

WG-4 – Analysis techniques

Machine learning in Alice and Atlas

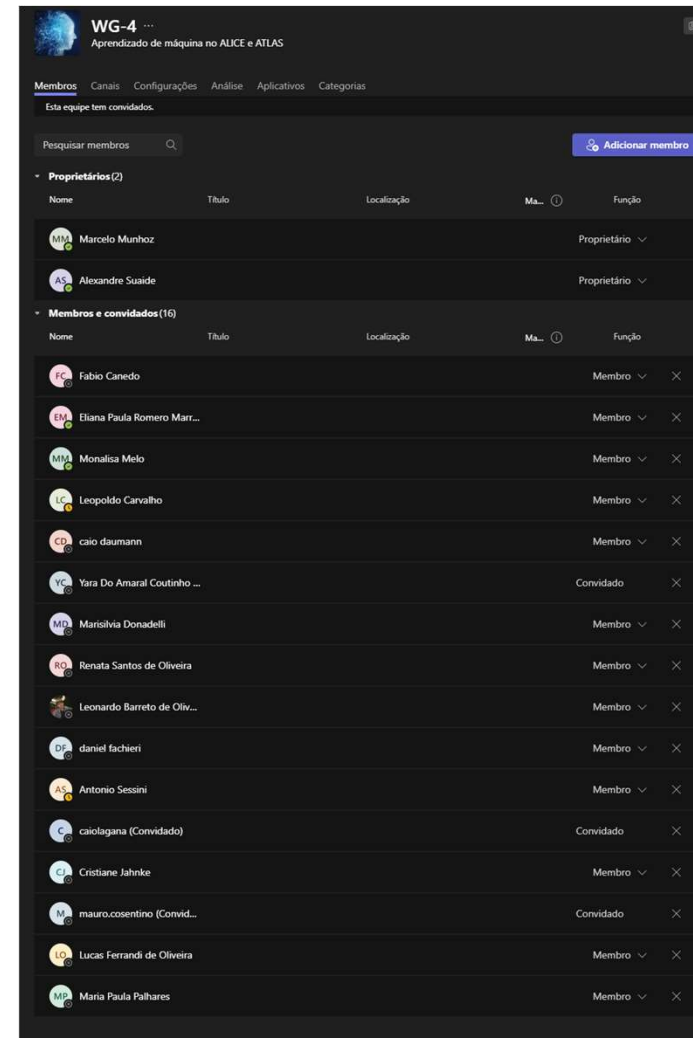
Alexandre Suaide

Objetivos do WG-4

- Desenvolver técnicas de análise e simulações envolvendo ML
 - Incorporar essas técnicas ao dia-a-dia das análises no Alice e Atlas
 - Grande interface com demais grupos
 - Grande envolvimento dos estudantes do projeto
- Infraestrutura computacional do projeto

Acesso à equipe no Teams

- 18 pessoas inscritas até o momento
- Dois canais criados para discussão de atividades
 - Infraestrutura computacional
 - Machine Learning



Análises que envolvem técnicas de ML

- Medidas da massa de Higgs para pares de tau no ATLAS
 - Caio Daumann
 - Estimativa de grandezas cinemáticas por conta de neutrinos impossíveis de serem medidos usando “Mixed Density Networks”
- Identificação de jatos no Alice usando ML
 - Antonio
 - Métodos a serem empregados ainda indefinidos
- Identificação de hádrons exóticos, X (3872)
 - Leopoldo
 - Uso de BDT (Boosted Decision Tree) em variáveis cinemáticas
- Identificação de hipernúcleos no Alice
 - Maria Paulo, Caio Laganá
 - Uso de redes neurais para identificação de hipernúcleos a partir de variáveis cinemáticas do decaimento desses núcleos

Ferramentas disponíveis para desenvolvimento

<http://jarvis.if.usp.br>

JupyterHub (python 3.7)

R-Studio

Streamlit server (em implementação)

