CENTRO BRASILEIRO DE PESQUISAS FÍSICAS – CBPF
COTEC · LAB-IA

# ATLAS-CBPF: INFRAESTRUTURA DE HPC PARA IA

Paulo José Russano

Versão 09/11/2026

# ROTEIRO

A. **Introdução** — A doação e o sistema

B. **Rede** — InfiniBand fat-tree, gerência e WAN

C. **Lustre + Alta Disponibilidade** — O filesystem paralelo /atlas

D. **Futuro** — Próximos passos e escala

# A Doação: Petrobras → CBPF

**14/11/2025**
Regulamento Petrobras de doação de HPCs

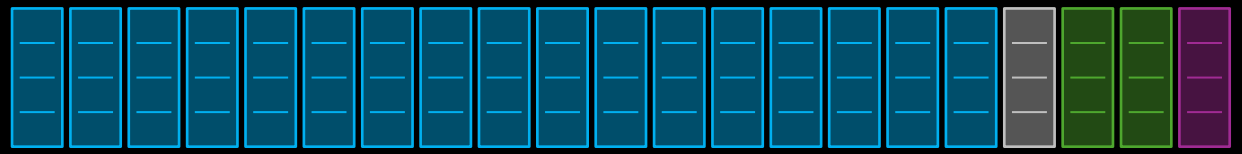
→

**12/12/2025**
CBPF habilitado (inscrição 14147/2025)

→

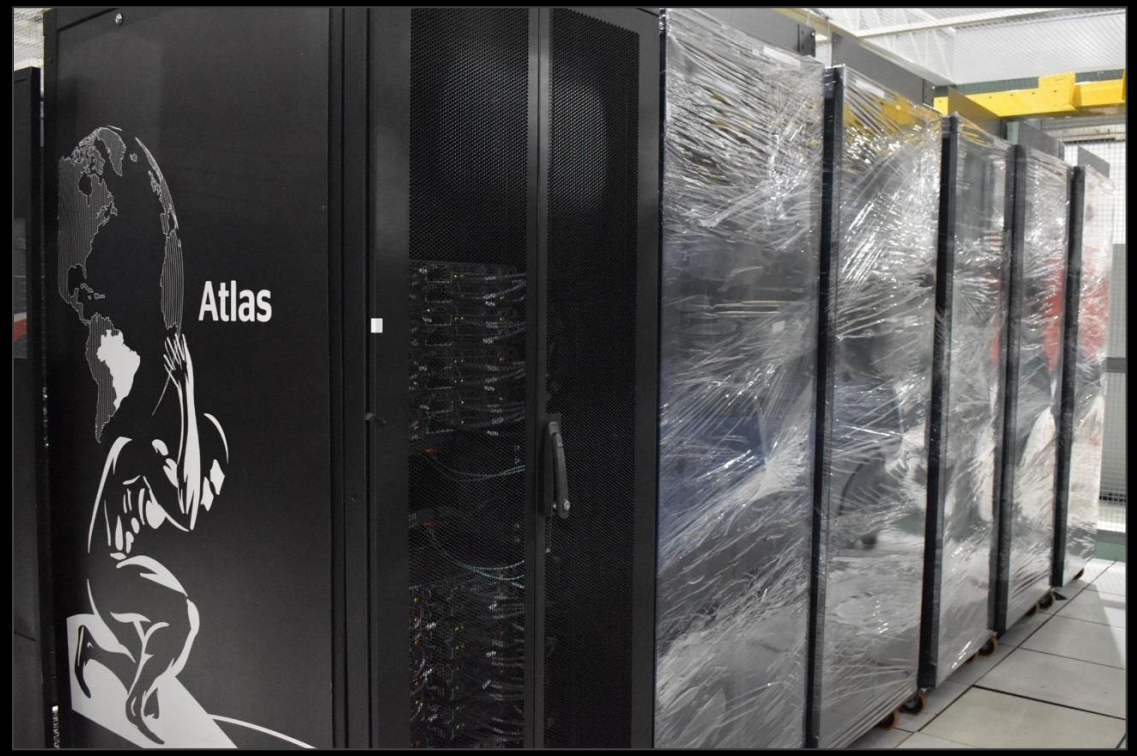
**30/12/2025**
Termo de Doação assinado

**Objeto do termo: cluster HPC de 21 bastidores**



17 processamento · 1 controle · 2 FS paralelo · 1 switches IB EDR

| 8,9 PFlops | 610 kVA | 1,3 PB |
|---|---|---|
| capacidade Rpeak | consumo nominal | FS paralelo (nominal) |




**Finalidade e obrigações**

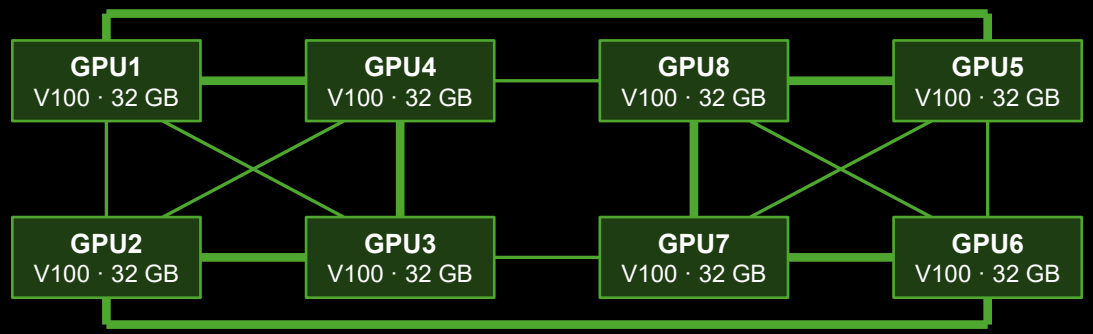
IA e Deep Learning para astronomia, física de altas energias e ciência de dados.

