

Aula 20 - Multimídia: QoS, DiffServ

Diego Passos

Universidade Federal Fluminense

Redes de Computadores II

Na Última Aula...

- RTP: *Real-Time Protocol*.
 - Define **estrutura** de pacote de áudio, vídeo.
 - *Timestamp*, # de sequência, codificação, ...
 - Roda **sobre UDP**.
 - Padrão, permite **interoperabilidade**.
 - **Não provê garantias de entrega**.
- RTCP: *Real-Time Control Protocol*.
 - Trabalha **em conjunto** com o RTP.
 - Pacotes de controle são enviados **periodicamente**.
 - Transmissores e receptores.
 - Estatísticas, informações ajudam em **sincronização, adaptação**.
- SIP: *Session Initiation Protocol*.
 - **Configuração** de chamadas.
 - **Localização** do destinatário.
 - Acordo sobre **codificações**.
 - **Gerenciamento** de chamadas.
 - Adição de **novos fluxos**.
 - Alteração de **codificação**.
 - Convide a **novos** usuários.
 - **Transferência** de chamadas.
 - **Chamada em espera**.
 - Utiliza servidores para:
 - Proxy.
 - Registro de localização de usuários.
 - Simples, mensagens de texto, filosofias da web.

Suporte da Rede às Aplicações Multimídia

Suporte da Rede às Aplicações Multimídia

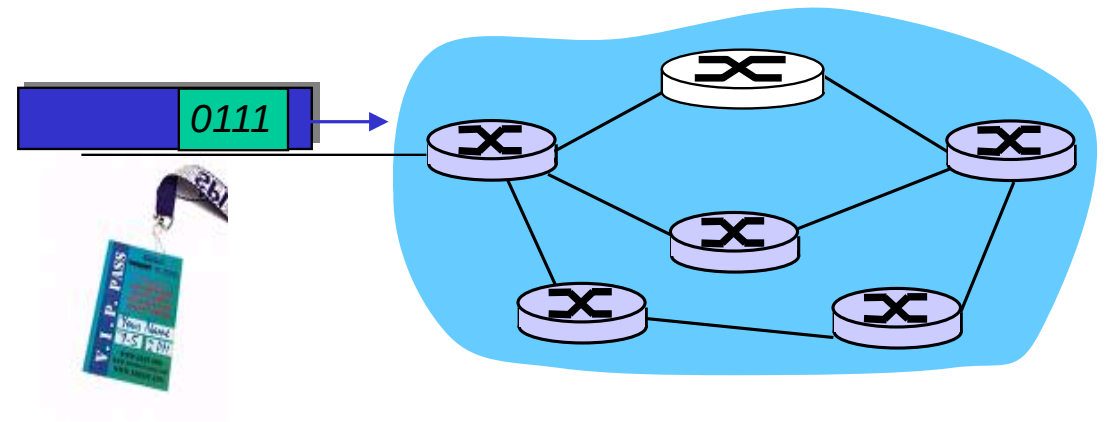
Abordagem	Granularidade	Garantias	Mecanismos	Complexidade	Implantação?
Melhor esforço	Todo tráfego tratado igualmente	Nenhuma ou soft	Nenhum suporte da rede (tudo feito pelas aplicações)	Baixa	Ubíqua
<i>Differentiated services</i>	“Classes” de tráfego	Nenhuma ou soft	Marcação de pacotes, escalonamento, políticas	Média	Alguma
QoS por conexão	Fluxos por conexão	Hard ou soft após admissão do fluxo	Marcação de pacotes, escalonamento, políticas, admissão de fluxos	Alta	Pouca ou nenhuma

Dimensionamento de Redes de Melhor Esforço

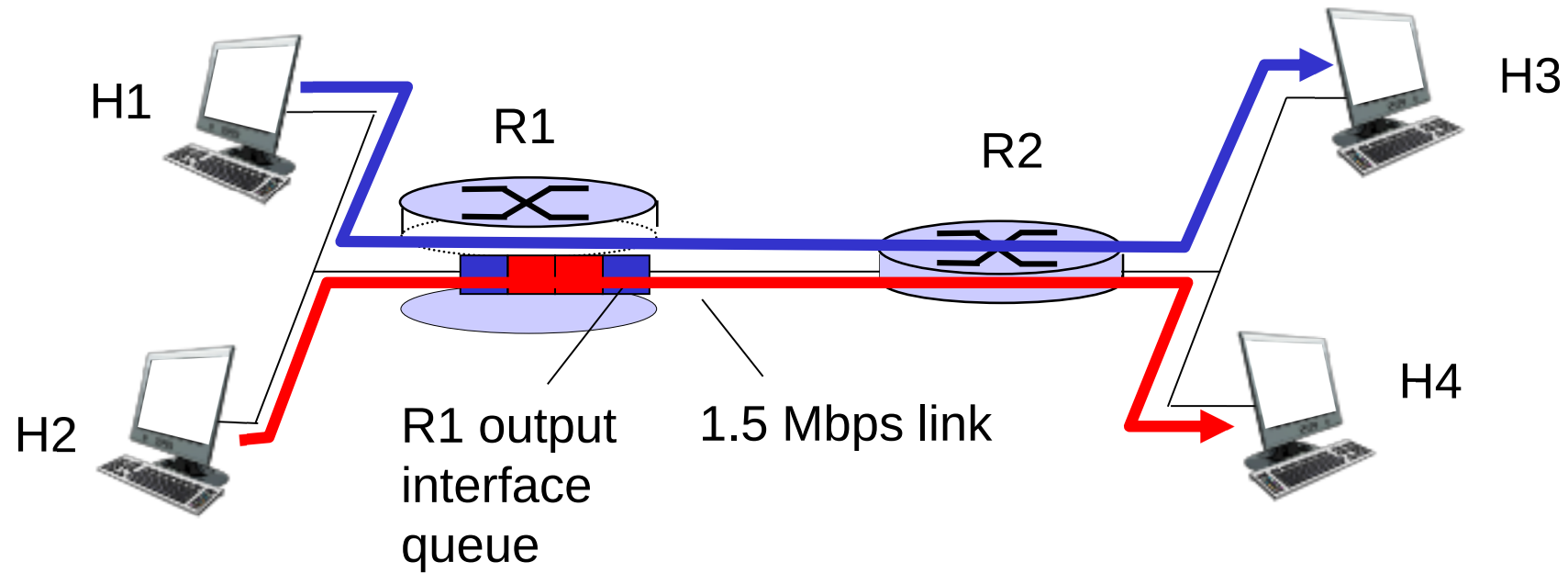
- **Abordagem:** implantar capacidade suficiente para os enlaces de forma que congestionamentos não ocorram; tráfego multimídia percorre a rede sem atrasos ou perdas.
 - Baixa complexidade de mecanismos de rede (utiliza modelo atual de melhor esforço).
 - Altos custos em termos de banda necessária.
- Desafios:
 - **Dimensionamento da rede:** quanta banda é “suficiente”?
 - **Estimativa da demanda de tráfego:** necessária para se determinar o quanto de banda é “suficiente” (para aquele tráfego).

