*SNMP*

Puede utilizar SNMP para apagar una interfaz en el router o comprobar la velocidad a la que funciona su interfaz Ethernet. SNMP puede incluso controlar la temperatura de su interruptor y que le avise cuando ya es demasiado alta. SNMP por lo general se asocia con la gestión de los routers, pero es importante entender que se puede utilizar para manejar muchos tipos de dispositivos.

Un manager es un servidor que ejecuta algún tipo de sistema de software que puede manejar las tareas de gestión de una red. Los managers se refieren a menudo como estaciones de administración de red (SMN). Un NMS es responsable de sondeo y la recepción de las traps de los agents en la red. Un pull, en el contexto de la gestión de la red, es el acto de realizar queries a un agent (router, switch, servidor Unix, etc.) por alguna pieza de información. Esta información puede ser utilizada después para determinar si se ha producido algún tipo de evento catastrófico. Un trap es una forma para decirle al agent que algo ha ocurrido. Los traps se envían de forma asincrónica, no en respuesta a las preguntas de los NMS.

La segunda entidad, el agente, es una pieza de software que se ejecuta en los dispositivos de red que está administrando. Puede ser un programa independiente (un demonio, en lenguaje Unix), o puede ser incorporado en el sistema operativo. La NMS puede consultar el estado de cada interfaz y tomar las medidas apropiadas si alguno de ellos está caído. Cuando el agente se da cuenta de que algo

Ha ocurrido, se puede enviar una captura de los NMS. Esta trampa se origina en el agente y se envía a la NMS, donde se maneja apropiadamente. Algunos dispositivos envían una "fuera de peligro" trampa correspondiente cuando hay una transición de un mal estado a un buen estado.

Las polls y los traps pueden ocurrir al mismo tiempo.

La Estructura de Gestión de Información y MIB

La Estructura de Gestión de Información (SMI) ofrece una manera de definir objetos y su comportamiento.

La Base de Información de Gestión (MIB) puede ser pensado como una base de datos de objetos que gestiona lo que rastrea el agente. Cualquier tipo de estado o información estadística que se puede acceder por el NMS se define en una MIB. El SMI proporciona una manera de definir los objetos gestionados, mientras que el MIB es la definición (utilizando la sintaxis SMI) de los objetos mismos.

Un agente puede poner en práctica muchas MIB, pero todos los agentes implementan una MIB particular, llamado MIB-II.

Una breve introducción a RMON

El RMON MIB fue diseñado para permitir que RMON se ejecute en un modo sin conexión que permitiendo reunir datos estadísticos sobre la red que está monitoreando sin necesidad de un NMS para consultar constantemente.

El concepto de Network Management

Un modelo para la gestión de la red llamado FCAPS, o gestión de fallos, gestión de configuración, gestión de contabilidad, gestión del rendimiento y de administración de seguridad.

* Gestión de fallos: El objetivo de la gestión de fallos es detectar, registrar y notificar a los usuarios de sistemas o redes de problemas. gestión de fallos indica que se seguirán los siguientes pasos para la resolución de fallos:

1. Aislar el problema usando herramientas que determinen los síntomas.
2. Resolver el problema.
3. Registrar el proceso que fue utilizado para detectar y resolver el problema.

