

Gestion des capacités et des performances : Livre blanc sur les pratiques recommandées

Table des matières

[Introduction](#)

[Présentation de la gestion des capacités et des performances](#)

[CPU](#)

[Fond de panier ou E/S](#)

[Mémoire](#)

[Taille des interfaces et des tuyaux](#)

[Mise en file d'attente, latence et gigue](#)

[Vitesse et distance](#)

[Caractéristiques des applications](#)

[Meilleures pratiques de gestion des capacités et des performances](#)

[Gestion de niveau de service](#)

[Analyse de simulation du réseau et des applications](#)

[Planification et tendances](#)

[Gestion des exceptions](#)

[Gestion QoS](#)

[Collecte et reporting des informations de capacité](#)

[Déterminer vos besoins](#)

[Définir un processus](#)

[Définir les zones de capacité](#)

[Définir les variables de capacité](#)

[Interprétation des données](#)

[Informations connexes](#)

Introduction

La haute disponibilité du réseau est une condition requise essentielle dans les grandes entreprises et les réseaux de fournisseur de service. Les responsables réseau doivent faire face à des défis croissants pour fournir une meilleure disponibilité, y compris les défis liés aux temps d'arrêt non programmés, au manque d'expertise, à l'insuffisance des outils, à la complexité des technologies, à la consolidation du marché et à la concurrence. Les outils de gestion de la capacité et des performances aident les responsables réseau à atteindre de nouveaux objectifs d'entreprise ainsi qu'une disponibilité et des performances du réseau homogènes.

Ce document traite des sujets suivants :

- Problèmes généraux de capacité et de performances, y compris les risques et les problèmes potentiels de capacité au sein des réseaux.
- Meilleures pratiques de gestion des capacités et des performances, y compris les analyses de simulation, la planification initiale, les tendances, la gestion des exceptions et la gestion de la

qualité de service.

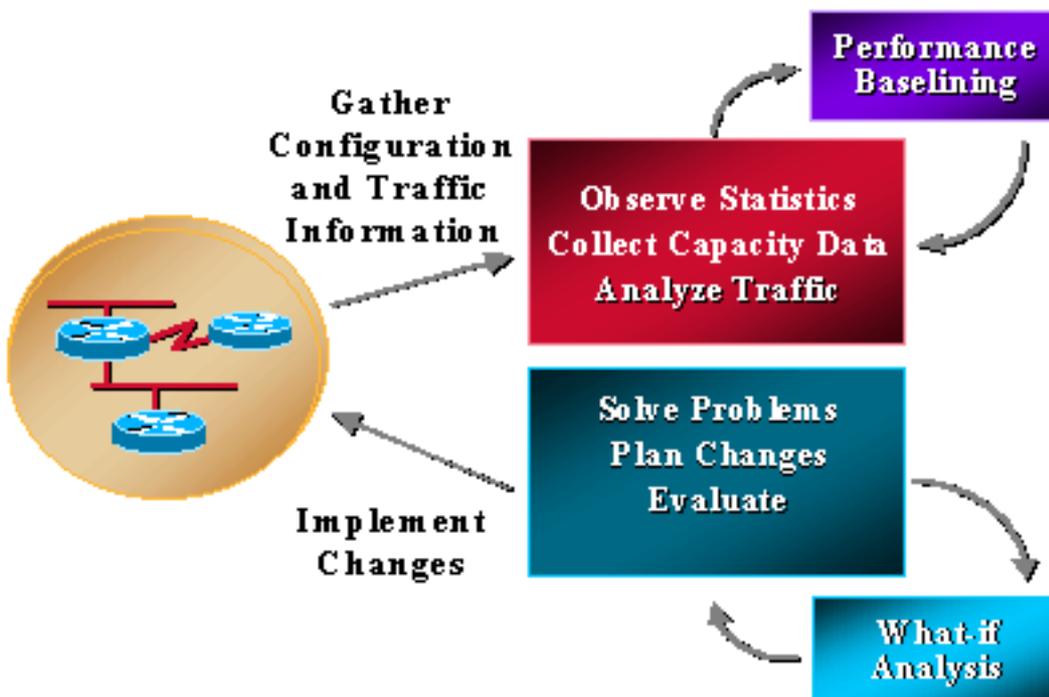
- Comment élaborer une stratégie de planification des capacités, y compris les techniques, les outils, les variables de base de données MIB et les seuils communs utilisés dans la planification des capacités.

Présentation de la gestion des capacités et des performances

La planification de la capacité est le processus qui consiste à déterminer les ressources réseau nécessaires pour empêcher un impact sur les performances ou la disponibilité des applications critiques de l'entreprise. La gestion des performances consiste à gérer le temps de réponse, la cohérence et la qualité du service réseau pour les services individuels et globaux.

