Guia sobre Observabilidade e Monitoramento

by Google

OPSERVICES

Introdução

Um bom monitoramento é um item essencial para equipes de alto desempenho. A pesquisa da DevOps Research and Assessment (DORA) mostra que uma solução abrangente de monitoramento e observabilidade, além de várias outras práticas técnicas, contribui positivamente para a entrega contínua.

A pesquisa da DORA definiu estes termos da seguinte maneira:

O **monitoramento** é uma ferramenta ou solução técnica que permite às equipes monitorar e entender o estado dos sistemas. Ele é baseado na coleta de conjuntos predefinidos de métricas ou registros.

A **observabilidade** é uma ferramenta ou uma solução técnica que permite às equipes depurar ativamente um sistema, explorando propriedades e padrões não definidos com antecedência.

PARA FAZER UM BOM TRABALHO COM MONITORAMENTO E OBSERVABILIDADE, SUAS EQUIPES PRECISAM TER:

- Relatórios sobre a integridade geral dos sistemas ("Meus sistemas estão funcionando?", "Meus sistemas têm recursos suficientes disponíveis?")
- Relatórios sobre o estado do sistema conforme a experiência dos clientes ("Meus clientes sabem se meu sistema está fora do ar e têm uma experiência ruim?")
- Monitoramento para as principais métricas de negócios e sistemas
- Ferramentas para ajudar você a entender e depurar seus sistemas em produção
- Ferramentas para encontrar informações sobre itens que você não sabia anteriormente (ou seja, é possível identificar o que não era conhecido)
- Acesso a ferramentas e dados que ajudam a rastrear, entender e diagnosticar problemas de infraestrutura no ambiente de produção, incluindo interações entre serviços

Sumário

A implementação dos dois modelos

As soluções de monitoramento e observabilidade foram criadas para:

- Fornecer indicadores líderes de interrupção ou degradação do serviço;
- Detectar interrupções, degradação do serviço, bugs e atividades não autorizadas;
- Ajudar a depurar interrupções, degradação do serviço, bugs e atividades não autorizadas;
- Identificar tendências de longo prazo para fins de planejamento de capacidade e negócios;
- Expor efeitos colaterais inesperados de alterações ou adição de recursos.

Assim como todos os recursos de DevOps, a instalação de uma ferramenta não é suficiente para atingir os objetivos, mas elas podem ajudar ou impedir o trabalho. Os sistemas de monitoramento não podem ser limitados a um único indivíduo ou equipe dentro de uma organização. Capacitar todos os desenvolvedores a serem proficientes no monitoramento ajuda a desenvolver uma cultura de tomada de decisão baseada em dados e melhora a depuração geral do sistema, reduzindo as interrupções.

Há algumas chaves para uma implementação eficaz do monitoramento e da observabilidade. Primeiro, seu monitoramento precisa informar o que está corrompido e ajudar você a entender o motivo, antes que muitos danos sejam feitos. A métrica principal no caso de uma interrupção ou degradação do serviço é o tempo de restauração (TTR, na sigla em inglês). Um dos principais fatores que contribuem com o TTR é a capacidade de entender rapidamente o que quebrou e o caminho mais rápido para restaurar o serviço (o que pode não envolver a correção imediata dos problemas básicos).

