En quoi sont les événements de performances

OnCommand Unified Manager 9.5

NetApp
August 02, 2024

This PDF was generated from https://docs.netapp.com/fr-fr/oncommand-unified-manager-95/performance-checker/reference-performance-event-analysis-and-notification.html on August 02, 2024.
Always check docs.netapp.com for the latest.

Sommaire

En quoi sont les événements de performances 1
Analyse et notification des événements de performance 1
Comment Unified Manager détermine l’impact sur les performances d’un événement 3
Les composants du cluster et les conflits 4
Rôles des charges de travail impliquées dans un événement de performance 5

En quoi sont les événements de performances

Les événements de performance sont des incidents liés aux performances des charges de travail sur un cluster. Ils vous aident à identifier les workloads avec des temps de réponse lents. Avec les événements de santé qui se sont produits en même temps, vous pouvez déterminer les problèmes qui pourraient avoir causé, ou contribué à, les délais de réponse lents.

Lorsque Unified Manager détecte plusieurs occurrences de la même condition d'événement pour le même composant de cluster, il traite toutes les occurrences comme un événement unique et non comme des événements distincts.

Analyse et notification des événements de performance

Les événements de performance vous signalent les problèmes de performances d'E/S au niveau d'une charge de travail de volume causée par des conflits dans un composant de cluster. Unified Manager analyse l'événement pour identifier toutes les charges de travail impliquées, le composant dans les conflits et si l'événement reste un problème à résoudre.

Unified Manager surveille la latence (temps de réponse) et les IOPS (opérations) des volumes d'un cluster. Lorsque d'autres charges de travail surfont un composant de cluster, par exemple, les conflits sont possibles et le composant ne peut pas fonctionner à un niveau optimal pour répondre aux demandes de charge de travail. Les performances des autres charges de travail qui utilisent le même composant peuvent être affectées, ce qui entraîne une augmentation des latences. Si la latence franchit le seuil de performance, Unified Manager déclenche un événement de performance et envoie une alerte e-mail vous avertie.

Analyse des événements

Unified Manager effectue les analyses suivantes, en s'appuyant sur les statistiques de performance des 15 derniers jours, pour identifier les workloads victime, les workloads dominants et le composant de cluster impliqué dans un événement :

- Identifie les charges de travail victimes dont la latence a dépassé le seuil de performance, qui est la limite supérieure de la plage attendue :
  - Pour les volumes des agrégats HDD ou Flash Pool (hybride), les événements sont déclenchés uniquement lorsque la latence dépasse 5 millisecondes (ms) et que les IOPS représentent plus de 10 opérations par seconde (OPS/s).
  - Pour les volumes situés sur des agrégats 100 % SSD ou des agrégats FabricPool (composites), les événements sont déclenchés uniquement lorsque la latence est supérieure à 1 ms et que les IOPS sont plus de 100 OPS/s.
- Identifie le composant de cluster dans les conflits.

  Si la latence des charges de travail victimes au niveau de l'interconnexion de cluster est supérieure à 1 ms, Unified Manager le traite comme important et déclenche un événement pour l'interconnexion de cluster.

- Identifie les charges de travail dominantes qui font l'objet d'une surutilisation du composant de cluster et qui l'entraînent des conflits.

- Classe les charges de travail impliquées, en fonction de leur déviation de l'utilisation ou de l'activité d'un composant du cluster, afin de déterminer les principaux changements d'utilisation du composant du cluster et les victimes les plus affectées.

Un événement peut se produire brièvement et se corriger après le composant qu'il utilise n'est plus en conflit. Un événement continu est un événement qui se produit de nouveau pour le même composant de cluster au cours d'un intervalle de cinq minutes et qui reste à l'état actif. Pour les événements continus, Unified Manager déclenche une alerte après avoir détecté le même événement à deux intervalles d'analyse consécutifs. Les événements qui ne sont pas résolus, qui ont un état nouveau, peuvent afficher différents messages de description en tant que charges de travail impliquées dans le changement d'événement.

