

# Contenido

[Introducción](#)

[prerrequisitos](#)

[Requisitos](#)

[Componentes Utilizados](#)

[Antecedentes](#)

[Mecanismo del trabajo](#)

[Colas de administración del tráfico AAAMGR](#)

[Limitaciones](#)

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

## Introducción

Este documento describe estrangular de la característica de los expedientes AAA (RADIUS) que soporta estrangular del acceso (autenticación y autorización) y de los registros de contabilidad que se envía al servidor de RADIUS.

Esta característica permite que un usuario configure la tarifa que estrangula apropiada para evitar la congestión de red y la inestabilidad cuando hay ancho de banda escaso para acomodar un suddenburst de los expedientes generados del router Cisco al servidor de RADIUS.

## Prerrequisitos

### Requisitos

No hay requisitos específicos para este documento.

### Componentes Utilizados

La información en este documento se basa en el platform ASR5k.

La información que contiene este documento se creó a partir de los dispositivos en un ambiente de laboratorio específico. Todos los dispositivos que se utilizan en este documento se pusieron en funcionamiento con una configuración verificada (predeterminada). Si la red está funcionando, asegúrese de haber comprendido el impacto que puede tener cualquier comando.

## Antecedentes

Cuando el aaamgr envía los mensajes de RADIUS al servidor de RADIUS a una alta velocidad (e.g cuando van un gran número de sesiones abajo al mismo tiempo, los mensajes de la parada que consideran para todo el las sesiones se generan al mismo tiempo) que el servidor de RADIUS puede no poder recibir los mensajes a tales altas velocidades. Para manejar esta condición necesitamos un mecanismo de control de velocidad eficaz en el aaamgr, de modo que el aaamgr envíe los mensajes a una tarifa óptima tales que el servidor de RADIUS es capaz de recibir todos los mensajes y se asegure de que no hay mensaje caído debido a la carga excesiva

en el servidor de RADIUS.

## Mecanismo del trabajo

Cuando el aaamgr envía los mensajes a la velocidad configurada al servidor de RADIUS, envía los mensajes uniformemente en todas partes cada segundo bastante que enviando todos los mensajes en una sola explosión. Dependiendo de la configuración, cada segundo se divide en varios slots de tiempo iguales (con un período específico por el slot). El período de tiempo mínimo de un slot podían ser 50 milisegundos.

La tarifa tiene que ser configurada tomando en cuenta éstos

- El índice de llamadas entrantes,
- Número de casos del aaamgr
- La tarifa en la cual el servidor de RADIUS puede recibir los mensajes y
- Intervalo del interims (para la configuración que considera)
- Algoritmo usado para seleccionar el servidor

Si el valor configurado para el valor de los servidores de autenticación es demasiado bajo, después habrá un cuello de la botella que lleva a congestión, que puede llevar a las llamadas que consiguen caída debido al descanso de la configuración de la sesión. Si un valor bajo se configura para los servidores de contabilidad, después mucha purga de los mensajes de las estadísticas será observada, debido desbordar de la cola.

Cuando se configura la característica, los slots de la cantidad de veces en un segundo y un período de tiempo de un segundo se computan y se salvan en el nivel del radio. Cuando un mensaje está listo para ser enviado al servidor de RADIUS, se marca si la cuota (número de mensajes para este slot de tiempo) ha alcanzado. Si el límite no se alcanza, se envía el mensaje, si es, después el mensaje se hace cola en la cola del nivel del servidor que se enviará en los slots de tiempo futuros. Cada servidor de RADIUS lleva a cabo los detalles sobre el número de mensajes enviados en el slot de hora actual y el tiempo en los cuales el slot de tiempo expira. Cuando los mensajes hechos cola se escogen de la cola del nivel del servidor, se ponen en el jefe de la cola del nivel del caso, asegurando la preferencia por los mensajes más antiguos que cualquier otro nuevo mensaje. Los mensajes de la cola del nivel del caso se escogen para mantener.

## AAAMGR hace cola

Hay dos tipos de colas de administración del tráfico en AAAMGR para los mensajes:

1. Colas de administración del tráfico llanas del caso
2. Colas de administración del tráfico llanas del servidor

Cuando se genera un mensaje, se hace cola inicialmente en la cola del nivel del caso para mantener.

La cola del nivel del caso se procesa por 25 milisegundos por cada 50 milisegundos. Cualquier mensaje que dequeued de la cola del nivel del caso será intentado para ser enviado al servidor de

RADIUS. Bajo algunas condiciones podemos no poder enviar los mensajes (ningún ancho de banda disponible o ningunos ID disponibles). En estos casos, los mensajes que fallaron la tentativa serán hechos cola en las colas de administración del tráfico del nivel del servidor. Para cada 50 milisegundos usted escoge tantos mensajes que tengan ID disponibles y también el ancho de banda disponible y los pone en el jefe de la cola del nivel del caso (estos mensajes son más viejos que cualquier otro mensaje que esté presente en la cola del nivel del caso).