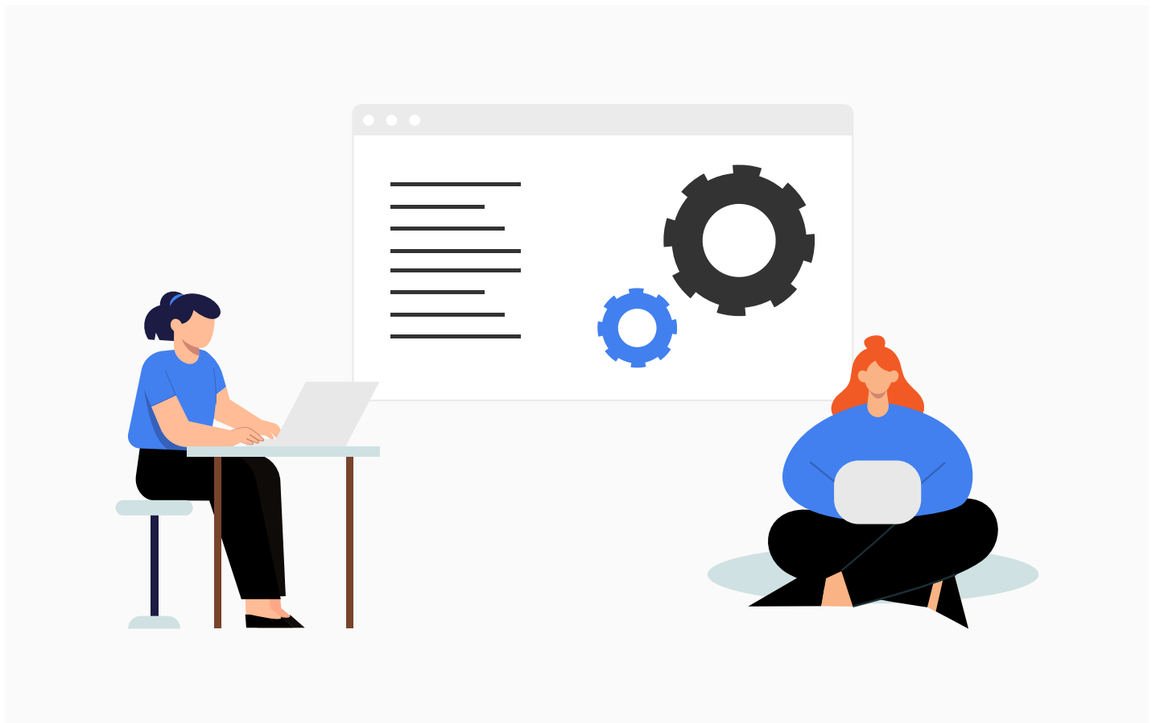
Existem duas maneiras de analisar um sistema de alto nível: o monitoramento de caixa preta, em que o estado interno e os mecanismos do sistema não são conhecidos, e o monitoramento de caixa branca, onde eles estão.

Monitoramento de caixa preta

Em um sistema de monitoramento de caixa preta (ou sintético), a entrada é enviada ao sistema em análise da mesma forma que um cliente pode fazer. Isso pode assumir a forma de chamadas HTTP para uma API pública ou chamadas RPC para um endpoint exposto ou pode estar chamando uma página da Web inteira como parte do processo de monitoramento.

O monitoramento de caixa preta é um método baseado em amostragem. O mesmo sistema responsável pelas solicitações do usuário é monitorado pelo sistema de caixa preta. Um sistema de caixa preta também pode fornecer cobertura da área de superfície do sistema de destino. Isso pode significar analisar cada método de API externo. Pense também em uma mistura representativa de solicitações para imitar melhor o comportamento real do cliente. Por exemplo, é possível executar 100 leituras e apenas uma gravação de uma determinada API.

É possível controlar esse processo com um sistema de programação para garantir que essas entradas sejam feitas com uma taxa suficiente para ganhar confiança na amostragem. Seu sistema também precisa conter um mecanismo de validação, que pode ser tão simples quanto verificar códigos de resposta ou corresponder a saída com expressões regulares, até renderizar um site dinâmico em um navegador sem comando e percorrer a árvore do DOM, procurando por elementos específicos. Depois que uma decisão é tomada (aprovada, com falha) em uma determinada sondagem, você precisa armazenar o resultado e os metadados para fins de relatório e alerta. Examinar um snapshot de uma falha e o contexto dela pode ser inestimável para diagnosticar um problema.



Monitoramento de caixa branca

O monitoramento e a observabilidade dependem de sinais enviados da carga de trabalho em análise para o sistema de monitoramento. Isso geralmente tem a forma dos três componentes mais comuns: métricas, registros e traces. Alguns sistemas de monitoramento também rastreiam e relatam eventos, que podem representar interações do usuário com um sistema inteiro ou alterações de estado no próprio sistema.

