Redes de Computadores

Camada de rede: NAT e PAT, QoS

Versão 1.0
Abril de 2017

Prof. Jairo

jairo@uni9.pro.br
professor@jairo.pro.br

http://www.jairo.pro.br/

Sumário

1 – NAT, DNAT e SNAT

1.1 – Descrição

O NAT (*Network Address Translation, tradução de endereço de rede*), é um método de tradução de endereço de rede que opera na camada 3, e que remapeia o endereço IP no pacote. O funcionamento básico consiste em modificar o endereço IP no cabeçalho do datagrama IP e/ou a porta TCP/UDP no cabeçalho de camada de transporte, quando este estiver sendo roteado.

Porém, a aplicabilidade do conceito NAT é muito extensa, diversificada e, em muitos casos, além da camada 3 pode envolver também camadas 4 (transporte) e 7 (aplicação). Muitas vezes DNAT (*Destination* NAT) e SNAT (*Source* NAT) também são tratados como NAT.

Por sua vez PAT está relacionado ao NAT, mas trata de tradução de portas TCP e não de endereços.

1.2 – NAT

O principal motivador da criação do NAT foi a percepção de que estava esgotando a disponibilidade dos endereços ipv4 para a internet (IPs públicos), isso em meados dos anos 1990.

Já naquela época, nas redes internas, passaram então a usar endereços IP das faixas reservadas (IPs privados). Com isto, para acesso à internet foi necessário NAT, técnica em que o roteador modifica o endereço IP do remetente no datagrama para que o pacote possa ser roteável na internet mesmo tendo origem numa LAN com endereços IP da faixa reservada.

Inicialmente, o conceito NAT era usado para a técnica de mascarar (*masquerade*) um endereço IP privado (LAN) num endereço IP público do roteador (IP externo). A tradução era de um para um, então para cada usuário na LAN que quisesse navegar na internet, era necessário também haver disponível um endereço público no roteador. Uma vez esgotados os endereços públicos, ninguém mais na LAN conseguia navegar na internet até que fosse liberado um IP público. A liberação do IP público ocorria automaticamente quando o usuário ficasse um determinado intervalo de tempo sem navegar na internet.

Na figura 1, abaixo. É mostrado uma única máquina da rede interna, com IP da faixa privada, navegando na internet com o IP externo (público) do roteador.