Provendo Múltiplas Classes de Serviço

- Até aqui: tiramos o máximo do serviço de melhor esforço.
 - Modelo de serviço do tipo “tamanho único”.
- Alternativa: múltiplas classes de serviço.
 - Particionamento do tráfego em classes.
 - Rede trata classes de tráfego diferentes de formas diferentes (analogia: serviço VIP vs. serviço normal).
- Granularidade:
 - Serviços diferenciados entre classes, **mas não entre conexões individuais**.
- História: bits ToS.

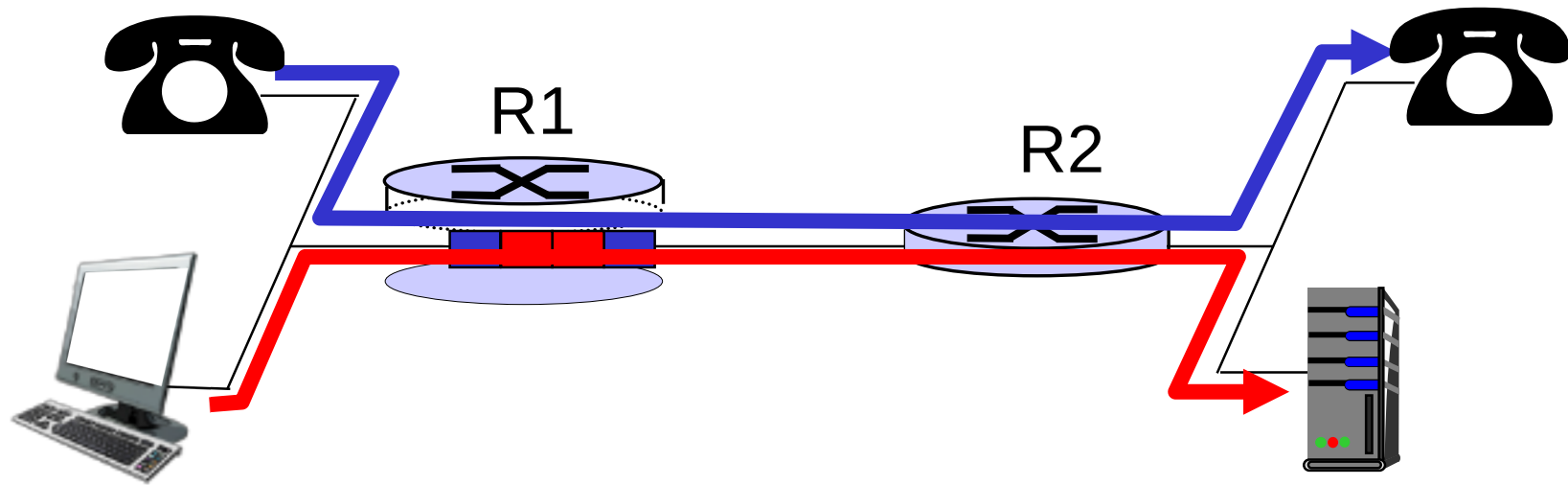


Múltiplas Classes de Serviço: Cenários



Cenário 1: Misturando HTTP e VoIP

- Exemplo: VoIP a 1 Mb/s e HTTP compartilham enlace de 1,5 Mb/s.
 - Rajadas do HTTP podem congestionar roteador, causando perda do áudio.
 - Queremos priorizar o áudio sobre o HTTP.

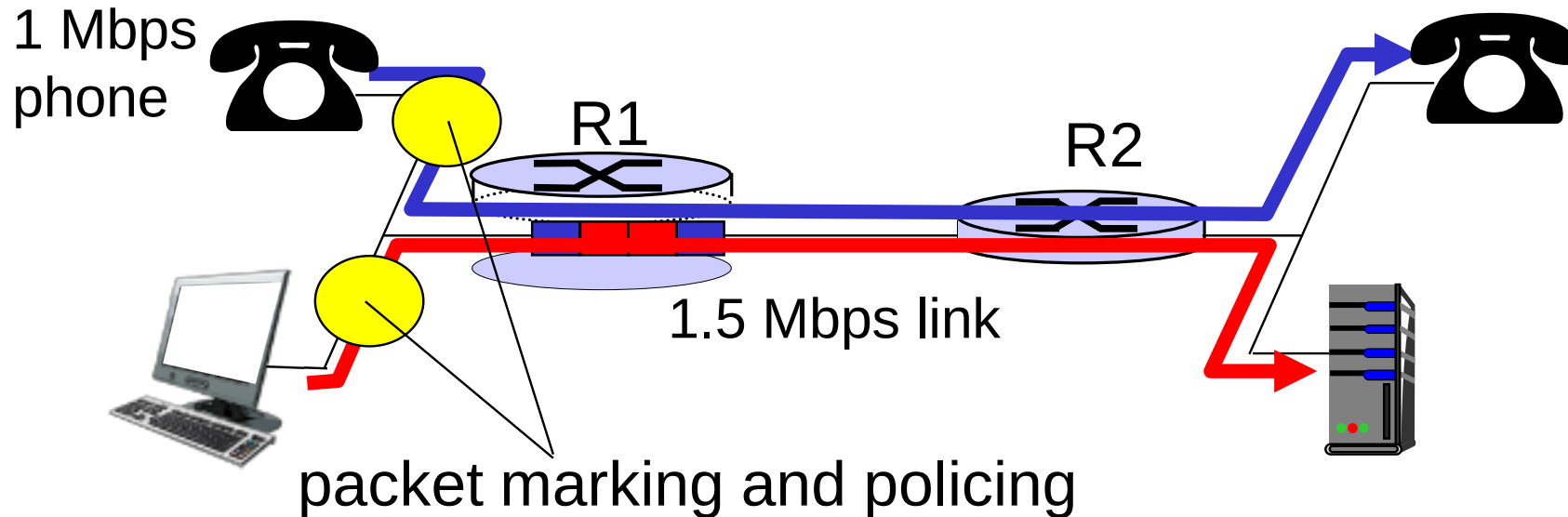


- Primeiro princípio:

Marcação de pacotes é necessária para que roteador diferencie pacotes de classes diferentes. Políticas diferentes também são necessárias para tratar pacotes de acordo.

Mais Princípios para Garantia de QoS (I)

- O que acontece se aplicações se comportam de forma inesperada (e.g., VoIP gera taxa maior que a declarada)?
 - Regulação: forçar fonte a aderir às alocações de banda.
- **Marcação e regulação** nas bordas da rede.

