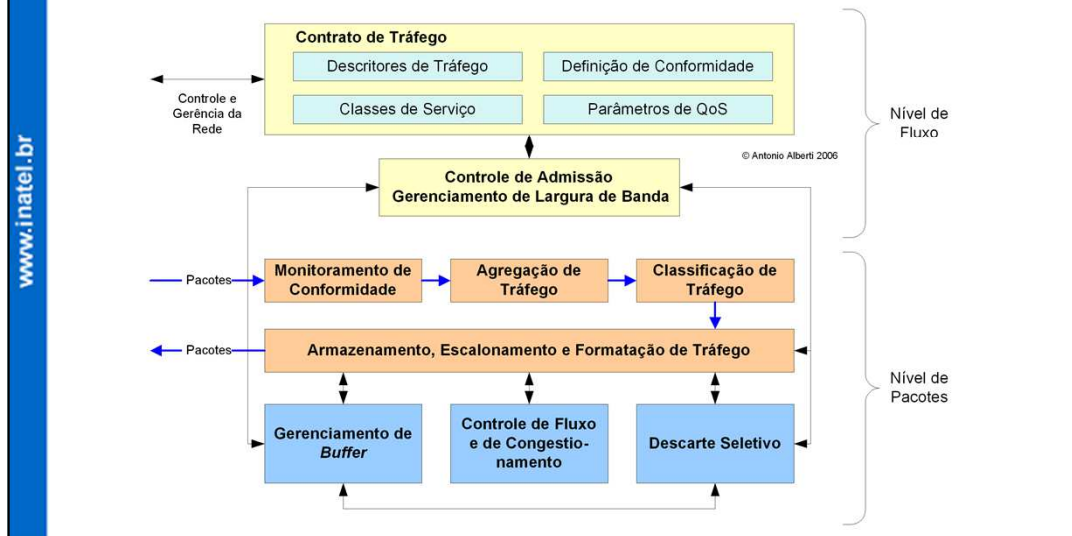


- QoS EM REDES DE PACOTES

Introdução

- Formas mais adequadas de provimento de QoS.



O contrato de tráfego é um acordo que define o comportamento do tráfego e o nível de serviço requerido por cada fluxo de tráfego, seja ele individual ou agregado. Pode fazer parte do acordo de nível de serviço (SLA – *Service Level Agreement*). O SLA é um acordo formal (contrato) assinado entre o provedor de serviço e o cliente deste serviço. Um elemento chave do contrato de tráfego são as classes de serviço que podem ser utilizadas.

Cada classe de serviço agrupa os pré-requisitos de qualidade de serviço e as características do fluxo de tráfego que será transportado, bem como a forma como este será policiado e/ou formatado. O comportamento esperado de cada fluxo de tráfego é especificado através de descritores de tráfego específicos para cada classe de tráfego. O contrato de tráfego é utilizado pela gerência ou sinalização da rede para diversas funções no nível de fluxo e de rede.

No nível de fluxo, pode ser utilizado para controle de admissão e gerenciamento de largura de banda. O controle de admissão avalia se um novo fluxo (ou conexão) pode ser admitido na rede sem afetar a QoS dos demais fluxos existentes. O controle de admissão pode negar a aceitação de um novo fluxo, caso a rede não possua recursos para atendê-lo. Vale ressaltar que a maior parte das tecnologias não fazem este controle.

O gerenciamento de largura de banda trabalha otimizando o uso dos recursos de transmissão. Em algumas tecnologias este gerenciamento trabalha em conjunto com o controle de admissão. Para garantir que a QoS seja mantida para todos os fluxos (ou conexões) da rede, esta deve verificar se o tráfego submetido está de acordo com o negociado no contrato de tráfego. Esta verificação é feita por funções de monitoramento de conformidade.

O tráfego do fluxo (ou conexão) que satisfaz as funções de monitoramento de conformidade é então enviado pela rede, onde deverá passar por vários pontos de multiplexação. Nestes pontos de multiplexação, a comutação de pacotes é utilizada, e estruturas de filas (ou *buffers*) são utilizados para armazenar temporariamente os pacotes que aguardam por transmissão. As estruturas de filas são servidas de acordo com um algoritmo de escalonamento, que deve ser projetado para atender aos pré-requisitos de QoS de todos fluxos.

