鍵を生成します。



鍵のパスフレーズを入力後、【公開鍵の保存】(id_rsa.pubファイルの保存)、【秘密鍵の保存】(id_rsa,id_rsa.pubを作成します)。



遺伝研スパコンへ公開鍵設置

<https://sc2.ddbj.nig.ac.jp/index.php/2014-09-17-05-42-33>

(2) 遺伝研ゲートウェイノードへの公開鍵の設置

作成したid_rsa.pubをファイルのアップロードフォームでアップロードして下さい。

・ファイルのアップロードフォーム

① クリックでウェブサイトログイン。
登録済アカウントとパスワードを入力

② 「ファイルを選択」でデスクトップに生成されたid_rsa.pubを選択

SSH公開鍵ファイル (必須)

ファイルを選択

公開鍵登録

③ 「公開鍵登録」ボタンを押下

画面下部 (赤枠) の【選択】を押し、ファイル選択画面寄り、アップロードするファイルを選択後、画面下部 (赤枠) の【公開鍵登録】ボタンを押して登録を行う。

遺伝研スパコンでジョブ投入

スパコン使用方法(イメージ)

- ①ゲートウェイノード(gw.ddbj.nig.ac.jp)にログインする
- ②qloginを実行しインタラクティブノードにログインする
- ③qloginしたホストからジョブをUGEに投入する
- ④UGEは負荷の低いノードでジョブを実行する
- ⑤ジョブ実行結果をlustreのホームディレクトリに出力する
- ⑥ジョブ実行結果を確認する



遺伝研スパコンへのログインから Thin nodeでのジョブ実行までの流れ

1. Connect to the gateway node using ssh.

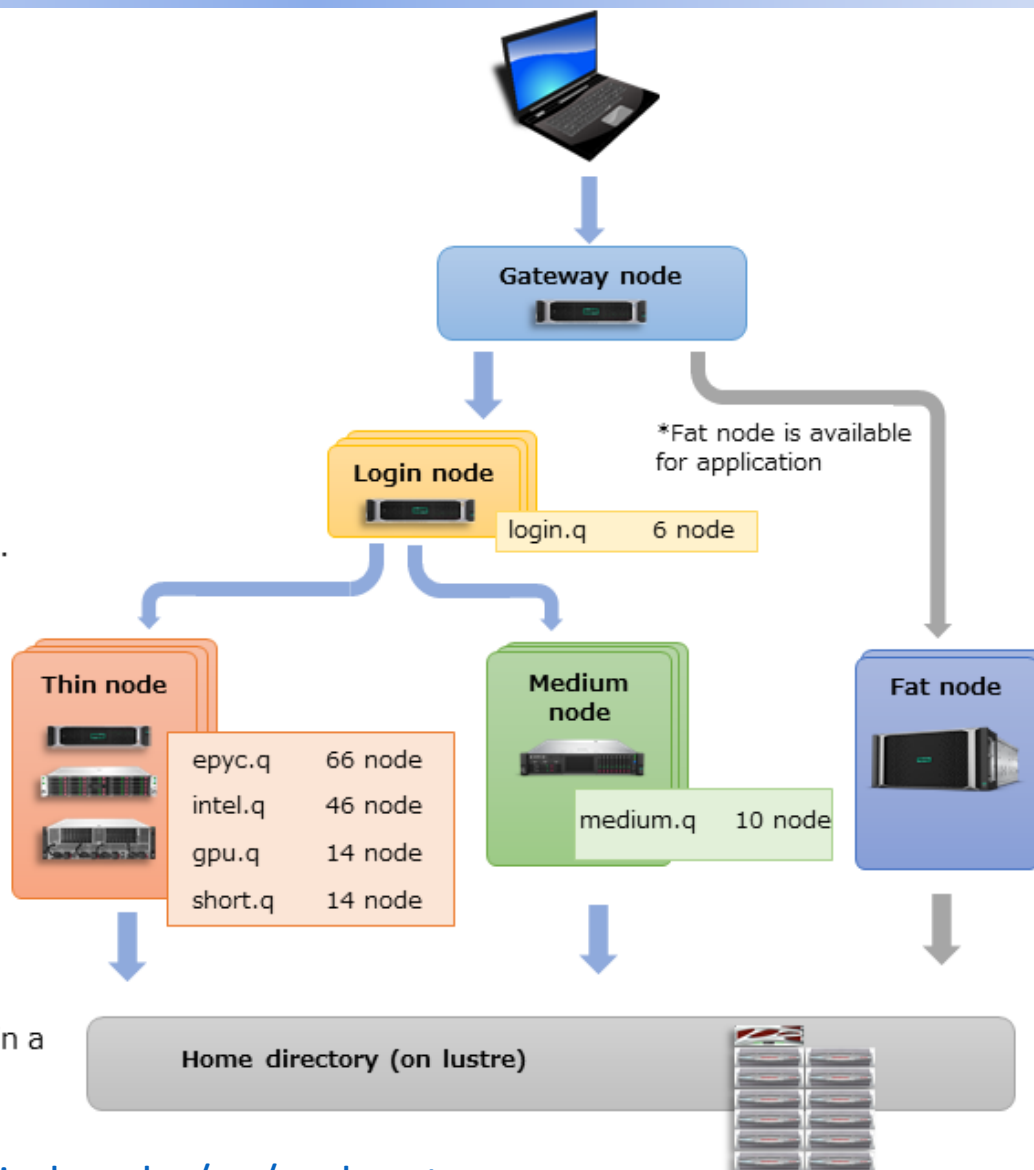
2. Login to a Login node.

