Gestion des performances : Livre blanc sur les pratiques recommandées

Table des matières

[Introduction](#)
[Informations générales](#)
[Facteurs critiques de succès](#)
[Indicateurs de gestion des performances](#)
[Flux du processus de gestion des performances](#)
[Développer un concept d'exploitation de gestion de réseau](#)
[Mesurer les performances](#)
[Effectuer une analyse proactive des pannes](#)
[Indicateurs de gestion des performances](#)
[Documentation des objectifs commerciaux de la gestion du réseau](#)
[Documenter les contrats de niveau de service](#)
[Créer une liste de variables pour la ligne de base](#)
[Examiner les analyses de référence et de tendances](#)
[Documentation d'une méthodologie d'analyse de scénarios](#)
[Documentation de la méthodologie utilisée pour améliorer les performances du réseau](#)
[Résumé](#)
[Informations connexes](#)

Introduction

La gestion des performances implique l'optimisation du temps de réponse du service réseau et la gestion de la cohérence et de la qualité de services réseau individuels et globaux. Le service le plus important est celui de mesurer le temps de réponse des utilisateurs et des applications. Pour la plupart des utilisateurs, le temps de réponse est l'indicateur de succès essentiel en ce qui a trait au rendement. Cette variable influence la perception du réseau par vos utilisateurs et les administrateurs des applications.

Informations générales

La planification de la capacité est le processus par lequel vous déterminez les besoins en ressources réseau futures afin d'éviter un impact sur les performances ou la disponibilité des applications critiques de l'entreprise. Dans le domaine de la planification de la capacité, la ligne de base du réseau (processeur, mémoire, tampons, octets entrants/sortants, etc.) peut affecter le temps de réponse. Par conséquent, gardez à l'esprit que les problèmes de performances sont souvent corrélés avec la capacité. Dans les réseaux, il s’agit généralement de la bande passante et des données qui doivent attendre dans les files d’attente avant de pouvoir être transmises sur le réseau. Dans les applications vocales, ce temps d'attente affecte presque certainement les

utilisateurs, car des facteurs tels que le délai et la gigue affectent la qualité de l'appel vocal.

Un autre problème majeur qui complique la gestion des performances est que, bien que la disponibilité élevée du réseau soit essentielle pour les réseaux des grandes entreprises et des fournisseurs de services, la tendance est de rechercher des gains économiques à court terme au risque de coûts plus élevés (souvent imprévus) à long terme. À chaque cycle budgétaire, les administrateurs réseau et le personnel chargé de la mise en oeuvre des projets peinent à trouver un équilibre entre performances et rapidité de mise en oeuvre. En outre, les administrateurs réseau sont confrontés à des défis tels que le développement rapide de produits afin de répondre à des créneaux de marché étroits, des technologies complexes, la consolidation des activités, des marchés concurrents, des temps d'arrêt imprévus, un manque d'expertise et souvent des outils insuffisants.

À la lumière de ces défis, comment les performances s'intègrent-elles dans le cadre de gestion du réseau ? La fonction principale d'un système de gestion de réseau idéal est d'optimiser les capacités opérationnelles d'un réseau. Une fois que vous acceptez ce point comme l'objectif ultime de la gestion du réseau, l'objectif principal de la gestion du réseau est de maintenir le fonctionnement du réseau à des performances optimales.

Un système de gestion de réseau idéal comprend les opérations principales suivantes :

- Informe l'opérateur de la détérioration imminente des performances.
- Fournit un routage alternatif et des solutions de contournement en cas de détérioration ou de défaillance des performances.
- Fournit les outils permettant d'identifier les causes de détérioration ou de défaillance des performances.
- Sert de station principale pour la résilience et la survie du réseau.
- Communication des performances en temps réel.

Selon cette définition de système idéal, la gestion des performances devient essentielle à la gestion du réseau. Ces problèmes de gestion des performances sont critiques :

- Performances utilisateur
- Performances des applications
- Planification de capacité
- Gestion proactive des pannes

Il est important de noter qu'avec les applications plus récentes comme la voix et la vidéo, les performances sont la variable clé du succès et si vous ne parvenez pas à obtenir des performances cohérentes, le service est considéré comme de faible valeur et échoue. Dans d'autres cas, les utilisateurs souffrent simplement de performances variables avec des délais d'attente intermittents des applications qui nuisent à la productivité et à la satisfaction des

utilisateurs.

Ce document détaille les problèmes de gestion des performances les plus critiques, notamment les facteurs de succès critiques, les indicateurs de performances clés et un schéma de processus général pour la gestion des performances. Il aborde également les concepts de disponibilité, de temps de réponse, de précision, d'utilisation et de planification des capacités, et inclut une brève discussion sur le rôle de l'analyse proactive des pannes dans la gestion des performances et le système de gestion de réseau idéal.

Facteurs critiques de succès

Les facteurs critiques de succès identifient les exigences des meilleures pratiques de mise en oeuvre. Pour être considéré comme un facteur de réussite critique, un processus ou une procédure doit améliorer la disponibilité ou l'absence de la procédure doit diminuer la disponibilité. En outre, le facteur de réussite critique doit être mesurable afin que l'organisation puisse déterminer l'étendue de sa réussite.

Note : Voir Indicateurs de gestion des performances pour plus d'informations.

Voici les principaux facteurs de succès de la gestion des performances :

- Collectez une ligne de base pour les données du réseau et des applications.
- Effectuez une analyse de simulation sur votre réseau et vos applications.
- Générer des rapports d'exception pour les problèmes de capacité.
- Déterminez la surcharge de gestion du réseau pour tous les services de gestion du réseau proposés ou potentiels.
- Analysez les informations de capacité.
- Examiner périodiquement les informations de capacité pour le réseau et les applications, ainsi que la ligne de base et les exceptions.
- Mettre en place des procédures de mise à niveau ou de réglage pour gérer les problèmes de capacité sur une base réactive et à long terme.

Indicateurs de gestion des performances

Les indicateurs de performance fournissent le mécanisme par lequel une organisation peut mesurer les facteurs de réussite critiques. Les indicateurs de rendement pour la planification du rendement comprennent :

- Documentez les objectifs commerciaux de gestion du réseau. Il peut s’agir d’un concept formel d’opérations pour la gestion du réseau ou d’un énoncé moins formel des caractéristiques et des objectifs requis.

- Créez des objectifs de niveau de service détaillés et mesurables.
- Fournir une documentation sur les contrats de niveau de service avec des tableaux ou des graphiques qui montrent le succès ou l'échec de la façon dont ces contrats sont respectés au fil du temps.
- Collectez une liste des variables de la ligne de base, telles que l'intervalle d'interrogation, la surcharge de gestion du réseau encourue, les seuils de déclenchement possibles, l'utilisation de la variable comme déclencheur d'une interruption et l'analyse des tendances utilisée pour chaque variable.
- Tenir une réunion périodique pour examiner l'analyse de la base de référence et des tendances.
- Disposer d'une méthodologie d'analyse par simulation documentée. Cela devrait inclure la modélisation et la vérification, le cas échéant.
- Lorsque les seuils sont dépassés, élaborez une documentation sur la méthodologie utilisée pour augmenter les ressources réseau. Un élément à documenter est le délai nécessaire pour ajouter de la bande passante WAN supplémentaire et une table des coûts.

