Entenda eventos de desempenho e alertas

Active IQ Unified Manager

NetApp
October 15, 2025

Índice

Entenda eventos de desempenho e alertas . . . 1
- Fontes de eventos de desempenho . . . 1
- Tipos de gravidade de eventos de desempenho . . . 2
- Alterações de configuração detectadas pelo Unified Manager . . . 2
- Tipos de políticas de limite de desempenho definidas pelo sistema . . . 3
  - Políticas de limite de cluster . . . 3
  - Políticas de limite de nó . . . 4
  - Políticas de limite agregado . . . 4
  - Políticas de limite de latência de carga de trabalho . . . 5
  - Políticas de limite de QoS . . . 5
- Análise e notificação de eventos de desempenho . . . 6
  - Análise de eventos . . . 6
  - Estado do evento . . . 7
  - Notificação de evento . . . 7
  - Interação de eventos . . . 7
- Como o Unified Manager determina o impacto no desempenho de um evento . . . 8
- Componentes do cluster e por que eles podem estar em disputa . . . 8
- Funções das cargas de trabalho envolvidas em um evento de desempenho . . . 11

Entenda eventos de desempenho e alertas

Eventos de desempenho são incidentes relacionados ao desempenho da carga de trabalho em um cluster. Eles ajudam você a identificar cargas de trabalho com tempos de resposta lentos. Juntamente com os eventos de saúde que ocorreram ao mesmo tempo, você pode determinar os problemas que podem ter causado ou contribuído para os tempos de resposta lentos.

Quando o Unified Manager detecta várias ocorrências da mesma condição de evento para o mesmo componente de cluster, ele trata todas as ocorrências como um único evento, não como eventos separados.

Você pode configurar alertas para enviar notificações por e-mail automaticamente quando ocorrerem eventos de desempenho de determinados tipos de gravidade.

Fontes de eventos de desempenho

Eventos de desempenho são problemas relacionados ao desempenho da carga de trabalho em um cluster. Eles ajudam você a identificar objetos de armazenamento com tempos de resposta lentos, também conhecidos como alta latência. Juntamente com outros eventos de saúde que ocorreram ao mesmo tempo, você pode determinar os problemas que podem ter causado ou contribuído para os tempos de resposta lentos.

O Unified Manager recebe eventos de desempenho das seguintes fontes:

- **Eventos de política de limite de desempenho definidos pelo usuário**

  Problemas de desempenho com base em valores de limite personalizados que você definiu. Você configura políticas de limite de desempenho para objetos de armazenamento; por exemplo, agregados e volumes, para que eventos sejam gerados quando um valor limite para um contador de desempenho for violado.

  Você deve definir uma política de limite de desempenho e atribuí-la a um objeto de armazenamento para receber esses eventos.

- **Eventos de política de limite de desempenho definidos pelo sistema**

  Problemas de desempenho com base em valores limite definidos pelo sistema. Essas políticas de limite estão incluídas na instalação do Unified Manager para cobrir problemas comuns de desempenho.

  Essas políticas de limite são ativadas por padrão, e você poderá ver eventos logo após adicionar um cluster.

- **Eventos de limite de desempenho dinâmico**

  Problemas de desempenho que são resultado de falhas ou erros em uma infraestrutura de TI ou de cargas de trabalho que utilizam excessivamente os recursos do cluster. A causa desses eventos pode ser um problema simples que se corrige sozinho ao longo do tempo ou que pode ser resolvido com um reparo ou alteração de configuração. Um evento de limite dinâmico indica que as cargas de trabalho em um sistema ONTAP estão lentas devido a outras cargas de trabalho com alto uso de componentes de cluster compartilhados.

Esses limites são ativados por padrão, e você poderá ver eventos após três dias de coleta de dados de um novo cluster.

Tipos de gravidade de eventos de desempenho

Cada evento de desempenho é associado a um tipo de gravidade para ajudar você a priorizar os eventos que exigem ação corretiva imediata.