3. Development processes and interactive processes are implemented on the login nodes.

4. Computation job is entered in each queue.

5. Execution of various jobs on various compute nodes.

6. Computation results are output in a shared home directory in the compute node group.





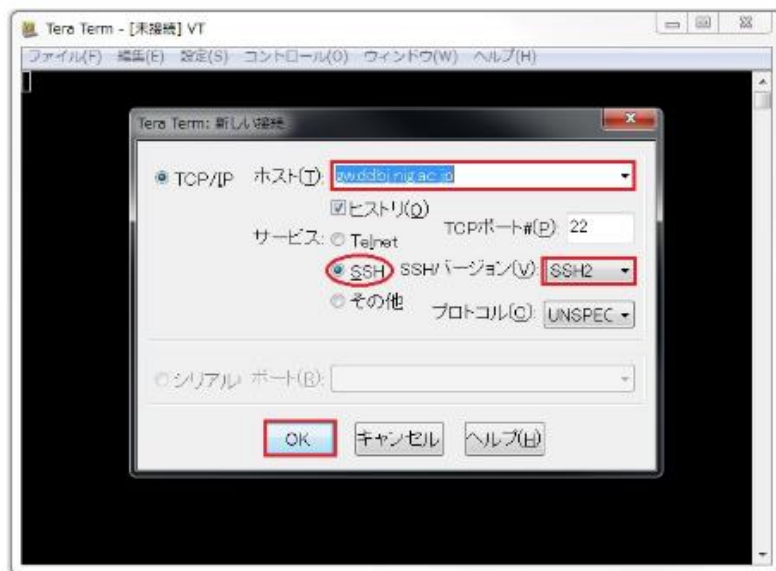
遺伝研スパコンへログイン実行①

<https://sc2.ddbj.nig.ac.jp/index.php/2014-09-17-05-42-33>

①接続情報入力

(3) 遺伝研ゲートウェイノードへのSSH接続確認

TeraTermを起動します。



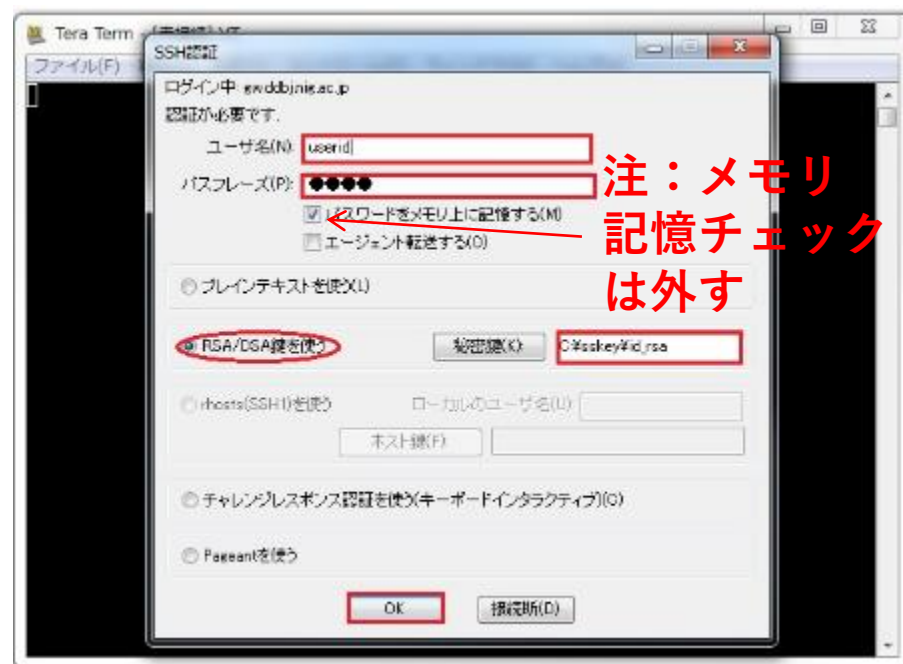
ホスト名：gw.ddbj.nig.ac.jp

ポート：22

サービス：SSH

サービスバージョンSSH2を設定し、
【OK】ボタンを押下する。

②パズフレーズ入力



スーパーコンピュータシステムにログインするユーザID、鍵生成時に指定したパズフレーズを入力する。

「R A S / I S A 鍵を使う」を選択し、作成した秘密鍵を選択し、【OK】ボタンを押下する。



遺伝研スパコンへログイン実行②

<https://sc2.ddbj.nig.ac.jp/index.php/2014-09-17-05-42-33>

