

Search. Observe. Protect.

Uma introdução ao APM

O que, por que e como.

elastic.co/pt



Índice

Por que precisamos de monitoramento de performance de aplicação (APM)?	3
As aplicações são a força motriz das empresas	3
Baixo desempenho equivale a perda de receita	4
As modernas aplicações nativas na nuvem são complexas	4
O que exatamente é o APM?	5
Visibilidade de ponta a ponta	5
Mostrar latência e erros	6
O APM além da produção	6
Como o APM funciona	7
Agentes de APM	7
Instrumentação e configuração	8
Análise	8
Principais termos	9
Rastreamento distribuído	9
Spans	10
Transações	10
Traces	11
Monitoramento de usuário real	11
Como escolher uma ferramenta de APM	12
Recursos técnicos	12
Facilidade de uso	13
Opções de implantação	13
Suporte para padrões abertos/dados abertos	13
Arquitetura e escalabilidade	14
Segurança	14
Recursos além do APM	15
Preços	15
Sobre a Elastic	16

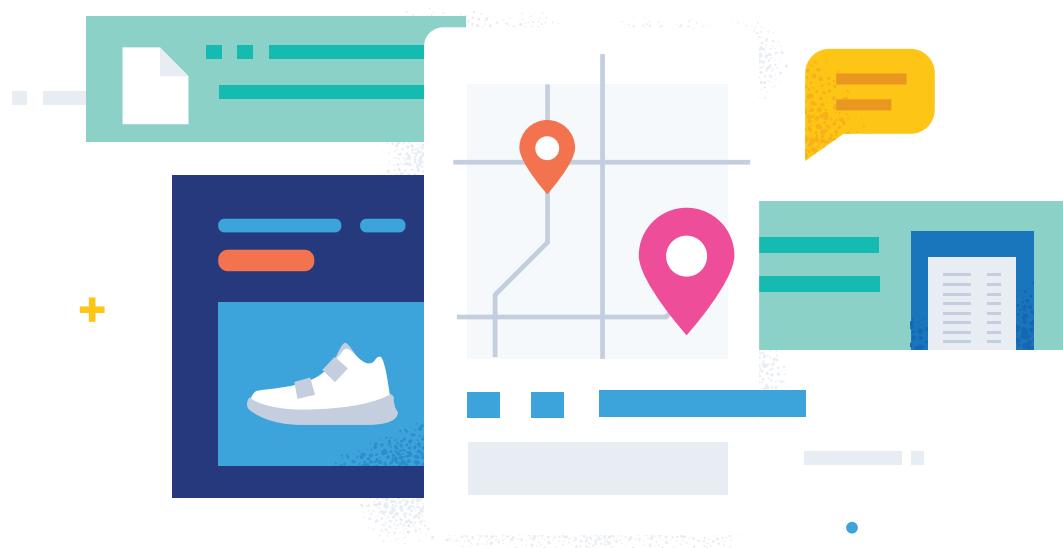


Por que precisamos de monitoramento de performance de aplicação (APM)?

Uma observação rápida: ao longo deste guia, abordaremos o [monitoramento de performance de aplicação \(APM\)](#) como um pilar da [observabilidade](#). Junto com os logs e as métricas, o APM desempenha um papel crucial na construção de sistemas observáveis.

As aplicações são a força motriz das empresas

As aplicações são as faces públicas das organizações modernas. É por meio delas que interagimos com produtos e serviços, seja uma loja de comércio eletrônico, um app de viagens compartilhadas ou as várias ferramentas de colaboração e produtividade que usamos diariamente. Quando adicionamos um par de fones de ouvido aos nossos carrinhos de compras digitais ou criamos a lista de reprodução perfeita para um jantar, estamos interagindo com aplicações e formando opiniões sobre elas. A percepção, a fidelidade e as avaliações da marca baseiam-se nessas experiências.



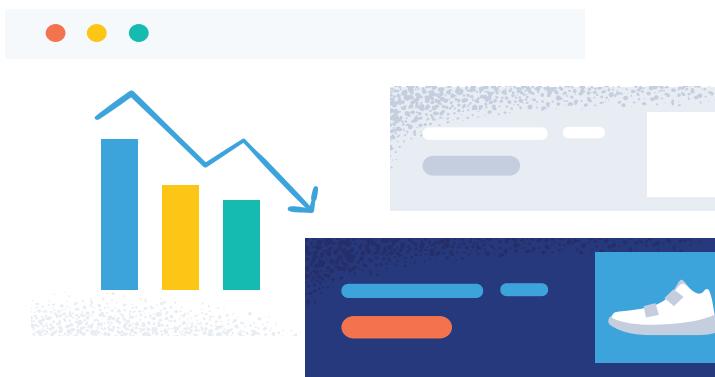
Baixo desempenho equivale a perda de receita

Pense nas aplicações que você usa para trabalhar — ou nos seus dispositivos pessoais em casa. O que vem à mente? Como é a facilidade de uso? Com que rapidez eles respondem? Com que frequência ocorrem erros? As respostas a essas perguntas influenciam a sua percepção sobre as empresas responsáveis por essas aplicações?