Lorsqu'un événement est résolu, il reste disponible dans Unified Manager dans le cadre de l'enregistrement des anciens problèmes de performances d'un volume. Chaque événement possède un ID unique qui identifie le type d'événement et les volumes, le cluster et les composants de cluster impliqués.

Un seul volume peut être impliqué dans plusieurs événements simultanément.

État de l'événement

Les événements peuvent être dans l'un des États suivants :

- **Actif**

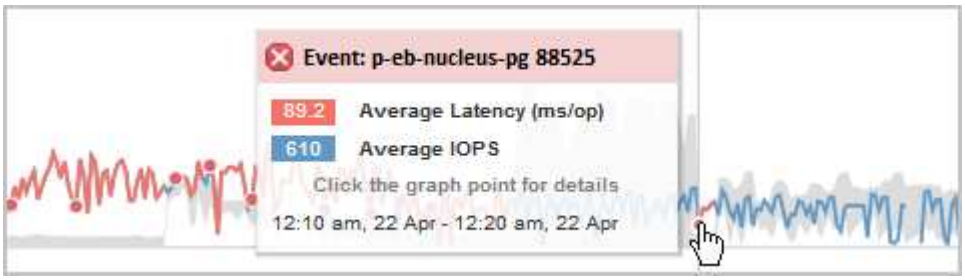
  Indique que l'événement de performance est actuellement actif (nouveau ou reconnu). Le problème à l'origine de l'incident n'a pas été corrigé lui-même ou n'a pas été résolu. Le compteur de performances de l'objet de stockage reste au-dessus du seuil de performance.

- **Obsolète**

  Indique que l'incident n'est plus actif. Le problème à l'origine de l'incident s'est corrigé ou a été résolu. Le compteur de performance de l'objet de stockage n'est plus au-dessus du seuil de performance.

Notification d'événement

Les alertes d'événement s'affichent sur la page tableaux de bord/Présentation, tableaux de bord/performances, page Détails performances/Volume et sont envoyées à des adresses e-mail spécifiées. Vous pouvez afficher des informations d'analyse détaillées sur un événement et obtenir des suggestions de résolution de cet événement sur la page Détails de l'événement.



Dans cet exemple, un événement est indiqué par un point rouge (●) Sur le graphique latence de la page Détails performances/Volume. Lorsque vous déplacez le curseur de la souris sur le point rouge, une fenêtre contextuelle contenant plus de détails sur l'événement et les options d'analyse s'affiche.

Interaction d'événement

Sur la page Performance/Volume Details, vous pouvez interagir avec les événements de l'une des manières suivantes :

- Le déplacement du pointeur sur un point rouge affiche un message indiquant l'ID de l'événement, la latence, le nombre d'opérations par seconde, ainsi que la date et l'heure de détection de l'événement.

  S'il y a plusieurs événements pour la même période, le message affiche le nombre d'événements, ainsi que la latence moyenne et les opérations par seconde pour le volume.

- Lorsque vous cliquez sur un seul événement, une boîte de dialogue affiche des informations plus détaillées sur l'événement, notamment les composants de cluster impliqués, similaires à la section Résumé de la page Détails de l'événement.

  Le composant en conflit est entouré et mis en évidence en rouge. Vous pouvez cliquer sur l'ID d'événement ou sur **Afficher l'analyse complète** pour afficher l'analyse complète sur la page Détails de l'événement. S'il existe plusieurs événements pour la même période, la boîte de dialogue affiche des détails sur les trois événements les plus récents. Vous pouvez cliquer sur un ID d'événement pour afficher l'analyse d'événement sur la page Détails de l'événement. S'il y a plus de trois événements pour la même période, le fait de cliquer sur le point rouge n'affiche pas la boîte de dialogue.