SSH認証

ログイン中: gw2.ddbj.nig.ac.jp

認証が必要です.

ユーザ名(N): kaminuma

パスワード(P): ●●●●●●●●

☐ パスワードをメモリ上に記憶する(M)

☐ エージェント転送する(O)

☐ プレインパスワードを使う(L)

☒ RSA/DSA/ECDSA/ED25519鍵を使う

秘密鍵(K): C:\Users\kaminu\

☐ rhosts(SSH1)を使う

☐ キーボードインタラクティブ認証を使う

☐ Pageantを使う

OK

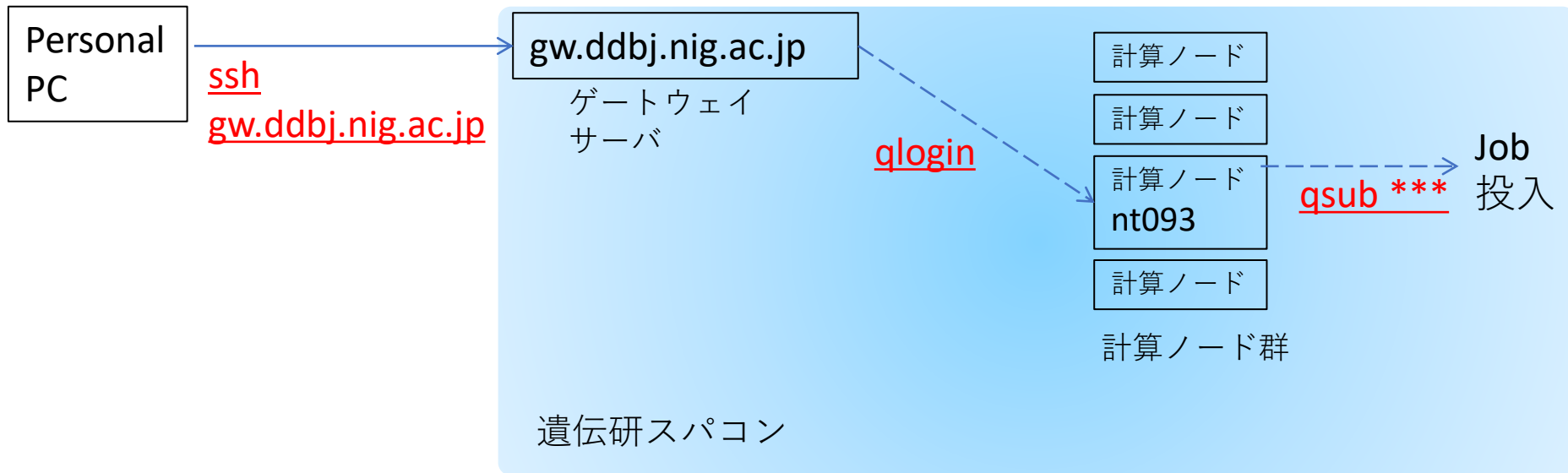
①ログインしたらすぐにqlogin実行

```
[kaminuma@gw2 ~]$  
[kaminuma@gw2 ~]$ qlogin  
Your job 10857340 ("QLOGIN") has been submitted  
waiting for interactive job to be scheduled ...  
Your interactive job 10857340 has been successfully scheduled.  
Establishing /home/geadmin2/UGER/utilbin/lx-amd64/qlogin_wrapper session to host nt093i ...  
Last login: Mon Jul 9 19:55:11 2018 from nt092i  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$
```

