

Treinamento prático para uso de HPC

Equipe de Dados Biológicos - LNBio



MINISTRY OF
SCIENCE TECHNOLOGY
AND INNOVATION



Computação de alto de desempenho

HPC - High Performance Computing



Grande quantidade de **processadores** - CPUs e GPUs



Grande quantidade de **memória RAM**

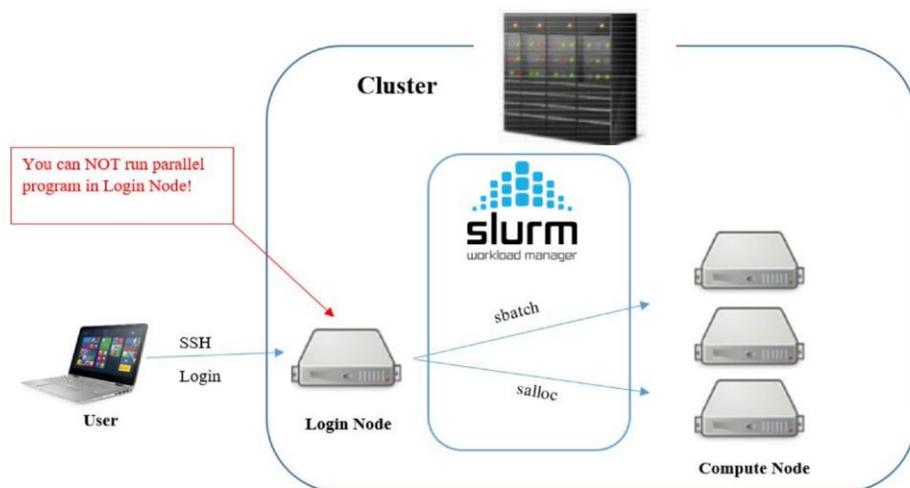


Grande capacidade de **armazenamento de dados** e acesso de alta velocidade



Rede de conexão de **alta velocidade**

Cluster de HPC



- Um **cluster** é um conjunto de computadores
- No cluster, cada computador é chamado de **nó**
- Há nós com funções específicas:
 - **Nó de acesso** (login node)
 - **Nós de processamento**
 - **Nós de armazenamento** (storage)

• No CNPEM

- Marvin – HPCC do LNBio
- TEPUI – HPCC do LNLIS
- Mendeleev – HPCC do LNNANO
- HPCC do LNBR

• CENAPAD/Unicamp

- cenapad.unicamp.br
- 10.280 CPUs
- 10 GPUs NVIDIA A100

• Santos Dumont – LNCC

- sdumont.lncc.br
- Top107
- 68.064 CPUs
- GPUs
 - 288 NVIDIA H100
 - 144 NVIDIA GH200
- Custo da expansão USD18.7M

Como utilizar?

Processamento em lote (batch)

- Usuário prepara um script
- Define os recursos que serão necessários para execução (quantos CPUs, quanta memória, vai usar gpu, quanto tempo)
- Submete para o sistema gerenciador de tarefas (SLURM)
- SLURM inicia a tarefa quando houver recursos disponíveis



CNPq

O custo de processamento e armazenamento é alto!

Como otimizar o uso em HPC:

- Multiusuário
- Flexível para atender diferentes demandas
- Reduzir recursos ociosos (reservados e subutilizados)

Sinais que um programa pode não ser adequado para HPC?

- Roda somente em Windows
- Roda bem em um laptop ou desktop
- É um programa de uso interativo (interface gráfica)
 - Há exceções:
 - GUI para preparação da tarefa (scipion, cryosparc, cellprofiler, FIJI)
 - GUI para dados muito grandes AVIZO/AMIRA, NVIDIA IndeX, Annotat3D



Marvin HPCC



Usuários



Nó de acesso
(login)



SLURM

Nós de processamento

gpu01marvin



GPUs NVIDIA
5 x A100 40GB
21 x MIG 5GB
256 CPUs
2TB RAM

c001marvin



256 CPUs
1TB RAM

Nós de armazenamento
(storage)



3 servidores dual node
Alta disponibilidade
300TB All Flash (SSD)
Proteção RAID 6
(8+2+spare)



O Marvin HPCC foi instalado em 2022 e custou USD350 mil. Inclui serviços de monitoramento 24x7 e reparo NBD.

Atividades

Roteiro

0. Preparação
1. Acessar e explorar
2. Manipular arquivos e diretórios
3. Enviar e baixar dados
4. Carregar programas
5. SLURM 1
6. SLURM 2
7. OpenOnDemand
8. Ajuda com LLMs



CNPq

Preparação

Objetivo

Preparar o ambiente para realizarmos as próximas atividades.

Ações

1. Abrir o Terminal (ou Prompt de Comandos)
2. Abrir um navegador:
 - marvindocs.cnpem.br
 - LLM de sua escolha (chatgpt, gemini, qwen, deepseek, ...)
 - marvin.cnpem.br (não tentem se conectar ainda)



Comandos



Acessar e explorar

Objetivo

Conectar ao cluster HPC (Marvin) e explorar o sistema de arquivos e diretórios.

