

ANÁLISE DE DESEMPENHO DE SISTEMAS DISTRIBUÍDOS EM AMBIENTES DE NUVEM

PERFORMANCE ANALYSIS OF DISTRIBUTED SYSTEMS IN CLOUD ENVIRONMENTS

Deyvid Yancey Guedes Silva¹

Gustavo Barros Lins²

RESUMO

O presente trabalho de conclusão de curso (TCC) tem como objetivo principal a análise de desempenho de sistemas distribuídos em ambientes de nuvem utilizando de três ferramentas: *Apache JMater*, *Prometheus* e *Grafana*. A análise de desempenho em sistemas distribuídos em nuvem é uma área de estudo fundamental para avaliar a capacidade de processamento, armazenamento e comunicação desses sistemas. Por meio da coleta e análise de métricas de desempenho, é possível identificar gargalos, mensurar a utilização dos recursos e identificar possíveis melhorias para otimizar o desempenho. O trabalho envolverá a aprofundamento bibliográfica de artigos científicos, livros e papéis técnicos que abordem a decomposição de andamento de sistemas distribuídos em ambientes de nuvem. Serão exploradas diferentes técnicas e metodologias utilizadas nessa área, como análise de carga, escalabilidade, latência, *throughput* e outros indicadores relevantes, para a medição dessas métricas serão utilizadas ferramentas tais como *Prometheus*, *Apache JMater* e *Grafana*. O resultado ajudará empresas e profissionais na tomada de decisões relacionadas à infraestrutura e configuração de sistemas distribuídos em nuvem.

Palavras-Chave: Análise de desempenho; Sistemas distribuídos; Ambiente de nuvem; Eficiência; Otimização.

ABSTRACT

This course completion work (TCC) has as its main objective the performance analysis of distributed systems in cloud environments. performance analysis in distributed cloud systems is a fundamental area of study to evaluate the processing, storage and communication capacity of these systems. through the collection and analysis of performance metrics and using tools such as *Apache JMater*, *Prometheus* and *Grafana*, it is possible to identify bottlenecks, measure the use of resources and identify possible improvements to optimize performance.

¹Acadêmico do Curso de Bacharelado em sistemas de informações da Associação Vitorienne de Educação, Ciência e Cultura, Centro Universitario FACOL - UNIFACOL; Deyvidy.silva@unifacol.edu.br.

²Professor Mestre e Orientador do Curso Bacharelado em Sistemas de Informações da Associação Vitoriense de Educação, Ciência e Cultura, centro Universitário FACOL - UNIFACOL; Gustavo.lins@unifacol.edu.br.

the work will involve a bibliographic review of scientific articles, books and technical documents that address the performance analysis of distributed systems in cloud environments. Different techniques and methodologies used in this area will be explored, such as load analysis, scalability, latency, throughput.

Keywords: *Performance analysis, Distributed systems, Cloud environment, Bottlenecks, Efficiency, Optimization.*

1 INTRODUÇÃO

Os sistemas distribuídos em ambientes de nuvem têm se tornado cada vez mais populares, uma vez que oferecem várias vantagens em relação a sistemas centralizados tradicionais, tais como escalabilidade, flexibilidade e disponibilidade. No entanto, garantir o desempenho desses sistemas pode ser um desafio, pois eles podem envolver uma grande quantidade de recursos computacionais, que precisam ser gerenciados de maneira eficiente.

Para avaliar o desempenho desses sistemas, é necessário levar em consideração diversos fatores, tais como a capacidade de processamento dos servidores, a latência de rede, a disponibilidade dos serviços e a escalabilidade. Além disso, é preciso escolher as métricas adequadas para avaliar o desempenho do sistema, como tempo de resposta, taxa de transferência, uso de CPU e memória, entre outras.

Coulouris **et al.** (2013) argumentam que sistemas distribuídos são essenciais para o funcionamento da Web, e que eles são usados em uma ampla gama de aplicações. Os autores discutem que o crescimento da internet e da web impõe novas demandas em termos de arquitetura, requisitos de projeto, variedade de aplicações e cargas de trabalho. Além disso, o artigo discute as principais muitas implementações em que sistemas distribuídos são utilizados.

A avaliação do desempenho de sistemas distribuídos em ambientes em nuvem é uma tarefa desafiadora, devido à complexidade dos sistemas e à variação das cargas de trabalho. Segundo Leonardo Menezes de Souza (2012), os sistemas de processamento distribuído,

como *grids*, *clusters* e *PPI*, têm sido alvo de estudos de análise de desempenho. Apesar dos avanços na área, ainda não existe um conjunto de padrões e métricas definidos, o que dificulta a comparação e interpretação dos resultados. Nesse contexto, o presente trabalho tem como objetivo realizar uma revisão bibliográfica da literatura existente sobre a análise de desempenho de sistemas distribuídos em ambientes em nuvem, visando identificar as principais metodologias, ferramentas e métricas utilizadas.