* Gestión de la Configuración: El objetivo de la gestión de la configuración es monitorear la red y la configuración del sistema, de manera que los efectos sobre el funcionamiento en red de varias versiones de hardware y software pueden ser rastreados y gestionados. Una serie de parámetros de configuración interesantes pueden estar interesados en la captura, incluyendo la versión del sistema operativo, el firmware, etc; Número de interfaces y velocidades de red, etc; Número de discos duros; Número de CPU; Cantidad de RAM.
* Gestión contable: El objetivo de la gestión de la contabilidad es asegurar que los recursos informáticos y de red se utilizan justamente por todos los grupos o individuos que acceden a ellos. A través de esta forma de regulación, problemas de red pueden minimizarse ya que los recursos se dividen en función de las capacidades.
* Gestión del Rendimiento: El objetivo de la gestión del rendimiento es medir e informar sobre diversos aspectos de la red o sistema de rendimiento.
  1. Registro de datos sobre el rendimiento.
  2. Los niveles de referencia se establecen con base en el análisis de los datos recogidos.
  3. Se establecen los umbrales de rendimiento. Cuando se superen estos umbrales, es indicativo de un problema que requiere atención.
* Gestión de la Seguridad: El objetivo de la gestión de la seguridad es doble. En primer lugar, deseamos controlar el acceso a algún recurso, tal como una red y sus anfitriones. En segundo lugar, deseamos ayudar a detectar y prevenir los ataques que pueden comprometer redes y hosts.
* Gestión de Seguridad: La gestión de seguridad abarca los sistemas de seguridad de red no sólo física, sino también la seguridad. La seguridad física incluye el acceso de tarjetas y sistemas de video vigilancia. El objetivo aquí es asegurar que sólo las personas autorizadas tengan acceso físico a los sistemas vulnerables. Varias herramientas y sistemas diseñados específicamente para este propósito. Estos incluyen firewalls; Sistemas de Detección de Intrusos (IDS); sistemas antivirus; La administración de políticas y sistemas de aplicación.

SNMP and UDP

SNMP utiliza el Protocolo de datagramas de usuario (UDP) como protocolo de transporte para el paso datos entre los gestores y agentes. UDP, fue elegido por el Protocolo de Control de Transmisión (TCP), ya que es sin conexión; es decir, hay un extremo a otro de conexión se hace entre el agente y el NMS cuando datagramas (paquetes) se envían de ida y vuelta. Este aspecto de la UDP hace que sea poco fiable ya que no hay reconocimiento de datagramas perdidos al nivel de protocolo. Todo depende de la aplicación SNMP para determinar si se pierden los datagramas y las retransmiten si así lo desea.

SNMP utiliza el puerto UDP 161 para el envío y la recepción de solicitudes y el puerto 162 para la recepción de las traps de los dispositivos gestionados.

Cuando ya sea un NMS o un agente desea realizar una función SNMP (por ejemplo, una solicitud o trampa), los siguientes eventos ocurren en la pila de protocolos:

* Aplicación: En primer lugar, la aplicación real de SNMP (NMS o agente) decide lo que va a hacer. La capa de aplicación proporciona servicios a un usuario final, tal como una información de estado de operador solicitante para un puerto en un swicth Ethernet.
* UDP: La siguiente capa, UDP, permite que dos hosts se puedan comunicar unos con otros. La cabecera UDP contiene, entre otras cosas, el puerto de destino del dispositivo al que está enviando la solicitud o trampa.
* IP: La capa IP intenta entregar el paquete SNMP a su destino previsto, tal como se especifica por su dirección IP.
* Media Access Control (MAC): La capa MAC se compone de los controladores de hardware y dispositivos reales que ponen sus datos en una pieza física de cable, tal como una tarjeta Ethernet. La capa MAC es también responsable de la recepción de paquetes de la red física y el envío de vuelta hasta la pila de protocolos de modo que puedan ser procesados por la capa de aplicación (SNMP, en este caso).

SNMP Communities

Un agente se configura con tres nombres de comunidad: sólo lectura, lectura y escritura, y la trampa. Los nombres de comunidad son esencialmente las contraseñas. La de comunidad de sólo lectura se le permite leer los valores de datos, pero no le permite modificarlos. Se le permite a la comunidad de lectura y escritura, leer y modificar los valores de los datos. Por último, la comunidad de captura podrá recibir notificaciones de trampas (asíncronos) desde el agente. Típicamente la comunidad de sólo lectura es pública y la comunidad de lectura-escritura es privada.

Además, dado que las comunidades SNMP se envían en texto plano, se puede configurar un agente para enviar una captura de autenticación de fallos SNMP cuando alguien intenta consultar con su dispositivo con la comunidad incorrecto.

SNMP Operations

La operación Get: La petición GET es iniciada por el NMS, que envía la solicitud al agente. El agente recibe la solicitud y la procesa. Si el agente tiene éxito en la recopilación de la información solicitada, se envía un GetResponse al SMN.