Comment Unified Manager détermine l'impact sur les performances d'un événement

Unified Manager utilise l'écart d'activité, d'utilisation, de débit d'écriture, de l'utilisation d'un composant du cluster ou de latence d'E/S (temps de réponse) pour une charge de travail afin de déterminer le niveau d'impact sur les performances d'une charge de travail. Ces informations déterminent le rôle de chaque charge de travail dans l'événement et leur classement sur la page Détails de l'événement.

Unified Manager compare les dernières valeurs analysées pour une charge de travail à la plage de valeurs attendue. La différence entre les valeurs analysées pour la dernière fois et la plage de valeurs attendue identifie les workloads pour lesquels les performances ont le plus été affectées par l'événement.

Supposons par exemple qu'un cluster contienne deux charges de travail : la charge De travail A et la charge de travail B. La plage prévue pour la charge de travail A est de 5-10 millisecondes par opération (ms/op) et sa latence réelle est généralement d'environ 7 ms/op. La plage attendue pour la charge de travail B est de 10-20 ms/op et sa latence réelle est généralement d'environ 15 ms/op. La latence prévue pour les deux charges de travail est très bonne. En raison des conflits sur le cluster, la latence des deux charges de travail augmente à 40 ms/activité, ce qui atteint le seuil de performances, qui correspond aux limites supérieures de la plage attendue et aux événements déclencheurs. L'écart de latence, entre les valeurs attendues et les valeurs supérieures au seuil de performances, pour la charge de travail A est d'environ 33 ms/op, et l'écart pour la charge de travail B est d'environ 25 ms/op. La latence des deux charges de travail atteint 40 ms/activité, mais la charge de travail A avait l'impact le plus important sur les performances, car elle avait l'écart de latence le plus élevé à 33 ms/opération.

Sur la page Détails de l'événement, dans la section diagnostic système, vous pouvez trier les charges de travail par variation de l'activité, de l'utilisation ou du débit d'un composant de cluster. Vous pouvez également trier les charges de travail par latence. Lorsque vous sélectionnez une option de tri, Unified Manager analyse l'écart en termes d'activité, d'utilisation, de débit ou de latence depuis que l'événement a été détecté à partir des valeurs attendues pour déterminer l'ordre de tri de la charge de travail. Pour la latence, les points rouges

(🔴) indiquez un seuil de performances franchissement par une charge de travail victime et l'impact qui en découle sur la latence. Chaque point rouge indique un niveau d'écart plus élevé de latence, ce qui vous aide à identifier les workloads victimes dont la latence a le plus été affectée par un événement.

Les composants du cluster et les conflits

Vous pouvez identifier les problèmes de performance du cluster lorsqu'un composant du cluster entre en conflit. Les performances des charges de travail de volumes qui utilisent le composant ralentissent et leur temps de réponse (latence) augmente pour les requêtes client, ce qui déclenche un événement dans Unified Manager.

Un composant en conflit ne peut pas se faire à un niveau optimal. Ses performances ont diminué, et la performance des autres composants et charges de travail du cluster, appelés *victimes*, peut avoir augmenté la latence. Pour mettre un composant à l'extérieur des conflits, vous devez réduire sa charge de travail ou augmenter sa capacité à gérer davantage de travail, de sorte que les performances puissent revenir à des niveaux normaux. Unified Manager collecte et analyse les performances des charges de travail toutes les cinq minutes. En effet, il ne détecte que lorsqu'un composant du cluster est constamment sur-utilisé. Les pics transitoires de surutilisation qui durent pendant une courte durée dans l'intervalle de cinq minutes ne sont pas détectés.

Par exemple, un agrégat de stockage peut être soumis à des conflits car une ou plusieurs charges de travail y sont en concurrence pour que leurs demandes d'E/S soient traitées. Des charges de travail peuvent être affectées sur l'agrégat, ce qui entraîne une baisse des performances. Pour réduire la quantité d'activité sur l'agrégat, plusieurs étapes sont possibles : déplacer une ou plusieurs charges de travail vers un agrégat moins occupé, par exemple, réduire la demande de charge de travail globale sur l'agrégat actuel. Pour un groupe de règles de qualité de service, vous pouvez ajuster la limite de débit ou déplacer les workloads vers un autre groupe de règles, de sorte que les charges de travail ne soient plus restreintes.

