

Por que AIOps? Desempenho da aplicação



Índice

- **Por que desempenho da aplicação?**
 - A Era da aplicação moderna
 - O serviço de hosting de aplicações modernas
 - O crescente ponto de pressão para o desempenho da aplicação
 - O compartilhamento eficaz de recursos de infraestrutura requer priorização
 - A obtenção de recursos eficaz da nuvem pública requer conhecimento da aplicação
- **Por que Turbonomic Application Resource Management for IBM Cloud Paks?**
 - O gerenciamento de recursos da aplicação (ARM) fornece desempenho e economia
 - Os novos requisitos para garantir o desempenho de aplicações modernas
- **Conclusão**
- **Sobre a Turbonomic, uma empresa IBM**

Por que desempenho da aplicação?

A Era da aplicação moderna

Na Era do aplicação moderna, em que mais de dois terços de todo o produto interno bruto (PIB) agora é digital,¹ o desempenho da aplicação é singularmente a mais alta prioridade para os CIOs porque a aplicação é o negócio. O fornecimento de aplicações é o motivo principal pela TI existir. O CIO, que comanda a TI, não tem opção a não ser fornecer desempenho da aplicação para garantir que o negócio nunca seja restringido pela TI. Na verdade, o CIO é considerado um fracasso ao não fornecer um desempenho da aplicação; ironicamente, é muito razoável para os CIOs exceder seus orçamentos planejados.

Em suma, falha em garantir desempenho da aplicação prejudica o negócio. O que é desafiador é que as aplicações que congestionam primeiro e por mais tempo são aquelas com os maiores picos de demanda, sendo essas frequentemente as aplicações mais valiosas.

O serviço de hosting de aplicações modernas

O investimento no desenvolvimento de aplicações excede em muito o custo de hosting dessas aplicações. Com esse investimento relativo, as empresas, consciente e voluntariamente, superprovisionam suas infraestruturas e ambientes em nuvem para minimizar o risco de desempenho da aplicação. Recursos de infraestrutura, tanto em datacenters quanto na nuvem, estão rapidamente se tornando comodidades descartáveis e, portanto, uma justificação de retorno sobre investimento (ROI) menos confiável. Além disso, o mau desempenho da aplicação cria desconfiança entre os proprietários de aplicações da linha de negócios (LOB) e entre as equipes de ITOps e CloudOps que fornecem o serviço de hosting de aplicações. Imagine as linhas de negócios investindo milhões de dólares com centenas de pessoas desenvolvendo novas experiências de usuário, colocando sua reputação em risco, apenas para que as equipes de operações ofereçam aos usuários finais uma experiência lenta de "roda girando". Todo aquele investimento para melhorar a experiência do usuário final é perdido.

O crescente ponto de pressão para o desempenho da aplicação

Escassez de aplicação ocorre quando a infraestrutura, local ou na nuvem, não consegue atender à demanda por aplicação

e seus usuários finais. Escassez de recursos de infraestrutura é a causa mais frequente de degradação de desempenho da aplicação. Em contraste, a arquitetura de código da aplicação, monitorada com as ferramentas de gerenciamento de desempenho da aplicação (APM), é raramente inferior a 10% da origem da degradação de desempenho da aplicação em ambientes de produção. Além disso, dado o foco do desenvolvimento de aplicação na qualidade do código ao usar processos aprimorados, como integração contínua, entrega contínua (CICD), controle de qualidade (QA), pré-produção e preparação, o impacto da qualidade do código no desempenho está se tornando cada vez mais raro.

Compartilhamento eficaz de recursos de infraestrutura requer priorização

No datacenter on-premises, aplicações sofrem escassez quando a demanda por recursos de infraestrutura compartilhada permanece não priorizada. Dada esta limitação, as cargas de trabalho de aplicações são tipicamente superprovisionadas e colocadas sem suficiente entendimento e contexto em relação aos recursos disponíveis. Ao compartilhar recursos, todas as aplicações tomam de um conjunto comum de recursos independentemente da utilização. Alocação de recursos superprovisionada ou mal dimensionada resulta em congestionamento constante, levando a violações do acordo de nível de serviço (SLA), resolução manual de problemas ineficiente, ajustes perpétuos de recursos e, conforme anteriormente observado, investimento não utilizado em desenvolvimento de aplicações.

As ferramentas de monitoramento reativas e de recurso único de hoje falham em entender o relacionamento entre aplicações e infraestrutura e, portanto, dependem da interpretação e da intervenção manual para solucionar o congestionamento de recursos. Quanto mais tempo se leva para oferecer uma resolução, mais provável a decisão sobre o que fazer será obsoleta.