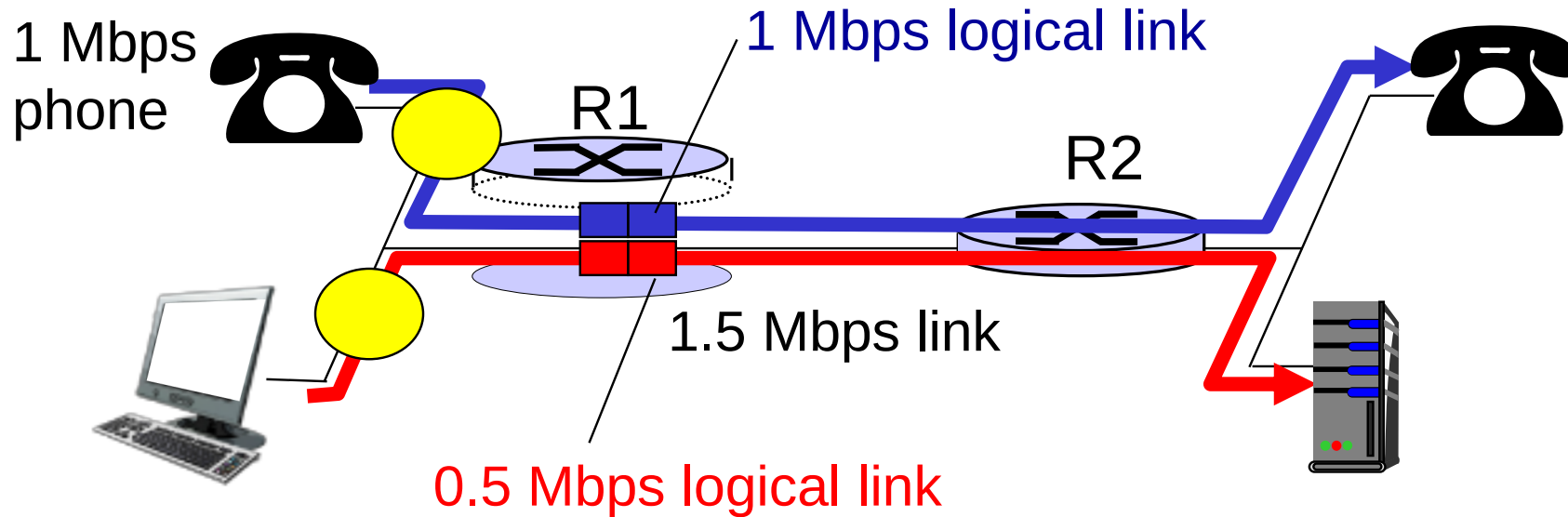


- Segundo princípio:

Prover proteção (isolamento) de uma classe em relação às demais.

Mais Princípios para Garantia de QoS (II)

- Alocação fixa (não compartilhável) de banda para fluxos: ineficiente, se fluxos não utilizam suas alocações.

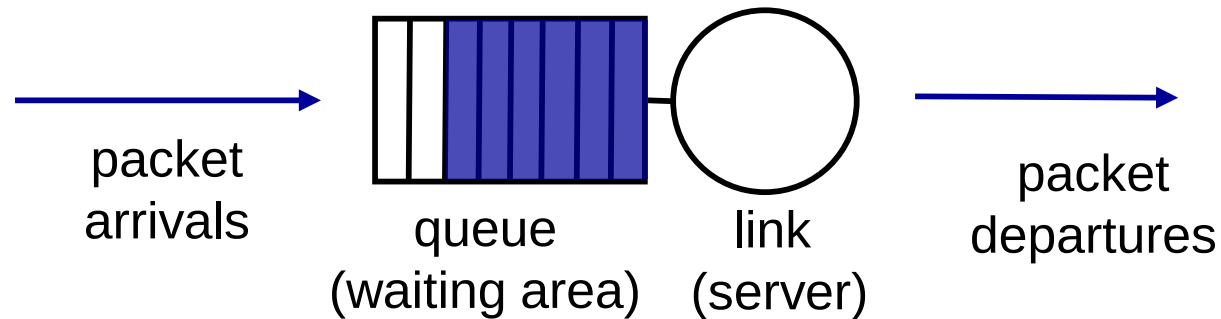


- Terceiro princípio:

Ao mesmo tempo em que se provê isolamento, é desejável utilizar os recursos da forma mais eficiente possível.

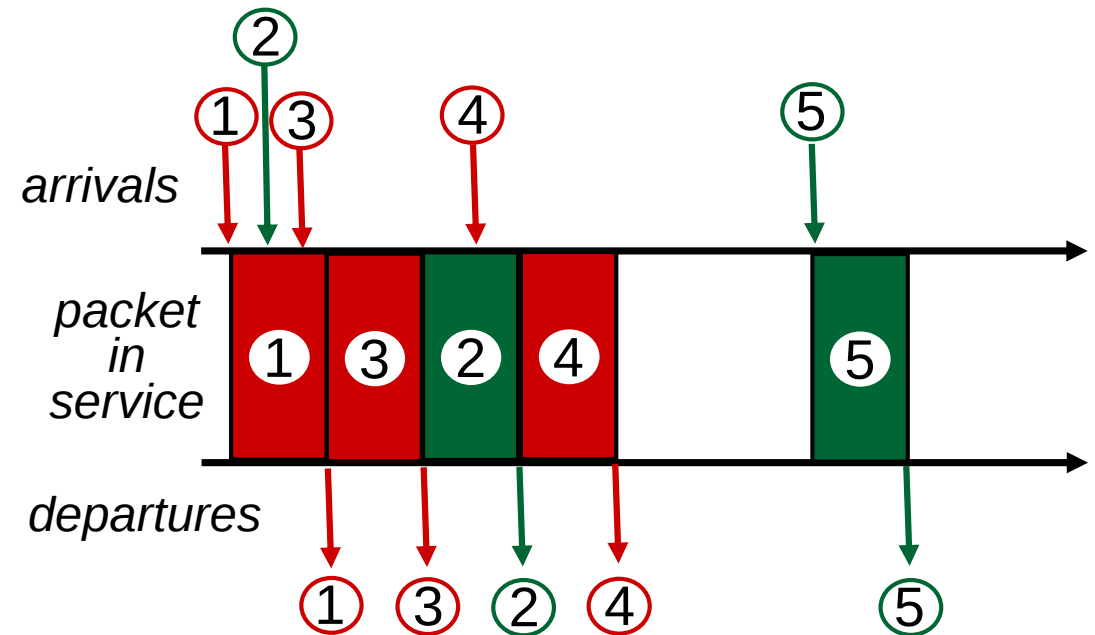
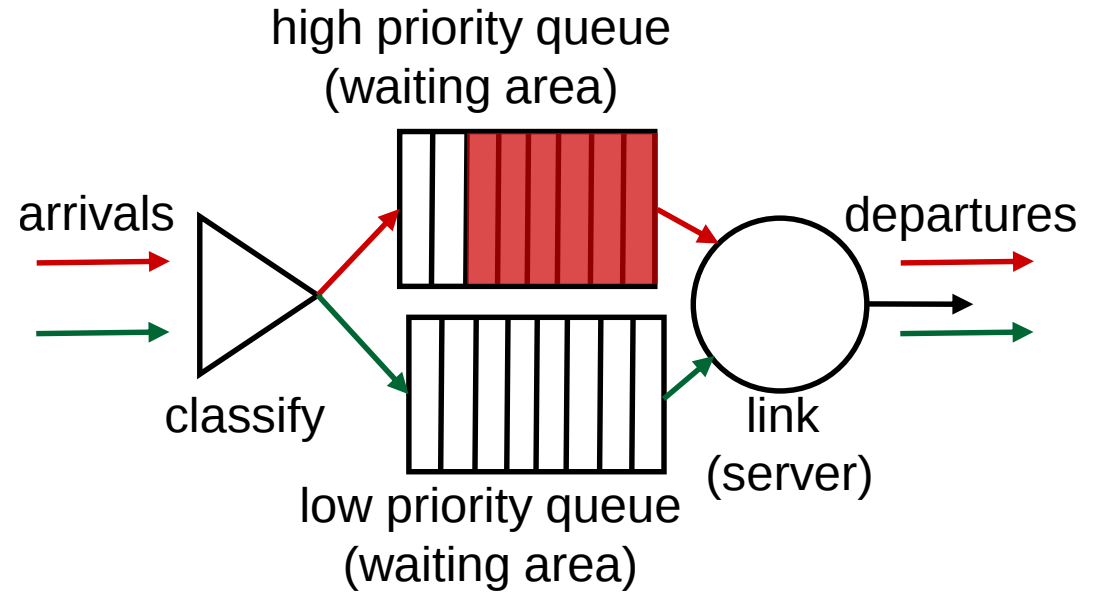
Mecanismos de Escalonamento e Regulação

- **Escalonamento:** escolha do próximo pacote a enviar pelo enlace.
- **FIFO (First In First Out):** envia na ordem de chegada à fila.
 - Exemplo desta política no mundo real?
 - **Política de descarte:** se pacote chega a uma fila cheia, quem é descartado?
 - **Tail Drop:** descarta pacote que acaba de chegar.
 - **Prioridade:** descarta com base em prioridades.
 - **Aleatório:** descarta aleatoriamente.



Políticas de Escalonamento (I)

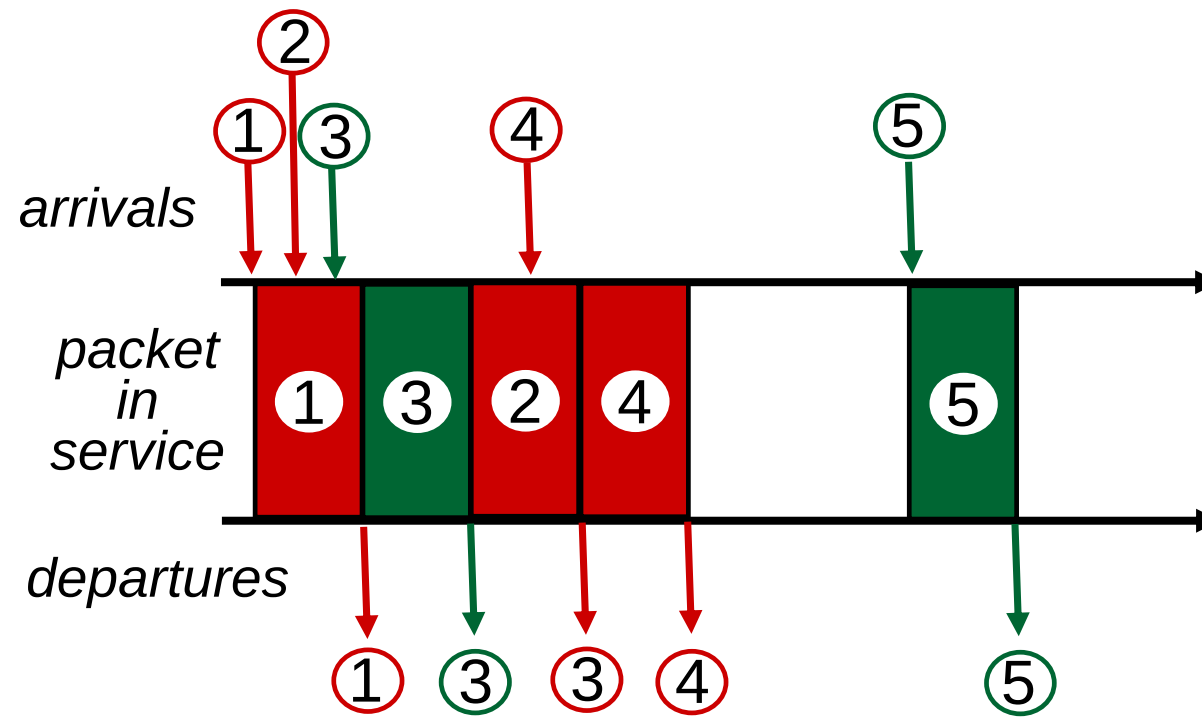
- **Priority Scheduling:** envia pacote enfileirado de mais alta prioridade.
 - Múltiplas **classes** com diferentes prioridades.
 - Classe pode depender de uma marcação ou outra informação de cabeçalho, *e.g.*, IP de origem/destino, número de porta, etc.
 - Exemplo do mundo real?



Políticas de Escalonamento (II)

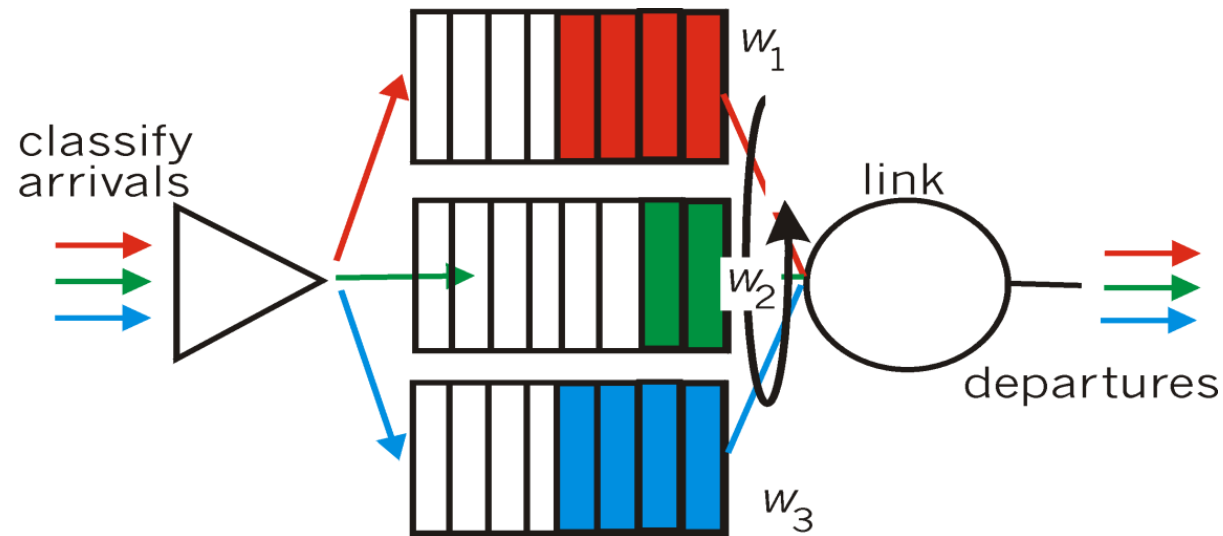
- **Escalonamento Round Robin (RR):**

- Múltiplas classes.
- Varre as filas das classes ciclicamente, enviando um pacote completo de cada classe (se disponível).
- Exemplo do mundo real?



Políticas de Escalonamento (III)

- *Weighted Fair Queuing (WFQ)*:
 - Round Robin generalizado.
 - Cada classe recebe uma fração ponderada de serviço em cada ciclo.
 - Exemplo do mundo real?

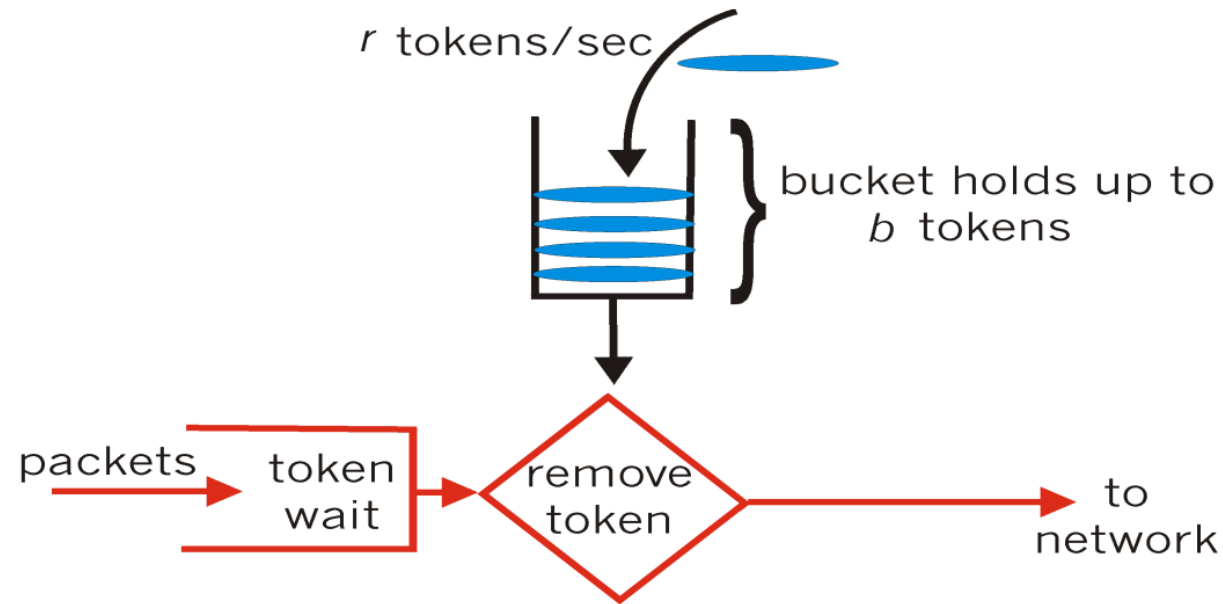


Mecanismos de Regulação

- **Objetivo:** limitar tráfego para que este não exceda parâmetros declarados.
- Três critérios comumente utilizados:
 - **Taxa média (de longo prazo):** quantos pacotes podem ser enviados por unidade de tempo (a longo prazo).
 - Questão fundamental: qual é o comprimento do intervalo considerado? 100 pacotes/s ou 6000 pacotes/min resultam na mesma média!
 - **Taxa máxima:** *e.g.*, 6000 pacotes/min em média; 1500 pacotes/s de taxa máxima.
 - **Tamanho (máximo) da rajada:** número máximo de pacotes enviados em sequência (sem períodos intermediários de inatividade).

Mecanismos de Regulação: Implementação

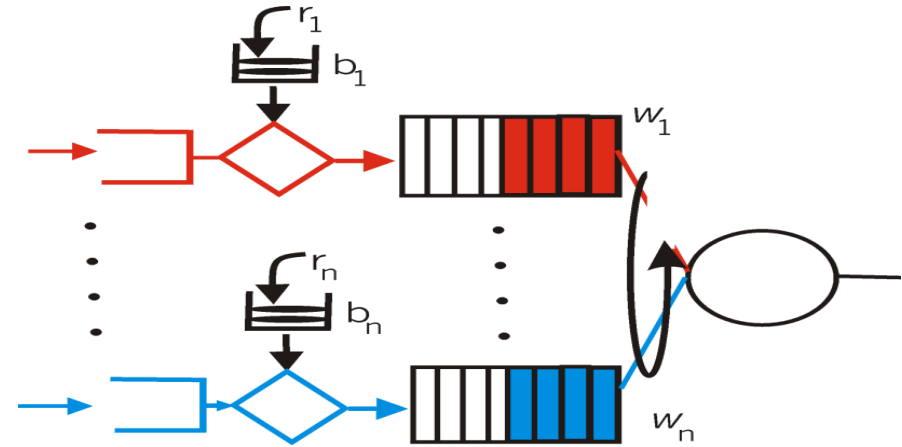
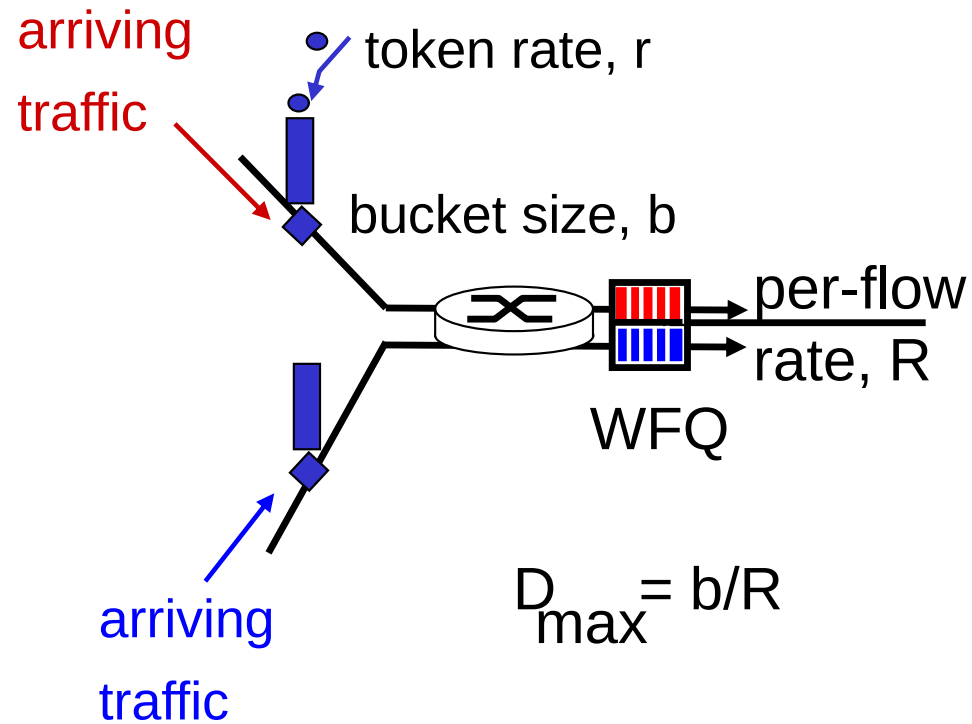
- *Token Bucket*: limita chegada a um **tamanho de rajada** e uma **taxa média**.