Unified Manager contrôle les composants de cluster suivants pour vous alerter en cas de conflit :

- **Réseau**

  Représente le temps d'attente des demandes d'E/S par les protocoles iSCSI ou Fibre Channel (FC) sur le cluster. Le temps d'attente correspond au temps passé à attendre la fin des transactions iSCSI Ready to Transfer (R2T) ou FCP Transfer Ready (XFER_RDY) pour que le cluster puisse répondre à une demande d'E/S. Si le composant réseau est en conflit, cela signifie qu'un temps d'attente élevé au niveau de la couche de protocole de bloc a un impact sur la latence d'une ou de plusieurs charges de travail.

- **Traitement réseau**

  Composant logiciel dans le cluster impliqué dans le traitement des E/S entre la couche de protocole et le cluster. Le traitement du réseau de traitement des nœuds a peut-être changé depuis la détection de l'événement. Si le composant de traitement de réseau est en conflit, son utilisation élevée au niveau du nœud de traitement réseau a un impact sur la latence d'une ou de plusieurs charges de travail.

- **Politique de qualité de service**

  Représente le groupe de règles de qualité de service (QoS) de stockage dont la charge de travail est membre. Si le composant de groupe de règles conflits, cela signifie que toutes les charges de travail du groupe de règles sont restreintes par la limite de débit définie, qui a un impact sur la latence d'une ou plusieurs de ces charges de travail.

- * Interconnexion de cluster*

La représente les câbles et adaptateurs avec lesquels les nœuds en cluster sont physiquement connectés. Si le composant d'interconnexion de cluster est en conflit, cela signifie un temps d'attente élevé pour les demandes d'E/S au niveau de l'interconnexion de cluster se répercute sur la latence d'une ou de plusieurs charges de travail.

- **Traitement de données**

  Composant logiciel dans le cluster impliqué dans le traitement des E/S entre le cluster et l'agrégat de stockage qui contient la charge de travail. Le traitement des données de traitement du nœud peut avoir changé depuis la détection de l'événement. Si le composant de traitement des données conflit, une utilisation élevée au niveau du nœud de traitement des données affecte la latence d'un ou de plusieurs workloads.

- **Ressources MetroCluster**

  La représente les ressources MetroCluster, y compris la NVRAM et les liens ISL, utilisés pour mettre en miroir les données entre les clusters dans une configuration MetroCluster. Si le composant MetroCluster rencontre des conflits, il s'agit d'un débit d'écriture élevé avec les charges de travail sur le cluster local ou d'un problème d'état de santé de la liaison ayant un impact sur la latence d'une ou de plusieurs charges de travail sur le cluster local. Si le cluster ne se trouve pas dans une configuration MetroCluster, cette icône n'est pas affichée.

- **Agrégat ou agrégat SSD**

  Agrégat de stockage sur lequel les charges de travail s'exécutent. Si le composant de l'agrégat est en conflit, une utilisation élevée de l'agrégat a un impact sur la latence d'une ou de plusieurs charges de travail. Un agrégat se compose de tous les disques durs ou d'un mélange de disques durs et de disques SSD (un agrégat Flash Pool). Un « agrégat SD » se compose de tous les SSD (un agrégat 100 % Flash), ou d'une combinaison de SSD et d'un niveau cloud (un agrégat FabricPool).

- * Latence cloud*

  Représente le composant logiciel du cluster impliqué dans le traitement des E/S entre le cluster et le niveau cloud sur lequel les données utilisateur sont stockées. Si le composant de latence dans le cloud conflits, une grande quantité de lectures sur les volumes hébergés sur le Tier cloud ont une incidence sur la latence d'un ou de plusieurs workloads.