Uma alternativa é usar continuamente a demanda dinâmica por aplicações para priorizar a alocação da solução completa de recursos de infraestrutura compartilhada. À medida que a curva de demanda por uma aplicação sobe, sua prioridade relativa de adquirir e preservar recursos aumenta. À medida que esta curva de demanda desce, sua prioridade relativa de adquirir e preservar recursos diminui. Essa redefinição de prioridade contínua da solução completa de recursos compartilhados minimiza a escassez, permite o compartilhamento fluido de recursos entre todas as aplicações e garante o desempenho da aplicação.

Obtenção de recursos eficazes na nuvem pública requer conhecimento da aplicação

Na nuvem pública, aplicações sofrem escassez quando instâncias e recursos permanecem insuficientes para atender a demanda por aplicação. Para abordar esse risco, os administradores da nuvem muitas vezes superprovisionam instâncias. Em ambos os casos, o fornecimento incorreto de instâncias resulta do conhecimento limitado do administrador da nuvem sobre os requisitos de recursos da demanda por aplicação. Além disso, desenvolvedores que focam na construção de aplicações estão tipicamente desinteressados na criação de perfis e na previsão de seus requisitos de recurso. Como resultado, estimativas aproximadas são feitas enquanto o provedor em nuvem coloca o risco, a carga e as consequências dessa estimativa no cliente.

Agravando a complexidade, a decisão sobre a instância ideal está cada vez mais desafiadora uma vez que um único fornecimento de instância do Amazon Elastic Compute Cloud (Amazon EC2) tem milhões de opções de configuração ao decidir o tipo de recurso, hardware subjacente, dimensionamento, geografias, precificação, reservas, planos de economia e assim por diante.

Além disso, alterações dinâmicas na demanda por aplicações requerem uma reavaliação contínua desta estimativa aproximada utilizada para selecionar instâncias de nuvem. Por fim, o gerenciamento da elasticidade de instâncias da aplicação em nuvem pública requer calibração contínua em tempo real, especialmente com a adoção de aplicações com base em contêiner de vida curta. Como resultado, a maioria das alocações de instâncias na nuvem pública é superprovisionada e gerenciada manual e reativamente como uma proteção contra riscos de desempenho e com uma completa falta de entendimento da demanda por recurso da aplicação. Além disso, o provedor em nuvem com base em monitoramento e as ferramentas da plataforma de gerenciamento de nuvem não têm o entendimento sobre a demanda por aplicação para garantir o desempenho da aplicação na nuvem. Eles estão restritos à visibilidade dos custos, do histórico de faturamento e da alocação de custos departamentais, nada relacionado a garantir o desempenho de aplicações em execução em uma nuvem pública.

Por que Turbonomic Application Resource Management for IBM Cloud Paks?

Turbonomic, uma empresa IBM, fornece a sustentação do serviço de hosting de aplicações modernas ao usar exclusivamente um entendimento da demanda por aplicação para fornecer ações de obtenção de recursos e de análise de desempenho contínuas para ajudar a garantir o desempenho da aplicação em tempo real, ao longo do tempo, 24x7x365.

O Turbonomic Application Resource Management for IBM Cloud® Paks usa um modelo de dados comum para que os clientes possam gerenciar com confiança suas aplicações hoje, bem como as aplicações modernas do futuro, independentemente se elas são executadas on premises, em nuvens públicas ou na borda.

O gerenciamento de recursos de aplicação (ARM) fornece desempenho e economia

O ARM da Turbonomic, usando análise de dados orientada por IA, ajuda a garantir o desempenho da aplicação por continuamente corresponder em tempo real aos requisitos de demanda da aplicação com tipos e tamanhos de recursos. Ações de obtenção de recursos incluem iniciar e parar, colocação inicial e contínua, ajuste de escala e redimensionamento.

Por contraste, centenas de ferramentas, e até mesmo planilhas, pretendem economizar o dinheiro do cliente, mas suas "recomendações" são muitas vezes baseadas em simples alertas de níveis que podem causar problemas de desempenho. Uma pesquisa de opinião recente relatou que 39% dos profissionais de operações financeiras (FinOps) assinalaram "fazer com que engenheiros tomem ação" como o seu maior desafio.² O motivo? A falta de confiança nas ferramentas existentes usadas para gerar essas "ações".

Os novos requisitos para garantir o desempenho de aplicações modernas

A plataforma AIOps da Turbonomic utiliza um modelo de dados comum abrangente que ingere, normaliza e gerencia todos os recursos compartilhados nos quais o desempenho da aplicação depende. Mais importante ainda, ela desenvolve a topologia do relacionamento da cadeia de suprimento, "pontuando", entre cada dependência de obtenção de recursos, da aplicação à infraestrutura.

