



**PUC Minas**

# Projeto de Redes de Computadores

---

**Congestionamento e QoS**

# Congestionamento e QoS

---

- Qualidade de Serviço (QoS)
- Controle de Congestionamento do TCP
- Modelos de QoS (IntServ e DiffServ)



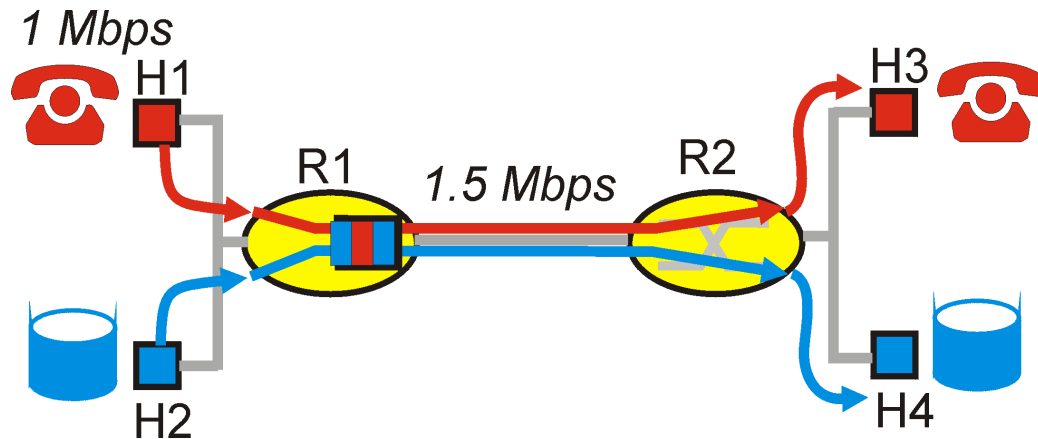
# Congestionamento e QoS – Qualidade de Serviço

---

- Atualmente a Internet oferece serviços de melhor esforço
  - Sem garantia de entrega
  - Sem controle de admissão
  - Sem classificação de tráfego
- O futuro das aplicações:
  - Qualidade de Voz sobre IP (VoIP)
  - Transmissão de vídeo em tempo real
  - Adaptações de congestionamento

# Congestionamento e QoS – Qualidade de Serviço

O que acontece em um enlace de 1.5Mbps quando ocorrer a transmissão simultânea de uma chamada de voz com uma transmissão de arquivo via FTP?



**Congestionamento** → Afetando os serviços de Voz

# Congestionamento e QoS – Qualidade de Serviço

---

- Qualidade de Serviço (QoS)

"Conjunto de técnicas e mecanismos que garantem o desempenho da rede, com o objetivo de oferecer um serviço previsível para um programa da camada de aplicação." (Forouzan)

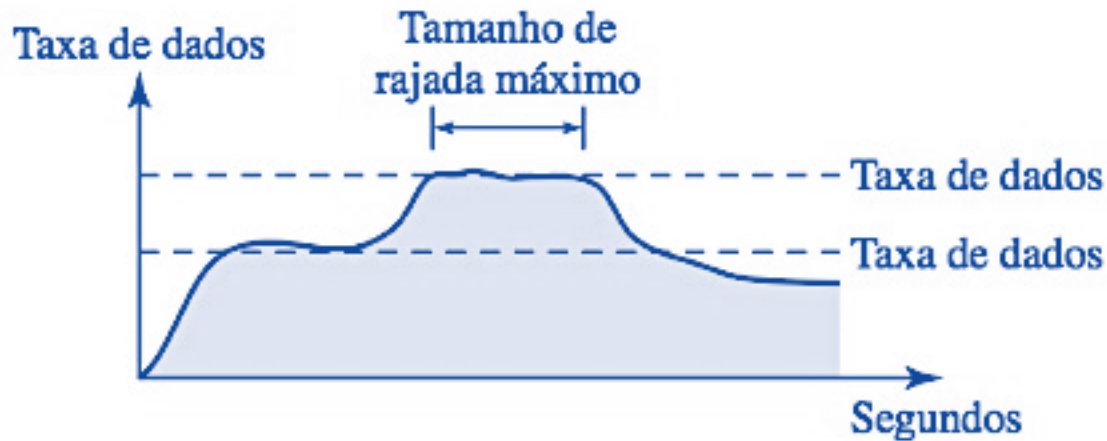
# Controle de Congestionamento e QoS

---

- Principal foco do controle de congestionamento e qualidade de serviços: tráfego de dados
- Controle de congestionamento: evitar o congestionamento do tráfego de dados
- Qualidade de serviços: criar um ambiente apropriado para o tráfego

# Controle de Congestionamento e QoS

- Descritores de tráfego: valores qualitativos que representam o fluxo de dados



# Controle de Congestionamento e QoS

---

- A taxa de dados média: é o número de bits transmitidos durante certo período, dividido pelo número de segundos desse período

$$\text{Taxa de dados média} = \frac{\text{volume de dados}}{\text{tempo}}$$

- A taxa de dados de pico define a taxa de dados máxima do tráfego
  - Indica a largura de banda de pico que a rede precisa para o tráfego passar sem mudar seu fluxo de dados



# Controle de Congestionamento e QoS

---

- O tamanho máximo de rajada, em geral, refere-se ao período máximo que o tráfego é gerado na taxa de pico
- Largura de banda efetiva: largura de banda que a rede precisa alocar para o fluxo de tráfego

# Perfis de tráfego – Taxa de bits constante

---

- É um modelo de tráfego de velocidade fixa que possui uma taxa de dados que não muda
- Nesse tipo de fluxo, a taxa de dados média e a taxa de dados de pico são iguais. O tamanho máximo de rajada não é aplicável
- Esse tipo de tráfego é muito fácil para uma rede lidar já que ele é previsível

# Perfis de tráfego – Taxa de bits variável

---

- A taxa de fluxo de dados muda com o tempo, com mudanças suaves em vez de repentinas e abruptas
- Nesse tipo de fluxo, a taxa de dados média e a taxa de dados de pico são diferentes. O tamanho máximo de rajada normalmente é um valor pequeno
- Tipo de tráfego mais difícil de ser tratado que o tráfego com taxa de bits constante, porém, geralmente, ele não precisa ser reformulado

# Perfis de tráfego – Taxa de bits em rajada

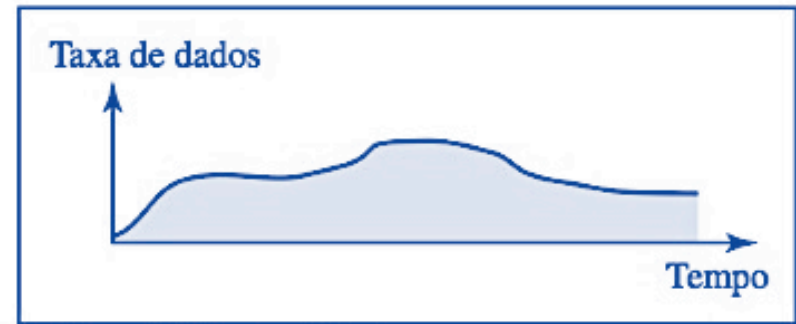
---

- A taxa de dados muda repentinamente em um espaço de tempo muito curto
- A taxa de bits média e a taxa de bits de pico são valores bem diferentes. O tamanho máximo de rajada é significativo
- Este é o tipo de tráfego mais difícil para uma rede lidar, pois o perfil é muito imprevisível

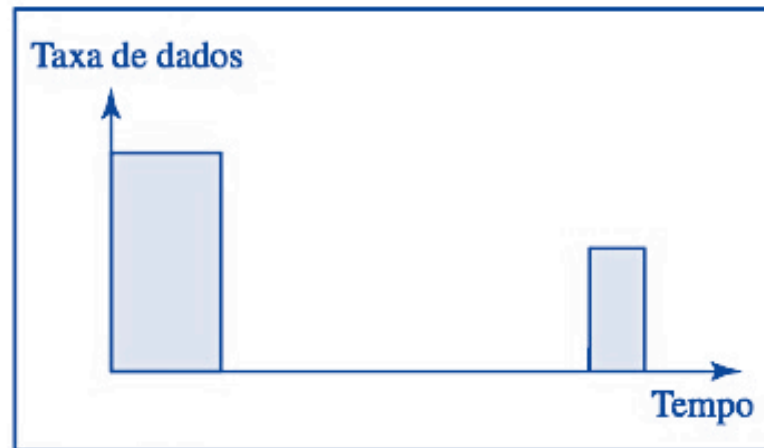
# Perfis de tráfego



**a. Taxa de bits constante**



**b. Taxa de bits variável**



**c. Em rajadas**

# Controle de Congestionamento e QoS

## Métricas comuns

### ■ Vazão (Throughput)

- Taxa efetiva de pacotes transferidos por tempo
- Largura de Banda = Capacidade do canal

### ■ Perdas

- Percentual de perda de pacotes no tempo

### ■ Atraso ou Latência

- Tempo gasto pelos pacotes para ir da origem ao destino (medido em ms)



### ■ Jitter

- Variação do retardo

Fonte: <http://labcisco.blogspot.com.br/2013/05/interpretacao-dos-resultados-do-ping.html>

# Controle de Congestionamento e QoS

## Sensibilidade das Aplicações

Tipo de Tráfego	Vazão	Perdas	Latência	Jitter
Voz	Muito baixa	Média	Alta	Alta
Comércio eletrônico	Baixa	Alta	Alta	Baixa
Transações	Baixa	Alta	Alta	Baixa
Correio eletrônico	Baixa	Alta	Baixa	Baixa
Acesso Remoto(Telnet)	Baixa	Alta	Média	Baixa
Navegação web casual	Baixa	Média	Média	Baixa
Navegação web crítica	Média	Alta	Alta	Baixa
Transferência de arquivos	Alta	Média	Baixa	Baixa
Videoconferência	Alta	Média	Alta	Alta
Multicast	Alta	Alta	Alta	Alta