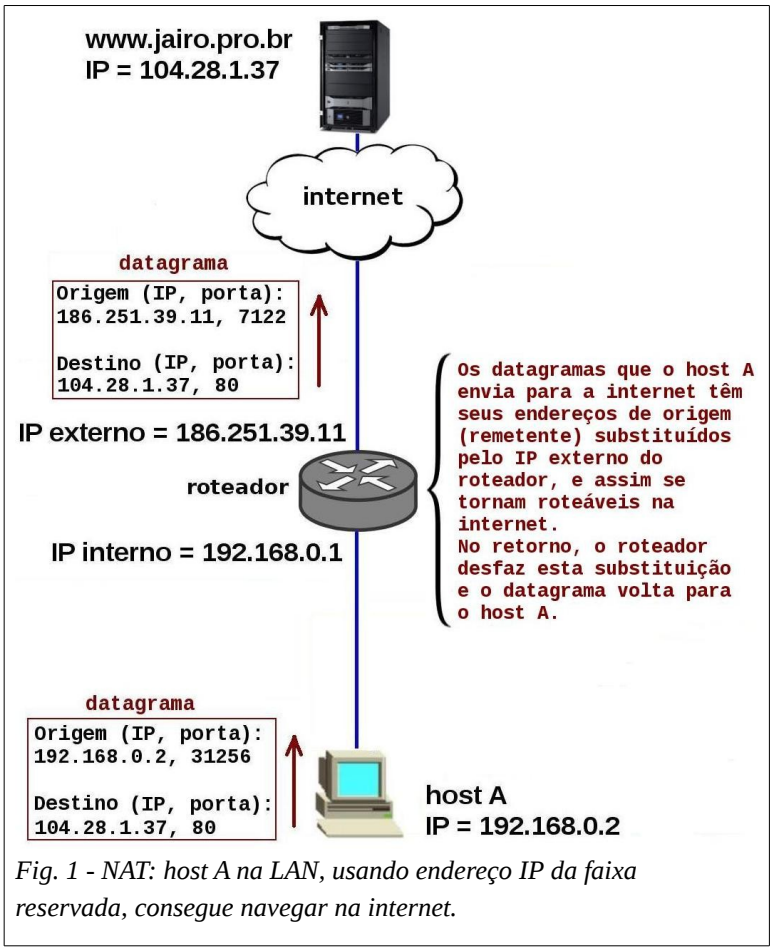


Fig. 1 - NAT: host A na LAN, usando endereço IP da faixa reservada, consegue navegar na internet.

Atualmente, o mapeamento de um para um é conhecido por *static* NAT (NAT estático).

Na figura 1, acima, no roteador as regras são um para um, portanto cada usuário na LAN irá navegar na internet com um endereço IP público diferente do outro.

Posteriormente, o conceito avançou e permitiu traduzir muitos para um, e assim foi possível traduzir todos os endereços IP privados na LAN para um único endereço IP público (IP externo).

Na figura 2, abaixo, nos datagramas com destino à internet os endereços IP da rede privada (LAN) são traduzidos para o endereço IP externo do roteador.

Atualmente, o mapeamento de muitos para um é conhecido por *dinamic* NAT (NAT dinâmico).

Um aspecto interessante sobre o NAT atualmente, é que ele impossibilita a comunicação fim a fim. Mas por outro lado, torna transparente a comunicação tanto para o *host* interno (na LAN), quanto externo (na internet). O *host* interno não sabe qual o IP público que está sendo usado na internet, e o *host* externo só vê o endereço IP público (externo) do roteador NAT. O único que sabe da conversão de endereços IP que está ocorrendo nos datagramas é o roteador NAT.

NAT, conforme apresentado acima, só vale para a versão 4 do protocolo IP, pois o que existe atualmente para ipv6 é muita discussão.