Retirada em até 60 dias no CPTI (Vargem Grande, RJ) · permanência mínima de 2 anos no acervo.

**136** nós × **8** GPUs V100 = **1.088** GPUs

# O Sistema: 136 Nós, 1.088 GPUs, 1,2 PB

## Nó GPU — Supermicro 4029GP-TVRT (4U)




**× 136** 8 nós por rack · 17 racks de processamento

**Processamento**
17 racks × 8 nós = 136 nós · 1.088 GPUs V100 32 GB

**Serviço — rack E01**
master (gateway/NAT) · login ×2 · portal · DTN · Slurm ×2 · deploy ×2 · VDI · desenvolvimento

**Storage Lustre — racks E02/E03**
MGS ×2 · MDS ×2 · OSS ×12 · 12 gavetas OST (88 HDD 1,8 TB) · 1 gaveta MDT (24 SSD 1,6 TB)

**Interconexão**
InfiniBand EDR 100 Gb/s: 4 spines SB7800 + 8 leafs SB7890

| 8,9 PFlops | 1,2 PB | 100 Gb/s |
|---|---|---|
| Rpeak (termo de doação) | úteis no Lustre /atlas | por porta InfiniBand EDR |

# Layout Físico: 20 Racks em Duas Fileiras




*gr10bNN = nome gravado no HCA (herdado da instalação original); não segue a ordem física. Diagrama esquemático.*

# ROTEIRO

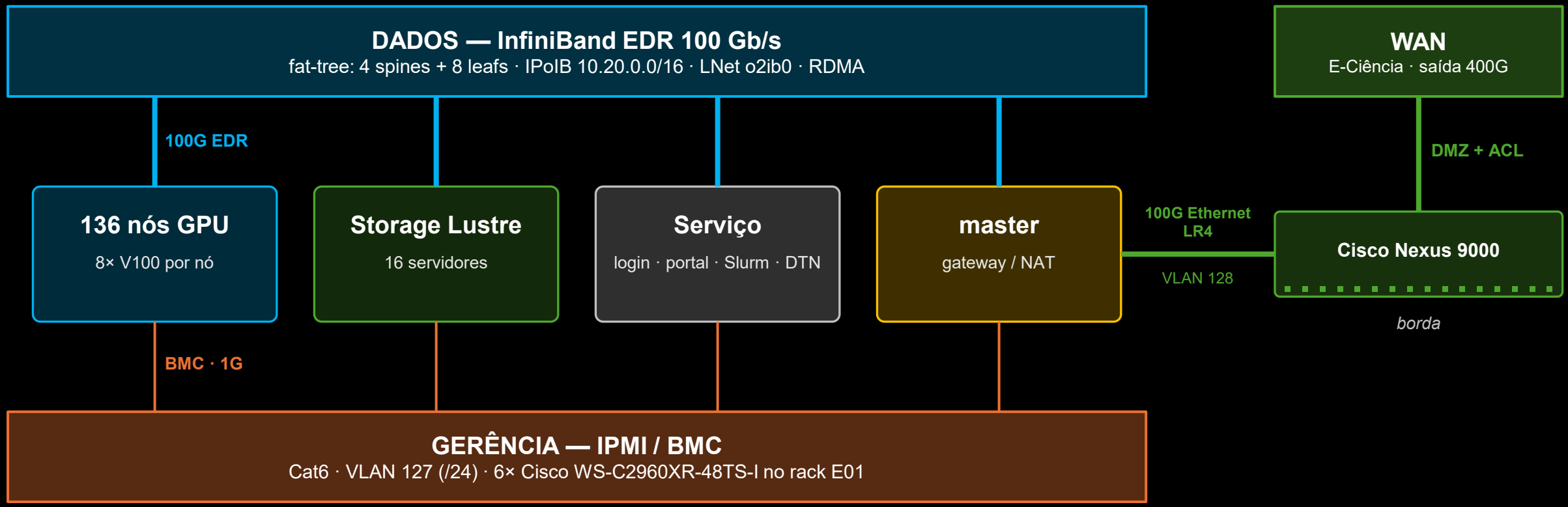
A. **Introdução** — A doação e o sistema