- **Crítico**

  Ocorreu um evento de desempenho que pode levar à interrupção do serviço se uma ação corretiva não for tomada imediatamente.

  Eventos críticos são enviados somente a partir de limites definidos pelo usuário.

- **Aviso**

  Um contador de desempenho para um objeto de cluster está fora da faixa normal e deve ser monitorado para garantir que não atinja a gravidade crítica. Eventos dessa gravidade não causam interrupção do serviço e pode não ser necessária ação corretiva imediata.

  Eventos de aviso são enviados a partir de limites definidos pelo usuário, definidos pelo sistema ou dinâmicos.

- **Informação**

  O evento ocorre quando um novo objeto é descoberto ou quando uma ação do usuário é executada. Por exemplo, quando qualquer objeto de armazenamento é excluído ou quando há alguma alteração de configuração, o evento com tipo de gravidade Informação é gerado.

  Eventos de informação são enviados diretamente do ONTAP quando ele detecta uma alteração de configuração.

Para mais informações, consulte os seguintes links:

- "O que acontece quando um evento é recebido"
- "Quais informações estão contidas em um e-mail de alerta"
- "Adicionando alertas"
- "Adicionar alertas para eventos de desempenho"

Alterações de configuração detectadas pelo Unified Manager

O Unified Manager monitora seus clusters em busca de alterações de configuração para ajudar você a determinar se uma alteração pode ter causado ou contribuído para um evento de desempenho. As páginas do Performance Explorer exibem um ícone de evento de alteração ( ) para indicar a data e a hora em que a alteração foi detectada.

Você pode revisar os gráficos de desempenho nas páginas do Performance Explorer e na página Análise de carga de trabalho para ver se o evento de alteração impactou o desempenho do objeto de cluster selecionado.

Se a alteração foi detectada no mesmo momento ou próximo a um evento de desempenho, a alteração pode ter contribuído para o problema, o que fez com que o alerta do evento fosse acionado.

O Unified Manager pode detectar os seguintes eventos de alteração, que são categorizados como eventos informativos:

- Um volume se move entre agregados.

  O Unified Manager pode detectar quando a movimentação está em andamento, concluída ou falhou. Se o Unified Manager estiver inativo durante uma movimentação de volume, quando ele voltar a funcionar, ele detectará a movimentação de volume e exibirá um evento de alteração para ela.

- O limite de taxa de transferência (MB/s ou IOPS) de um grupo de políticas de QoS que contém uma ou mais cargas de trabalho monitoradas muda.

  Alterar um limite de grupo de políticas pode causar picos intermitentes na latência (tempo de resposta), o que também pode acionar eventos para o grupo de políticas. A latência retorna gradualmente ao normal e quaisquer eventos causados pelos picos se tornam obsoletos.

- Um nó em um par de HA assume ou devolve o armazenamento de seu nó parceiro.

  O Unified Manager pode detectar quando a operação de aquisição, aquisição parcial ou devolução foi concluída. Se a aquisição for causada por um nó em pânico, o Unified Manager não detectará o evento.

- Uma operação de atualização ou reversão do ONTAP foi concluída com sucesso.

  A versão anterior e a nova versão são exibidas.

Tipos de políticas de limite de desempenho definidas pelo sistema

O Unified Manager fornece algumas políticas de limite padrão que monitoram o desempenho do cluster e geram eventos automaticamente. Essas políticas são habilitadas por padrão e geram eventos de aviso ou informação quando os limites de desempenho monitorados são violados.

As políticas de limite de desempenho definidas pelo sistema não estão habilitadas nos sistemas Cloud Volumes ONTAP, ONTAP Edge ou ONTAP Select .

Se estiver recebendo eventos desnecessários de quaisquer políticas de limite de desempenho definidas pelo sistema, você poderá desabilitar os eventos para políticas individuais na página Configuração de Eventos.

Políticas de limite de cluster

As políticas de limite de desempenho do cluster definidas pelo sistema são atribuídas, por padrão, a cada cluster monitorado pelo Unified Manager:

- **Desequilíbrio de carga do cluster**

  Identifica situações em que um nó está operando com uma carga muito maior do que outros nós no cluster e, portanto, potencialmente afetando as latências da carga de trabalho.