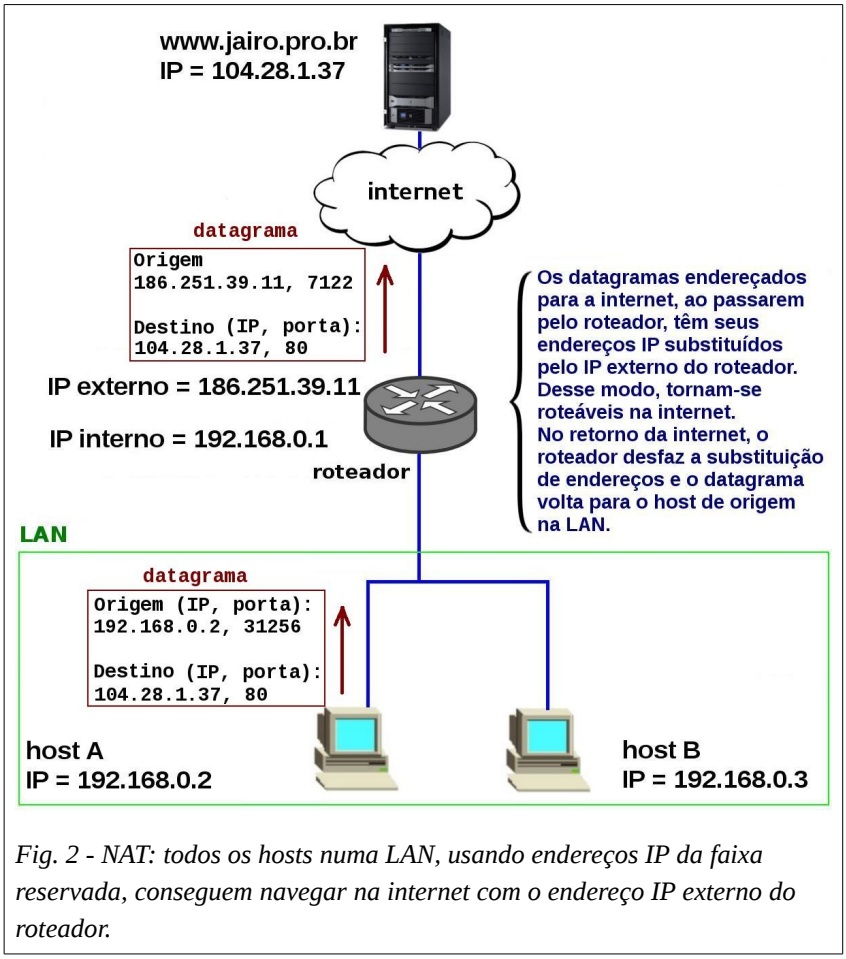


Fig. 2 - NAT: todos os hosts numa LAN, usando endereços IP da faixa reservada, conseguem navegar na internet com o endereço IP externo do roteador.

No passado, a ampla adoção do NAT ocorreu sem a contrapartida de padronização desta técnica, por isso existe hoje um grande número de produtos NAT que seguem definições diferentes entre si. Esta falta de padronização foi devido ao foco estar mais centrado na nova versão do protocolo IP (ipv6), que resolveria o problema de rápido crescimento da internet. O surpreendente é ver que hoje, transcorridos cerca de 20 anos, o ipv6 ainda está numa fase inicial de implantação e é o NAT que tem garantido o crescimento da internet.

E pior, a mesma resistência que houve no passado para padronizar o NAT está ocorrendo de novo agora, quando se busca interconectar as LANs em ipv4 com a internet em ipv6.

1.3 – Conceitos adicionais: *masquerading*, DNAT, SNAT

Masquerading: *IP masquerade* (IP mascarado) é uma técnica de mascaramento de endereço IP similar ao NAT um para muitos. É basicamente apenas uma questão de nomenclatura diferente.

DNAT: *Destination Network Address Translation* (*tradução de endereço de rede de destino*) é uma técnica para transparentemente alterar o endereço IP de destino (e não o endereço de origem

como no NAT) do datagrama ao passar pelo roteador, e depois, quando do retorno do datagrama, desfazer esta alteração. Tem semelhança com *port forward*, a diferença é que atua no endereço IP de destino e não na porta do serviço. No entanto, em muitos casos, DNAT também pode modificar a porta de destino. A figura 3, abaixo, ilustra um caso de uso do DNAT, onde o host C tenta acessar o serviço web em 172.16.1.8 mas é transparentemente redirecionado para o serviço web em 10.10.1.5.

SNAT: *Source Network Address Translation* (tradução de endereço de rede na origem) é também uma técnica de NAT conforme descrito acima, porém normalmente somente é usada para modificar o endereço de destino do datagrama originado dentro de uma rede interna que esteja sendo encaminhado para a internet.

Nos conceitos acima para DNAT e SNAT, é importante notar que a definição é dependente do referencial, faz diferença se a conexão parte de um *host* dentro de uma LAN ou o oposto.

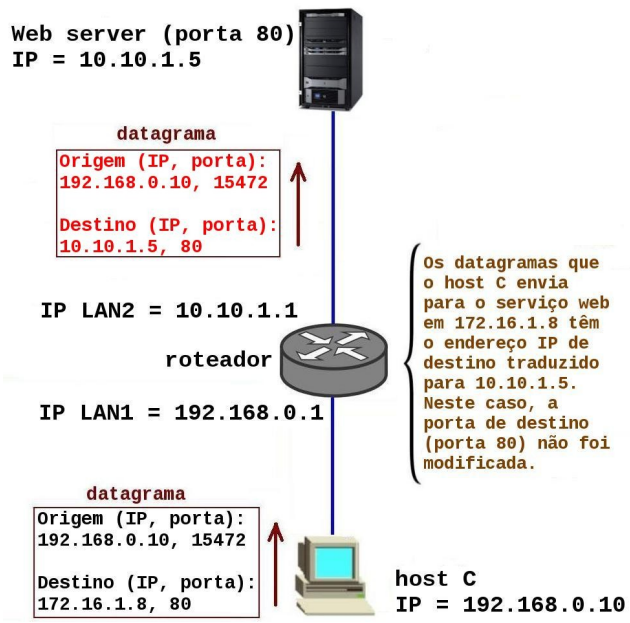


Fig. 3 - DNAT: o host C endereça o datagrama para o IP de destino 172.16.1.8, porém o roteador faz DNAT e modifica o endereço de destino para 10.10.1.5.