Uno de los elementos en la solicitud GET es una variable binding. Una variable de unión, o var-bind, es una lista de objetos MIB que permite a los destinatarios de una solicitud ver lo que el autor quiere saber.

La Operación GetBulk: la operación GetBulk, permite a una aplicación de gestión para recuperar una gran parte de una tabla a la vez. La operación GetBulk, por otro lado, le dice al agente que envíe de vuelta como respuesta todo lo que puede. Esto significa que las respuestas incompletas son posibles.

Set Operation: El comando set se utiliza para cambiar el valor de un objeto gestionado o para crear una nueva fila en una tabla. Los objetos que se definen en la MIB como lectura-escritura o de lectura, sólo puede ser modificado o creado con este comando. Es posible para un NMS configurar más de un objeto a la vez.

Métricas de medición

Delay (Retardo): Delay se puede cuantificar, ya sea como un solo sentido de demora o delay de ida y vuelta como (tiempo de ida y vuelta o RTT). En este caso, el RTT se puede determinar si el remitente indica su timestamp cuando se los envía (el timestamp se realiza en los datos del paquete de prueba) y resta este valor del timestamp correspondiente cuando recibe la respuesta de prueba. La medición de delay en un sentido requiere que el remitente y el receptor tengan sincronizados los relojes locales de tal manera que el delay unidireccional pueda ser determinado en el receptor, si el receptor también recibe el timestamp en el paquete de prueba; la diferencia entre el timestamp del envío y el de recepción es el delay unidireccional. RTT es más fácil de implementar y mide el delay unidireccional

Para aplicaciones como la videoconferencia interactiva puede no importar en qué dirección el delay es experimentado; si el exceso es experimentado en absoluto, entonces el servicio se verá afectado. Si se producen violaciones en el SLA por delay, sin embargo, el RTT oculta los detalles de la dirección en la que se produjo el problema que causa la violación. El delay puede proporcionar un número de indicadores importantes del rendimiento de la red.

* El retardo mínimo. El retardo de red mínimo es el retardo de "línea de base", de red que proporciona una indicación del retraso que experimentará el tráfico cuando la ruta del origen al destino está cargada ligeramente. Este se compone de retardo de propagación, retardo de conmutación, y el retardo de serialización. Los valores de retraso por encima del mínimo proporcionan una indicación de la congestión experimentada a lo largo del camino.
* Umbral superado: puede ser útil para contar el número de paquetes de prueba, de un total que experimentó un retraso de más de un umbral definido, establecido para indicar cuando un paquete llegó demasiado tarde para ser útil.
* Delay promedio: El retardo promedio puede ser interesante para propósitos de tendencia, pero para fines de comparación deben registrarse junto con la desviación estándar de la muestra; las mayores desviaciones estándar de lo normal pueden ser indicativos de problemas falsos más que de una tendencia.
* Delay-Jitter: se considera generalmente que es la variación del retardo de ida para dos paquetes consecutivos. Afortunadamente, para el cálculo del jitter no hay necesidad de conocer los retrasos individuales de un solo sentido: en su lugar, esto se puede calcular a partir de la diferencia entre marcas de tiempo tomadas en los dispositivos individuales. Las estadísticas más importantes que informen con respecto a la jitter son el jitter percentil elevado, umbral superado, y jitter promedio.
* Pérdida de paquetes: Con el fin de determinar la pérdida de paquetes es necesario que haya una manera de distinguir entre un paquete perdido y un paquete con un retardo grande pero finito. En métricas adicionales, las cuales describen patrones de pérdida y se pueden utilizar para analizar el posible impacto en las aplicaciones:

1. período de pérdida. El período de pérdida define la frecuencia y la longitud (pérdida de ráfaga) de la pérdida una vez que comienza.
2. distancia pérdida. La distancia pérdida define la separación entre los periodos de pérdida.

* Ancho de banda y Throughput: La aplicación de throughput depende de muchos factores, que pueden variar ampliamente dependiendo de implementaciones de sistema de extremo y los perfiles de tráfico.
* Disponibilidad: La disponibilidad de servicios IP se define en general, ya sea como disponibilidad de la red o como la disponibilidad del servicio.
  1. Disponibilidad de la red: disponibilidad de la red bidireccional o conectividad entre dos dispositivos de monitorización activos se pueden determinar usando pruebas de envío desde un emisor a un receptor y luego de vuelta al remitente; por cada respuesta recibida con éxito la red se considera disponible y no para cada recibido la red se considera no disponible.
  2. Disponibilidad del servicio: Es un indicador compuesto definido cuando un servicio está disponible entre un punto de entrada especificado y un punto de salida especificada dentro de los límites de las métricas de SLA comprometidos para el servicio.
* Calidad de Experiencia: "calidad de la experiencia" o QOE. La medida QOE más común es la "puntuación media de opinión" o MOS, que proporciona una medida numérica subjetiva del QOE de una llamada de voz.

La sincronización de reloj

La forma más precisa para sincronizar los relojes de los dispositivos de red consiste en sincronizar cada dispositivo con una "capa 1" fuente de reloj externa precisa como un reloj GPS o radio reloj. No sería viable para el usuario final o ubicaciones de sucursales pequeñas. Un enfoque alternativo es la de distribuir el tiempo de estrato 1 usando un protocolo, tal como el protocolo de tiempo de red (NTP). NTP sincroniza los relojes entre dispositivos de red mediante el intercambio de mensajes con timestamp entre un servidor y sus clientes. NTP busca precisión a largo plazo a expensas de la exactitud de corto plazo; lo hará, por ejemplo, disminuir o acelerar el reloj interno para ajustar el reloj local progresivamente a lo que cree que es el verdadero tiempo.

Debido a las limitaciones y los costos de entre dispositivos de sincronización de reloj, un modelo de implementación común es la de distribuir el tiempo a partir de una fuente de reloj de capa 1 a todos los dispositivos dentro de un punto de presencia (POP). La sincronización de los routers de acceso a través de NTP en general no es lo suficientemente precisa y el uso de capa 1 fuentes de reloj en estos lugares generalmente no es viable, por lo tanto, los informes de SLA de los enlaces de acceso de POP de acceso del router es comúnmente reportado como RTT en lugar de un solo sentido demora.

¿Cuál es la calidad de servicio?

Consideramos que un servicio es una descripción del tratamiento global del tráfico de un cliente a través de un dominio particular.

Podemos definir la calidad de servicio en términos de los requisitos fundamentales para una aplicación, la cual se puede definir en términos de las métricas de SLA para el desempeño del servicio IP.

Calidad de servicio, implica algo más que asegurar que un servicio de red sea capaz de soportar los requisitos de SLA de las aplicaciones que tiene el objetivo de apoyar.

A un alto nivel de calidad de servicio que podemos describir en términos de los objetivos que se está tratando de lograr, que definen efectivamente un problema de optimización de tratar de maximizar la satisfacción del usuario final (utilidad o eficacia) y reducir al mínimo costo. Maximizar la satisfacción del usuario requiere que las aplicaciones de usuario final trabajar con eficacia. Se minimizan los costos requiere que no sobre diseñar la red con el fin de entregar la calidad de la llamada, lo que puede requerir la necesidad de diferenciar los niveles de servicio ofrecidos a diferentes aplicaciones.

Servicio de mejor esfuerzo

"Mejor esfuerzo" describe un servicio de red que intenten entregar tráfico a su destino, pero que no ofrece ninguna garantía de entrega, y por lo tanto es, sin compromiso alguno de retardo, jitter, pérdida, y el rendimiento. Donde una red soporta múltiples clases de servicio al mismo tiempo, mejor esfuerzo se utiliza a menudo para referirse a los servicios que ofrece a los compromisos de SLA más bajas.

El conjunto de herramientas de calidad de servicio

Plano de datos: los mecanismos de plano de datos de calidad de servicio se aplican en los nodos de la red y pueden afectar directamente el comportamiento de reenvío de paquetes. Estos mecanismos se pueden clasificar en función de las características de comportamiento primitivas que se imparten al tráfico:

* Clasificación: es el proceso de clasificación de un flujo de tráfico agregado en un número de clases constituyentes.
* Marcado: la demarcación de tráfico es el proceso de establecer explícitamente el valor de los campos asignados para la clasificación de calidad de servicio en los encabezados de paquete IP o MPLS para que el tráfico posteriormente se puede identificar fácilmente.
* Aplicación de la tasa máxima: política utilizado para hacer cumplir una tasa máxima para una clase de tráfico.
* Priorización: algunos tipos de tráfico sobre los demás logrando así retardo y jitter.
* Aseguramiento mínima tasa de: técnicas de planificación que pueden ser utilizadas para proporcionar diferentes clases de tráfico con diferentes garantías mínimas de ancho de banda.
* Plano de control: mecanismos que normalmente se ocupan del control de admisión y reserva de recursos, se implementan típicamente como procesos de software, tales como los protocolos de enrutamiento.

**Policing and Metering**

Policing es un mecanismo que puede ser usado para asegurar que un flujo de tráfico no exceda una velocidad máxima definida.

Una simple implementación es one-rate token bucket que tiene una profundidad definida máxima (normalmente en bytes), conocido como la ráfaga B, y una tasa R (normalmente en bps) a la que el cubo está llenado de fichas de tokens de bytes.

Cuando un mecanismo de token bucket se aplica a un flujo de tráfico, y un paquete de esa corriente llega, el tamaño de paquete b se compara con el número de fichas actualmente en el bucket. Si hay al menos tantas fichas de bytes en el cubo como hay en el paquete, entonces usamos la terminología que el paquete ha sido "conforme"; Si hay menos fichas en el cubo que bytes en el paquete, entonces el paquete ha "superado" la cantidad de fichas del cubo. Si el paquete se ajusta a continuación, entonces se decrementa de la cubeta la misma cantidad de fichas que el tamaño del paquete.

Las acciones más sencillas son transmitir el paquete si se ajusta y descartar el paquete si es superior; aplicado de esta forma la política de token bucket aplicaría una tasa máxima de R y B de ráfagas en el flujo de tráfico.

Es importante señalar que un token bucket policer nunca retrasa el tráfico, mientras que si lo hace un shaper. Es decir, no hay paquetes almacenados en el cubo; sólo hay fichas en el cubo. Por lo tanto, como un token bucket no retrasa el tráfico, no puede volver a ordenar o priorizar el tráfico como un planificador.

Single Rate Three Color Marker

Los "tres colores" se refieren a los tres estados posibles que son resultados de la SR-MTC, y que se describen usando una combinación de colores "semáforo".

SR-TCM utiliza dos cubetas de fichas. Los cubos se definen como C y E (para comprometida y en exceso, o se conforman y exceden) con el tamaño máximo de ráfaga CBS y EBS, respectivamente. Ambos cubos están llenos de tokens. Cuando el SR-TCM se aplica a un flujo de tráfico y un paquete de esa corriente llega:

* El tamaño del paquete B se compara con el número de fichas actualmente en el cubo C. Si hay al menos tantas fichas como en el cubo C como el tamaño del paquete, entonces el paquete se ha ajustado a la definición SR-MTC y el cubo C sólo es decrementa en tokens igual al número de bytes en el paquete. Usando el "semáforo" de color verde.
* sino, si no hay tantas fichas en el cubo C como bytes en el paquete, entonces el paquete ha superado el cubo C de la definición SR-TCM y ahora se compara con el cubo E. Si hay al menos tantas fichas de bytes en el cubo E como hay en el paquete, entonces el cubo de E solamente se decrementa en tokens igual al número de bytes en el paquete. Usando el tráfico de color amarillo claro.
* Si no hubiera ni tantas fichas en el cubo C o E como bytes en el paquete, entonces usamos la terminología que el paquete ha "violado" la política de SR-TCM. Se utiliza el color de la luz roja del semáforo y no se disminuye ninguna cubeta

**Two Rate Three Color Marker**

TR-TCM utiliza dos cubos de fichas; sin embargo, una diferencia notable con respecto al caso anterior, los cubos se llenan a diferentes velocidades. C se llena a la tasa de información comprometida (CIR), mientras que P se llena al tipo de información pico (PIR).

Cuando el TR-TCM se aplica a un flujo de tráfico y un paquete de esa corriente llega:

