

# Índice

[Introdução](#)

[Pré-requisitos](#)

[Requisitos](#)

[Componentes Utilizados](#)

[Informações de Apoio](#)

[Mecanismo do trabalho](#)

[Filas AAAMGR](#)

[Limitações](#)

[Cisco relacionado apoia discussões da comunidade](#)

## Introdução

Este documento descreve o estrangulamento da característica dos registros AAA (RAIO) que apoia o estrangulamento do acesso (authentication e autorização) e dos registros de contabilidade que é enviada ao servidor Radius.

Esta característica permite que um usuário configure a taxa de estrangulamento apropriada para evitar o congestionamento de rede e a instabilidade quando há uma insuficiente largura de banda para acomodar um suddenburst dos registros gerados do roteador Cisco ao servidor Radius.

## Pré-requisitos

### Requisitos

Não existem requisitos específicos para este documento.

### Componentes Utilizados

A informação neste documento é baseada no platfrom ASR5k.

As informações neste documento foram criadas a partir de dispositivos em um ambiente de laboratório específico. Todos os dispositivos utilizados neste documento foram iniciados com uma configuração (padrão) inicial. Se a sua rede estiver ativa, certifique-se de que entende o impacto potencial de qualquer comando.

## Informações de Apoio

Quando o aaamgr envia as mensagens do raio ao servidor Radius em uma taxa alta (por exemplo quando um grande número sessões vão para baixo ao mesmo tempo, mensagens explicando da parada para o todo o as sessões são geradas ao mesmo tempo) que o servidor Radius não pode receber as mensagens em tais taxas altas. Para segurar esta circunstância nós precisamos um mecanismo de controle de taxa eficaz no aaamgr, de modo que o aaamgr envie a mensagens em uma taxa ótima tais que o servidor Radius é capaz de receber todas as mensagens e se assegure de que nenhuma mensagem seja deixado cair devido à carga excedente no servidor

Radius.

## Mecanismo do trabalho

Quando o aaamgr envia mensagens na taxa configurada ao servidor Radius, envia mensagens uniformemente por toda parte cada segundo um pouco do que enviando todas as mensagens em uma única explosão. Segundo a configuração, cada segundo é dividido em diversos timeslot iguais (com um período de tempo específico pelo entalhe). O período de tempo mínimo de um entalhe podiam ser os milissegundos dos 50 pés.

A taxa tem que ser configurada levando em conta estes

- A taxa de chamadas recebidas,
- Número de exemplos do aaamgr
- A taxa em que o servidor Radius pode receber as mensagens e
- Intervalo dos temporários (para a configuração explicando)
- Algoritmo usado para a seleção de servidor

Se o valor configurado para o valor dos Authentication Server é demasiado baixo, a seguir haverá um pescoço da garrafa que conduz a congestão, que pode conduzir aos atendimentos que obtêm deixado cair devido ao intervalo da instalação da sessão. Se um valor baixo é configurado para servidores de contabilidade, a seguir muita remoção de mensagens da contabilidade estará observada, devido transbordar da fila.

Quando a característica é configurada, os entalhes do número de vezes num segundo e o período de tempo de um segundo estão computados e armazenados a nível do raio. Quando uma mensagem está pronta para ser enviado ao servidor Radius, verifica-se se a quota (número de mensagens para este timeslot) alcance. Se o limite não é alcançado, a mensagem está enviada, se é, a seguir a mensagem está enfileirada na fila do nível do server a ser enviada nos timeslot futuros. Cada servidor Radius guarda detalhes sobre o número de mensagens enviadas no entalhe de horas atual e no tempo em que o timeslot expira. Quando as mensagens enfileiradas são escolhidas da fila do nível do server, estão postas na cabeça da fila do nível do exemplo, assegurando a preferência para mensagens mais antiga do que toda a outra mensagem nova. As mensagens da fila do nível do exemplo são escolhidas prestando serviços de manutenção.

## AAAMGR enfileira-se

Há dois tipos de filas em AAAMGR para mensagens:

1. Filas niveladas do exemplo
2. Filas niveladas do server

Quando uma mensagem é gerada, está enfileirada inicialmente na fila do nível do exemplo para prestar serviços de manutenção.

A fila do nível do exemplo é processada por 25 milissegundos para o cada milissegundos dos 50 pés. Toda a mensagem que dequeued da fila do nível do exemplo será tentada ser enviada ao servidor Radius. Sob algumas circunstâncias nós não podemos poder enviar as mensagens (nenhuma largura de banda disponível ou nenhuns ID disponíveis). Nesses casos, as mensagens

que falharam a tentativa serão enfileiradas nas filas do nível do server. Para cada milissegundo dos 50 pés você escolhe tantas como mensagens que têm os ID disponíveis e igualmente largura de banda disponível e põe-nos na cabeça da fila do nível do exemplo (estas mensagens são mais velhas do que toda a outra mensagem que estar presente na fila do nível do exemplo).

Quando há um controle de taxa para mensagens explicando, e se há muitas mensagens da contabilidade na fila do nível do exemplo, a seguir todo o mensagem de autenticação novo vai à cauda da fila do nível do exemplo. Para obter processado lhe tem que esperar toda a mensagem da contabilidade (que precede a mensagem nova do AUTH) para ser enviado ao servidor Radius ou para ser movido para a fila do nível do server. É um comportamento existente e não é alterado. Assim pode fazer com que um atraso pequeno para que a mensagem nova do AUTH obtenha processado.

## Exemplo

Baseado na MAX-taxa com valor de 5, você pode enviar cinco mensagens em 1 segundos e ter não respondido proeminente das mensagens da autenticação RADIUS 256 (configuração MAX-proeminente do padrão) pelo aaamgr para o servidor de autenticação RADIUS. Caso que há mais do que as mensagens 5, em 1 segundo as mensagens estão postas na fila até que o servidor AAA responder aos pedidos existentes.

Caso que você alcança as mensagens da autenticação RADIUS 256 enviadas de um aaamgr para o server, os pedidos permanecendo estarão postos na fila até que o servidor AAA responder aos pedidos existentes. Entrará outra vez na mesma fila que aquele da MAX-taxa. A mensagem está pegurada da fila somente quando você tem um entalhe livre. O entalhe livre vem em quando você recebe uma resposta para a mensagem ou quando ele intervalos.

## Limitações

Desde que Cisco ASR5K é um sistema distribuído com sessmgr/aaamgr independente emparelha o processamento dos atendimentos, o estrangulamento da taxa poderia ser executado somente para exemplos independentes do aaamgr. É teórico estender a taxa de uma instância única à caixa inteira de Cisco ASR5K no conjunto apenas multiplicando o número total de exemplos com a MAX-taxa de cada exemplo.

Este número é apenas o limite superior absoluto em uma encenação do dia ensolarado. Você não pode tratar Cisco ASR5K como uma caixa preta e não pode supor que todos os atendimentos devem suceder se o valor calculado considerado no sistema não cruzou o limite superior.

A max-taxa do raio é amarrada com outros parâmetros internos e externos relativos ao sistema. Veja por favor o impacto previsto se uma das circunstâncias não é estado conforme.

### Condições

Distribuição uniforme dos atendimentos do demuxmgr a todos os sessmgrs

Distribuição uniforme de IMSI (isto é só no caso

### Impacto se não encontrado com

Se a distribuição de chamada não é uniforme, a seguir as mensagens do raio podem seja enfileirado para alguns exemplos. Assim mesmo que limite teórico da MAX-taxa não seja alcançado, os atendimentos serão deixados cair para os exemplos onde mensagens são enfileiradas.

o arredondamento robin da Mediação-contabilidade é baseado no roteamento IMSI-baseado.

da contabilidade da mediação do arredondamento robin)

Nenhumas explosões repentinas dos atendimentos vir-em

Os servidores Radius devem responder a tempo

Neste caso, com base na distribuição IMSI, algum grupo de server pode ser preferido sobre outro baseados na lógica de roteamento, fila pôde ser constituído para aqueles server que conduzem a atendimento-gota.

Se há uma explosão de atendimentos novos, a seguir as mensagens recentemente geradas do raio estarão enfileiradas outra vez no sistema. Antes que as requisições RADIUS possam serem processadas. O tempo de instalação da sessão por condução expirada para chamar gotas.

Quando intervalos das requisições RADIUS devido às questões de servidor, lá? o II seja outra vez acúmulo da fila porque os pedidos novos não serão mandados a menos que o atual que espera uma resposta for removido do sistema. A taxa em que as mensagens programadas-para fora serão removidas do sistema é dependente das configurações M proeminentes e do intervalo também.

Em muitos casos nós podemos ver que os pedidos do acesso não estão processados por todas as tarefas ativas do aaamgr. Isso significa que nós estamos tendo a distribuição de chamada desigual dentro das tarefas do sessmgr e mais adicional sobre, não todos os exemplos do aaamgr são envolvidos no Processamento de chamadas.

A distribuição de chamada não é baseada no mecanismo restrito do arredondamento robin que é se há as chamadas recebidas que 10 irão aos sessmgrs 10 em um algoritmo monótona.

A distribuição de chamada é baseada nestes quatro fatores principais

- **active\_session\_count**
- **cpu\_load**
- **Round\_trip\_delay** (demuxmgr? sessmgr? demuxmgr)
- **mais outstanding\_add\_request** (demux ao sessmgr)

Esta é a implementação atual. A max-taxa é apenas um limite superior, mas devido à natureza distribuída de nossa arquitetura, você pode? t extrapola-a diretamente à carga do chassi. O comportamento depende da carga em um AAAMgr dado em um dado momento.

A fila da MAX-taxa do raio deve ser usada **para monitorar o estado do sistema**. Se há um **acúmulo da fila**, então significa que uma destas 4 (refira a tabela) circunstâncias não está estado conforme com e você deve identificar a causa de raiz para o mesmos.

\*\* o ponto inicial da fila da MAX-taxa podia ser executado e constantemente monitorado.