- **SnapMirror de synchronisation**

  Représente le composant logiciel du cluster impliqué dans la réplication des données utilisateur depuis le volume primaire vers le volume secondaire dans une relation SnapMirror synchrone. Si le composant SnapMirror synchrone entre en conflit, l'activité des opérations SnapMirror synchrone a un impact sur la latence d'un ou de plusieurs workloads.

Rôles des charges de travail impliquées dans un événement de performance

Unified Manager utilise des rôles pour identifier la participation d'une charge de travail en cas de performance. Les rôles sont les victimes, les taureaux et les requins. Une charge de travail définie par l'utilisateur peut être une victime, un tyran et un requin en même temps.

Rôle	Description
Victime	Charge de travail définie par l'utilisateur dont les performances ont diminué en raison des autres charges de travail, appelées « bullies », qui sont sur-utilisées lors de l'utilisation d'un composant du cluster. Seules les charges de travail définies par l'utilisateur sont identifiées comme victimes. Unified Manager identifie les charges de travail victimes en fonction de leur écart de latence, où la latence réelle, pendant un événement, a été considérablement améliorée par rapport à la plage de latence prévue.
Intimider	Une charge de travail définie par l'utilisateur ou définie par le système dont l'utilisation excessive d'un composant de cluster a entraîné une diminution des performances d'autres charges de travail, appelées « victimes ». Unified Manager identifie les workloads dominants en fonction de leur déviation par l'utilisation d'un composant de cluster, où l'utilisation réelle, au cours d'un événement, a considérablement augmenté à partir de sa plage d'utilisation prévue.
Requin	Charge de travail définie par l'utilisateur, avec l'utilisation la plus élevée d'un composant de cluster, et non pas toutes les charges de travail impliquées dans un événement. Unified Manager identifie les charges de travail Shark en fonction de leur utilisation d'un composant de cluster pendant un événement.

Les charges de travail d'un cluster peuvent partager la plupart des composants du cluster, tels que les agrégats de stockage et la CPU pour le traitement du réseau et des données. Lorsqu'une charge de travail, par exemple un volume, augmente l'utilisation d'un composant de cluster au point que le composant ne peut pas répondre efficacement aux exigences de la charge de travail, le composant engendre des conflits. La charge de travail sur-utilisation d'un composant de cluster est un phénomène tyran. Les autres charges de travail qui partagent ces composants, et dont la performance est impactée par le tyran, sont les victimes. L'activité provenant des charges de travail définies par le système, telles que la déduplication ou les copies Snapshot, peut également créer des « brimades ».

Lorsqu'Unified Manager détecte un événement, il identifie tous les workloads et composants de cluster impliqués, notamment les workloads dominants qui ont causé l'événement, le composant de cluster en conflit et les workloads victimes dont les performances ont diminué en raison de l'augmentation de l'activité des workloads dominants.

> Si Unified Manager ne peut pas identifier les charges de travail dominantes, cette alerte s'applique uniquement aux charges de travail victimes et au composant de cluster concerné.

Unified Manager est capable d'identifier les charges de travail victimes de charges de travail dominantes. Il peut également y avoir une identification lorsque ces mêmes charges de travail deviennent des charges de travail dominantes. Un workload peut être un tyran à lui-même. Par exemple, une charge de travail élevée au ralenti par une limite de groupe de règles entraîne le restriction de toutes les charges de travail du groupe de règles, y compris de celles-ci. Une charge de travail dominante ou victime dans un événement de performance

continu peut changer son rôle ou ne plus y participer. Sur la page Détails performances/Volume, dans le tableau liste d’événements, lorsque le volume sélectionné change son rôle de participant, la date et l’heure du changement de rôle s’affichent.

Informations sur le copyright

Copyright © 2024 NetApp, Inc. Tous droits réservés. Imprimé aux États-Unis. Aucune partie de ce document protégé par copyright ne peut être reproduite sous quelque forme que ce soit ou selon quelque méthode que ce soit (graphique, électronique ou mécanique, notamment par photocopie, enregistrement ou stockage dans un système de récupération électronique) sans l'autorisation écrite préalable du détenteur du droit de copyright.

Les logiciels dérivés des éléments NetApp protégés par copyright sont soumis à la licence et à l'avis de non-responsabilité suivants :

CE LOGICIEL EST FOURNI PAR NETAPP « EN L'ÉTAT » ET SANS GARANTIES EXPRESSES OU TACITES, Y COMPRIS LES GARANTIES TACITES DE QUALITÉ MARCHANDE ET D'ADÉQUATION À UN USAGE PARTICULIER, QUI SONT EXCLUES PAR LES PRÉSENTES. EN AUCUN CAS NETAPP NE SERA TENU POUR RESPONSABLE DE DOMMAGES DIRECTS, INDIRECTS, ACCESSOIRES, PARTICULIERS OU EXEMPLAIRES (Y COMPRIS L'ACHAT DE BIENS ET DE SERVICES DE SUBSTITUTION, LA PERTE DE JOUISSANCE, DE DONNÉES OU DE PROFITS, OU L'INTERRUPTION D'ACTIVITÉ), QUELLES QU'EN SOIENT LA CAUSE ET LA DOCTRINE DE RESPONSABILITÉ, QU'IL S'AGISSE DE RESPONSABILITÉ CONTRACTUELLE, STRICTE OU DÉLICTUELLE (Y COMPRIS LA NÉGLIGENCE OU AUTRE) DÉCOULANT DE L'UTILISATION DE CE LOGICIEL, MÊME SI LA SOCIÉTÉ A ÉTÉ INFORMÉE DE LA POSSIBILITÉ DE TELS DOMMAGES.

NetApp se réserve le droit de modifier les produits décrits dans le présent document à tout moment et sans préavis. NetApp décline toute responsabilité découlant de l'utilisation des produits décrits dans le présent document, sauf accord explicite écrit de NetApp. L'utilisation ou l'achat de ce produit ne concède pas de licence dans le cadre de droits de brevet, de droits de marque commerciale ou de tout autre droit de propriété intellectuelle de NetApp.

Le produit décrit dans ce manuel peut être protégé par un ou plusieurs brevets américains, étrangers ou par une demande en attente.

LÉGENDE DE RESTRICTION DES DROITS : L'utilisation, la duplication ou la divulgation par le gouvernement sont sujettes aux restrictions énoncées dans le sous-paragraphe (b)(3) de la clause Rights in Technical Data-Noncommercial Items du DFARS 252.227-7013 (février 2014) et du FAR 52.227-19 (décembre 2007).

Les données contenues dans les présentes se rapportent à un produit et/ou service commercial (tel que défini par la clause FAR 2.101). Il s'agit de données propriétaires de NetApp, Inc. Toutes les données techniques et tous les logiciels fournis par NetApp en vertu du présent Accord sont à caractère commercial et ont été exclusivement développés à l'aide de fonds privés. Le gouvernement des États-Unis dispose d'une licence limitée irrévocable, non exclusive, non cessible, non transférable et mondiale. Cette licence lui permet d'utiliser uniquement les données relatives au contrat du gouvernement des États-Unis d'après lequel les données lui ont été fournies ou celles qui sont nécessaires à son exécution. Sauf dispositions contraires énoncées dans les présentes, l'utilisation, la divulgation, la reproduction, la modification, l'exécution, l'affichage des données sont interdits sans avoir obtenu le consentement écrit préalable de NetApp, Inc. Les droits de licences du Département de la Défense du gouvernement des États-Unis se limitent aux droits identifiés par la clause 252.227-7015(b) du DFARS (février 2014).

Informations sur les marques commerciales

NETAPP, le logo NETAPP et les marques citées sur le site http://www.netapp.com/TM sont des marques déposées ou des marques commerciales de NetApp, Inc. Les autres noms de marques et de produits sont des marques commerciales de leurs propriétaires respectifs.