Flux du processus de gestion des performances

Ces étapes fournissent un flux de processus de haut niveau pour la gestion des performances :

1. Développer un concept d'exploitation de gestion de réseau
    a. Définissez les fonctions requises : Objectifs en matière de services, d'évolutivité et de disponibilité
    b. Définition des objectifs de disponibilité et de gestion du réseau
    c. Définir les SLA et les mesures de performances
    d. Définition des SLA
2. Mesurer les performances
    a. Collecte des données de base du réseau
    b. Mesurer la disponibilité
    c. Mesurer le temps de réponse
    d. Précision des mesures
    e. Utilisation des mesures
    f. Planification de capacité

3. Effectuer une analyse proactive des pannes
    a. Utiliser des seuils pour une gestion proactive des pannes
    b. Implémentation de la gestion du réseau
    c. Mesures de fonctionnement du réseau

Développer un concept d'exploitation de gestion de réseau

Avant de définir les variables détaillées de performances et de capacité d'un réseau, vous devez examiner le concept global d'exploitation pour la gestion du réseau au sein de votre organisation. Lorsque vous définissez ce concept global, il fournit une base commerciale sur laquelle vous pouvez construire des définitions précises des fonctionnalités souhaitées dans votre réseau. Si vous ne parvenez pas à développer un concept opérationnel pour la gestion du réseau, cela peut entraîner un manque d'objectifs ou des objectifs qui changent constamment en raison des demandes des clients.

En règle générale, la première étape de la phase de définition du système du programme de gestion du réseau consiste à produire le concept d'exploitation de gestion du réseau. L'objectif est de décrire les caractéristiques globales du système d'un point de vue opérationnel. Ce document sert à coordonner les objectifs commerciaux globaux (non quantitatifs) des opérations réseau, de l'ingénierie, de la conception, des autres unités commerciales et des utilisateurs finaux. Le présent document vise à définir les activités de planification opérationnelle à long terme pour la gestion et l'exploitation du réseau. Il fournit également des directives pour l'élaboration de toute la documentation sur les définitions subséquentes, comme les accords de niveau de service. Cet ensemble initial de définitions ne peut évidemment pas se focaliser trop étroitement sur la gestion de problèmes réseau spécifiques, mais sur les éléments qui mettent l'accent sur l'importance pour l'ensemble de l'organisation et par rapport aux coûts qui doivent être gérés également. Certains objectifs sont les suivants :

- Identifiez les caractéristiques essentielles à une utilisation efficace de l'infrastructure réseau.
- Identifiez les services/applications pris en charge par le réseau.
- Lancez la gestion de bout en bout des services.
- Mettre en place des mesures basées sur les performances pour améliorer le service global.
- Collecter et distribuer les informations de gestion des performances.
- Soutenir l'évaluation stratégique du réseau avec les commentaires des utilisateurs.

En d'autres termes, le concept d'exploitation de la gestion du réseau doit se concentrer sur les objectifs généraux de l'entreprise et sur votre philosophie d'atteinte de ces objectifs. Les principaux éléments sont les définitions de niveau supérieur de la mission, les objectifs de la mission, les objectifs du système, la participation organisationnelle et la philosophie opérationnelle globale.

En tant qu'administrateur réseau, vous êtes en mesure d'unifier les attentes de performances souvent incohérentes de vos utilisateurs. Par exemple, si la principale exigence du réseau est le transfert de fichiers volumineux d'un emplacement à un autre, vous devez vous concentrer sur le haut débit et moins sur les temps de réponse des utilisateurs interactifs. Veillez à ne pas limiter votre vision des performances à moins que vous ne considériez une variété de problèmes. Par exemple, lorsque vous testez un réseau, examinez les niveaux de charge qui sont utilisés. La charge est souvent basée sur de très petits paquets et le débit sur de très grands paquets. Chacun de ces tests de performances peut donner une image très positive, mais en fonction de la charge de trafic réseau, les tests peuvent ne pas donner une image fidèle des performances. Étudiez les performances du réseau dans autant de conditions de charge de travail que possible et documentez les performances.

En outre, bien que de nombreuses organisations d'administration de réseaux disposent de techniques d'alarme efficaces pour avertir les techniciens d'une panne de périphérique, il est beaucoup plus difficile de définir et de mettre en oeuvre un processus d'évaluation des performances des applications de bout en bout. Par conséquent, bien que le centre d'exploitation du réseau (NOC) puisse répondre rapidement à un routeur ou un commutateur en panne, les conditions réseau susceptibles de nuire aux performances du réseau et d'affecter la perception de l'utilisateur peuvent facilement passer inaperçues jusqu'à ce que cette perception devienne négative. Aussi difficile soit-il, ce second processus peut apporter des avantages considérables à l'organisation de l'entreprise et à la gestion du réseau.

Enfin, veillez à ne pas créer d'attentes irréalistes en matière de performances réseau. Des attentes irréalistes sont généralement créées lorsque vous ne comprenez pas correctement les détails des protocoles réseau ou des applications. Souvent, les performances médiocres ne sont pas dues à une défaillance du réseau, mais plutôt à une conception inadéquate des applications. La seule façon de documenter et de mesurer les performances des applications consiste à disposer d'une base des performances du réseau avant l'installation des applications.

Définissez les fonctions requises : Objectifs en matière de services, d'évolutivité et de disponibilité

La première étape de la gestion des performances, de la planification continue des capacités et de la conception du réseau consiste à définir les fonctionnalités et/ou les services requis. Cette étape nécessite que vous compreniez les applications, les flux de trafic de base, le nombre d'utilisateurs et de sites et les services réseau requis. La première utilisation de ces informations consiste à déterminer le degré de criticité de l'application par rapport aux objectifs de l'organisation. Vous pouvez également appliquer ces informations pour créer une base de connaissances à utiliser dans la conception logique afin de comprendre les exigences en matière de bande passante, d'interface, de connectivité, de configuration et de périphérique physique. Cette étape initiale permet à vos architectes réseau de créer un modèle de votre réseau.

Créez des objectifs d'évolutivité de la solution afin d'aider les ingénieurs réseau à concevoir des réseaux qui répondent aux exigences de croissance future et de garantir que les conceptions proposées ne subissent pas de contraintes de ressources en raison de la croissance ou de l'extension du réseau. Les contraintes de ressources peuvent inclure :

- Trafic global

- Volume
- Nombre de routes
- Nombre de circuits virtuels
- Nombre de voisins
- Domaines de diffusion
- Débit des périphériques
- capacité média

Les planificateurs réseau doivent déterminer la durée de vie requise de la conception, les extensions ou les sites attendus pendant toute la durée de vie de la conception, le volume de nouveaux utilisateurs et le volume ou la modification du trafic attendu. Ce plan permet de s'assurer que la solution proposée répond aux exigences de croissance pendant la durée de vie prévue de la conception.