Remarque : les problèmes de performances sont généralement liés à la capacité. Les applications sont plus lentes, car la bande passante et les données doivent attendre dans les files d'attente avant d'être transmises sur le réseau. Dans les applications vocales, des problèmes tels que le délai et la gigue affectent directement la qualité de l'appel vocal.

La plupart des entreprises collectent déjà des informations sur les capacités et travaillent de manière cohérente pour résoudre les problèmes, planifier les modifications et mettre en oeuvre de nouvelles fonctionnalités de capacité et de performances. Cependant, les entreprises n'effectuent pas systématiquement d'analyses de tendances et de scénarios. L'analyse de scénarios est le processus qui permet de déterminer l'impact d'une modification du réseau. Les tendances sont le processus consistant à établir des références de répartition des problèmes de capacité et de performances du réseau et à examiner les références des tendances du réseau afin de comprendre les besoins futurs de mise à niveau. La gestion de la capacité et des performances doit également inclure la gestion des exceptions lorsque des problèmes sont identifiés et résolus avant que les utilisateurs n'appellent, et la gestion de la qualité de service lorsque les administrateurs réseau planifient, gèrent et identifient des problèmes de performances de service individuels. Le schéma suivant illustre les processus de gestion des capacités et des performances.



La gestion de la capacité et des performances a également ses limites, généralement liées au

processeur et à la mémoire. Voici les sujets de préoccupation potentiels :

- CPU
- Fond de panier ou E/S
- Mémoire et tampons
- Taille des interfaces et des tuyaux
- Mise en file d'attente, latence et gigue
- Vitesse et distance
- Caractéristiques des applications

Certaines références à la planification de la capacité et à la gestion du rendement mentionnent aussi ce qu'on appelle le « plan de données » et le « plan de contrôle ». Le plan de données est simplement constitué de problèmes de capacité et de performances liés aux données qui traversent le réseau, tandis que le plan de contrôle implique des ressources requises pour maintenir le bon fonctionnement du plan de données. Les fonctionnalités du plan de contrôle incluent la surcharge de service, comme le routage, le Spanning Tree, les messages de maintien de la connexion des interfaces et la gestion SNMP du périphérique. Ces exigences de plan de contrôle utilisent le processeur, la mémoire, la mise en mémoire tampon, la mise en file d'attente et la bande passante, tout comme le trafic qui traverse le réseau. La plupart des exigences du plan de contrôle sont également essentielles au fonctionnement global du système. S'ils ne disposent pas des ressources dont ils ont besoin, le réseau tombe en panne.

CPU

Le processeur est généralement utilisé par le plan de contrôle et le plan de données sur n'importe quel périphérique réseau. Pour la gestion de la capacité et des performances, vous devez vous assurer que le périphérique et le réseau disposent d'un processeur suffisant pour fonctionner à tout moment. Un processeur insuffisant peut souvent réduire un réseau, car des ressources insuffisantes sur un périphérique peuvent avoir un impact sur l'ensemble du réseau. Un CPU insuffisant peut également augmenter la latence, car les données doivent attendre d'être traitées lorsqu'il n'y a pas de commutation matérielle sans le CPU principal.

Fond de panier ou E/S

Le fond de panier ou E/S fait référence à la quantité totale de trafic qu'un périphérique peut gérer, généralement décrite en termes de taille de bus ou de capacité de fond de panier. Un fond de panier insuffisant entraîne normalement l'abandon de paquets, ce qui peut entraîner des retransmissions et un trafic supplémentaire.

Mémoire

La mémoire est une autre ressource qui a des exigences de plan de données et de plan de contrôle. La mémoire est requise pour les informations telles que les tables de routage, les tables ARP et d'autres structures de données. Lorsque les périphériques manquent de mémoire, certaines opérations sur le périphérique peuvent échouer. L'opération peut affecter les processus du plan de contrôle ou du plan de données, selon la situation. Si les processus du plan de contrôle échouent, l'ensemble du réseau peut se dégrader. Par exemple, cela peut se produire lorsque de la mémoire supplémentaire est requise pour la convergence de routage.

Taille des interfaces et des tuyaux

Les tailles d'interface et de canal correspondent à la quantité de données pouvant être envoyées simultanément sur une connexion. On parle souvent à tort de vitesse de connexion, mais les données ne se déplacent pas à des vitesses différentes d'un périphérique à un autre. La vitesse du silicium et les capacités matérielles permettent de déterminer la bande passante disponible en fonction du support. En outre, les mécanismes logiciels peuvent « réguler » les données pour se conformer aux allocations de bande passante spécifiques pour un service. En général, cela se retrouve dans les réseaux des fournisseurs de services pour les réseaux Frame Relay ou ATM qui ont des capacités de débit de 1,54 kbit/s à 155 mbit/s et plus. Lorsque la bande passante est limitée, les données sont mises en file d'attente dans une file d'attente de transmission. Une file d'attente de transmission peut avoir différents mécanismes logiciels pour hiérarchiser les données dans la file d'attente ; toutefois, lorsque des données se trouvent dans la file d'attente, il doit attendre les données existantes avant de pouvoir les transférer hors de l'interface.