As métricas são simplesmente medidas realizadas dentro de um sistema, representando o estado desse sistema de maneira mensurável. Elas são quase sempre numéricas e tendem a assumir a forma de contadores, distribuições e medidores. Em alguns casos, as métricas de string fazem sentido, mas geralmente as métricas numéricas são usadas devido à necessidade de realizar cálculos matemáticos nelas para formar estatísticas e visualizações.

Os registros podem ser considerados como arquivos somente anexos que representam o estado de uma única linha de execução em um único momento. Esses registros podem ser uma única string, como "Usuário pressionou o botão X", ou uma entrada de registro estruturada que inclui metadados, como a hora do evento, o servidor que o processou e outros elementos ambientais. Às vezes, um sistema que não pode gravar registros estruturados produzirá uma string semiestruturada, como [timestamp] [server] message [code], que pode ser analisada após o fato, conforme necessário. As entradas de registro tendem a ser gravadas usando uma biblioteca de cliente, como log4j, structlog, bunyan, log4net ou Nlog. O processamento de registros pode ser um método muito confiável para

produzir estatísticas que podem ser consideradas confiáveis, já que podem ser reprocessadas com base em registros armazenados imutáveis, mesmo se o sistema de processamento de registros for problemático. Além disso, os registros podem ser processados em tempo real para produzir métricas com base em registros.

Os traces são compostos por períodos, que são usados para acompanhar um evento ou uma ação do usuário por meio de um sistema distribuído. Um período pode mostrar o caminho de uma solicitação por meio de um servidor, enquanto outro período pode ser executado em paralelo, ambos com o mesmo período pai. Juntos, eles formam um trace, que costuma ser visualizado em um gráfico de cascata semelhante ao usado em ferramentas de criação de perfil. Isso permite que os desenvolvedores entendam o tempo gasto em um sistema, em muitos servidores, filas e saltos de rede. Um framework comum para isso é o OpenTelemetry, que foi formado com base no OpenCensus e no OpenTracing (links em inglês).

Métricas, registros e traces podem ser relatados ao sistema de monitoramento pelo servidor em medição ou por um agente adjacente que pode testemunhar ou inferir algo sobre o sistema.

Instrumentação

Para usar um sistema de monitoramento, seu sistema precisa ser instrumentado. Ou seja, o código precisa ser adicionado a um sistema para expor o estado interno dele. Por exemplo, se um programa simples contém um pool de conexões com outro serviço, convém acompanhar o tamanho desse pool e o número de conexões não utilizadas a qualquer momento. Para fazer isso, um desenvolvedor precisa escrever um código na lógica do pool de conexão para acompanhar quando as conexões são formadas ou destruídas, quando elas são lançadas e quando são retornadas. Isso pode assumir a forma de entradas de registro ou eventos para cada um deles. Também é possível usar um medidor que aumente ou diminua de acordo com o tamanho da fila ou um contador que aumente sempre que uma conexão for criada ou sempre que um pool for expandido.

Correlação

As métricas podem ser coletadas no aplicativo, bem como nos sistemas de base, como JVM, SO convidado, hipervisor, SO do nó e o próprio hardware. À medida que você avança em uma pilha, talvez comece a misturar métricas que são compartilhadas entre as cargas de trabalho. Por exemplo, se uma máquina atender a vários aplicativos, a observação do uso do disco pode não corresponder diretamente ao sistema em observação. No entanto, a correlação de problemas entre aplicativos em um sistema compartilhado pode ajudar você a identificar um fator de contribuição (como um disco lento).

Medir um sistema distribuído significa ter capacidade de observabilidade em muitos lugares e poder vê-los juntos. Isso significa um front-end e o banco de dados dele ou um aplicativo para dispositivos móveis executado em um dispositivo de cliente, um balanceador de carga em nuvem e um conjunto de microsserviços. A capacidade de conectar dados de todas essas fontes em um só lugar é um requisito fundamental para as ferramentas de observabilidade.

Computação

Depois de coletar dados de várias fontes para seu sistema, você gera estatísticas e agrega dados em vários realms. Pode ser coortes de usuários, regiões da sua cobertura de computação ou localizações geográficas dos seus clientes. Ser capaz de desenvolver essas estatísticas rapidamente com base em eventos brutos é muito vantajoso, mas pode ser caro tanto em termos de armazenamento quanto de capacidade de computação em tempo real.