Si vous n'étudiez pas l'évolutivité de la solution, vous serez peut-être contraint de mettre en oeuvre des modifications de conception réactives majeures. Cette modification de conception peut inclure une hiérarchie supplémentaire, des mises à niveau de support ou des mises à niveau matérielles. Dans les entreprises qui dépendent de cycles budgétaires assez précis pour leurs achats de matériel importants, ces changements peuvent constituer un frein majeur à la réussite globale. En termes de disponibilité, les réseaux peuvent connaître des limitations de ressources inattendues qui entraînent des périodes de non disponibilité et des mesures réactives.

Les tests d'interopérabilité et d'interopérabilité peuvent être essentiels à la réussite des déploiements de nouvelles solutions. L'interopérabilité peut faire référence à différents fournisseurs de matériel, ou à différentes topologies ou solutions qui doivent s'interconnecter pendant ou après une mise en oeuvre de réseau. Les problèmes d'interopérabilité peuvent inclure la signalisation matérielle via la pile de protocoles vers des problèmes de routage ou de transport. Les problèmes d'interopérabilité peuvent survenir avant, pendant ou après la migration d'une solution réseau. La planification de l'interopérabilité doit inclure la connectivité entre différents périphériques et les problèmes de topologie susceptibles de se produire lors des migrations.

La comparaison de solutions consiste à comparer différentes conceptions potentielles par rapport à d'autres pratiques relatives aux exigences de la solution. Cette pratique permet de s'assurer que la solution est la mieux adaptée à un environnement particulier et que le processus de conception ne repose pas sur des préjugés personnels. La comparaison peut inclure différents facteurs tels que le coût, la résilience, la disponibilité, le risque, l'interopérabilité, la facilité de gestion, l'évolutivité et les performances. Tous ces éléments peuvent avoir un impact majeur sur la disponibilité globale du réseau une fois la conception mise en oeuvre. Vous pouvez également comparer les supports, la hiérarchie, la redondance, les protocoles de routage et des fonctionnalités similaires. Créez un graphique avec des facteurs sur l'axe des abscisses et des solutions potentielles sur l'axe des ordonnées afin de résumer les comparaisons de solutions. La comparaison détaillée des solutions dans un environnement de laboratoire permet également

d'étudier objectivement les nouvelles solutions et fonctionnalités en fonction des différents facteurs de comparaison.

Dans le cadre du concept d'exploitation de la gestion du réseau, il est essentiel de définir les objectifs du réseau et des services pris en charge d'une manière compréhensible par tous les utilisateurs. Les activités qui suivent l'élaboration du concept opérationnel sont fortement influencées par la qualité de ce document.

Voici les objectifs de performance standard :

- Temps de réponse
- Utilisation
- Débit
- Capacité (débit maximal)

Bien que ces mesures puissent être triviales pour un LAN simple, elles peuvent être très difficiles sur un réseau de campus commuté ou un réseau d'entreprise multifournisseur. Lorsque vous utilisez un concept bien pensé de plan d'opérations, chacun des objectifs de performance est défini d'une manière mesurable. Par exemple, le temps de réponse minimal pour l'application « x » est de 500 Ms ou moins pendant les heures de pointe. Elle définit les informations permettant d'identifier la variable, la manière de la mesurer et la période de la journée sur laquelle l'application de gestion du réseau doit se concentrer.

Définition des objectifs de disponibilité et de gestion du réseau

Les objectifs de disponibilité définissent le niveau de service ou les exigences de niveau de service pour un service réseau. Cela permet de garantir que la solution répond aux exigences de disponibilité finale. Définir différentes classes de service pour une organisation particulière et détailler les exigences réseau pour chaque classe qui sont appropriées aux exigences de disponibilité. Différentes zones du réseau peuvent également nécessiter différents niveaux de disponibilité. Un objectif de disponibilité plus élevé peut nécessiter une redondance et des procédures d'assistance accrues. Lorsque vous définissez un objectif de disponibilité pour un service réseau particulier et que vous mesurez la disponibilité, votre organisation réseau peut comprendre les composants et les niveaux de service requis pour atteindre les SLA projetés.

Définissez des objectifs de facilité de gestion afin de vous assurer que la gestion globale du réseau ne manque pas de fonctionnalités de gestion. Pour définir des objectifs de facilité de gestion, vous devez comprendre le processus d'assistance et les outils de gestion réseau associés de votre entreprise. Les objectifs de gérabilité doivent inclure la connaissance de la manière dont les nouvelles solutions s'intègrent dans le modèle d'assistance et d'outil actuel, avec des références aux différences potentielles ou aux nouvelles exigences. Cela est essentiel à la disponibilité du réseau, car la capacité à prendre en charge de nouvelles solutions est primordiale pour la réussite du déploiement et pour atteindre les objectifs de disponibilité.

Les objectifs de facilité de gestion doivent comprendre toutes les informations importantes sur la

base MIB ou les outils réseau nécessaires pour prendre en charge un réseau potentiel, la formation requise pour prendre en charge le nouveau service réseau, les modèles de recrutement pour le nouveau service et toute autre exigence d'assistance. Souvent, ces informations ne sont pas découvertes avant le déploiement et la disponibilité globale pâtit du manque de ressources affectées à la prise en charge de la nouvelle conception de réseau.

Définir les SLA et les mesures de performances

Les SLA et les mesures de performances permettent de définir et de mesurer les performances des nouvelles solutions réseau afin de s'assurer qu'elles répondent aux exigences de performances. Les performances de la solution proposée peuvent être mesurées à l'aide d'outils de surveillance des performances ou d'une simple requête ping sur l'infrastructure réseau proposée. Les SLA de performances doivent inclure le volume moyen attendu de trafic, le volume de pointe de trafic, le temps de réponse moyen et le temps de réponse maximal autorisé. Ces informations peuvent ensuite être utilisées ultérieurement dans la section relative à la validation de la solution et aident à déterminer les performances et la disponibilité requises du réseau.

Définition des SLA

La définition du service pour les utilisateurs ou les clients constitue un aspect important de la conception du réseau. Les entreprises appellent ces accords de niveau de service tandis que les fournisseurs de services les appellent gestion des niveaux de service. La gestion des niveaux de service inclut généralement des définitions des types et de la gravité des problèmes, ainsi que des responsabilités du centre d'assistance, telles que le chemin d'escalade et le délai avant l'escalade à chaque niveau d'assistance, le délai avant le début du travail sur le problème et le délai avant la fermeture des cibles en fonction de la priorité. Les autres facteurs importants sont le service fourni dans le domaine de la planification de la capacité, de la gestion proactive des pannes, de la notification de la gestion des changements, des seuils, des critères de mise à niveau et du remplacement du matériel.

Lorsque les entreprises ne définissent pas les niveaux de service à l'avance, il devient difficile d'améliorer ou d'obtenir les besoins en ressources identifiés à une date ultérieure. Il devient également difficile de comprendre quelles ressources ajouter pour faciliter la prise en charge du réseau. Dans de nombreux cas, ces ressources ne sont appliquées qu'après la découverte de problèmes.