Funções de controle de congestionamento são utilizadas para evitar e gerenciar situações de congestionamento na rede. Embora o controle de admissão verifique se novos fluxos podem ser aceitos, situações de congestionamento nos componentes da rede ainda podem acontecer. O controle de congestionamento atua reduzindo o tráfego das fontes a medida que o nível de congestionamento está aumentando.

Outras funções relacionadas ao controle de congestionamento são:

- Gerenciamento de *Buffer*.
- Descarte Seletivo.

O Gerenciamento de *Buffer* distribui os recursos de armazenamento de forma justa entre os fluxos (ou conexões), de forma a melhorar a QoS na rede. Decide quando pacotes devem ser descartados para evitar ou sair de uma situação de congestionamento. A escolha de quais pacotes devem ser removidos é feita pelo mecanismo de descarte seletivo, em função das orientações providas pelo gerenciamento de *buffer*.

Tem-se ainda as funções de:

- Agregação de Tráfego - **Agrega** o tráfego de várias fontes em um único fluxo de tráfego agregado.

- Classificação de Tráfego - **Classifica** o tráfego entrante (agregado ou não) em função das classes de tráfego existentes.

Introdução

- Mecanismos para gerência de tráfego:
 - Classes de Serviço
 - Parâmetros de QoS
 - Descritores de Tráfego
 - Banda Efetiva
 - Classificação de Tráfego
 - Agregação de Tráfego
 - Controle de Admissão de Conexões
 - Monitoramento de Conformidade
 - Armazenamento e Escalonamento
 - Gerenciamento de *Buffer*
 - Descarte Seletivo

Descritores de Tráfego

- Servem para descrever o tráfego que um determinado cliente está enviando para a rede.
 - Taxa de Pico
 - Taxa máxima da fonte em pacotes por segundo.
 - Taxa Sustentável
 - Taxa média sustentada pela fonte em pacotes por segundo.
 - Tamanho Máximo de Surto
 - Quantidade máxima de pacotes que pode ser transmitida na taxa de pico sem violar a taxa sustentável.
 - Tolerância de Surto
 - Tempo em que a fonte pode transmitir a uma taxa igual a de pico.
 - Taxa Mínima

Controle de Admissão

- Determina se um novo fluxo pode ou não ser aceito na rede.
 - Para tanto é necessário que em cada equipamento um **algoritmo de controle de admissão de conexões (CAC)** seja **consultado**, de forma a verificar se os pré-requisitos de QoS desejados por este fluxo podem ser **aceitos** sem que a QoS dos fluxos já existentes seja **deteriorada**.
- Dependendo das regras utilizadas, os algoritmos de CAC podem ser:
 - **Conservativos** (Alocação de recursos além do necessário).
 - **Agressivos** (Alocação de recursos iguais ou menores que o necessário).

Monitoramento de Conformidade

- Verifica se o tráfego submetido para a rede está de acordo com o contrato de tráfego negociado.
- Para isso, a rede atua sobre os pacotes não conformes por meio da função de policiamento de tráfego (TP – *Traffic Policing*). Dentre os principais algoritmos de policiamento estão:
 - *Leaky Bucket*
 - *Virtual Scheduling*
 - *Token Bucket*

Armazenamento e Escalonamento

- Estruturas de Filas
 - Armazenam os pacotes nos nós da rede.
- Escalonadores – Escolhem a ordem apropriada para retirar os pacotes das estrutura de filas.
 - Escalonamento Baseado em Prioridades (*Priority-based scheduling*).
 - Escalonamento com Divisão Justa (*Fair-share scheduling*).
 - Escalonamento com Regulação de Tráfego (*Traffic shaping*).

Os diversos fluxos de tráfego existentes em uma rede compartilham os recursos desta rede (largura de faixa, armazenamento, etc.) e tentam utilizá-los simultaneamente. Devido ao compartilhamento de recursos, estruturas de filas (*Queueing Structures*) são usadas para armazenar temporariamente os pacotes.

Tipicamente, as estruturas de filas possuem diversas filas divididas em grupos lógicos, cada um dos quais servidos por um escalonador. Exemplos:

Filas por porta de saída, filas por classe de serviço, filas por fluxo individual.