Mise en file d'attente, latence et gigue

La mise en file d'attente, la latence et la gigue affectent également les performances. Vous pouvez régler la file d'attente de transmission pour affecter les performances de différentes manières. Par exemple, si la file d'attente est volumineuse, les données attendent plus longtemps. Lorsque les files d'attente sont petites, les données sont abandonnées. Cela s'appelle le « taldrop » et est acceptable pour les applications TCP, car les données seront retransmises. Cependant, la voix et la vidéo ne fonctionnent pas bien avec la suppression de file d'attente ou même une latence de file d'attente significative nécessitant une attention particulière à la bande passante ou aux tailles de canalisation. Un délai de file d'attente peut également se produire avec les files d'attente d'entrée si le périphérique ne dispose pas de ressources suffisantes pour transférer immédiatement le paquet. Cela peut être dû au processeur, à la mémoire ou aux tampons.

La latence décrit le temps de traitement normal entre le moment où elle est reçue et celui où le paquet est transféré. Les commutateurs et routeurs de données modernes normaux présentent une latence extrêmement faible (< 1 ms) dans des conditions normales sans contraintes de ressources. Les périphériques modernes équipés de processeurs de signal numérique pour convertir et compresser les paquets vocaux analogiques peuvent prendre plus de temps, même jusqu'à 20 ms.

La gigue décrit l'écart entre les paquets pour les applications de diffusion en continu, y compris la voix et la vidéo. Si les paquets arrivent à des moments différents avec des intervalles de temps entre paquets différents, la gigue est élevée et la qualité vocale se dégrade. La gigue est principalement un facteur de délai de mise en file d'attente.

Vitesse et distance

La vitesse et la distance sont également un facteur de performance du réseau. Les réseaux de données ont une vitesse de transfert de données cohérente basée sur la vitesse de la lumière. C'est environ 100 miles par milliseconde. Si une entreprise exécute une application client-serveur au niveau international, elle peut s'attendre à un délai de transfert de paquets correspondant. La vitesse et la distance peuvent être un facteur important dans les performances des applications lorsque celles-ci ne sont pas optimisées pour les performances du réseau.

Caractéristiques des applications

Les caractéristiques des applications sont le dernier domaine qui affecte la capacité et les performances. Des problèmes tels que la petite taille des fenêtres, les keepalives d'application et

la quantité de données envoyées sur le réseau par rapport à ce qui est requis peuvent affecter les performances d'une application dans de nombreux environnements, en particulier les WAN.

Meilleures pratiques de gestion des capacités et des performances

Cette section présente en détail les cinq principales meilleures pratiques de gestion des capacités et des performances :

- [Gestion de niveau de service](#)
- [Analyse de simulation du réseau et des applications](#)
- [Planification et tendances](#)
- [Gestion des exceptions](#)
- [Gestion QoS](#)

Gestion de niveau de service

La gestion des niveaux de service définit et régule les autres processus de gestion des performances et de la capacité requis. Les administrateurs réseau comprennent qu'ils ont besoin d'une planification des capacités, mais ils sont confrontés à des contraintes budgétaires et de personnel qui empêchent une solution complète. La gestion des niveaux de service est une méthodologie éprouvée qui aide à résoudre les problèmes de ressources en définissant un livrable et en créant une responsabilité bidirectionnelle pour un service lié à ce livrable. Pour ce faire, vous pouvez procéder de deux manières :

- Créer un accord de niveau de service entre les utilisateurs et l'organisation réseau pour un service qui inclut la gestion de la capacité et des performances. Le service comprendrait des rapports et des recommandations pour maintenir la qualité du service. Cependant, les utilisateurs doivent être prêts à financer le service et toutes les mises à niveau requises.
- L'organisation réseau définit son service de gestion de la capacité et des performances, puis tente de le financer et procède à des mises à niveau au cas par cas.

Quoi qu'il en soit, l'organisation réseau doit commencer par définir un service de planification de la capacité et de gestion des performances qui inclut les aspects du service qu'elle peut actuellement fournir et ceux qui sont prévus à l'avenir. Un service complet inclurait une analyse de simulation pour les modifications du réseau et des applications, une planification initiale et des tendances pour les variables de performances définies, la gestion des exceptions pour les variables de capacité et de performances définies et la gestion de la qualité de service.