B. **Rede** — InfiniBand fat-tree, gerência e WAN

C. **Lustre + Alta Disponibilidade** — O filesystem paralelo /atlas

D. **Futuro** — Próximos passos e escala

# TRÊS REDES, TRÊS FUNÇÕES

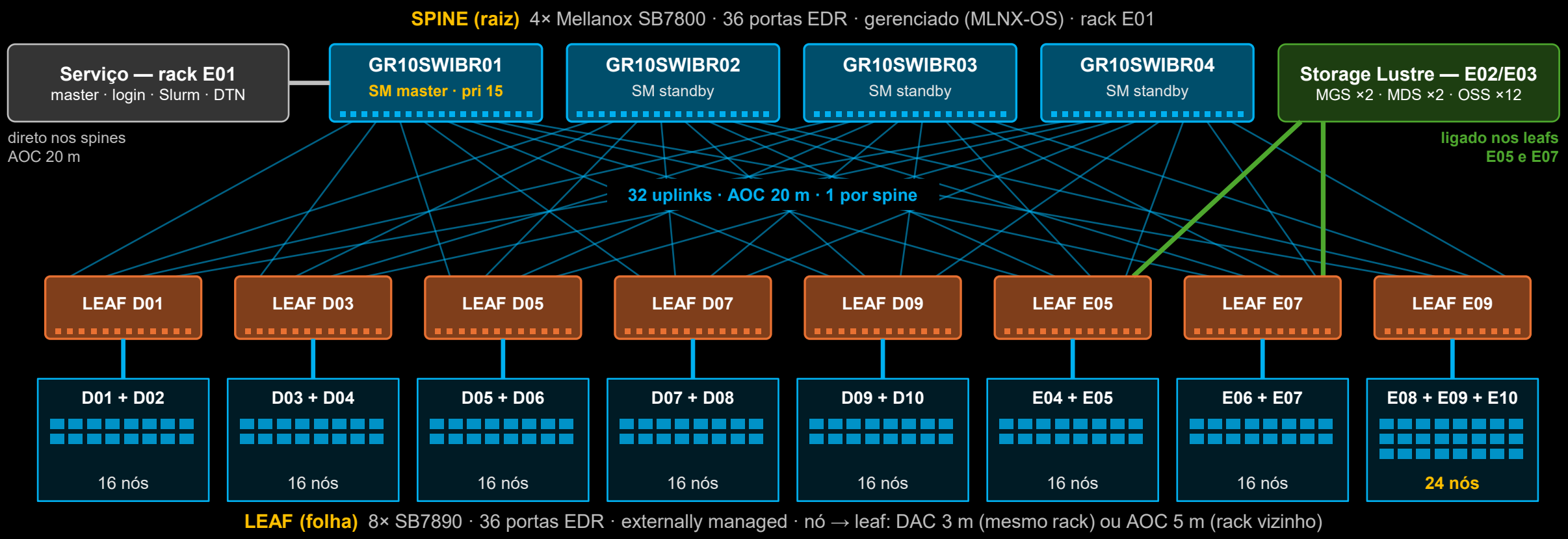



**Sem Ethernet de dados**
Escalonador, deploy, Lustre e RDMA trafegam na InfiniBand.

**Gerência isolada**
BMC de todos os equipamentos em VLAN própria, fora da malha de dados.

**Saída dos nós internos**
NAT no master pela InfiniBand; só a borda fala com a WAN.

# INFINIBAND EDR: FAT-TREE DE 2 NÍVEIS

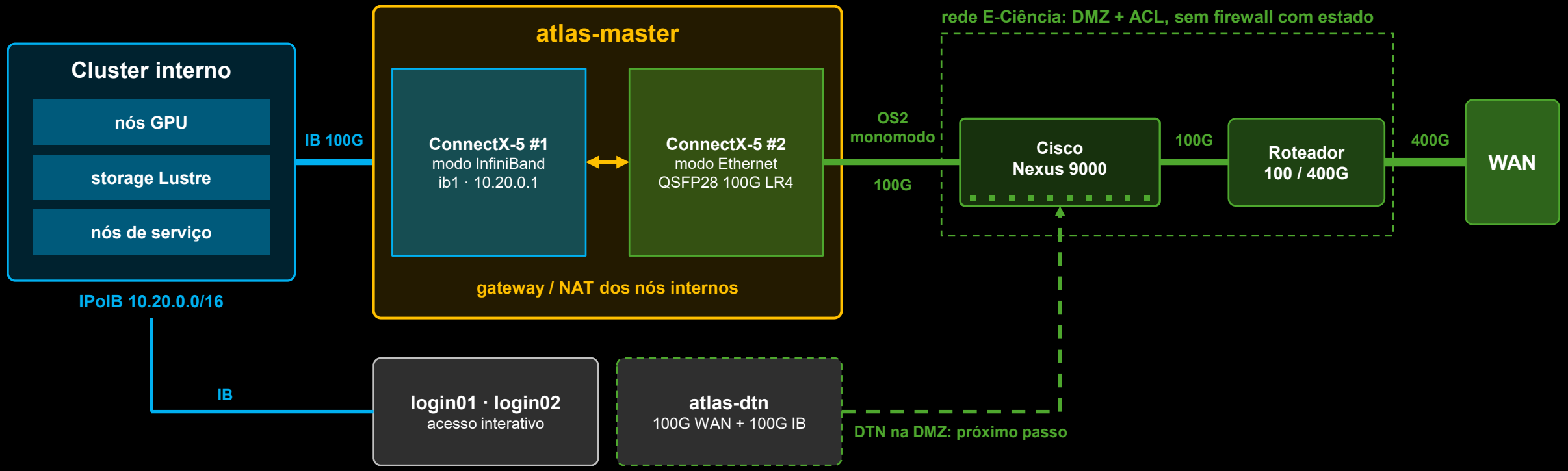



**Fat-tree de 2 níveis**
Cada leaf sobe com 4 uplinks, um por spine: no máximo 3 switches entre dois nós.

**Projeto × execução**
Projeto: 2 portas por nó. Executado: 1 porta (single-rail) — faltou cabo DAC.