Escalonadores são necessários para extrair pacotes de estruturas de filas e transmiti-los de forma a melhorar ou garantir a QoS na rede. Portanto, um escalonador é um mecanismo de gerenciamento de tráfego responsável pela arbitração entre filas a serem servidas e pela alocação de largura de banda para estas filas. Um escalonador robusto protege os fluxos de tráfego de fontes ou usuários maliciosos (ou mal comportados) sem depender somente da função de policiamento. Ele também permite o uso eficiente da largura de faixa disponível sob qualquer situação de variação de carga na rede.

Controle de Congestionamento

- Os objetivos destes mecanismos são:
 - Maximizar a eficiência do uso dos recursos de armazenamento. Para isso é utilizado um algoritmo de gerenciamento de estrutura de filas (BM – *Buffer Management*).
 - Distribuir os recursos de armazenamento de forma justa entre os fluxos ou conexões, de forma que a QoS de cada fluxo seja respeitada.
 - Prevenir que fluxos afetem a QoS um dos outros.
 - Decidir quais pacotes devem ser descartados quando a ocupação da estrutura de filas ultrapassa um determinado valor. Para isso é utilizado um algoritmo de descarte seletivo de células (SD – *Selective Discard*).

Para evitar que a QoS das conexões da rede seja afetada pelo congestionamento das estruturas de filas, mecanismos de controle de congestionamento precisam ser implementados em cada nó da rede.

O objetivo é administrar de forma eficiente o espaço disponível em uma estrutura de filas e isolar o tráfego destinado a diferentes filas. A eficiência é conseguida através do compartilhamento do espaço físico disponível no maior número possível de filas. Alguns dos algoritmos de gerenciamento oferecem isolamento naturalmente, enquanto outros precisam ser acoplados a algoritmos de descarte seletivo mais inteligentes de maneira a prevenir que uma fila da estrutura utilize mais recursos do que o permitido e acabe estorvando outras filas.

Descarta pacotes para evitar ou reduzir o nível de congestionamento na rede. Exemplos de mecanismos de descarte:

Descarte Baseado na Prioridade dos Pacotes

Descarte Baseado na Classe de Serviço

Descarte Baseado em Parâmetro de QoS

Descarte Aleatório

Descarte Baseado na Ocupação das Filas

- CLASSIFICAÇÃO DE TRÁFEGO.

Classificação de Tráfego

- Identificação do tipo de tráfego transportado por cada pacote.
 - IP PRECEDENCE
 - Arquitetura DiffServ
 - Cabeçalho MAC IEEE 801.D
- Pode ser feita pelas fontes externas, dispositivos de borda ou dispositivos de backbone.

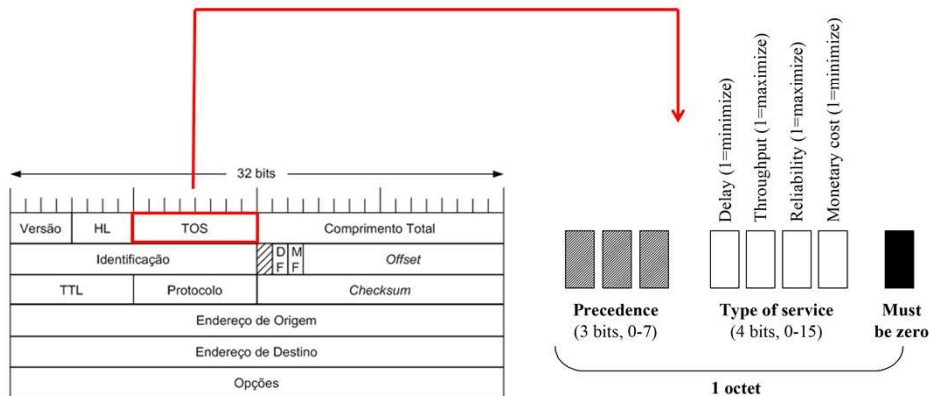
Um dos mecanismos para se alcançar o nível de QoS adequado aos diversos tipos de tráfego em uma rede de pacotes é diferenciar o tratamento que os nós da rede dispensam a cada tipo de tráfego. A classificação dos pacotes, que corresponde à identificação do tráfego transportado por cada pacote, é portanto uma das técnicas fundamentais para se obter QoS em uma rede de pacotes transportando voz.