Ações

1. Login no Marvin utilizando `ssh`
2. Ver o diretório de aterrisagem com `pwd`
3. Acessar o manual dos comandos com `man`
4. Listar os arquivos do diretório com `ls`
5. Mudar de diretório com `cd`
6. Ver o conteúdo de arquivos com `cat` e `less`

Comandos

```
ssh <username>@marvin.cnpem.br  
pwd  
man <comando>  
cd  
ls  
cat  
less
```

Dicas

- `Tab` para completar paths e comandos
- `↑ ↓` Para repetir comandos
- O `man` e o `less` tem atalhos como `/` [buscar], `n` [próxima ocorrência], `q` [sair]

Manipulando arquivos e diretórios

Objetivo

Criar, mover, copiar e apagar arquivos e diretórios.

Ações

1. Criar uma pasta "treinamento_hpc" e pastas de atividades com **mkdir**
2. Criar e editar um arquivo "hello.txt" dentro de uma pasta com o **nano**
3. Copiar, mover e renomear pastas e arquivos com **cp** e **mv**
4. Apagar arquivos e pastas com **rm** e **rm -r**
5. Criar atalhos com **ln -s**

Comandos

```
mkdir  
nano  
cp  
mv  
rm  
ln
```

Dicas

- Não use ESPAÇO em nomes, use **_**
- O nano é um editor básico, outras opções famosas são o vim e o emacs

Enviando e baixando dados

Objetivo

Mover arquivos e diretórios entre o computador local e o remoto (Marvin).

Ações

- Baixar arquivos da internet direto no Marvin com **wget**
- Descompactar e compactar pastas com **tar** e **zip**
- Copiar arquivos do computador local para o HPC e vice-versa com **scp**



Comandos

```
wget  
scp  
zip  
tar
```

```
# opcionais para cópia remota  
sftp  
rsync
```

Dicas

- Ctrl + Shift + V **cola** no terminal
- Ctrl + Shift + C **copia** do terminal

Carregando programas

Objetivo

Explorar e ativar módulos de programas disponíveis para todos os usuários.

Ações

1. Ver os módulos disponíveis.
2. Carregar um módulo.
3. Executar algum programa para ver o help.



Comandos

```
module avail
module spider
module help <programa>
module load <programa>
module unload

# opcionais
ml load conda
conda create ...
```

Recapitulando

- ✓ Login
- ✓ Manipulação local de arquivos e pastas
- ✓ Movimentação entre o computador local e o HPC
- ✓ Ativação de módulos



Usuários

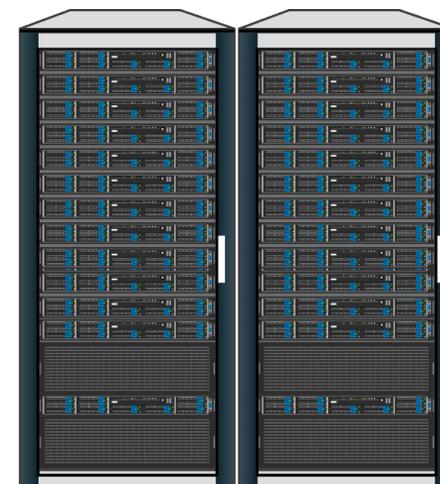


Nó de login



SLURM

Nós de processamento



CNPq

SLURM

Objetivo

Conhecer os comandos básicos do SLURM

Ações

1. Ver as partições (filas) existentes com **sinfo**
2. Ver as tarefas na fila com **squeue**
3. Submeter uma tarefa ao SLURM com **srun** e **salloc + srun**
4. Cancelar uma tarefa na fila com **scancel**

Comandos

```
sinfo  
squeue  
srun  
salloc + srun  
scancel
```

Dicas

- Ctrl + C cancela a execução de um programa

SLURM 2

Objetivo

Preparar scripts de tarefas.

Ações

1. Criar scripts com diferentes recursos e para filas diferentes



Comandos

```
sbatch
```

Atividade 7

OpenOnDemand

Objetivo

Explorar o sistema OpenOnDemand.

Ações

1. Abrir no navegador a página marvin.cnpem.br
2. Navegar pelas abas disponíveis



CNPem

Comandos



Ajuda de LLMs

Objetivo

Usar o "contexto" disponível em marvin.cnpem.br para obter ajuda.

Ações

1. Baixar o arquivo **llms.txt** do site marvindocs.cnpem.br
2. Subir o arquivo para a LLM de sua preferência
3. Dicas de conversa:
 - Crie um header de um job para rodar o <programa>
 - Quero copiar uma pasta de um projeto que está no HPC para o meu computador Windows. Quero salvá-la comprimida.

Comandos



Revisão

Roteiro

0. Preparação
1. Acessar e explorar
2. Manipular arquivos e diretórios
3. Enviar e baixar dados
4. Carregar programas
5. SLURM 1
6. SLURM 2
7. OpenOnDemand
8. Ajuda com LLMs



CNPq

Obrigado

Equipe de Dados Biológicos - DFS/LNBio



Sign up to receive
newsletters
about **CNPEM**
and its units

cnpem.br



CNPEM
Brazilian Center for Research
in Energy and Materials

MINISTRY OF
SCIENCE TECHNOLOGY
AND INNOVATION