2 – PAT ou NAPT (*Port Address Translation*)

Um conceito relacionado a NAT, porém ligeiramente diferente, é o PAT (*Port Address Translation*, tradução de endereço de porta) ou NAPT (*Network Address Port Translation*, tradução de endereço de porta de rede).

PAT é técnica de tradução de endereço de porta. No exemplo ao lado, o host D tenta acessar o servidor web no IP 10.10.1.6:80 (porta web padrão, 80), porém na realidade o serviço está na porta 8000.

Para que o cliente possa então acessar o serviço, ao passarem pelo roteador os pacotes de dados têm a porta traduzida de 80 para 8000.

É importante notar que, na prática, muitas vezes a tarefa de tradução pode envolver simultaneamente NAT e PAT, ou seja, o roteador pode alterar tanto o endereço IP de destino quanto a porta de destino do datagrama.

Outro aspecto a ser considerado, é que as portas de origem e de destino são definidas em campos do cabeçalho TCP ou UDP (camada 4), e não do cabeçalho IP (camada 3).

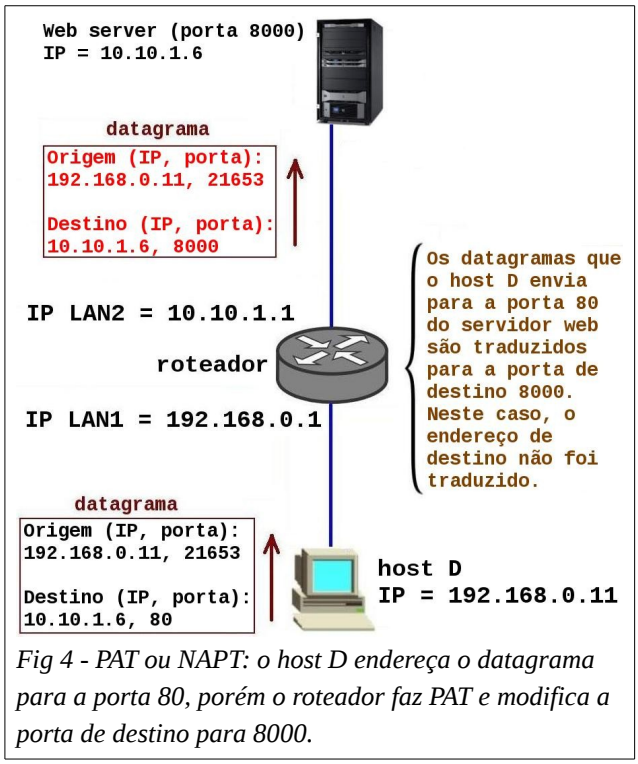


Fig 4 - PAT ou NAPT: o host D endereça o datagrama para a porta 80, porém o roteador faz PAT e modifica a porta de destino para 8000.

3 – QoS (Quality of Service)

3.1 – Descrição

Qualidade do Serviço (*Quality of Service* – QoS) se refere à priorização do tráfego de pacotes e mecanismos de controle de reserva de recursos da rede para um ou mais fluxos de dados em detrimento dos demais. QoS é a habilidade de prover diferentes prioridades a diferentes aplicações, a diferentes usuários ou a diferentes fluxos de dados.

QoS é usado para favorecer um tipo de tráfego ou de serviço em detrimento de outro. Deste modo, diferentes tipos de tráfegos e serviços são tratados diferenciadamente. Por exemplo, favorecer o tráfego de voz sobre IP (VoIP) em detrimento da transferência de arquivos via FTP (*File Transfer Protocol*).