Fonte: <http://www.gta.ufrj.br>

# Controle de Congestionamento e QoS

---

## Quatro Princípios Básicos

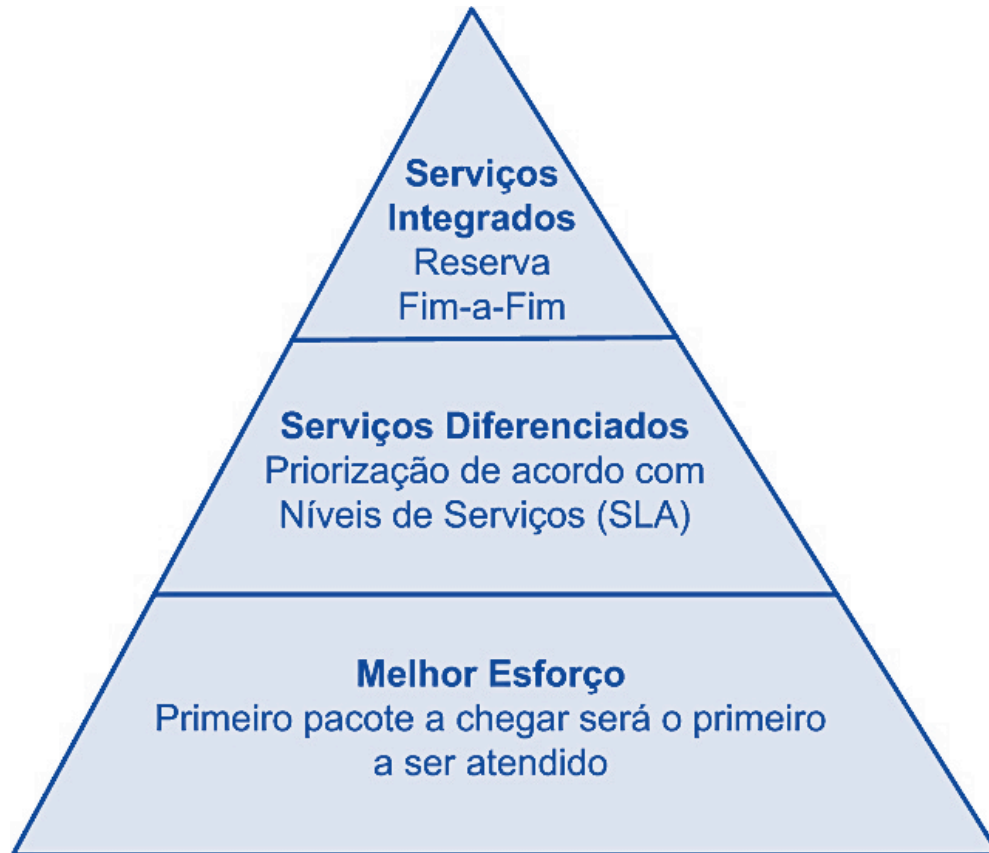
- Marcação: Marque os pacotes para que o roteador possa diferenciar diferentes tipos de classe
- Isolação: Separe uma classe de outra
- Alocação: Aloque uma porção de banda para cada classe de serviço (política)
- Admissão: A capacidade do link estourou? Bloqueie!



# Controle de Congestionamento e QoS

---

## Tipos de Serviço



# Controle de Congestionamento e QoS

---

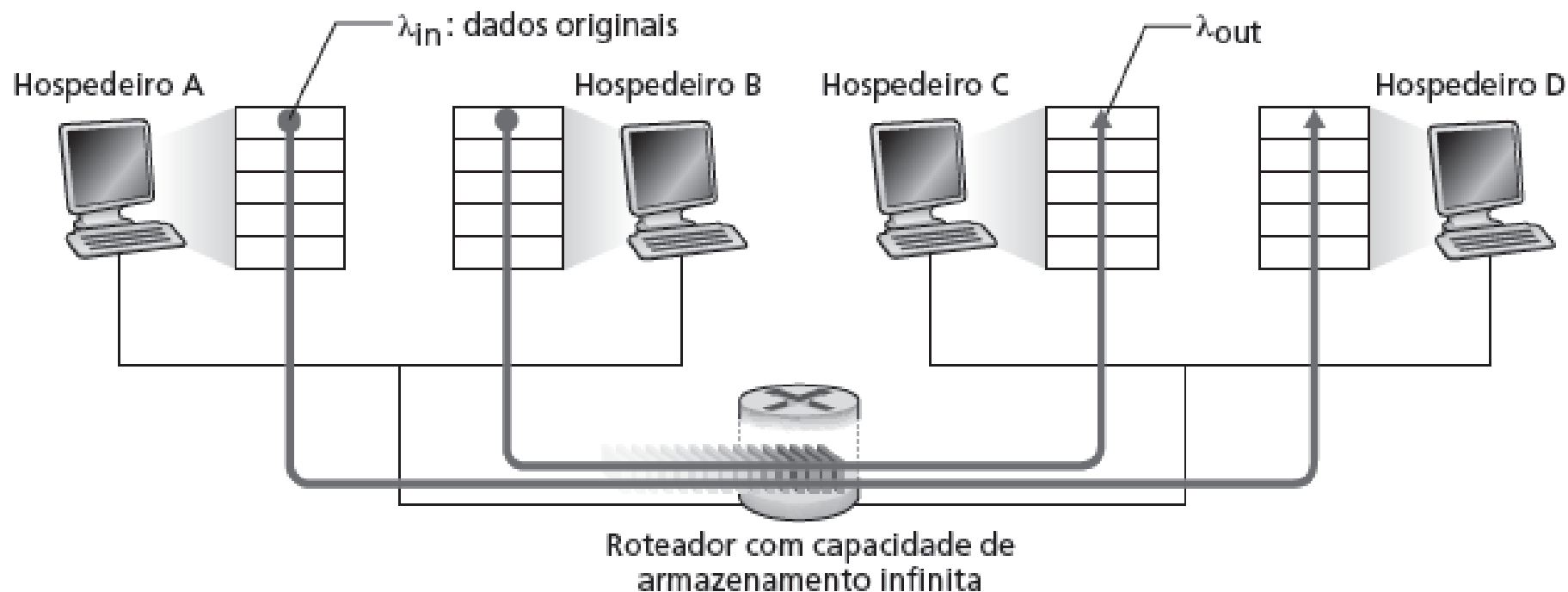
## O que é Congestionamento

- Informalmente: diversas fontes enviando dados a uma taxa acima da taxa que a rede consegue tratar
- É diferente de controle de fluxo
- Sintomas:
  - pacotes perdidos (estouro de buffer nos roteadores)
  - longos atrasos (enfileiramento nos buffers do roteador)
- É um dos maiores problemas da rede

# Controle de Congestionamento e QoS

## Causas e Custos de Congestionamento – CENÁRIO 1

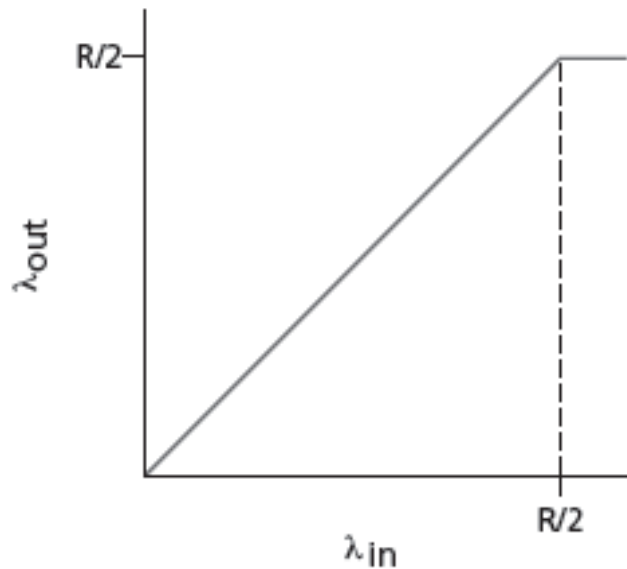
- Dois remetentes/destinatários e um roteador com buffer infinito
- Sem retransmissão de dados perdidos



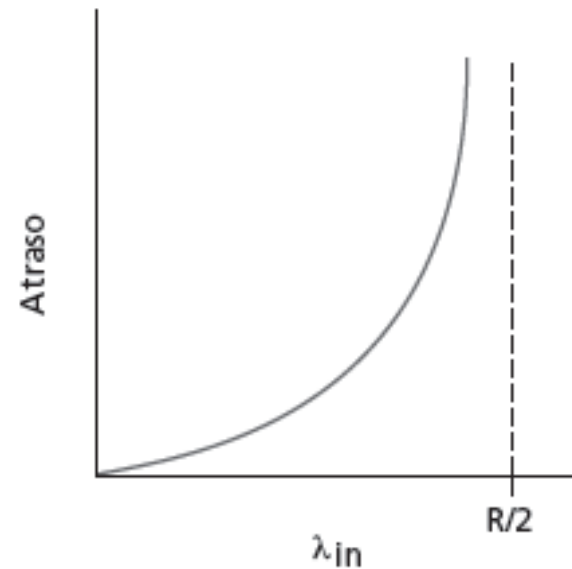
# Controle de Congestionamento e QoS

## Causas e Custos de Congestionamento – CENÁRIO 1

- Vazão atinge a capacidade nominal do canal (Toda a largura de banda)
- Aumento do tempo de atraso dos pacotes que ficam retidos por conta do congestionamento



a.

