User Group Meeting

December 2025

OIST

Storage Update – Storage Outages!

The Next Cluster!

Using Flash

Slack Is Discontinued

◆◆
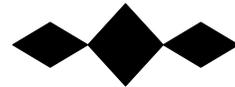
Say Hi to Patricia, our New Bioinformatics Specialist!

● ● ●

Bioinformatics user group:
When do we delete old databases?

New Bucket

- **60PB total capacity**
  ...but not all goes to research storage

Installation

done {
- ✔ Initial sync of Bucket to new system
- ✔ Finish hardware installation

Needs downtime {
- ☐ Final Sync of new and changed data
- ☐ Unmount old, mount new system
  - Bucket itself, also other file systems

do later
- ☐ Switch monitoring and analytics over
  - **"Storage Inspector"** will show stale data for some time.
- ☐ Also other internal bits and pieces

New Bucket

- **60PB total capacity**
  ...but not all goes to research storage

New!!

Installation

Needs downtime {
- ☐ Final Sync of new and changed data
- ☐ Unmount old, mount new system
  - Bucket itself, also other file systems

Complete downtime: **Feb 02 – Feb 06**

- Bucket
- Deigo and Saion
- Other storage systems

We will *try* to be faster, but may need the entire week

Please Plan ahead

New Bucket

- 60PB total capacity
  ...but not all goes to research storage

Installation

- ☐ Switch monitoring and analytics over
  - **"Storage Inspector"** will show stale data for some time.
- ☐ Also other internal bits and pieces

There may be a second outage

~end of February

- Most likely to only impact specific services
  → We will contact affected users directly

Post-DE·i·GO

Deigo is 6 years old (Time flies!)

- Installed early 2020

    right when COVID started 🤧

- Vendor support ~5-7 years typical
- Older nodes are energy inefficient
- Repairs no longer cost-effective
- Difficult to physically move

DC2 – Data Center 2

DC2

- Purpose-built as DC
- Separated space for storage, two clusters
- Direct access for heavy transport
- Set up for liquid cooling
- Sufficient power

Post-DE·i·GO

Rough Timeline

- Now: New storage

- April— : Tender process for new cluster
- Autumn: Decision

- Early 2027: Installation starts
- Mid-2027: Available to you

Design: CPU-only, more cores/node, more memory.

Name: ?

Using /Flash (and /Work)

Design Capacity: **500 TB**

Allocated Capacity: **1100 TB**

Flash is *only* for running jobs.

Never keep data on Flash.

Do *you* have a job running on Deigo?

NO?

↓

you should have

0 bytes

on Flash

GPU Node Update:

V100 GPU nodes are upgraded

- Rocky 8.8, CUDA 12.8 etc.
- Access the nodes with the "v100" gres parameter:

```
--gres=gpu:v100:1
```

- Remaining nodes will be upgraded soon(ish)

Matlab:

Newer (2024+) matlab starts *very slowly* on the cluster when interactive

- Not a bug
- We're supposed to use "-batch" or remote execution on the cluster

---

Geneiuous Prime is Discontinued
Last day is August 2026

IBM SPSS Maintenance Discontinued
No more Version Updates; Current version will stay available

The Slack OIST account will be discontinued

We have Teams for internal use
yes, yes — they all suck
>2x Teams over Slack

Slack at OIST:

- For connecting with external users
- Discount over unit accounts

But:

We have no cost control

- We pay per user, but you can invite as many as you like

  → We can't know the cost beforehand

End Date: **December 11, 2026**

The Slack OIST account will be discontinued

We have Teams for internal use

    yes, yes — they all suck

    >2x Teams over Slack

Slack at OIST:

- For connecting with external users
- Discount over unit accounts

But:

We have no cost control

- We pay per user, but you can invite as many as you like

  → We can't know the cost beforehand

End Date: **December 11, 2026**

IT now provides "**Access Packages**"

- Per-project Teams/Sharepoint workspaces w. external access

Repeated price increases:
no longer much cheaper than unit accounts

The Slack OIST account will be discontinued

We have Teams for internal use
  yes, yes — they all suck
  >2x Teams over Slack

Slack at OIST:

- For connecting with external users
- Discount over unit accounts

But:

We have no cost control

- We pay per user, but you can invite as many as you like
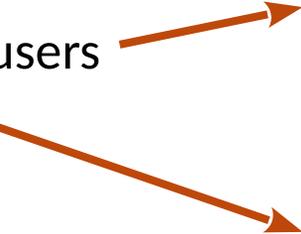 → We can't know the cost beforehand

End Date: **December 11, 2026**

IT now provides "**Access Packages**"

- Per-project Teams/Sharepoint workspaces w. external access

Repeated price increases:
no longer much cheaper than unit accounts

Faculty voted (80% in favour) to Discontinue the OIST Slack contract in 2026

The Slack OIST account will be discontinued

End Date: **December 11, 2026**

We have Teams for internal use

IT provides "**Access Packages**"

- Per-project Teams/Sharepoint workspaces w. external access

Faculty voted (80% in favour) to Discontinue the OIST Slack contract in 2026

Teams is common at OIST, no worse than Slack.

Units can use Slack with their own budget if they want

Migration:

- Slack migration tools are limited
- Individual accounts can (perhaps) not be transferred
- Slack charges 8 million yen to migrate a unit workspace

→ Export your data, start fresh

BioinfoUgrp database retention policy

- Storage space usage by databases:
  - DFAM (repeats): 2.5 TiB, 3 versions
  - BLAST: 2.5 TiB, 5 versions
  - Diamond: 0.6 TiB , 1 version

- We do not have the space to both update databases regularly and keep them for a long time.

- Our database retention policy is:
  - Keep for 2 years.
  - Announce removals on Teams channel.
  - Really delete the files 6 months after removing from the search path.

- Please load the database modules to help us collect usage statistics!

- By the way, please pay attention that our BLAST databases only contain curated sequences; for including the latest submissions you may want to download uncurated sequences as well, especially when focusing on some taxonomic clades.

- We want to hear your feedback!

Schedule

Dec 15 – 23:
Open Hours
Mon, Tue
With Patricia!

February 2 – 6:

Clusters and the Bucket storage
are **unavailable**

HPC Training:

Intro to HPC 1: January 19, 13-15
Command line: January 21, 13-15
Intro to HPC 2: January 22, 13-15

All in Room C700, 13-15; or Zoom

Bioinfo Training:

January/February:
    R on the cluster: mini analysis

February:
    Running nf-core pipeline
    on the clusters