

WG 4

Computação e IA

Alexandre Suaide

maio/2024

Upgrade do cluster - storage

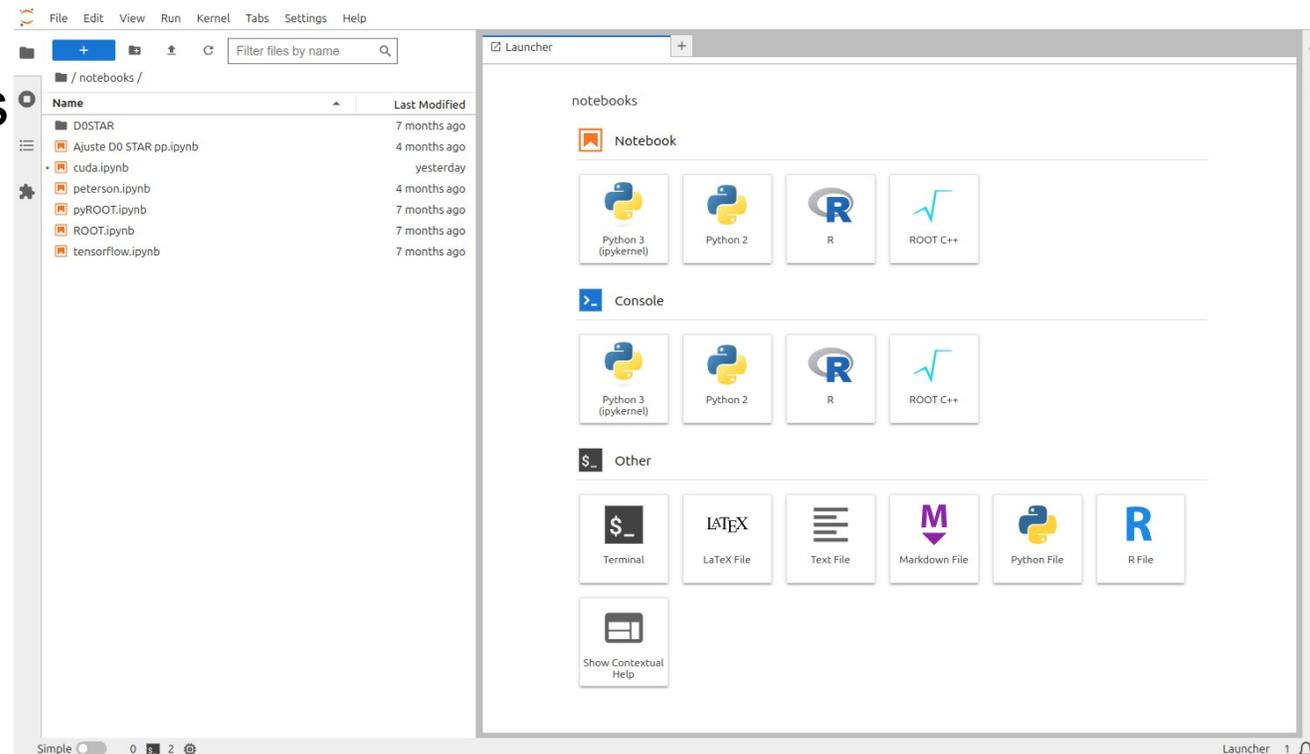
- Compra de 1.4 PB de STORAGE em 2023
- Espaço já disponível
 - Alguns ajustes necessários para disponibilizar para o GRID



| SaoPaulo | | | | | | | | | | | | |
|------------------------|---------------------------|------|----------------------|-----------------|----------------|--------|------------------|------|-------------------------|---------------|-----------------|--------|
| AliEn SE | | | Catalogue statistics | | | | | | Storage-provided inform | | | |
| SE Name | AliEn name | Tier | Size | Used | Free | Usage | No. of files | Type | Size | Used | Free | Usage |
| 1. SaoPaulo - EOS | ALICE::SaoPaulo::EOS | 2 | 892 TB | 289.2 TB | 602.8 TB | 32.42% | 5,208,363 | FILE | 892.4 TB | 665.9 TB | 226.5 TB | 74.62% |
| 2. SaoPaulo - EOS_TEST | ALICE::SaoPaulo::EOS_TEST | 2 | 892 TB | 10 B | 892 TB | - | 1 | FILE | 881.4 TB | 60.56 GB | 881.4 TB | 0.007% |
| Total | | | 1.742 PB | 289.2 TB | 1.46 PB | | 5,208,364 | | 1.732 PB | 666 TB | 1.082 PB | |