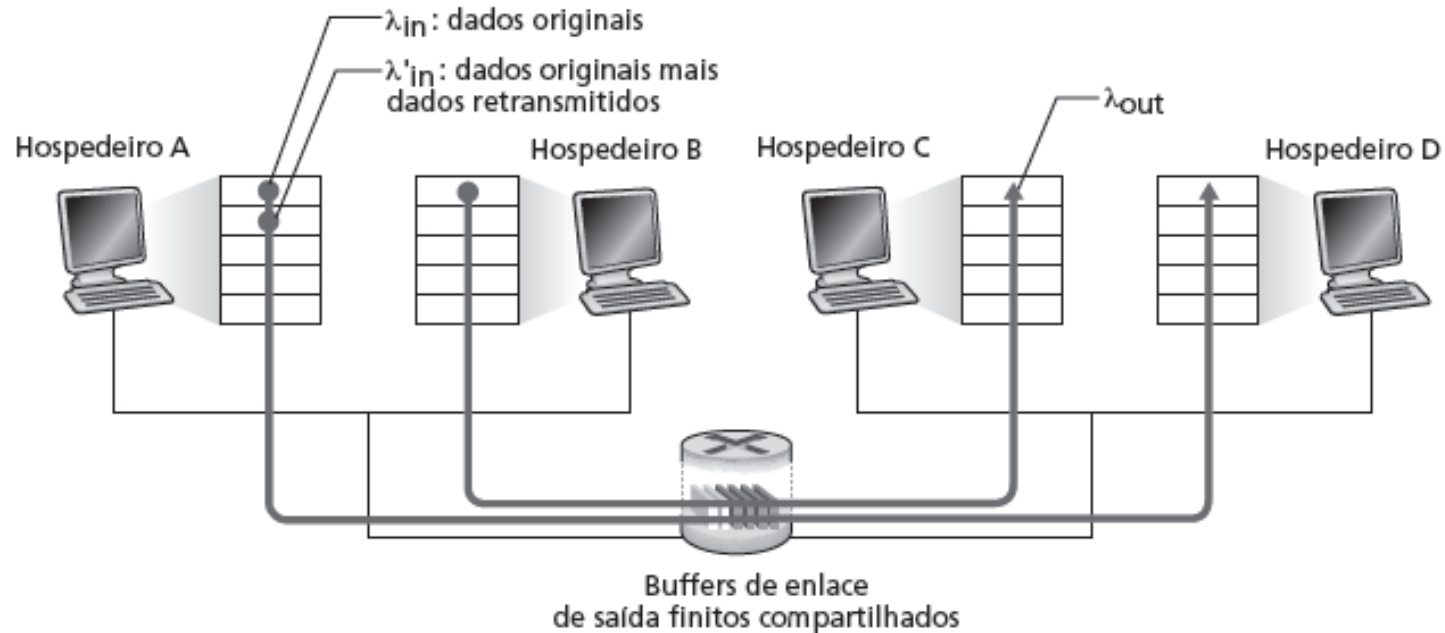


b.

# Controle de Congestionamento e QoS

## Causas e Custos de Congestionamento – CENÁRIO 2

- Dois remetentes/destinatários e um roteador com buffer finito
- Retransmissão do pacote perdido pelo remetente



# Controle de Congestionamento e QoS

---

## Técnicas

- Controle fim-a-fim
  - Não há *feedbacks* explícito da rede
  - O congestionamento é deduzido pela perda/atraso observados do sistema final
  - Técnica adotada pela camada de transporte do TCP
- Controle assistido pela Rede
  - Roteadores oferecem feedback aos sistemas finais:
    - ✓ único bit indicando congestionamento (SNA, DECbit, TCP/IP ECN, ATM)

# Controle de Congestionamento e QoS

---

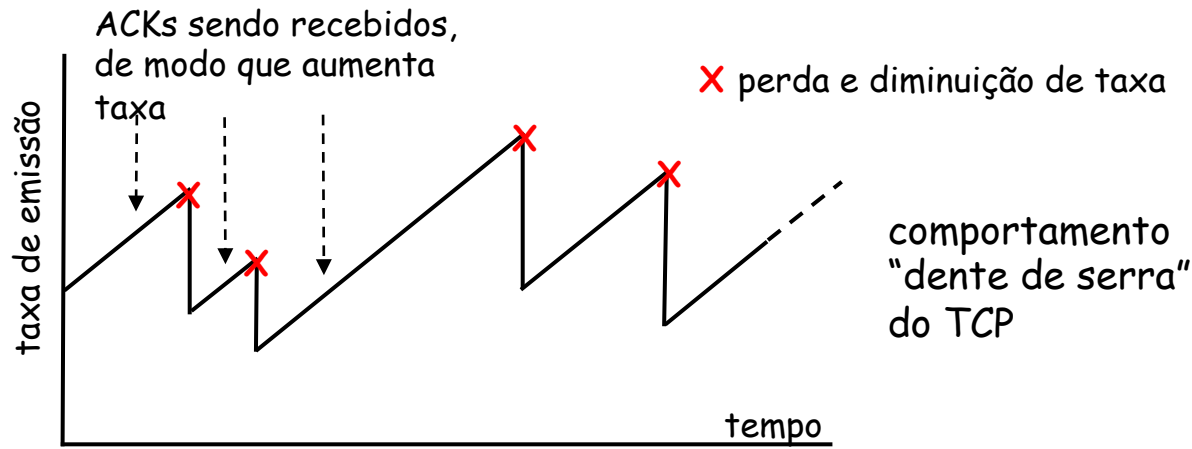
## Controle fim-a-fim

- Nenhum pacote será enviado informando que ocorreu congestionamento: deduzido em função da perda de pacote e atrasos. Destinatário não envia confirmação antes do time-out do remetente estourar

# Controle de Congestionamento e QoS

## Controle de congestionamento do TCP – Partida Lenta

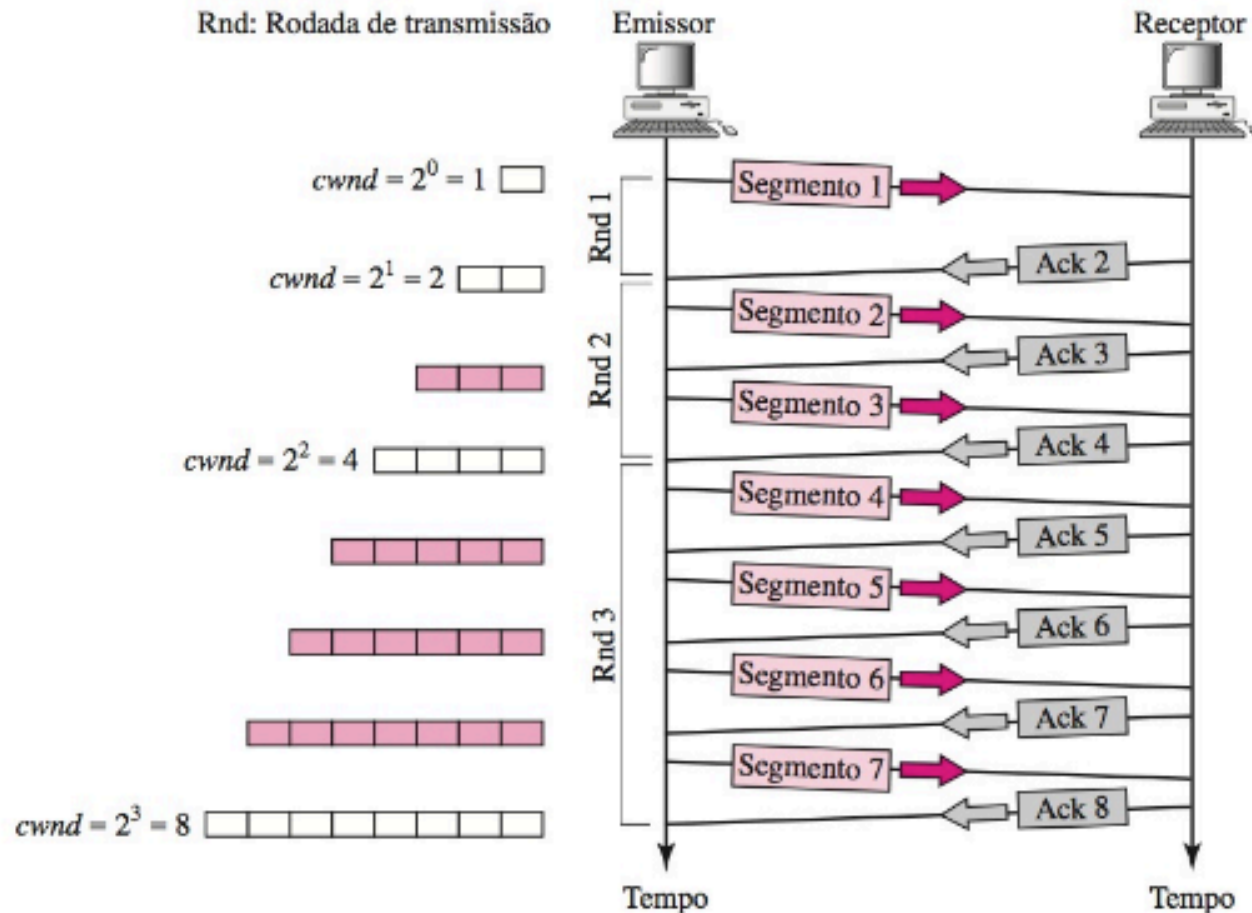
- Ocorre o aumento da taxa de transmissão no recebimento do ACK até que ocorrem perdas o que resulta em diminuição da taxa de transmissão
- Continua a aumentar no ACK e diminui na perda (pois largura de banda disponível está mudando, dependendo de outras conexões na rede)