Analyse de simulation du réseau et des applications

Effectuer une analyse de simulation du réseau et des applications pour déterminer le résultat d'une modification planifiée. En l'absence d'analyse de scénarios, les entreprises prennent des risques importants pour modifier la réussite et la disponibilité globale du réseau. Dans de nombreux cas, les modifications du réseau ont entraîné un effondrement congestif, entraînant de nombreuses heures d'interruption de la production. En outre, un nombre impressionnant d'introductions d'applications échouent et ont un impact sur les autres utilisateurs et applications. Ces défaillances se poursuivent dans de nombreuses entreprises de réseau, mais elles peuvent être totalement évitées grâce à quelques outils et à quelques étapes de planification supplémentaires.

Vous avez normalement besoin de quelques nouveaux processus pour effectuer une analyse de simulation de qualité. La première étape consiste à identifier les niveaux de risque pour tous les changements et à exiger une analyse par simulation plus approfondie pour les changements à risque élevé. Le niveau de risque peut être un champ obligatoire pour toutes les soumissions de modifications. Des changements de niveau de risque plus élevés exigeraient alors une analyse de simulation définie du changement. Une analyse de simulation de réseau détermine l'impact des modifications du réseau sur l'utilisation du réseau et les problèmes de ressources du plan de contrôle du réseau. Une analyse de simulation d'application déterminerait la réussite de l'application du projet, les besoins en bande passante et les problèmes de ressources réseau. Les tableaux suivants sont des exemples d'affectation des niveaux de risque et des exigences de test correspondantes :

Niveau de risque	Définition	Recommandations de planification du changement
1	<ul style="list-style-type: none"> • Impact potentiel élevé sur un grand nombre d'utilisateurs (plus de 500) ou de services critiques en raison de l'introduction de nouveaux produits, logiciels, topologies ou fonctionnalités. • Le changement implique un temps d'arrêt du réseau prévu. 	<ul style="list-style-type: none"> • Validation du laboratoire de la nouvelle solution. La validation en laboratoire inclut des tests et une validation documentés de la solution, ainsi qu'une analyse de scénarios montrant l'impact sur l'infrastructure existante. Nous recommandons des pilotes de solutions. Les nouvelles solutions nécessitent la rédaction d'un document d'assistance opérationnelle. • Réaliser une

		<p>étude de conception Cisco NSA.</p> <ul style="list-style-type: none"> • Créez un plan de repli. • Créer un plan de mise en oeuvre. • Créer un processus de modification.
2	<ul style="list-style-type: none"> • Impact potentiel élevé sur un grand nombre d'utilisateurs (plus de 500) ou de services critiques en raison d'une augmentation importante du trafic ou des utilisateurs, de modifications du réseau fédérateur ou de modifications du routage. • La modification peut nécessiter un certain temps d'arrêt. 	<ul style="list-style-type: none"> • Effectuer une analyse de simulation pour déterminer l'impact sur l'environnement existant (doit être effectuée dans un environnement de laboratoire). • Tester et vérifier les modifications de routage pour en vérifier la fonctionnalité. • Créez un plan de repli. • Révision de la conception pour les modifications majeures du routage ou du backbone. • Créer un plan de mise en oeuvre. • Créer un processus de modification.
3	<ul style="list-style-type: none"> • Impact potentiel moyen sur un plus petit nombre d'utilisateurs ou de services 	<ul style="list-style-type: none"> • Effectuer une analyse technique de la

	<p>professionnels en raison de toute modification non standard.</p> <ul style="list-style-type: none"> • Inclut un nouveau produit, un nouveau logiciel, une nouvelle topologie, l'ajout de fonctionnalités pour de nouveaux utilisateurs, une augmentation du trafic ou une topologie non standard. • La modification peut nécessiter un certain temps d'arrêt. 	<p>nouvelle solution (peut nécessiter une validation en laboratoire).</p> <ul style="list-style-type: none"> • Créer un plan de mise en oeuvre. • Créer un processus de modification.
4	<ul style="list-style-type: none"> • Réduction de l'impact potentiel sur les services ou les utilisateurs. • Inclut l'ajout de nouveaux modèles de modules réseau standard, tels que des commutateurs/concentrateurs de serveur ou de construction sur des routeurs. • Inclut l'ajout de nouveaux sites WAN ou de services d'accès supplémentaires éprouvés. • Tous les changements de niveau de risque 3 ont été prouvés techniquement dans l'environnement de production. • La modification peut nécessiter un certain temps d'arrêt. 	<ul style="list-style-type: none"> • Créer un plan de mise en oeuvre. • Créer un processus de modification.
5	<ul style="list-style-type: none"> • Aucun impact sur les utilisateurs ou les services. • Inclut l'ajout d'utilisateurs individuels au réseau et les modifications de configuration standard telles que le mot de passe, la bannière, SNMP ou d'autres paramètres de configuration standard. 	<ul style="list-style-type: none"> • Processus de modification facultatif.

- | | | |
|--|----------------------|--|
| | • Pas de temps mort. | |
|--|----------------------|--|

Une fois que vous avez défini l'emplacement où vous avez besoin de l'analyse par simulation, vous pouvez définir le service.

