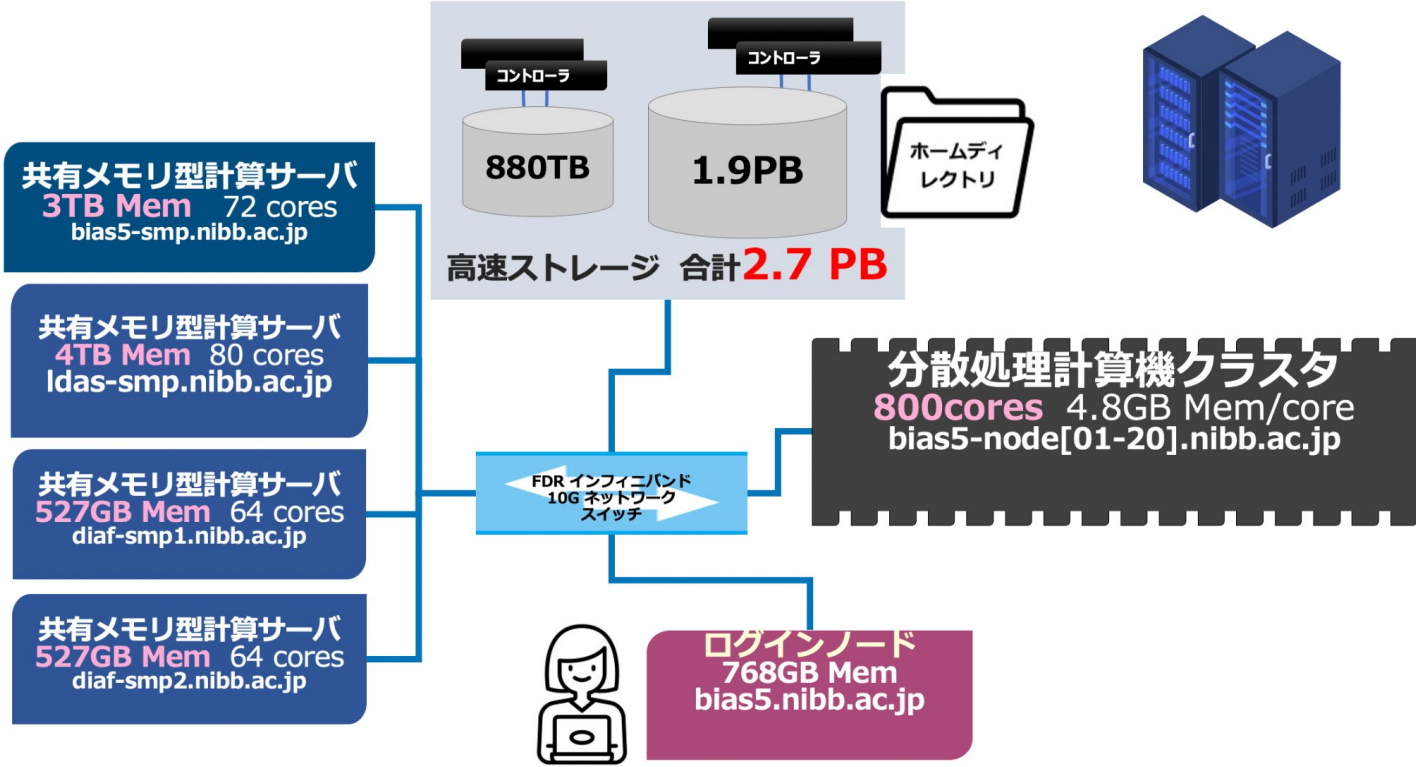


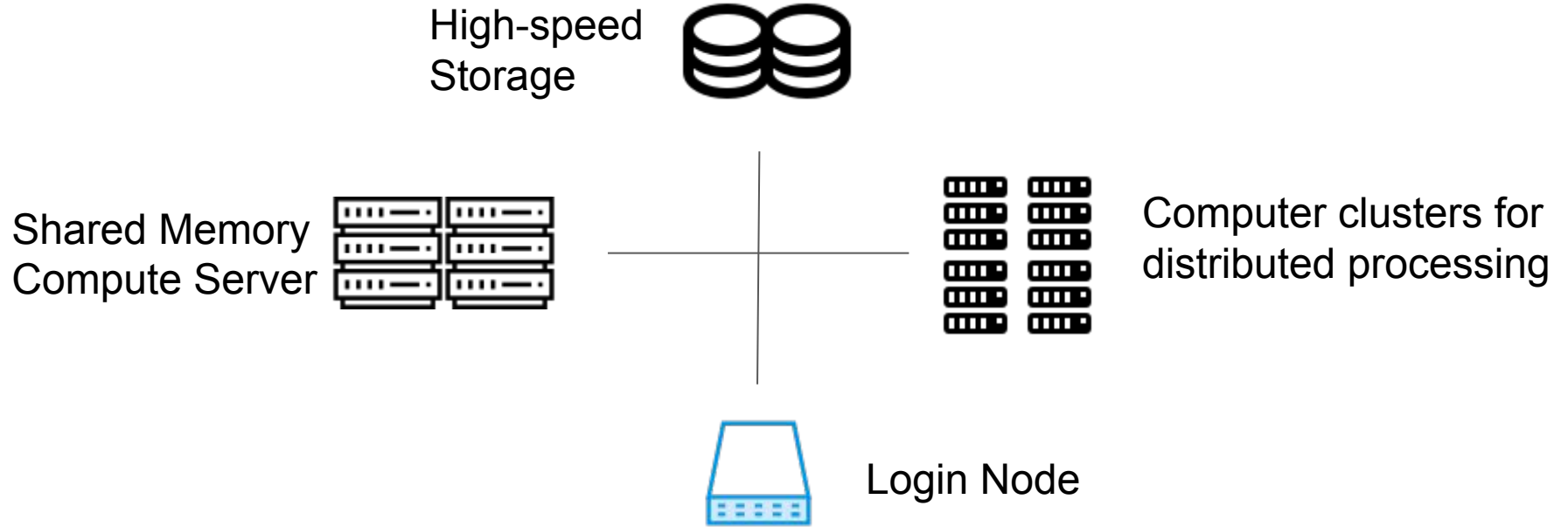
BIAS5 overview

NIBB Data Integration and Analysis Facility
Takanori NAKAMURA

Biological Information Analysis System (BIAS5)



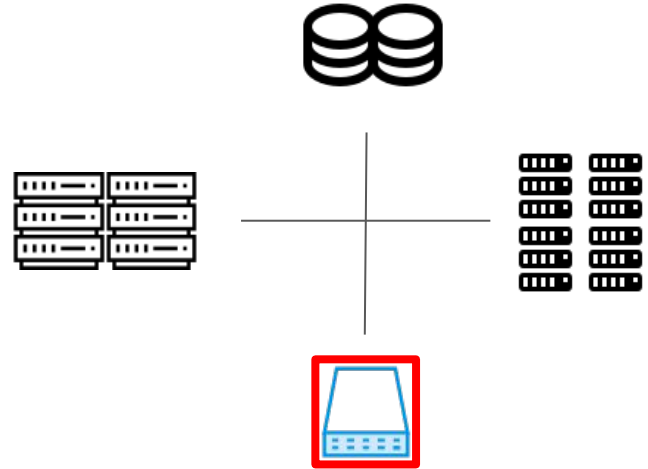
BIAS5 overview



Login Node "bias5-login"

The server where the user first login with SSH command.

You can view and write your own files. The file entity is on high speed storage.



Please do not run the program on the login node.

Run it on another compute server using the job management system (PBS) described below.