Ele faz isso comparando o valor da capacidade de desempenho usada para todos os nós em um cluster para ver se algum nó excedeu o valor limite de 30% por mais de 24 horas. Este é um evento de alerta.

- **Desequilíbrio de capacidade do cluster**

  Identifica situações em que um agregado tem uma capacidade utilizada muito maior do que outros agregados no cluster e, portanto, pode afetar o espaço necessário para as operações.

  Ele faz isso comparando o valor da capacidade usada para todos os agregados no cluster para ver se há uma diferença de 70% entre quaisquer agregados. Este é um evento de alerta.

Políticas de limite de nó

As políticas de limite de desempenho de nó definidas pelo sistema são atribuídas, por padrão, a cada nó nos clusters monitorados pelo Unified Manager:

- **Limite de capacidade de desempenho utilizada ultrapassado**

  Identifica situações em que um único nó está operando acima dos limites de sua eficiência operacional e, portanto, potencialmente afetando as latências da carga de trabalho.

  Ele faz isso procurando por nós que estão usando mais de 100% de sua capacidade de desempenho por mais de 12 horas. Este é um evento de alerta.

- **Par de nós HA superutilizado**

  Identifica situações em que os nós em um par HA estão operando acima dos limites de eficiência operacional do par HA.

  Ele faz isso observando o valor da capacidade de desempenho usada para os dois nós no par HA. Se a capacidade de desempenho combinada usada pelos dois nós exceder 200% por mais de 12 horas, um failover do controlador afetará as latências da carga de trabalho. Este é um evento informativo.

- **Fragmentação do disco do nó**

  Identifica situações em que um disco ou discos em um agregado são fragmentados, tornando lentos os principais serviços do sistema e potencialmente afetando as latências da carga de trabalho em um nó.

  Ele faz isso observando certas taxas de operação de leitura e gravação em todos os agregados em um nó. Esta política também pode ser acionada durante a ressincronização do SyncMirror ou quando erros são encontrados durante operações de limpeza de disco. Este é um evento de alerta.

  

  A política “Fragmentação de disco de nó” analisa agregados somente de HDD; agregados de Flash Pool, SSD e FabricPool não são analisados.

Políticas de limite agregado

A política de limite de desempenho agregado definida pelo sistema é atribuída por padrão a cada agregado nos clusters monitorados pelo Unified Manager:

- **Discos agregados superutilizados**

  Identifica situações em que um agregado está operando acima dos limites de sua eficiência operacional, afetando potencialmente as latências da carga de trabalho. Ele identifica essas situações procurando por

agregados em que os discos no agregado são utilizados em mais de 95% por mais de 30 minutos. Essa política multicondição executa a seguinte análise para ajudar a determinar a causa do problema:

- Um disco no agregado está passando por atividade de manutenção em segundo plano?

  Algumas das atividades de manutenção em segundo plano pelas quais um disco pode estar passando são reconstrução de disco, limpeza de disco, ressincronização do SyncMirror e reparação.

- Há um gargalo de comunicação na interconexão Fibre Channel da prateleira de disco?
- Há pouco espaço livre no agregado? Um evento de aviso é emitido para esta política somente se uma (ou mais) das três políticas subordinadas também forem consideradas violadas. Um evento de desempenho não é acionado se apenas os discos no agregado forem utilizados em mais de 95%.

A política “Discos agregados superutilizados” analisa agregados somente de HDD e agregados de Flash Pool (híbridos); agregados de SSD e FabricPool não são analisados.

Políticas de limite de latência de carga de trabalho

As políticas de limite de latência de carga de trabalho definidas pelo sistema são atribuídas a qualquer carga de trabalho que tenha uma política de Nível de Serviço de Desempenho configurada com um valor de “latência esperada” definido:

- **Limite de latência de LUN/volume de carga de trabalho violado, conforme definido pelo nível de serviço de desempenho**

  Identifica volumes (compartilhamentos de arquivos) e LUNs que excederam seu limite de “latência esperada” e que estão afetando o desempenho da carga de trabalho. Este é um evento de alerta.