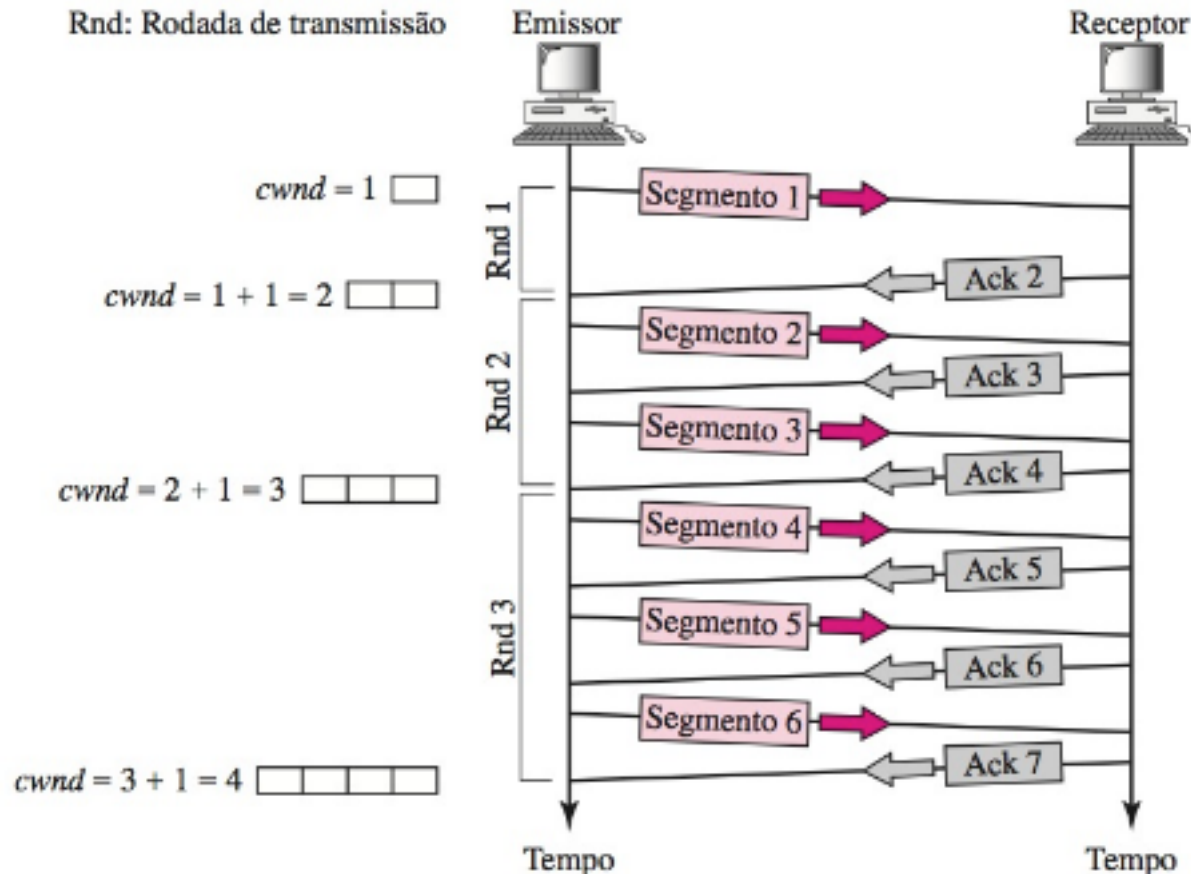
# Controle de Congestionamento e QoS

## Partida Lenta



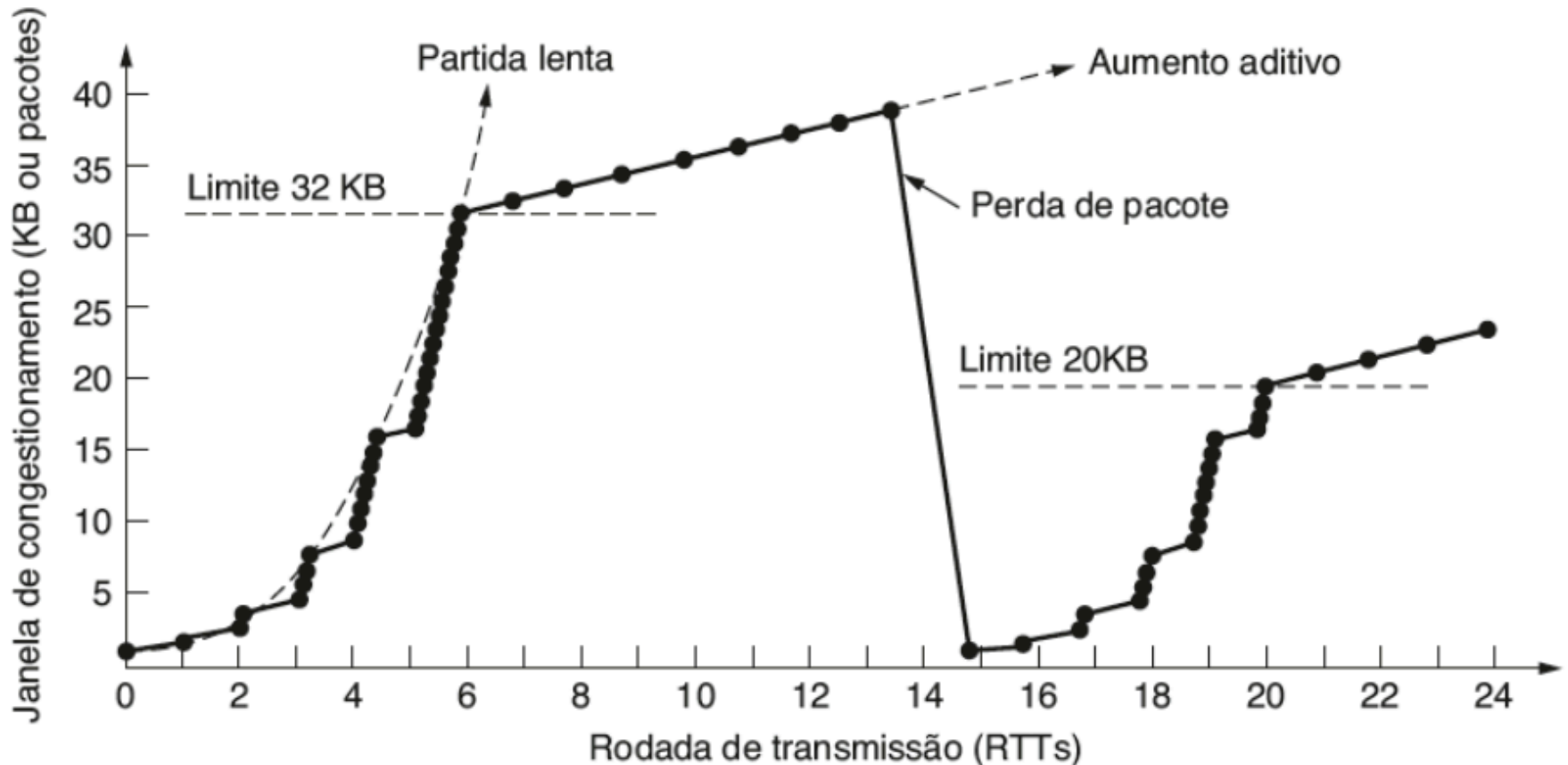
# Controle de Congestionamento e QoS

## Aumento aditivo



# Controle de Congestionamento e QoS

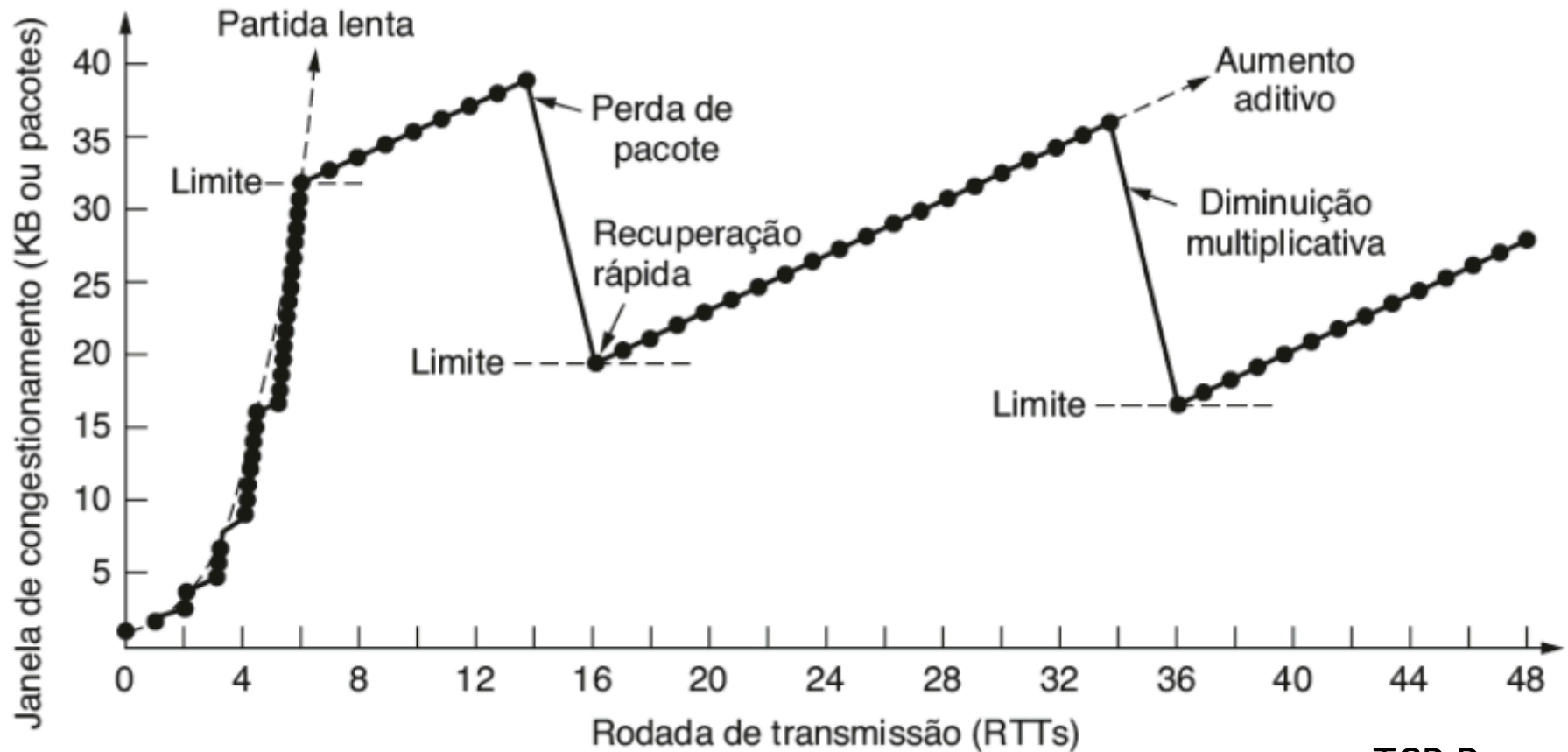
## Diminuição multiplicativa por esgotamento do timer



