

T1 - Quality of Service

João Pedro Moreira Antunes, Vinícius França Mibielli

6 de maio de 2025

1 Introdução

Este trabalho explora diversas técnicas de Qualidade de Serviço (QoS) aplicadas em redes, com foco em otimizar o gerenciamento de pacotes e garantir o desempenho adequado para diferentes tipos de tráfego, abordando métodos de agendamento como filas por prioridade e ponderadas, técnicas de formatação como Balde Furado e com Tokens, e a implementação de reserva de recursos para banda mínima; cada técnica é detalhada em termos de funcionamento, vantagens e desvantagens, e aplicada em um cenário prático com TCLink, utilizando sua estrutura hierárquica para implementar filas, classes e filtros, demonstrando a configuração para requisitos de tráfego específicos e visando fornecer uma visão geral das técnicas implementadas e sua aplicação no contexto da rede.

2 Técnicas de QoS

2.1 Scheduling

O Scheduling, técnica comum para otimizar a qualidade de serviço, estabelece uma ordem prioritária para o gerenciamento de pacotes nos roteadores. Este trabalho implementou filas por prioridade e ponderada, ampliando o suporte a diversas especificações.

2.1.1 Fila por prioridade

Filas são criadas com diferentes níveis de prioridade (Alta, Média, Baixa, etc.). O scheduler sempre atende a fila de maior prioridade primeiro, esvaziando-a antes de passar para a próxima. Desvantagem: Filas de baixa prioridade podem sofrer "starvation" (nunca serem atendidas se houver tráfego constante nas filas de alta prioridade).

- **Qdisc:** Uma fila do tipo prio com 3 níveis de prioridade foi aplicada à classe raiz da rede, identificada pelo handle 50: .
- **Filter:** Tráfego classificado com base nas portas de destino: pacotes RTP com maior prioridade (fila 0), tráfego entre h3 e h4 classificado com prioridade média (fila 1) e os demais pacotes vão para a fila de menor prioridade (fila 2).

2.1.2 Fila ponderada

Tenta alocar a largura de banda de forma "justa" entre diferentes fluxos ou classes, baseado em pesos defi-

nidos. Fluxos com maior peso recebem uma porção maior da banda. Garante que nenhuma fila seja completamente ignorada.

- **Class e Classid:** São criadas duas subclasses, 5:10 e 5:20, subordinadas da classe raiz (5:1), com diferentes tamanhos de largura de banda.
- **Qdisc:** Aplicação de uma fila do tipo pfifo para cada subclasse, identificada por um único handle.
- **Filter:** Priorização de banda é feita com base nas portas: RTP com classe de maior banda, tráfego entre h3 e h4 com classe intermediária e os demais fluxos possuem menor prioridade.

2.2 Formatação do Tráfego

O traffic shaping é essencial pra controlar a velocidade da internet. Ele ajuda a evitar que a rede fique muito lenta em momentos de pico, garantindo que a velocidade contratada seja entregue, mesmo que ela varie ou tenha limites. O único problema é que usar essa técnica pode deixar a rede um pouco mais lenta.

2.2.1 Balde Furado (TBF)

Uma das opções de formatação foi a utilização do método Leaky Bucket (LB) para modelar o tráfego de saída a uma taxa constante. Pacotes de entrada são armazenados em um buffer de tamanho fixo e liberados a uma taxa invariável pré-definida. O estouro do buffer resulta em descarte de pacotes. Os parâmetros chave configurados foram a capacidade do buffer e a taxa de saída, permitindo o controle da vazão do tráfego.

- O TBF é configurado com burst = 32kbit e latency=400ms como padrão, limitando a taxa de transmissão das classes.
- Seu funcionamento é semelhante ao da Fila Ponderada, porém com foco em limitar o uso de banda, especialmente para o tráfego gerado entre h3 e h4.
- Para Balde Furado com Tokens, ele é ajustado com burst=256kbit e latency=50ms, possibilitando rajadas no tráfego quando necessário.

2.2.2 Balde Furado com Tokens (TBF+)

A outra opção disponível para a formatação do tráfego foi o Token Bucket, configurado com uma taxa média (r) e um tamanho de rajada (b). O algoritmo opera pela geração e consumo de tokens que representam permissão de envio. A falta de tokens para um pacote disparou a ação de formatação, permitindo a análise do tráfego em relação aos limites de banda impostos.

2.3 Reserva de Recursos

Foi também implementada uma Reserva de Recursos que permite que os fluxos de dados solicitem e recebam recursos de rede específicos (como banda e buffers) através de sinalização e controle de admissão. Essa abordagem foi utilizada para assegurar garantias de Qualidade de Serviço (QoS) para os fluxos de dados mais importantes.