Vous pouvez effectuer une simulation de réseau à l'aide d'outils de modélisation ou d'un laboratoire qui imite l'environnement de production. Les outils de modélisation sont limités par la capacité de l'application à comprendre les problèmes de ressources des périphériques et, comme la plupart des modifications apportées au réseau concernent de nouveaux périphériques, l'application peut ne pas comprendre l'effet de la modification. La meilleure méthode consiste à créer une représentation du réseau de production dans un laboratoire et à tester le logiciel, la fonctionnalité, le matériel ou la configuration souhaités sous charge à l'aide de générateurs de trafic. La fuite de routes (ou d'autres informations de contrôle) du réseau de production dans le laboratoire améliore également l'environnement du laboratoire. Testez les besoins en ressources supplémentaires avec différents types de trafic, y compris le trafic SNMP, de diffusion, de multidiffusion, chiffré ou compressé. Grâce à ces différentes méthodologies, analysez les besoins en ressources des périphériques dans des situations de stress potentiel telles que la convergence de routage, le battement de liaison et le redémarrage des périphériques. Les problèmes d'utilisation des ressources incluent des domaines de ressources de capacité normale tels que le processeur, la mémoire, l'utilisation du fond de panier, les tampons et la mise en file d'attente.

Les nouvelles applications doivent également effectuer une analyse de simulation pour déterminer la réussite des applications et les besoins en bande passante. Vous effectuez généralement cette analyse dans un environnement de travaux pratiques à l'aide d'un analyseur de protocole et d'un simulateur de délai WAN pour comprendre l'effet de la distance. Vous n'avez besoin que d'un PC, d'un concentrateur, d'un périphérique de retard WAN et d'un routeur de TP connectés au réseau de production. Vous pouvez simuler la bande passante dans les travaux pratiques en limitant le trafic à l'aide d'un formatage de trafic générique ou d'une limitation de débit sur le routeur test. L'administrateur réseau peut travailler en collaboration avec le groupe d'applications pour comprendre les besoins en bande passante, les problèmes de fenêtrage et les problèmes de performances potentiels de l'application dans les environnements LAN et WAN.

Effectuer une analyse de simulation d'application avant de déployer une application métier. Si vous ne le faites pas, le groupe d'applications blâme le réseau pour ses performances médiocres. Si vous avez besoin d'une analyse de simulation d'application pour les nouveaux déploiements via le processus de gestion des modifications, vous pouvez empêcher les déploiements infructueux et mieux comprendre les augmentations soudaines de la consommation de bande passante pour les besoins client-serveur et par lots.

Planification et tendances

La planification initiale et les tendances permettent aux administrateurs réseau de planifier et d'effectuer des mises à niveau du réseau avant qu'un problème de capacité ne provoque des pannes ou des problèmes de performances. Comparer l'utilisation des ressources au cours de périodes successives ou extraire les informations dans le temps dans une base de données et permettre aux planificateurs d'afficher les paramètres d'utilisation des ressources pour la dernière heure, le dernier jour, la dernière semaine, le dernier mois et la dernière année. Dans un cas comme dans l'autre, une personne doit examiner les renseignements une fois par semaine, deux fois par semaine ou tous les mois. Le problème avec la planification initiale et les tendances est qu'elles nécessitent une quantité considérable d'informations à examiner dans les grands réseaux.

Vous pouvez résoudre ce problème de plusieurs façons :

- Développez une capacité importante et passez à l'environnement LAN pour que la capacité ne pose pas de problème.
- Divisez les informations sur les tendances en groupes et concentrez-vous sur la haute disponibilité ou les zones critiques du réseau, telles que les sites WAN critiques ou les LAN de data center.
- Les mécanismes de signalement peuvent mettre en évidence les domaines qui dépassent un certain seuil d'attention particulière. Si vous commencez par implémenter des zones de disponibilité critiques, vous pouvez réduire considérablement la quantité d'informations requises pour la révision.

Avec toutes les méthodes précédentes, vous devez toujours revoir les informations sur une base périodique. La planification initiale et les tendances constituent un effort proactif et si l'entreprise ne dispose que de ressources pour une assistance réactive, les personnes ne liront pas les rapports.