Mesurer les performances

La gestion du rendement est un terme générique qui englobe la configuration et la mesure de secteurs de rendement distincts. Cette section décrit les six concepts suivants de la gestion des performances :

- Collecte des données de base du réseau
- Mesurer la disponibilité
- Mesurer le temps de réponse

- Précision des mesures
- Utilisation des mesures
- Planification de capacité

Collecte des données de base du réseau

La plupart des intranets d'entreprise disposent d'une bande passante suffisante. Cependant, sans données adéquates, vous ne pourrez peut-être pas exclure la possibilité que le réseau soit encombré, car cela contribue aux performances médiocres des applications. L'un des indices de congestion ou d'erreurs est que les performances médiocres sont intermittentes ou dépendent de l'heure de la journée. Un exemple de cette condition est lorsque la performance est adéquate tard dans la soirée, mais très lent le matin et pendant les heures de pointe d'affaires.

Une fois que vous avez défini le concept d'opérations de gestion du réseau et les données d'implémentation nécessaires, il est nécessaire de collecter ces données au fil du temps. Ce type de collecte constitue la base de la ligne de base du réseau.

Effectuer une planification initiale du réseau actuel avant le déploiement d'une nouvelle solution (application ou modification de l'IOS) et après le déploiement afin de mesurer les attentes définies pour la nouvelle solution. Cette ligne de base permet de déterminer si la solution répond aux objectifs de performances et de disponibilité, et d'évaluer la capacité. Un rapport de base type sur un routeur/commutateur inclut des problèmes de capacité liés au processeur, à la mémoire, à la gestion de la mémoire tampon, à l'utilisation des liaisons/supports et au débit. Il existe d'autres types de données de référence que vous pouvez également inclure, en fonction des objectifs définis dans le concept d'opérations. Par exemple, une ligne de base de disponibilité démontre une stabilité/disponibilité accrue de l'environnement réseau. Effectuer une comparaison de base entre les anciens et les nouveaux environnements afin de vérifier les exigences de la solution.

Une autre ligne de base spécialisée est la ligne de base des applications, qui est utile lorsque vous établissez une tendance des besoins du réseau d'applications. Ces informations peuvent être utilisées à des fins de facturation et/ou de budgétisation dans le cycle de mise à niveau. Les niveaux de référence des applications peuvent également être importants dans le domaine de la disponibilité des applications par rapport aux services préférés ou aux qualités de service par application. Les informations de base des applications sont principalement constituées de la bande passante utilisée par les applications par période. Certaines applications de gestion du réseau peuvent également servir de référence pour les performances des applications. Une répartition du type de trafic (Telnet ou FTP) est également importante pour la planification. Dans certaines entreprises, les zones du réseau où les ressources sont limitées sont surveillées afin de détecter les meilleurs correspondants. Les administrateurs réseau peuvent utiliser ces informations pour budgéter, planifier ou régler le réseau. Lorsque vous réglez le réseau, vous pouvez modifier les paramètres de qualité de service ou de file d'attente du service réseau ou de l'application.

Mesurer la disponibilité

L’une des principales mesures utilisées par les administrateurs réseau est la disponibilité. La disponibilité est la mesure du temps pendant lequel un système ou une application réseau est disponible pour un utilisateur. Du point de vue du réseau, la disponibilité représente la fiabilité des composants individuels d’un réseau.

Par exemple, afin de mesurer la disponibilité, vous pouvez coordonner les appels téléphoniques du centre d'assistance avec les statistiques collectées à partir des périphériques gérés. Cependant, les outils de disponibilité ne peuvent pas déterminer toutes les raisons de l'échec.

La redondance du réseau est un autre facteur à prendre en compte lorsque vous mesurez la disponibilité. La perte de redondance indique une dégradation du service plutôt qu'une panne totale du réseau. Le résultat peut être un temps de réponse plus lent et une perte de données due à l'abandon de paquets. Il est également possible que les résultats se manifestent dans d'autres domaines de la mesure du rendement, comme l'utilisation et le temps de réponse.

Enfin, si vous appliquez un contrat de niveau de service, vous devez tenir compte des interruptions programmées. Ces pannes peuvent être le résultat de déplacements, d'ajouts et de modifications, d'arrêts d'usine ou d'autres événements que vous ne souhaitez pas signaler. Il ne s'agit pas seulement d'une tâche difficile, mais également d'une tâche manuelle.

Mesurer le temps de réponse

Le temps de réponse du réseau est le temps nécessaire au trafic pour se déplacer entre deux points. Des temps de réponse plus lents que la normale, observés lors d'une comparaison de base ou qui dépassent un seuil, peuvent indiquer un encombrement ou une défaillance du réseau.

Le temps de réponse est la meilleure mesure de l'utilisation du réseau par le client et peut vous aider à évaluer l'efficacité de votre réseau. Quelle que soit la source de la lenteur de la réponse, les utilisateurs sont frustrés en raison du trafic retardé. Dans les réseaux distribués, de nombreux facteurs affectent le temps de réponse, tels que :

- Encombrement du réseau
- Route inférieure à la destination souhaitée (ou aucune route)
- Périphériques réseau sous-alimentés
- Défaillances du réseau, telles qu'une tempête de diffusion
- Bruit ou erreurs CRC

Dans les réseaux qui utilisent la mise en file d’attente liée à la QoS, la mesure du temps de réponse est importante afin de déterminer si les types de trafic corrects circulent sur le réseau comme prévu. Par exemple, lorsque vous implémentez le trafic vocal sur des réseaux IP, les paquets vocaux doivent être livrés à temps et à un débit constant afin de maintenir une bonne qualité vocale. Vous pouvez générer du trafic classé comme trafic vocal afin de mesurer le temps de réponse du trafic tel qu'il apparaît aux utilisateurs.

Vous pouvez mesurer le temps de réponse afin d'aider à résoudre les conflits entre les serveurs

d'applications et les administrateurs réseau. Les administrateurs réseau sont souvent présumés coupables lorsqu'une application ou un serveur semble être lent. L'administrateur réseau doit prouver que le réseau n'est pas le problème. La collecte de données sur le temps de réponse fournit un moyen incontestable de prouver ou de réfuter que le réseau est la source de problèmes d'application.

Dans la mesure du possible, vous devez mesurer le temps de réponse tel qu'il apparaît aux utilisateurs. Un utilisateur perçoit la réponse comme le temps qui s'écoule entre le moment où il appuie sur Entrée ou clique sur un bouton et l'affichage de l'écran. Ce temps écoulé comprend le temps nécessaire à chaque périphérique réseau, à la station de travail de l'utilisateur et au serveur de destination pour traiter le trafic.

Malheureusement, la mesure à ce niveau est presque impossible en raison du nombre d'utilisateurs et du manque d'outils. En outre, lorsque vous intégrez le temps de réponse de l'utilisateur et du serveur, cela n'apporte que peu de valeur lorsque vous déterminez la croissance future du réseau ou que vous résolvez des problèmes réseau.