  Ele faz isso procurando por cargas de trabalho que excederam o valor de latência esperado em 30% do tempo durante a hora anterior.

Políticas de limite de QoS

As políticas de limite de desempenho de QoS definidas pelo sistema são atribuídas a qualquer carga de trabalho que tenha uma política de taxa de transferência máxima de QoS ONTAP configurada (IOPS, IOPS/TB ou MB/s). O Unified Manager aciona um evento quando o valor da taxa de transferência da carga de trabalho é 15% menor que o valor de QoS configurado:

- **Limite máximo de QoS IOPS ou MB/s**

  Identifica volumes e LUNs que excederam seu limite máximo de IOPS ou MB/s de taxa de transferência de QoS e que estão afetando a latência da carga de trabalho. Este é um evento de alerta.

  Quando uma única carga de trabalho é atribuída a um grupo de políticas, isso é feito procurando por cargas de trabalho que excederam o limite máximo de taxa de transferência definido no grupo de políticas de QoS atribuído durante cada período de coleta da hora anterior.

  Quando várias cargas de trabalho compartilham uma única política de QoS, isso é feito adicionando o IOPS ou MB/s de todas as cargas de trabalho na política e verificando esse total em relação ao limite.

- **QoS Peak IOPS/TB ou IOPS/TB com limite de tamanho de bloco**

  Identifica volumes que excederam seu limite de taxa de transferência de pico de IOPS/TB de QoS

adaptável (ou IOPS/TB com limite de tamanho de bloco) e que estão afetando a latência da carga de trabalho. Este é um evento de alerta.

Ele faz isso convertendo o limite de pico de IOPS/TB definido na política de QoS adaptável em um valor máximo de IOPS de QoS com base no tamanho de cada volume e, em seguida, procura volumes que excederam o IOPS máximo de QoS durante cada período de coleta de desempenho da hora anterior.

Esta política é aplicada a volumes somente quando o cluster é instalado com o software ONTAP 9.3 e posterior.

Quando o elemento “tamanho do bloco” é definido na política de QoS adaptável, o limite é convertido em um valor máximo de QoS em MB/s com base no tamanho de cada volume. Em seguida, ele procura volumes que excederam o QoS máximo de MB/s durante cada período de coleta de desempenho da hora anterior.

Esta política é aplicada a volumes somente quando o cluster é instalado com o software ONTAP 9.5 e posterior.

Análise e notificação de eventos de desempenho

Eventos de desempenho notificam você sobre problemas de desempenho de E/S em uma carga de trabalho causados por contenção em um componente do cluster. O Unified Manager analisa o evento para identificar todas as cargas de trabalho envolvidas, o componente em contenção e se o evento ainda é um problema que você precisa resolver.

O Unified Manager monitora a latência de E/S (tempo de resposta) e IOPS (operações) para volumes em um cluster. Quando outras cargas de trabalho usam excessivamente um componente do cluster, por exemplo, o componente fica em contenção e não consegue ter um desempenho ideal para atender às demandas da carga de trabalho. O desempenho de outras cargas de trabalho que usam o mesmo componente pode ser afetado, causando aumento de latência. Se a latência ultrapassar o limite de desempenho dinâmico, o Unified Manager acionará um evento de desempenho para notificá-lo.

Análise de eventos

O Unified Manager executa as seguintes análises, usando os 15 dias anteriores de estatísticas de desempenho, para identificar as cargas de trabalho das vítimas, as cargas de trabalho dos agressores e o componente do cluster envolvido em um evento:

- Identifica cargas de trabalho de vítimas cuja latência ultrapassou o limite de desempenho dinâmico, que é o limite superior da previsão de latência:
  - Para volumes em agregados híbridos de HDD ou Flash Pool (camada local), os eventos são acionados somente quando a latência é maior que 5 milissegundos (ms) e o IOPS é maior que 10 operações por segundo (ops/seg).
  - Para volumes em agregados totalmente SSD ou agregados FabricPool (camada de nuvem), os eventos são acionados somente quando a latência é maior que 1 ms e o IOPS é maior que 100 ops/s.
- Identifica o componente do cluster em contenção.