A partir dessa revisão bibliográfica, será possível compreender as principais técnicas e abordagens adotadas para avaliar o desempenho desses sistemas e identificar as melhores práticas para garantir sua eficiência e escalabilidade. Espera-se, com isso, contribuir para o desenvolvimento de sistemas distribuídos em ambientes em nuvem mais robustos e confiáveis, capazes de atender às demandas cada vez mais complexas do mercado.

No contexto de tudo que foi tratado aqui, surge o questionamento que norteia essa pesquisa: quais as principais metodologias e ferramentas utilizadas para a análise de desempenho de sistemas distribuídos em ambientes de nuvem?

Com base nessa pergunta, foi estabelecida a seguinte hipótese: A análise de tais métricas podem e devem contribuir para uma boa utilização e um bom desempenho de tais sistemas.

Diante desses dados percebe-se diante da enorme demanda desses tipos de sistemas a importância de manter o bom funcionamento e o alto desempenho dos sistemas.

Para tanto, foram delineados os seguintes objetivos específicos: revisar as metodologias e ferramentas utilizadas para a análise de desempenho de sistemas distribuídos em ambientes de nuvem, identificar as métricas de desempenho mais relevantes para os mesmos e descrever os principais desafios encontrados na garantia do desempenho dos sistemas.

Apesar das vantagens oferecidas pelos sistemas distribuídos em ambientes de nuvem, como escalabilidade e flexibilidade, a garantia de um desempenho eficiente permanece um desafio crítico. O aumento na complexidade e no uso de recursos nesses sistemas levanta questões sobre como medir, monitorar e otimizar o desempenho. Como objetivo geral, este trabalho de pesquisa irá abordar essa lacuna, focando na análise as principais metodologias e ferramentas utilizadas para a análise de desempenho de sistemas distribuídos em ambientes em nuvem.

Na primeira seção será realizada uma fundamentação teórica onde serão mapeados os principais conceitos relacionados ao tema, incluindo sistemas distribuídos, computação em nuvem, análise de desempenho e métricas utilizadas na avaliação de desempenho de sistemas distribuídos em ambientes em nuvem.

Na segunda seção o assunto abordado será as métricas para avaliação de desempenho, nessa sessão, seria apresentado as principais métricas utilizadas para avaliar o desempenho de sistemas distribuídos em ambientes em nuvem, tais como latência, throughput e disponibilidade.

Na terceira seção serão abordadas as metodologias e ferramentas para análise de desempenho, nessa seção, será apresentada uma revisão bibliográfica das principais metodologias e ferramentas utilizadas para a análise de desempenho de sistemas distribuídos em ambientes em nuvem. Serão discutidos as vantagens e desafios de cada metodologia e ferramenta, visando identificar as melhores práticas para a análise de desempenho.

2 FUNDAMENTAÇÃO TEÓRICA

Existe uma grande necessidade do conhecimento sobre ferramentas que podem nos trazer medidas exatas sobre o desempenho desses sistemas, mas primeiro temos que saber o que são esses sistemas.

2.1 sistemas distribuídos

Um sistema distribuído é um conjunto de computadores independentes interconectados que trabalham juntos como uma única entidade lógica. Ao contrário de um sistema centralizado, em que todas as tarefas são executadas em um único computador, um sistema distribuído divide o processamento, o armazenamento e as tarefas entre vários computadores interligados em rede.

Marcelo M. Gonçalves propõe que Sistemas distribuídos são parte de uma categoria de computação conhecida como computação paralela (*parallel computing*), classificada conforme o grau de paralelismo oferecido pelo hardware(Gonçalves, 2022)..esse tipo de sistema, cada computador, também conhecido como nó, possui seus próprios recursos de hardware e software. Esses nós se comunicam e coordenam suas atividades por meio de troca

de mensagens e compartilhamento de dados, colaborando para realizar tarefas complexas e atingir objetivos comuns.

Luiz Antônio Vivacqua Corrêa Meyer fala que realização da computação paralela envolve a fragmentação do problema em unidades menores, as quais podem ser processadas simultaneamente por diversos processadores (Meyer, 2019). A distribuição das tarefas em um sistema distribuído traz vários benefícios, como aumento de desempenho, escalabilidade, confiabilidade e disponibilidade. Ao dividir a carga de trabalho entre vários nós, o sistema distribuído pode lidar com grandes volumes de dados e realizar cálculos complexos de maneira mais eficiente e rápida. Além disso, a distribuição dos recursos reduz a dependência de um único ponto de falha, aumentando a confiabilidade e a disponibilidade do sistema.

Com a diversidade de usos empregados a sistemas distribuídos, eles são empregados em muitas áreas, como processamento paralelo, sistema de gerenciamento de banco de dados distribuídos, entre outros. São fundamentais para lidar com problemas complexos que exigem processamento intensivo, cooperação entre várias entidades e em grande escala.

No entanto, a concepção e implementação de sistemas distribuídos também apresentam desafios, como garantir a consistência dos dados, gerenciar a comunicação entre os nós, lidar com a sincronização e a concorrência e garantir a segurança e a integridade do sistema. Portanto, é necessário um planejamento cuidadoso e o uso de algoritmos e protocolos adequados para garantir o funcionamento eficiente e confiável de um sistema distribuído.