Computer clusters for distributed processing

“bias5-node[01-20]”

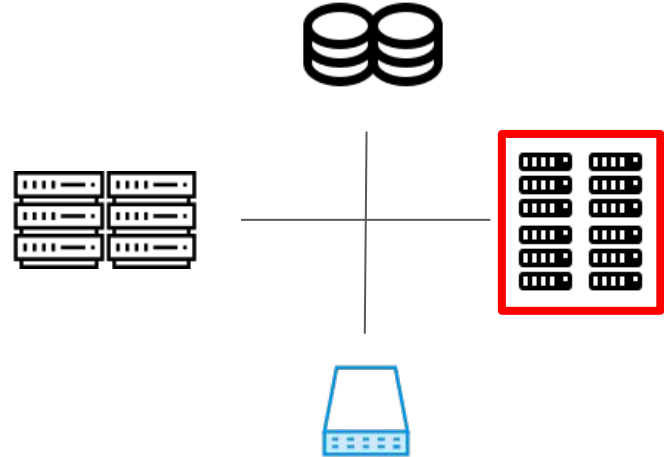
A group of servers that perform massively parallel computations.

Users do not login directly.

To use it, you will have to throw a job using PBS.

Recommended for programs that require extensive CPU processing.

Each node does not have much archive capacity.
Please be careful when keeping logs on a node.



CPU: Intel Xeon Gold 6140
(2.3 GHz) 72 cores /
Memory : 3TB

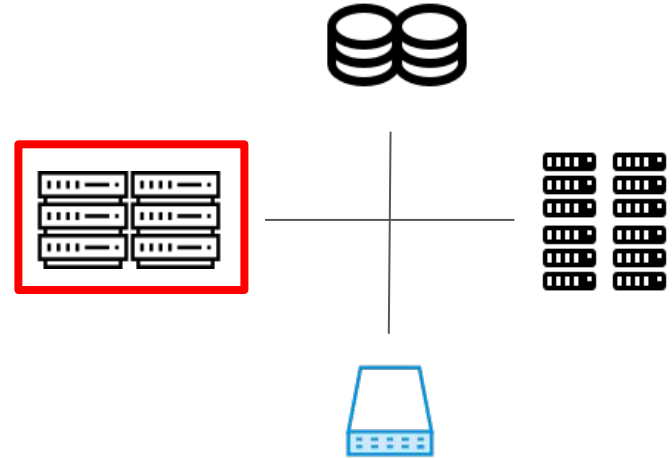
Shared Memory Compute Server “bias5-smp”, “ldas-smp”

A group of servers that perform calculations that use large amounts of memory. Users do not log in directly.

To use it, you will have to throw a job using PBS.

Recommended for programs that require extensive file processing.

The size of memory to be used can be specified in PBS.



bias5-smp

CPU: Intel Xeon (2.3 GHz) 72 cores

Memory : 3TB

ldas-smp

CPU: Intel Xeon (2.4 GHz) 80 cores

Memory: 4TB

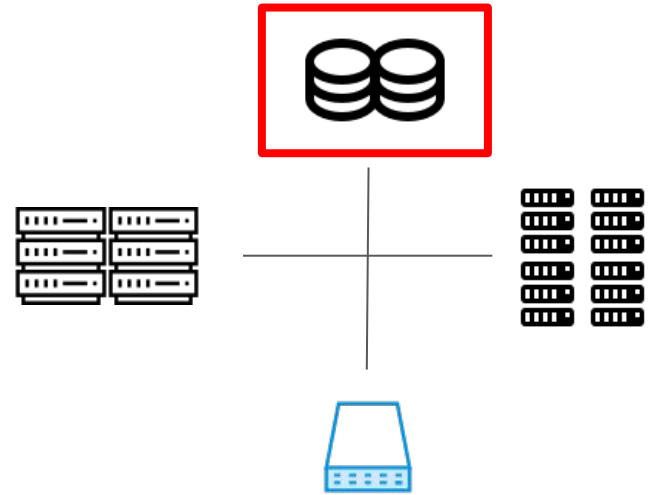
High-speed Storage

This is a group of storage devices, including the user's home directory, where data is stored. Users do not log in directly; it is mounted and available when they enter the login node.