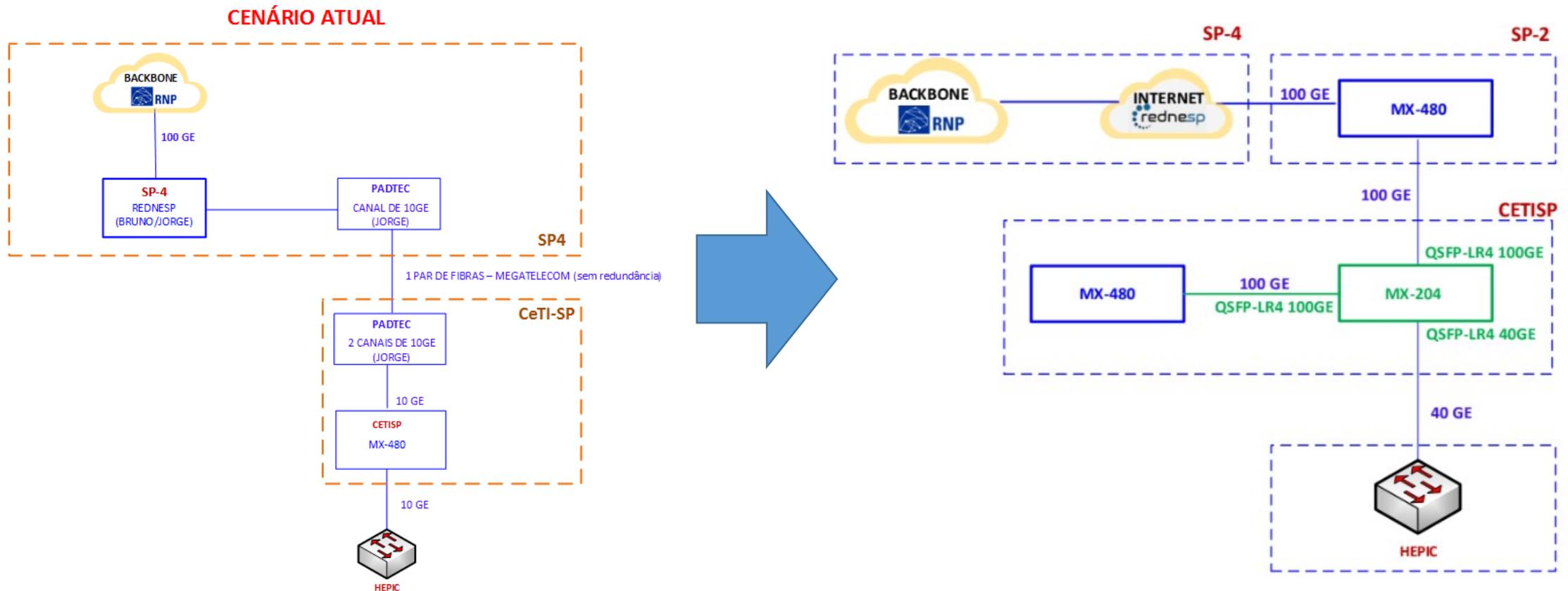
Upgrade do cluster - JARVIS

- <http://jarvis.if.usp.br:8000>
 - Notebook com kernels diversos
 - Acesso também via ssh
- Pequeno cluster com GPUs
 - 4 x NVIDIA RTX 3060
 - CUDA cores: 1792
 - Threads: 43008
 - CPU clock: 1777 MHz
 - Memória: 12 GB
- Basta ter conta no SAMPA



