

Neste artigo, um sistema de detecção de anomalias é projetado para detectar e diagnosticar as anomalias em microserviços através de mono-torneamento e análise dos dados de desempenho em tempo real dos mesmos. O ADS proposto consiste em um módulo de monitoramento que coleta os dados de desempenho de recipientes, um módulo de processamento de dados baseado no mod- els de aprendizagem de máquinas e um módulo de injeção de falhas integrado para o treinamento desses modelos. O módulo de injeção de falhas também é usado para avaliar a detecção de anomalias e o desempenho do diagnóstico de nosso ADS.

Atualmente, cada vez mais aplicações Web são desenvolvidas em projetos de microserviços para melhor escalabilidade, flexibilidade e confiabilidade. Com a adoção de uma arquitetura de microserviços, muitos benefícios podem ser obtidos. Cada VM criada tem seu próprio sistema operacional . Container é outra tecnologia emergente para virtualização que está ganhando popularidade sobre as VMs devido a seu peso leve, alto desempenho e maior escalabilidade .

E os contêineres criados compartilham o sistema operacional . O desenvolvimento de tecnologias de virtualização, especialmente a tecnologia de contêineres, contribuiu para a ampla adoção da arquitetura de microserviços nos últimos anos. E os prestadores de serviços começam a colocar maiores exigências sobre a confiabilidade desses microserviços. Eles podem incluir vários aspectos tais como requisitos de desempenho e propriedades de confiabilidade .

A detecção de anomalias pode nos ajudar a identificar padrões incomuns que não se conformam aos padrões esperados e o diagnóstico de anomalias pode nos ajudar a localizar a causa raiz de uma anomalia. Como a detecção e o diagnóstico de anomalias exigem uma grande quantidade de dados históricos, os prestadores de serviços têm que instalar muitas ferramentas de monitoramento em sua infra-estrutura para coletar dados de desempenho de seus serviços em tempo real. Em primeiro lugar, para os microserviços baseados em contêineres, que métricas devem ser monitoradas. Neste documento, um sistema de detecção de anomalias é proposto e pode abordar estes dois principais desafios de forma eficiente.

ADS proposto dá um tipo proto-tipo para os prestadores de serviços para detectar e diagnosticar anomalias para microserviços baseados em contêineres com monitoramento de desempenho.

## **Backgroud**

Os contêineres são virtualizações leves em nível de SO que nos permitem executar uma aplicação e suas dependências em um processo isolado de recursos. Com cada vez mais aplicações e serviços implantados em ambientes hospedados em nuvem, a arquitetura de microserviços depende muito do uso da tecnologia de contêineres. A detecção de anomalias é a identificação de itens, eventos ou observações que não estão em conformidade com um padrão esperado ou outros itens em um conjunto de dados . A correlação entre a carga de trabalho e o desempenho da aplicação deve ser estável e flutua significativamente quando as falhas são acionadas .

## **Trabalho Relacionado**

Com a ampla adoção da arquitetura de microserviços e tecnologias de contêineres, o monitoramento e a avaliação do desempenho se tornam um tema quente para os pesquisadores de contêineres. Em , os autores avaliaram o desempenho dos microserviços baseados em contêineres em dois modelos diferentes com os dados de desempenho da CPU e da rede. Em , os autores forneceram uma comparação de desempenho entre um ambiente Linux nativo, containers Docker e KVM. Eles tiraram uma conclusão de que o uso do docker poderia alcançar uma melhoria por forma de acordo com as métricas de desempenho coletadas por suas

ferramentas de benchmarking.

Em , os autores apresentaram sua abordagem de detecção de anomalias para os vícios dos serviços de nuvem. Os dados de desempenho de cada VM foram coletados e então, processados para detectar possíveis anomalias com base em técnicas de aprendizagem de máquinas. Em , os autores prestaram atenção ao desempenho do sistema. Para detectar anomalias, eles construíram modelos com métricas históricas de desempenho e os compararam com modelos monitorados on-line.

Este documento tem como objetivo criar um ADS que possa detectar e diagnosticar anomalias para microserviços baseados em contêineres com monitoramento de desempenho.

## **Sistema de Detecção de Anomalias**

Esta seção apresenta uma visão geral de nosso sistema de detecção de anomalias. Primeiro, o módulo de monitoramento coleta os dados de monitoramento do desempenho do sistema alvo. Em seguida, o módulo de processamento de dados analisará os dados coletados e detectará as anomalias. O módulo de injeção de falhas simula falhas de serviço e reúne conjuntos de dados de monitoramento de desempenho que representam condições normais e anormais.

Os dados de desempenho do sistema alvo são coletados e processados para a detecção de possíveis anomalias. O ADS proposto tem que aprender com os dados de monitoramento de desempenho com os próprios modelos de aprendizagem da máquina.

## **Agente de monitoramento**

Para aplicações implantadas em tais ambientes baseados em contêineres, os dados de monitoramento de desempenho devem ser coletados de várias camadas de uma aplicação . Ao processar os dados de desempenho de um contêiner, podemos dizer se o contêiner funciona bem. Ao coletar os dados de desempenho de todos os recipientes relacionados, podemos obter os dados de desempenho total de um micro-serviço específico. E também podemos saber se um microserviço é anômalo ao processar esses dados de desempenho do serviço.

## **Processamento de dados**

Tarefas de processamento de dados O processamento de dados nos ajuda a detectar e diagnosticar anomalias - mentiras. Modelos de processamento de dados As técnicas de detecção de anomalias são baseadas em algoritmos de aprendizagem ma-chine. Todos estes algoritmos podem ser aplicados para classificar os comportamentos do sistema alvo com dados de monitoramento de desempenho. Para detectar diferentes tipos de anomalias que podem levar a SLAVs, são utilizados algoritmos de aprendizagem super-visuais.

Primeiramente, amostras de dados de desempenho rotulados que representam diferentes comportamentos de serviço são coletadas e armazenadas em um banco de dados. Estas amostras são chamadas de dados de treinamento. Em seguida, o módulo de processamento de dados treina os modelos de classificação com estes dados de treinamento. Para coletar mais dados de desempenho em diferentes tipos de erros, um módulo de injeção de falhas é implantado e ele injetará diferentes falhas em containers.

A segunda fase é a fase de detecção. Uma vez que o modelo é treinado, alguns dados de desempenho em tempo real podem ser coletados e transferidos para o módulo de processamento de dados como entradas, e o módulo de processamento de dados pode detectar anomalias que ocorrem no sistema com o modelo treinado. Para a validação do módulo de

processamento de dados, alguns erros serão injetados no sistema alvo, e então o módulo de processamento de dados usa os dados de desempenho em tempo real para detectar estes erros.