

Administración de capacidad y rendimiento: Informe oficial de Mejores Prácticas

ID del Documento: 20769

Actualizado: De oct el 04 de 2005



[Descarga PDF](#)



[Imprimir](#)

[Feedback](#)

Productos Relacionados

- [Alta disponibilidad](#)
- [QoS Congestion Management \(queueing\)](#)

Contenido

[Introducción](#)

[Descripción de la Administración de capacidad y rendimiento](#)

[CPU](#)

[Backplane o entrada-salida](#)

[Memoria](#)

[Tamaños de la interfaz y del tubo](#)

[Datos en espera, tiempo de espera, y jitter](#)

[Velocidad y distancia](#)

[Características de la aplicación](#)

[Mejores prácticas de la Administración de capacidad y rendimiento](#)

[Service Level Management](#)

[Análisis del qué si de la red y de la aplicación](#)

[Baselining y el tender](#)

[Administración de excepciones](#)

[Administración de QoS \(Calidad de servicio\)](#)

[Recogiendo y señalando la información de capacidad](#)

[Determine sus necesidades](#)

[Defina un proceso](#)

[Defina las áreas de capacidad](#)

[Defina las variables de capacidad](#)

[Interprete los datos](#)

[Información Relacionada](#)

[Discusiones relacionadas de la comunidad del soporte de Cisco](#)

Introducción

Una alta disponibilidad de red es un requisito de misión crítica dentro de las redes del proveedor de servicios y las grandes empresas. Los administradores de red se enfrentan a cada vez mayores desafíos para proporcionar la más alta disponibilidad, incluido el tiempo de inactividad no programado, la falta de experiencia, herramientas insuficientes, tecnologías complejas, la consolidación comercial y los mercados de la competencia. La administración de capacidad y rendimiento ayudan a los administradores de red a alcanzar nuevos objetivos comerciales mundiales y una disponibilidad y funcionamiento de la red constantes.

Este documento examina los temas siguientes:

- Capacidad general y problemas de rendimiento, incluyendo los riesgos y los problemas de la capacidad potencial dentro de las redes.
- Mejores prácticas de la Administración de capacidad y rendimiento, incluyendo los análisis del qué si, el baselining, tender, la administración de excepciones, y la Administración de QoS (Calidad de servicio).
- Cómo desarrollar una estrategia de planificación de capacidad, incluyendo las técnicas comunes, las herramientas, las variables MIB, y los umbrales usados en la planificación de capacidad.

Descripción de la Administración de capacidad y rendimiento

La planificación de capacidad es el proceso de determinar a los recursos de red requeridos prevenir un funcionamiento o un efecto en la disponibilidad en las aplicaciones comerciales cruciales. La Administración del rendimiento es la práctica manejo del tiempo de respuesta del servicio de red, del estado coherente, y de los servicios del calidad para el individuo y totales.

Note: Los problemas de rendimiento se relacionan generalmente con la capacidad. Las aplicaciones son más lentas porque el ancho de banda y los datos deben esperar en las colas de administración del tráfico antes de ser transmitido a través de la red. En las Aplicaciones de voz, los problemas como la fluctuación y retraso afectan directamente a la calidad de la llamada de voz.

La mayoría de las organizaciones recogen ya cierta información y trabajo capacidad-relacionados constantemente para solucionar los problemas, planean los cambios, y implementan la nueva capacidad y la funcionalidad de rendimiento. Sin embargo, las organizaciones no realizan rutinario los análisis el tender y del qué si. El análisis del qué si es el proceso de determinar la influencia de un cambio de la red. El tender es el proceso de la ejecución consiste las líneas de fondo de la capacidad de la red y los problemas de rendimiento y el repaso de las líneas de fondo para la red tiende para entender los requisitos para la actualización futuros. La Administración de capacidad y rendimiento debe también incluir la administración de excepciones donde se identifican los problemas y resuelto antes de los usuarios llame adentro, y la Administración de QoS (Calidad de servicio) donde los administradores de la red planean, manejan, e identifican los problemas de rendimiento del servicio individual. El gráfico siguiente ilustra los procesos de la Administración de capacidad y rendimiento.

La Administración de capacidad y rendimiento también tiene sus limitaciones, relacionadas típicamente con el CPU y la memoria. Los siguientes son áreas potenciales por preocupación:

- CPU
- Backplane o entrada-salida
- Memoria y buffers
- Tamaños de la interfaz y del tubo
- Datos en espera, tiempo de espera, y jitter
- Velocidad y distancia
- Características de la aplicación

Algunas referencias a la planificación de capacidad y a la Administración del rendimiento también mencionan algo llamado el “avión de los datos” y el “avión del control”. El avión de los datos es simplemente capacidad y problemas de rendimiento implicados con los datos que atraviesan la red mientras que el avión del control significa los recursos requeridos mantener la funcionalidad adecuada del avión de los datos. Las funciones planas del control incluyen los gastos indirectos del servicio tales como encaminamiento, atravesando - árbol, las señales de mantenimiento de la interfaz, y administración de SNMP del dispositivo. Estos requerimientos de planos del control utilizan el CPU, la memoria, mitigar, la espera, y el ancho de banda apenas como el tráfico que atraviesa la red. Muchos de los requerimientos de planos del control son también esenciales para la funcionalidad general del sistema. Si no tienen los recursos que necesitan, la red falla.

CPU

El CPU es utilizado típicamente por el avión del control y los datos acepillan en cualquier dispositivo de red. En Administración de capacidad y rendimiento, usted debe asegurarse de que el dispositivo y la red tengan suficiente CPU a funcionar siempre. La CPU insuficiente puede derrumbarse a menudo una red porque los recursos inadecuados en un dispositivo pueden afectar toda la red. La CPU insuficiente puede también aumentar el tiempo de espera puesto que los datos deben esperar para ser procesado cuando no hay Hardware Switching sin la CPU principal.

Backplane o entrada-salida

El backplane o la entrada-salida refiere a la cantidad total de tráfico que un dispositivo puede manejar, descrita generalmente en términos de tamaño del BUS o capacidad del panel de conexiones. Placa de interconexiones insuficiente da lugar normalmente a los paquetes perdidos, que pueden llevar a las retransmisiones y al tráfico adicional.

Memoria

La memoria es otro recurso que tiene requerimientos de planos del avión y del control de los datos. La memoria se requiere para la información tal como tablas de ruteo, tablas ARP, y otras estructuras de datos. Cuando los dispositivos funcionados con de la memoria, algunas operaciones en el dispositivo pueden fallar. La operación podía afectar a los procesos planos del control o a los procesos planos de los datos, dependiendo de la situación. Si los procesos planos del control fallan, toda la red puede degradar. Por ejemplo, esto puede suceder cuando memoria extra se requiere para la convergencia de la encaminamiento.

Tamaños de la interfaz y del tubo

Los tamaños de la interfaz y del tubo refieren a la cantidad de datos que se puedan enviar simultáneamente en cualquier una conexión. Esto a menudo se refiere incorrectamente como la velocidad de una conexión, pero los datos no viajan realmente a diversas velocidades a partir de

un dispositivo a otro. La velocidad del silicio y la ayuda de la capacidad del hardware determinan el ancho de banda disponible basado en los media. Además, los mecanismos del software pueden los datos de la “válvula reguladora” para ajustarse a las asignaciones de ancho de banda específicas para un servicio. Usted ve típicamente esto en las redes del proveedor de servicios para el Frame Relay o las atmósferas que intrínsecamente tienen capacidades de la velocidad de 1.54kpbs a 155mbps y más arriba. Cuando hay limitaciones de ancho de banda, los datos se hacen cola en una cola de transmisión. Una cola de transmisión puede tener diversos mecanismos del software para dar prioridad a los datos dentro de la cola; sin embargo, cuando hay datos en la cola, debe esperar los datos existentes antes de que pueda remitir a los datos hacia fuera la interfaz.

Datos en espera, tiempo de espera, y jitter

Los Datos en espera, tiempo de espera, y están inquietos también funcionamiento de la influencia. Usted puede ajustar la cola de transmisión para afectar al funcionamiento en las maneras diferentes. Por ejemplo, si la cola es grande, después los datos esperan más de largo. Cuando las colas de administración del tráfico son pequeñas, se caen los datos. Esto se llama taildrop y es aceptable para las aplicaciones TCP puesto que los datos serán retransmitidos. Sin embargo, la Voz y el vídeo no se realizan bien con la caída de la cola o aún la latencia de cola significativa que requiere la especial atención a los tamaños del ancho de banda o del tubo. El retardo de la cola puede también ocurrir con las colas de entrada si el dispositivo no tiene los recursos suficientes para remitir inmediatamente el paquete. Esto puede ser debido al CPU, a la memoria, o a los buffers.

El tiempo de espera describe el tiempo de procesamiento normal a partir del tiempo que se recibe hasta que el tiempo se remite el paquete. Los switches de datos modernos normales y el Routers tienen extremadamente - latencia baja (< 1ms) en condiciones normales sin las restricciones de recursos. Los dispositivos Modern con los procesadores de señales digitales a convertir y los paquetes de la voz analógica de la compresión pueden durar, incluso hasta 20ms.

El jitter describe la brecha entre paquetes para las aplicaciones de flujo continuo, incluyendo la Voz y el vídeo. Si los paquetes llegan los momentos diferentes con diversa sincronización de la brecha entre paquetes, después el jitter es alto y la Calidad de voz degrada. El jitter es principalmente un factor de retardo de colocación en cola.

Velocidad y distancia

La velocidad y la distancia es también un factor en el rendimiento de la red. Las redes de datos tienen una velocidad constante del reenvío de datos basada en la velocidad de la luz. Ésta es aproximadamente 100 millas por milisegundo. Si una organización está ejecutando una aplicación del servidor cliente internacionalmente, después pueden contar con un retardo correspondiente del reenvío de paquete. La velocidad y la distancia pueden ser un enorme factor en el rendimiento de la aplicación cuando las aplicaciones no se optimizan para el rendimiento de la red.

Características de la aplicación

Las características de la aplicación son el área más reciente que afecta a la capacidad y al funcionamiento. Los problemas tales como pequeños tamaños de la ventana, Keepalives de la aplicación, y la cantidad de datos enviados sobre la red contra se requiere qué pueden afectar al funcionamiento de una aplicación en muchos entornos, especialmente los WAN.

Mejores prácticas de la Administración de capacidad y rendimiento

Esta sección discute las cinco mejores prácticas principales de la Administración de capacidad y rendimiento detalladamente:

- [Service Level Management](#)
- [Análisis del qué si de la red y de la aplicación](#)
- [Baselining y el tender](#)
- [Administración de excepciones](#)
- [Administración de QoS \(Calidad de servicio\)](#)

Service Level Management

El Service Level Management define y regula otros procesos requeridos de la Administración de capacidad y rendimiento. Los administradores de la red entienden que necesitan la planificación de capacidad, pero hacen frente a presupuestar y a proveer de personal los apremios que previenen una solución completa. El Service Level Management es una metodología probada que las ayuda con los problemas del recurso definiendo una responsabilidad de dos partes entregable y que crea para un servicio atados a ese entregable. Usted puede lograr esto de dos maneras:

- Cree un Service Level Agreement entre los usuarios y la organización de la red para un servicio que incluya la Administración de capacidad y rendimiento. El servicio incluiría los informes y las recomendaciones de mantener la calidad del servicio. Sin embargo, los usuarios deben ser preparados para financiar el servicio y cualquier actualización requerida.
- La organización de la red define su servicio de Administración de capacidad y rendimiento y entonces intenta la financiación para ese servicio y la actualiza caso por caso.

En cualquier caso, la organización de la red debe comenzar definiendo una planificación de capacidad y un servicio de administración de rendimiento que incluya qué aspectos del servicio pueden proporcionar actualmente y qué se planea en el futuro. Un servicio completo incluiría un análisis del qué si para los cambios de la red y los cambios de la aplicación, el baselining y tender para las variables del rendimiento definidas, la administración de excepciones para la capacidad y las variables del rendimiento definidas, y la Administración de QoS (Calidad de servicio).

Análisis del qué si de la red y de la aplicación

Realice un análisis del qué si de la red y de la aplicación para determinar el resultado de un cambio planificado. Sin un análisis del qué si, las organizaciones llevan los riesgos significativos el cambio exitoso y la disponibilidad de la red total. En muchos casos, los cambios de la red han dado lugar al colapso congestivo que causaba a muchos el tiempo muerto de las horas de producción. Además, una cantidad sorprendente de introducciones de aplicación falla y causa el impacto a otros usuarios y aplicaciones. Estos errores continúan en muchas organizaciones de la red, con todo son totalmente evitables con algunas herramientas y algunos pasos de planificación adicionales.

Usted necesita normalmente algunos nuevos procesos realizar un análisis del qué si de la calidad. El primer paso es identificar los Niveles de riesgo para todos los cambios y requerir un análisis más profundizado del qué si para cambios con riesgo más altos. El Nivel de riesgo puede

ser un campo obligatorio para todas las presentaciones del cambio. Cambios más altos del Nivel de riesgo entonces requerirían un análisis definido del qué si del cambio. Un análisis del qué si de la red determina la influencia de los cambios de la red en los problemas del recurso de plano de control de la utilización de la red y de la red. Un análisis del qué si de la aplicación determinaría el éxito de la aplicación del proyecto, los requerimientos de ancho de banda, y cualquier problema de los recursos de red. Las tablas siguientes son ejemplos de los requerimientos de prueba de la asignación y de la correspondencia del Nivel de riesgo:

Nivel de riesgos	Definición	Cambie las recomendaciones de las hojas de operación (planning)
1	<ul style="list-style-type: none"> • Impacto de gran potencial al número grande de los usuarios (500+) o de servicio crucial para empresas debido al producto nuevo, al software, a la topología, o a la introducción de la característica. • El cambio implica el tiempo muerto previsto de la red. 	<ul style="list-style-type: none"> • Valide el laboratorio de la nueva solución. La Validación de Lab incluye la prueba de solución documentada y el análisis de la validación y del qué si que muestran el impacto a la infraestructura existente. Recomendamos a las soluciones piloto. Las nuevas soluciones requieren la realización de un documento de soporte de las operaciones. • Realice la revisión del diseño de Cisco NSA. • Cree el plan de retroceso. • Cree el plan de

		<p>instrumentación.</p> <ul style="list-style-type: none"> • Cree el proceso de cambio.
2	<ul style="list-style-type: none"> • Impacto de gran potencial al número grande de los usuarios (500+) o servicio crucial para empresas debido a un aumento grande del tráfico o usuarios, cambios de estructura básica, o cambios de ruteo. • El cambio puede requerir un cierto tiempo muerto. 	<ul style="list-style-type: none"> • Realice el análisis del qué si para determinar el impacto al entorno existente (debe ser hecho en el ambiente de laboratorio). • Cambios de ruteo de la prueba y del estudio para las funciones. • Cree el plan de retroceso. • Realice la revisión del diseño para el ruteo principal o los cambios de estructura básica. • Cree el plan de instrumentación. • Cree el proceso de cambio.
3	<ul style="list-style-type: none"> • Efecto potencial medio a un número más pequeño de usuarios o de servicio comercial debido a cualquier cambio no estándar. • Incluye el producto nuevo, software, topología, incorporación de características o los usuarios nuevos, tráfico 	<ul style="list-style-type: none"> • Realice la análisis de ingeniería de la nueva solución (puede requerir la Validación de Lab). • Cree el plan de instrumentación

	<p>aumentado, o topología no estándar.</p> <ul style="list-style-type: none"> • El cambio puede requerir un cierto tiempo muerto. 	<p>n.</p> <ul style="list-style-type: none"> • Cree el proceso de cambio.
4	<ul style="list-style-type: none"> • Baje el servicio potencial o el impacto en el usuario. • Incluye agregar los nuevos módulos de red de plantilla estándar, tales como edificio o switches del servidor/Hubs en el Routers. • Incluye que saca a colación los nuevos sitios PÁLIDOS o los servicios de acceso acreditado adicionales. • Todos los cambios del Nivel de riesgo 3 se han probado técnico en el entorno de producción. • El cambio puede requerir un cierto tiempo muerto. 	<ul style="list-style-type: none"> • Cree el plan de instrumentación. • Cree el proceso de cambio.
5	<ul style="list-style-type: none"> • Ningún impacto del usuario o del servicio. • Incluye agregar a los usuarios individuales a los cambios de la red y de configuración estándar tales como contraseña, banner, SNMP, u otros parámetros de configuración estándar. • Ningún tiempo muerto. 	<ul style="list-style-type: none"> • Proceso de cambio opcional.

Una vez que usted define donde usted necesita el análisis del qué si, usted puede definir el servicio.

Usted puede realizar un análisis del qué si de la red con las herramientas de modelado o con un laboratorio que imite el entorno de producción. Las herramientas de modelado son limitadas por como de bien la aplicación entiende los problemas del recurso del dispositivo y puesto que la mayoría de los cambios de la red son nuevos dispositivos, la aplicación puede no entender el efecto del cambio. El mejor método es construir una cierta representación de la red de producción en un laboratorio y probar el software deseado, la característica, el soporte físico, o la configuración bajo carga usando los generadores de tráfico. Escaparse las rutas (o la otra información de control) de la red de producción en el laboratorio también aumenta el ambiente de laboratorio. Pruebe los requisitos de recurso adicional con diversos tipos de tráfico, incluyendo el SNMP, el broadcast, el Multicast, cifrado, o el Tráfico comprimido. Con todas estas diversas

metodologías, analice los requisitos de recurso del dispositivo durante las situaciones potenciales de stress tales como convergencia de Routes, link inestable, y recomienzos del dispositivo. Los problemas de utilización de recursos incluyen las áreas normales de los recursos de capacidad tales como CPU, memoria, utilización de backplane, buffers, y Datos en espera.

Las nuevas aplicaciones deben también realizar un análisis del qué si para determinar el éxito de la aplicación y los requerimientos de ancho de banda. Usted realiza normalmente este análisis en un ambiente de laboratorio usando un analizador de protocolo y un Simulador de retraso WAN. para entender el efecto de la distancia. Usted necesita solamente un PC, un concentrador, un dispositivo de retraso WAN, y a un router del laboratorio conectado en la red de producción. Usted puede simular el ancho de banda en el laboratorio estrangulando el tráfico usando el Control de tráfico genérico o la limitación de la tarifa en el router de prueba. El administrador de la red puede trabajar conjuntamente con el grupo de aplicaciones para entender los requerimientos de ancho de banda, los problemas de la visualización en una ventana, y los problemas de rendimiento potencial para la aplicación en el LAN y los entornos WAN.

Realice un análisis del qué si de la aplicación antes de desplegar cualquier aplicación comercial. Si usted no hace esto, el grupo de aplicaciones culpa la red por el rendimiento pobre. Si usted puede requerir de alguna manera un análisis del qué si de la aplicación para las nuevas implementaciones vía el Proceso de administración de cambio, usted puede ayudar a prevenir los despliegues fallidos y a entender mejor los aumentos en el consumo de ancho de banda súbitos para el servidor del cliente y los requisitos del lote.

Baselining y el tender

El baselining y el tender permiten que los administradores de la red planeen y las actualizaciones de la red completas antes de que un problema de capacidad cause el tiempo muerto o los problemas de rendimiento de la red. Compare la utilización de recursos durante los períodos de tiempo sucesivos o destile la información abajo en un cierto plazo en una base de datos y permita que los planificadores vean los parámetros de la utilización de recursos por la hora, el día, la semana, el mes, y el año más pasados. En ambos casos, alguien debe revisar la información sobre un semanal, bisemanal, o la base mensual. El problema con el baselining y tender es que requiere una cantidad impresionante de información revisar en las Redes grandes.

Usted puede resolver este problema es varias maneras:

- El un montón de la estructura de capacidad y de transferencia en el entorno LAN así que la capacidad es no un problema.
- Divida la información de la tendencia en los grupos y concentre en la alta disponibilidad o las áreas críticas de la red, tales como sitios WAN críticos o centro de datos LAN.
- Los mecanismos de generación de informes pueden resaltar las áreas que bajan sobre cierto umbral que requiere especial atención. Si usted implementa las áreas de disponibilidad crítica primero, usted puede reducir perceptiblemente la cantidad de información requerida para el estudio.

Con todos los métodos anteriores, usted todavía necesita revisar la información sobre una forma periódica. El baselining y el tender es un esfuerzo proactivo y si la organización tiene solamente recursos para el soporte reactivo, los individuos no leerán los informes.

Muchas soluciones de administración de red proporcionan la información y los gráficos en las variables de los recursos de capacidad. Desafortunadamente, la mayoría de la gente utiliza solamente estas herramientas para el soporte reactivo a un problema existente; esto derrota el

propósito del baselining y de tender. Dos herramientas que son eficaces en proporcionar a la información de la tendencia de la capacidad para las redes de Cisco son el producto de las Integridades de la red Concord y los Productos INS EnterprisePRO. En muchos casos, las organizaciones de la red funcionan con los lenguajes de la secuencia de comandos simples para recoger la información de capacidad. Abajo están algunos informes de ejemplo que fueron recogidos vía el script para la utilización del vínculo, la utilización de la CPU, y el funcionamiento del ping. Otras variables de recurso que pueden ser importantes tender incluyen la memoria, la profundidad de espera en cola, el volumen de broadcast, el buffer, la notificación de congestión de Frame Relay, y la utilización de backplane. Refiera a éstos tabla para la información sobre la utilización del vínculo y la utilización de la CPU:

Utilización del vínculo

Recurso	Dirección	Segmento	Uso promedio (%)	Utilización pico (%)
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

Utilización de la CPU

Recurso	Dirección de sondeo	Uso promedio (%)	Utilización pico (%)
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

Utilización del vínculo

Recurso	Dirección	AvResT (ms) 09-09-98	AvResT (ms) 09-09-98	AvResT (ms) 09-09-98	AvResT (ms) 10-01-98
AADR01	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
APRR01	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
ASYR0					503.4

1S					
AZWRT 01	10.177.32 .1	460.1		444.7	
BEJR0 1	10.195.18 .1	1023.7	1064.6	1184	1021.9

Administración de excepciones

La administración de excepciones es una metodología valiosa por la capacidad y problemas de rendimiento de identificación y de resolución. La idea es recibir las infracciones de la notificación de capacidad y del umbral de rendimiento para investigar y reparar inmediatamente el problema. Por ejemplo, un administrador de la red pudo recibir una alarma para CPU elevada en un router. El administrador de la red puede registrar en el router para determinar porqué el CPU es tan alto. Ella puede entonces realizar una cierta configuración reparadora que reduzca el CPU o crea una lista de acceso que previene el tráfico que causa el problema, especialmente si el tráfico no aparece ser negocio crítico.

Usted puede configurar la administración de excepciones para más aspectos críticos bastante simplemente usando los comandos de configuración de RMON en un router o con herramientas más avanzadas tales como Servicio Netsys nivele al administrador conjuntamente con el SNMP, el RMON o los datos de NetFlow. La mayoría de las herramientas de administración de red tienen la capacidad para fijar los umbrales y las alarmas en las infracciones. El aspecto importante del proceso de la administración de excepciones es proporcionar cerca de la notificación del tiempo real del problema. Si no, el problema puede desaparecer antes de cualquier persona notó que la notificación fue recibida. Esto se puede hacer dentro de un NOC si la organización tiene supervisión constante. Si no, recomendamos la notificación del localizador.

El ejemplo de configuración siguiente proporciona la notificación del levantamiento y del umbral descendente para CPU del router a un archivo del registro que se pueda revisar sobre una base consistente. Usted puede configurar los comandos rmon similares para las Violaciones de umbral críticas de la utilización del vínculo u otros umbrales 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
```

Administración de QoS (Calidad de servicio)

La Administración de la calidad de servicio implica el crear y el monitorear de las clases de tráfico específicas dentro de la red. Un tráfico proporciona más rendimiento coherente para los grupos de aplicaciones específicos (definidos dentro de las clases de tráfico). Los parámetros de modelado del tráfico proporcionan la gran flexibilidad en el priorización y el formar EL tráfico para las clases de tráfico específicas. Estas características incluyen las capacidades tales como Committed Access Rate (CAR), Weighted Random Early Detection (WRED), y Datos en espera cargados feria basados clase. Las clases de tráfico se crean normalmente sobre la base del SLA de rendimiento para más aplicaciones críticas del negocio y de los requerimientos de la aplicación específicos tales como Voz. El tráfico no crítico o del NON-negocio también sería controlado de una manera que no puede afectar a las aplicaciones de mayor prioridad y a los servicios.

Crear las clases de tráfico requiere una comprensión de la línea de fondo de la utilización de la red, de los requerimientos de la aplicación específicos, y de las prioridades de la aplicación comercial. Los requerimientos de la aplicación incluyen el conocimiento de los tamaños de paquetes, de los problemas del descanso, de los requisitos del jitter, de los requisitos de la explosión, de los requisitos del lote, y de los problemas de rendimiento general. Con este conocimiento, los administradores de la red pueden crear los planes y las configuraciones del formar EL tráfico que proporcionan un rendimiento de la aplicación más constante a través de una variedad de topologías LAN/WAN.

Por ejemplo, una organización tiene una conexión ATM 10-megabit entre dos sitios importantes. El link a veces se congestiona de las transferencias de archivos grandes, que causa la degradación del rendimiento para el Transaction Processing en línea y la Calidad de voz pobre o inutilizable.

La organización configura cuatro diversas clases de tráfico. La Voz fue dada la prioridad más alta y permitida mantener esa prioridad incluso si repartió sobre la tarifa estimada del volumen de tráfico. La clase de la aplicación crítica fue dada la prioridad más alta siguiente pero no fue permitida repartir sobre el tamaño del link total menos los requisitos estimados del ancho de banda de voz. Cuando reparte, será caída. El tráfico de la transferencia de archivos fue dado simplemente una prioridad baja y el resto del tráfico cupo en alguna parte en el centro.

La organización ahora tiene que realizar la Administración de QoS (Calidad de servicio) en este link para determinar cuánto tráfico está tomando cada clase y para medir el funcionamiento dentro de cada clase. Si la organización no puede hacer esto, el hambre puede ocurrir para algunas clases o el SLA de rendimiento no se puede resolver dentro de una clase determinada.

El manejo de las configuraciones de QoS sigue siendo una tarea difícil debido a la falta de herramientas. Un método es utilizar al administrador del rendimiento de Internet de Cisco (IPM) para enviar diverso tráfico a través del link que cae en cada uno de las clases de tráfico. Usted podría entonces monitorea el funcionamiento para cada clase y el IPM proporciona tender, el análisis en tiempo real, y el análisis del salto por el salto para establecer claramente las áreas problemáticas. Otros pueden todavía confiar en un más método manual como la investigación de los Datos en espera y de los paquetes perdidos dentro de cada clase de tráfico basada en las estadísticas de la interfaz. En algunas organizaciones, estos datos se pueden recoger vía el SNMP o analizar en una base de datos para líneas de base y tender. Algunas herramientas también existen en el mercado que envían los tipos de tráfico específicos a través de la red para determinar el funcionamiento para un servicio determinado o una aplicación.

[Recogiendo y señalando la información de capacidad](#)

La recogida y señalar de la información de capacidad se deben conectar a las tres áreas recomendadas de la administración de capacidad:

- Análisis del qué si, que se centra alrededor del cambio de la red y cómo el cambio afecta al entorno
- Baselining y el tender
- Administración de excepciones

Dentro de cada uno de estas áreas, desarrolle un plan de recopilación de información. En el caso de los análisis del qué si de la red o de la aplicación, usted necesita las herramientas imitar el entorno de red y entender la influencia de los problemas potenciales en relación con del recurso del cambio dentro del avión del control de dispositivo o del avión de los datos. En el caso del

baselining y de tender, usted necesita las instantáneas para dispositivos y los links que muestran la utilización del recurso actual. Usted entonces revisa los datos en un cierto plazo para entender los requisitos para la actualización potenciales. Esto permite que los administradores de la red planeen correctamente las actualizaciones antes de que se presente la capacidad o los problemas de rendimiento. Cuando se presentan los problemas, usted necesita la administración de excepciones alertar a los administradores de la red así que pueden ajustar la red o reparar el problema.

Este proceso se puede dividir en los pasos siguientes:

1. Determine sus necesidades.
2. Defina un proceso.
3. Defina las áreas de capacidad.
4. Defina las variables de capacidad.
5. Interprete los datos.

[Determine sus necesidades](#)

Desarrollar un plan de Administración de capacidad y rendimiento requiere la comprensión de la información que usted necesita y el propósito de esa información. Parta el plan en tres áreas requeridas: uno por cada uno para el análisis del qué si, el baselining/el tender, y la administración de excepciones. Dentro de cada uno de estas áreas, descubra qué recursos y herramientas están disponibles y cuál es necesario. Mucho el fall de la organización con las implementaciones de la herramienta porque consideran la tecnología y las características de las herramientas pero no consideran la gente y la experiencia necesarias para manejar las herramientas. Incluya la gente y la experiencia en su plan, así como las mejoras del proceso requeridas. Esta gente puede incluir a los administradores de sistema para manejar las estaciones de administración de red, los administradores de la base de datos para ayudar con la Administración de base de datos, los administradores entrenados para utilizar y para monitorear las herramientas, y a los administradores de la red de alto nivel para determinar las directivas, los umbrales, y los requisitos de la recopilación de información.

[Defina un proceso](#)

Usted también necesita un proceso asegurarse de que la herramienta esté utilizada con éxito y constantemente. Usted puede requerir las mejoras del proceso para definir qué administradores de la red deben hacer cuando ocurren las Violaciones de umbral o qué proceso a seguir para el baselining, tendiendo, y actualizando la red. Una vez que usted determina los requisitos y los recursos para la planificación de capacidad acertada, usted puede considerar la metodología. Muchas organizaciones eligen externalizar este tipo de funciones a una organización de servicios de red tal como INS o construir la experiencia interna porque consideran el servicio una capacidad de base.

[Defina las áreas de capacidad](#)

El plan para la planificación de capacidad debe también incluir una definición de las áreas de capacidad. Éstas son áreas de la red que puede compartir una estrategia de planificación de capacidad común: por ejemplo, el LAN corporativo, las oficinas de campo PÁLIDAS, los sitios WAN críticos, y el acceso dial in. La definición de diversas áreas es útil por varias razones:

- Diversas áreas pueden tener diversos umbrales. Por ejemplo, el ancho de banda de LAN es

mucho más barato que el ancho de banda WAN así que los umbrales de utilización deben ser más bajos.

- Diversas áreas pueden requerir monitorear diversas variables MIB. Por ejemplo, los contadores FECN y BECN en el Frame Relay son críticos en los problemas de capacidad de comprensión del Frame Relay.
- Puede ser más difícil o largo actualizar algunas áreas de la red. Por ejemplo, los circuitos internacionales pueden tener plazos de ejecución mucho más largos y necesitar una correspondencia de alto nivel de las hojas de operación (planning).

[Defina las variables de capacidad](#)

La área importante siguiente está definiendo las variables para monitorear y los valores de umbral que requieren la acción. La definición de las variables de capacidad depende perceptiblemente de los dispositivos y de los media usados dentro de la red. En los parámetros generales tales como CPU, la memoria y la utilización del vínculo tienen valores. Sin embargo, otras áreas pueden ser importantes para las Tecnologías o los requisitos específicos. Éstos pueden incluir las profundidades de espera en cola, el funcionamiento, la notificación de congestión de Frame Relay, la utilización de backplane, la utilización del almacén intermedio, las estadísticas de Netflow, el volumen de broadcast, y los datos RMON. Tenga en cuenta sus planes a largo plazo, pero el comienzo con solamente algunas áreas claves de ayudar a asegurar el éxito.

[Interprete los datos](#)

La comprensión de los datos recogidos es también dominante a proporcionar un servicio de alta calidad. Por ejemplo, muchas organizaciones no entienden completamente los niveles del pico y del Uso promedio. El diagrama siguiente muestra un pico del parámetro de la capacidad basado en un intervalo de obtención minucioso de 5 SNMP, (mostrado en el verde).

Aunque el valor señalado era menos que los picos del umbral (mostrado en el rojo) todavía pueden ocurrir dentro del intervalo de obtención que están sobre el valor de umbral (mostrado en el azul). Esto es significativo porque durante el intervalo de obtención, la organización puede experimentar los valores pico que afectan al funcionamiento o a la capacidad de la red. Tenga cuidado de seleccionar un intervalo de obtención significativo que sea útil y que no cause los gastos indirectos excesivos.

Otro ejemplo es Uso promedio. Si los empleados están solamente en la oficina a partir del ocho a cinco, pero el Uso promedio es 7X24, la información puede ser engañosa.

[Información Relacionada](#)

- [Soporte Técnico - Cisco Systems](#)

¿Era este documento útil? [Sí](#) [ningún](#)

Gracias por su feedback.

[Abra un caso de soporte](#) (requiere un [contrato de servicios con Cisco](#).)

Discusiones relacionadas de la comunidad del soporte de Cisco

[La comunidad del soporte de Cisco](#) es un foro para que usted haga y conteste a las preguntas, las sugerencias de la parte, y colabora con sus pares.

Refiera a los [convenios de los consejos técnicos de Cisco](#) para la información sobre los convenios usados en este documento.

Actualizado: De oct el 04 de 2005

ID del Documento: 20769