A classificação do tráfego em si não é uma ferramenta direta para se obter QoS, mas sim uma técnica auxiliar que permitirá a implementação de outras técnicas, como por exemplo políticas de priorização da transmissão ou descarte de determinados tipos de pacote, em função do tráfego transportado.

A classificação do tráfego pode ser feita pacote a pacote (analisando a característica do tráfego de cada pacote) ou sessão a sessão (quando o transmissor negocia uma classificação fim a fim antes de transmitir). A política de classificação dos pacotes é definida pelo operador da rede, e pode se basear em diversos critérios, tais como: tipo de tráfego contido no pacote, endereço da porta física, endereço MAC, endereço IP de fonte ou destino, porta de aplicação, etc.

IP Precedence

- IP Precedence.
 - Referente ao campo ToS (*Type of Service*) do cabeçalho IP.



A classificação dos pacotes pode ser feita pelas fontes de tráfego externas, pelos dispositivos de borda, ou pelos dispositivos de backbone da rede. Quando a classificação é feita pela fonte (ou por uma outra rede à downstream), a rede pode aceitar a classificação recebida ou reclassificar o tráfego de acordo com a sua própria política. Os critérios para a classificação de tráfego podem ser tão amplos quanto “tráfego destinado à rede X” ou tão estreitos quanto “tráfego do fluxo Y”. No backbone da rede a granularidade da classificação tende a ser menor, em função do grande número de fluxos existentes.

Os três primeiros bits do campo ToS do cabeçalho IPv4, denominados bits de Precedência, definem a prioridade relativa do datagrama, estabelecendo até seis classes de serviço distintas. Os quatro bits seguintes, denominados Type of Service, consistem de 4 flags para vários tipos de serviços, e o último bit não é utilizado.

IP Precedence

- A RFC 791 define o significado dos oito possíveis valores dos bits de precedência.

Valor dos bits de Precedência	Tipo de tráfego
000 – 0	Routine
001 – 1	Priority
010 – 2	Immediate
011 – 3	Flash
100 – 4	Flash Override
101 – 5	Critical
110 – 6	Internetwork control
111 – 7	Network control

Embora a RFC 791 defina determinados tipos de tráfego, outras definições proprietárias pode ser feitas. Os valores 6 e 7 mostrados na Tabela 2 são reservados para informações de controle da rede, tais como atualizações de roteamento.

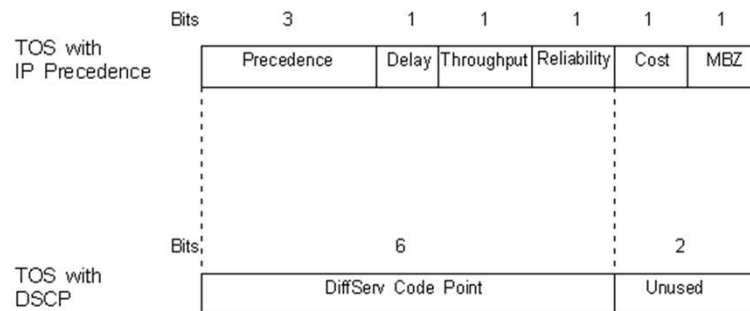
Não há nenhuma definição de como o roteador deve se comportar quando recebe um pacote com uma das prioridades mostradas na tabela. A RFC 791 descreve os bits de Precedência como “uma medida de importância” do pacote.

