

## **PROJETO – DE – PESQUISA**

### **Programa de Iniciação Científica e Tecnológica CBPF**

Nome do pesquisador ou tecnologista (orientador interno):

**Marcelo Portes de Albuquerque**

Coorientador:

**Wagner Philippe Calazans**

Coordenação:

**COTEC – Coordenação de Desenvolvimento Tecnológico**

Nome do pesquisador ou tecnologista (coorientador/colaborador externo, se houver):

**Não há**

Instituição de Pesquisa Externa (se houver):

**UFF – Universidade Federal Fluminense**

Título do projeto:

**Explorando Padrões Preditivos para Manutenções Preventivas através do Aprendizado de Máquina**

Palavras-chave:

**Aprendizado de máquina, análise de de dados**

Área de conhecimento:

**Ciência da Computação**

Pré-requisito desejado (se houver):

**Conhecimento da área de aprendizado de máquina**

Possibilidade de orientação remota: ( X ) Sim ( ) Não

Resultante principal do Projeto:

(X) Publicação (horizonte de 4 anos).

(X) Preparação do bolsista para área científica.

( ) Produto tecnológico.

(X) Produto educacional ou didático.

Rio de Janeiro, 25 de abril de 2024

# Projeto

## Introdução:

A ampla adoção de técnicas de monitoramento de rede e serviços resultou na geração massiva de dados. No entanto, devido à complexidade das redes e à quantidade abundante de dados distribuídos, poucos avanços foram feitos na análise e inferência do comportamento dos dispositivos de rede. Este projeto propõe investigar técnicas de Big Data Analytics para analisar os dados coletados por ferramentas de monitoramento amplamente utilizadas pela comunidade, como Cacti, Zabbix e Splunk, entre outras. O objetivo é desenvolver um sistema capaz de fornecer informações relevantes para a predição de falhas de hardware e software, visando a redução do tempo de inatividade do parque computacional e a simplificação da operação de rede.

## Desenvolvimento:

Reduzir o tempo de inatividade do parque é um desafio significativo para provedores de serviços de todo tipo, sejam eles acadêmicos, comerciais ou grandes usuários corporativos e acadêmicos, como o Centro Brasileiro de Pesquisas Físicas. As principais causas desse tempo de inatividade são as falhas conhecidas como "hard failures", que ocorrem quando algum elemento ou componente do ambiente não consegue fornecer conectividade. Isso pode incluir falhas em interfaces de rede, equipamentos ou até mesmo a interrupção do sinal devido a uma fibra ótica rompida. Antecipar essas falhas e tomar medidas corretivas pode reduzir significativamente o tempo de inatividade.

Os parâmetros de operação desses equipamentos são registrados via rede e em uma estação de monitoramento por meio de protocolos e ferramentas estabelecidas, como SNMP e Syslog. No entanto, a análise desses dados e a criação de inferências sobre o comportamento de rede e clusters ainda são áreas pouco exploradas.

Este projeto visa pesquisar e desenvolver ferramentas para prever "hard failures" usando técnicas de Big Data Analytics, com o objetivo de minimizar o tempo de inatividade. Isso garantirá um acesso mais longo aos recursos computacionais do CBPF, seja para pesquisa científica, como o acesso a grandes experimentos internacionais, ou para fornecer serviços à comunidade científica por meio de projetos colaborativos.

Para isso, serão estudadas e mapeadas as informações coletadas pelas

ferramentas de monitoramento de rede, permitindo a compreensão de quais dados

são mais adequados para análise. Estratégias serão desenvolvidas para processar o grande volume de dados e extrair correlações e inferências. Serão desenvolvidos algoritmos dinâmicos e adaptativos para processar esses dados e fazer previsões com base em padrões e tendências. Isso incluirá a identificação e o desenvolvimento de um algoritmo ótimo para a previsão de falhas.

### **Referência Biblioteca:**

Observação: Se necessário informar referências técnico/científicas que possam fundamentar o trabalho proposto.

- [1] YU, Shui et al. Networking for big data. Ed.: Chapman and Hall/CRC, 2015.
- [2] MARCOS, P. et al. DYNAMIXA Dynamic Agreement Marketplace on Internet eXchange Points. In: CoNEXT Student Workshop. 2016.
- [3] MORAES, L. F. M. D.; ALBUQUERQUE, M. P. D.; RIBEIRO FILHO, J. L. Infraestrutura Redes de Alta Velocidade no Rio de Janeiro: história e estado da arte. In: (Ed.). A História da Telessaúde da Cidade para o Estado do Rio de Janeiro. Rio de Janeiro: EdUERJ, 2015. ISBN 978-85-7511395-0.
- [4] SHUAN, Lam Hai et al. Network Equipment Failure Prediction with Big Data Analytics. International Journal of Advances in Soft Computing & Its Applications, v. 8, n. 3, 2016.
- [5] IEEE Xplore. 2019. Disponível em:  
<<https://ieeexplore.ieee.org/abstract/document/9615039/>>.  
Acesso em: 20 de abril de 2024.