Ao escolher as ferramentas e planejar a instrumentação, você precisa considerar a cardinalidade e a dimensionalidade. Esses dois aspectos da coleta de métricas podem afetar muito sua capacidade de escalonar a observação de um sistema.

Cardinalidade é a medida de valores distintos em um sistema. Por exemplo, um campo como cpu-utilization tende a precisar de um intervalo entre 0 e 100. No entanto, se você acompanhar o identificador exclusivo de um usuário, todos eles serão distintos, portanto, se você tiver 1 milhão de usuários, terá uma cardinalidade de 1 milhão. Isso faz uma enorme diferença.

Dimensionabilidade é a capacidade de gravar mais do que um único valor junto com um carimbo de data/hora, como você talvez tenha em um banco de dados de séries temporais simples que apoia um sistema de monitoramento. Gravar o valor de um contador, por exemplo, solicitações enviadas, pode, inicialmente, registrar apenas o valor do número de solicitações enviadas até esse momento, como `{time=x, value=y}`. No entanto, assim como nos registros estruturados, talvez você também queira registrar alguns dados ambientais, resultando em algo como: `{time=x, value=y, server=foo, cluster=123, environment=prod, service=bar}`.

Combinar alta cardinalidade e alta dimensionalidade pode resultar em um aumento significativo nos requisitos de computação e armazenamento, até o momento em que o monitoramento talvez não funcione como esperado. Isso precisa ser compreendido por desenvolvedores que gravam dados e metadados gerados dinamicamente em sistemas de monitoramento.

Como aprender e aprimorar

Parte de como operar um sistema é o aprendizado com falhas e erros. O processo de gravação de retrospectivas ou *post mortems* com ações corretivas é bem documentado. Um resultado desse processo é o desenvolvimento de um monitoramento aprimorado.

É fundamental para uma organização em rápida evolução permitir que os sistemas de monitoramento sejam atualizados de maneira rápida e eficiente por qualquer pessoa na organização. A configuração do monitoramento também é essencial neste tipo de projeto. Portanto, as alterações precisam ser acompanhada por meio de revisão e aprovação, assim como no desenvolvimento e na entrega de código.

Manter sua configuração de monitoramento em um sistema de controle de versões é um bom primeiro passo para permitir acesso amplo ao sistema, enquanto mantém o controle sobre essa parte essencial do sistema. Desenvolver a automação em torno da implantação da configuração de monitoramento por meio de um pipeline de automação também pode melhorar sua capacidade de garantir que essas configurações sejam válidas e aplicadas de maneira consistente.

Depois de tratar a configuração de monitoramento como código, todas essas melhorias podem ser realizadas por meio de um processo de automação de implantação, de preferência o mesmo sistema usado pelo restante da equipe.

Armadilhas comuns na implementação do monitoramento e da observabilidade

Ao desenvolver um sistema de monitoramento e observabilidade para sua organização, geralmente não há uma solução simples pronta para uso. Qualquer sistema de monitoramento adequado exigirá uma compreensão profunda de cada componente que você quer medir, bem como a manipulação direta do código para instrumentar esses sistemas. Evite ter uma única pessoa monitorando ou uma equipe dedicada que seja a única responsável pelo sistema. Isso não apenas evitará um único ponto de falha, como também aumentará sua capacidade de entender e melhorar o sistema como uma organização inteira. O monitoramento e a observabilidade precisam ser incorporados ao conhecimento do valor de referência de todos os seus desenvolvedores. Uma armadilha comum é que a equipe de operações, NOC ou outra equipe semelhante seja a única autorizada a fazer alterações em um sistema de monitoramento. Isso precisa ser evitado e substituído por um sistema que siga padrões de CD.