Vous pouvez utiliser les périphériques réseau et les serveurs pour mesurer le temps de réponse. Vous pouvez également utiliser des outils comme ICMP pour mesurer les transactions, bien qu'il ne prenne pas en compte les retards introduits dans un système lorsque les couches supérieures le traitent. Cette approche résout le problème de la connaissance des performances du réseau.

À un niveau simplifié, vous pouvez chronométrer la réponse aux requêtes ping de la station de gestion du réseau vers des points clés du réseau, tels qu'une interface mainframe, un point d'extrémité d'une connexion de fournisseur de services ou des adresses IP d'utilisateur clé, afin de mesurer le temps de réponse. Le problème avec cette méthode est qu'elle ne reflète pas précisément la perception de l'utilisateur du temps de réponse entre sa machine et la machine de destination. Il collecte simplement des informations et signale le temps de réponse du point de vue de la station de gestion du réseau. Cette méthode masque également les problèmes de temps de réponse saut par saut sur l'ensemble du réseau.

Une alternative à l'interrogation centrée sur le serveur consiste à répartir l'effort plus près de la source et de la destination que vous souhaitez simuler pour la mesure. Utiliser des interrogateurs de gestion de réseau distribués et mettre en oeuvre la fonctionnalité Cisco IOS Service Assurance Agent (SAA). Vous pouvez activer SAA sur des routeurs afin de mesurer le temps de réponse entre un routeur et un périphérique de destination tel qu'un serveur ou un autre routeur. Vous pouvez également spécifier un port TCP ou UDP, qui force le trafic à être transféré et dirigé de la même manière que le trafic simulé.

Grâce à l'intégration de la voix, de la vidéo et des données sur les réseaux multiservices, les clients mettent en oeuvre la hiérarchisation de la qualité de service sur leur réseau. Les mesures ICMP ou UDP simples ne reflètent pas avec précision le temps de réponse, car différentes applications reçoivent des priorités différentes. En outre, avec la commutation par étiquette, le routage du trafic peut varier en fonction du type d'application contenu dans un paquet spécifique. Ainsi, une requête ping ICMP peut recevoir des priorités différentes dans la façon dont chaque routeur la gère et peut recevoir des routes différentes et moins efficaces.

Dans ce cas, la seule façon de mesurer le temps de réponse est de générer du trafic qui

ressemble à l'application ou à la technologie concernée. Cela oblige les périphériques réseau à gérer le trafic comme ils le feraient pour le trafic réel. Vous pourrez peut-être atteindre ce niveau avec SAA ou en utilisant des sondes tierces sensibles aux applications.

Précision des mesures

La précision est la mesure du trafic d'interface qui ne génère pas d'erreur et peut être exprimée en termes de pourcentage qui compare le taux de réussite au taux de paquets total sur une période donnée. Vous devez d'abord mesurer le taux d'erreur. Par exemple, si deux paquets sur 100 génèrent une erreur, le taux d'erreur est de 2 % et le taux de précision est de 98 %.

Avec les technologies réseau antérieures, en particulier dans le domaine étendu, un certain niveau d’erreurs était acceptable. Cependant, avec les réseaux à haut débit et les services WAN actuels, la transmission est beaucoup plus précise et les taux d'erreur sont proches de zéro, sauf en cas de problème réel. Voici quelques-unes des causes courantes des erreurs d’interface :

- Câblage non conforme
- Interférences électriques
- Matériel ou logiciel défectueux

Utilisez un taux de précision réduit pour déclencher une enquête plus approfondie. Vous découvrirez peut-être qu’une interface particulière présente des problèmes et décidera que les erreurs sont acceptables. Dans ce cas, vous devez ajuster le seuil de précision pour cette interface afin de refléter les cas où le taux d'erreur est inacceptable. Le taux d'erreur inacceptable a peut-être été signalé dans une base de référence antérieure.

Les variables décrites dans ce tableau sont utilisées dans les formules de précision et de taux d'erreur :

Notation	Description
siErreursEntrantes	Delta (ou différence) entre deux cycles d'interrogation qui collectent l'objet snmp ifInErrors, qui représente le nombre de paquets entrants avec une erreur.
siPaquetsInUcast	Delta entre deux cycles d'interrogation qui collectent l'objet snmp ifInUcastPkts, qui représente le nombre de paquets de monodiffusion entrants.
siPaquetsInNUcast	Delta entre les deux cycles d'interrogation qui collectent l'objet snmp ifInNUcastPkts, qui représente le nombre de paquets entrants non monodiffusion (multidiffusion et

	diffusion).

La formule du taux d'erreur est généralement exprimée en pourcentage :

Taux d'erreur = (ifInErrors) *100

(ifInUcastPkts + (ifInNUcastPkts)

Notez que les erreurs sortantes ne sont pas prises en compte dans les formules de taux d'erreur et d'exactitude. En effet, un périphérique ne doit jamais sciemment placer des paquets comportant des erreurs sur le réseau et les taux d'erreur d'interface sortante ne doivent jamais augmenter. Par conséquent, le trafic entrant et les erreurs sont les seules mesures d'intérêt pour les erreurs d'interface et la précision.

La formule de précision prend le taux d'erreur et le soustrait de 100 (encore une fois, sous la forme d'un pourcentage) :

Précision = 100 - (ifInErrors) *100

(ifInUcastPkts + (ifInNUcastPkts)

Ces formules reflètent l'erreur et la précision en termes de compteurs génériques d'interface MIB II (RFC 2233). Le résultat est exprimé en termes de pourcentage qui compare les erreurs au nombre total de paquets vus et envoyés. Le taux d'erreur qui en résulte est soustrait de 100, ce qui produit le taux de précision. Un taux de précision de 100 % est parfait.

Puisque les variables MIB II sont stockées en tant que compteurs, vous devez prendre deux cycles d'interrogation et calculer la différence entre les deux (d'où le Delta utilisé dans l'équation).

Utilisation des mesures

L'utilisation mesure l'utilisation d'une ressource particulière dans le temps. La mesure est généralement exprimée sous la forme d'un pourcentage dans lequel l'utilisation d'une ressource est comparée à sa capacité opérationnelle maximale. Grâce aux mesures d'utilisation, vous pouvez identifier l'encombrement (ou l'encombrement potentiel) sur tout le réseau. Vous pouvez également identifier les ressources sous-utilisées.

L'utilisation est la mesure principale permettant de déterminer le niveau de remplissage des canaux (liaisons) du réseau. Mesurez l'UC, l'interface, la mise en file d'attente et d'autres mesures de capacité liées au système afin de déterminer le degré de consommation des ressources du système réseau.

Une utilisation élevée n'est pas nécessairement mauvaise. Une faible utilisation peut indiquer des flux de trafic à des endroits inattendus. À mesure que les lignes deviennent surutilisées, les effets peuvent devenir importants. Une surexploitation se produit lorsqu'il y a plus de trafic en file

d'attente pour passer sur une interface qu'il ne peut en gérer. Des sauts soudains dans l'utilisation des ressources peuvent indiquer une condition de panne.