- “Balde” pode armazenar b tokens.
- Tokens são gerados a taxa de r tokens/s (a menos que o balde esteja cheio).
- Durante um intervalo de comprimento t : número de pacotes admitidos é menor ou igual a $(r \cdot t + b)$.

Regulação e Garantias de QoS



- Token Bucket e WFQ combinados proveem garantia de um limite superior no atraso.
 - i.e., **garantia de QoS!**

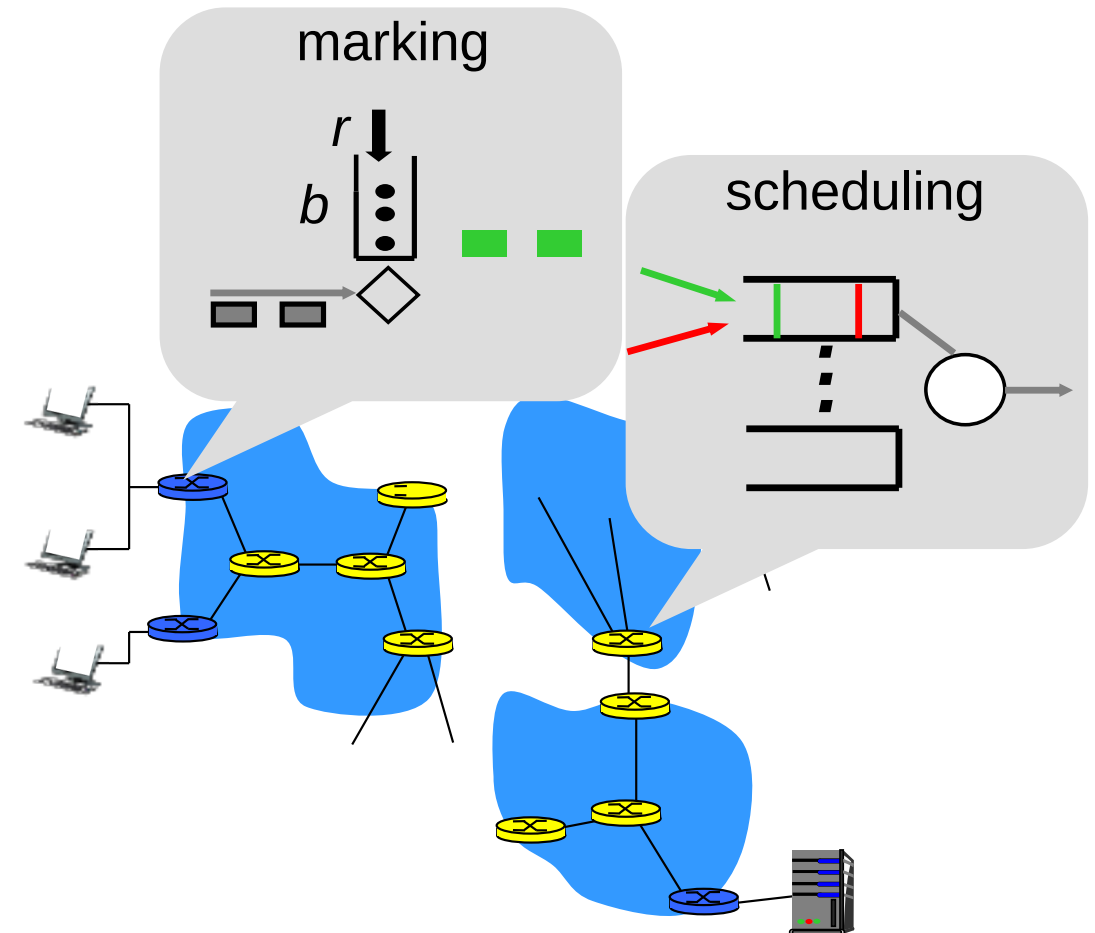


Differentiated Services

- Desejamos classes de serviço “qualitativas”.
 - “Se comporta como um cabo”.
 - Diferenciação relativa de serviços: Platinum, Gold, Silver.
- **Escalabilidade**: funções simples no núcleo da rede, funções complexas nas bordas (ou hosts).
 - Sinalização, manutenção de estado do roteador por fluxo são difíceis com muitos fluxos.
- Não define classes de serviço, mas provê componentes funcionais para construí-las.