Um antipadrão comum na criação de alertas nos sistemas de monitoramento é tentar enumerar todas as condições de erro possíveis e gravar um alerta para cada uma delas. Chamamos isso de alerta baseado na causa e você precisa evitá-lo o máximo possível. Em vez disso, concentre-se no alerta com base em sintaxe, que só avisa quando um sintoma voltado para o usuário está visível ou é previsto para ocorrer em breve. Você ainda consegue observar sistemas que não são voltados para o usuário, mas eles não alertarão diretamente os engenheiros de plantão se não houver sintomas voltados para

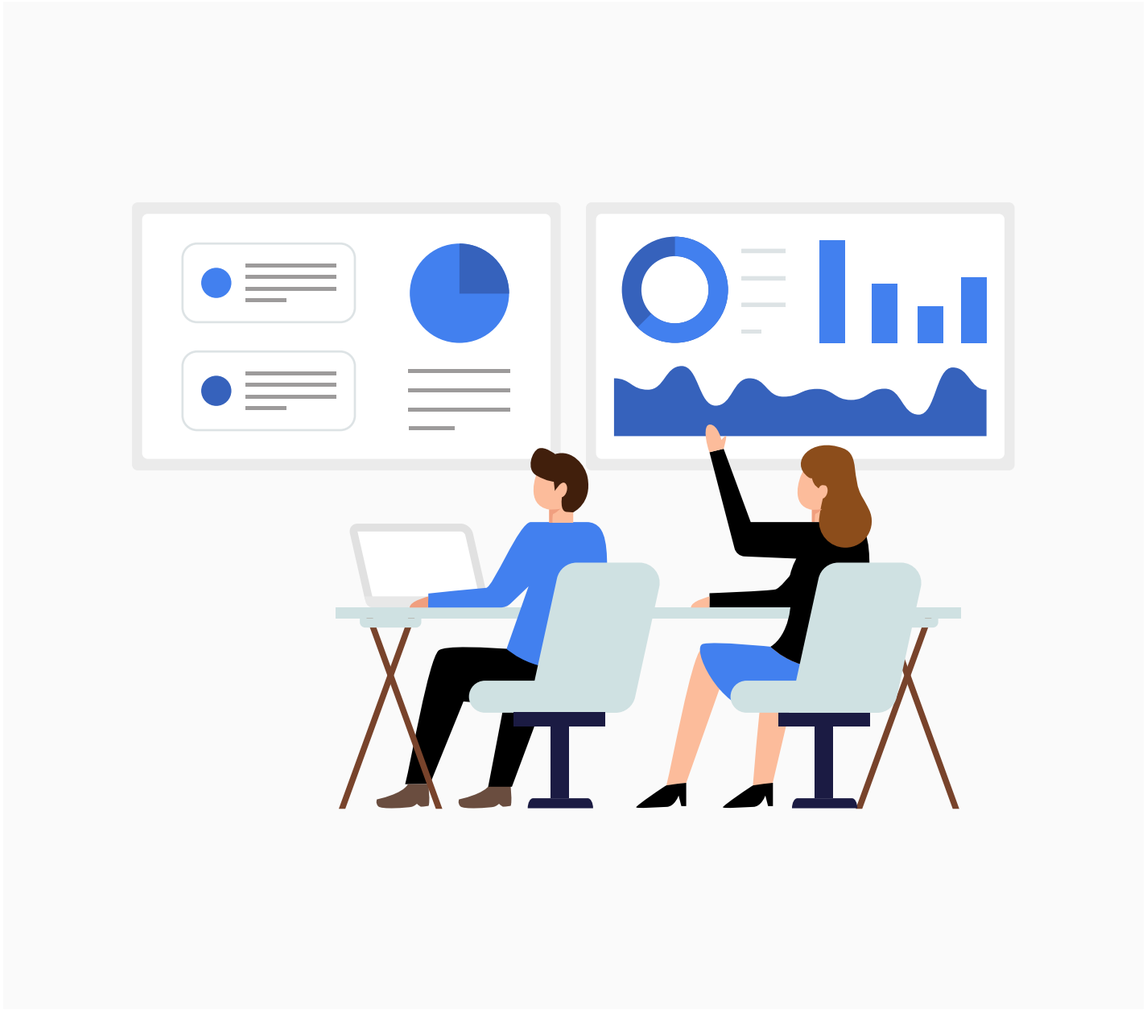
o usuário. O termo voltado para o usuário também pode incluir usuários internos da sua organização.