TCP Tahoe

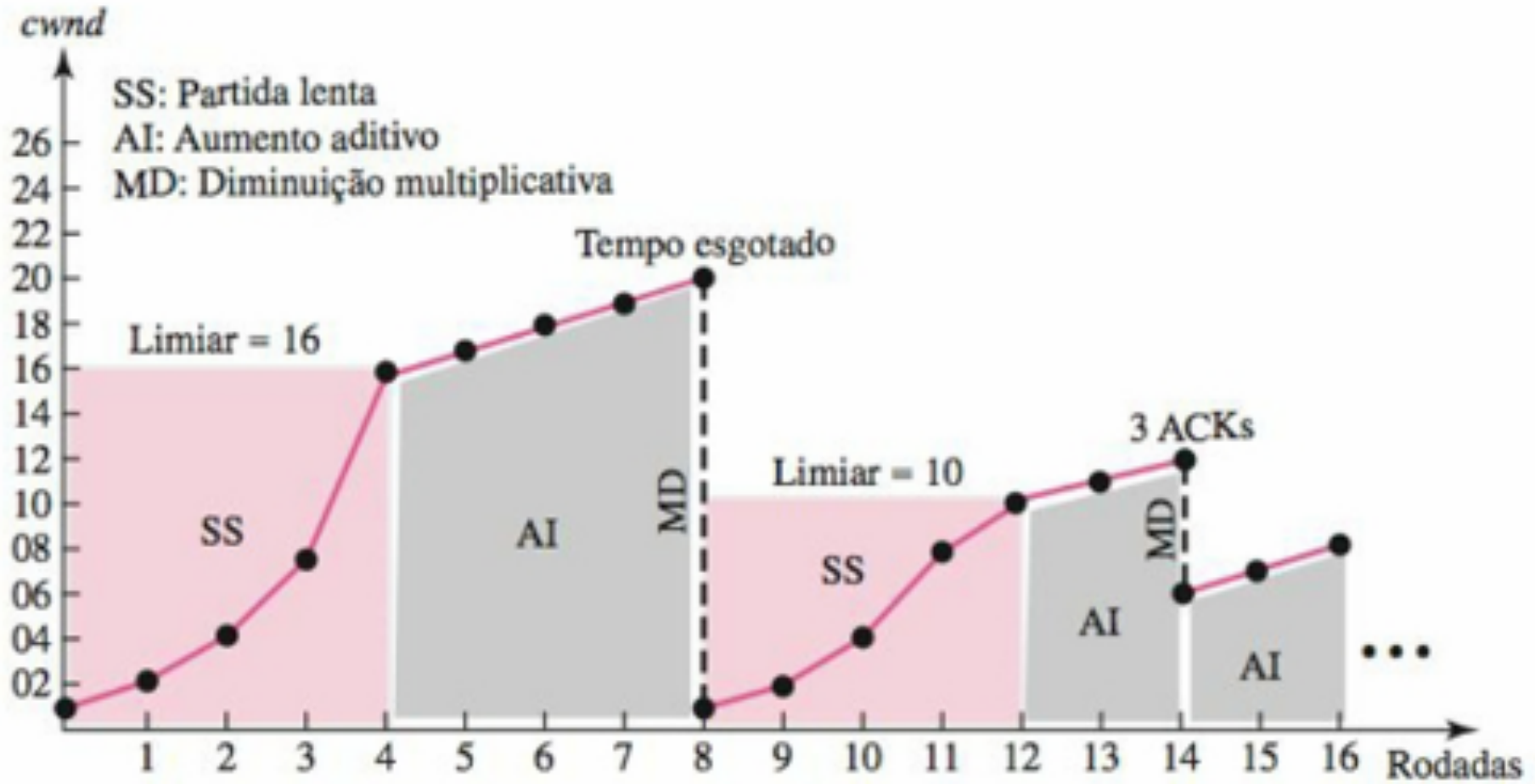
# Controle de Congestionamento e QoS

## Diminuição multiplicativa por recebimento de 3 ACK's



TCP Reno

# Controle de Congestionamento e QoS



# Controle de congestionamento no Frame Relay

---

- O congestionamento em uma rede Frame Relay diminui o *throughput* e aumenta o atraso. Um *throughput* elevado e um atraso pequeno são os principais objetivos do protocolo Frame Relay
- O Frame Relay não possui controle de fluxo
- O Frame Relay permite que o usuário transmita dados em rajadas

# Controle de congestionamento no Frame Relay

---

- Frame Relay usa 2 bits no frame para alertar explicitamente a origem e o destino da presença de congestionamento
  - O bit BECN (*Backward Explicit Congestion Notification* - notificação de congestionamento explícito no sentido inverso)
  - O bit FECN (*Forward explicit congestion notification* - notificação de congestionamento explícito no sentido direto)

# Controle de congestionamento no Frame Relay

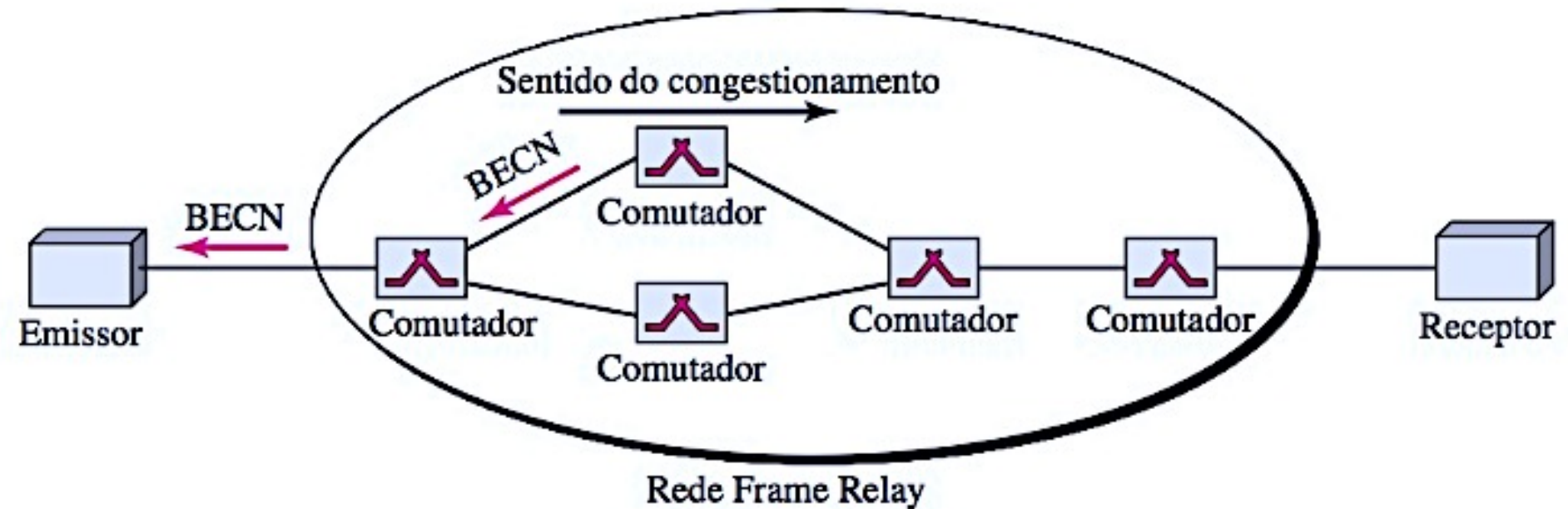
---

- **BECN:** alerta o emissor sobre o congestionamento na rede
- Como isso é feito? Os frames estão trafegando na direção inversa do emissor
  1. o comutador pode usar frames de resposta do receptor (modo full-duplex) ou
  2. o comutador pode usar uma conexão predefinida para transmitir frames especiais para essa finalidade específica