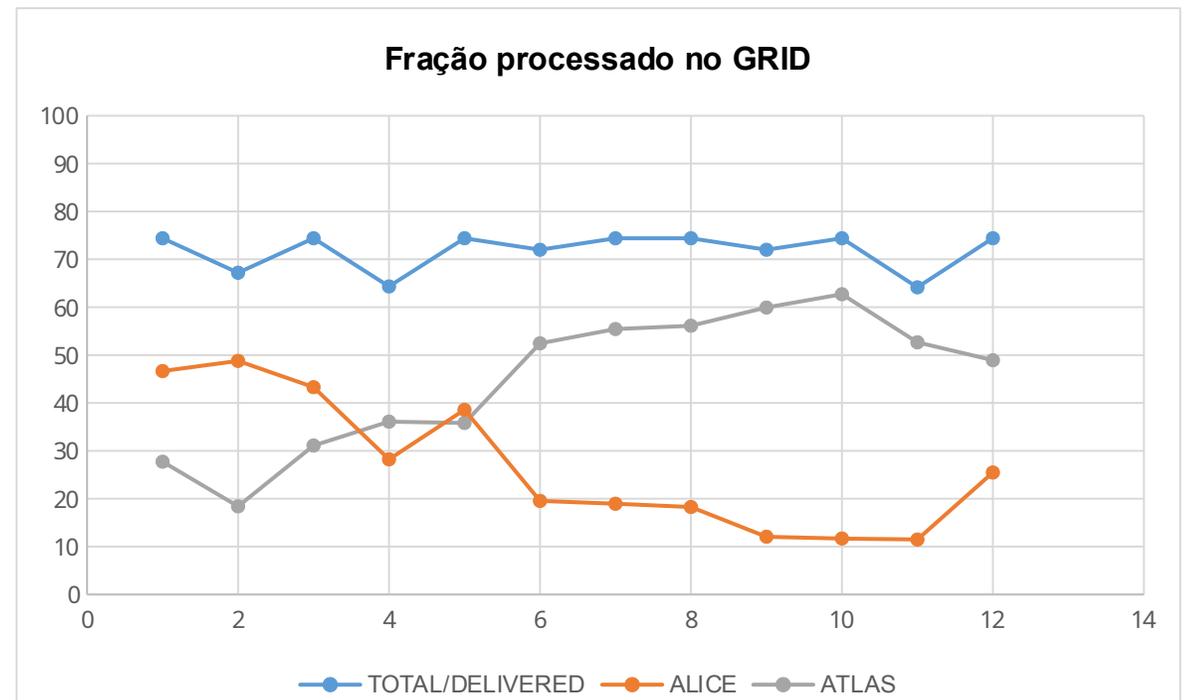
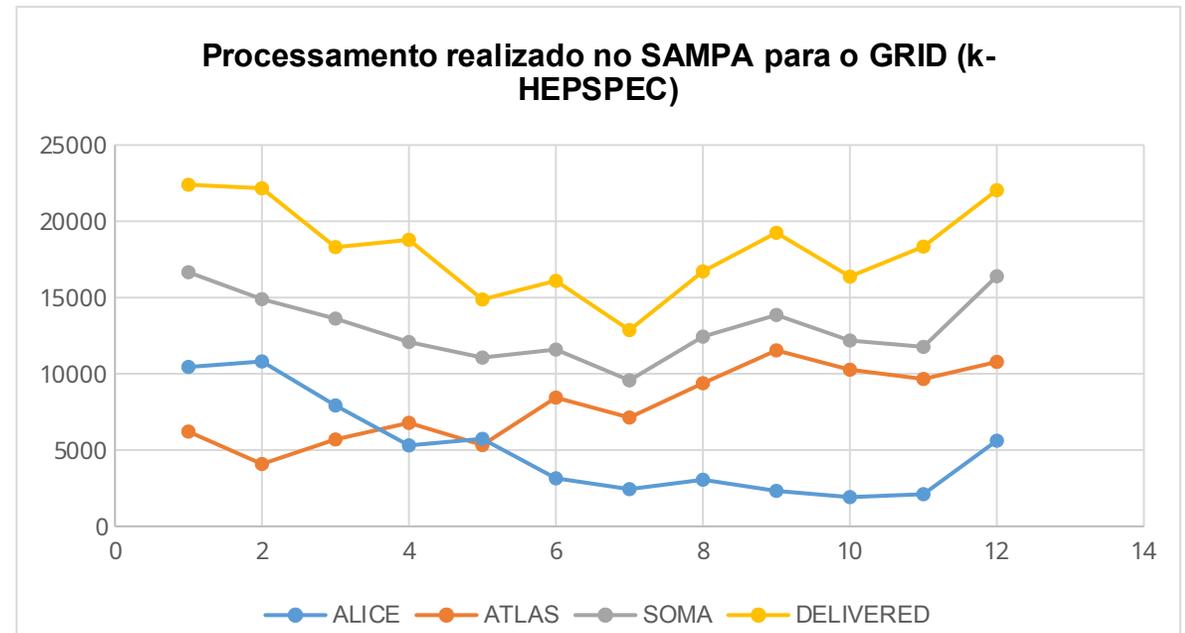
Upgrade do cluster - conectividade