Cuando hay un control de velocidad para los mensajes que consideran, y si hay muchos mensajes de las estadísticas en la cola del nivel del caso, después cualquier nuevo mensaje de autenticación va a la cola de la cola del nivel del caso. Para conseguir procesado le tiene que esperar todo el mensaje de las estadísticas (que precede el nuevo mensaje auth) para ser enviado al servidor de RADIUS o para ser movido a la cola del nivel del servidor. Es un comportamiento existente y no se modifica. Puede hacer tan un pequeño retardo para el nuevo mensaje auth conseguir procesado.

### Ejemplo:

De acuerdo con la MAX-tarifa con el valor de 5, usted puede enviar cinco mensajes en 1 segundo y tener por contestar excepcional de los mensajes de la autenticación de RADIUS 256 (configuración MAX-excepcional predeterminada) por el aaamgr hacia el servidor de autenticación de RADIUS. En caso de que haya más de 5 mensajes, en 1 segundo los mensajes se ponen en la cola hasta que el servidor de AAA responde a las peticiones existentes.

En caso de que usted alcance los mensajes de la autenticación de RADIUS 256 enviados a partir de un aaamgr hacia el servidor, que siguen habiendo las peticiones serán puestas en la cola hasta que el servidor de AAA responde a las peticiones existentes. Entrará otra vez la misma cola que la de la MAX-tarifa. El mensaje se coge de la cola solamente cuando usted tiene un slot libre. El slot libre viene en cuando usted recibe una respuesta para el mensaje o cuando él los time-out.

## Limitaciones

Puesto que es Cisco ASR5K un sistema distribuido con sessmgr/aaamgr independiente empareja el proceso de las llamadas, el estrangular de la tarifa se podría implementar solamente para los casos independientes del aaamgr. Es teórico ampliar el índice de una instancia única al cuadro entero de Cisco ASR5K en conjunto apenas multiplicando el número total de casos con la MAX-tarifa de cada caso.

Este número es apenas el límite superior absoluto en un escenario del día soleado. Usted no puede tratar Cisco ASR5K como caja negra y no puede asumir que todas las llamadas deben tener éxito si el valor calculado considerado en el sistema no cruzó el límite superior.

La max-tarifa del radio se ata con otros parámetros internos y externos relacionados con el sistema. Vea por favor el impacto previsto si una de las condiciones no se cumple.

### Condiciones

Distribución uniforme de las llamadas del demuxmgr a todos los sessmgrs

Distribución uniforme de IMSI (esto está sólo

### Impacto si no hecho frente con

Si la distribución de llamadas no es uniforme, después los mensajes de RADIUS pueden hágase cola para algunos casos. Tan aunque el límite teórico de la MAX-tarifa no se alcanza, las llamadas serán caídas los casos donde se hacen cola los mensajes.

las Mediación-estadísticas circulares se basan en la

en caso de las estadísticas circulares de la mediación)

Ningunas explosiones súbitas de las llamadas venir-en

Los servidores de RADIUS deben responder a tiempo

encaminamiento IMSI-basada.

En este caso, sobre la base de la distribución IMSI, un conjunto de los servidores se puede preferir sobre otros basados en la lógica de la encaminamiento, cola pudo ser urbanizado para esos servidores que llevan el llamada-descenso.

Si hay una explosión de las nuevas llamadas, después los mensajes de RADIUS nuevamente generados serán hechos cola otra vez en el sistema. Para el momento en que se procesen los nuevos pedidos de RADIUS. El tiempo de configuración de la sesión puede ser el llevar expirado para llamar los descensos.

¿Cuando time-out de los pedidos de RADIUS debido a los problemas de servidor, allí? el ll sea otra vez acumulación la cola, porque las nuevas peticiones no serán enviadas a menos que la actual que cuenta con una respuesta se que el sistema. La tarifa en la cual los mensajes sincronizados-h fuera serán quitados del sistema es dependiente en las configuraciones MAX-excepcionales y time-out también.

En muchos casos podemos ver que las peticiones del acceso no son procesadas por todas las tareas activas del aaamgr. Eso significa que estamos teniendo distribución de llamadas desigual dentro de las tareas del sessmgr y más futuro encendido, no todos los casos del aaamgr están implicados en el Procesamiento de Llamadas.

La distribución de llamadas no se basa en el mecanismo estricto del ordenamiento cíclico que es si hay 10 llamadas entrantes que irán a 10 sessmgrs en un algoritmo monotónico.

La distribución de llamadas se basa en estos cuatro factores principales

- **active\_session\_count**
- **cpu\_load**
- **¿Round\_trip\_delay** (demuxmgr? ¿sessmgr? demuxmgr)
- **el más outstanding\_add\_request** (demux al sessmgr)

Ésta es la implementación actual. ¿La max-tarifa es apenas un límite superior, pero debido a la naturaleza distribuida de nuestra arquitectura, usted puede? t la extrapola directamente a la carga del chasis. El comportamiento depende de la carga en un AAAMgr dado en un momento dado.

La cola de la MAX-tarifa del radio se debe utilizar **para monitorear el estatus del sistema**. Si hay una **acumulación de la cola**, entonces significa que una de estas 4 (refiera a la tabla) condiciones no está cumplido con y usted debe identificar la causa raíz para lo mismo.

\*\* el umbral de la cola de la MAX-tarifa podía ser implementado y ser monitoreado constantemente.