Por padrão (*default*), os roteadores manuseiam o tráfego de pacotes numa base primeiro-que-chega (*first-come*), primeiro-que-sai (*first-served*). Não há como controlar retardos (*delays*) na entrega de pacotes, pois são causados tanto pela distância que o pacote precisa viajar quanto pela sobrecarga nos equipamentos ao longo do caminho. A latência da rede aumenta à medida que o volume de tráfego aumenta, e se a carga num determinado linque aumentar rapidamente a ponto de causar transbordamento de fila (*queue overflow*), isso resulta em congestionamento com consequente perda de pacotes de dados. Usando protocolo de transporte TCP, se ocorrer congestionamento num determinado roteador, este vai sinalizar para reduzir a taxa de transmissão com que está recebendo estes pacotes, e assim aliviar o congestionamento.

Sem QoS, a comunicação VoIP é menos confiável que a telefonia convencional, pois a rede não garante que os pacotes não sejam perdidos, que não haja demora nesta entrega e que estes pacotes sejam entregues na mesma ordem sequencial em que foram enviados. Normalmente VoIP usa UDP e não TCP, pois UDP causa menos sobrecarga (*overhead*) na rede, e a recuperação de congestionamento TCP implica em muita latência devido à retransmissão de pacotes perdidos.

A implementação do QoS envolve, basicamente, configurar os roteadores (camada 3) e *switches* (camada 2) por onde passa o tráfego do serviço que necessita mais qualidade.

Roteadores submetidos a um alto volume de tráfego introduzem latência que excede os limiares (*thresholds*) permitidos para VoIP ou para transmissão de vídeo (*streaming video*). Mas se forem usados os mecanismos QoS, pacotes de determinados serviços podem ser transmitidos à frente de qualquer tráfego massivo enfileirado no mesmo linque, isso mesmo quando estiver ocorrendo transbordamento de fila e perda de pacotes de dados.

3.2 – Definições QoS

As definições de QoS variam em acordo com o serviço e com a abordagem escolhida. Características e métricas típicas incluem **largura de banda** (*bandwidth*), **latência** (*delay* ou *latency*), **variação da latência** (*jitter*) e **confiabilidade** (*reliability*).

- **Largura de banda** (*bandwidth*): determina a taxa de transferência necessária para o fluxo de dados. Normalmente, *stream* vídeo consome muito mais recursos da rede que *stream* áudio. Tipicamente, um *stream* áudio consome cerca de 64 Kbps de largura de banda, já um *stream* vídeo de 1920x1080 com 60 *frames* por segundo consome cerca de 2 Mbps. A largura de banda pode oscilar e ter um pico, pode ser sustentada ou pode ser mínima;
- **Latência** (*delay* ou *latency*): é o tempo decorrido entre o envio (origem) e o recebimento do pacote (destinatário). Se for regular, um valor típico de 100 ms de latência não é problema, pois a consistência na chegada dos pacotes tende a ser mais importante do que o tempo gasto no caminho;
- **Variação da latência** (*jitter*): é a variação do tempo decorrido entre o envio e recebimento da mensagem. Este tempo variável, aleatório e imprevisível é devido à competição dos outros usuários pelo mesmo linque de transmissão. Os efeitos de variação de latência podem ser mitigados estocando brevemente num *buffer*[1] os pacotes de *stream* que estão chegando, onde permanecem até a reprodução do áudio ou vídeo. O uso de *buffer* aumenta a demora no recebimento da mensagem, mas melhora a qualidade da transmissão pela redução de interrupções, pois aumenta a probabilidade de que cada pacote já esteja à mão no momento de reproduzir o *stream*. Em *stream* áudio, a interrupção aparece como som “cortado”, em *stream* vídeo como tela congelada. Outro problema é receber alguns pacotes fora de ordem sequencial, alguns pacotes muito atrasados ou mesmo alguns pacotes nunca chegarem;
- **Confiabilidade** (*reliability*): determina os erros de transmissão e a perda de pacotes. A perda de pacotes é causada tanto pelo tráfego excessivo na rede quanto falhas de *hardware* ou roteadores incorretamente configurados. Se ocorrer tráfego excessivo, o próprio roteador pode começar a descartar pacotes. Em todos estes casos, a qualidade da comunicação será afetada. Baixa confiabilidade geralmente é o pior problema para a Qualidade do Serviço.