Quando as aplicações não funcionam conforme o esperado, há impactos diretos nos negócios. A receita e a confiança do cliente estão em jogo.

Se a experiência de um carrinho de compras digital for interrompida, os usuários irão para outra loja, resultando em vendas perdidas. Se houver lentidão constante nas respostas ou se os erros forem frequentes, até mesmo usuários fiéis poderão decidir parar de usar o produto ou serviço. Cada organização sabe o que o tempo de inatividade significa para seus resultados financeiros.

O tempo médio de detecção (MTTD) e o tempo médio de resposta (MTTR) têm seu preço.



As modernas aplicações nativas na nuvem são complexas

A evolução no desenvolvimento e na entrega das aplicações, aliada à complexidade resultante dessa mudança, impõe a necessidade de um monitoramento abrangente do desempenho com mais urgência.

A adoção de arquiteturas de microsserviços mudou efetivamente a maneira como as aplicações eram construídas e mantidas, oferecendo entrega contínua e uma escala que antes era inimaginável. Mas a natureza distribuída e poliglota (escrita usando diferentes linguagens/frameworks) dessas arquiteturas gera mais complexidade. A entrega contínua e os princípios de automação possibilitaram envios frequentes de código, mas tornaram a capacidade de acompanhar o impacto no desempenho mais crítica para que os problemas possam ser corrigidos rapidamente ou as alterações sejam revertidas se uma implantação não sair conforme o planejado.

Os princípios de desenvolvimento nativos da nuvem permitem que as equipes forneçam excelentes experiências digitais aos usuários continuamente, reajam rapidamente ao feedback e façam ajustes conforme o necessário. O que os usuários não veem são todas essas engrenagens trabalhando em segundo plano para atender a cada uma de suas solicitações. Então, como as equipes ficam de olho nessas aplicações?

Como humanos, não podemos processar esse volume de dados de forma significativa. Precisamos de uma visão organizada com um caminho claro para o insight. É aí que entra o APM.

O que exatamente é o APM?

APM é a solução para coletar, monitorar e analisar o desempenho de ponta a ponta e o comportamento das transações por meio das suas aplicações (e dos serviços dos quais elas dependem). O APM acompanha a experiência do usuário desde o momento em que ele começa a interagir com a aplicação até quando ele atinge os resultados desejados.

Por causa de seus vínculos diretos com a experiência do usuário, o APM ajuda a conectar a TI às metas de negócios. A relação pode não ser exatamente de 1 para 1, mas tem uma correlação positiva. Ao fornecer um insight contínuo do desempenho das suas aplicações, o APM permite que você seja proativo, em vez de reagir aos problemas após uma chuva de reclamações dos clientes no seu sistema de tíquetes ou até mesmo nas redes sociais.

Visibilidade de ponta a ponta

Depois que uma aplicação é instrumentada, os dados resultantes informam exatamente o que está acontecendo dentro dela. O APM acompanha as transações durante toda a jornada de forma que cada solicitação e resposta seja registrada e medida, independentemente do grau de complexidade da arquitetura.

Interligando todos esses trases, o APM oferece uma visão completa do desempenho da sua aplicação — desde uma visão panorâmica dos seus serviços (e de como eles estão interagindo) até insights no nível do código.



Mas espere aí: nós não temos logs para nos dizerem o que está acontecendo dentro das nossas aplicações? Os logs realmente nos fornecem dados valiosos com informações contextuais sobre os tipos de erros ou outros eventos que ocorrem. O que os logs não mostram é a jornada de ponta a ponta da experiência do usuário, que é onde os trases do APM (e o rastreamento distribuído) preenchem os detalhes. A combinação dos trases com os logs e as métricas proporcionará uma visibilidade unificada do seu ecossistema.

Mostrar latência e erros

O APM monitora dois atributos em torno da experiência do usuário: a duração da transação e os erros. Quanto tempo normalmente levam as solicitações HTTP? Quantos erros 500 costumamos ver ao longo de uma hora?

Problemas de latência ocorrem quando uma resposta a uma solicitação está demorando mais do que o normal. Um exemplo: um usuário clica para ver um produto, e a resposta leva mais do que alguns segundos para carregar.

Os erros, é claro, significam que ocorreu um resultado indesejado. Isso acontece quando a solicitação não é concluída com sucesso.

O objetivo final de qualquer investigação é corrigir os problemas. Para fazer isso de forma eficaz, as equipes realizam análises de causa raiz para determinar exatamente qual componente está causando o problema. Podemos identificar rapidamente qual serviço é o gargalo de desempenho? Podemos encontrar o método ou a chamada de função exato(a) que é responsável por esse problema? Com o APM, a resposta para ambas as perguntas é “sim”.

O APM além da produção

Costumamos falar sobre o APM no contexto da experiência do usuário com aplicações em produção, mas o APM entra em ação antes disso, com impactos em várias equipes e ambientes. Por exemplo, o APM fornece aos desenvolvedores o feedback de que precisam para criar o código que leva a um ciclo de desenvolvimento-teste-implantação mais rápido.

As empresas estão usando o APM em todos os estágios para ajudar as equipes a desenvolver, depurar, implantar e monitorar aplicações com mais eficiência. Se as organizações usarem o APM em muitas equipes, poderão melhorar os processos e criar fluxos de trabalho mais eficientes, resultando em mais tempo para inovação (e menos tempo apagando incêndios).

Já vimos os benefícios; agora vamos ver como o APM funciona em detalhes.

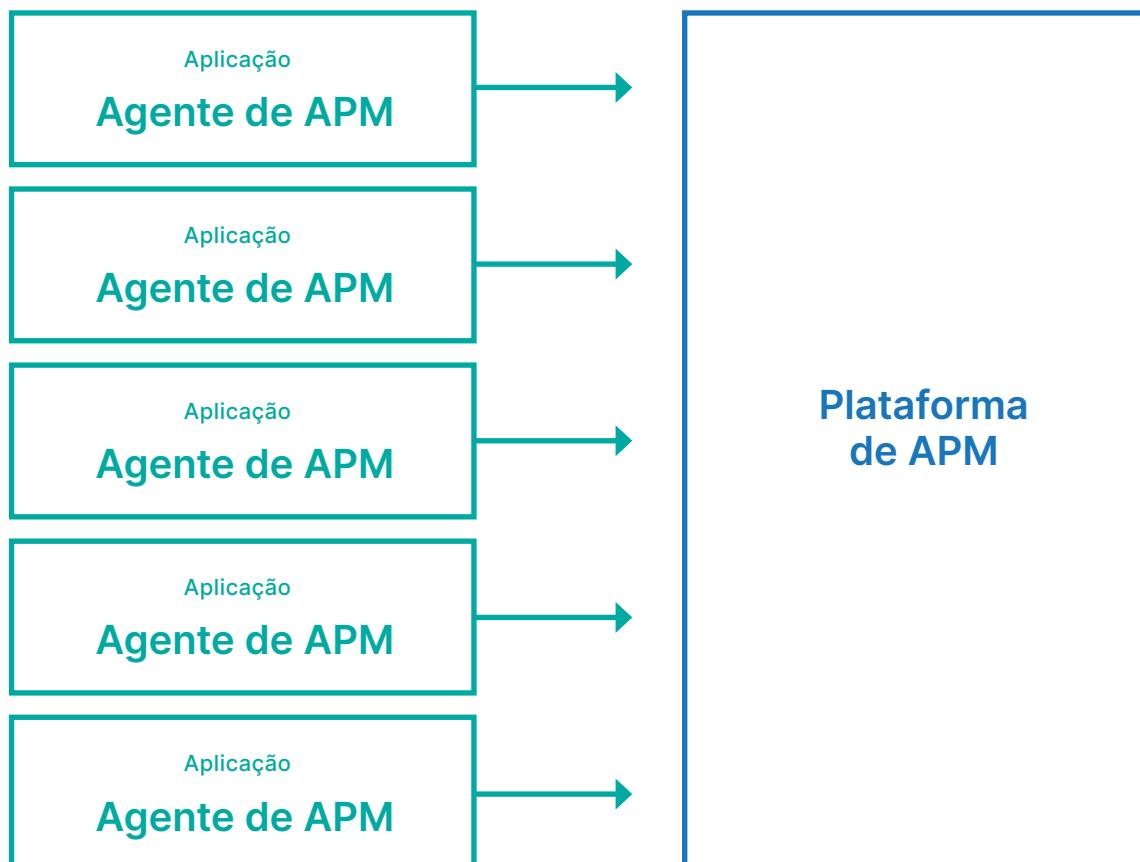
Como o APM funciona

O APM instrumenta cada aspecto do seu sistema que você gostaria de monitorar (hosts, containers, aplicações etc.) e depois envia os dados de desempenho coletados ao local de sua escolha para análise.

Agentes de APM

Um agente de APM é uma biblioteca, plugin ou extensão que monitora as métricas de desempenho descritas acima. Dependendo do que você precisa monitorar (e da linguagem na qual está escrito), pode ser necessário um ou vários agentes. Depois de identificar tudo o que gostaria de monitorar, você implantará agentes de APM em cada uma dessas partes. Embora os agentes variem de acordo com o fornecedor, a maioria dos agentes instrumenta o seu código, coleta dados de desempenho e envia esses dados para um servidor ou coletor.

Servidor de borda



Instrumentação e configuração

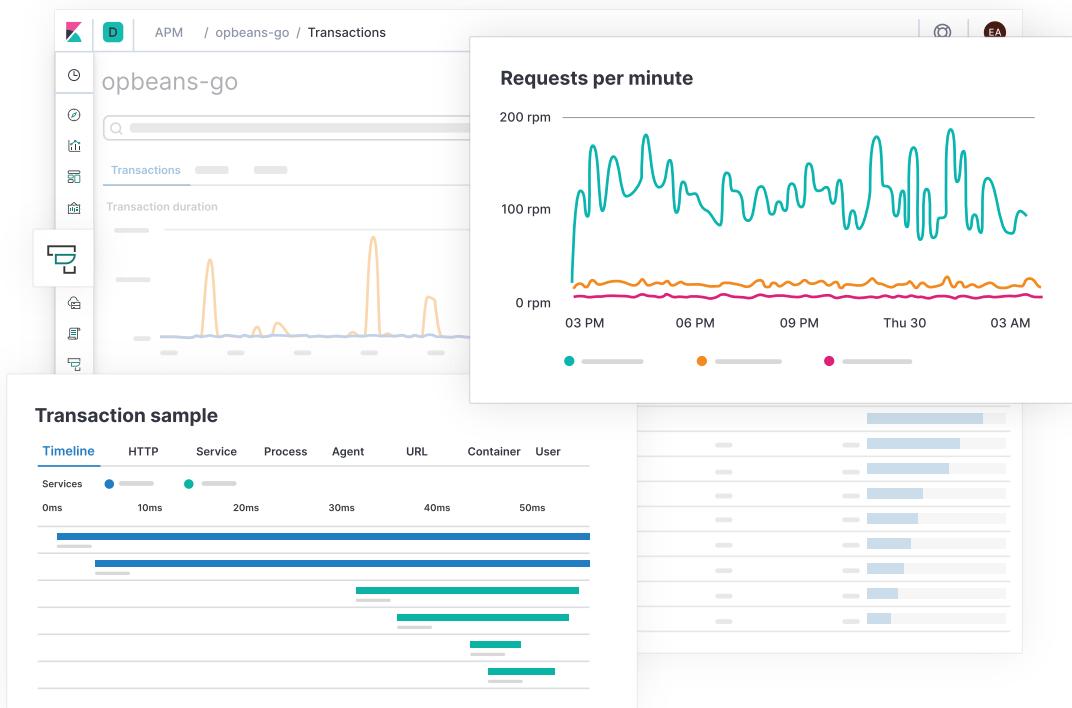
Instrumentação é o processo de estender o código da sua aplicação para relatar dados de trases. Após instalar um ou mais agentes necessários, você pode instrumentar as partes específicas da sua aplicação que gostaria de analisar para coletar e enviar dados de desempenho para o local de sua escolha, geralmente um endpoint determinado pela sua ferramenta de APM.

Você provavelmente vai querer configurar coisas como nomes de ambiente, taxas de amostragem, instrumentações e métricas para ajudar suas equipes a identificar e analisar facilmente os dados que estão sendo transmitidos para a sua ferramenta de APM. Geralmente, isso é feito usando a UI e a API da ferramenta ou diretamente nas variáveis do ambiente.

Análise

Depois que os dados de desempenho são enviados para o local de sua escolha, você já pode analisá-los. A maioria das ferramentas tem uma UI que ajudará a identificar erros, problemas de latência e outras anomalias que estão afetando seus usuários.

Esse é um ótimo lugar para começar a sua investigação sobre os problemas relatados ou, melhor ainda, detectar os problemas antes que os usuários sejam afetados por eles. Identifique precisamente quais serviços (até o nível do código) estão enfrentando problemas para acelerar a análise de causa raiz e baixar o MTTR.





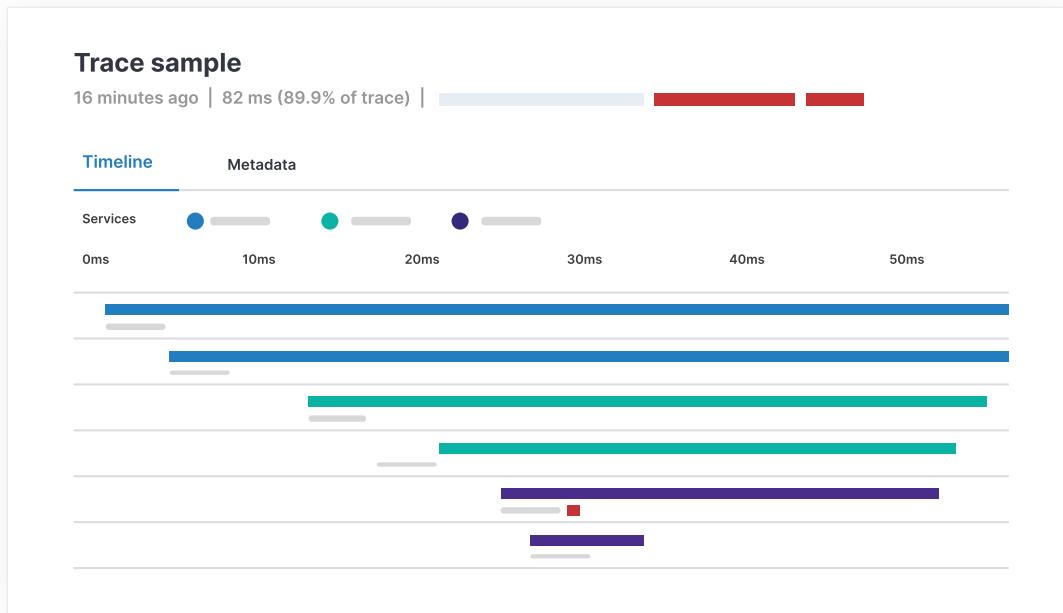
Principais termos

Cada ferramenta de APM pode ter definições ligeiramente diferentes para esses termos, mas, em geral, elas capturam as mesmas informações para fornecer ao usuário um insight sobre o desempenho de suas aplicações.

Rastreamento distribuído

Ao rastrear todas as solicitações, desde a solicitação inicial da Web até o serviço de frontend e as consultas feitas aos serviços de backend, o rastreamento distribuído permite que você analise o desempenho em toda a sua arquitetura de microserviços em uma única visualização.

O rastreamento distribuído facilita a detecção de gargalos, exibindo eventos completos por serviço e depois por solicitação dentro desse serviço. A detecção de erros e outros problemas de maneira prática no nível do código acelera as investigações e o MTTR.



Spans

Cada unidade ou parte do fluxo de trabalho é chamada de span. Os spans são o que você normalmente veria na visualização em cascata de uma ferramenta de análise de APM, geralmente representada na forma de barras horizontais. Esses segmentos são o elemento central do rastreamento distribuído. Os spans medem do início ao fim de uma atividade e contêm informações sobre a execução de um caminho de código específico.

Atributos comuns de um span:

- Hora de início
- Hora de término
- Um nome
- Um tipo

Transações

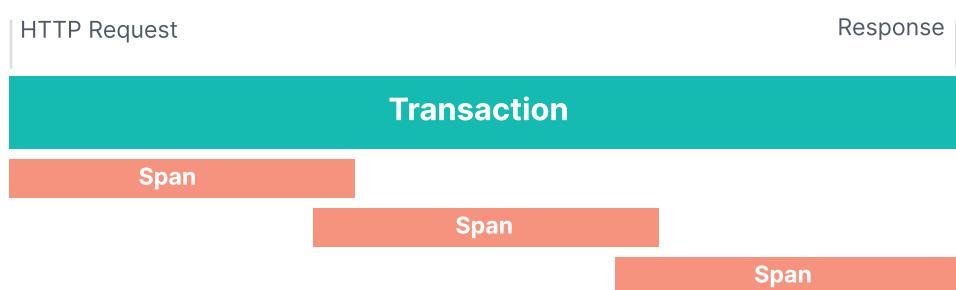
As transações são um tipo de span que descreve um evento e também pode incluir vários spans.

Alguns exemplos de transações:

- Uma solicitação para o seu servidor
- Um trabalho em lote
- Um trabalho em segundo plano

As transações têm atributos adicionais associados a elas, como dados sobre o ambiente no qual o evento é registrado:

- Serviço: ambiente, framework, linguagem etc.
- Host: arquitetura, nome do host, IP etc.
- Processo: args, PID, PPID etc.
- URL: completa, domínio, porta, consulta etc.
- Usuário: (se fornecido) e-mail, ID, nome de usuário etc.

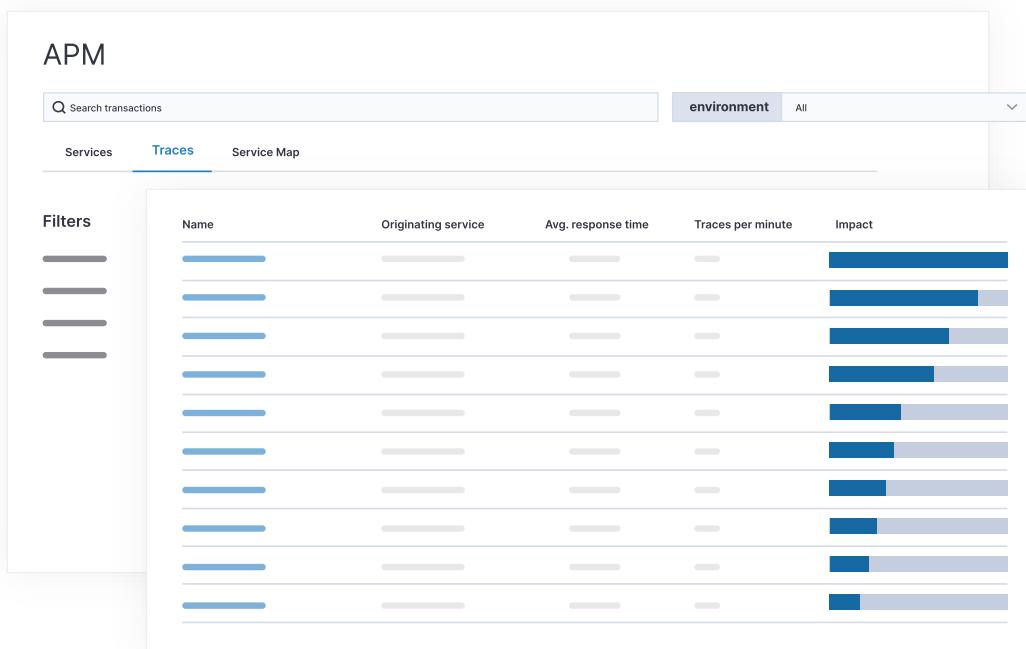


Traces

Os traces medem o status e a duração de uma solicitação feita por meio da sua aplicação. Juntos, spans e transações formam um trace, ligando os pontos desde uma solicitação do usuário até a resposta.

Você verá traces individuais para ações completas, como processamento de um pagamento, processamento de um pedido concluído e atualização do status de envio.

Quando combinados com os logs e as métricas da aplicação (e outros aspectos da sua infraestrutura), os traces proporcionam visibilidade completa de todo o seu ecossistema.



Monitoramento de usuário real

O monitoramento de usuário real (RUM) captura a interação direta do usuário com os clientes em ambientes de produção. Por exemplo, o RUM mostra exatamente como um usuário interagiu com o seu website, juntamente com os tempos de resposta e os erros incorridos. Isso é realizado por um agente no navegador do usuário que coleta os dados de desempenho e os envia para a ferramenta de APM para análise.

Como escolher uma ferramenta de APM

Como em qualquer processo de seleção de ferramentas, comece definindo claramente os seus requisitos.

Ao listar os requisitos, leve em consideração seu crescimento e necessidades futuras. Planeje uma prova de conceito (POC) para garantir que a ferramenta realmente atenda às suas necessidades. Os recursos demonstrados e divulgados pelo fornecedor funcionam conforme o esperado no seu ambiente? Um POC ajuda a evitar surpresas e custos ocultos posteriormente no processo.

Recursos técnicos

Crie uma lista de verificação dos recursos necessários pensando nas suas necessidades técnicas específicas. Seja abrangente e detalhado(a) em seus requisitos. Embora esta não seja uma lista exaustiva, aqui estão algumas áreas a serem consideradas:

Supporte para linguagens e frameworks

- A ferramenta de APM dá suporte para as linguagens e os frameworks amplamente implantados no seu ambiente?

Amplitude e profundidade da visibilidade

- Ela fornece visibilidade do desempenho no nível do código?
- Ela dá visibilidade de ponta a ponta da sua stack de tecnologia?

Recursos de busca, analítica e exploração visual

- Ela inclui uma UI organizada para dar suporte a fluxos de trabalho investigativos, tanto proativos quanto reativos?
- Ela tem ferramentas visuais, como gráficos em cascata e mapas de dependência, para analisar o desempenho de aplicações distribuídas?
- Ela fornece uma linguagem de consulta flexível (e rápida) para permitir a investigação ad-hoc?
- Ela oferece um framework de visualização flexível que vai além dos dashboards padrão do fornecedor?

Integração ao seu ecossistema mais amplo

- Ela se integra perfeitamente a outras ferramentas padrão nos seus ambientes — por exemplo, PagerDuty para resposta a incidentes, ServiceNow para gerenciamento de casos ou Slack para notificações? Ou às suas estruturas de CI/CD para ajudar a acompanhar automaticamente o impacto de envios de código e novas implantações no desempenho?

Facilidade de uso

Enquanto estiver elaborando os requisitos técnicos, avalie a facilidade de uso e a acessibilidade dos recursos. Pense nas várias pessoas que interagirão com a ferramenta e leve em consideração suas respectivas experiências, expectativas e fluxos de trabalho ao responder à pergunta sobre a facilidade de uso. Expanda sua avaliação além da experiência inicial, pensando na sobrecarga operacional contínua.

- Com que rapidez você pode instrumentar uma aplicação e ir do zero ao insight?
- A UI é intuitiva e fácil de usar?
- Quanto esforço é necessário para implantar a solução?
- É fácil atualizar a plataforma? E quanto aos agentes?
- Os agentes podem ser gerenciados centralmente a partir de um único local?
- Qual é o esforço envolvido na ampliação (ou redução)?

Opções de implantação

É importante que a ferramenta de APM escolhida seja compatível com as suas preferências de consumo de software.

- Você prefere uma opção de SaaS para reduzir custos operacionais e administrativos?
- Você quer uma opção que possa ajudar a reduzir os gastos anuais comprometidos no provedor de serviços em nuvem escolhido (AWS, Azure, Google Cloud etc.)?
- Você precisa de uma solução autogerenciada porque a nuvem não é uma opção por motivos de custo ou conformidade?
- Você tem uma estratégia multinuvem ou híbrida e quer executar sua solução de APM mais perto da sua carga de trabalho para reduzir latências ou custos de transferência de dados?

Supporte para padrões abertos/dados abertos

O espaço da observabilidade está começando a se desenvolver e abraçar padrões abertos como o OpenTelemetry (formado pela fusão de dois outros padrões abertos: OpenTracing e OpenCensus) como um framework padrão de instrumentação com neutralidade de fornecedor. O objetivo dessas iniciativas é ajudar os desenvolvedores a fazer a portabilidade de sua aplicação facilmente para uma solução de APM diferente, sem a necessidade de reinstrumentar.

Se você tem aplicações existentes instrumentadas usando um padrão aberto como o Jaeger, a mudança para uma ferramenta compatível pode acelerar e simplificar a migração. Os padrões abertos também ajudam a preparar o seu investimento para o futuro.

Arquitetura e escalabilidade

Avalie se a ferramenta de APM que você selecionou tem uma base e uma arquitetura robustas para lidar com sua escala atual e futura. Sua capacidade de investigar e resolver problemas de desempenho das aplicações com rapidez e eficácia depende do desempenho da sua ferramenta de APM. Não deixe de fazer o teste de carga da ferramenta de APM durante o processo de avaliação (ou analise os dados de referência se isso não for possível) para ter certeza de que ela pode lidar com o volume previsto de ingestão e consulta no seu ambiente sem arriar.

- Ela é construída sobre uma arquitetura simples? Ou a base é toda remendada, com risco de acabar se despedaçando?
- Ela dá suporte para alta disponibilidade?
- Ela pode ser ampliada facilmente para lidar com picos no volume de dados de monitoramento?
- Existe algum limite para o volume (aplicações, métricas, consultas etc.) com o qual ela consegue lidar?

Segurança

O processo de análise de segurança deve ser uma parte central do processo de avaliação da ferramenta. Não deixe de considerar estes dois ângulos:

1. O compromisso do fornecedor do APM com a segurança na forma como a ferramenta é construída e entregue
 - Os agentes de APM implantados nas suas aplicações precisam de privilégios excessivos?
 - O tráfego entre os componentes internos é devidamente criptografado e protegido?
 - A ferramenta tem as certificações exigidas (especialmente em SaaS)?
 - A ferramenta usa extensões de terceiros? Elas são seguras?
2. A capacidade de proteger e controlar o acesso à ferramenta de APM
 - A ferramenta se integra ao seu sistema de gerenciamento de acesso de identidade corporativa?
 - A ferramenta dá suporte para controle de acesso por função, com um modelo de permissões granular?

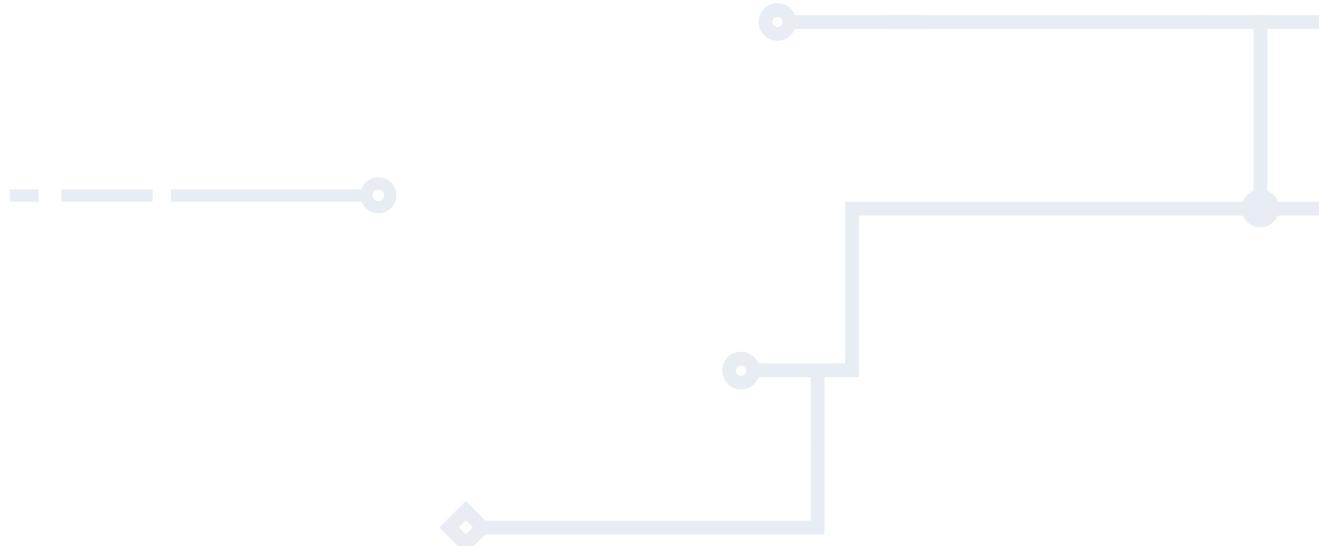
Recursos além do APM

Ao avaliar as ferramentas, pense além dos limites típicos do APM. A necessidade de consolidação de ferramentas e visibilidade unificada está motivando muitas organizações a unificar o APM com os logs e as métricas — os outros dois “pilares da observabilidade” — para ajudar a agilizar a investigação e reduzir o MTTR. Avalie se a ferramenta de APM dá suporte para essa visão unificada, exibindo logs e métricas no contexto dos dados de trace da aplicação. E quanto à ingestão de dados fora do cenário típico de TI, como sentimento social ou solicitações de suporte ao cliente?

Preços

Finalmente, é importante pesar cuidadosamente as opções de preços para garantir que a ferramenta de sua escolha não force você a comprometer sua visibilidade ou suas metas de monitoramento. Como acontece com muitos dos outros critérios, você precisará levar em consideração o uso atual e futuro (e a arquitetura) no processo. Há muita variação na forma como as soluções de APM são especificadas — por número de agentes, número de hosts, recursos de hardware etc. Além disso, alguns fornecedores também impõem custos adicionais quando você ultrapassa certos limites (por exemplo, o número de containers ou métricas). Aqui estão algumas perguntas a serem feitas sobre os modelos de preços:

- O modelo de preços está alinhado às suas necessidades de negócios e opções de arquitetura?
- Como os custos serão ampliados com o crescimento planejado e a evolução arquitetônica (por exemplo, de monólito para microsserviços)?
- Existe um nível gratuito? O que está incluído no nível gratuito? Existem limitações de uso?
- Qual nível de suporte está incluído?





Sobre a Elastic

Na qualidade de plataforma de logging open source mais popular, a Elastic proporciona velocidade, escala e relevância ao APM.

Quando você está investigando um problema que está afetando os usuários, cada segundo é valioso. Com o Elastic APM, seus dados de desempenho são salvos como um índice no Elasticsearch, permitindo que as equipes busquem e encontrem gargalos em tempo real. O Elastic APM também oferece mapas de serviços baseados em machine learning, opções de alerta customizadas e muito mais para que você possa criar melhores experiências digitais para seus usuários. Acesse elastic.co/pt/apm e saiba mais.

O APM é uma das peças do quebra-cabeça. Unifique seus logs, métricas e traces de APM em uma plataforma e conte com a verdadeira observabilidade de todo o seu ecossistema. Acesse elastic.co/pt/observability para saber mais.

Inicie a sua avaliação gratuita hoje mesmo