Lorsqu'une interface devient encombrée, le périphérique réseau doit soit stocker le paquet dans une file d'attente, soit l'abandonner. Si un routeur tente de stocker un paquet dans une file d'attente complète, le paquet est abandonné. Les paquets abandonnés se produisent lorsque le trafic est transféré d'une interface rapide à une interface plus lente. Ceci est indiqué dans la formule Q = u / (1-u) où u est l'utilisation, et Q est la profondeur moyenne de la file d'attente (trafic aléatoire supposé). Ainsi, des niveaux d'utilisation élevés sur les liaisons entraînent des profondeurs de file d'attente moyennes élevées, ce qui est une latence prévisible si vous connaissez la taille du paquet. Certains fournisseurs de rapports réseau indiquent que vous pouvez commander moins de bande passante et payer moins cher pour votre WAN. Cependant, les implications de latence apparaissent lorsque vous exécutez des liaisons WAN à 95 % d'utilisation. En outre, à mesure que les réseaux sont migrés vers la VoIP, les administrateurs réseau peuvent devoir modifier leurs politiques et exécuter des liaisons WAN avec un taux d'utilisation d'environ 50 %.

Lorsqu'un paquet est abandonné, le protocole de couche supérieure peut forcer une retransmission du paquet. Si plusieurs paquets sont abandonnés, un trafic de nouvelle tentative excessif peut en résulter. Ce type de réaction peut entraîner des sauvegardes sur des périphériques situés plus loin dans la ligne. Afin de résoudre ce problème, vous pouvez définir différents degrés de seuils.

La principale mesure utilisée pour l'utilisation du réseau est l'utilisation des interfaces. Utilisez les formules décrites dans ce tableau selon que la connexion que vous mesurez est en mode bidirectionnel non simultané ou bidirectionnel simultané :

Notation	Description
siOctets entrants	Delta (ou différence) entre deux cycles d'interrogation qui collectent l'objet snmp ifInOctets, qui représente le nombre d'octets entrants du trafic.
siOctetsSortants	Delta entre deux cycles d'interrogation qui collectent l'objet snmp ifOutOctets qui représente le nombre d'octets sortants du trafic.
siVitesse	Vitesse de l'interface indiquée dans l'objet snmp ifSpeed. Notez que ifSpeed peut ne pas refléter précisément la vitesse d'une interface WAN.

Les connexions de réseau local partagé tendent à être en mode bidirectionnel non simultané, principalement parce que la détection des conflits nécessite qu'un périphérique écoute avant de transmettre. Les connexions WAN sont généralement en mode bidirectionnel simultané, car la

connexion est point à point ; les deux périphériques peuvent transmettre et recevoir en même temps, car ils savent qu'un seul autre périphérique partage la connexion.

Puisque les variables MIB II sont stockées en tant que compteurs, vous devez prendre deux cycles d'interrogation et calculer la différence entre les deux (d'où le Delta utilisé dans l'équation).

Pour les supports bidirectionnels non simultanés, utilisez la formule suivante pour l'utilisation de l'interface :

(ifInOctets + ifOutOctets) * 8 * 100

(nombre de secondes en) * ifSpeed

Pour les supports bidirectionnels simultanés, le calcul de l'utilisation est plus complexe. Par exemple, avec une connexion série T-1 complète, la vitesse de la ligne est de 1,544 Mbits/s. Cela signifie qu'une interface T-1 peut à la fois recevoir et transmettre 1,544 Mbits/s pour une bande passante combinée possible de 3,088 Mbits/s.

Lorsque vous calculez la bande passante d'interface pour les connexions bidirectionnelles simultanées, vous pouvez utiliser cette formule dans laquelle vous prenez la plus grande des valeurs in et out et générez un pourcentage d'utilisation :

max(ifInOctets, (ifOutOctets) * 8 * 100

(nombre de secondes en) * ifSpeed

Cependant, cette méthode masque l'utilisation de la direction qui a la valeur la plus faible et fournit des résultats moins précis. Une méthode plus précise consiste à mesurer séparément l'utilisation des entrées et l'utilisation des sorties, par exemple :

Utilisation en entrée = ifInOctets *8 * 100

(nombre de secondes en) * ifSpeed

ET

Utilisation de la sortie = ifOutOctets *8 * 100

(nombre de secondes en) * ifSpeed

Bien que ces formules soient quelque peu simplifiées, elles ne tiennent pas compte de la surcharge associée à un protocole particulier. Il existe des formules plus précises pour gérer les aspects uniques de chaque protocole. Par exemple, la RFC 1757 contient des formules

d'utilisation Ethernet qui prennent en compte la surcharge de paquets. Cependant, l'équipe de haute disponibilité a constaté que les formules générales présentées ici peuvent être utilisées de manière fiable sur les interfaces LAN et WAN dans la plupart des cas.

Planification de capacité

Comme indiqué précédemment, la planification de la capacité est le processus par lequel vous déterminez les besoins futurs probables en ressources réseau afin d'éviter un impact sur les performances ou la disponibilité des applications critiques de l'entreprise. Reportez-vous à la section Gestion des capacités et des performances : Livre blanc sur les meilleures pratiques pour plus d'informations sur ce sujet.

Effectuer une analyse proactive des pannes

L'analyse proactive des pannes est essentielle à la gestion des performances. Le même type de données collectées pour la gestion des performances peut être utilisé pour l'analyse proactive des pannes. Cependant, la synchronisation et l'utilisation de ces données diffèrent entre la gestion proactive des pannes et la gestion des performances.

La gestion proactive des pannes est le moyen par lequel le système de gestion de réseau idéal peut atteindre les objectifs que vous avez fixés. La relation avec la gestion des performances se fait par le biais de la ligne de base et des variables de données que vous utilisez. La gestion proactive des pannes intègre des événements personnalisés, un moteur de corrélation d'événements, la signalisation des incidents et l'analyse statistique des données de base afin de relier la gestion des pannes, des performances et des changements dans un système de gestion du réseau idéal et efficace.

Lorsque l'interrogation des données de performances est normalement effectuée toutes les 10, 15, voire 30 minutes, la reconnaissance d'une condition de panne doit être effectuée à un intervalle de temps beaucoup plus court. Une méthode de gestion proactive des pannes consiste à utiliser des alarmes et des groupes d'événements RMON. Vous pouvez définir des seuils sur vos périphériques qui ne sont pas interrogés par des périphériques externes afin que les seuils soient beaucoup plus courts. Une autre méthode, qui n'est pas traitée dans ce document, est par l'utilisation d'un système de gestion distribuée qui permet d'effectuer des sondages au niveau local avec agrégation de données au niveau d'un gestionnaire de gestionnaires.

Utiliser des seuils pour une gestion proactive des pannes

Le seuil est le processus par lequel vous définissez des points d'intérêt dans des flux de données spécifiques et générez des événements lorsque des seuils sont déclenchés. Utilisez vos données de performances réseau pour définir ces seuils.

Il existe différents types de seuils, dont certains sont plus applicables à certains types de données. Les seuils ne s'appliquent qu'aux données numériques. Par conséquent, convertissez toutes les données textuelles en valeurs numériques discrètes. Même si vous ne connaissez pas toutes les chaînes de texte possibles pour un objet, vous pouvez toujours énumérer les chaînes « intéressantes » et affecter toutes les autres chaînes à une valeur définie.