O emissor pode responder a esse alerta simplesmente reduzindo a taxa de dados



# Controle de congestionamento no Frame Relay

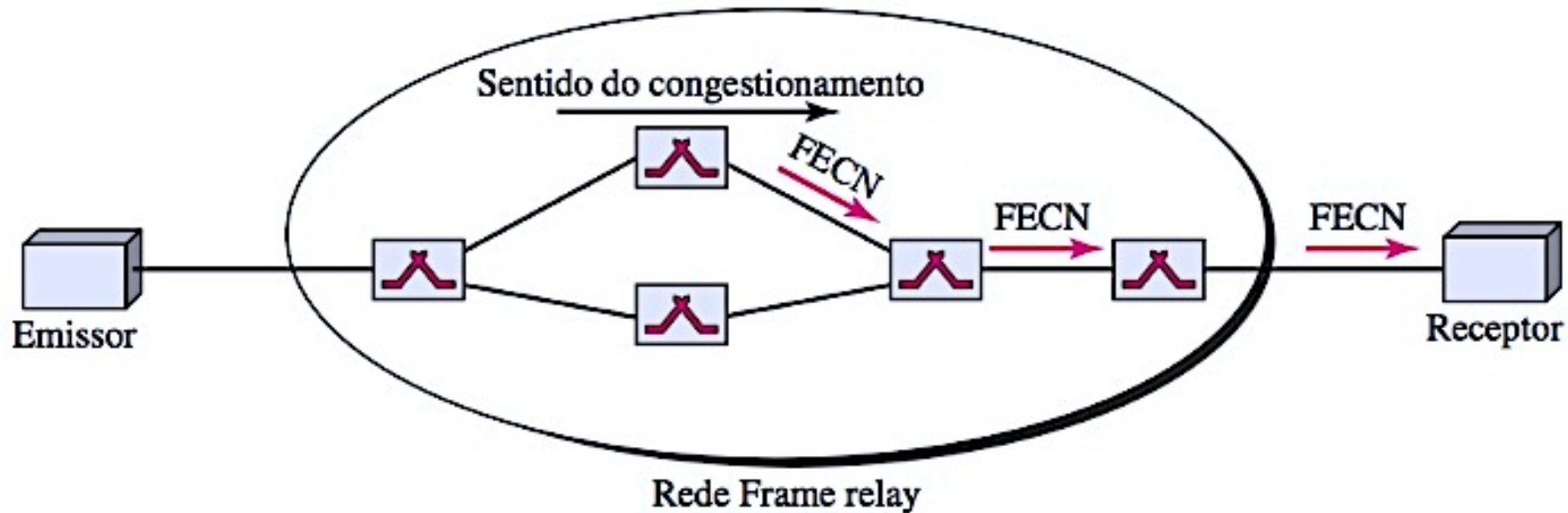


# Controle de congestionamento no Frame Relay

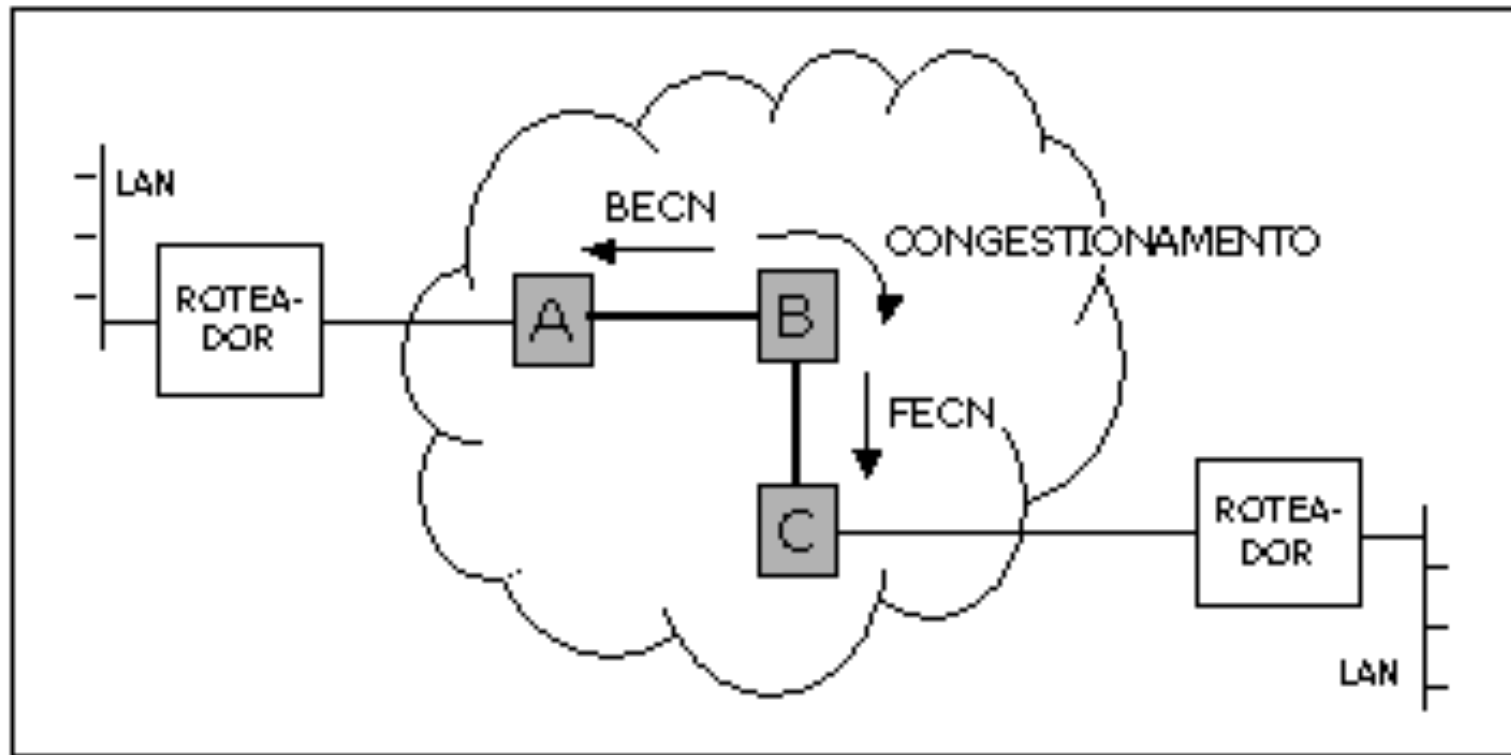
---

- **FECN:** alerta o receptor sobre congestionamentos na rede
- O protocolo Frame Relay supõe que o emissor e o receptor estejam se comunicando e estejam usando algum tipo de controle de fluxo em um nível mais alto
- Se houver um mecanismo de confirmação no nível mais alto, o receptor pode atrasar as confirmações forçando o emissor a diminuir o ritmo

# Controle de congestionamento no Frame Relay



# Controle de congestionamento no Frame Relay



# Técnicas para melhorar QoS

---

- Programação
- Formatação de tráfego
- Controle de admissão
- Reserva de recursos

# Técnicas para melhorar QoS - Programação

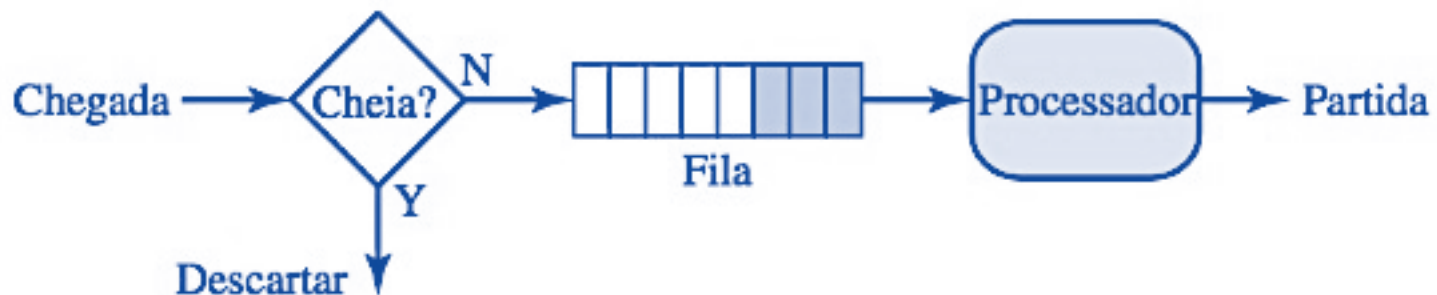
---

- Pacotes de fluxos diferentes chegam em um comutador ou roteador para processamento
- Uma boa técnica de programação trata os diferentes fluxos de maneira equilibrada e apropriada
  - Formação de filas FIFO
  - Formação de filas por prioridade
  - Formação de filas ponderadas

# Técnicas para melhorar QoS - Programação

## ■ Formação de filas FIFO

- Os pacotes aguardam em um buffer até que o nó (roteador ou comutador) esteja pronto para processá-los
- Se a velocidade média de chegada > velocidade média de processamento: fila cheia - novos pacotes que entram serão descartados



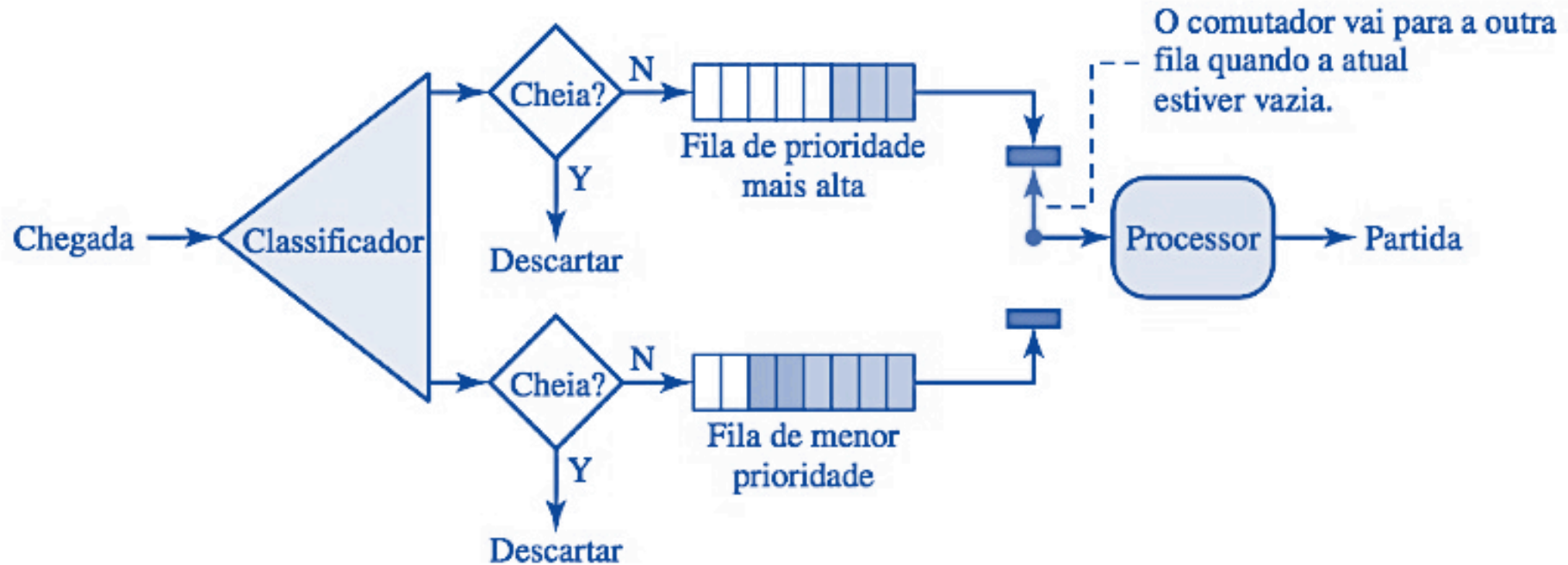
# Técnicas para melhorar QoS - Programação