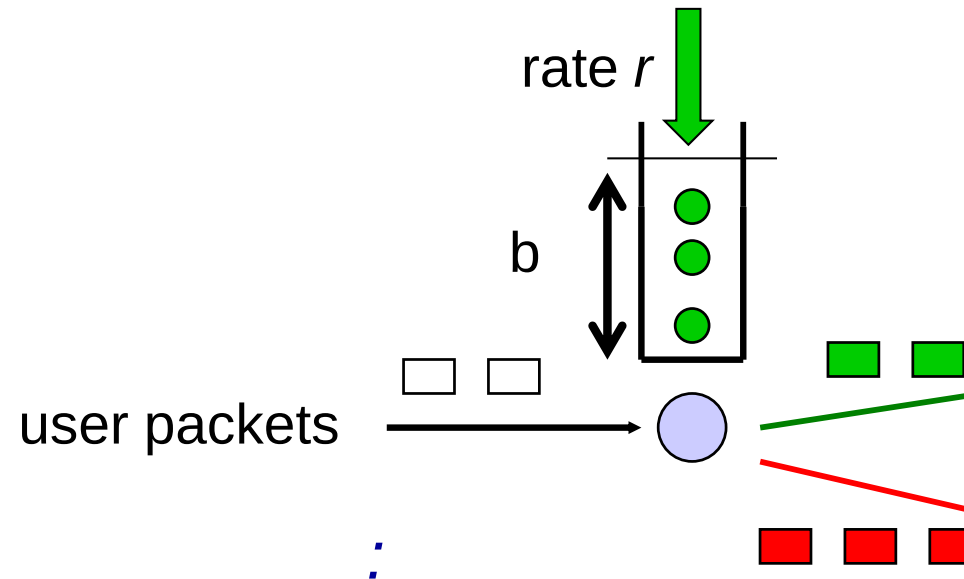
Arquitetura Diffserv

- Roteador de borda: 
 - Gerenciamento de tráfego **por fluxo**.
 - Marca pacotes como *in-profile* ou *out-of-profile*.
- Roteador de núcleo: 
 - Gerenciamento de tráfego por **classe**.
 - Bufferização e escalonamento baseados nas marcações feitas na borda da rede.
 - Preferência dada a pacotes *in-profile*, em relação aos *out-of-profile*.



Marcação de Pacotes no Roteador de Borda

- **Perfil:** taxa pré-negociada r , tamanho do *bucket* b .
- Marcação de pacote na borda baseada em perfil **por fluxo**.



- **Possíveis usos da marcação:**

- Marcação baseada em classe: pacotes de diferentes classes recebem diferentes marcações.
- Marcação intra-classe: porção do fluxo em conformidade com as garantias recebe marcação diferente da porção que não está em conformidade.

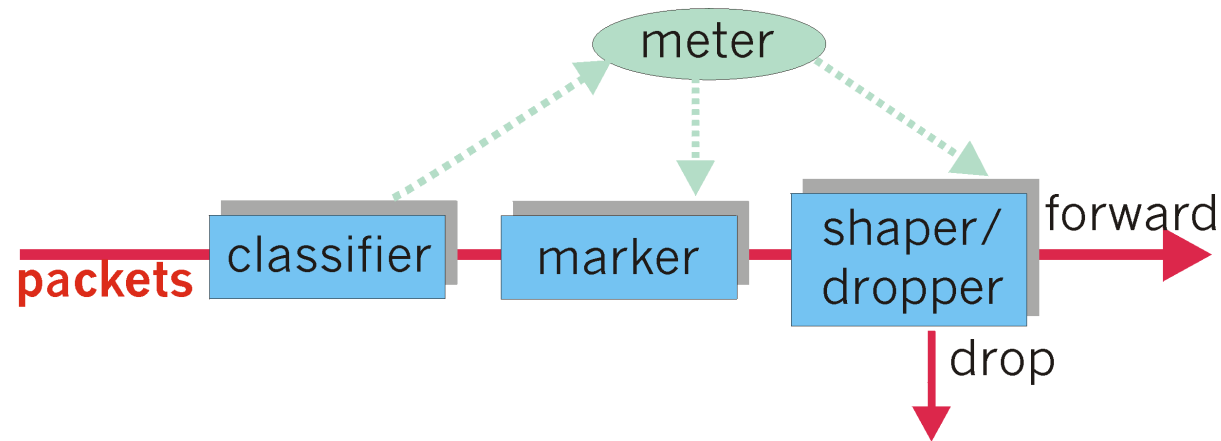
Marcação de Pacotes no Diffserv: Detalhes

- Pacote é marcado no campo *Type of Service* (TOS) do IPv4 e *Traffic Class* no IPv6.
- 6 bits são usados para o *Differentiated Service Code Point* (DSCP).
 - Determina PHB que o pacote receberá.
 - Atualmente, 2 bits não são usados.



Classificação, Condicionamento

- Pode ser desejável limitar a taxa de injeção de tráfego de alguma classe:
 - Usuário declara perfil do tráfego (e.g., taxa, tamanho da rajada).
 - Tráfego é medido e formatado se não está em conformidade com o perfil.



Per-Hop Behavior (PHB)

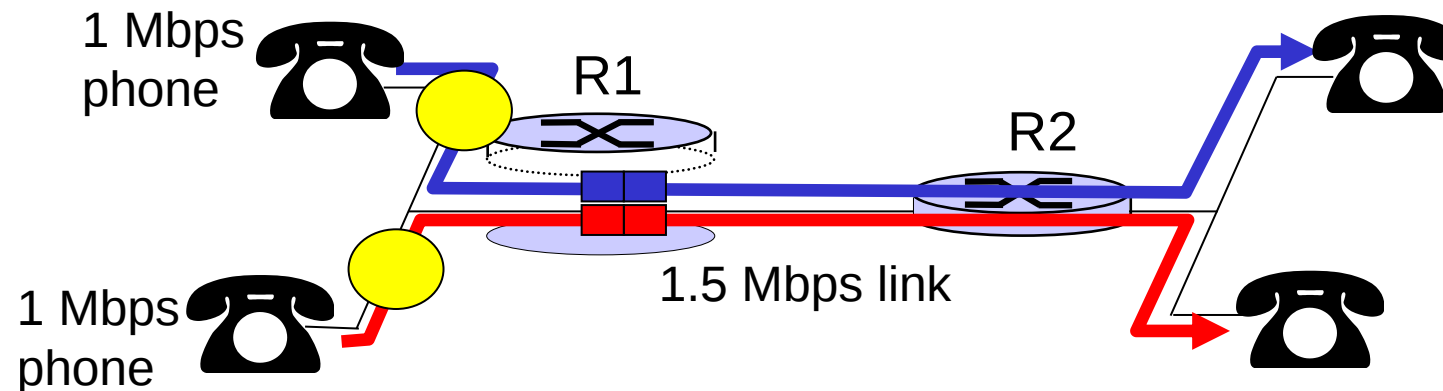
- Resulta em diferenças no comportamento observável do desempenho do encaminhamento.
- PHB **não** especifica quais mecanismos devem ser usados para garantir o desempenho.
- Exemplos:
 - A classe A ganha x% da banda do enlace de saída durante intervalos de comprimento especificado.
 - Pacotes da classe A devem ser transmitidos antes dos pacotes da classe B.

Encaminhamento PHB

- PHBs propostas:
 - **Expedited Forwarding:** taxa de envio de pacotes de uma classe é igual ou maior que determinada taxa.
 - Enlaces lógicos com uma taxa mínima garantida.
 - **Assured Forwarding:** 4 classes de tráfego.
 - Cada uma garante uma banda mínima.
 - Cada uma com três partições de preferências de descarte.