- Reserva de recursos vai garantir que o tráfego do RTP tenha uma banda mínima mesmo em situações de alto congestionamento.
- Por meio da Hierarquia HTB, uma subclasse é criada com rate mínimo para o tráfego RTP. Caso as outras classes estejam ociosas, parte da banda será temporariamente emprestada via parâmetro ceil, garantindo estabilidade para a transmissão prioritária.

3 Funcionamento do Sistema

3.1 Inicialização e limpeza de dados

O sistema inicia perguntando ao usuário se deseja excluir o arquivo resultados.csv, que armazena os valores de jitter e latência média da transmissão de vídeo via RTP.

Limpar o arquivo csv? [1] Não [2] Sim

Escolha: ☐

3.2 Seleção de técnicas de QoS

Em seguida, é apresentado um menu com opções de técnicas de QoS, permitindo ao usuário aplicar uma ou mais abordagens para melhorar a qualidade da transmissão de vídeo.

MENU DE OPÇÕES QoS
Escalonamento: [1] Nenhum [2] Prioritário [3] Ponderado
Controle de Tráfego: [1] Nenhum [2] Balde Furado [3] Balde Furado com Tokens
Reserva de Recursos: [1] Não [2] Sim

Escolha o tipo de escalonamento: 3

Escolha o controle de tráfego: 2

Reserva de recursos ativa?: 2 ☐

3.3 Configuração da rede

O usuário pode personalizar a largura de banda do link e definir a quantidade de fluxos de tráfego de fundo (UDP e TCP) gerados com o iperf, simulando diferentes níveis de congestionamento.

Largura de banda: Mínimo: 2Mbits Padrão: 10Mbits Máximo: 50Mbits
Instâncias paralelas: Aumentam o congestionamento na rede. Cada instância possui um tráfego UDP de 100Mbits e TCP de 50Mbits com duração de 20 segundos.

Escolha a largura de banda da rede: 10

Escolha o número de instâncias paralelas entre h3 e h4: 30 ☐

3.4 Visualização da interface

A configuração da interface do switch 1 é exibida, permitindo ao usuário verificar os parâmetros ativos da rede.

[QoS] Aplicando escalonamento 'ponderado' como filha de HTB em s1-eth3 de s1...
[QoS] Aplicando 'Balde Furado' como filha de HTB em s1-eth3 de s1...
[QoS] Aplicando 'Reserva de Recursos' como filha de HTB em s1-eth3 de s1...
Configuração atual do tc em s1-eth3 de s1:

3.5 Coleta e armazenamento dos dados

Ao final da simulação, os dados de latência média e jitter do tráfego RTP são coletados e armazenados no arquivo resultados.csv. Jitter: é a variação do tempo de entrega dos pacotes. Valores altos de jitter influenciam na qualidade de voz e vídeo. Latência média: é o tempo médio que um pacote leva para ir de um ponto a outro na rede. Dita a velocidade da comunicação.

- **Jitter:** é a variação do tempo de entrega dos pacotes. Valores altos de jitter influenciam na qualidade de voz e vídeo.
- **Latência média:** é o tempo médio que um pacote leva para ir de um ponto a outro na rede. Dita a

velocidade da comunicação.

Observação: Para simular um maior nível de congestionamento na rede, são utilizados valores mais altos de taxa de envio (rate) no iperf e tamanhos maiores de pacotes RTP (1500).

4 Conclusão

1	Tecnica	Latência Média (ms)	Jitter (ms)	Banda Larga (Mbits)	Nº de Instâncias
2	nenhum+nenhum+não	274.997	168.519	10	10
3	ponderado+tlb+sim	1.396	2.495	10	10

A análise prática das técnicas de Qualidade de Serviço aplicadas evidenciou de forma significativa os benefícios do uso combinado de estratégias de escalonamento, controle de tráfego e reserva de recursos. Com uma rede operando sob largura de banda de 10 Mbit/s e 10 instâncias de tráfego de fundo, observou-se que a ausência de qualquer técnica resultou em uma latência média extremamente elevada (274,997 ms) e um jitter igualmente alto (168,519 ms), indicando grande influência negativa na qualidade do serviço para o tráfego RTP. Enquanto que, para a aplicação conjunta das técnicas de fila ponderada, Balde Furado com Tokens (TBF+) e reserva de recursos reduziu drasticamente a latência média, que caiu para apenas 1,396 ms, e o jitter para 2,495 ms, mostrando uma melhoria de mais de 98% na estabilidade e no tempo de entrega dos pacotes. A escolha adequada de combinações de técnicas pode garantir o desempenho da transmissão RTP mesmo em situações de alto congestionamento.