Em contraste, as ferramentas manuais e fragmentadas não garantem o desempenho uma vez que elas monitoram recursos em isolamento e têm limitada perspectiva da aplicação, se alguma. O desempenho da aplicação requer o entendimento e o gerenciamento de todas as dependências de recursos em precisamente as quantidades, ordem e prazo ideais.

Conclusão

Na era da aplicação moderna, a aplicação é o negócio. À medida que as aplicações se tornam mais complexas, com mais dependências e ambientes mais diferentes e distribuídos, o risco para o desempenho da aplicação e a experiência do usuário cresce exponencialmente. A única maneira de abordar esses desafios é implementar uma abordagem focada na aplicação que avalie continuamente a demanda por aplicação, o fornecimento dos recursos disponíveis e gera recomendações acionáveis nas quais tanto os proprietários de operações quanto de aplicações podem confiar. Ao longo do tempo, à medida que a confiança aumenta, a próxima fase é automatizar essas ações. O resultado será aplicações de alto desempenho, enorme redução em gastos com TI e a oportunidade de despertar a inovação do negócio. Essa é a mágica do AIOps.

Mais de dois terços do PIB global agora é digital.¹

- A entrega de aplicações é o motivo principal pela TI existir.
- Os CIOs devem assegurar que o avanço do negócio nunca seja restringido pela TI.
- O desenvolvimento da aplicação é 3 vezes o custo do hosting de aplicações.
- As empresas superprovisionam para minimizar o risco de degradação de desempenho da aplicação.
- Escassez de recursos de infraestrutura é a origem mais frequente de degradação de desempenho da aplicação.
- Aplicações sofrem escassez quando a demanda por recursos de infraestrutura compartilhada não é priorizada.
- As ferramentas de monitoramento focadas em recursos de hoje, sem um entendimento da demanda por aplicações, podem somente reagir a situações negativas.
- Redefinição de prioridades contínua da solução completa de recursos compartilhados, com base na demanda por aplicação, é a única maneira de assegurar desempenho da aplicação.

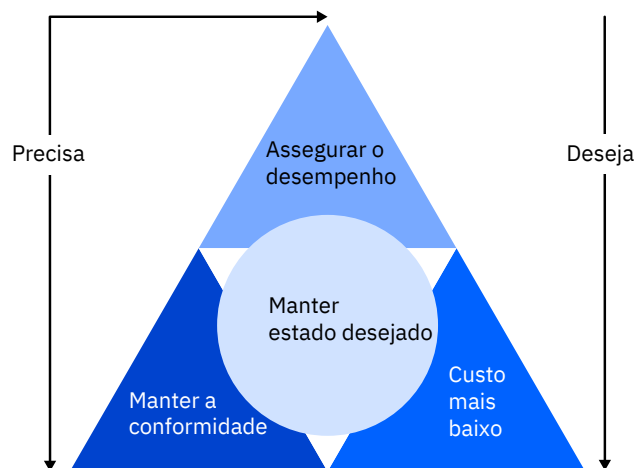


Figura 1. As aplicações estão no estado desejado quando o desempenho é garantido ao mesmo tempo que a conformidade é mantida no custo mais baixo.

- Administradores de nuvem têm limitado conhecimento a respeito dos requisitos de recursos de uma aplicação.
- Desenvolvedores de aplicações estão focados na lógica de negócios e adiam decisões sobre recursos para a equipe de TI.
- Escolher os recursos ideais na nuvem pública é complexo com milhões de opções de configuração.
- A maioria das seleções de instância da nuvem é superprovisionada para minimizar o risco de degradação de desempenho.
- O dimensionamento ideal de instâncias em nuvem pública requer conhecimento da demanda por aplicação.

Sobre a Turbonomic, uma empresa IBM

A Turbonomic, uma empresa IBM, fornece o software Application Resource Management (ARM), que é utilizado por clientes para ajudar a garantir o desempenho e o controle da aplicação ao obter recursos de aplicações dinamicamente em ambientes híbridos e multinuvem. O gerenciamento de desempenho de rede (NPM) da Turbonomic fornece soluções modernas de monitoramento e análise de dados para ajudar a garantir o contínuo desempenho da rede em escala entre as rede de diversos fornecedores para empresas, operadoras e provedores de serviços gerenciados.

Para obter informações adicionais, acesse ibm.com/cloud/turbonomic.