Ao gerar alertas, você precisa considerar como eles são entregues. Seus alertas precisam ter vários caminhos para os engenheiros de plantão, incluindo, entre outros: entrega de SMS, aplicativos para dispositivos móveis dedicados, chamadas telefônicas automatizadas ou e-mail. Uma armadilha comum é enviar alertas por e-mail para uma equipe inteira por meio de uma lista de distribuição de e-mails. Isso pode resultar rapidamente em alertas ignorados devido à difusão de responsabilidade.

Outra falha comum é uma baixa proporção de sinal-ruído. Se muitos alertas não forem acionáveis ou resultarem em nenhuma melhoria, a equipe perderá facilmente os alertas significativos e possivelmente muito importantes, um problema conhecido como deterioração do alarme. Qualquer método para silenciar ou suprimir algum conjunto de alertas precisa ser rastreado com muito cuidado para garantir que não seja muito amplo ou aplicado com muita frequência.

Ao criar painéis de monitoramento para visualizar métricas, um erro comum é passar muito tempo selecionando o "Painel perfeito". Isso é semelhante ao erro de alerta baseado em causa acima. Em uma equipe de alto desempenho, o sistema em observação muda tão rapidamente que qualquer tempo gasto em gerenciar um painel estará atrasado antes de ser concluído. Em vez disso, é importante que os membros da equipe criem rapidamente um painel ou outro conjunto de visualizações que atenda às necessidades deles.

Não separar as métricas voltadas para produtos ou executivos (como a taxa de aquisição de usuários e o rastreamento de receita) dos painéis operacionais ou de integridade do serviço também é um problema muito comum, já que ambas são importantes, mas distintas. É altamente recomendável mantê-las separadas.



Como medir o monitoramento e a observabilidade

Ao implementar um sistema de monitoramento e observabilidade na sua organização acompanhe algumas métricas internas para ver o desempenho. Veja algumas opções que podem ser acompanhadas com uma pesquisa mensal ou possivelmente analisando automaticamente os *post mortems* ou registros de alerta.

ALTERAÇÕES FEITAS NO MONITORAMENTO DA CONFIGURAÇÃO

Quantas solicitações de pull ou alterações por semana são feitas no repositório que contém a configuração de monitoramento? Com que frequência essas alterações são enviadas ao sistema de monitoramento? Diariamente? Por lote? Imediatamente em PR?

ALERTAS "FORA DO HORÁRIO"

Qual porcentagem dos alertas é gerenciada à noite? Algumas empresas globais têm um modelo de suporte 24 horas, o que não torna isso um problema, mas pode indicar que não foi dada atenção suficiente aos principais indicadores de falhas. Alertas noturnos regulares podem causar cansaço e equipes esgotadas.

CONFIRMAÇÃO DE ALERTA

Que porcentagem de alertas é confirmada dentro do prazo acordado (como 5 minutos e 30 minutos)? Às vezes, eles são acoplados ou podem ser rastreados por uma métrica como fall-through de alerta quando uma pessoa secundária de plantão é notificada para um alerta.

SALDO DE ALERTA DA EQUIPE

Se você tiver equipes em diferentes locais responsáveis por um serviço, os alertas serão distribuídos de maneira justa e endereçadas por todas as equipes? Em caso negativo, por quê?

FALSOS POSITIVOS

Quantos alertas resultaram em nenhuma ação ou foram marcados como "Trabalhando como pretendido"? Os alertas que não são acionáveis e que não ajudaram a prever falhas precisam ser excluídos.

FALSOS NEGATIVOS

Quantas falhas do sistema ocorreram sem alerta ou alerta depois do esperado? Com que frequência os post-mortems incluem adicionar alertas novos (baseados em sintomas)?

CRIAÇÃO DE ALERTAS

Quantos alertas são criados por semana (total ou agrupados por gravidade, equipe etc.)?

SILENCIAMENTO DE ALERTAS E DURAÇÃO DO SILENCIAMENTO

Quantos alertas há em um estado estável ou suprimido por semana? Quantos deles foram adicionados a este pool, quantos foram removidos? Se o silenciamento de alertas tiver um sistema de expiração, quantos silenciamentos serão estendidos para durar mais do que o esperado inicialmente? Qual é o período médio e máximo de silêncio? (Um engraçado é "quantos silêncios são efetivamente 'infinitos'?")