De nombreuses solutions de gestion de réseau fournissent des informations et des graphiques sur les variables de ressources de capacité. Malheureusement, la plupart des gens n'utilisent ces outils que pour une assistance réactive à un problème existant ; cela va à l'encontre des objectifs de planification initiale et de tendance. Les produits Concord Network Health et INS EnterprisePRO sont deux outils efficaces pour fournir des informations sur les tendances de capacité pour les réseaux Cisco. Dans de nombreux cas, les organisations réseau utilisent des langages de script simples pour collecter les informations de capacité. Vous trouverez ci-dessous des exemples de rapports collectés via Script pour l'utilisation de la liaison, l'utilisation du CPU et les performances des requêtes ping. D'autres variables de ressources qui peuvent être importantes pour la tendance incluent la mémoire, la profondeur de file d'attente, le volume de diffusion, la mémoire tampon, la notification d'encombrement Frame Relay et l'utilisation du fond de panier. Reportez-vous à ces tableaux pour obtenir des informations sur l'utilisation des liaisons et du CPU :

Utilisation des liaisons

Ressource	Adresse	Segment	Utilisation moyenne (%)	Utilisation maximale (%)
JTKR01S2	10.2.6.1	128 Kbps	66.3	97.6
JYKR01S0	10.2.6.2	128 Kbps	66.3	97.8
FMCR18S4/4	10.2.5.1	384 Kbps	51.3	109.7
PACR01S3/1	10.2.5.2	384 Kbps	51.1	98.4

Utilisation du processeur

Ressource	Adresse d'interrogation	Utilisation moyenne (%)	Utilisation maximale (%)
FSTR01	10.28.142.1	60.4	80
NERT06	10.170.2.1	47	86
NORR01	10.73.200.1	47	99

RTCR01	10.49.136.1	42	98
--------	-------------	----	----

Utilisation des liaisons

Ressource	Adresse	AvResT (mS) 09-09-98	AvResT (mS) 09-09-98	AvResT (mS) 09-09-98	AvResT (mS) 10-01-98
ADR01	10.190.56.1	469.1	852.4	461.1	873.2
ABNR01	10.190.52.1	486.1	869.2	489.5	880.2
AVR01	10.190.54.1	490.7	883.4	485.2	892.5
ASAR01	10.196.170.1	619.6	912.3	613.5	902.2
ASRR01	10.196.178.1	667.7	976.4	655.5	948.6
ASYR01S					503.4
AZWRT01	10.177.32.1	460.1		444.7	
BEJR01	10.195.18.1	1023.7	1064.6	1184	1021.9

Gestion des exceptions

La gestion des exceptions est une méthodologie précieuse pour identifier et résoudre les problèmes de capacité et de performances. L'idée est de recevoir une notification des violations des seuils de capacité et de performances afin d'enquêter immédiatement et de résoudre le problème. Par exemple, un administrateur réseau peut recevoir une alarme pour un processeur élevé sur un routeur. L'administrateur réseau peut se connecter au routeur pour déterminer pourquoi le processeur est si élevé. Elle peut alors effectuer une configuration corrective qui réduit le CPU ou créer une liste d'accès empêchant le trafic qui cause le problème, surtout si le trafic ne semble pas être critique pour l'entreprise.

Vous pouvez configurer la gestion des exceptions pour des problèmes plus critiques simplement en utilisant les commandes de configuration RMON sur un routeur ou en utilisant des outils plus avancés tels que Netsys service level manager en conjonction avec les données SNMP, RMON ou Netflow. La plupart des outils de gestion du réseau permettent de définir des seuils et des alarmes en cas de violation. L'aspect important du processus de gestion des exceptions est de fournir une notification en temps quasi réel du problème. Dans le cas contraire, le problème peut disparaître avant que quiconque ne remarque que la notification a été reçue. Cela peut se faire au sein d'un centre d'exploitation du réseau si l'organisation a une surveillance constante. Dans le cas contraire, nous vous recommandons de téléavertir.

L'exemple de configuration suivant fournit une notification de seuil croissant et décroissant pour le processeur du routeur vers un fichier journal qui peut être révisé de manière cohérente. Vous pouvez configurer des commandes RMON similaires pour les violations de seuil d'utilisation de

liaison critique ou d'autres seuils SNMP.

```
rmon event 1 trap CPUtrap description
"CPU Util >75%"rmon event 2 trap CPUtrap description
"CPU Util <75%"rmon event 3 trap CPUtrap description
"CPU Util >90%"rmon event 4 trap CPUtrap description
"CPU Util <90%"rmon alarm 75 lsystem.56.0 10 absolute rising-threshold
75 1 falling-threshold 75 2rmon alarm 90 lsystem.56.0 10 absolute rising-threshold
90 3 falling-threshold 90 4
```

Gestion QoS

La gestion de la qualité de service implique la création et la surveillance de classes de trafic spécifiques au sein du réseau. Un trafic offre des performances plus homogènes pour des groupes d'applications spécifiques (définis dans des classes de trafic). Les paramètres de mise en forme du trafic offrent une grande souplesse dans la hiérarchisation et la mise en forme du trafic pour des classes spécifiques de trafic. Ces fonctionnalités incluent des fonctionnalités telles que le débit d'accès garanti (CAR), la détection précoce aléatoire pondérée (WRED) et la mise en file d'attente pondérée équitable basée sur les classes. Les classes de trafic sont généralement créées en fonction des SLA de performances pour les applications stratégiques et les exigences spécifiques des applications, telles que la voix. Le trafic non critique ou non commercial serait également contrôlé de manière à ne pas affecter les applications et les services prioritaires.