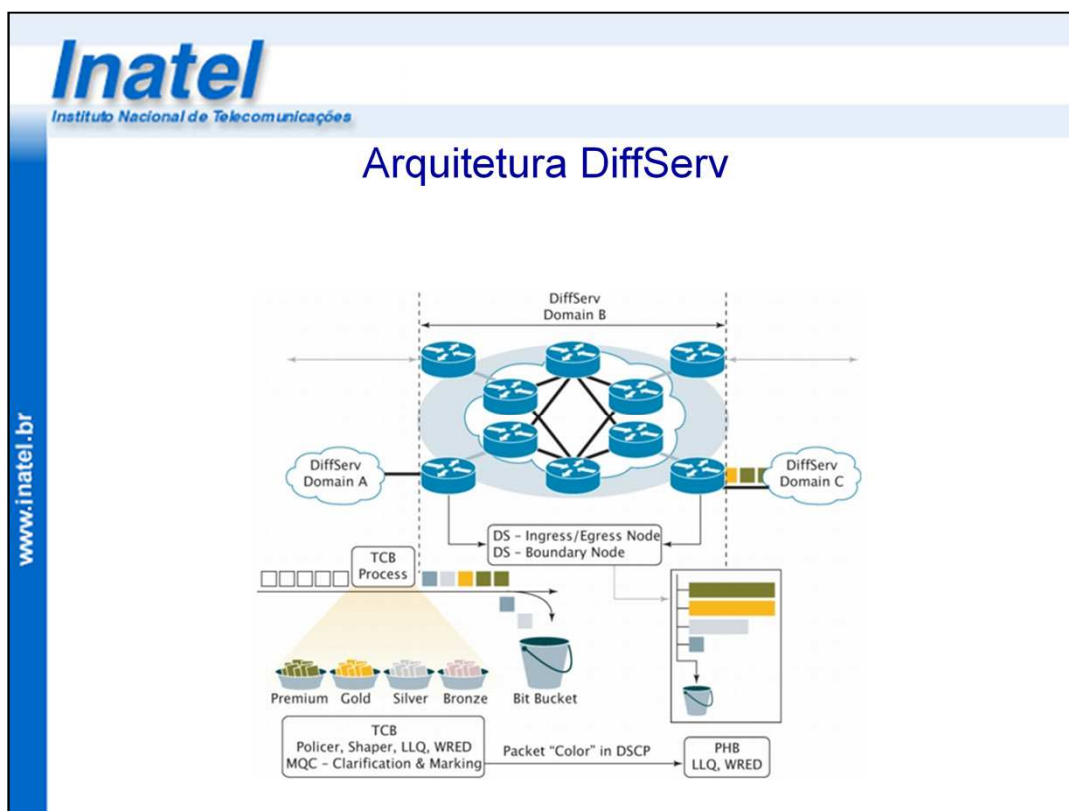
De modo que cada elemento subsequente da rede possa prover o serviço com base na política estabelecida, a Precedência do pacote é definida o mais próximo possível da borda da rede. Podemos ver a definição da Precedência como uma função de borda que permitirá aos dispositivos do núcleo, ou backbone, da rede executarem funções visando obter QoS. Por exemplo, os roteadores no backbone da rede podem utilizar os bits de precedência para determinar a ordem de transmissão ou a probabilidade de descarte dos pacotes.

Além dos bits de precedência, os bits Type of Service também podem ser utilizados para definir o tratamento a ser dispensado ao pacote. Os bits de Precedência usualmente guiam o comportamento por hop (descarte, enfileiramento, prioridade e transmissão), enquanto os bits Type of Service guiam o comportamento inter-rede (seleção de rota, por exemplo).

Arquitetura DiffServ

- A Arquitetura de Serviços Diferenciados (*DiffServ*) irá redefinir o uso dos 8 bits do campo ToS do cabeçalho IP, para indicar aos dispositivos da rede que pacotes devem ser manuseados de forma especial.





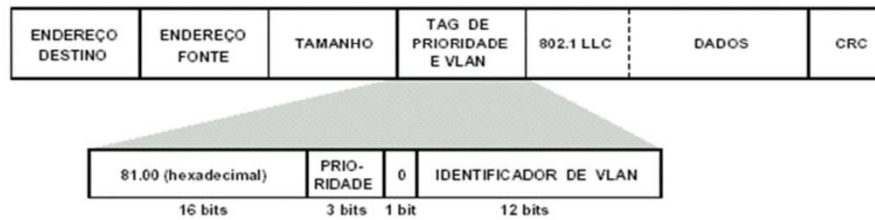
Domínios DiffServ: conjuntos de roteadores onde se aplicam uma determinada classificação e diferenciação de tráfego, definidas pela gerência da rede.

Roteadores de borda: roteadores que classificam fluxos que entram num domínio DiffServ. Essa tarefa de classificação e condicionamento de fluxos de entrada pode ser complexa. Por esses roteadores também passam fluxos que saem do domínio, para serem entregues a sistemas finais.