- Ativação do link de 40 Gbps dependendo de alguns detalhes
 - Toda infraestrutura aparentemente já pronta



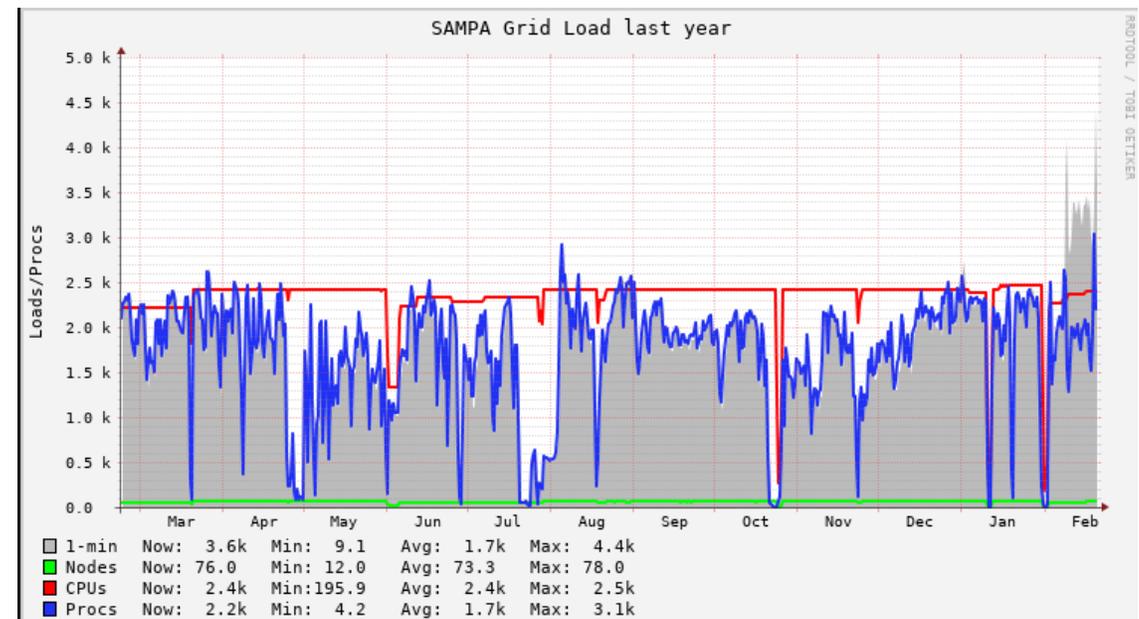
Processamento de dados

- Balanço entre processamento do ATLAS e do ALICE foram alterando ao longo do ano



Processamento de dados

- Aproximadamente 70% de carga ao longo do último ano
- Contudo, alta taxa de erro de jobs do ALICE
 - Após ajustes, aproximadamente 50%
 - Memória/disco



Select site:

MonALISA information Version: 22.11.05 (JDK 12.0.1)
Running on: sampacvo.if.usp.br
Administrator: <Douglas Vieira <dvieira@gmail.com> sampa-admin <sampa-admin@if.usp.br>>

Service health NTP: SYNC, offset: 0.011s

Services status ClusterMonitor: n/a Proxies status AllEn proxy: OK (1 day, 23:23)
AllEn: PackMan: n/a Delegated proxy: n/a (n/a)
CE: n/a Proxy server: OK (57 days, 06:26)
CE info: Proxy of the machine: OK (18:08)
Max running jobs: 320
Max queued jobs: 10

Current jobs status Assigned: 0 Accounting Success jobs: 390 (profile) Site averages Active nodes: 1773
Running: 155 Error jobs: 10788 + 43 expired (last 24h) Average KSI2k/core: 0
Saving: 44 kSI2k units: 0 / pledged