Storage systems are differentiated by file paths according to their use, and each has its own quota (capacity limit).

We accept capacity changes.

Please contact the Data Integration and Analysis Facility.



Storage Capacity

880TB (A) + 1.9PB (B)

How to use

First, log in to the login node (`bias5.nibb.ac.jp`).

- The login node only performs non-resource consuming processes such as program creation, and job execution is always done through the PBS.
- To use distributed processing clusters and shared memory servers, submit jobs to their respective dedicated queues.
- For more information, please visit our page <https://biaswiki.nibb.ac.jp/menu>.

Login to bias5

Login to bias5 is done using the ssh command.

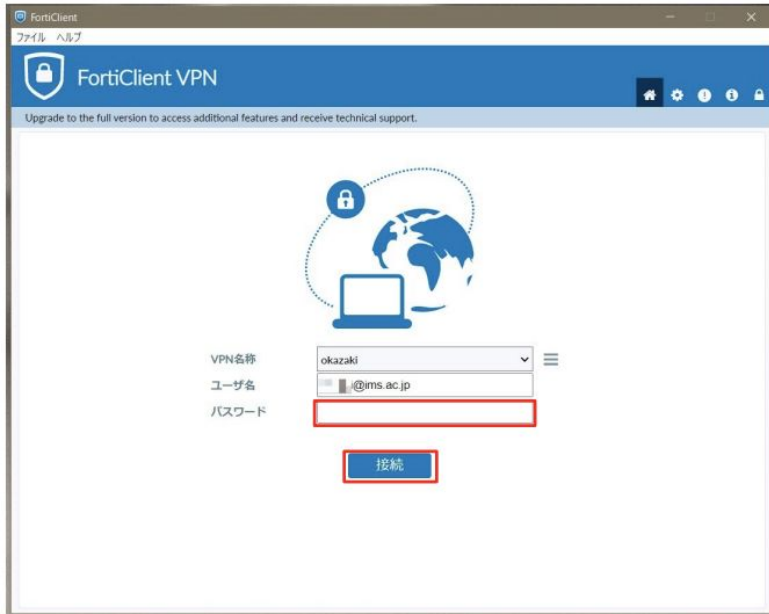
Users on the internal network can login only with the ssh command.

Users on the external network use the VPN to belong to the internal network and then login using the ssh command.

As for VPN, a specific client is used.

VPN client (FortiVPN / FortiCleint)

For external users:
Use FortiVPN.



For Internal users (NIBB users):
It is not necessary to use a VPN if it is an internal network.
Use FortiClient to connect from external network.
(SAML authentication is required)



Login to bias5

To login to bias5 using the ssh command, execute the command as follows;

```
$ ssh user-account@bias5.nibb.ac.jp
```

"user-account" is the name of the account to be used on bias5.

After executing the command, you will be asked for a password, enter your password.

Example:

```
$ ssh unixr@bias5.nibb.ac.jp
unixr@bias5.nibb.ac.jp's password:
Last login: xxx xxx xx xx:xx:xx 20xx from xxxxx.nibb.ac.jp
[unixr@bias5-login ~]$
```

Logout from bias5

To log out of bias5, run the exit command.

```
$ exit
```

After logging out, you will be returned to the previous terminal where you logged into bias5.

Example

```
[unixr@bias5-login ~]$ exit  
logout
```

```
Connection to bias5.nibb.ac.jp closed. ←接続が解除されたメッセージが出る。
```

```
$
```

Remote file copy (scp)

To send a file from your local terminal to bias5, use the scp command.
This allows files to be exchanged over the network.