La création de classes de trafic nécessite une compréhension de base de l'utilisation du réseau, des exigences spécifiques des applications et des priorités des applications métier. Les exigences des applications incluent la connaissance de la taille des paquets, des problèmes de délai d'attente, des exigences de gigue, des exigences de rafale, des exigences de lot et des problèmes de performances globaux. Grâce à ces connaissances, les administrateurs réseau peuvent créer des plans et des configurations de formatage du trafic qui offrent des performances applicatives plus homogènes sur diverses topologies LAN/WAN.

Par exemple, une entreprise dispose d'une connexion ATM de 10 mégabits entre deux sites principaux. La liaison est parfois encombrée par des transferts de fichiers volumineux, ce qui entraîne une dégradation des performances pour le traitement des transactions en ligne et une qualité vocale médiocre ou inutilisable.

L'organisation a configuré quatre classes de trafic différentes. La voix a reçu la priorité la plus élevée et a pu conserver cette priorité même si elle dépassait le débit de trafic estimé. La classe d'application critique a reçu la priorité la plus élevée suivante, mais elle n'a pas été autorisée à dépasser la taille totale de liaison moins les besoins estimés en bande passante vocale. Lorsqu'elle éclate, elle est abandonnée. Le trafic de transfert de fichiers a simplement reçu une priorité plus faible et tous les autres trafics se situent quelque part au milieu.

L'entreprise doit maintenant effectuer une gestion QoS sur cette liaison pour déterminer la quantité de trafic prise par chaque classe et mesurer les performances au sein de chaque classe. Si l'organisation ne le fait pas, il peut arriver que certaines classes souffrent de la famine ou que les SLA de performance ne soient pas respectés dans une classe particulière.

La gestion des configurations QOS reste une tâche difficile en raison du manque d'outils. Une méthode consiste à utiliser le gestionnaire de performances Internet (IPM) de Cisco pour envoyer un trafic différent sur la liaison qui appartient à chacune des classes de trafic. Vous pouvez alors surveiller les performances de chaque classe et IPM fournit des analyses de tendances, en temps réel et saut par saut pour identifier les zones problématiques. D'autres peuvent encore utiliser une

méthode plus manuelle, comme l'analyse de la mise en file d'attente et des paquets abandonnés dans chaque classe de trafic en fonction des statistiques d'interface. Dans certaines entreprises, ces données peuvent être collectées via SNMP ou analysées dans une base de données pour les lignes de base et les tendances. Il existe également sur le marché des outils qui envoient des types de trafic spécifiques sur le réseau pour déterminer les performances d'un service ou d'une application spécifique.

Collecte et reporting des informations de capacité

La collecte et la communication d'informations sur les capacités devraient être liées aux trois domaines recommandés de la gestion des capacités :

- Analyse de simulation, axée sur les modifications apportées au réseau et leur impact sur l'environnement
- Planification et tendances
- Gestion des exceptions

Dans chacun de ces domaines, élaborer un plan de collecte d'information. Dans le cas d'analyses de simulation de réseau ou d'application, vous avez besoin d'outils pour imiter l'environnement réseau et comprendre l'impact de la modification par rapport à des problèmes potentiels de ressources dans le plan de contrôle du périphérique ou le plan de données. Dans le cas de la planification initiale et des tendances, vous avez besoin d'instantanés pour les périphériques et les liaisons montrant l'utilisation actuelle des ressources. Vous examinerez ensuite les données au fil du temps pour comprendre les exigences potentielles de mise à niveau. Cela permet aux administrateurs réseau de planifier correctement les mises à niveau avant que des problèmes de capacité ou de performances ne surviennent. En cas de problème, vous devez gérer les exceptions pour alerter les administrateurs réseau afin qu'ils puissent régler le réseau ou résoudre le problème.

Ce processus peut être divisé en plusieurs étapes :

1. Déterminez vos besoins.
2. Définissez un processus.
3. Définir les zones de capacité.
4. Définissez les variables de capacité.
5. Interpréter les données.

Déterminer vos besoins

L'élaboration d'un plan de gestion de la capacité et du rendement exige de comprendre l'information dont vous avez besoin et le but de cette information. Divisez le plan en trois zones obligatoires : une pour l'analyse de scénarios, la planification initiale/les tendances et la gestion des exceptions. Dans chacun de ces domaines, découvrez quelles ressources et quels outils sont disponibles et quels sont les besoins. De nombreuses entreprises échouent dans le déploiement d'outils parce qu'elles tiennent compte de la technologie et des fonctionnalités des outils, mais pas des personnes et de l'expertise requises pour gérer les outils. Incluez les personnes et l'expertise requises dans votre plan, ainsi que les améliorations apportées aux processus. Ces personnes peuvent être des administrateurs système chargés de gérer les stations de gestion du réseau, des administrateurs de bases de données chargés de l'administration des bases de données, des administrateurs formés à l'utilisation et à la surveillance des outils et des administrateurs réseau de niveau supérieur chargés de déterminer les politiques, les seuils et les exigences en matière

de collecte d'informations.