Storages status

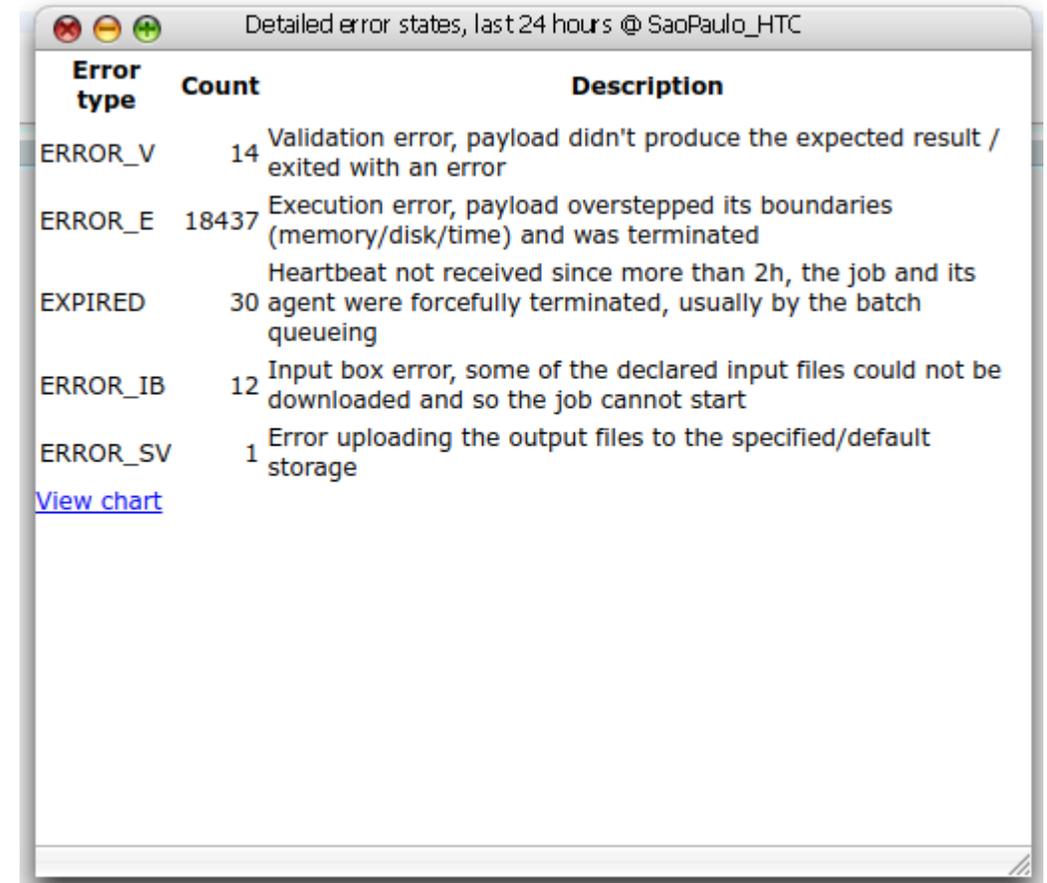
| Name | Status | Size | Used | Free | Usage | No of files | Type | ADD test |
|------|--------|------|------|------|-------|-------------|------|----------|
|------|--------|------|------|------|-------|-------------|------|----------|

VoBox health CPUs: x MHz CPU usage (last 1h avg) Load: User: 1.959% System: 0.19% Int: 0%
Mem usage: % of IOWait: 0% Soft int: 0.006%
Processes: 220 Idle: 97.84% Nice: 0%
Sockets: 71 TCP / 28 UDP Uptime: Steal: 0.001%

| AllEn LDAP var | VoBox path | Size | Used | Free | Use% |
|----------------|------------|------|------|------|------|
| TMP | | n/a | n/a | n/a | n/a |
| LOG | | n/a | n/a | n/a | n/a |
| CACHE | | n/a | n/a | n/a | n/a |

Processamento de dados

- Muitos erros por conta de falta de recursos em disco/memória/tempo
 - Vamos discutir isso mais a frente. Necessidade de ação urgente



| Error type | Count | Description |
|------------|-------|--|
| ERROR_V | 14 | Validation error, payload didn't produce the expected result / exited with an error |
| ERROR_E | 18437 | Execution error, payload overstepped its boundaries (memory/disk/time) and was terminated |
| EXPIRED | 30 | Heartbeat not received since more than 2h, the job and its agent were forcefully terminated, usually by the batch queueing |
| ERROR_IB | 12 | Input box error, some of the declared input files could not be downloaded and so the job cannot start |
| ERROR_SV | 1 | Error uploading the output files to the specified/default storage |