ALERTAS NÃO ACIONÁVEIS

Qual porcentagem de alertas foi considerada "não acionável"? Ou seja, o engenheiro alertado não conseguiu tomar qualquer medida imediatamente, devido à incapacidade de entender a implicação do alerta ou devido a um problema conhecido. Os alertas não acionáveis são uma fonte conhecida de trabalho.

USABILIDADE: ALERTAS, RUNBOOKS, PAINÉIS

Quantos gráficos há nos painéis? Quantas linhas por gráfico? As equipes entendem os gráficos? Há texto explicativo para ajudar novos engenheiros? As pessoas precisam rolar e procurar muito para encontrar as informações de que precisam? Os engenheiros podem navegar de alertas a manuais e painéis de maneira eficaz? Os alertas são nomeados de modo que apontem os engenheiros para a direção certa? Isso pode ser medido por pesquisas da equipe ao longo do tempo.

MTTD, MTTR, IMPACTO

O resultado é a hora de detectar, de resolver e de causar impacto. Avalie a "área sob a curva" do tempo em que a interrupção afetou os clientes vezes o número de clientes afetados. Isso pode ser estimado ou feito com mais precisão com ferramentas.

Ao acompanhar algumas ou todas essas métricas, você começará a entender melhor como seus sistemas de monitoramento e observabilidade estão funcionando para sua organização. Analisar essas medições por produto, equipe operacional ou outros métodos fornecerá insights não só sobre a integridade dos seus produtos, mas também sobre os processos e as pessoas.

Conclusão

As arquiteturas de software se transformaram muito rapidamente nos últimos anos. Hoje, existe um abismo tecnológico e metodológico entre as organizações que passaram pelo processo de transformação digital e as que ainda não iniciaram este processo. E isso pegou muitos profissionais e empresas de surpresa! O monitoramento e a observabilidade estão em uma relação simbiótica, ou seja, a observabilidade é alcançada quando os dados são disponibilizados a partir do sistema que você deseja monitorar. Já o monitoramento é a tarefa real de coletar e exibir esses dados.

A utilização do conceito de observabilidade se dá, principalmente, a partir da adoção em escala da nuvem. Embora se espere que as aplicações sejam altamente disponíveis e resilientes a falhas em ambos os ambientes (local e na nuvem), os métodos usados para alcançar tais objetivos são distintos.

Depois de ter tornado o sistema observável e de ter coletado os dados usando uma ferramenta de monitoração, é preciso realizar a análise manualmente ou automaticamente. Sem uma análise significativa, não faz sentido a observabilidade e o monitoramento. Quanto melhores forem as suas capacidades de análise, mais valiosos serão os seus investimentos em observabilidade e monitoramento.

No contexto do processo de integração e entrega contínua e no desenvolvimento de aplicações de microsserviços, o foco atual da observabilidade diz respeito à melhoria da qualidade dos dados de monitoramento, tornando esses serviços observáveis e fornecendo dados internos para aumentar as medições externas. Os dados de monitoramento de alta fidelidade resultantes melhoram a qualidade do feedback no loop de CI/CD.

Fontes de pesquisa

https://cloud.google.com/architecture/devops/devops-measurement-monitoring-and-observability?hl=pt-br

https://www.devops-research.com/research.html

https://cloud.google.com/devops?hl=pt-br

GOSTOU DO NOSSO CONTEÚDO E

QUER AJUDA PARA IMPLEMENTAR UM PROJETO DE OBSERVABILIDADE?

Somos especializados em desenvolver projetos para a coleta e visualização de dados de clientes de médio e grande porte.

ACESSE O NOSSO SITE

Sobre a OpServices

Empresa especializada em **monitoração e visualização de dados**, com quase 20 anos de mercado. É reconhecida por desenvolver soluções e serviços de monitoramento da infraestrutura e da rede de qualquer ambiente de TI.

Conheça nossas redes sociais: