Gestion des performances grâce à la capacité en termes de performances et aux informations d'IOPS disponibles

Active IQ Unified Manager

NetApp
March 26, 2025

This PDF was generated from https://docs.netapp.com/fr-fr/active-iq-unified-manager-97/online-help/concept-what-performance-capacity-used-is.html on March 26, 2025. Always check docs.netapp.com for the latest.

Sommaire

Gestion des performances grâce à la capacité en termes de performances et aux informations d'IOPS disponibles . . . 1
- Quelle est la capacité des performances utilisée . . . 1
- Signification de la capacité en termes de performances utilisée . . . 2
- Les IOPS disponibles . . . 3
- Affichage des valeurs de capacité des nœuds et des performances des agrégats utilisées . . . 4
- Affichage des valeurs d'IOPS disponibles du nœud et de l'agrégat . . . 5
- Affichage des graphiques des compteurs de capacité de performances pour identifier les problèmes . . . 6
  - Description de la tâche . . . 6
  - Étapes . . . 7
- Conditions du seuil de performance utilisé par la capacité de performance . . . 8
- Utilisation du compteur de performances utilisé pour gérer les performances . . . 9

Gestion des performances grâce à la capacité en termes de performances et aux informations d'IOPS disponibles

Performance Capacity indique le débit que vous pouvez obtenir d'une ressource sans dépasser les performances utiles de cette ressource. Lorsqu'il est utilisé des compteurs de performances existants, la capacité de performances est le point où l'utilisation maximale est atteint depuis un nœud ou un agrégat avant que la latence ne devienne un problème.

Unified Manager collecte les statistiques de capacité des performances à partir des nœuds et des agrégats de chaque cluster. *Performance Capacity used* représente le pourcentage de capacité de performance actuellement utilisée et *performance free* représente le pourcentage de capacité de performance qui est toujours disponible.

Tandis que la fonction libération de la capacité qui constitue un pourcentage de ressource encore disponible, *Available IOPS* représente le nombre d'IOPS pouvant être ajoutés à la ressource avant d'atteindre la capacité de performance maximale. Cette mesure vous permet d'ajouter des charges de travail d'un nombre prédéterminé d'IOPS à une ressource.

Le contrôle des informations sur la capacité en matière de performances présente les avantages suivants :

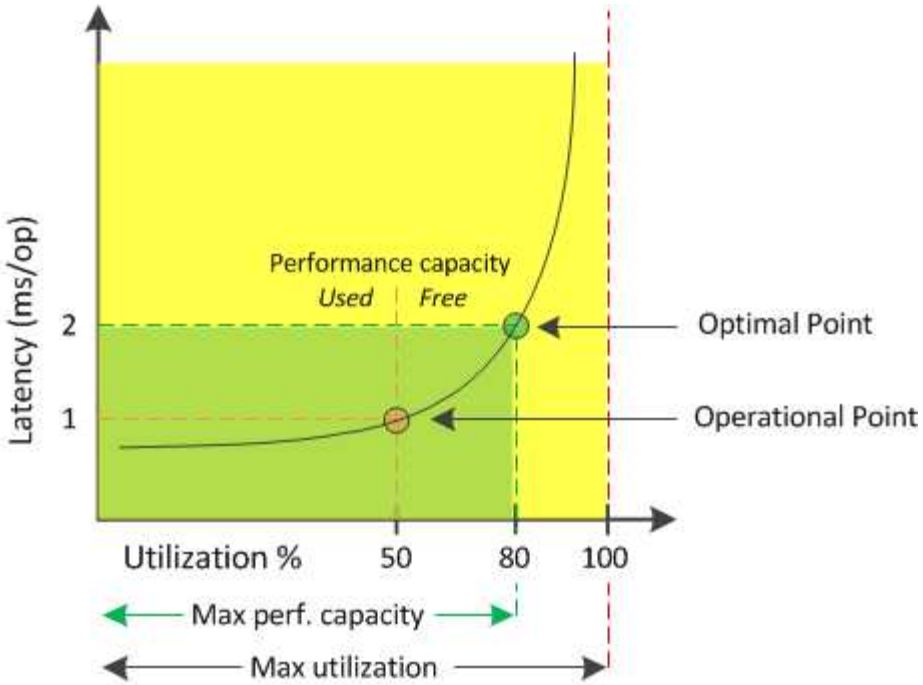
- Aide au provisionnement et à l'équilibrage des flux de travail.
- Vous aide à éviter de surcharger un nœud ou de repousser ses ressources au-delà du point optimal, réduisant ainsi la nécessité de résoudre le problème.
- Elle vous permet de déterminer avec plus de précision les endroits où un équipement de stockage supplémentaire peut être nécessaire.

Quelle est la capacité des performances utilisée

Le compteur de performances utilisé vous permet de déterminer si les performances d'un nœud ou d'un agrégat atteignent un point où les performances peuvent se dégrader si les charges de travail augmentent. Il peut également vous indiquer si un nœud ou un agrégat est actuellement utilisé pendant des périodes spécifiques. Les performances utilisées sont similaires aux taux d'utilisation, mais la première fournit des informations supplémentaires sur les capacités de performances disponibles dans une ressource physique pour une charge de travail spécifique.

Les données de performance et de capacité sont disponibles uniquement lorsque les nœuds d'un cluster sont installés avec le logiciel ONTAP 9.0 ou une version ultérieure.

La capacité à performances optimales utilisées est au point de permettre à un nœud ou un agrégat d'optimiser l'utilisation et la latence (temps de réponse) et d'être utilisé de manière efficace. Un exemple de latence par rapport à une courbe d'utilisation est présenté pour un agrégat dans la figure suivante.



Dans cet exemple, le *point opérationnel* identifie que l'agrégat fonctionne actuellement à 50 % d'utilisation avec une latence de 1.0 ms/op. En se basant sur les statistiques collectées par l'agrégat, Unified Manager détermine que des performances supplémentaires sont disponibles pour cet agrégat. Dans cet exemple, le *optimal point* est identifié comme le point où l'agrégat est à 80% d'utilisation avec une latence de 2.0 ms/op. Vous pouvez donc ajouter davantage de volumes et de LUN à cet agrégat, afin que vos systèmes soient utilisés plus efficacement.

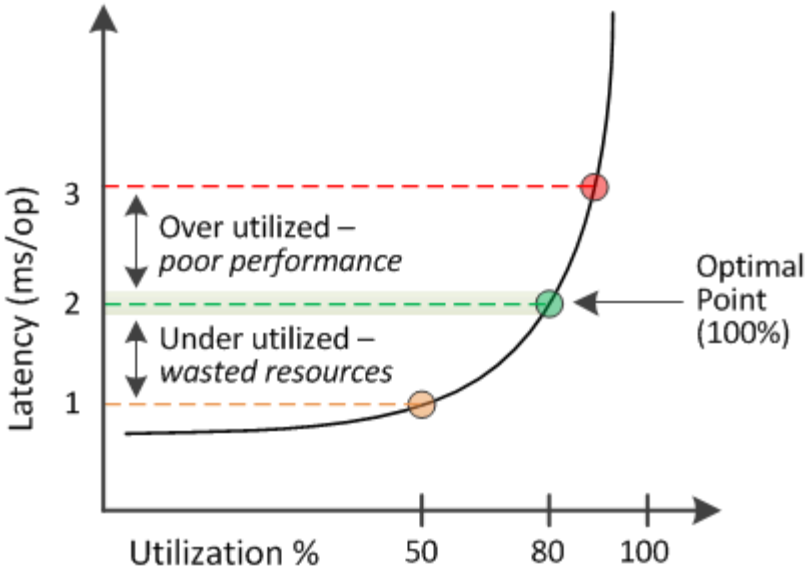
Le compteur de capacités de performances utilisé devrait être un nombre supérieur au compteur de « utilisation », car des capacités de performances augmentent l'impact sur la latence. Par exemple, si un nœud ou un agrégat est utilisé à 70 %, la valeur de la capacité de performances peut se situe dans la plage de 80 à 100 %, selon la valeur de latence.

Dans certains cas, le compteur d'utilisation peut cependant être plus élevé sur la page Tableau de bord. Cet aspect est normal car le tableau de bord actualise les valeurs de compteur actuelles à chaque période de collecte. Il n'affiche pas les moyennes sur un certain temps, comme les autres pages de l'interface utilisateur Unified Manager. Le compteur de performances utilisé est le mieux utilisé comme indicateur de performance moyenne sur une période de temps, alors que le compteur d'utilisation est le mieux utilisé pour déterminer l'utilisation instantanée d'une ressource.

Signification de la capacité en termes de performances utilisée

La valeur de performance utilisée permet d'identifier les nœuds et les agrégats actuellement sur-utilisés ou sous-utilisés. Vous pouvez ainsi redistribuer les charges de travail afin d'améliorer l'efficacité de vos ressources de stockage.

La figure suivante montre la courbe de latence par rapport à l'utilisation d'une ressource et identifie, avec des points de couleur, trois zones où le point opérationnel actuel peut être localisé.



- Un pourcentage de capacité de performances utilisé égal à 100 est au point optimal.

  À ce stade, les ressources sont utilisées efficacement.

- Un pourcentage de capacité de performances utilisé supérieur à 100 indique que le nœud ou l'agrégat est sur-exploité et que les charges de travail bénéficient de performances sous-optimales.

  Aucune nouvelle charge de travail ne doit être ajoutée à la ressource et la redistribution des charges de travail existantes peut s'avérer nécessaire.

- Un pourcentage de capacité de performances utilisé inférieur à 100 indique que le nœud ou l'agrégat est sous-utilisé, et que les ressources ne sont pas utilisées efficacement.

  Il est possible d'ajouter davantage de charges de travail à une ressource.

Contrairement à l'utilisation, le pourcentage de capacité haute performance utilisé peut être supérieur à 100 %. Il n'y a pas de pourcentage maximal, mais les ressources sont généralement comprises entre 110 et 140 % lorsqu'elles sont sur-exploitées. Des pourcentages plus élevés indiquerait une ressource avec des problèmes graves.

Les IOPS disponibles

Le compteur IOPS disponibles identifie le nombre restant d'IOPS pouvant être ajouté à un nœud ou à un agrégat avant que la ressource n'atteigne sa limite. Le nombre total d'IOPS qu'un nœud peut fournir est basé sur les caractéristiques physiques du nœud—par exemple, le nombre de processeurs, la vitesse du processeur et la quantité de RAM. Le nombre total d'opérations d'E/S par seconde qu'un agrégat peut fournir dépend des propriétés physiques des disques, par exemple un disque SATA, SAS ou SSD.

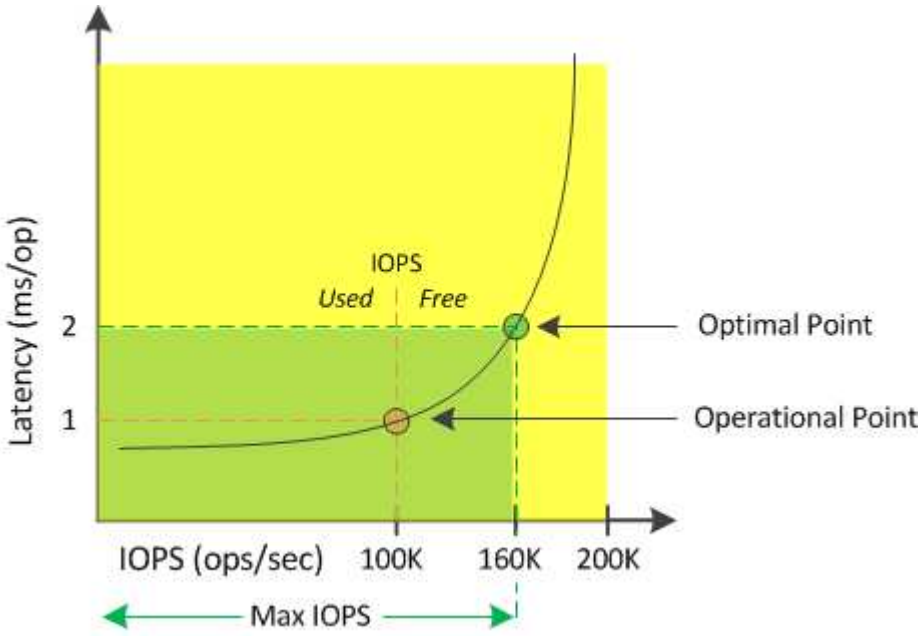
Le compteur de performances disponible affiche le pourcentage de ressource qui reste disponible, mais le compteur IOPS disponible indique le nombre exact d'IOPS (charges de travail) à ajouter à une ressource avant d'atteindre la capacité de performance maximale.

Par exemple, si vous utilisez une paire de systèmes de stockage FAS2520 et FAS8060, une valeur sans

capacité en termes de performances de 30 % indique que vous bénéficiez de capacité de performances gratuites. Toutefois, cette valeur n'apporte aucune visibilité sur le nombre de charges de travail que vous pouvez déployer sur ces nœuds. Le compteur IOPS disponible peut indiquer que vous disposez de 500 000 IOPS disponibles sur le système FAS8060, mais que seuls 100 000 IOPS sont disponibles sur le système FAS2520.

Les données IOPS disponibles sont uniquement disponibles lorsque les nœuds d'un cluster sont installés avec le logiciel ONTAP 9.0 ou une version ultérieure.

La figure suivante présente un exemple de latence par rapport aux courbes IOPS pour un nœud.



Le nombre maximal d'IOPS qu'une ressource peut fournir est le nombre d'IOPS lorsque la capacité de performance utilisée compteur est de 100 % (le point optimal). Le point opérationnel identifie que le nœud fonctionne actuellement à 100 000 IOPS avec une latence de 1.0 ms/op. En fonction des statistiques collectées à partir du nœud, Unified Manager détermine que les IOPS maximales du nœud sont de 160 000, ce qui signifie que 60 000 IOPS sont disponibles ou libres. Vous pouvez donc ajouter des workloads à ce nœud afin que vos systèmes soient utilisés plus efficacement.

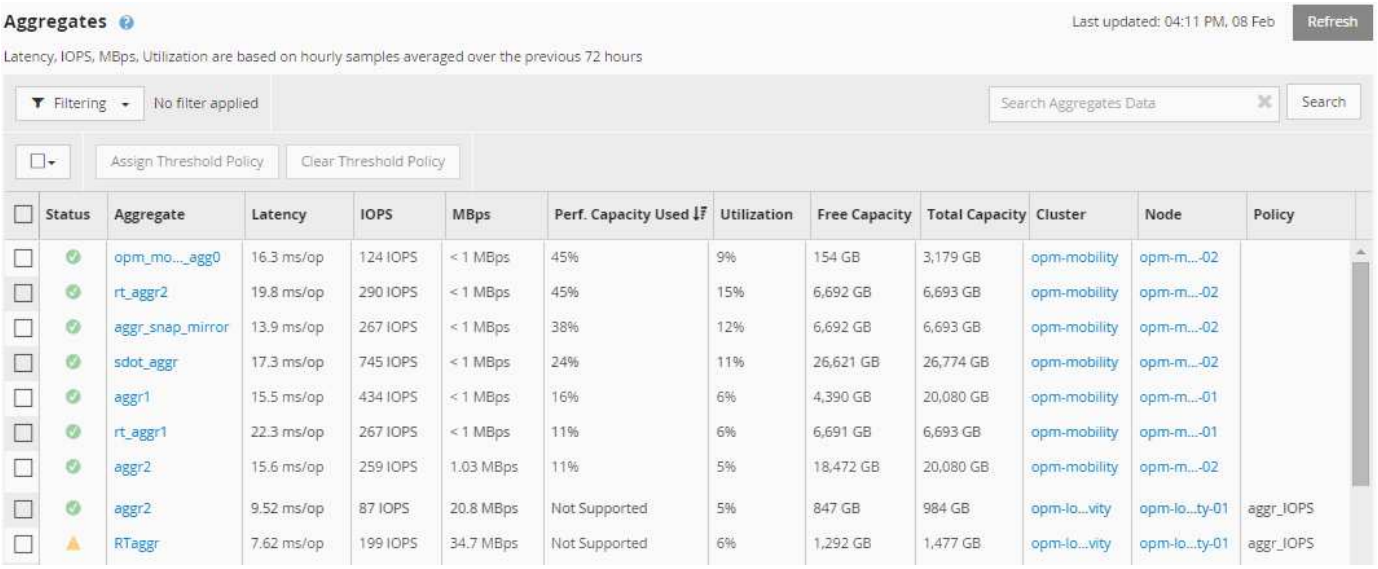
Lorsque l'activité de la ressource est minimale, la valeur des IOPS disponibles est calculée en supposant qu'une charge de travail générique s'appuie sur environ 4,500 000 IOPS par cœur de processeur. En effet, Unified Manager n'a pas de données pour estimer avec précision les caractéristiques de la charge de travail traitée.

Affichage des valeurs de capacité des nœuds et des performances des agrégats utilisées

Vous pouvez contrôler les valeurs de performance de la capacité utilisée pour tous les nœuds ou pour tous les agrégats d'un cluster, ou afficher les détails d'un nœud ou agrégat unique.

Les valeurs utilisées relatives à la capacité de performance apparaissent dans le tableau de bord, les pages Inventaire des performances, la page collaborateurs les plus performants, la page Créer une stratégie de seuil,

les pages Explorateur de performances et les graphiques détaillés. Par exemple, la page performances : tous les agrégats fournit une colonne Perf. La capacité utilisée pour afficher la valeur des performances utilisées pour tous les agrégats.



Status	Aggregate	Latency	IOPS	MBps	Perf. Capacity Used	Utilization	Free Capacity	Total Capacity	Cluster	Node	Policy
	opm_mo..._agg0	16.3 ms/op	124 IOPS	< 1 MBps	45%	9%	154 GB	3,179 GB	opm-mobility	opm-m...-02	
	rt_aggr2	19.8 ms/op	290 IOPS	< 1 MBps	45%	15%	6,692 GB	6,693 GB	opm-mobility	opm-m...-02	
	aggr_snap_mirror	13.9 ms/op	267 IOPS	< 1 MBps	38%	12%	6,692 GB	6,693 GB	opm-mobility	opm-m...-02	
	sdot_aggr	17.3 ms/op	745 IOPS	< 1 MBps	24%	11%	26,621 GB	26,774 GB	opm-mobility	opm-m...-02	
	aggr1	15.5 ms/op	434 IOPS	< 1 MBps	16%	6%	4,390 GB	20,080 GB	opm-mobility	opm-m...-01	
	rt_aggr1	22.3 ms/op	267 IOPS	< 1 MBps	11%	6%	6,691 GB	6,693 GB	opm-mobility	opm-m...-01	
	aggr2	15.6 ms/op	259 IOPS	1.03 MBps	11%	5%	18,472 GB	20,080 GB	opm-mobility	opm-m...-02	
	aggr2	9.52 ms/op	87 IOPS	20.8 MBps	Not Supported	5%	847 GB	984 GB	opm-lo...vity	opm-lo...ty-01	aggr_IOPS
	RTaggr	7.62 ms/op	199 IOPS	34.7 MBps	Not Supported	6%	1,292 GB	1,477 GB	opm-lo...vity	opm-lo...ty-01	aggr_IOPS

L'état « N/A » s'affiche lorsque les nœuds ne sont pas installés avec le logiciel ONTAP 9.0 ou version ultérieure.

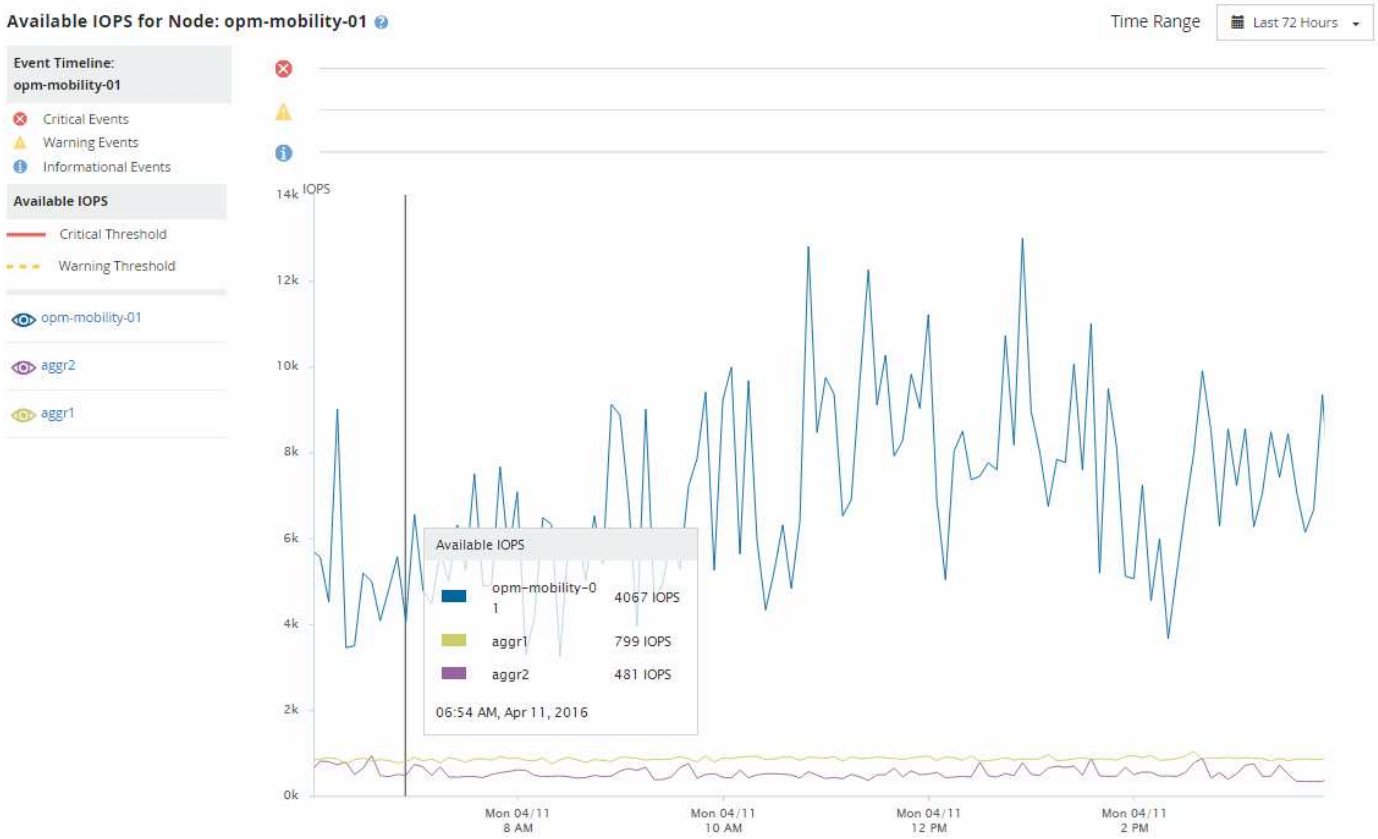
Le contrôle du compteur de performances utilisé vous permet d'identifier ce qui suit :

- Que les nœuds ou les agrégats de n'importe quel cluster disposent d'une capacité haute performance utilisée
- Que des nœuds ou des agrégats de tous les clusters disposent d'événements de capacité de performances active
- Les nœuds et les agrégats qui présentent la capacité de performance la plus élevée et la plus faible au sein d'un cluster
- Les valeurs des compteurs de latence et d'utilisation associées à des nœuds ou des agrégats qui possèdent des valeurs de capacité haute performance utilisée
- L'impact des valeurs de performance utilisées pour les nœuds d'une paire haute disponibilité sera affecté en cas de panne de l'un des nœuds
- Les volumes et les LUN les plus sollicités d'un agrégat dont les capacités de performances sont élevées

Affichage des valeurs d'IOPS disponibles du nœud et de l'agrégat

Vous pouvez contrôler les valeurs d'IOPS disponibles pour tous les nœuds ou pour tous les agrégats d'un cluster, ou afficher les détails d'un seul nœud ou agrégat.

Les valeurs d'IOPS disponibles s'affichent dans les graphiques de la page de l'explorateur de performances. Par exemple, lorsque vous affichez un nœud dans la page Performance/Node Explorer, vous pouvez sélectionner le compteur « IOPS disponibles » dans la liste afin de comparer les valeurs d'IOPS disponibles pour plusieurs agrégats de ce nœud.



La surveillance du compteur IOPS disponible vous permet d'identifier :

- Les nœuds ou les agrégats qui disposent des valeurs d'IOPS les plus élevées pour déterminer l'emplacement où les futurs workloads peuvent être déployés.
- Pour identifier les ressources dont vous devez contrôler les problèmes de performance futurs, les nœuds ou les agrégats disposant des valeurs d'IOPS les plus faibles.
- Les volumes et les LUN les plus sollicités d'un agrégat qui dispose d'une faible valeur des IOPS disponibles.

Affichage des graphiques des compteurs de capacité de performances pour identifier les problèmes

Vous pouvez consulter les diagrammes de capacité des performances utilisés pour les nœuds et les agrégats sur la page de l'explorateur de performances. Vous pouvez ainsi consulter des données détaillées sur la capacité de performances pour les nœuds et les agrégats sélectionnés pendant un délai spécifique.

Description de la tâche

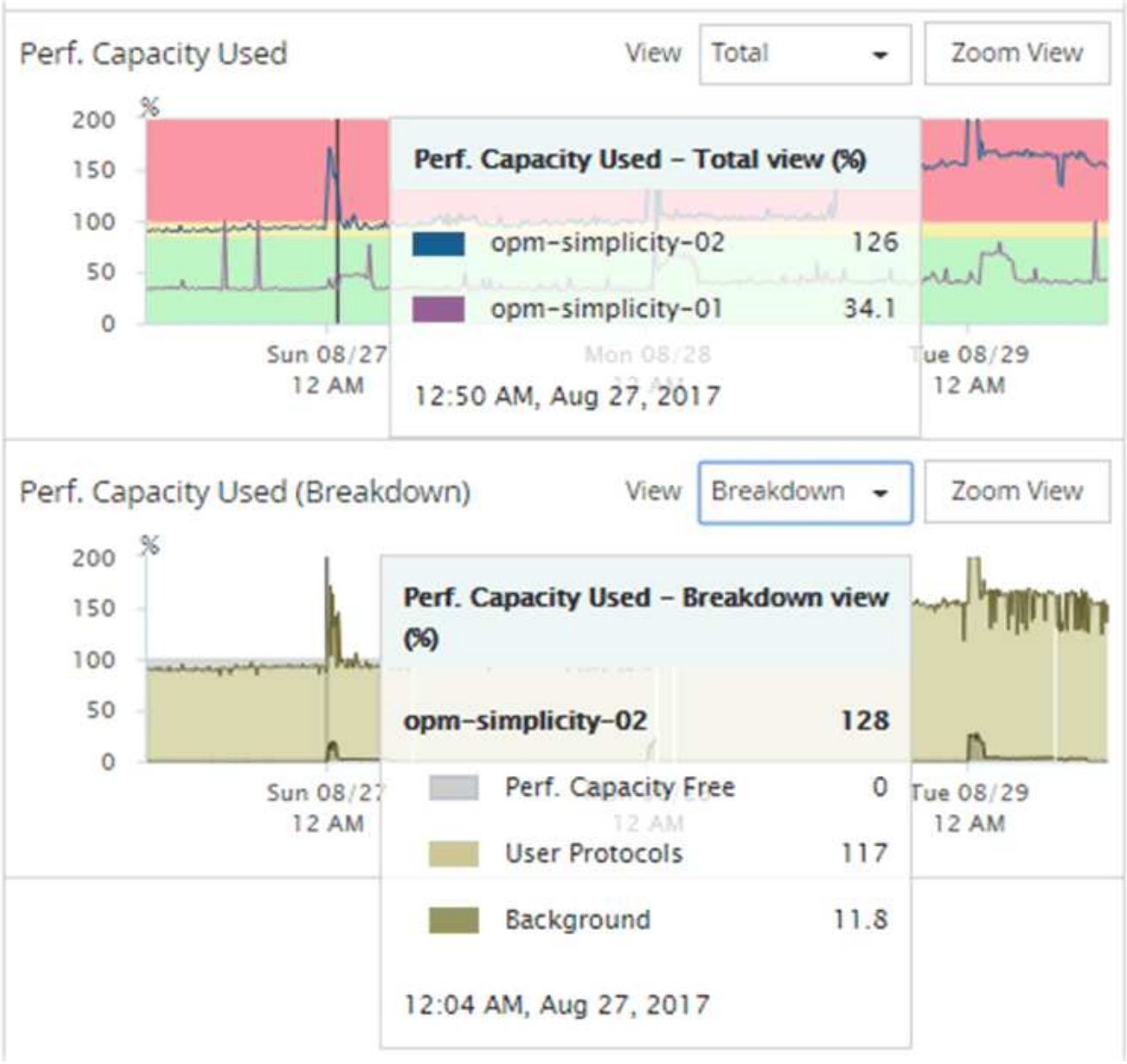
Le tableau des compteurs standard affiche les valeurs de capacité de performances utilisées pour les nœuds ou agrégats sélectionnés. Le tableau des compteurs d'analyse affiche les valeurs de capacité de performance totale pour l'objet racine, séparées en utilisation en fonction des protocoles utilisateur par rapport aux processus du système en arrière-plan. En outre, la capacité des performances libres est également indiquée.

Comme certaines activités d'arrière-plan associées à la gestion du système et des données sont identifiées comme des charges de travail utilisateur et classées comme protocoles utilisateur, le pourcentage de protocoles utilisateur peut paraître artificiellement élevé lors de l'exécution de ces processus. Ces processus s'exécutent généralement autour de minuit lorsque le niveau d'utilisation du cluster est bas. Si vous voyez un pic d'activité du protocole utilisateur vers minuit, vérifiez si les tâches de sauvegarde du cluster ou d'autres activités en arrière-plan sont configurées pour s'exécuter à ce moment-là.

Étapes

1. Sélectionnez l'onglet **Explorer** à partir d'une page de noeud ou d'agrégat **Landing**.
2. Dans le volet **compteur graphiques**, cliquez sur **choisir les graphiques**, puis sélectionnez **Perf. Tableau capacité utilisée**.
3. Faites défiler vers le bas jusqu'à ce que vous puissiez afficher la carte.

   Les couleurs du graphique standard indiquent lorsque l'objet se trouve dans la plage optimale (jaune), lorsque l'objet est sous-utilisé (vert) et lorsque l'objet est surutilisé (rouge). Le tableau décomposition affiche des détails détaillés sur la capacité de performances pour l'objet racine uniquement.



4. Si vous souhaitez afficher l'un ou l'autre des graphiques en format plein format, cliquez sur **vue Zoom**.

   Vous pouvez ainsi ouvrir plusieurs diagrammes dans une fenêtre distincte afin de comparer les valeurs de capacité utilisée avec les valeurs d'IOPS ou de Mo/sec sur la même période.

Conditions du seuil de performance utilisé par la capacité de performance

Vous pouvez créer des règles de seuils de performances définies par l'utilisateur, de sorte que les événements se déclenchent lorsque la capacité de performance utilisée pour un nœud ou un agrégat dépasse le seuil de capacité de performance défini défini.

En outre, les nœuds peuvent être configurés avec une règle de seuil « prise en charge de la capacité de performances utilisée ». Cette règle de seuil total les statistiques de performance utilisées pour les deux nœuds d'une paire HA afin de déterminer si la capacité de l'un des nœuds serait insuffisante si l'autre nœud

tombe en panne. Étant donné que la charge de travail lors du basculement est la combinaison des charges de travail des deux nœuds partenaires, la même capacité de performance utilisée dans la règle de basculement peut être appliquée aux deux nœuds.

> Cette équivalence des performances utilisée est généralement vrai entre les nœuds. Cependant, en cas de trafic croisé sur l'un des nœuds via son partenaire de basculement, la capacité de performance totale utilisée lors de l'exécution de toutes les charges de travail sur un nœud partenaire par rapport aux autres nœuds partenaires peut légèrement différente en fonction du nœud défaillant.

Les conditions de performance utilisées peuvent également être utilisées comme paramètres de seuil de performance secondaire pour créer une règle de seuil de combinaison lors de la définition de seuils pour les LUN et les volumes. La condition de performance utilisée est appliquée à l'agrégat ou au nœud sur lequel réside le volume ou la LUN. Par exemple, vous pouvez créer une stratégie de seuil de combinaison à l'aide des critères suivants :

Objet de stockage	Compteur de performances	Seuil d'avertissement	Seuil critique
Durée	Volumétrie	Latence	15 ms/op
25 ms/op	20 minutes	Agrégat	Capacité utilisée

Les stratégies de seuil de combinaison n'entraînent la génération d'un événement que lorsque les deux conditions sont enfreintes pendant toute la durée.

Utilisation du compteur de performances utilisé pour gérer les performances

Généralement, les entreprises souhaitent profiter d'une capacité de performance utilisée inférieure à 100 et pouvoir utiliser les ressources de manière efficace tout en réservant une capacité de performance supplémentaire pour supporter les pics de demande. Vous pouvez utiliser des règles de seuil pour personnaliser l'envoi d'alertes pour des valeurs de capacité haute performance utilisée.

Vous pouvez établir des objectifs spécifiques en fonction de vos exigences de performances. Par exemple, les entreprises de services financiers pourraient réserver davantage de capacité de rendement pour garantir l'exécution opportune des opérations. Ces entreprises peuvent vouloir définir des seuils de capacité pour les performances utilisés dans une plage de 70-80 %. Dans ce secteur, les entreprises de fabrication dont les marges sont plus faibles pourraient choisir de réserver une capacité moins performante si elles risquent de mieux gérer les coûts IT. Ces entreprises peuvent définir des seuils de capacité pour les performances utilisés dans les plages de 85-95 %.

Lorsque la valeur de performance capacité utilisée dépasse le pourcentage défini dans une règle de seuil définie par l'utilisateur, Unified Manager envoie un e-mail d'alerte et l'ajoute à la page Event Inventory. Les problèmes potentiels sont ainsi gérés avant qu'ils n'affectent les performances. Ces événements peuvent également servir d'indicateurs pour déplacer et modifier les charges de travail au sein de vos nœuds et agrégats.

Informations sur le copyright

Copyright © 2025 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.