2.2 Computação em nuvem

Segundo Eduardo Savarese Neto a computação em nuvem consiste na disponibilização de serviços computacionais conforme a necessidade, por meio da internet.(Neto, 2019), também conhecida como "*cloud computing*" em inglês, refere-se a um modelo de provisionamento e entrega de serviços de computação através da Internet. Em vez de executar aplicativos ou armazenar dados localmente em um computador pessoal, ou servidor local, a computação em nuvem permite que os usuários acessem recursos computacionais, como armazenamento, servidores, banco de dados e software, por meio de provedores de serviços em nuvem.

A nuvem é uma metáfora para a Internet, e os serviços em nuvem são fornecidos por meio de uma infraestrutura de servidores distribuídos geograficamente, geralmente mantida por grandes provedores como *Amazon Web Services (AWS)*, *Microsoft Azure*, *Google Cloud Platform* e outros. Esses provedores de serviços em nuvem oferecem uma variedade de modelos de implantação, como nuvem pública, nuvem privada e nuvem híbrida, para atender às necessidades das organizações e indivíduos.

Em um artigo acadêmico Bruna Thomé, Eduardo Hentges e Dalvan Griebler definem que a computação em nuvem oferece a conveniência de acessar recursos computacionais, como servidores, armazenamento, redes, serviços e aplicativos, de maneira ágil e sob demanda. Esses recursos podem ser provisionados de forma eficiente e disponibilizados ao usuário sem a necessidade de intervenção gerencial, isso elimina a necessidade de investimentos iniciais em infraestrutura e permite uma escalabilidade flexível, onde os recursos podem ser facilmente aumentados ou reduzidos de acordo com a demanda. Além disso, a computação em nuvem oferece maior disponibilidade, confiabilidade e segurança dos dados, com provedores de serviços em nuvem geralmente assumindo a responsabilidade pela manutenção e segurança da infraestrutura.

Diversos serviços são oferecidos na computação em nuvem, como infraestrutura como serviço (*IaaS*), plataforma como serviço (*PaaS*) e software como serviço (*SaaS*). No modelo *IaaS*, os usuários têm acesso a recursos de infraestrutura, como servidores virtuais, redes e armazenamento. No modelo *PaaS*, é fornecida uma plataforma de desenvolvimento e execução de aplicativos, com ferramentas e ambiente de execução pré-configurados. Já no modelo *SaaS*, os usuários têm acesso a aplicativos completos, disponibilizados através da nuvem.

Em resumo, a computação em nuvem revoluciona a forma como os recursos computacionais são acessados e utilizados, proporcionando maior flexibilidade, eficiência e economia de custos para indivíduos e organizações.

2.3 Sistemas Distribuídos em Ambientes de Nuvem

Sistemas distribuídos em ambientes de nuvem são sistemas de computação que utilizam recursos distribuídos em diferentes locais geográficos, conectados por meio de uma rede, geralmente a Internet. Esses sistemas são executados em ambientes de nuvem,

compostos por servidores, dispositivos de armazenamento e recursos de rede, fornecidos por provedores de serviços em nuvem, como a Amazon Web Services, Microsoft Azure e Google *Cloud Platform*. Os sistemas distribuídos em nuvem permitem que aplicativos complexos sejam executados em larga escala, com alto desempenho e escalabilidade, além de permitir o compartilhamento de recursos e a tolerância a falhas.

De acordo com Li e Li (2020), os sistemas distribuídos em ambientes de nuvem têm atraído cada vez mais atenção devido às vantagens que oferecem, como escalabilidade, disponibilidade e redução de custos de infraestrutura. Além disso, esses sistemas são capazes de lidar com grandes volumes de dados em tempo real, o que é essencial para diversas aplicações modernas.

Entende-se que sistemas distribuídos em ambiente de nuvem são fundamentais para a sustentação das aplicações modernas. Alguns exemplos são, mas não se limitam a: armazenamento de dados em larga escala: serviços de armazenamento em nuvem como Amazon S3 e Google Cloud *Storage* utilizam sistemas distribuídos para armazenar e gerenciar grandes volumes de dados, *Machine Learning* em larga escala: serviços de machine *learning* em nuvem, como Amazon SageMaker e Google Cloud ML, usam sistemas distribuídos para treinar e implantar modelos de machine *learning* em larga escala, entre outros

A computação em nuvem é uma evolução dos serviços e produtos de tecnologia da informação sob demanda, também chamada de Utility Computing [Brantner et al. 2008]. A ideia de compartilhar recursos de computação entre diversos usuários remotos já era explorada há muito tempo, principalmente em grandes organizações que utilizavam unidades centrais de processamento para processar grandes volumes de dados.

Com o desenvolvimento da internet e das redes de computadores, a computação distribuída tornou-se mais acessível e viável para um número crescente de usuários. A partir dos anos 90, surgiram as primeiras soluções de computação em nuvem, como o serviço de e-mail Hotmail, que utilizava servidores distribuídos para prover o acesso à caixa de entrada dos usuários.

Desde então, a computação em nuvem vem evoluindo rapidamente, impulsionada pela crescente demanda por serviços de processamento, armazenamento e gerenciamento de dados em larga escala. A partir dos anos 2000, surgiram os primeiros provedores de serviços em

nuvem, como a *Amazon Web Services*, que oferecem infraestrutura virtualizada e serviços de plataforma para desenvolvedores e usuários finais.

Atualmente, a computação em nuvem é amplamente utilizada por empresas de todos os tamanhos e setores, sendo vista como uma solução flexível, escalável e econômica para as demandas crescentes por recursos computacionais. A tendência é que os sistemas distribuídos em ambientes de nuvem continuem a evoluir e se tornem ainda mais integrados e automatizados, permitindo uma maior eficiência na gestão e uso dos recursos de computação em larga escala.

3 MÉTRICAS DE ANÁLISE DE DESEMPENHO

A análise de sistemas distribuídos em ambientes de nuvem envolve a consideração de várias métricas para avaliar o desempenho, eficiência e confiabilidade desses sistemas. Essas métricas são essenciais para entender e monitorar o funcionamento dos sistemas distribuídos em nuvem, garantindo sua eficácia e capacidade de atender às demandas dos usuários.

Uma das métricas mais importantes é a latência, que mede o tempo que uma solicitação leva para ser processada e obter uma resposta. A latência é especialmente crítica em aplicações sensíveis ao tempo, como transações financeiras e jogos online, onde cada milissegundo conta para uma experiência fluida.

Existem várias causas para a latência em sistemas distribuídos na nuvem. Uma delas é a distância física entre os servidores, especialmente em ambientes de nuvem pública, onde os data centers podem estar localizados em regiões geograficamente distantes. Quanto maior a distância, maior será o tempo necessário para que os dados percorram essa distância, resultando em latência.

Outra métrica fundamental é o *throughput*, que indica a quantidade de trabalho que um sistema distribuído pode realizar em um determinado período. É uma medida da capacidade de processamento do sistema e está diretamente relacionada à quantidade de solicitações que ele pode lidar simultaneamente.

Em um ambiente de nuvem, onde os recursos estão distribuídos em diferentes servidores e data centers, o *throughput* pode ser afetado por vários fatores. Um deles é a capacidade de processamento dos servidores e a disponibilidade de recursos computacionais,

como CPU, memória e armazenamento. Se um sistema distribuído não estiver dimensionado corretamente para lidar com a carga de trabalho, o *throughput* pode ser comprometido.

A escalabilidade também é uma métrica relevante, avaliando a capacidade do sistema distribuído de lidar com um aumento na carga de trabalho. A escalabilidade pode ser horizontal, envolvendo a adição de mais nós ou recursos computacionais para lidar com uma demanda maior, ou vertical, que consiste em melhorar a capacidade de um nó existente para lidar com uma carga de trabalho mais intensa.

A disponibilidade refere-se à capacidade do sistema de estar operacional e acessível para os usuários durante o tempo desejado. Em ambientes de nuvem, onde os recursos estão distribuídos em vários servidores e data centers, a disponibilidade pode ser afetada por falhas de hardware, problemas de rede, atualizações de software ou até mesmo ataques cibernéticos. Para garantir alta disponibilidade, é necessário implementar redundância em diferentes níveis, como servidores, data centers e regiões geográficas. Isso significa que, se um componente ou servidor falhar, outros assumirão a carga de trabalho, garantindo a continuidade do serviço.

A confiabilidade, por sua vez, diz respeito à capacidade do sistema de fornecer resultados corretos e consistentes, mesmo em face de possíveis falhas ou interrupções. Isso envolve garantir a integridade dos dados, proteção contra perda ou corrupção de informações e a capacidade de recuperação em caso de falhas. Em sistemas distribuídos, a confiabilidade pode ser alcançada por meio de técnicas como replicação de dados, controle de transações, mecanismos de detecção e correção de falhas, backups regulares e planos de recuperação de desastres.

Por fim, a eficiência de uso de recursos é uma métrica que avalia o aproveitamento eficiente dos recursos computacionais disponíveis em um sistema distribuído, como CPU, memória e armazenamento. É importante garantir que os recursos sejam utilizados de forma otimizada, evitando ociosidade ou sobrecarga.

Uma abordagem eficiente em sistemas distribuídos é a distribuição equilibrada da carga de trabalho entre os diferentes componentes ou servidores. Isso pode ser alcançado por meio de técnicas como balanceamento de carga, onde as requisições são distribuídas inteligentemente, levando em consideração a capacidade e a carga atual de cada servidor. Dessa forma, evita-se sobrecarregar alguns componentes enquanto outros permanecem subutilizados.

Essas métricas são essenciais para a análise e monitoramento contínuo dos sistemas distribuídos em ambientes de nuvem, permitindo a otimização e o ajuste necessário para garantir um desempenho adequado, eficiência e confiabilidade, atendendo às demandas dos usuários e proporcionando uma experiência de qualidade.

4 FERRAMENTAS E METODOLOGIAS

4.1 Apache JMeter

De acordo com Lucas (2019), o JMeter é uma ferramenta de teste de desempenho gratuita e de código aberto que pode ser usada para testar a performance de aplicações web. É amplamente utilizada para análise de desempenho e testes de carga em sistemas distribuídos e aplicações web. Sua funcionalidade abrangente e flexibilidade o tornam uma escolha popular para avaliar a capacidade de sistemas em lidar com cargas de trabalho variáveis e situações de estresse.

O principal objetivo do Apache JMeter é simular cenários do mundo real, nos quais um grande número de usuários acessa uma aplicação ou sistema ao mesmo tempo. Isso ajuda a identificar gargalos, pontos fracos e possíveis problemas de desempenho, garantindo que a aplicação possa suportar a demanda do mundo real.

A ferramenta oferece diversos recursos para criação e execução de testes de carga. Os usuários podem criar planos de teste nos quais especificam os cenários, as ações dos usuários e as métricas a serem monitoradas. Os testes podem incluir a simulação de diferentes tipos de solicitações, como acessos a páginas, uploads de arquivos, submissões de formulários e muito mais.

O Apache *JMeter* suporta uma variedade de protocolos, incluindo *HTTP*, *HTTPS*, *FTP*, *JDBC*, *SOAP*, entre outros. Isso permite testar uma ampla gama de aplicações, desde websites até sistemas de banco de dados e serviços web.

Uma característica notável do *JMeter* é a capacidade de realizar testes distribuídos, nos quais várias instâncias da ferramenta podem ser executadas em diferentes máquinas, coordenando-se para criar uma carga de trabalho mais realista e intensa.

Além disso, o *JMeter* fornece recursos de relatórios detalhados, que incluem métricas como tempo de resposta, taxa de transferência, latência, número de erros e muito mais. Segundo Samuel Lucas (2019)O teste de escalabilidade se concentra na capacidade de um

sistema em atender a requisitos futuros de eficiência, que podem estar além dos requisitos exigidos atualmente sendo assim de suma importância, entre outros. Esses relatórios ajudam a identificar gargalos e anomalias de desempenho, permitindo que as equipes de desenvolvimento e operações tomem ações corretivas.

Em resumo, o Apache *JMeter* é uma ferramenta essencial para a análise de desempenho em sistemas distribuídos, especialmente em ambientes de nuvem. Sua capacidade de simular cenários de carga realista, suportar diversos protocolos e fornecer relatórios detalhados fazem dele uma escolha poderosa para garantir que as aplicações e sistemas possam lidar com cargas de trabalho reais e proporcionar uma experiência confiável aos usuários.

4.2 Prometheus

O *Prometheus* é uma ferramenta de código aberto projetada para monitorar e coletar métricas em ambientes de sistemas distribuídos e em nuvem. Segundo Evandro F. Souza (2018) O monitoramento de aplicações e servidores é uma parte importante do dia-a-dia do desenvolvedor de *software*. Isso inclui diversos tipos de análises, desde o monitoramento contínuo de possíveis exceções até o uso de CPU, memória e armazenamento do servidor. Ele se destaca por sua capacidade de coletar, armazenar e consultar métricas em tempo real, o que proporciona aos administradores e desenvolvedores *insights* valiosos sobre o desempenho e o estado dos sistemas complexos.

Uma das características distintivas do *Prometheus* é o seu modelo de dados multidimensional. Cada métrica coletada é identificada por meio de rótulos (*labels*) que descrevem suas origens e características específicas. Isso permite uma segmentação detalhada das métricas, facilitando a análise e a visualização de dados.

A ferramenta oferece flexibilidade na coleta de dados, permitindo a obtenção de métricas de várias fontes, como serviços, aplicativos e sistemas operacionais. Para facilitar isso, o *Prometheus* dispõe de agentes de coleta (*exporters*) específicos para tecnologias populares, tais como bancos de dados, servidores web e serviços em nuvem.

A consulta de dados é realizada por meio da linguagem *PromQL*, que possibilita aos usuários buscar e analisar métricas de maneira versátil. Essa funcionalidade é particularmente útil para criar painéis e gráficos personalizados que apresentam informações específicas de interesse.

Com recursos de alertas e notificações, o *Prometheus* permite que regras de alerta sejam definidas com base em condições específicas das métricas. Quando uma condição é atingida, a ferramenta dispara alertas, que podem ser integrados a sistemas de notificação como Slack, e-mails ou outras plataformas.

Sua integração natural com a plataforma Grafana amplia ainda mais suas capacidades, permitindo a criação de painéis de monitoramento visualmente atraentes e informativos.

Graças à sua escalabilidade e à comunidade ativa de desenvolvedores e usuários, o Prometheus é uma escolha robusta para monitorar e analisar o desempenho de sistemas distribuídos em ambientes de nuvem. Ele oferece uma maneira eficaz de coletar, armazenar e visualizar métricas em tempo real, facilitando a detecção e solução de problemas de maneira ágil e precisa.

4.3 Grafana

A Grafana, uma plataforma de visualização e monitoramento, desempenha um papel complementar ao Prometheus ao oferecer uma maneira intuitiva de criar painéis visuais e gráficos informativos a partir das métricas coletadas. Sua interface amigável e recursos poderosos permitem a transformação eficaz de dados em *insights* visuais claros e compreensíveis.

Um destaque da Grafana é sua capacidade de se conectar a diversas fontes de dados, como bancos de dados, sistemas de monitoramento e serviços em nuvem. Essa versatilidade possibilita a integração fluida de diferentes tipos de métricas e informações em um único painel.

Lucas ferreira da silva (2022) discute em seu artigo que a visualização de gráficos pelo Grafana por meio de dashboards é dinâmica e pode ser compartilhada com e por toda a equipe. É possível configurar alertas com base nas métricas, as quais são analisadas de forma contínua a fim de notificar o usuário conforme as regras por ele definidas. A plataforma oferece uma variedade ampla de tipos de painéis e gráficos, incluindo linhas, barras, tabelas, mapas e medidores. Essa diversidade possibilita a representação visual de diferentes conjuntos de dados de maneira eficaz e personalizada.

A personalização e configuração flexível são características-chave da Grafana. Os usuários podem ajustar a aparência, o layout e as configurações dos painéis para atender às

suas preferências. Além disso, é possível criar alertas personalizados com base em critérios específicos.

A Grafana também oferece recursos interativos, como a capacidade de explorar dados de forma dinâmica e realizar *drill-down* para obter *insights* detalhados. A plataforma suporta o compartilhamento de painéis, possibilitando a colaboração entre equipes e a comunicação eficaz dos *insights* obtidos.

A integração com alertas e notificações é um ponto forte da Grafana. Ela permite a configuração de alertas com base nas métricas dos painéis. Quando uma métrica ultrapassa um limite predefinido, a plataforma emite notificações para ação imediata.

Além disso, a Grafana é apoiada por uma comunidade ativa de desenvolvedores e oferece um marketplace com uma variedade de *plugins* e painéis personalizados. Esses recursos ampliam a funcionalidade da plataforma e possibilitam integrações com diferentes fontes de dados e sistemas.

Em resumo, a Grafana desempenha um papel crucial na análise e visualização de métricas coletadas por ferramentas como o *Prometheus* em ambientes de sistemas distribuídos em nuvem. Sua habilidade de transformar dados em painéis interativos e personalizados facilita a compreensão e a comunicação dos insights, auxiliando na identificação de tendências e tomadas de decisão informadas para otimização de desempenho.

5 RECOMENDAÇÃO DE ESTUDO FUTURO PARA TCC: AVALIAÇÃO DE DESEMPENHO EM SISTEMAS DE ARQUIVOS DISTRIBUÍDOS UTILIZANDO A FERRAMENTA IOzone.

Um estudo futuro relevante para um TCC é a avaliação do desempenho em sistemas de arquivos distribuídos, com foco no uso da ferramenta *IOzone*. Esta ferramenta é reconhecida por realizar *benchmarks* em sistemas locais e distribuídos. Ela oferece flexibilidade para executar várias operações de entrada/saída de dados e gera dados tabulados para facilitar a análise. O estudo poderia abordar a diversidade tecnológica dos sistemas de arquivos distribuídos, escalabilidade, confiabilidade, otimização de configuração, desenvolvimento de metodologias de teste e seus benefícios práticos para empresas. A pesquisa tem o potencial de beneficiar a comunidade acadêmica e a indústria de TI,

promovendo o desenvolvimento de sistemas de arquivos distribuídos mais eficientes e confiáveis.

6 METODOLOGIA

A metodologia de pesquisa empregada neste estudo foi a revisão bibliográfica, um método sistemático e estruturado para identificar, avaliar e consolidar informações pertinentes ao tópico central da revisão. A revisão bibliográfica visa minimizar qualquer viés potencial e garantir imparcialidade na coleta e análise dos dados.

Para a condução da revisão bibliográfica, foram estabelecidos os seguintes critérios para a inclusão e exclusão dos materiais a serem examinados. Os seguintes termos-chave específicos foram utilizados para delimitar a pesquisa: análise de desempenho, sistemas distribuídos, ambiente de nuvem, Gargados, Eficiência, Otimização.

A busca foi realizada em três engines de busca: Scielo, CAPES e Google Academic. Os resultados foram inicialmente filtrados através título e data de publicação, sendo de 2010 até 2023 nos respectivos idiomas português e inglês, selecionando-se apenas aqueles que diziam respeito ao tópico de interesse. Além disso, foram selecionados alguns artigos de sites para melhor exemplificação do conteúdo do estudo. Em seguida, os resumos de cada trabalho foram lidos para verificar se eles continham as informações necessárias. Por fim, os trabalhos que continham o que se procurava foram selecionados para leitura completa.

Os resultados foram os seguintes:

- Google Academic: 11.000 resultados
- CAPES: 52 resultados
- Scielo: 38 resultados

Após aplicado os devidos filtros Foram selecionados 17 artigos com as informações e relevâncias necessárias para validar a revisão bibliográfica.

A partir dos materiais selecionados, informações relevantes foram extraídas e os resultados foram sintetizados de maneira organizada.

Em resumo, a revisão bibliográfica foi a metodologia de pesquisa escolhida para este estudo, permitindo uma análise criteriosa e organizada das informações disponíveis sobre o assunto em foco. Essa abordagem proporciona uma base sólida para a construção de

conhecimento e a formulação de conclusões fundamentadas nas melhores fontes bibliográficas disponíveis.

7 CONSIDERAÇÕES FINAIS

Em um mundo cada vez mais dependente de sistemas distribuídos em ambientes de nuvem, a análise de desempenho desses sistemas torna-se crucial para garantir uma experiência satisfatória aos usuários finais. As métricas discutidas anteriormente desempenham um papel fundamental nesse contexto, permitindo uma avaliação abrangente e precisa do desempenho dos sistemas distribuídos.

A métrica de latência, por exemplo, desempenha um papel crucial em aplicações sensíveis ao tempo, como transações financeiras e jogos online. Reduzir a latência é essencial para proporcionar uma experiência fluida e responsiva aos usuários. Através da análise de latência, é possível identificar gargalos na rede, otimizar a comunicação entre os componentes distribuídos e melhorar a eficiência do sistema.

Já o *throughput*, que mede a quantidade de trabalho que um sistema distribuído pode realizar em um determinado período, é um indicador-chave da capacidade de processamento do sistema. Ao analisar o *throughput*, é possível identificar a capacidade máxima do sistema, bem como possíveis gargalos de processamento que podem afetar a eficiência e a escalabilidade do sistema.

A escalabilidade, por sua vez, é uma métrica essencial para sistemas distribuídos em nuvem. Ela mede a capacidade do sistema de lidar com um aumento na carga de trabalho, seja através da adição de mais recursos ou da melhoria dos existentes. Uma análise de escalabilidade permite identificar os limites do sistema e tomar decisões estratégicas para garantir que ele possa lidar com picos de demanda sem comprometer o desempenho.

Além disso, a disponibilidade e a confiabilidade são métricas críticas para sistemas distribuídos em nuvem. A disponibilidade refere-se à capacidade do sistema de estar operacional e acessível aos usuários, sem interrupções indesejadas. Já a confiabilidade diz respeito à capacidade do sistema de fornecer resultados corretos e consistentes, mesmo em face de possíveis falhas. A análise dessas métricas permite identificar pontos de falha, implementar estratégias de redundância e garantir uma alta disponibilidade e confiabilidade do sistema.

Essa pesquisa, que teve como objetivo apresentar as principais abordagens de análise de desempenho de sistemas distribuídos em ambientes de nuvem, chega ao seu término trazendo importantes considerações para a área de computação em nuvem.

Os autores destacaram a importância do desempenho eficiente de sistemas distribuídos na nuvem para manter a satisfação dos usuários e a competitividade das organizações. Eles abordaram conceitos de análise de desempenho, como a medição de escalabilidade, latência e *throughput* em ambientes distribuídos na nuvem, e discutiram ferramentas de monitoramento, como *Prometheus* e Grafana, que auxiliam na coleta de métricas e na análise de desempenho em tempo real. Esses elementos são cruciais para garantir a qualidade e eficiência desses sistemas na nuvem.

No entanto, a pesquisa destacou que a análise de desempenho não deve se limitar apenas a métricas técnicas, mas também deve considerar a experiência do usuário e a capacidade de fornecer serviços confiáveis e disponíveis.

Concluindo, é evidente que a análise de desempenho de sistemas distribuídos na nuvem é fundamental para atender às demandas atuais. Para pesquisas futuras, pode-se aplicar as metodologias e ferramentas discutidas em ambientes reais e desenvolver novas abordagens que considerem aspectos de qualidade do serviço e segurança. Espera-se que esta pesquisa possa ser utilizada como um guia para estudos adicionais na área de análise de desempenho de sistemas distribuídos em ambientes de nuvem.

Por fim, a eficiência de uso de recursos é uma métrica relevante para otimizar o aproveitamento dos recursos computacionais disponíveis no ambiente de nuvem. Ao analisar a eficiência de uso de recursos, é possível identificar oportunidades de otimização, como dimensionamento adequado de recursos, alocação eficiente de tarefas e melhor aproveitamento dos recursos disponíveis.

Em suma, a análise de desempenho em sistemas distribuídos em ambientes de nuvem é um processo complexo e abrangente, que envolve a consideração de diversas métricas. Através dessa análise, é possível identificar áreas de melhoria, otimizar o desempenho, garantir a disponibilidade e confiabilidade do sistema e proporcionar uma experiência de qualidade aos usuários. As métricas discutidas são ferramentas essenciais para avaliar, monitorar e aprimorar o desempenho dos sistemas distribuídos em ambientes de nuvem, contribuindo para o sucesso das organizações nesse cenário tecnológico em constante evolução.

REFERÊNCIAS

Barros, Pedro Gustavo. "O Que é um Sistema Distribuído?". R7, 06 de junho de 2013. Disponível em: <https://tecnologia.culturamix.com/dicas/o-que-e-um-sistema-distribuido>. Acesso em: 6 de abril de 2023.

Calegari, Gabriel. "Arquiteturas de Sistemas Distribuídos." Gabriel Calegari, 20 de janeiro de 2021. Disponível em: <https://calegari.dev/posts/arquiteturas-de-sistemas-distribuidos/>. Acesso em: 6 de maio de 2023.

Cika, Petr, e Vlastimil Clupek. "Stress Tester and Network Emulator in Apache JMeter." IEEE, 2019. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9017650/authors#authors>. Acesso em: 30 de agosto de 2023.

Gonçalves, Marcelo. **Sistema Distribuídos: Conceito e Definição**. 2022. Disponível em: <https://medium.com/@marcelomg21/sistemas-distribu%C3%ADdos-conceito-e-defini%C3%A7%C3%B5es-f2baa4efc88d>. Acesso em: 29 de março de 2023.

Meyer, Luiz. **Uma Visão Geral dos Sistemas Distribuídos de Cluster e Grid e suas Ferramentas Para o Processamento Paralelo de Dados**. In: XX CONGRESSO BRASILEIRO DE AUTOMÁTICA, 20, 2019, [Local da Conferência]. Anais eletrônicos... Disponível em: https://www.ibge.gov.br/confest_e_confega/pesquisa_trabalhos/CD/palestras/368-1.pdf. Acesso em: 10 de abril de 2023.

Neto, Eduardo. **Computação em Nuvem: O que é, Como funciona e a Importância**. 2019. Disponível em: <https://fia.com.br/blog/computacao-em-nuvem/>. Acesso em: 15 de maio de 2023.

Queiroz, Gabriel F.C. **Um estudo comparativo de simuladores para sistemas distribuídos**. 1ª ed., no. 1, 15 de junho de 2018, Rio de Janeiro, Rio de Janeiro, Brasil. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/cadinf/article/viewFile/25988/28251>. Acesso em: 9 de junho de 2023.

Schmeler, Elgio; GEYER, Fernando. **Escalonamento em Sistemas Distribuídos**. 1998. Disponível em: https://gravatai.ulbra.tche.br/~elgio/ufrgs/public_html/trabs-html/escalona/EscProc.html. Acesso em: 7 de abril de 2023.

Silva, Lucas Ferreira da. **QA Metrics: Integração das métricas de qualidade de software, em ambiente Docker, para exibição de dashboards Grafana alimentado pelo banco de dados temporal InfluxDB via Newman**. 2022. 40 f. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) - Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Recife, 2022.

Silva, F. H. R. **Um estudo sobre os benefícios e os riscos de segurança na utilização de Cloud Computing**. 2010. 15f. Artigo científico de conclusão de curso apresentado no Centro Universitário Augusto Motta, UNISUAM-RJ.

Souza, L. M. (2012). **Análise de desempenho de sistemas de processamento distribuído e web services**. Dissertação de Mestrado, Universidade Federal de São Carlos, São Carlos, SP. Disponível em:

<https://www.uece.br/wp-content/uploads/sites/51/2020/02/LEONARDO-MENEZES-DE-SO UZA.pdf> Acesso em: 24 de julho de 2023.

Taurion, Cezar. **Computação em nuvem: transformando o mundo da tecnologia da informação**. 1ª edição. Rio de Janeiro: Brasport Livros, 2009.

Thomé, Bruna, et al. “**Computação em Nuvem: Análise Comparativa de Ferramentas Open Source para IaaS**.” Vol. 1, no. 1.

http://hiperfcloud.setrem.com.br/wp-content/uploads/2017/03/THOME_ERRC_2013.pdf. Acesso em: 10 de junho de 2023.

Westphall, Carla. **Um Esquema de Autorização Para a Segurança em Sistemas Distribuídos de Larga Escala**. 2000. Disponível em:

<https://repositorio.ufsc.br/bitstream/handle/123456789/111744/182491.pdf?sequence=1>. Acesso em: 10 de abril de 2023.

Zettler, Kev. **O que é um Sistema Distribuído?**. 2022. Disponível em:

<https://www.atlassian.com/br/microservices/microservices-architecture/distributed-architecture>. Acesso em: 5 de abril de 2023.

Souza F. Evandro (2018). **Prometheus — Monitorando a saúde da sua aplicação**. Medium. Disponível em:

<https://medium.com/tech-grupozap/prometheus-monitorando-a-sa%C3%BAde-da-sua-aplica%C3%A7%C3%A3o-bd9b3b63e7b1>. Acesso em: 29 de setembro de 2023

Lucas, S. (2022, 14 de julho). **Primeiros passos com testes de performance e JMeter**. Medium. Disponível em:

<https://medium.com/@samuellucas/primeiros-passos-com-testes-de-performance-e-jmeter-a96b4db360ab> Acesso em: 29 de setembro de 2023.