Garantias de QoS por Fluxo

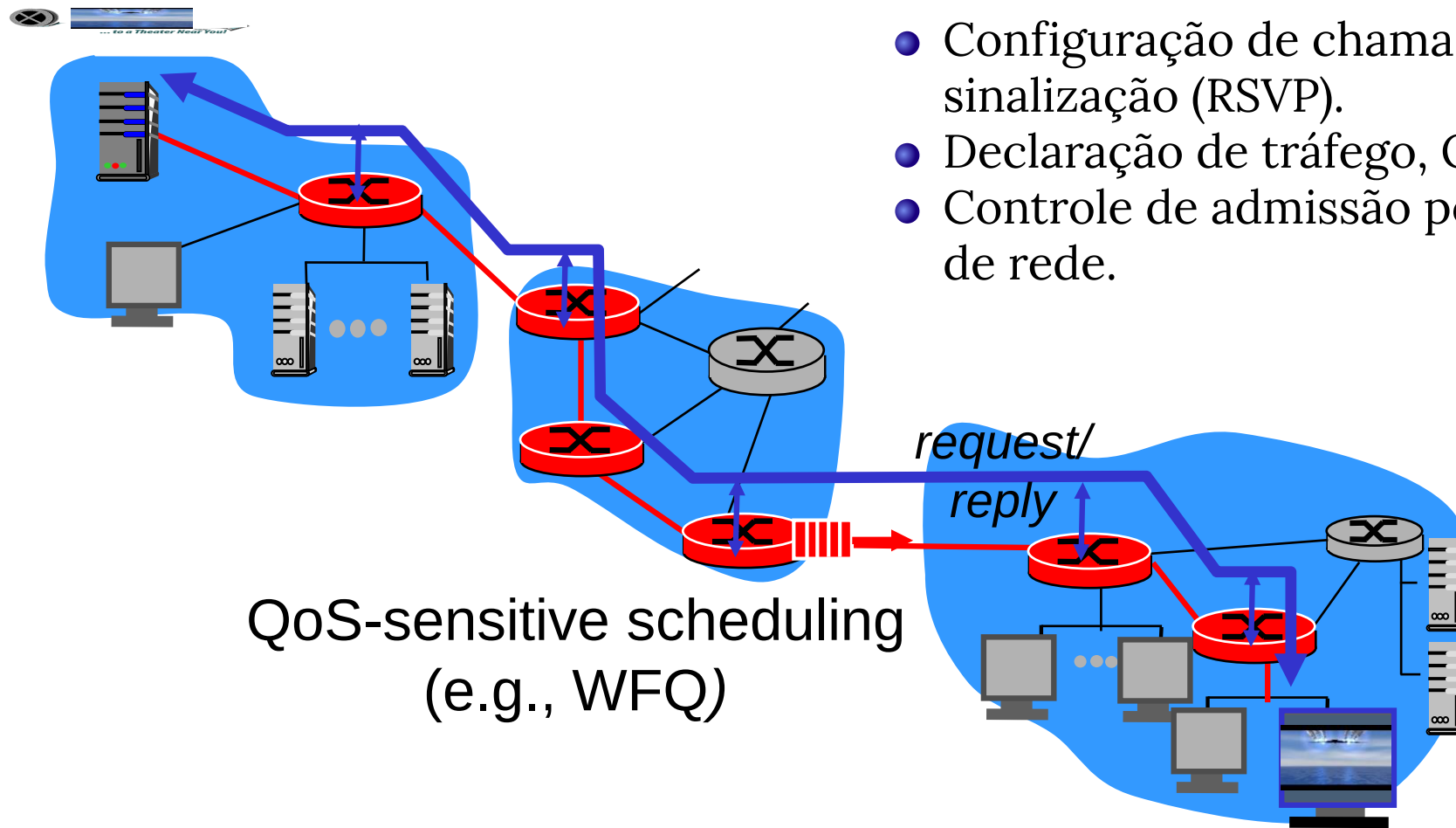
- **Fato básico da vida:** não é possível suportar demanda de tráfego superior à capacidade dos enlaces.



- Quarto princípio:

Controle de Admissão: fluxo declara suas necessidades, rede pode bloquear a chamada (e.g., sinal de ocupado) se não é capaz de atender à demanda.

Cenário de Garantia de QoS



- Reserva de recursos:

- Configuração de chamada, sinalização (RSVP).
- Declaração de tráfego, QoS.
- Controle de admissão por elemento de rede.

Resumo da Aula (I)...

- Dimensionamento:
 - Criar **capacidade** compatível com a **demanda**.
 - Mas quanto é **suficiente**?
- Classes de Serviço:
 - Tráfego é dividido em **classes**.
 - Classes recebem **tratamentos diferentes**.
 - **Escalabilidade**.
- Marcação de pacotes:
 - Pacotes recebem marcas.
 - Identificação de classe.
- Isolamento:
 - Classes não devem se interferir.
 - Mas **recursos não usados devem ser aproveitados**.
- Escalonamento:
 - Política de escolha de pacotes para uso do enlace.
 - FIFO, *Priority Scheduling*, Round Robin, WFQ, ...
- Mecanismos de regulação:
 - Garantem que tráfego atende parâmetros declarados.
 - e.g., *Token Bucket*.

Resumo da Aula (II)...

- Diffserv: arquitetura para diferenciação de serviços.
 - **Escalabilidade:** maior esforço nas bordas.
 - Marcação de pacotes, condicionamento.
 - Roteadores de núcleo: obedecem ao PHB.
 - Aplicam políticas de compartilhamento de banda, *buffer*.
 - Políticas diferentes para classes diferentes.
- Garantias de QoS por fluxo: necessita de **controle de admissão**.
 - Garantir que rede/enlace possui **capacidade suficiente** para atender a **todos os fluxos**.
 - Configuração de chamada.
 - Cada elemento da rede deve prover garantias.

Leitura e Exercícios Sugeridos

- Classes de serviço, escalonamento e regulação:
 - Páginas 464 a 474 do Kurose (Seção 7.5 até Subseção 7.5.2, inclusive).
 - Exercícios de fixação 13 e 15 do capítulo 7 do Kurose.
 - Problemas 20 a 27 do capítulo 7 do Kurose.
- Diffserv, QoS por fluxo:
 - Páginas 474 a 483 do Kurose (Final da Seção 7.5 e Seção 7.6).
 - Exercício de fixação 16 do capítulo 7 do Kurose.

Próxima Aula...

- Iniciamos o último capítulo da disciplina: gerência.
- Na próxima aula:
 - Conceitos básicos.
 - O protocolo SNMP.
 - MIBs.