Um Sistema de Monitoramento de Anomalias em Containers de Docker Baseado em Floresta de Isolamento Otimizado

Este artigo propõe um sistema online de detecção de anomalias em contêineres, monitorando e analisando as métricas de recursos multidimensionais dos contêineres, com base no algoritmo florestal de isolamento otimizado. Além disso, o algoritmo ele pode identificar métricas anormais de recursos e ajustar automaticamente o período de monitoramento para reduzir o atraso de monitoramento e a sobrecarga do sistema. Os resultados experimentais demonstram o desempenho e eficiência do sistema na detecção das anomalias típicas em contêineres, tanto em ambientes simulados quanto em ambientes de nuvens reais.

W a popularidade das plataformas de computação em nuvem, cada vez mais empresas têm seus próprios dados centers, prestando serviços a clientes com necessidades diferentes. Uma das tecnologias-chave no centro de dados é a virtualização. Entretanto, com a crescente aplicação em larga escala de grupos de contêineres, a questão da segurança e estabilidade dos contêineres também tem chamado cada vez mais a atenção. Cloud que se baseia em cluster de contêineres e máquinas virtuais levou à invalidação de milhares de websites e aplicativos .

Como os recipientes continuam a subir e descer, um dos desafios é como monitorar múltiplos recursos ao mesmo tempo em um ambiente dinâmico com uma baixa sobre-cabeça.

O método baseado em estatísticas pressupõe que os dados obedecem a alguns modelos de distribuição padrão e descobre os valores anômalos que se desviam da distribuição. Como a maioria dos modelos é baseada em suposições univariadas, eles não são aplicáveis a dados multidimensionais. Eles identificam aberrações estimando a densidade dos dados locais ou calculando a mudança de ângulo. Entretanto, ambos incorrem em uma grande sobrecarga de cálculo quando o tamanho da amostra de dados é grande.

Os sistemas de monitoramento existentes geralmente adotam um período fixo de mono-torção para consultar a anormalidade do sistema. Quando o período de monitoramento é muito pequeno, o sistema de monitoramento pode localizar rapidamente as anormalidades. No entanto, isto resulta em uma sobrecarga enorme do sistema quando há muitos objetos de monitoramento. Quando o período de monitoramento é grande, o atraso do monitoramento também aumentará.

Portanto, é necessário adotar um período de monitoramento adequado de acordo com o estado de funcionamento do sistema. Anomalia através do monitoramento da métrica dos recursos de contêineres. Este documento propõe um sistema de monitoramento de anomalias de contêineres baseado em florestas de isolamento otimizadas. O sistema primeiro obtém cada taxa de utilização de recursos de cada contêiner na máquina hospedeira de forma não intrusiva.

Quando dados suficientes de monitoramento são coletados, o valor da anomalia de cada dado de monitoramento é calculado usando a floresta de isolamento otimizada, que leva em conta as características da carga de trabalho da aplicação do contêiner. Especificamente, o sistema atribui um peso a cada métrica de recurso. Se uma aplicação de contêiner depende muito de uma métrica de recurso , o sistema atribuirá um grande valor a essa métrica de recurso. De forma correspondente, mudamos a seleção aleatória de recursos para a seleção ponderada de recursos ao escolher um fea- ture dos dados para dividir o conjunto de dados no algoritmo de isolamento florestal.

Assim, se um recurso métrico com grande peso estiver em um estado anormal, será mais fácil

escolhê-lo como a característica para dividir o conjunto de dados. Então, o sistema identifica a causa da anomalia através da análise dos registros do contêiner. Ao mesmo tempo, o sistema pode aumentar ou diminuir o período de monitoração de acordo com o grau de anomalias.

Primeiro, o Docker tem maior desempenho e eficiência do que

Implementamos tanto o sistema quanto algo- rithm e os avaliamos tanto em ambientes simulados como em nuvens comerciais reais em uma grande variedade de casos de anomalias em termos de detecção de acesso, atraso de monitoramento e análise de log.

ANTECEDENTES E TRABALHO RELACIONADO

Em seguida, elaboramos o trabalho relacionado ao sistema de monitoramento e aos métodos de detecção de anomalias.

Os passos para a construção de um iTree são os seguintes

Primeiro, obtemos n amostras dos dados N como as amostras de treinamento para esta árvore. Em terceiro lugar, repetimos o processo acima nos itens de dados à esquerda e à direita até atingir a condição de rescisão. Uma é que os dados em si não podem ser divididos , e a outra é que a altura da árvore atinge o toro2 . Suponha que o comprimento do caminho entre cada dado x e o nó raiz é h, a média de todos os h é Eh.

s é o valor da anomalia dos dados x nas amostras n de um conjunto de dados. Ao contrário da virtualização em camada dura das máquinas virtuais, Docker não tem emulação de hardware, e implementa a virtualização no nível do sistema operacional .

s= 2

Ele pode obter parâmetros individuais e dados históricos de utilização de recursos para cada contêiner. Além disso, os dados dos gráficos são apenas uma janela deslizante de um minuto. Não há função de armazenamento de dados, e não há função de alarme. Arquitetura do sistema.

Se a maioria dos s estiver próxima a 0,5, todo o conjunto de dados é considerado como não tendo valores aberrantes óbvios.

Sistema de Monitoramento

Ganglia é um projeto de monitoramento de clusters de código aberto ini- tiatizado pela UC Berkeley. O Gmond é instalado na máquina física monitorada e é responsável pelo monitoramento da coleta de dados. Gmetad é responsável pela coleta de dados sobre os nós gmond e gmetad. O web front end pode mostrar dados em tempo real de todo o sistema de monitoramento.

Nagios é um sistema de monitoramento que monitora o status operacional do sistema e as informações da rede. propuseram um método simples de monitoração de recipientes que utiliza o próprio API do estivador para obter recursos e armazená-los no banco de dados. O método estima o desvio de stan- dard de um parâmetro de monitoramento de recursos.