②計算ノードnt093が割り当てられた事が判る



sshコマンドでスパコンにログインしたら、
必ずqloginで計算ノードに移動



```
[kaminuma@gw2 ~]$  
[kaminuma@gw2 ~]$ qlogin  
Your job 10857340 ("QLOGIN") has been submitted  
waiting for interactive job to be scheduled ...  
Your interactive job 10857340 has been successfully scheduled.  
Establishing /home/geadmin2/UGER/utilbin/lx-amd64/qlogin_wrapper session to host nt093i ...  
  
Last login: Mon Jul  9 19:55:11 2018 from nt092i  
  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$  
[kaminuma@nt093 ~]$
```

※gwサーバ(スパコン入口)で計算を始めたら他のユーザに大迷惑！



(初回のみ) 環境設定ファイルを ホームディレクトリにコピー

■ bashの環境設定 2種類の設定ファイル

設定ファイル	実行	記載内容
.bash_profile	ログイン時に一回実行	環境変数 (特にPATH変数の設定)
.bashrc	シェル起動時に一回実行 (ジョブ投入毎に実行)	alias設定など

■ bashの環境設定をコピー

※bの前のドットに注意

```
cp /home/kaminuma/.bashrc_dsc .bashrc
source ~/.bashrc
cp /home/kaminuma/.bash_profile_dsc .bash_profile
source ~/.bash_profile
```

<.bashrcの中身>

```
alias ls='ls -aCF --color'
```




qsub コマンドでジョブ投入

https://sc2.ddbj.nig.ac.jp/images/stories/meetingdoc/20120510/ja/ja_Overview_of_UGE-.pdf

■Shell Scriptを作成（test.shで保存）

```
#!/bin/sh
#$ -S /bin/bash
#$ -cwd

pwd
hostname
sleep 10
echo 'error message here' 1>&2
```

（シェルは**bash**利用）
（現ディレクトリに出力）

■qsubコマンドでジョブ投入

```
$ cp /home/kaminuma/tmp_data/test.sh .
$ qsub test.sh
```

■ジョブ終了後に標準出力、標準エラー出力の記録ファイルが生成される

```
$ ls -l
$ cat ~/test.sh.o325
/home/kaminuma/tmp_data
nt030
$ cat ~/test.sh.e325
error message here
```



qstat コマンドでジョブの状態確認、
qdel コマンドでジョブを消去。

```
[kaminuma@nt095 ~]$ qstat -u "*" |grep medium
```

JobID	UserID	JobName	State	Start Time	End Time	NodeName
10740562	0.74882	STR	r	06/12/2018 14:49:32		month_medium.q@nm1i
10740557	0.48541	STR	Rr	06/27/2018 20:34:17		month_medium.q@nm6i
10740558	0.48541	STR	Rr	06/27/2018 20:34:17		month_medium.q@nm2i
10740559	0.48541	STR	Rr	06/27/2018 20:34:17		month_medium.q@nm2i
10740560	0.48541	STR	Rr	06/27/2018 20:34:17		month_medium.q@nm2i
10740561	0.48541	STR	Rr	06/27/2018 20:34:17		month_medium.q@nm2i
10819782	0.47339	MINion_SSP	r	06/28/2018 16:15:29		month_medium.q@nm2i
10845054	0.38353	calc_bfs_e	r	07/03/2018 21:01:45		month_medium.q@nm2i
10741097	0.38229	pilon	Rr	07/03/2018 19:45:28		month_medium.q@nm2i
10846363	0.37116	MINion_SMI	r	07/04/2018 11:12:34		month_medium.q@nm2i
10846441	0.37044	job_vert	r	07/04/2018 12:12:38		month_medium.q@nm1i
10857100	0.34157	platanus.f	r	07/09/2018 11:13:43		month_medium.q@nm6i
10853329	0.33607	job_vert	r	07/06/2018 11:55:50		month_medium.q@nm2i
10853331	0.33606	job_vert	r	07/06/2018 11:56:51		month_medium.q@nm2i
10853342	0.33605	job_vert	r	07/06/2018 11:57:21		month_medium.q@nm2i
10853363	0.33605	job_vert	r	07/06/2018 11:57:36		month_medium.q@nm2i
10853374	0.33604	job_vert	r	07/06/2018 11:58:21		month_medium.q@nm7i
10855736	0.33227	lastal-LMP	r	07/06/2018 20:12:05		month_medium.q@nm2i
10855737	0.33227	lastal-LMP	r	07/06/2018 20:12:05		month_medium.q@nm2i
10857116	0.28616	lastal-LMP	r	07/09/2018 12:13:27		month_medium.q@nm2i
10857117	0.28616	lastal-LMP	r	07/09/2018 12:13:27		month_medium.q@nm2i
10857076	0.28538	round1_run	r	07/09/2018 10:18:53		month_medium.q@nm2i
10857577	0.28339	test	ki	07/10/2018 12:03:58		month_medium.q@nm1i
10857809	0.26724	q_SRAWGS	r	07/10/2018 18:29:44		month_medium.q@nm8i
10857909	0.26016	umap_py_EA	o r	07/10/2018 21:20:08		month_medium.q@nm8i
10857910	0.26016	umap_py_AL	o r	07/10/2018 21:20:08		month_medium.q@nm8i
10857911	0.26016	umap_py_BB	o r	07/10/2018 21:20:23		month_medium.q@nm7i
10857912	0.26015	umap_py_BB	o r	07/10/2018 21:20:38		month_medium.q@nm4i
10857913	0.26015	umap_py_AL	o r	07/10/2018 21:20:38		month_medium.q@nm4i
10857914	0.26015	umap_py_EA	o r	07/10/2018 21:20:53		month_medium.q@nm4i
10857960	0.25754	stargenome	r	07/11/2018 00:57:59		month_medium.q@nm1i

medium node(メモリ
中規模計算ノード)
でのジョブとユーザ
をチェック

```
$ qsub test.sh
```

Your job 327 ("test.sh") has been submitted

```
$ qdel -u kaminuma
```

kaminuma has registered the job 327 for deletion



投入済ジョブは一定時間が経つと強制終了される

■キュー制御による実行時間制限

Phase3システム 2019年3月導入

キュー名	ジョブスロット数 (キュー全体)	ジョブスロット数 (ノード当たり)	メモリ (キュー全体)	メモリ (ノード当たり)	実行時間の 上限	用途	キュー指定のオプション
epyc.q	4,224	64	33.792TB	512GB	62日	実行時間2ヶ月間で他にリソース要求が無い時	指定なし または -l epyc
intel.q	1,472	32	17.664TB	384GB	62日	Intel Xeonを使用	-l intel
gpu.q	384	8	2.688TB	192GB	62日	GPUを使用	-l gpu -l cuda=n (nは1-4の整数) 詳細は こちら をご覧ください。
medium.q	800	80	30TB	3TB	62日	Medium計算ノードを使用	-l medium
login.q	258	64	3.072TB	512GB	無期限	ゲートウェイノードからqloginする際に利用 AMD CPU搭載ノードのみ	
login_gpu.q	48	24	768GB	384GB	無期限	GPU、Intel CPUを使用する際にqloginして利用	-l gpu
short.q	744	16	2.688TB	192GB	3日	短時間ジョブ向け	-l short

■ジョブの実行上限時間の指定

```
qsub -l d_rt=192:00:00 -l s_rt=192:00:00 test.sh
```

(実行上限時間に192時間(8日×24時間)を指定)