¿Qué eventos de rendimiento son

OnCommand Unified Manager 9.5

NetApp
April 02, 2025

Tabla de contenidos

¿Qué eventos de rendimiento son

Los eventos de rendimiento son incidentes relacionados con el rendimiento de la carga de trabajo en un clúster. Le ayudan a identificar cargas de trabajo con tiempos de respuesta lentos. Junto con los eventos de salud que ocurrieron al mismo tiempo, usted puede determinar los problemas que podrían haber causado, o contribuido a, los tiempos de respuesta lentos.

Cuando Unified Manager detecta varias apariciones de la misma condición de evento para el mismo componente del clúster, trata todas las ocurrencias como un solo evento, no como eventos independientes.

Análisis y notificación de eventos de rendimiento

Los eventos de rendimiento le notifican problemas de rendimiento de I/o en una carga de trabajo de volumen debido a la contención en un componente de clúster. Unified Manager analiza el evento para identificar todas las cargas de trabajo involucradas, el componente en disputa y si el evento sigue siendo un problema que podría necesitar resolver.

Unified Manager supervisa la latencia de I/o (tiempo de respuesta) y las IOPS (operaciones) para volúmenes en un clúster. Cuando otras cargas de trabajo realizan un uso excesivo de un componente del clúster, por ejemplo, el componente es objeto de disputa y no puede funcionar en un nivel óptimo para satisfacer las demandas de las cargas de trabajo. El rendimiento de otras cargas de trabajo que utilizan el mismo componente puede verse afectado, lo que provoca el aumento de las latencias. Si la latencia supera el umbral de rendimiento, Unified Manager activa un evento de rendimiento y envía una alerta de correo electrónico para notificarle.

Análisis de eventos

Unified Manager realiza los siguientes análisis utilizando los 15 días anteriores de estadísticas de rendimiento para identificar las cargas de trabajo víctimas, las cargas de trabajo abusivas y el componente del clúster implicados en un evento:

- Identifica las cargas de trabajo víctimas cuya latencia ha superado el umbral de rendimiento, que es el límite superior del rango esperado:
  - Para los volúmenes en agregados de HDD o Flash Pool (híbridos), los eventos solo se activan cuando la latencia es superior a 5 milisegundos (ms) y las IOPS son superiores a 10 operaciones por segundo (OPS/s).
  - Para volúmenes en agregados íntegramente de SSD o agregados de FabricPool (compuestos), los eventos solo se activan cuando la latencia es superior a 1 ms y el IOPS es superior a 100 OPS/s.
- Identifica el componente del clúster en disputa.

  Si la latencia de las cargas de trabajo víctimas en la interconexión del clúster es superior a 1 ms, Unified Manager lo considera importante y activa un evento para la interconexión del clúster.

- Identifica las cargas de trabajo abusivas que están sobreutilizando el componente del clúster y que hacen que estén en contención.

- Clasifica las cargas de trabajo involucradas, en función de su desviación en utilización o actividad de un componente del clúster, para determinar qué elementos agresores tienen el mayor cambio de uso del componente del clúster y qué víctimas son las más afectadas.

Es posible que se produzca un evento solo durante un breve momento y, a continuación, corregirlo después de que el componente que está utilizando ya no sea objeto de disputa. Un evento continuo es un evento que se vuelve a producir para el mismo componente del clúster en un intervalo de cinco minutos y permanece en el estado activo. En el caso de eventos continuos, Unified Manager activa una alerta tras detectar el mismo evento en dos intervalos de análisis consecutivos. Los eventos que no se resuelven, que tienen el estado de nuevo, pueden mostrar diferentes mensajes de descripción según las cargas de trabajo involucradas en el cambio del evento.

Cuando se resuelve un evento, este sigue disponible en Unified Manager como parte del registro de problemas de rendimiento anteriores de un volumen. Cada evento tiene un ID único que identifica el tipo de evento y los componentes de volúmenes, clúster y clúster implicados.

Un único volumen puede participar en más de un evento a la vez.

Estado del evento

Los eventos pueden estar en uno de los siguientes estados:

- **Activo**

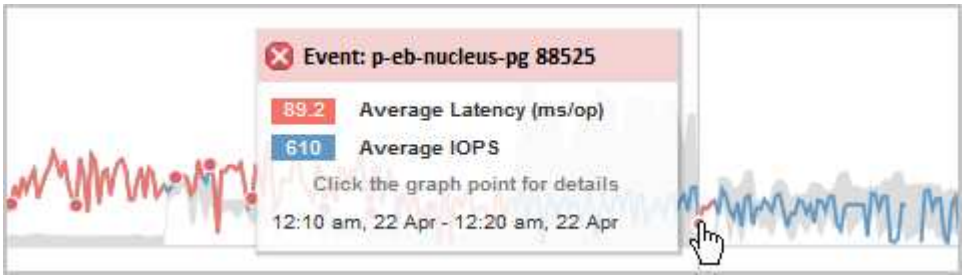
  Indica que el evento de rendimiento está activo (nuevo o reconocido). El problema que causa el evento no se ha corregido solo o no se ha resuelto. El contador de rendimiento del objeto de almacenamiento sigue por encima del umbral de rendimiento.

- **Obsoleto**

  Indica que el evento no está activo. El problema que causa el evento se ha corregido solo o se ha resuelto. El contador de rendimiento del objeto de almacenamiento ya no está por encima del umbral de rendimiento.

Notificación de eventos

Las alertas de eventos se muestran en la página Paneles/Descripción general, la página Paneles/rendimiento, la página Detalles de rendimiento/volumen y se envían a direcciones de correo electrónico especificadas. Puede ver información detallada de análisis sobre un evento y obtener sugerencias para resolverlo en la página de detalles del evento.



En este ejemplo, un evento se indica mediante un punto rojo (🔴) En el gráfico latencia de la página Performance/Volume Details. Al pasar el cursor del ratón por encima del punto rojo, se muestra una ventana emergente con más detalles sobre el evento y las opciones para analizarlo.

Interacción de eventos

En la página Performance/Volume Details, puede interactuar con los eventos de las siguientes maneras:

- Si mueve el puntero por un punto rojo, se muestra un mensaje que muestra el identificador del evento, junto con la latencia, el número de operaciones por segundo y la fecha y la hora en que se detectó el evento.

  Si existen varios eventos para el mismo período de tiempo, el mensaje muestra el número de eventos, junto con la latencia media y las operaciones por segundo del volumen.

- Al hacer clic en un solo evento se muestra un cuadro de diálogo que muestra información más detallada sobre el evento, incluidos los componentes del clúster implicados, de forma similar a la sección Summary en la página de detalles Event.

  El componente objeto de la contención está en un círculo y se resalta en rojo. Puede hacer clic en el ID del evento o en **Ver análisis completo** para ver el análisis completo en la página de detalles del evento. Si hay varios eventos para el mismo período de tiempo, el cuadro de diálogo muestra detalles acerca de los tres eventos más recientes. Puede hacer clic en un ID de evento para ver el análisis de eventos en la página de detalles Event. Si hay más de tres eventos para el mismo período de tiempo, al hacer clic en el punto rojo no se muestra el cuadro de diálogo.

La forma en que Unified Manager determina el impacto en el rendimiento de un evento

Unified Manager utiliza la desviación de la actividad, la utilización, el rendimiento de escritura, el uso de componentes del clúster o la latencia de l/o (tiempo de respuesta) en una carga de trabajo para determinar el nivel de impacto en el rendimiento de la carga de trabajo. Esta información determina el rol de cada carga de trabajo del evento y su clasificación en la página de detalles Event.

Unified Manager compara los valores del último análisis de una carga de trabajo con el rango esperado de valores. La diferencia entre los valores que se analizaron por última vez y el rango esperado de valores identifica las cargas de trabajo cuyo rendimiento tuvo un mayor impacto en el evento.

Por ejemplo, supongamos que un clúster contiene dos cargas de trabajo: La carga de trabajo A y la carga de trabajo B. El rango esperado para la carga de trabajo A es de 5-10 milisegundos por operación (ms/op) y su latencia real suele ser de aproximadamente 7 ms/op. El rango esperado para la carga de trabajo B es de 10-20 ms/op y su latencia real suele ser de aproximadamente 15 ms/op. Ambas cargas de trabajo están dentro del rango esperado de latencia. Debido a una contención en el clúster, la latencia de ambas cargas de trabajo aumenta a 40 ms/op, que cruza el umbral de rendimiento, que es los límites superiores del rango esperado y activa los eventos. La desviación en latencia, de los valores esperados a los valores por encima del umbral de rendimiento, para la carga De trabajo A es aproximadamente de 33 ms/op y la desviación de la carga de trabajo B es de unos 25 ms/op. La latencia de ambas cargas de trabajo se dispara hasta los 40 ms/op, pero la carga De trabajo A tuvo un impacto en el rendimiento mayor porque tuvo la mayor desviación de latencia en 33 ms/op.

En la página de detalles Event, en la sección System Diagnosis, se pueden ordenar las cargas de trabajo por su desviación de la actividad, la utilización o el rendimiento de un componente del clúster. También es posible ordenar las cargas de trabajo por latencia. Cuando se selecciona una opción de ordenación, Unified Manager analiza la desviación de la actividad, la utilización, el rendimiento o la latencia desde que se detectó el evento de los valores esperados para determinar el orden de clasificación de la carga de trabajo. Para la latencia, los

puntos rojos (🔴) indica un umbral de rendimiento que cruza una carga de trabajo víctima y el impacto posterior en la latencia. Cada punto rojo indica un nivel más alto de desviación en latencia, lo que ayuda a identificar las cargas de trabajo víctimas cuya latencia más se vio afectada por un evento.

Los componentes del clúster y el motivo por los que pueden estar en contención

Puede identificar los problemas de rendimiento del clúster cuando un componente del clúster entra en disputa. El rendimiento de las cargas de trabajo de volúmenes que utilizan el componente disminuye y aumenta su tiempo de respuesta (latencia) para las solicitudes de clientes, lo que activa un evento en Unified Manager.

Un componente que está en disputa no puede ejecutarse en un nivel óptimo. Su rendimiento ha disminuido, y el rendimiento de otros componentes del clúster y las cargas de trabajo, llamado *Victimas*, puede haber aumentado la latencia. Para eliminar un componente de la contención, debe reducir su carga de trabajo o aumentar su capacidad para gestionar más trabajo, de modo que el rendimiento pueda volver a los niveles normales. Dado que Unified Manager recopila y analiza el rendimiento de la carga de trabajo en intervalos de cinco minutos, solo detecta cuando un componente de clúster se utiliza de forma coherente en exceso. No se detectan picos transitorios de sobreutilización que duran solo una corta duración dentro del intervalo de cinco minutos.

Por ejemplo, un agregado de almacenamiento puede ser objeto de disputa porque una o más cargas de trabajo en él están compitiendo por sus solicitudes de I/O. Otras cargas de trabajo del agregado pueden verse afectadas, lo que provoca una disminución del rendimiento. Para reducir la cantidad de actividad del agregado, hay diferentes pasos que se pueden realizar, como mover una o más cargas de trabajo a un agregado menos ocupado, para reducir la demanda general de la carga de trabajo en el agregado actual. En el caso de un grupo de políticas de calidad de servicio, puede ajustar el límite de rendimiento o mover cargas de trabajo a otro grupo de políticas, de modo que las cargas de trabajo ya no se aceleren.

Unified Manager supervisa los siguientes componentes del clúster para alertarle cuando los hay en disputa:

- **Red**

  Representa el tiempo de espera de las solicitudes de I/o de los protocolos de iSCSI o los protocolos de Fibre Channel (FC) en el clúster. El tiempo de espera transcurrido para que las transacciones iSCSI Ready to Transfer (R2T) o FCP Transfer Ready (XFER_RDY) finalicen antes de que el clúster pueda responder a una solicitud de I/O. Si el componente de red es objeto de disputa, significa que un tiempo de espera elevado en la capa de protocolo de bloqueo está afectando a la latencia de una o más cargas de trabajo.

- **Procesamiento de red**

  Representa el componente de software del clúster involucrado en el procesamiento de I/o entre la capa de protocolo y el clúster. Es posible que el nodo que gestiona el procesamiento de red haya cambiado desde que se detectó el evento. Si el componente de red es objeto de disputa, significa que un uso elevado en el nodo de procesamiento de red está afectando a la latencia de una o más cargas de trabajo.

- **Política de QoS**

  Representa el grupo de políticas de calidad de servicio del almacenamiento al que pertenece la carga de trabajo. Si el componente del grupo de políticas es objeto de disputa, significa que el límite de rendimiento establecido está acelerando todas las cargas de trabajo del grupo de políticas, lo que afecta a la latencia de una o más cargas de trabajo.

- **Interconexión en cluster**

  Representa los cables y los adaptadores que conectan de forma física los nodos en clúster. Si el componente de interconexión del clúster es objeto de disputa, significa que un tiempo de espera elevado para las solicitudes de I/o en la interconexión del clúster está afectando a la latencia de una o más cargas de trabajo.

- **Procesamiento de datos**

  Representa el componente de software del clúster involucrado en el procesamiento de I/o entre el clúster y el agregado de almacenamiento que contiene la carga de trabajo. Es posible que el nodo que gestiona el procesamiento de datos haya cambiado desde que se detectó el evento. Si el componente de procesamiento de datos es objeto de disputa, significa que un uso elevado en el nodo de procesamiento de datos está afectando a la latencia de una o más cargas de trabajo.

- **Recursos de MetroCluster**

  Representa los recursos de MetroCluster, incluidos NVRAM y los vínculos interswitch (ISL), que se usan para reflejar datos entre los clústeres de una configuración de MetroCluster. Si el componente MetroCluster es objeto de disputa, significa que el alto rendimiento de escritura de las cargas de trabajo del clúster local o un problema de estado del enlace afectan a la latencia de una o más cargas de trabajo del clúster local. Si el clúster no está en una configuración MetroCluster, este icono no se muestra.

- **Operaciones globales o agregados de SSD**

  Representa el agregado de almacenamiento en el que se ejecutan las cargas de trabajo. Si el componente de agregado es objeto de disputa, significa que un uso elevado en el agregado está afectando a la latencia de una o más cargas de trabajo. Un agregado está formado por todos los HDD, o una combinación de HDD y SSD (un agregado de Flash Pool). Un «agregado SD» está compuesto por todos los SSD (un agregado all-flash) o una combinación de SSD y un nivel de cloud (un agregado de FabricPool).

- **Latencia de cloud**

  Representa el componente de software del clúster involucrado en el procesamiento de I/o entre el clúster y el nivel de cloud en el que se almacenan los datos del usuario. Si el componente de latencia del cloud es objeto de disputa, significa que una gran cantidad de lecturas de volúmenes que están alojados en el nivel de cloud están afectando a la latencia de una o más cargas de trabajo.

- **SnapMirror** sincronizado

  Representa el componente de software del clúster involucrado en la replicación de datos de usuario del volumen primario al secundario en una relación de SnapMirror Synchronous. Si el componente Sync SnapMirror es objeto de disputa, significa que la actividad de las operaciones de SnapMirror Synchronous está afectando a la latencia de una o más cargas de trabajo.

Funciones de las cargas de trabajo involucradas en un evento de rendimiento

Unified Manager utiliza roles para identificar la implicación de una carga de trabajo en un evento de rendimiento. Entre las funciones se encuentran las víctimas, los agresores y los tiburones. Una carga de trabajo definida por el usuario puede ser víctima, abusador o tiburón al mismo tiempo.

Función	Descripción
Víctima	Carga de trabajo definida por el usuario cuyo rendimiento ha disminuido debido a otras cargas de trabajo, denominadas verdugos, que usan en exceso un componente de clúster. Solo las cargas de trabajo definidas por el usuario se identifican como víctimas. Unified Manager identifica las cargas de trabajo víctimas en función de su desviación en la latencia, donde la latencia real, durante un evento, aumentó significativamente con respecto al rango esperado de latencia.
Matón	Una carga de trabajo definida por el usuario o el sistema cuyo uso excesivo de un componente del clúster ha provocado la reducción del rendimiento de otras cargas de trabajo denominadas víctimas. Unified Manager identifica cargas de trabajo problemáticas según su desviación en el uso de un componente del clúster, donde el uso real, durante un evento, ha aumentado considerablemente desde su rango de uso esperado.
IBM	Carga de trabajo definida por el usuario con el mayor uso de un componente del clúster en comparación con todas las cargas de trabajo involucradas en un evento. Unified Manager identifica cargas de trabajo de tiburón en función de su uso de un componente de clúster durante un evento.

Las cargas de trabajo de un clúster pueden compartir muchos de los componentes del clúster, como agregados de almacenamiento y la CPU para el procesamiento de datos y red. Cuando una carga de trabajo, como un volumen, aumenta el uso de un componente del clúster hasta el punto de que el componente no puede satisfacer de forma eficiente las demandas de las cargas de trabajo, el componente es objeto de disputa. La carga de trabajo que utiliza en exceso un componente del clúster es un problema. Las otras cargas de trabajo que comparten esos componentes y cuyo rendimiento afecta al abusador son las víctimas. La actividad de las cargas de trabajo definidas por el sistema, como la deduplicación o las copias Snapshot, también puede convertirse en «bullying».

Cuando Unified Manager detecta un evento, identifica todas las cargas de trabajo y los componentes del clúster implicados, incluidas las cargas de trabajo abusivas que causaron el evento, el componente de clúster que está en disputa y las cargas de trabajo víctimas cuyo rendimiento ha disminuido debido al aumento de la actividad de las cargas de trabajo abusivas.

Si Unified Manager no puede identificar las cargas de trabajo abusivas, solo envía alertas sobre las cargas de trabajo víctimas y el componente de clúster correspondiente.

Unified Manager puede identificar cargas de trabajo que son víctimas de cargas de trabajo abusivas y, además, identificar cuándo esas mismas cargas de trabajo pasan a ser cargas de trabajo abusivas. Una carga de trabajo puede ser un problema para sí misma. Por ejemplo, una carga de trabajo de alto rendimiento que está acelerando por el límite de un grupo de políticas provoca que se aceleren todas las cargas de trabajo del grupo de políticas, incluso la propia. Una carga de trabajo que sea intimidatada o víctima en un evento de

rendimiento continuo puede cambiar su función o ya no ser participante en el evento. En la página Performance/Volume Details, en la tabla Events List, cuando el volumen seleccionado cambia su rol de participante, se muestra la fecha y hora del cambio de rol.