Définir un processus

Vous avez également besoin d'un processus pour vous assurer que l'outil est utilisé avec succès et de manière cohérente. Vous devrez peut-être améliorer les processus pour définir ce que les administrateurs réseau doivent faire en cas de franchissement de seuil ou le processus à suivre pour la planification initiale, les tendances et la mise à niveau du réseau. Une fois que vous avez déterminé les besoins et les ressources nécessaires à une planification de capacité réussie, vous pouvez envisager la méthodologie. De nombreuses entreprises choisissent d'externaliser ce type de fonctionnalité à une organisation de services réseau telle qu'INS ou de développer l'expertise en interne parce qu'elles considèrent le service comme une compétence de base.

Définir les zones de capacité

Le plan de planification des capacités devrait également inclure une définition des zones de capacité. Voici les zones du réseau qui peuvent partager une stratégie commune de planification des capacités : par exemple, le réseau local de l'entreprise, les bureaux extérieurs de réseau étendu, les sites WAN critiques et l'accès commuté. La définition de différentes zones est utile pour plusieurs raisons :

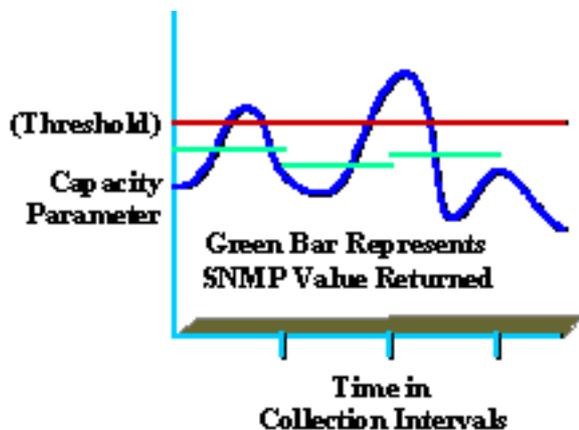
- Différentes zones peuvent avoir des seuils différents. Par exemple, la bande passante LAN est beaucoup moins chère que la bande passante WAN. Les seuils d'utilisation doivent donc être inférieurs.
- Différentes zones peuvent nécessiter la surveillance de différentes variables MIB. Par exemple, les compteurs FECN et BECN dans Frame Relay sont essentiels pour comprendre les problèmes de capacité de relais de trames.
- La mise à niveau de certaines zones du réseau peut être plus difficile ou plus longue. Par exemple, les circuits internationaux peuvent avoir des délais d'exécution beaucoup plus longs et nécessiter un niveau de planification plus élevé.

Définir les variables de capacité

L'étape suivante consiste à définir les variables à surveiller et les valeurs de seuil nécessitant une action. La définition des variables de capacité dépend de manière significative des périphériques et des supports utilisés sur le réseau. En général, les paramètres tels que le processeur, la mémoire et l'utilisation des liaisons sont importants. Cependant, d'autres domaines peuvent être importants pour des technologies ou des exigences spécifiques. Il peut s'agir de la profondeur des files d'attente, des performances, de la notification d'encombrement Frame Relay, de l'utilisation du fond de panier, de la mémoire tampon, des statistiques Netflow, du volume de diffusion et des données RMON. Gardez à l'esprit vos plans à long terme, mais commencez par quelques domaines clés pour assurer la réussite.

Interprétation des données

La compréhension des données collectées est également essentielle pour fournir un service de haute qualité. Par exemple, de nombreuses entreprises ne comprennent pas parfaitement les niveaux d'utilisation moyens et de pointe. Le schéma suivant illustre un pic de paramètre de capacité basé sur un intervalle de collecte SNMP de 5 minutes (en vert).



Même si la valeur signalée était inférieure au seuil (indiqué en rouge), des pics supérieurs à la valeur seuil (indiquée en bleu) peuvent encore se produire dans l'intervalle de collecte. Cela est important car pendant l'intervalle de collecte, l'entreprise peut connaître des valeurs maximales qui affectent les performances ou la capacité du réseau. Veillez à sélectionner un intervalle de collecte significatif qui soit utile et qui n'entraîne pas de surcharge excessive.

Un autre exemple est l'utilisation moyenne. Si les employés sont au bureau de huit à cinq personnes seulement, mais que l'utilisation moyenne est de 7 heures sur 24, l'information peut être trompeuse.

Informations connexes

- [Support technique - Cisco Systems](#)

À propos de cette traduction

Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.