---

- **Formação de filas por prioridade**
- Os pacotes recebem um nível de prioridade e cada nível de prioridade tem sua própria fila
- Os pacotes na fila de maior prioridade são processados primeiro. Os pacotes na fila de menor prioridade são processados por último



# Técnicas para melhorar QoS - Programação



Se existir um fluxo contínuo em uma fila de alta prioridade, os pacotes nas filas de menor prioridade jamais terão uma chance de serem processados: inanição

# Técnicas para melhorar QoS - Programação

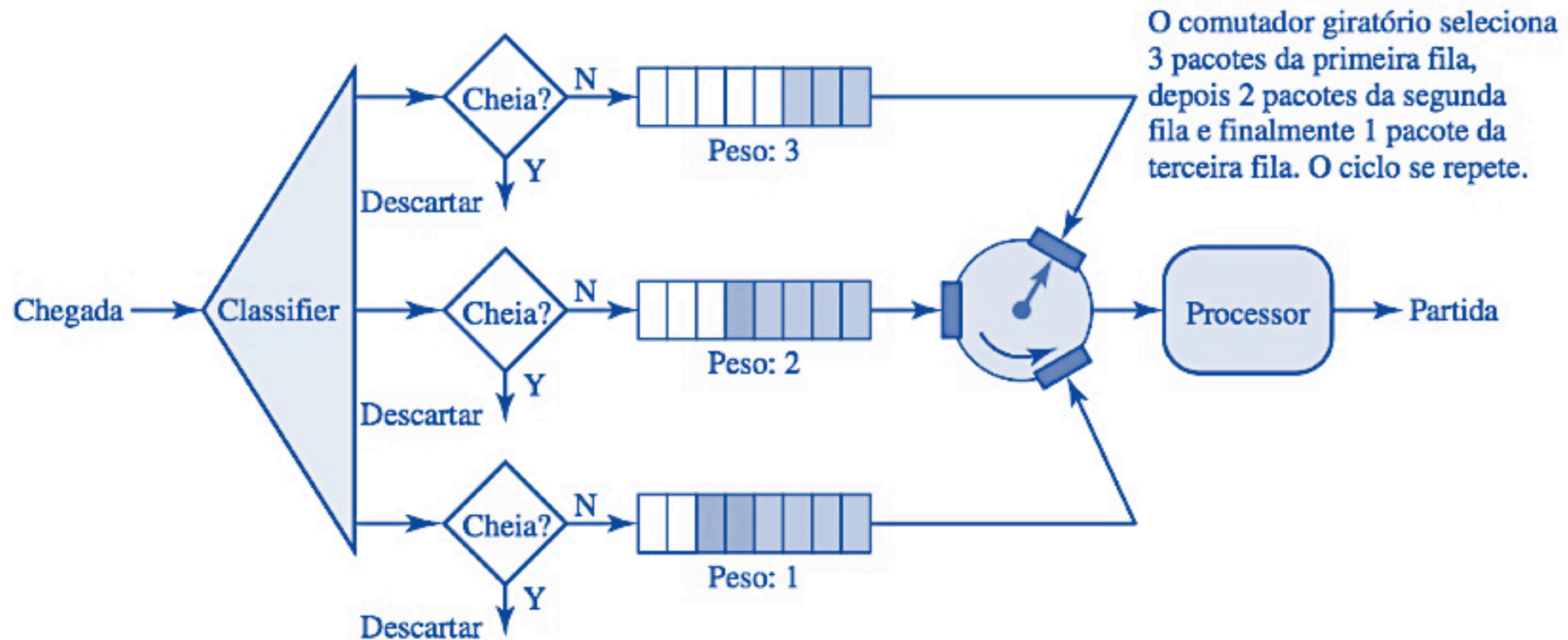
---

## ■ **Formação de filas ponderadas**

- Os pacotes recebem diferentes classificações e são admitidos em filas diferentes
- As filas são ponderadas em termos de prioridade das filas: prioridade maior significa peso maior
- O sistema processa pacotes em cada fila, em rodízio, com o número de pacotes selecionados de cada fila com base no peso correspondente

# Técnicas para melhorar QoS - Programação

## ■ Formação de filas ponderadas



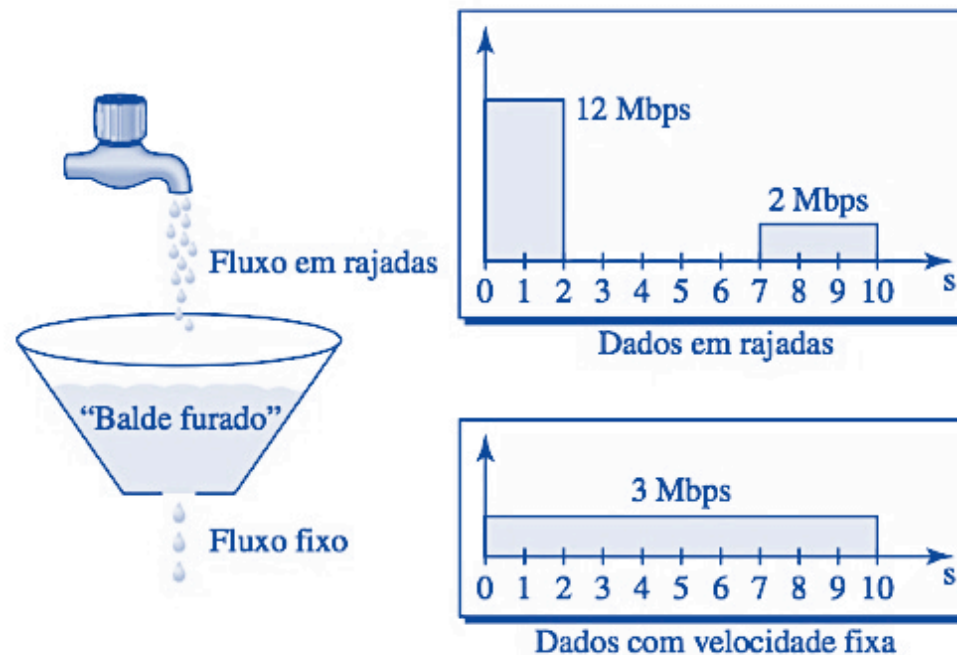
# Técnicas para melhorar QoS - Formatação

---

- Formatação de tráfego: mecanismo para controlar a quantidade e a velocidade com que o tráfego é transmitido pela rede
  - Balde furado
  - Balde de fichas

# Técnicas para melhorar QoS - Formatação

- Balde furado: conjuntos de blocos em rajadas são armazenados no balde e transmitidos a uma velocidade média



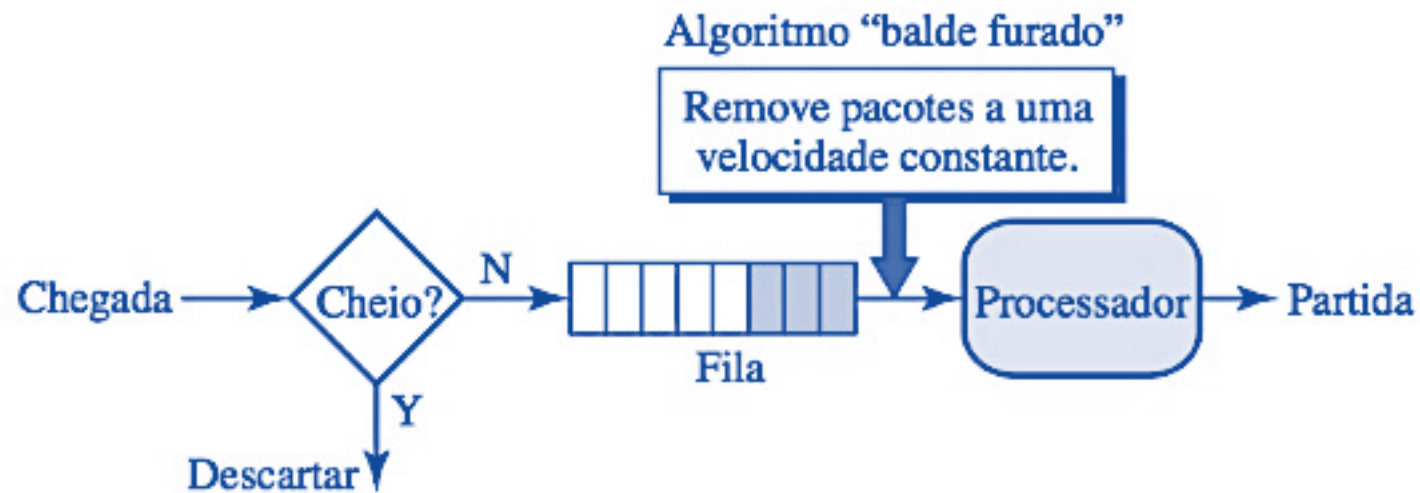
# Técnicas para melhorar QoS - Formatação

---

- Implementação: fila FIFO
- Tráfego formado por pacotes de tamanho fixo: o processo elimina um número fixo de pacotes da fila a cada instante de clock
- Tráfego formado por pacotes de comprimento variável: a velocidade de saída fixa deve se basear no número de bytes ou bits

# Técnicas para melhorar QoS - Formatação

- Implementação técnica Balde furado



# Técnicas para melhorar QoS - Formatação

---

- Balde de fichas: permite que hosts ociosos acumulem crédito, na forma de fichas, para o futuro
- Para cada período do clock, o sistema envia  $n$  fichas para o balde. O sistema elimina uma ficha para cada célula (ou byte) de dados transmitidos
- Permite tráfego em rajadas a uma velocidade máxima regular