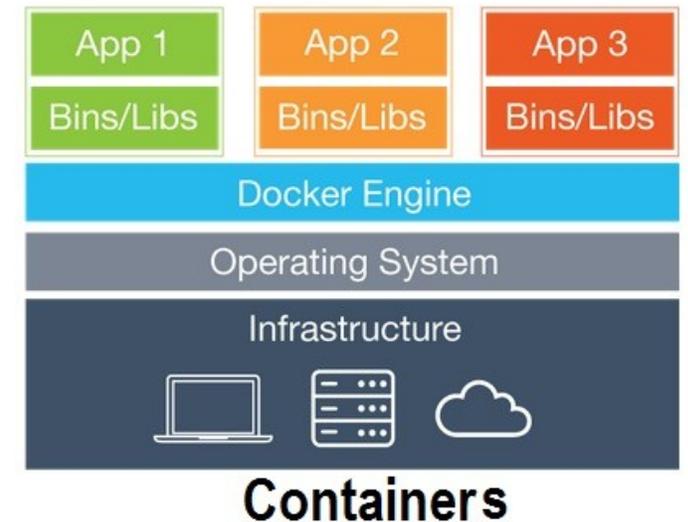
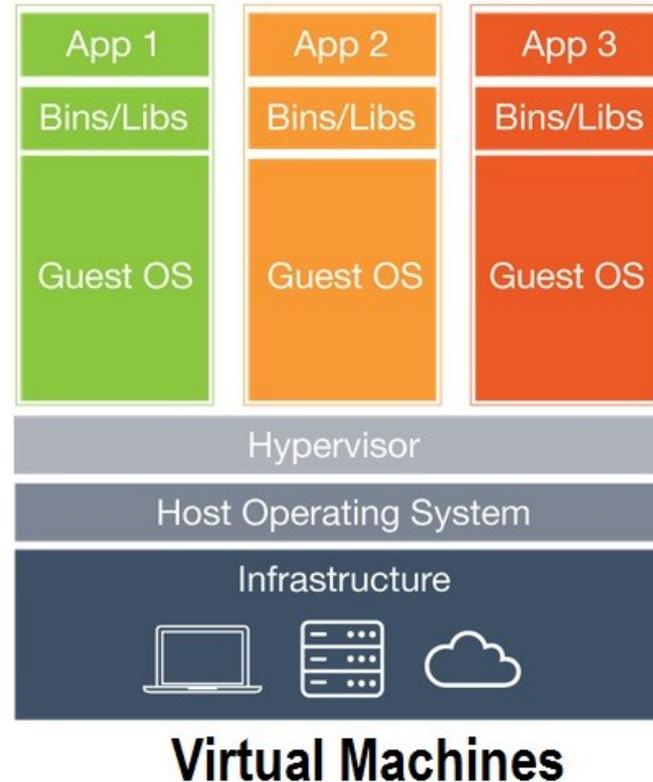
[View chart](#)

Upgrade de hardware urgente

- Muitos desses jobs não rodam por falta de espaço temporário em disco local no WN e memória RAM/core
 - Hoje: 600 GB/WN de espaço no WN (bem menos disponível para job)
 - Quase não é suficiente para 1 Job do ATLAS ou ALICE
 - Hoje: máquinas que variam entre 2 GB/core até 4 GB/core
- Ação:
 - Atualização para 2 TB/WN (2 x SSD 1TB)
 - Custo
 - Discos de prateleira: R\$ 1000,00/disco x 2 x 48 WN ~ R\$ 100k
 - Solução já testada
 - Memória RAM → 8 GB/Core? (ainda a ser verificada a viabilidade técnica)
 - Aprox. 2400 cores x 4GB = 19200 GB
 - Custo de RAM aprox R\$ 1000,00 /64 GB ~ R\$ 300k

Atualizações e upgrades

- Atualização dos sistemas do cluster de RedHat para ALMA Linux em preparação e andamento
 - Todos WN já em containers
 - Essa arquitetura permite o uso de recursos oportunistas



Federação de clusters do IFUSP

- Agregar recursos computacionais oportunistas do IFUSP e outros sites para uso coletivo
 - Canalizar recursos de HPC para beneficiar os pesquisadores
 - Flexibilização
 - Descentralização
 - Independência de administração
 - Tornar acessível o financiamento e a participação de interessados
 - Alocação racional do dinheiro público
- Projeto FINEP sendo submetido para essa infra (aprox R\$ 10M)