Il existe deux classes de seuils pour les deux classes de données numériques : continu et discret. Les seuils continus s'appliquent aux données continues ou chronologiques, telles que les données stockées dans les compteurs ou les jauges SNMP. Des seuils discrets s'appliquent aux objets énumérés ou à toute donnée numérique discrète. Les objets booléens sont des valeurs énumérées avec deux valeurs : vrai ou faux. Les données discrètes peuvent également être appelées données d'événements car les événements marquent la transition d'une valeur à la suivante.

Les seuils continus peuvent déclencher des événements lorsque l'objet de série chronologique dépasse la valeur spécifiée du seuil. La valeur de l'objet s'élève au-dessus du seuil ou en dessous. Il peut également être utile de définir des seuils d'augmentation et de baisse distincts. Cette technique, appelée mécanisme d'hystérésis, permet de réduire le nombre d'événements générés à partir de cette classe de données. Le mécanisme d'hystérésis permet de réduire le volume d'événements générés par des seuils sur des données de séries chronologiques à variation rapide. Ce mécanisme peut être utilisé avec n'importe quelle technique de seuil sur des données de séries temporelles.

Le volume d'événements est réduit par une alarme qui est générée pour suivre la valeur d'un objet. Des seuils de montée et de descente sont attribués à cette alarme. L'alarme n'est déclenchée que lorsque le seuil ascendant est franchi. Une fois ce seuil franchi, une alarme montante n'est pas à nouveau générée tant que le seuil descendant n'est pas franchi. Et le même mécanisme empêche la génération de seuils descendants jusqu'à ce que le seuil ascendant soit à nouveau franchi. Ce mécanisme peut réduire considérablement le volume d'événements et n'élimine pas les informations nécessaires pour déterminer si une panne existe.

Les données de séries chronologiques peuvent être représentées sous forme de compteurs, où chaque nouveau point de données est ajouté à la somme des points de données précédents, ou sous forme de jauge, où les données sont représentées sous forme de taux sur un intervalle de temps. Il existe deux formes différentes de seuils continus applicables à chaque type de données : les seuils continus absolus et les seuils continus relatifs. Utilisez des seuils continus absolus avec des jauges et des seuils continus relatifs avec des compteurs.

Pour déterminer les valeurs de seuil de votre réseau, procédez comme suit :

1. Sélectionnez les objets.
2. Sélectionnez les périphériques et les interfaces.
3. Déterminez les valeurs de seuil pour chaque type d'objet ou d'objet/interface.
4. Déterminez la gravité de l'événement généré par chaque seuil.

Un travail considérable est nécessaire pour déterminer les seuils à utiliser sur quels objets (et pour quels périphériques et interfaces). Heureusement, si vous avez collecté une base de données de performances, vous avez déjà effectué une grande partie de ce travail. En outre, la NSA et le programme de service haute disponibilité (HAS) peuvent formuler des recommandations qui vous aident à définir des objets et à créer des plages. Toutefois, vous devez adapter ces recommandations à votre réseau.

Comme vous avez collecté des données de performances pour le réseau, le programme HAS vous recommande de regrouper vos interfaces par catégories. Cela simplifie la définition des seuils, car vous devrez peut-être déterminer des seuils pour le type de support de chaque catégorie plutôt que pour chaque périphérique et objet sur ce périphérique. Par exemple, vous pouvez définir des seuils différents pour les réseaux Ethernet et FDDI. Il est généralement admis que vous pouvez exécuter des réseaux FDDI avec une utilisation plus proche de 100 % de celle d'un segment Ethernet partagé. Cependant, l'Ethernet bidirectionnel simultané peut être exécuté beaucoup plus près de l'utilisation à 100 %, car il n'est pas sujet à des collisions. Vous pouvez définir des seuils de collision très bas pour les liaisons bidirectionnelles simultanées, car vous ne devriez jamais voir de collision.

Vous pouvez également tenir compte de la combinaison de l'importance de l'interface et de la catégorie/gravité du type de seuil. Utilisez ces facteurs pour définir la priorité de l'événement et, par conséquent, l'importance de l'événement et son attention par le personnel d'exploitation du réseau.

On ne saurait trop insister sur le regroupement et la catégorisation des périphériques et des interfaces réseau. Plus vous êtes en mesure de regrouper et de classer les événements de seuil, plus il est facile d'intégrer les événements de seuil dans votre plate-forme de gestion du réseau. Utilisez la planification initiale comme principale ressource pour ces informations. Reportez-vous à la section Gestion des capacités et des performances : Livre blanc sur les meilleures pratiques pour plus d'informations.

Implémentation de la gestion du réseau

L'entreprise doit disposer d'un système de gestion de réseau mis en oeuvre capable de détecter les valeurs de seuil définies et de générer des rapports sur ces valeurs pour des périodes spécifiques. Utilisez un système de gestion de réseau RMON qui peut archiver les messages de seuil dans un fichier journal pour une révision quotidienne ou une solution de base de données plus complète qui permet de rechercher des exceptions de seuil pour un paramètre donné. Les informations doivent être mises à la disposition du personnel et du responsable de l'exploitation du réseau de manière continue. La mise en oeuvre de la gestion du réseau doit inclure la capacité à détecter les pannes logicielles/matérielles ou les remontées, la fiabilité de l'interface, le CPU, l'utilisation des liaisons, les échecs de mise en file d'attente ou de tampon, le volume de diffusion, les transitions de porteuse et les réinitialisations d'interface.

Mesures de fonctionnement du réseau

Enfin, la gestion proactive des pannes qui chevauche la gestion des performances concerne les mesures d'exploitation du réseau. Ces mesures fournissent des données précieuses pour l'amélioration des processus de gestion des pannes. Au minimum, ces mesures devraient comprendre une ventilation de tous les problèmes qui se sont produits au cours d'une période donnée. La ventilation doit inclure des informations telles que :

- Nombre de problèmes qui se produisent par priorité d'appel
- Durée minimale, maximale et moyenne de fermeture pour chaque priorité

- Répartition des problèmes par type de problème (matériel, panne logicielle, configuration, alimentation, erreur utilisateur)
- Répartition du temps de fermeture pour chaque type de problème
- Disponibilité par groupe de disponibilité ou SLA
- Fréquence à laquelle vous répondez aux exigences SLA ou les manquez

Le centre d'assistance dispose souvent d'un système de génération de rapports capable de générer des mesures ou des rapports. L'utilisation d'un outil de surveillance de la disponibilité constitue un autre moyen de collecter ces données. Les mesures globales devraient être disponibles sur une base mensuelle. L'amélioration des processus basée sur la discussion doit être mise en oeuvre afin d'améliorer les exigences des contrats de niveau de service manqués ou afin d'améliorer la façon dont certains types de problèmes sont traités.

Indicateurs de gestion des performances

Les indicateurs de performance fournissent le mécanisme par lequel une organisation mesure les facteurs de réussite critiques.

Documentation des objectifs commerciaux de la gestion du réseau

Ce document peut être un concept formel d'opérations pour la gestion du réseau ou un énoncé moins formel des caractéristiques et des objectifs requis. Cependant, le document doit aider l’administrateur réseau à mesurer le succès.

Ce document constitue la stratégie de gestion du réseau de l'entreprise et doit coordonner les objectifs commerciaux globaux (non quantitatifs) des opérations réseau, de l'ingénierie, de la conception, des autres unités commerciales et des utilisateurs finaux. Cette orientation permet à l'organisation de former les activités de planification à long terme pour la gestion et l'exploitation du réseau, ce qui inclut le processus de budgétisation. Il fournit également des conseils pour l’acquisition d’outils et le chemin d’intégration requis pour atteindre les objectifs de gestion du réseau, tels que les SLA.

Ce document stratégique ne peut pas se focaliser de manière trop étroite sur la gestion de problèmes de réseau spécifiques, mais sur les éléments importants pour l'organisation dans son ensemble, qui incluent des questions budgétaires. Exemple :

- Identifier un plan complet avec des objectifs réalisables.
- Identifiez chaque service/application d'entreprise nécessitant une assistance réseau.
- Identifiez les indicateurs de performances nécessaires pour mesurer le service.
- Planifier la collecte et la distribution des données de mesure des performances.
- Identifiez l'assistance requise pour l'évaluation du réseau et les commentaires des

utilisateurs.

- avoir des objectifs de niveau de service documentés, détaillés et mesurables ;

Documenter les contrats de niveau de service

Afin de documenter correctement les SLA, vous devez définir entièrement les mesures des objectifs de niveau de service. Cette documentation doit être mise à la disposition des utilisateurs pour évaluation. Il fournit la boucle de rétroaction pour s'assurer que l'organisation de gestion du réseau continue à mesurer les variables nécessaires pour maintenir le niveau d'accord de service.

Les SLA sont des documents « vivants », car l'environnement commercial et le réseau sont par nature dynamiques. Ce qui fonctionne aujourd'hui pour mesurer un SLA peut devenir obsolète demain. Ce n'est que lorsqu'ils mettent en place une boucle de rétroaction des utilisateurs et agissent sur ces informations que les opérations réseau peuvent maintenir les numéros de haute disponibilité requis par l'entreprise.

Créer une liste de variables pour la ligne de base

Cette liste inclut des éléments tels que l'intervalle d'interrogation, la surcharge de gestion du réseau encourue, les seuils de déclenchement possibles, l'utilisation de la variable comme déclencheur d'un déroutement et l'analyse des tendances utilisée pour chaque variable.

Ces variables ne se limitent pas aux mesures nécessaires pour atteindre les objectifs de niveau de service mentionnés ci-dessus. Au minimum, elles doivent inclure les variables suivantes : état du routeur, état du commutateur, informations de routage, données spécifiques à la technologie, utilisation et délai. Ces variables sont interrogées périodiquement et stockées dans une base de données. Des rapports peuvent ensuite être générés à partir de ces données. Ces rapports peuvent aider le personnel chargé des opérations de gestion et de la planification du réseau de différentes manières :

- Les problèmes réactifs peuvent souvent être résolus plus rapidement avec une base de données historique.
- Les rapports sur le rendement et la planification des capacités nécessitent ce type de données.
- Les objectifs de niveau de service peuvent être mesurés par rapport à ceux-ci.

Examiner les analyses de référence et de tendances

Le personnel de gestion du réseau doit organiser des réunions pour examiner périodiquement des rapports spécifiques. Cela permet d'obtenir des commentaires supplémentaires, ainsi qu'une approche proactive des problèmes potentiels sur le réseau.

Ces réunions devraient inclure à la fois le personnel opérationnel et le personnel de planification. Les planificateurs ont ainsi l'occasion de recevoir une analyse opérationnelle des données de base et des tendances. Il met également le personnel opérationnel « dans la boucle » pour une

partie de l'analyse de planification.

Les objectifs de niveau de service constituent un autre type de point à inclure dans ces réunions. À mesure que des seuils objectifs sont approchés, le personnel d'administration du réseau peut prendre des mesures pour éviter de manquer un objectif et, dans certains cas, ces données peuvent être utilisées comme justification budgétaire partielle. Les données peuvent indiquer où les objectifs de niveau de service doivent être atteints si les mesures appropriées ne sont pas prises. De plus, comme ces objectifs ont été identifiés par les services et les applications de l'entreprise, ils sont plus faciles à justifier sur le plan financier.

Procéder à ces examens toutes les deux semaines et tenir une réunion analytique plus approfondie toutes les six à douze semaines. Ces réunions vous permettent d'aborder des questions à court et à long terme.

Documentation d'une méthodologie d'analyse de scénarios

Une analyse de simulation implique la modélisation et la vérification de solutions. Avant d'ajouter une nouvelle solution au réseau (soit une nouvelle application, soit une modification de la version de Cisco IOS), documentez certaines alternatives.

La documentation de cette analyse inclut les principales questions, la méthodologie, les ensembles de données et les fichiers de configuration. Le point principal est que l'analyse par simulation est une expérience que quelqu'un d'autre devrait pouvoir recréer avec les informations fournies dans le document.

Documentation de la méthodologie utilisée pour améliorer les performances du réseau

Cette documentation inclut une bande passante WAN supplémentaire et un tableau des coûts qui permet d'augmenter la bande passante pour un type de liaison particulier. Ces informations aident l'entreprise à évaluer le temps et l'argent nécessaires à l'augmentation de la bande passante. La documentation officielle permet aux experts en performances et en capacité de découvrir comment et quand augmenter les performances, ainsi que le calendrier et les coûts d'une telle entreprise.

Examiner périodiquement cette documentation, peut-être dans le cadre de l'examen trimestriel du rendement, afin de s'assurer qu'elle demeure à jour.

Résumé

Le seul moyen d'atteindre les objectifs du système de gestion de réseau idéal est d'intégrer activement les composants de la gestion des performances dans le système. Cet objectif doit inclure l'utilisation de mesures de disponibilité et de temps de réponse liées à un système de notification lorsque les seuils sont dépassés. Elle devrait inclure l'utilisation d'une ligne de base pour la planification de la capacité, avec des liens vers un modèle heuristique pour la mise en service et la génération de rapports sur les exceptions. Il peut être équipé d'un moteur de modélisation ou de simulation intégré qui permet de mettre à jour le modèle en temps réel et de

fournir un niveau de planification et de dépannage via des simulations logicielles.

Bien qu'une grande partie de ce système puisse sembler un idéal impossible à atteindre, chacun des composants est actuellement disponible. De plus, les outils d'intégration de ces composants existent également dans des programmes comme MicroMuse. Nous devrions continuer à travailler à cet idéal, car il est plus réaliste aujourd'hui que jamais.

Informations connexes

- Livres blancs sur les technologies de haute disponibilité
- Livre blanc sur la gestion des capacités et des performances
- Assistance et documentation techniques - 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.

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.