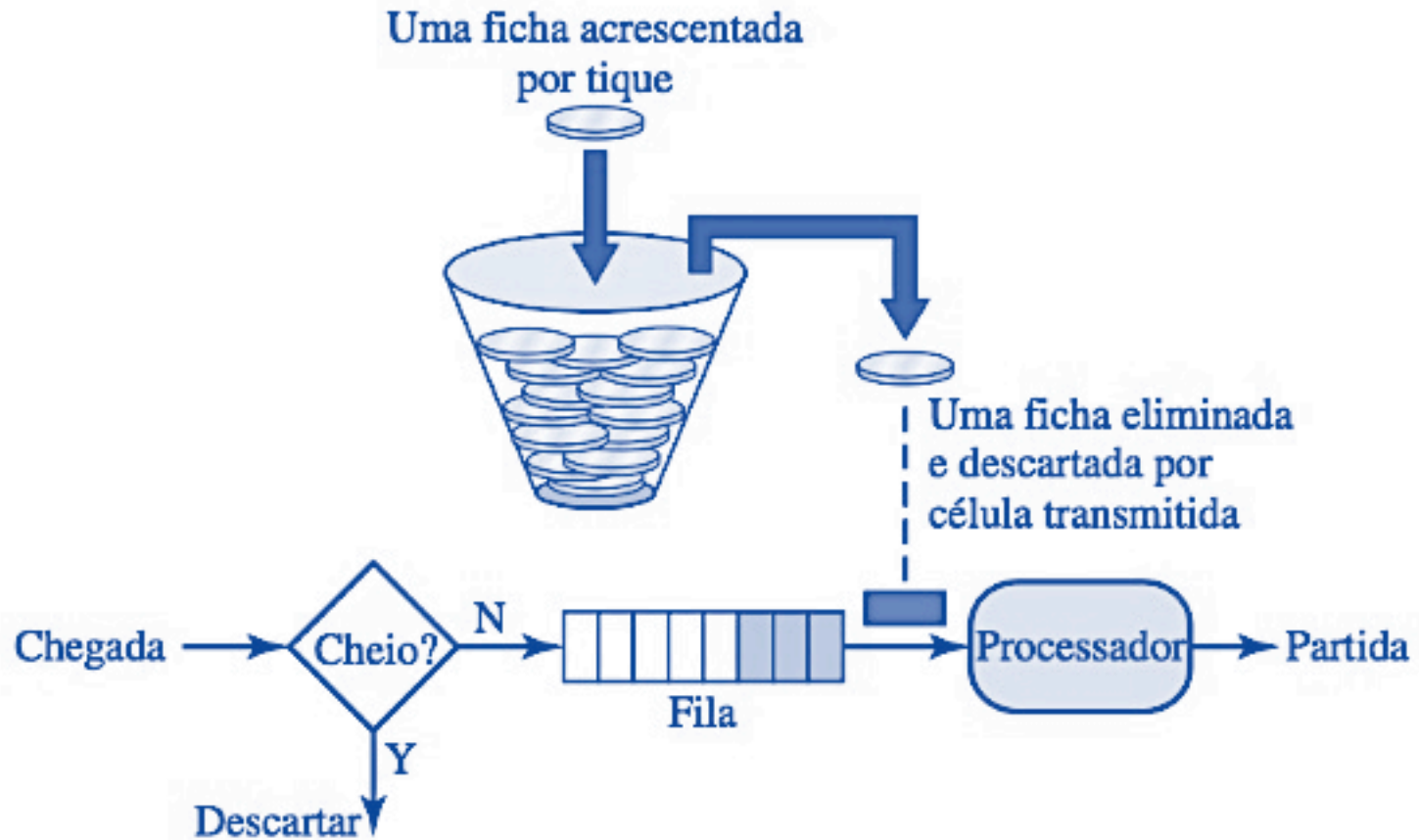


# Técnicas para melhorar QoS - Formatação

---

- Implementação: Contador
  - Ficha inicializada em zero
  - Cada vez que for adicionada uma ficha, o contador é incrementado em 1
  - Cada vez que uma unidade de dados for enviada, o contador é decrementado de 1 unidade
  - Quando o contador chegar a zero, o host não poderá enviar dados

# Técnicas para melhorar QoS - Formatação



# Técnicas para melhorar QoS – Controle Admissão

---

- Mecanismo usado por um roteador, ou um comutador, para aceitar ou rejeitar um fluxo
- Baseia-se em parâmetros predefinidos - especificações de fluxo

# Técnicas para melhorar QoS – Reserva de Recursos

---

- Um fluxo de dados precisa de recursos como:
  - ✓ buffers,
  - ✓ largura de banda
  - ✓ tempo de CPU, etc
- QoS é significativamente melhorada se esses recursos forem reservados de antemão

# Serviços Integrados

---

- IP: entrega de dados *best effort*
- Cada usuário recebe o mesmo nível de atendimento
- **Serviços Integrados (IntServ) são um modelo de QoS baseado em fluxos desenvolvido para o IP**

# Serviços Integrados

---

- IP: protocolo de comutação de pacotes e datagramas sem conexão
- **RSVP (*Resource Reservation Protocol*, protocolo de reserva de recursos)**
- Especificação de fluxo: Rspec (especificação de recursos) e Tspec (especificação de tráfego)

# Serviços Integrados

---

- Foram definidas duas classes de serviços para os Serviços Integrados
  - Serviços garantidos: projetado para tráfego em tempo real, que precisa de um atraso mínimo garantido de ponta a ponta
  - Serviços controlados por carga: projetado para aplicações que podem aceitar atrasos, mas que são sensíveis a uma rede sobrecarregada e com perigo de perda de pacotes.  
Ex.: transferência de arquivos, e-mail e acesso à Internet

# RSVP

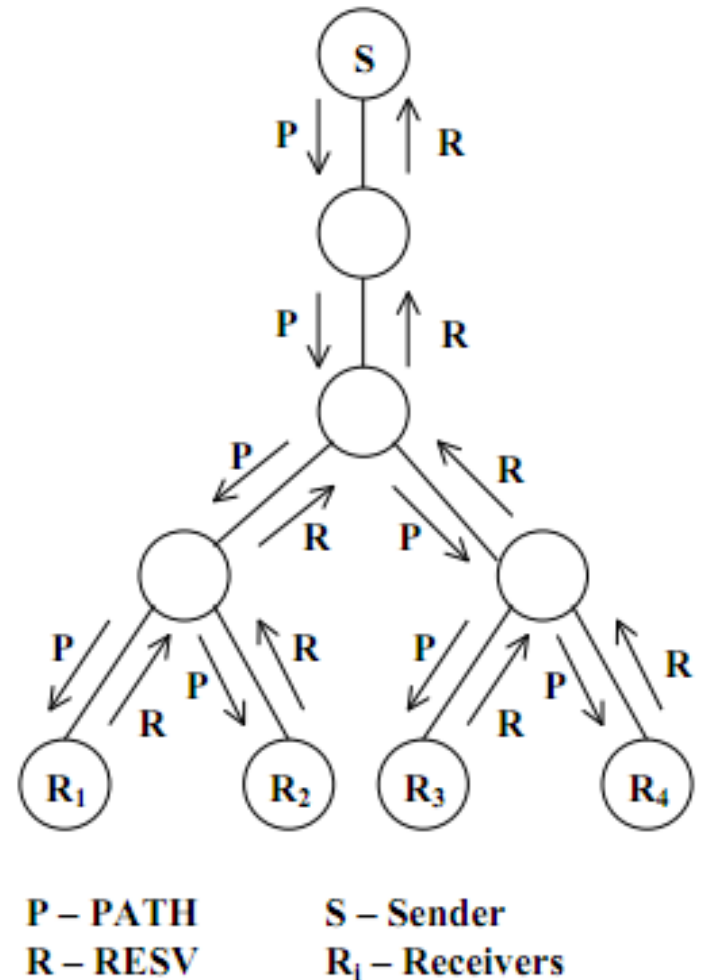
---

- Receptores fazem a reserva
  - Mensagem Path: parte do emissor e atinge todos os receptores de uma rota multicast. No caminho, a mensagem Path armazena as informações necessárias para os receptores
  - Mensagem Resv: trafega em direção ao emissor e faz uma reserva de recursos nos roteadores que suportam RSVP



# RSVP

- Reservas de recursos do RSVP:
  - Emissores enviam mensagens PATH que fornecem aos receptores informações sobre as características do tráfego e do percurso  
  
→ → PATH → →
  - Receptores solicitam reservas de recursos através de mensagens RESV no percurso inverso das mensagens PATH  
  
← ← RESV ← ←



# Serviços Diferenciados - DiffServ

---

- DiffServ ou Serviços Diferenciados é um método utilizado na tentativa de conseguir qualidade de serviço (QoS) em grandes redes, como a Internet
- Negociação para todos os pacotes de dados com base nos chamados acordos de nível de serviço (SLA)
- **Os Serviços Diferenciados são um modelo de QoS baseado em classes, projetado para o IP**

# Serviços Diferenciados - DiffServ

---

- Acordos especificam quais as classes de tráfego serão servidas, as garantias de cada classe e o volume de dados
- Utiliza o campo ToS (*Type of Service*) do cabeçalho IPv4 e *Traffic Class* do cabeçalho IPv6, para atribuir diferentes prioridades aos pacotes

# Serviços Diferenciados - DiffServ

---

- Comportamento por salto – PHB (*Per-Hop Behaviour*)
- DE PHB: entrega *best effort*
- EF PHB: (*Expedited forwarding PHB*), conexão virtual da origem até o destino
- AF PHB (*Assured Forwarding PHB*) entrega o pacote com elevada garantia, desde que o tráfego na classe não exceda o perfil de tráfego do nó

# Serviços Diferenciados - DiffServ