Se a latência das cargas de trabalho das vítimas na interconexão do cluster for maior que 1 ms, o Unified Manager tratará isso como significativo e acionará um evento para a interconexão do cluster.

- Identifica as cargas de trabalho agressivas que estão usando excessivamente o componente do cluster e fazendo com que ele fique em contenção.
- Classifica as cargas de trabalho envolvidas, com base no desvio na utilização ou atividade de um componente do cluster, para determinar quais agressores têm a maior alteração no uso do componente do cluster e quais vítimas são as mais impactadas.

Um evento pode ocorrer apenas por um breve momento e depois se corrigir depois que o componente que ele está usando não estiver mais em disputa. Um evento contínuo é aquele que ocorre novamente para o mesmo componente do cluster dentro de um intervalo de cinco minutos e permanece no estado ativo. Para eventos contínuos, o Unified Manager aciona um alerta após detectar o mesmo evento durante dois intervalos de análise consecutivos.

Quando um evento é resolvido, ele permanece disponível no Unified Manager como parte do registro de problemas de desempenho anteriores de um volume. Cada evento tem um ID exclusivo que identifica o tipo de evento e os volumes, cluster e componentes de cluster envolvidos.

Um único volume pode estar envolvido em mais de um evento ao mesmo tempo.

Estado do evento

Os eventos podem estar em um dos seguintes estados:

- **Ativo**

  Indica que o evento de desempenho está ativo no momento (novo ou reconhecido). O problema que causou o evento não foi corrigido ou não foi resolvido. O contador de desempenho do objeto de armazenamento permanece acima do limite de desempenho.

- **Obsoleto**

  Indica que o evento não está mais ativo. O problema que causou o evento foi corrigido ou resolvido. O contador de desempenho do objeto de armazenamento não está mais acima do limite de desempenho.

Notificação de evento

Os eventos são exibidos na página Painel e em muitas outras páginas na interface do usuário, e alertas para esses eventos são enviados para endereços de e-mail especificados. Você pode visualizar informações detalhadas de análise sobre um evento e obter sugestões para resolvê-lo na página Detalhes do evento e na página Análise de carga de trabalho.

Interação de eventos

Na página Detalhes do evento e na página Análise de carga de trabalho, você pode interagir com eventos das seguintes maneiras:

- Mover o mouse sobre um evento exibe uma mensagem que mostra a data e a hora em que o evento foi detectado.

Se houver vários eventos no mesmo período de tempo, a mensagem mostrará o número de eventos.

- Clicar em um único evento exibe uma caixa de diálogo que mostra informações mais detalhadas sobre o evento, incluindo os componentes do cluster envolvidos.

  O componente em disputa é circulado e destacado em vermelho. Você pode clicar em **Ver análise completa** para ver a análise completa na página de detalhes do evento. Se houver vários eventos no mesmo período, a caixa de diálogo mostrará detalhes sobre os três eventos mais recentes. Você pode clicar em um evento para visualizar a análise do evento na página Detalhes do evento.

Como o Unified Manager determina o impacto no desempenho de um evento

O Unified Manager usa o desvio na atividade, utilização, taxa de transferência de gravação, uso de componentes do cluster ou latência de E/S (tempo de resposta) de uma carga de trabalho para determinar o nível de impacto no desempenho da carga de trabalho. Essas informações determinam a função de cada carga de trabalho no evento e como elas são classificadas na página Detalhes do evento.

O Unified Manager compara os últimos valores analisados para uma carga de trabalho com o intervalo esperado (previsão de latência) de valores. A diferença entre os últimos valores analisados e o intervalo esperado de valores identifica as cargas de trabalho cujo desempenho foi mais impactado pelo evento.