**Malha preservada**
Switches não zerados: Subnet Manager e roteamento originais mantidos.

# BORDA: WAN DE 100G E SCIENCE DMZ




**Science DMZ**
Caminho de dado só com DMZ + ACL. Teto atual 100G; a porta de 400G já existe no roteador.

**Papéis separados**
master: gateway/NAT · login: acesso interativo · DTN: dados em massa para o Lustre.

**Lição da fibra**
Flap = óptica monomodo em fibra multimodo. 1G: LX → SX (−13 → −5 dBm). 100G: fibra OS2 nova.

# ROTEIRO

A. **Introdução** — A doação e o sistema

B. **Rede** — InfiniBand fat-tree, gerência e WAN

C. **Lustre + Alta Disponibilidade** — O filesystem paralelo /atlas

D. **Futuro** — Próximos passos e escala

# Por Que Lustre: Dividir para Paralelizar

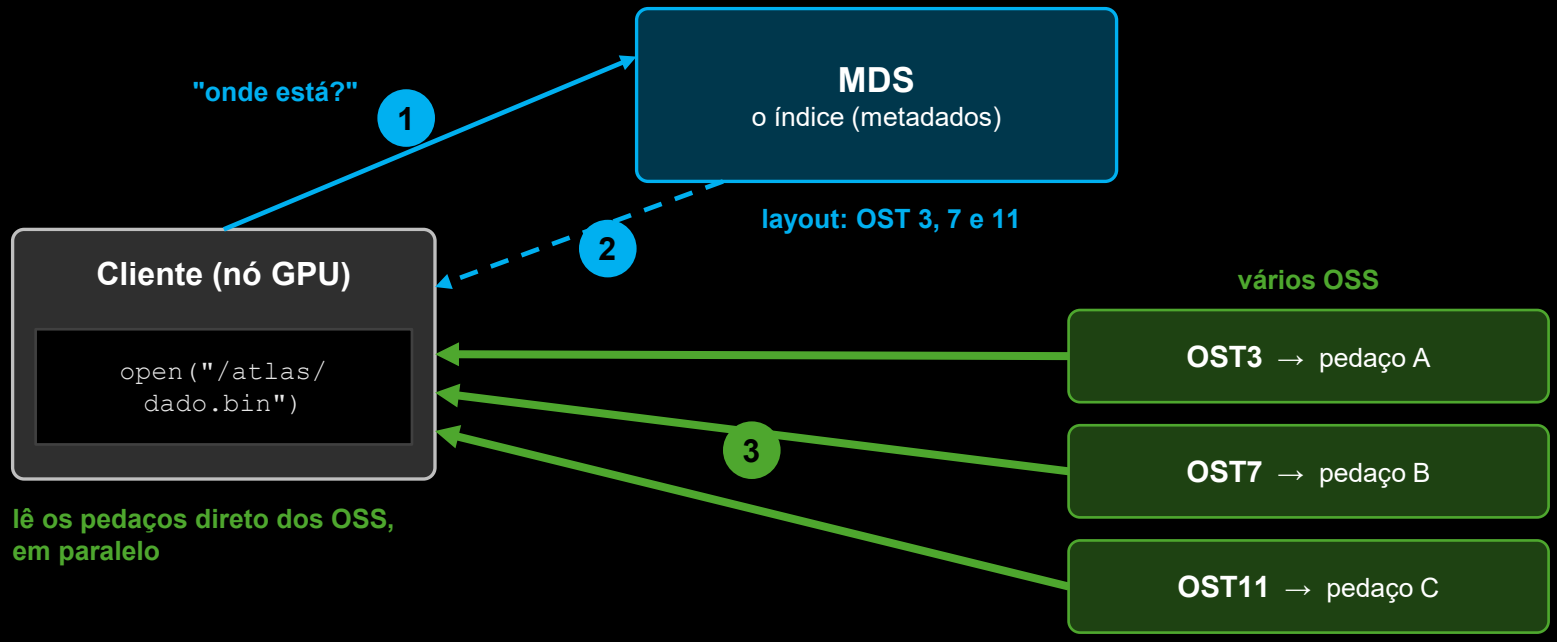
**MGS ×2**
A lista telefônica: quais servidores e alvos existem. mgs1 ativo, mgs2 com a réplica.

**MDS + MDT**
O índice: nomes, permissões e o layout de cada arquivo. Em SSD.

**OSS + OST**
As estantes: guardam e servem o conteúdo, em paralelo.

**Cliente**
Monta o /atlas como disco local: nós GPU, login e serviço.

## O caminho de uma leitura



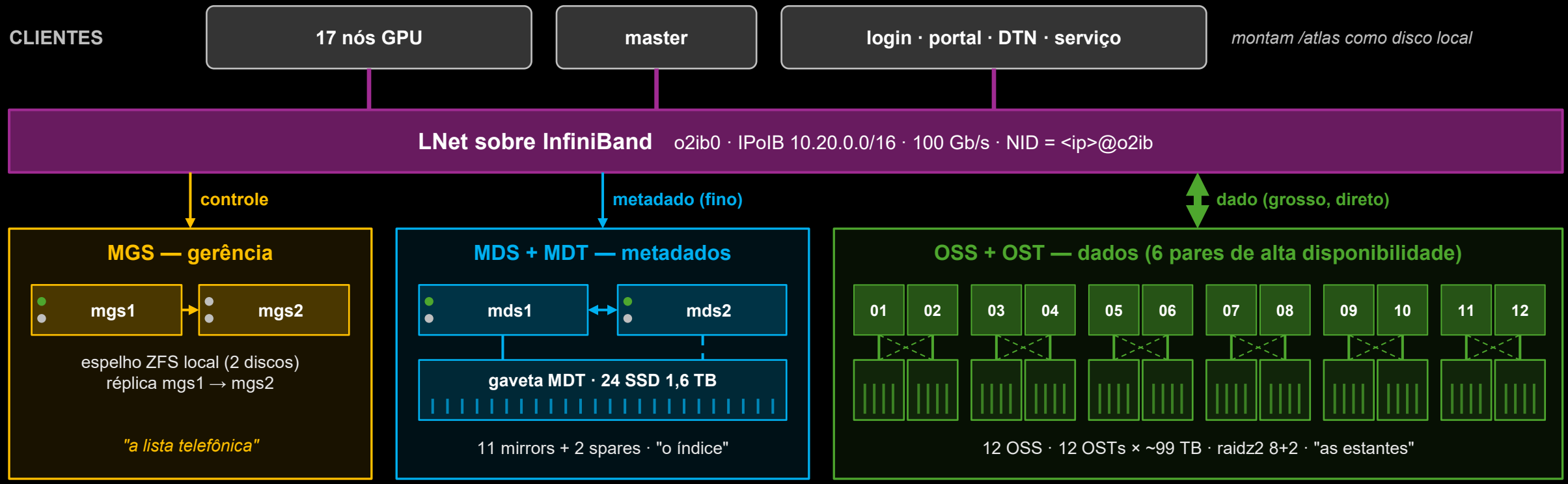

O MDS só participa de **1–2**. O dado pesado é cliente ↔ OSS e nunca passa pelo índice.

**Servidor único: um funil**
N clientes dividem a banda e os discos de uma só caixa.

**Lustre: a banda soma**
Arquivo fatiado (striping) entre os OSTs: 12 OSS entregam juntos.

**No Atlas**
1,2 PB sob /atlas, idêntico em todos os nós que o montam.

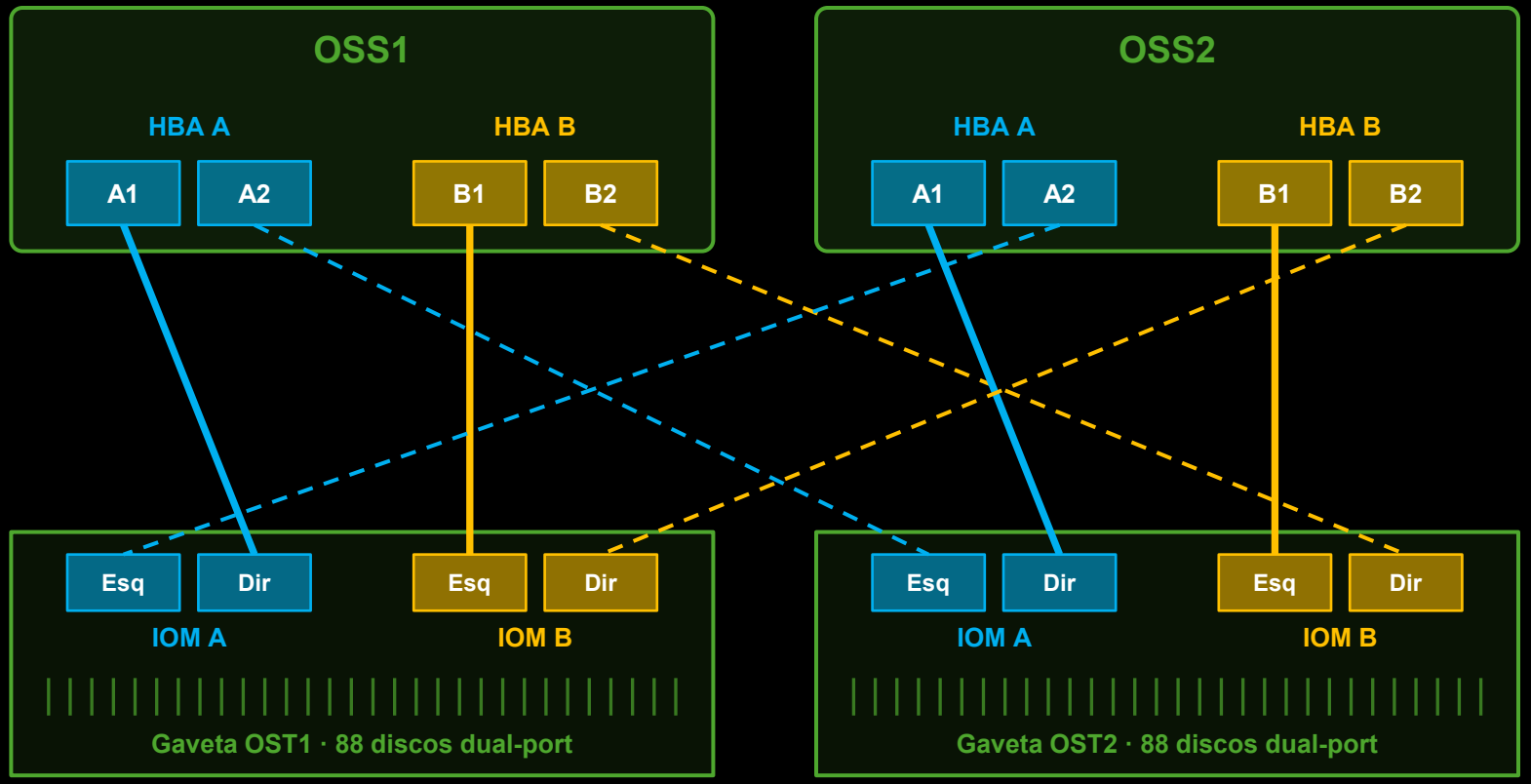
# Arquitetura do /atlas




**Gavetas de disco (JBOD SAS)** sem rede InfiniBand · ligadas aos servidores por cabo SAS, por dois caminhos: a base do HA

| 16 | 12 × 98,8 T | 15,7 T | 1,2 PB |
|---|---|---|---|
| servidores Lustre | OSTs (lfs df) | MDT em SSD (lfs df) | úteis em /atlas |

# SAS Cruzado: a Base Física do HA




### A regra, igual nos 6 pares

HBA A → IOM A das duas gavetas; HBA B → IOM B das duas gavetas.

Porta 1 → gaveta própria; porta 2 → gaveta do par.

### Resultado

Cada disco com 2 caminhos, um por HBA, em IOMs diferentes. Sobrevive à queda de uma HBA, uma IOM ou um OSS.

**176**
discos por par

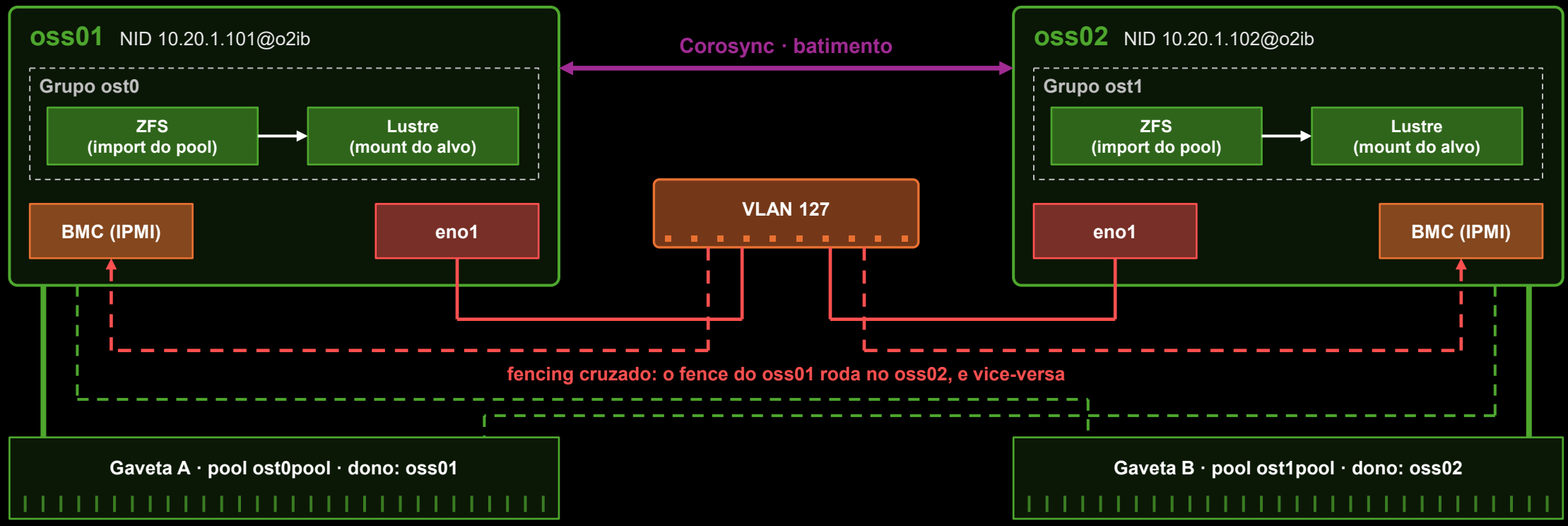
**352**
caminhos SAS por par

Conferência após cabear cada OSS

```
multipath -ll | grep -c '^mpath'
```

→ 176, cada disco com 2 caminhos

# Um Par de Alta Disponibilidade por Dentro

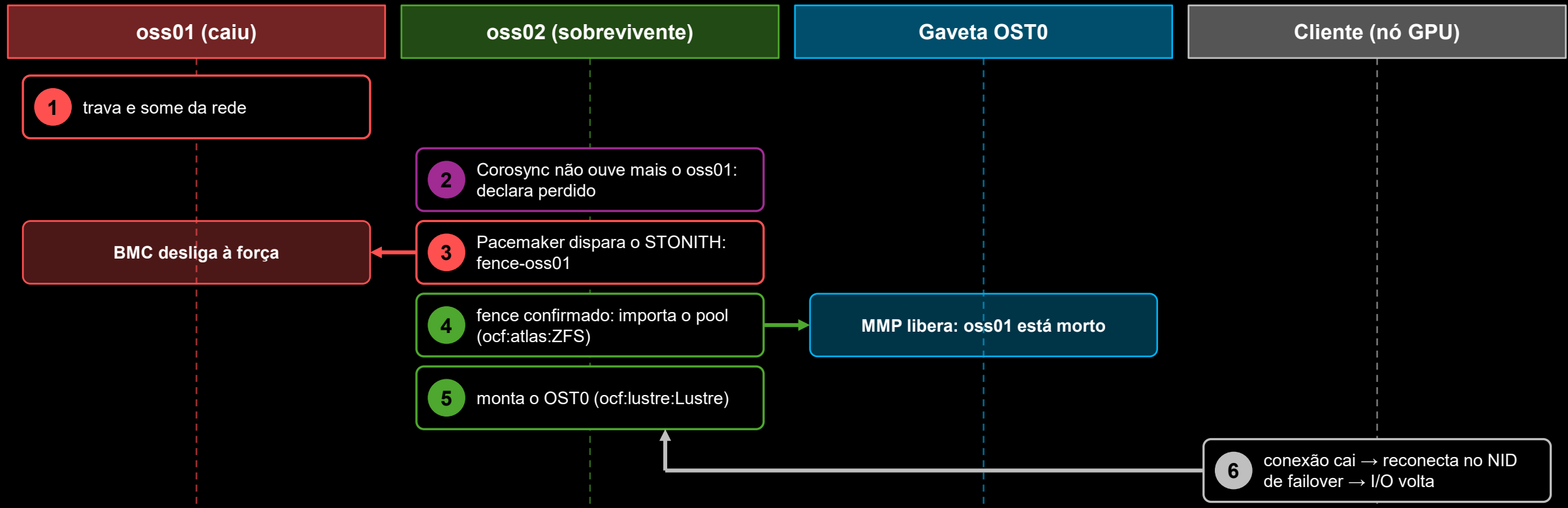



**Operação normal**
Cada OSS serve 1 OST; o parceiro é o caminho de failover.

**Fencing direto**
NIC eno1 confinada: sem gateway, zona drop. Sem NAT.

**MMP (multihost=on)**
O pool recusa ser importado pelos dois ao mesmo tempo.

# Um Failover, Passo a Passo




**Fence antes de importar**
A ordem 3 → 4 é o que impede o split-brain; o MMP é a reserva.

**Visto pelo cliente**
A aplicação espera e segue: o segundo NID já era conhecido.

**Quando o oss01 volta**
Stickiness 200: o OST fica no oss02 até o failback manual.

# LNet: Todo Serviço Tem Dois Endereços




**Roda no cliente**

```
mount -t lustre 10.20.1.1@o2ib:10.20.1.2@o2ib:/atlas /atlas
```

**Roda no nó dono do alvo, com o alvo desmontado**

```
tunefs.lustre --servicenode=10.20.1.111@o2ib --servicenode=10.20.1.112@o2ib ost10pool/ost10
```

### Efeito no failover

Se o dono do alvo some, o cliente já conhece o segundo NID e reconecta sozinho, sem intervenção.

### LNet no boot

O lnet.service padrão só sobe lo: um drop-in de systemd sobe o o2ib0 antes dos mounts.

# Validação: Capacidade, Discos e Testes

## Resultado: /atlas montado com 1,2 P úteis

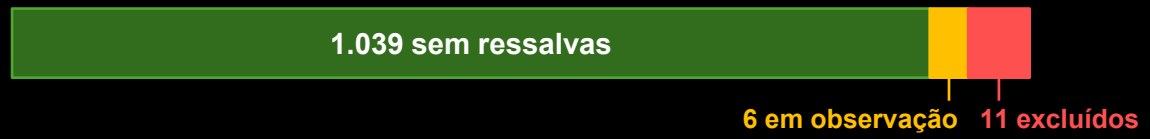
```
/dev/md123                                1023M  6.0M 1017M   1% /boot/efi
tmpfs                                       38G     0   38G   0% /run/user/1000
10.20.1.1@o2ib:10.20.1.2@o2ib:/atlas  1.2P  216M  1.2P   1% /atlas
```

**1,2 P**
216 M usados · 1%

*df -h no master: /atlas montado pelos dois NIDs do MGS.*

| | | |
|---|---|---|
| **12 × 98,8 T**<br>OSTs + MDT de 15,7 T (lfs df) | **1.056**<br>HDD SAS 1,8 TB: ≈ 1,9 PB brutos | **352**<br>caminhos SAS por par de OSS |
| **0**<br>erros não corrigidos (SMART) | **11 + 6**<br>discos excluídos + em observação | **17 / 17**<br>nós GPU montando o /atlas |

## Saúde dos 1.056 discos (laudo SMART)




Pools criados sem os 11 excluídos, sem esperar a reposição; os 20 de reposição reforçam os spares.

## Testes executados

- **Integridade** Escrita de 2 GB com checksum conferido após drop de cache.
- **Paralelismo** Striping nos 12 OSTs; os 12 registrados no MDT.
- **Failover do MDT** Fence real do mds1: o MDT migrou sozinho.
- **Adoção pelo Pacemaker** MDT e 6 pares OSS virados sem downtime.
- **Promoção do MGS** Provada ao vivo 2×; o FS seguiu em leitura e escrita.
- **Liga e desliga** Sequência completa de down/up por playbook.

# ROTEIRO

A. **Introdução** — A doação e o sistema

B. **Rede** — InfiniBand fat-tree, gerência e WAN

C. **Lustre + Alta Disponibilidade** — O filesystem paralelo /atlas

D. **Futuro** — Próximos passos e escala

# Onde Estamos e Próximos Passos

## Em produção

- Malha InfiniBand 100G em fat-tree
- Gerência na VLAN 127, sem credencial de fábrica
- WAN de 100G ativa
- Lustre /atlas de 1,2 PB em HA
- 12 nós de serviço instalados
- 17 nós GPU montando o /atlas

## Em implantação

- Slurm: fila única de GPU
- 6 QoS derivadas dos grupos LDAP
- Login pelo LDAP institucional (SSSD)
- Homes em /atlas/home
- Failover manual do controlador Slurm