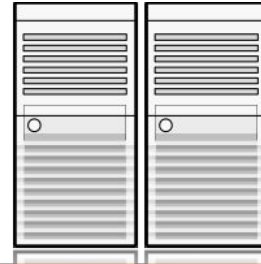
```
$ scp [copyfile-path] user-account@bias5.nibb.ac.jp:[file-path]
```

```
$ scp file1.dat user-account@bias5.nibb.ac.jp:temp/data/
```

The above command will copy the file named file1.dat in the local current directory to temp/data/file1.dat in the bias5 home directory.

Remote file copy (scp)

Opposite to the previous step, sending a file from bias5 to your local area is also done using the scp command. Note where the scp command is executed from and the source/destination of the file.



bias5.nibb.ac.jp

```
$ scp user-account@bias5.nibb.ac.jp:[copyfile-path] [local-path]
```

```
$ scp user-account@bias5.nibb.ac.jp:temp/data/result.dat ~/Desktop/result
```

The above command will copy the file temp/data/result.dat in bias5's home directory to your local ~/Desktop/result/result.dat.

Quota

- A quota (limit) is placed on the home directory. The default is 3TB.
- If you are using a large amount of disk space temporarily, use the scratch volume (/scratch/username) where no quota is applied. However, files in the scratch volume will be deleted after 30 days.
- Files that you want to keep long-term as backups or archives should be saved in the Save volume.
- To check the current usage, run the following command;

```
$ mmlsquota gpfsA:home
```

File limits

name	path	per user	to use
home	<i>/home/user_name</i>	3TB(max: 6TB)	your current directory
scratch	<i>/scratch/user_name</i>	No limit	temporarily (30 days limit)
save	<i>/home/user_name/save</i>	6TB	storage or archive

The home directory can be increased up to 6TB by applying for disk space expansion.

Job Management System (PBS)

A job management system is a system that accepts and manages a large number of processing requests (jobs) and allocates computing resources such as CPU and memory appropriately for sequential execution.

In bias5, PBS (Portable Batch System) is used as a job management system.

Basically, all jobs should be run through PBS.

Do not run them on login nodes.

Several queues are available depending on the scale and type of calculation. By selecting these queues appropriately, work can be performed efficiently.

Appendix; Queue setting

Queue	small	medium	large	blast	smps	smpm	smpi
Job	short time	medium time	large time	BLAST only	medium memory	large memory	exlarge memory
use on	node (01-20)	node (01-20)	node(01-20)	node(01-20)	shared memory server	shared memory server	shared memory server
max time	6 hours	72 hours	no limit	72 hours	no limit	no limit	no limit

reference : https://biaswiki.nibb.ac.jp/menu/index.php/ジョブ管理システム_PBSを用いた計算機の利用

Appendix: Installed biotechnology-related software (part of)

Software for next-generation sequencing analysis mapping, RNA-Seq analysis,

Bowtie2, HISAT2, BWA, StringTie, Trinity-RNAseq, salmon, Velvet, ABySS, AllPaths-LG

samtools, bamtools, BEDtools, cutadapt, SRA-toolkit

GeneMark, Augustus, BRAKER

lastz, MUMmer, BLAT

ClustalW, Muscle, MAFFT

Phylip, PhyML, MrBayes

InterProScan, HMMER, MEME

DBGET

EMBOSS, BLAST+, Diamond,

more info: <https://biaswiki.nibb.ac.jp/menu/index.php/ソフトウェア>

Appendix: Available biotechnology-related databases (part of)

- BLAST path: `/bio/db/blast/db`
- diamond path: `/bio/db/diamond/db`
- flat files path: `/bio/db/ideas`

name of DB	remark	data format	update
Genbank, EMBL	nucleotide sequence	DBGET	Regularly
gss, HTGS, dbsts, patnt, env_nt	nucleotide sequence	FASTA, BLAST	Regularly
RefSeq(Nuc+Pep)	Nucleotide sequence + Protein amino acid sequence	DBGET	Daily
NCBI nr-nt	non-redundant nucleotide sequence	FASTA, BLAST	Regularly

more info: <https://biaswiki.nibb.ac.jp/menu/index.php/分子生物学データベース>