Roteadores de núcleo: roteadores internos de um domínio DiffServ, que encaminham fluxos já classificados. Esses roteadores verificam a conformidade desses fluxos às regras definidas pela gerência da rede para as classes de serviço. Assim, pacotes desses fluxos estão sujeitos a priorizações, enfileiramentos, e mesmo descartes, dependendo das restrições de QoS que foram definidas. Essas regras compõem o que se chama de *Comportamento por Salto* (PHB - *Per-Hop Behaviour*), representada no diagrama contido na figura acima.

IEEE 802.1.D

- Cabeçalho estendido IEEE 802.1.D



O padrão IEEE 802.1.D, que inclui o 802.1.p e 802.1.Q, define uma extensão de 32 bits ao cabeçalho MAC das redes Ethernet e Token Ring. Doze bits deste espaço são utilizados como rótulos para redes VLAN; contudo, três bits dentro do cabeçalho 802.1.Q (definidos no padrão 802.1.p) são utilizados para sinalização de classe de serviço na rede, permitindo que se estabeleçam oito níveis de prioridades para classificação de tráfego, mas sem informação de Tipo de Serviço para indicar, por exemplo, elegibilidade para descarte.

IEEE 802.1.D

- O campo PRIORIDADE define a priorização de cada VLAN (*Virtual Local Area Network*).

Valor	Nome	Exemplos	Características
7	Controle da rede	RIP, OSPF, BGP4	Crítico para a operação da rede
6	Voz	NetMeeting audio	Sensível a latência e jitter; largura de faixa pequena
5	Vídeo	PictureTel, Indeo	Grande largura de faixa; sensível a jitter.
4	Carga controlada	SNA Transactions	Tempo de resposta previsível, aplicações sensíveis a latência
3	Esforço excelente	SAP, SQL	Tráfego crítico (negócios) que toleram atrasos.
2	Melhor esforço	Melhor esforço	Melhor esforço
1	<default>	<default>	<default>
0	Background	FTP backups	Insensível a latência

O padrão **IEEE802.1p** estende o quadro **IEEE802** de modo a incorporar a priorização de tráfego e o identificador de **LAN** virtual, com um *tag* de **32** bits acrescentado a um quadro IEEE802. Caso o *switch* implemente o padrão **IEEE802.1p** (ele é capaz de decodificar o campo *Ethertype* = **81.00hex**), os bits de priorização e identificação de **VLAN** são processados, caso contrário, o *switch* descarta o quadro.

Tal *switch* **IEEE802.1p** possui um classificador que decide o tipo de tráfego (caso não venha explícito no próprio quadro) baseando-se no tipo de protocolo transportado (em redes **IP**, combinando endereço fonte e destino, e portas **TCP** ou **UDP**). Neste caso, como o *switch* vai inspecionar o tipo de quadro, ele é chamado *switch* **L3** ou *switch* **L4**) e aloca o quadro em uma das setes filas de precedência. A seguir, o escalonador decide que quadro será transmitido de modo a garantir o nível de qualidade adequado ao tipo de tráfego.

- PRIORIZAÇÃO DE TRÁFEGO – ESCALONAMENTO.

- Escalonamento e priorização de tráfego.
 - First-In, First-Out - FIFO
 - Priority Queueing - PQ
 - Weighted Round Robin - WRR
 - Weighted Fair Queueing - WFQ
 - Modified Deficit Round Robin - MDRR

De modo a absorver situações momentâneas de congestionamento na rede, onde a taxa de chegada de pacotes excede a capacidade do enlace de saída, os nós de uma rede de pacotes (ex: roteadores em uma rede IP) possuem buffers especiais para armazenamento temporário dos pacotes, denominados filas. A disciplina de despacho define a forma como o nó da rede irá servir os pacotes armazenados nas filas. Quando a rede transporta simultaneamente tráfego de voz e dados, deve-se associar níveis de prioridade distintos aos dois tipos de tráfego, com a disciplina de despacho priorizando o tráfego de voz, de modo a minimizar o atraso que estes pacotes sofrem em cada nó da rede.

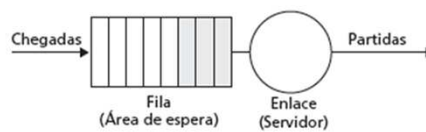
Inatel

Instituto Nacional de Telecomunicações

www.inatel.br

First-In, First-Out – FIFO

- Neste tipo de fila não há priorização de tráfego. Existe uma única fila de saída, os pacotes recebidos são armazenados e enviados na mesma ordem em que chegaram.
- Trata-se da implementação mais simples para filas. A maior parte dos roteadores trabalham com FIFO.

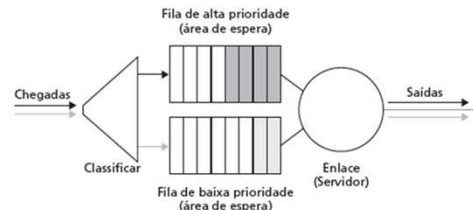


Neste tipo de fila, fontes de tráfego mal comportadas podem consumir toda a largura de faixa disponível, tráfegos em rajada podem causar atrasos inaceitáveis em tráfegos sensíveis a atraso, e pacotes pertencentes a tráfegos de maior importância podem ser perdidos devido a overflow do buffer, causado possivelmente por tráfegos de menor importância.

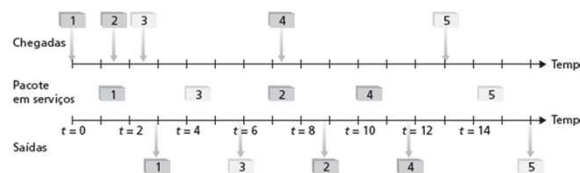
Este tipo de disciplina de despacho não é, portanto, adequada para aplicações onde há o tráfego simultâneo de voz e dados, como VoIP.

Priority Queueing – PQ

- No escalonamento prioritário, transmite-se pacotes da fila com prioridade mais alta primeiro. Portanto é feita uma classificação de tráfego na entrada do roteador.



- A ordem de saída não será necessariamente igual à ordem de entrada de pacotes.



Nesta técnica existem filas distintas para diferentes classes de prioridades. Por exemplo, podemos ter quatro níveis de prioridade (alta, média, normal e baixa), com uma fila associada a cada nível.

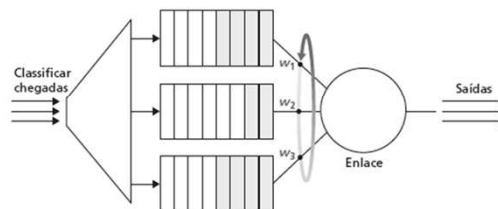
Os pacotes, uma vez classificados, são encaminhados para a fila com nível de prioridade correspondente. Durante a transmissão, o algoritmo de despacho dá tratamento preferencial absoluto à fila de maior prioridade, em detrimento das filas de menor prioridade. Ou seja, no momento da transmissão o algoritmo sempre busca um pacote na fila de alta prioridade; caso não haja pacote nesta fila, um pacote da fila de média prioridade é buscado, e assim sucessivamente até se chegar à fila de baixa prioridade.

Um volume de tráfego elevado de maior prioridade pode reter os tráfegos de menor prioridade por um tempo inaceitavelmente elevado; este fenômeno é conhecido como “starvation”. No pior caso, devido a um grande volume de tráfego de alta prioridade, por exemplo, o tráfego de baixa prioridade pode nunca ser transmitido. Para se evitar esta situação, pode-se utilizar ferramentas de controle para restringir a taxa de chegada de tráfego de alta prioridade.

A técnica Priority Queueing dá melhor resultado quando o tráfego de mais alta prioridade consome a menor quantidade de largura de faixa do canal, ou seja, possui menor volume.

Weighted Round Robin - WRR

- Na técnica WRR define-se várias classes de prioridade e associa-se a cada classe um percentual da banda do canal.



- As filas de maior prioridade são atendidas até que não haja mais tráfego ou até que o limite da banda alocada a cada fila seja alcançado.

No algoritmo Weighted Round Robin (WRR) os pacotes são classificados de acordo com o seu nível de prioridade. O algoritmo de despacho privilegia a transmissão dos pacotes de mais alta prioridade. No entanto, diferentemente do sistema de Filas com Prioridades, o tratamento de cada nível de prioridade não se dá de forma exaustiva.

No WRR define-se um percentual de banda associado a cada classe de prioridade. As filas são servidas em ordem decrescente de prioridade. Entretanto, contrário à operação da Fila com Prioridades, onde uma fila só começa a ser atendida após o atendimento pleno de todas as filas de maior prioridade, na técnica WRR o serviço passa para a próxima fila se a fila corrente ficar vazia ou se o percentual de banda atribuído a ela for ultrapassado. O controle da banda pode ser feito através de um contador de bytes associado a cada fila, que permite calcular o percentual de banda que cada fila está ocupando. Como se garante um percentual de banda a cada classe de fila, não há o problema de starvation.

O percentual de banda alocado a cada fila pode ser definido a partir de um peso associado a cada classe de fila, através da fórmula: $P_i = (W_i/S) \times B$. Onde P_i representa a taxa (em bps) alocada para o tráfego de classe i ; W_i representa o peso associado à classe de tráfego i ; S é o somatório dos pesos atribuídos a todas as classes de tráfego; e B é a banda total do canal.

Esta mesma filosofia de tratamento das filas aplica-se aos sistemas denominados Custom Queueing (CQ) e Class Based Queueing (CBQ).

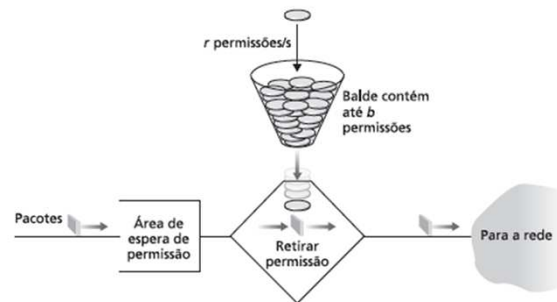
- POLICIAMENTO E CONFORMAÇÃO DE TRÁFEGO.

Policiamento e conformação de tráfego

- As funções de policiamento e conformação de tráfego usualmente identificam as violações no tráfego de uma mesma maneira. Elas diferem, contudo, na forma como respondem a estas violações.
- A função de policiamento usualmente descarta o tráfego que não está conforme ou o define como elegível para descarte.
- A função de conformação tipicamente atrasa o tráfego em excesso, através de mecanismos de enfileiramento, retendo os pacotes e liberando-os de maneira tal que o fluxo de saída esteja dentro dos parâmetros definidos.

Leaky Bucket – Balde Furado

- O leaky bucket, por meio de permissões, regula o fluxo de tráfego de uma determinada aplicação.



$b \equiv$ permissões de tráfego

$r \equiv$ taxa de permissões (permissões/segundo)

O Leaky Bucket é uma definição formal de uma taxa de transferência e se baseia em três parâmetros:

Comprimento de rajada: especifica a máxima rajada de dados que pode ser enviada dentro de um intervalo de tempo.

Taxa média: especifica quantos bits podem ser enviados por unidade de tempo, em média.

Intervalo de tempo: também chamado de intervalo de medida, especifica o intervalo de tempo em que se define o comprimento de rajada e a taxa média.

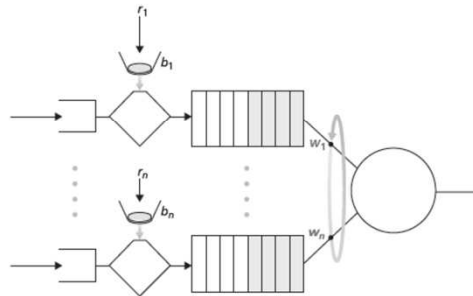
Leaky Bucket

- De acordo com o algoritmo do *Leaky Bucket*:
 - Como pode haver no máximo b permissões no balde, o tamanho máximo da rajada para um fluxo regulado pela técnica é de b pacotes.
 - Como a taxa de geração de permissões é r , o número máximo de pacotes que pode entrar na rede para qualquer momento t será:

$$\text{maxPacote} = r \cdot t + b$$

Leaky Bucket e Enfileiramento Justo Ponderado

- Junção do sistema de regulação e encaminhamento.



- O atraso máximo de um pacote neste sistema será:

$$d_{max} = \frac{b_1}{R \left(\frac{w_1}{\sum w_j} \right)}$$