Por exemplo, suponha que um cluster contenha duas cargas de trabalho: Carga de Trabalho A e Carga de Trabalho B. A previsão de latência para a Carga de Trabalho A é de 5 a 10 milissegundos por operação (ms/op) e sua latência real geralmente é em torno de 7 ms/op. A previsão de latência para a Carga de Trabalho B é de 10 a 20 ms/op e sua latência real geralmente é em torno de 15 ms/op. Ambas as cargas de trabalho estão dentro da previsão de latência. Devido à contenção no cluster, a latência de ambas as cargas de trabalho aumenta para 40 ms/op, cruzando o limite de desempenho dinâmico, que é o limite superior da previsão de latência, e acionando eventos. O desvio na latência, dos valores esperados para os valores acima do limite de desempenho, para a Carga de Trabalho A é de cerca de 33 ms/op, e o desvio para a Carga de Trabalho B é de cerca de 25 ms/op. A latência de ambas as cargas de trabalho atingiu o pico de 40 ms/op, mas a carga de trabalho A teve o maior impacto no desempenho porque teve o maior desvio de latência em 33 ms/op.

Na página Detalhes do evento, na seção Diagnóstico do sistema, você pode classificar as cargas de trabalho por desvio em atividade, utilização ou taxa de transferência para um componente do cluster. Você também pode classificar cargas de trabalho por latência. Quando você seleciona uma opção de classificação, o Unified Manager analisa o desvio na atividade, utilização, taxa de transferência ou latência desde que o evento foi detectado em relação aos valores esperados para determinar a ordem de classificação da carga de trabalho. Para a latência, os pontos vermelhos (🔴) indicam um limite de desempenho ultrapassado por uma carga de trabalho da vítima e o impacto subsequente na latência. Cada ponto vermelho indica um nível maior de desvio na latência, o que ajuda a identificar as cargas de trabalho das vítimas cuja latência foi mais impactada por um evento.

Componentes do cluster e por que eles podem estar em disputa

Você pode identificar problemas de desempenho do cluster quando um componente do cluster entra em contenção. O desempenho das cargas de trabalho que usam o

componente fica lento e seu tempo de resposta (latência) para solicitações do cliente aumenta, o que aciona um evento no Unified Manager.

Um componente que está em disputa não pode ter um desempenho ideal. Seu desempenho diminuiu, e o desempenho de outros componentes do cluster e cargas de trabalho, chamados de *vítimas*, podem ter aumentado a latência. Para tirar um componente da disputa, você deve reduzir sua carga de trabalho ou aumentar sua capacidade de lidar com mais trabalho, para que o desempenho possa retornar aos níveis normais. Como o Unified Manager coleta e analisa o desempenho da carga de trabalho em intervalos de cinco minutos, ele detecta apenas quando um componente do cluster é usado em excesso de forma consistente. Picos transitórios de uso excessivo que duram apenas um curto período dentro do intervalo de cinco minutos não são detectados.

Por exemplo, um agregado de armazenamento pode estar em disputa porque uma ou mais cargas de trabalho nele estão competindo para que suas solicitações de E/S sejam atendidas. Outras cargas de trabalho no agregado podem ser impactadas, causando diminuição de desempenho. Para reduzir a quantidade de atividade no agregado, há diferentes etapas que você pode seguir, como mover uma ou mais cargas de trabalho para um agregado ou nó menos ocupado, para diminuir a demanda geral de carga de trabalho no agregado atual. Para um grupo de políticas de QoS, você pode ajustar o limite de taxa de transferência ou mover cargas de trabalho para um grupo de políticas diferente, para que as cargas de trabalho não sejam mais limitadas.

O Unified Manager monitora os seguintes componentes do cluster para alertá-lo quando eles estão em contenção:

- **Rede**

  Representa o tempo de espera de solicitações de E/S pelos protocolos de rede externos no cluster. O tempo de espera é o tempo gasto esperando que as transações “transferência pronta” sejam concluídas antes que o cluster possa responder a uma solicitação de E/S. Se o componente de rede estiver em contenção, isso significa que o alto tempo de espera na camada de protocolo está afetando a latência de uma ou mais cargas de trabalho.