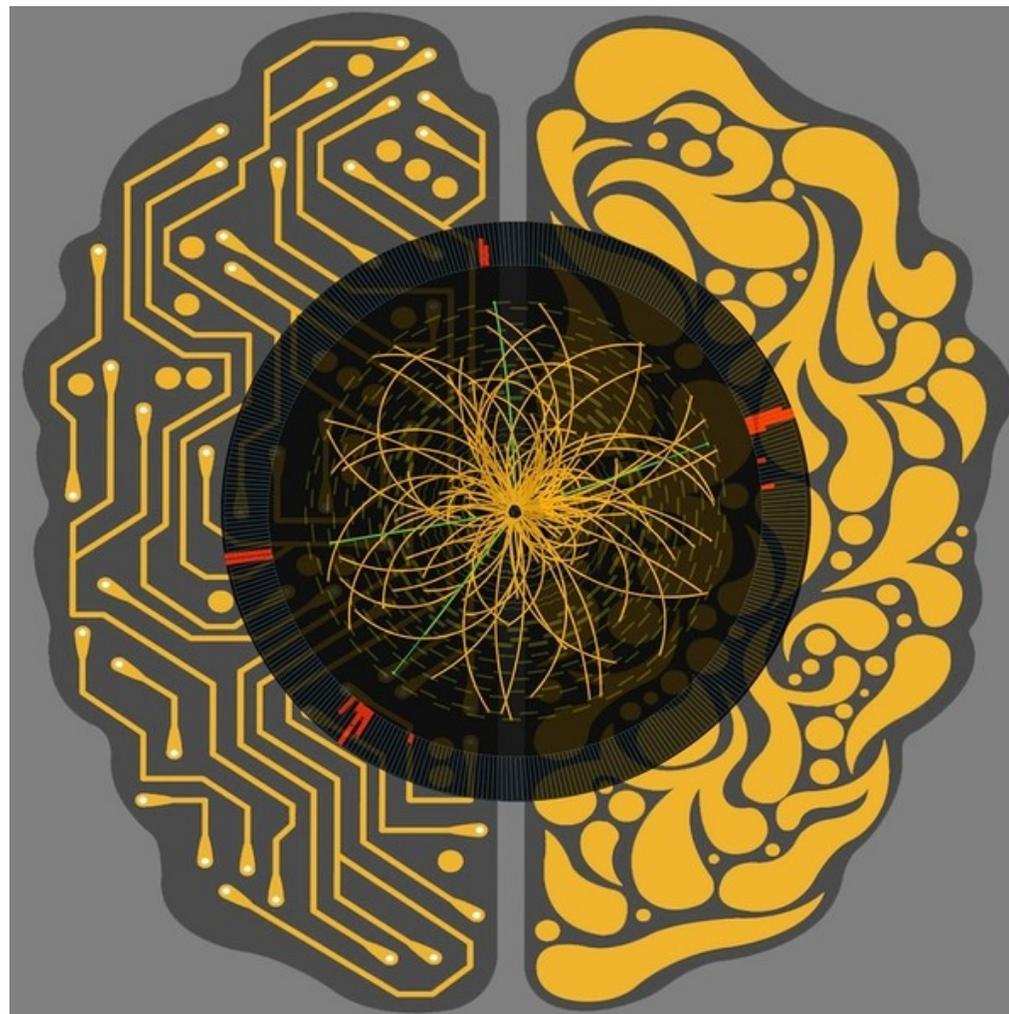
| Responsável | Departamento/Grupo | Qtdd | CPU | Freq | RAM |
|-----------------------------|--------------------|------|------------------------------------|----------|-----------|
| Prof. Oscar Eboli | FMA | 5 | 40 Cores | 2.20GHz | 251GB |
| Grupo HEPIC | FNC/SAMPA | 48 | 32 Cores | 1.70GHz | 125GB |
| Grupo HEPIC | FNC/SAMPA | 16 | 48 cores | 1.80GHz | 188GB |
| Grupo HEPIC | FNC/SAMPA | 9 | 32 Cores | | 125GB |
| Prof. Helena | FMT | 1 | 24 Cores | 1.20GHz | 65GB |
| Prof. Helena | FMT | 2 | 32 Cores | 3.30GHz | 264GB |
| Prof. Eric | FMT | 1 | 96 Cores | 2.10GHz | 512GB |
| Prof. Luis Gregório | FMT | 4 | 32 Cores | 2.13GHz | 8 GB |
| Prof. Frederique | FMA | 1 | 2x20 Cores | 2.1GHz | 128GB |
| Wanderley | FAP | 1 | 12 Cores | 1.9 GHz | 125.6 GB |
| Wanderley | FAP | 1 | 40 Cores | 2.9 GHz | 125.6 GB |
| Wanderley | FAP | 1 | 224 Tensor cores + 3804 CUDA cores | | |
| Wanderley | FAP | 1 | 12 Cores | 800 MHz | 125.6 GB |
| Wanderley | FAP | 1 | 16 Cores | 1.7 GHz | 94.28 GB |
| Wanderley | FAP | 1 | 16 Cores | 1.7 GHz | 94.28 GB |
| Wanderley | FAP | 1 | 16 Cores | 1.7 GHz | 94.28 GB |
| Wanderley | FAP | 1 | 16 Cores | 1.7 GHz | 94.28 GB |
| Wanderley | FAP | 1 | 16 Cores | 1.7 GHz | 94.28 GB |
| Wanderley | FAP | 1 | 32 Cores | 1107 MHz | 62.89 GB |
| Prof. Airton Depmann | FEP/Grenac | 1 | 2x10 Cores | 2.20GHz | 32 GB |
| Prof. Airton Depmann | FEP/Grenac | 6 | 2x10 Cores | 2.20GHz | 32 GB |
| Profa. Kaline | FGE | 7 | Total 207 cores + 2GPUs | | 563 GB |
| Profa. Kaline | FGE | 12 | Total 304 cores | | 544 GB |
| Profa. Kaline | FGE | 1 | 40 cores | | 125,42 GB |
| Profa. Kaline | FGE | 1 | 8 cores | | 30 GB |
| Profa. Kaline | FGE | 1 | 23 cores | | 94 GB |
| Profa. Kaline | FGE | 1 | 48 cores + 2GPUs | | 125 GB |
| Prof. Henrique M.J. Barbosa | FAP | 1 | 12 cores | | 64GB |
| Prof. Henrique M.J. Barbosa | FAP | 4 | 12 cores | | 24Gb |
| Prof. Henrique M.J. Barbosa | FAP | 4 | 20 cores | | 64GB |
| Prof. Caetano | FMT/SAMPA | 31 | 4x64 cores | 2.30GHz | 128GB |
| Prof. Caetano | FMT/SAMPA | 5 | 2x14 Cores | 2.20GHz | 96GB |
| Prof. Caetano | FMT/SAMPA | 9 | 2x18Cores | 2.20GHz | 96GB |
| Prof. Caetano | FMT/SAMPA | 7 | 4x12Cores | 2.20GHz | 192GB |
| Prof. Caetano | FMT/SAMPA | 1 | 10 Cores | | |
| Prof. Caetano | FMT/SAMPA | 2 | 72Cores | | |
| Prof. Caetano | FMT/SAMPA | 2 | GPUs | | |
| Prof Airton | FEP | 1 | 2x10 Cores | 2.20GHz | 32 GB |
| Prof Airton | FEP | 2 | 2x10 Cores | 2.20GHz | 32Gb |
| Prof Airton | FEP | 4 | 2x4 Cores | 2.60GHz | 32GB |
| Prof Airton | FEP | 1 | 2x20 Cores | 2.10GHz | 96GB |
| Prof. Aléssio | FEP | 1 | 2x24 Cores | | 192GB |
| Marco Leite | FNC | 3 | 8 Cores | 2.20GHz | 32GB |
| Marco Leite | FNC | 1 | 8 Cores | 2.20GHz | 32GB |

Em resumo

- Muitas coisas acontecendo/para acontecer no SAMPA em breve
- É importante que usuários fiquem atentos às mensagens nesse período
 - Vamos tentar agendar algumas "oficinas" para bom uso dessas novas facilidades
- Obrigado ao Ricardo Romão e ao Fábio Pádoa que têm sido essenciais nesse processo.

Uso de IA em projetos do grupo

- Vários alunos do grupo participando do curso de IA em física
 - Espero que isso motive o pessoal a pensar nessas ferramentas
- Vários projetos nasceram nos últimos meses com foco em aprendizado de máquina
 - IA em jatos (1 + 1?)
 - Hipernúcleos (2)
 - Evolução hidrodinâmica (1 + 2?)
- Resultados para esses projetos são previstos para o próximo relatório da FAPESP em 2025



Para o relatório atual...

- Descrição do desempenho do cluster e necessidade de atualização
 - Solicitação de aditivo
 - Upgrades necessários (ou parte deles)
 - Compra de mais processamento
- Sobre as atividades de análise
 - Descrição dessas análises que estão iniciando com a perspectiva de resultados iniciais em 2025 com possíveis apresentações em conferências