Outras ferramentas e/ou atualizações
podem ser discutidas na equipe

JARVIS

High Energy Physics and Instrumentation Center

Selecione uma das opções abaixo

JupyterHub

R-Studio

WebROOT

HEPIC web page

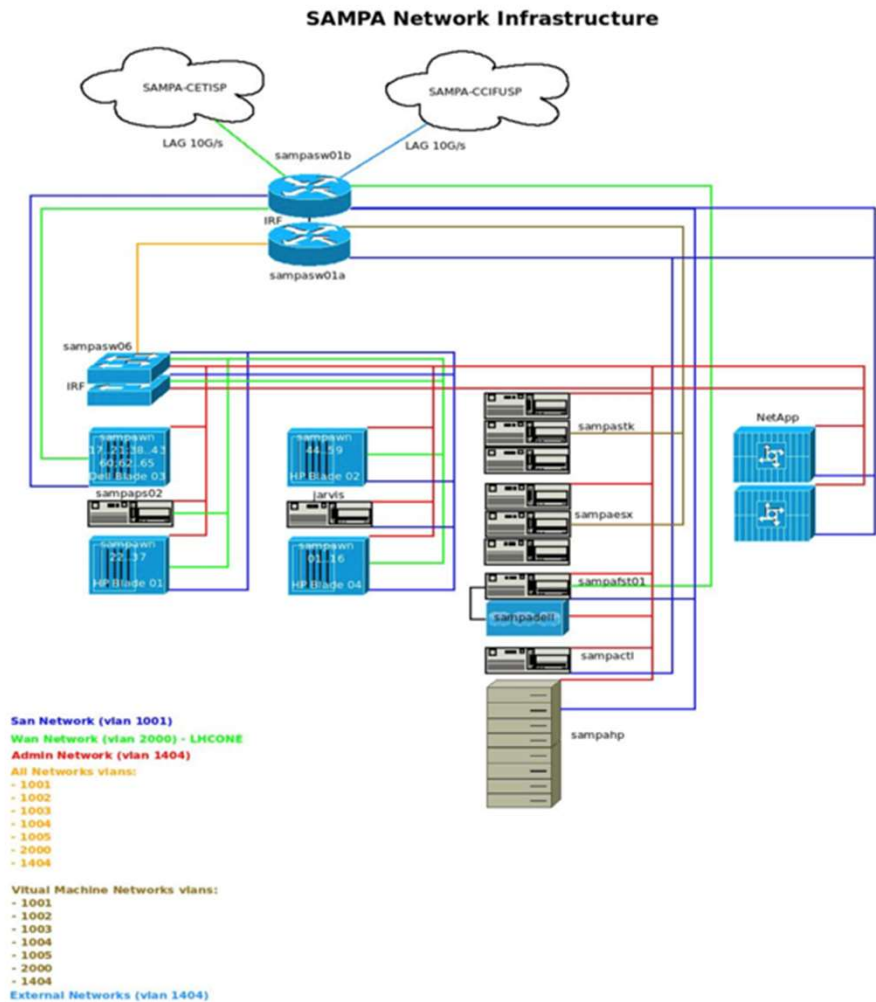


Dinâmica do grupo

- Reuniões a cada 3 semanas (sextas-feiras às 14:00) para discutir as análises em andamento
 - Presencial + virtual
 - Último dia 24 – Caio
 - 15/outubro – Maria Paula
- Participem!

Infraestrutura Cluster SAMPA

- Configuração atual
 - 2408 cores
 - 18.7 kHS06
 - 0.85 TB de armazenamento
 - 2 GPU Tesla K20
- Uso
 - Processamento em GRID do ALICE e ATLAS
 - Análise de dados e simulações
 - Processamento em geral
 - Desenvolvimento



Cronograma de upgrades

	Requested Processing Power (kHS06)	Extra Compute Modules for the SAMPA cluster	Processing Power per Compute Module (kHS06)
2022	20,7	4	0.62
2023	23,8	6	0.62
2024	27,4	5	0.75
2025	31,5	6	0.75
2026	36,3	6	0.90
TOTAL		27	

Table 2: Expect grown of the processing power of the computer cluster assuming 15% per year increase in the demand and 20% evolution every 2 years of the processing power of the compute modules

	Requested Disk Storage (PB)	Number of Extra Disks for the SAMPA Cluster	Single Disk Capacity (TB)
2022	2,12	91	14
2023	2,44	23	14
2024	2,80	23	16
2025	3,22	27	16
2026	3,71	27	18
TOTAL		191	

Table 3: Expect grown of the storage power of the computer cluster assuming 15% per year increase in the demand and 20% evolution every 2 years of single-disk capacity