- **Processamento de rede**

  Representa o componente de software no cluster envolvido com o processamento de E/S entre a camada de protocolo e o cluster. O nó que manipula o processamento da rede pode ter mudado desde que o evento foi detectado. Se o componente de processamento de rede estiver em contenção, isso significa que a alta utilização no nó de processamento de rede está afetando a latência de uma ou mais cargas de trabalho.

  Ao usar um cluster All SAN Array em uma configuração ativa-ativa, o valor de latência de processamento da rede é exibido para ambos os nós para que você possa verificar se os nós estão compartilhando a carga igualmente.

- **Limite Máximo de QoS**

  Representa a configuração máxima (pico) de taxa de transferência do grupo de políticas de Qualidade de Serviço (QoS) de armazenamento atribuído à carga de trabalho. Se o componente do grupo de políticas estiver em contenção, isso significa que todas as cargas de trabalho no grupo de políticas estão sendo limitadas pelo limite de taxa de transferência definido, o que está impactando a latência de uma ou mais dessas cargas de trabalho.

- **Limite mínimo de QoS**

  Representa a latência para uma carga de trabalho que está sendo causada pela configuração mínima

(esperada) de taxa de transferência de QoS atribuída a outras cargas de trabalho. Se o QoS mínimo definido em determinadas cargas de trabalho usar a maior parte da largura de banda para garantir a taxa de transferência prometida, outras cargas de trabalho serão limitadas e terão mais latência.

- **Interconexão de Cluster**

  Representa os cabos e adaptadores com os quais os nós agrupados estão fisicamente conectados. Se o componente de interconexão do cluster estiver em contenção, isso significa que o alto tempo de espera para solicitações de E/S na interconexão do cluster está afetando a latência de uma ou mais cargas de trabalho.

- * Data Processing*

  Representa o componente de software no cluster envolvido com o processamento de E/S entre o cluster e o agregado de armazenamento que contém a carga de trabalho. O nó que manipula o processamento de dados pode ter mudado desde que o evento foi detectado. Se o componente de processamento de dados estiver em contenção, isso significa que a alta utilização no nó de processamento de dados está afetando a latência de uma ou mais cargas de trabalho.

- **Ativação de Volume**

  Representa o processo que rastreia o uso de todos os volumes ativos. Em ambientes grandes onde mais de 1000 volumes estão ativos, esse processo rastreia quantos volumes críticos precisam acessar recursos por meio do nó ao mesmo tempo. Quando o número de volumes ativos simultâneos excede o limite máximo recomendado, alguns dos volumes não críticos sofrerão latência, conforme identificado aqui.

- * Recursos do MetroCluster *

  Representa os recursos do MetroCluster , incluindo NVRAM e links entre switches (ISLs), usados para espelhar dados entre clusters em uma configuração do MetroCluster . Se o componente MetroCluster estiver em contenção, isso significa que há alta taxa de transferência de gravação de cargas de trabalho no cluster local ou um problema de integridade do link que está afetando a latência de uma ou mais cargas de trabalho no cluster local. Se o cluster não estiver em uma configuração MetroCluster , este ícone não será exibido.

- **Operações agregadas ou SSD agregadas**

  Representa o agregado de armazenamento no qual as cargas de trabalho estão sendo executadas. Se o componente agregado estiver em contenção, isso significa que a alta utilização no agregado está afetando a latência de uma ou mais cargas de trabalho. Um agregado consiste em todos os HDDs ou uma mistura de HDDs e SSDs (um agregado Flash Pool) ou uma mistura de HDDs e uma camada de nuvem (um agregado FabricPool). Um “Agregado SSD” consiste em todos os SSDs (um agregado all-flash) ou uma mistura de SSDs e uma camada de nuvem (um agregado FabricPool).

- **Latência da Nuvem**

  Representa o componente de software no cluster envolvido com o processamento de E/S entre o cluster e a camada de nuvem na qual os dados do usuário são armazenados. Se o componente de latência da nuvem estiver em contenção, isso significa que uma grande quantidade de leituras de volumes hospedados na camada de nuvem estão impactando a latência de uma ou mais cargas de trabalho.

- **Sincronizar SnapMirror**

  Representa o componente de software no cluster envolvido na replicação de dados do usuário do volume primário para o volume secundário em um relacionamento síncrono do SnapMirror . Se o componente de sincronização SnapMirror estiver em contenção, isso significa que a atividade das operações síncronas do

SnapMirror está impactando a latência de uma ou mais cargas de trabalho.

Funções das cargas de trabalho envolvidas em um evento de desempenho

O Unified Manager usa funções para identificar o envolvimento de uma carga de trabalho em um evento de desempenho. Os papéis incluem vítimas, valentões e tubarões. Uma carga de trabalho definida pelo usuário pode ser uma vítima, um agressor e um tubarão ao mesmo tempo.

Papel	Descrição
Vítima	Uma carga de trabalho definida pelo usuário cujo desempenho diminuiu devido a outras cargas de trabalho, chamadas de bullies, que estão usando excessivamente um componente do cluster. Somente cargas de trabalho definidas pelo usuário são identificadas como vítimas. O Unified Manager identifica as cargas de trabalho das vítimas com base no desvio de latência, onde a latência real, durante um evento, aumentou muito em relação à previsão de latência (intervalo esperado).
Valentão	Uma carga de trabalho definida pelo usuário ou pelo sistema cujo uso excessivo de um componente do cluster fez com que o desempenho de outras cargas de trabalho, chamadas vítimas, diminuísse. O Unified Manager identifica cargas de trabalho agressivas com base no desvio no uso de um componente do cluster, onde o uso real, durante um evento, aumentou muito em relação ao intervalo de uso esperado.
Tubarão	Uma carga de trabalho definida pelo usuário com o maior uso de um componente de cluster em comparação com todas as cargas de trabalho envolvidas em um evento. O Unified Manager identifica cargas de trabalho tubarões com base no uso de um componente de cluster durante um evento.

As cargas de trabalho em um cluster podem compartilhar muitos dos componentes do cluster, como agregados e a CPU para processamento de rede e dados. Quando uma carga de trabalho, como um volume, aumenta o uso de um componente de cluster a ponto de o componente não conseguir atender às demandas de carga de trabalho com eficiência, o componente está em contenção. A carga de trabalho que utiliza excessivamente um componente do cluster é um problema. As outras cargas de trabalho que compartilham esses componentes e cujo desempenho é afetado pelo agressor são as vítimas. Atividades de cargas de trabalho definidas pelo sistema, como desduplicação ou cópias de instantâneos, também podem evoluir para "bullying".

Quando o Unified Manager detecta um evento, ele identifica todas as cargas de trabalho e componentes de cluster envolvidos, incluindo as cargas de trabalho agressivas que causaram o evento, o componente de

cluster que está em contenção e as cargas de trabalho vítimas cujo desempenho diminuiu devido ao aumento da atividade das cargas de trabalho agressivas.

Se o Unified Manager não conseguir identificar as cargas de trabalho agressivas, ele alertará apenas sobre as cargas de trabalho das vítimas e o componente do cluster envolvido.

O Unified Manager pode identificar cargas de trabalho que são vítimas de cargas de trabalho de intimidação e também identificar quando essas mesmas cargas de trabalho se tornam cargas de trabalho de intimidação. Uma carga de trabalho pode ser um agressor para si mesma. Por exemplo, uma carga de trabalho de alto desempenho que está sendo limitada por um limite de grupo de políticas faz com que todas as cargas de trabalho no grupo de políticas sejam limitadas, incluindo ela mesma. Uma carga de trabalho que é um agressor ou uma vítima em um evento de desempenho em andamento pode mudar sua função ou não ser mais um participante do evento.

Informações sobre direitos autorais

Informações sobre marcas comerciais