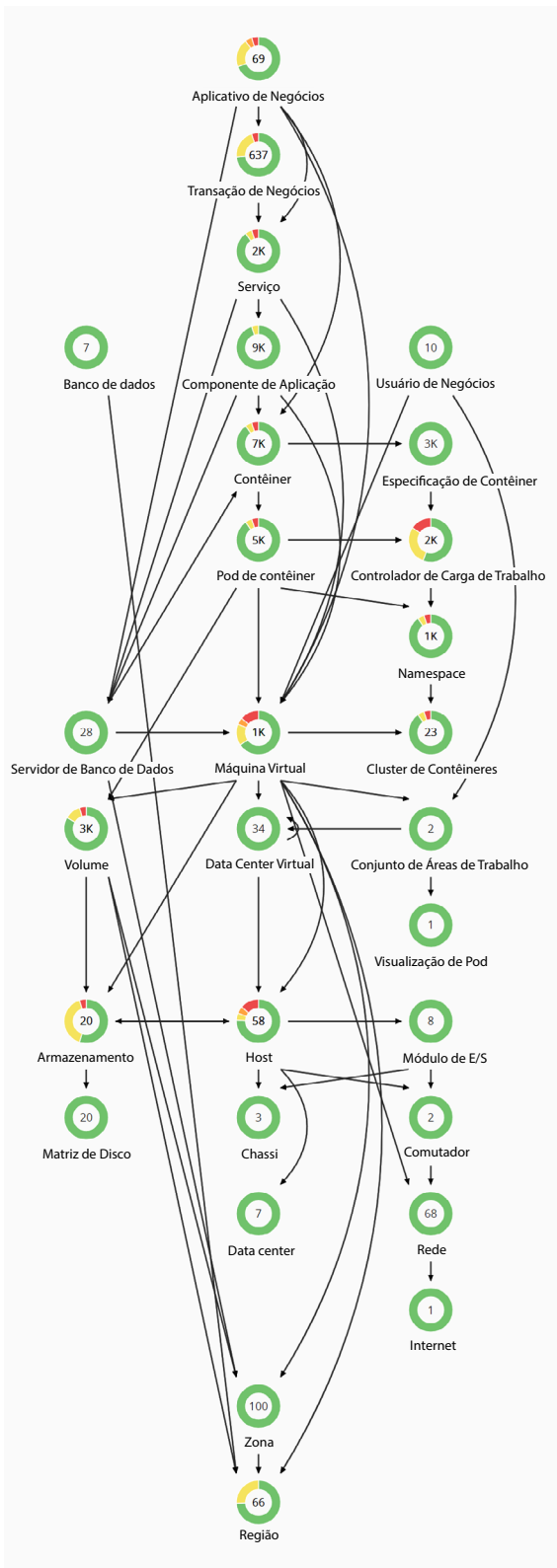


Figura 2. A Turbonomic desenvolve automaticamente a topologia de relacionamento da cadeia de suprimento.

© Copyright IBM Corporation 2021

IBM Brasil Ltda
Rua Tutóia, 1157
CEP 04007-900
São Paulo – SP
Brasil

Produzido nos Estados Unidos da América
Novembro de 2021

IBM, o logotipo IBM e IBM Cloud são marcas comerciais ou registradas da International Business Machines Corporation nos Estados Unidos e/ou em outros países. Outros nomes de serviços e produtos podem ser marcas comerciais da IBM ou de outras empresas. Uma lista atual de marcas comerciais da IBM está disponível em: ibm.com/trademark.

Turbonomic é uma marca registrada da Turbonomic Inc., uma empresa da IBM.

Este documento estava atualizado na data de publicação inicial e pode ser alterado pela IBM a qualquer momento. Nem todas as ofertas estão disponíveis nos países onde a IBM opera.

AS INFORMAÇÕES NESTE DOCUMENTO SÃO OFERECIDAS NO ESTADO EM QUE SE ENCONTRAM ("AS IS") SEM QUALQUER GARANTIA, EXPLÍCITA OU IMPLÍCITA, INCLUINDO SEM QUAISQUER GARANTIAS DE COMERCIALIZABILIDADE, ADEQUAÇÃO A UM PROPÓSITO ESPECIAL E QUALQUER GARANTIA OU CONDIÇÃO DE NÃO VIOLAÇÃO. Os produtos IBM são garantidos de acordo com os termos e condições dos acordos sob os quais são fornecidos.

¹ IDC Reveals 2021 Worldwide Digital Transformation Predictions, IDC, 29 de outubro de 2020.

² State of FinOps Report 2021, The FinOps Foundation, 2021.