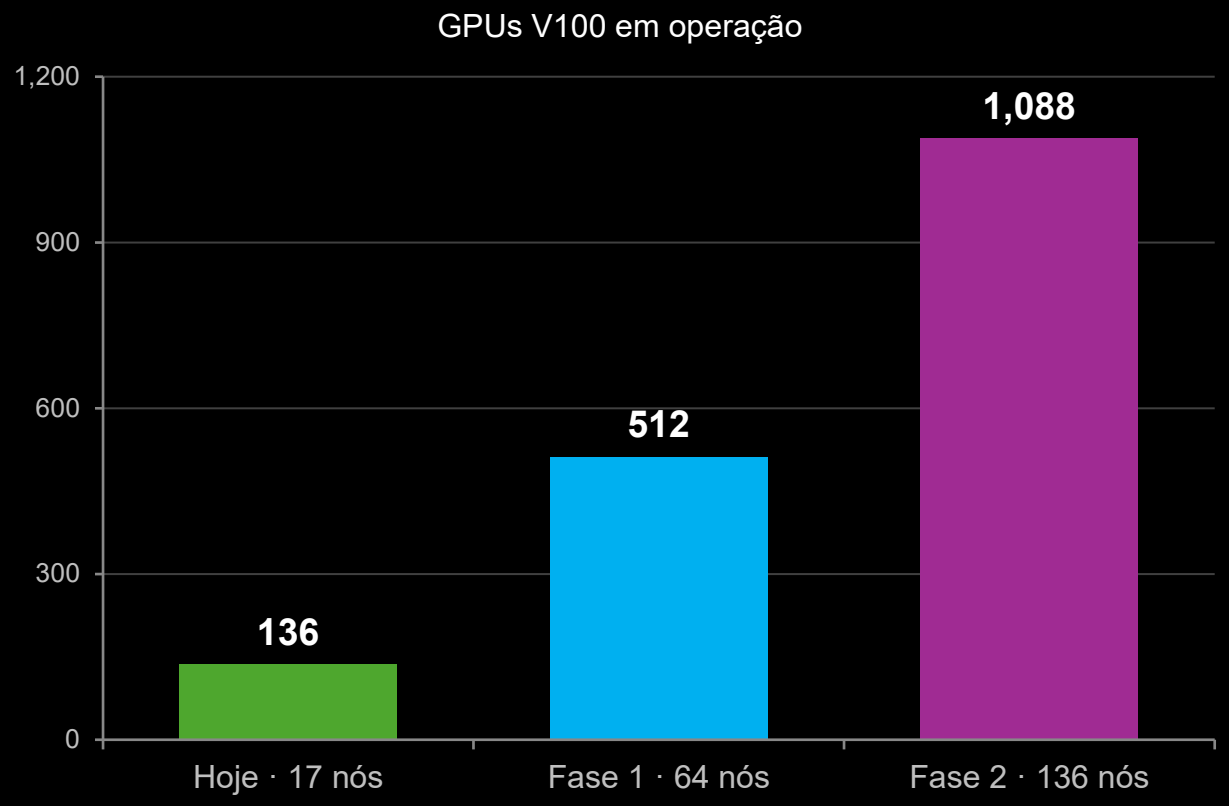
## Próximos passos

- Portal Open OnDemand com SSO
- DTN na Science DMZ (100G)
- Home com DoM e quota; scratch com purga
- Backup do home + snapshots do Lustre
- IOR + mdtest com vários clientes
- Firmware dos BMCs e monitoramento

## Escala

- 2ª porta InfiniBand por nó (dual-rail)
- Fase 1: elétrica → 512 GPUs
- Fase 2: UPS e climatização → 1.088 GPUs
- WAN a 400G
- Zona Atlas: infraestrutura dedicada no RSDC

# ESCALA: DE 17 A 136 NÓS




**610 kVA**
consumo nominal do cluster completo

**12,5%**
dos nós em operação hoje (17 de 136)

**Hoje** **17 nós · 136 GPUs**
Um nó ligado por rack de processamento. Rede, storage e serviço completos.

**Fase 1** **64 nós · 512 GPUs**
Infraestrutura elétrica do Atlas na obra elétrica do CBPF (FINEP 1182/25, RELE-CBPF), com a climatização do data center.

**Fase 2** **136 nós · 1.088 GPUs**
Atlas completo, com UPS e climatização dedicados. Recursos em captação.

**Em paralelo**
Dual-rail: 2ª porta InfiniBand por nó, que depende de cabos DAC.

# ZONA ATLAS: INFRAESTRUTURA NO RSDC




## Zona Atlas no RSDC

Área dedicada ao Atlas e a futuros supercomputadores desse porte, com infraestrutura separada do restante do data center.

## Sistemas do modelo

- Corredor frio confinado entre as fileiras D e E
- Energia crítica pela UPS: racks críticos
- Energia de lote por rede + gerador: racks de lote
- VRF: evaporadoras, condensadoras e linhas frigoríficas
- Insuflamento por dutos e difusores; retorno pelo forro
- Eletrocalhas de energia e encaminhamento óptico

# Considerações Finais

- **Da doação à operação:** em menos de um ano, o Atlas foi recebido, instalado e colocado em serviço, com a pilha de software montada do zero no CBPF.
- **Rede:** malha InfiniBand em fat-tree preservada e documentada; gerência e borda separadas da malha de dados.
- **Storage:** Lustre de 1,2 PB em alta disponibilidade: metadados e dados sobrevivem à queda de um servidor.
- **Método:** infraestrutura como código, com procedimentos de failover e de liga/desliga testados.
- **Futuro:** abrir o Atlas aos grupos de pesquisa com escalonador e portal e escalar com infraestrutura dedicada no RSDC.

| < 1 ano | 1,2 PB | 17 → 136 |
|---|---|---|
| do termo de doação ao storage em produção | em alta disponibilidade no /atlas | nós GPU: a escala depende de energia |

CENTRO BRASILEIRO DE PESQUISAS FÍSICAS – CBPF
COTEC · LAB-IA

**Origem:** doação Petrobras — Termo de 30/12/2025
**Sistema:** 136 nós · 1.088 GPUs V100 · InfiniBand EDR 100 Gb/s
**Storage:** Lustre /atlas — 1,2 PB em alta disponibilidade

# ATLAS-CBPF: INFRAESTRUTURA DE HPC PARA IA

Paulo José Russano

Versão 09/11/2026