Com estas definições, podemos dizer que, para um tráfego selecionado, Qualidade do Serviço é a capacidade de uma rede em prover melhores serviços pela atribuição de prioridade a este tráfego, devido a inclusão de uma largura de banda selecionada e com controle da latência e de sua variação.

Além disso, Qualidade do Serviço, QoS, é um nome genérico sob o qual estão numerosos conceitos, tecnologias e implementações, os quais providenciam diferentes mecanismos para resolver questões fundamentais de gerenciamento de largura de banda. O ponto chave para QoS é assegurar que a rede aloque recursos baseado nas necessidades de tráfego e na política do negócio.

Um aspecto importante nesta área é a diferença entre Classe de Serviço, CoS, e Qualidade do Serviço, QoS. Enquanto CoS é frequentemente parte de uma implementação QoS, CoS sozinho é capaz de prover um valioso mecanismo de capacidade de alocação para muitas, se não a maior parte

1 Buffer: é espaço de armazenamento temporário na memória RAM.

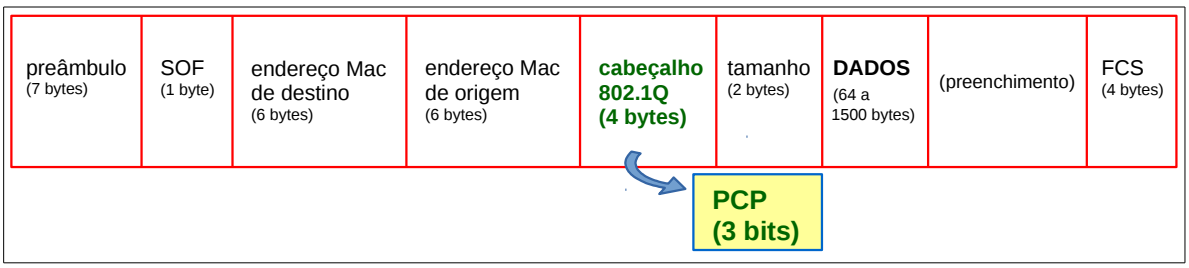
das redes de hoje.

3.3 – Classe de Serviço (*Class of Service* - CoS)

Classe de Serviço (CoS) é uma maneira de gerenciar o tráfego de rede agrupando juntos tipos similares de tráfegos, e depois tratando cada tipo como uma classe com seu próprio nível de prioridade de serviço. Ao contrário do gerenciamento do tráfego com QoS, a CoS isoladamente não consegue garantir um nível de serviço em termos de largura de banda e demora na entrega, apenas oferece um bom desempenho. Mas por outro lado, CoS é simples de gerenciar e mais escalável à medida que a rede cresce tanto na estrutura quanto no volume de tráfego. Mas quando comparado ao QoS, o controle de tráfego com CoS é limitado, especialmente nos casos em que toda a largura de banda estiver sendo gasta com tráfego priorizado pelo CoS.

Para CoS, são usadas três tecnologias: **Marcação de camada 2** (*layer 2 tagging*, VLAN), **Tipo do Serviço** (*Type of Service* – ToS) e **Serviços Diferenciados** (*Differentiated Services* – DiffServ). A primeira trabalha na camada 2 (enlace), e as outras duas na camada 3 (rede).

- **Marcação de camada 2** (*layer 2 tagging*, VLAN): no cabeçalho do quadro (*frame*), usa três bits do campo de VLAN (IEEE802.1Q) para especificar um Ponto de Código de Prioridade (*Priority Code Point* - PCP). Abaixo segue o esquema do cabeçalho de camada 2 (enlace) ethernet, onde é mostrado o campo PCP:



O padrão IEEE802.1Q suporta LANs virtuais (VLANs) em rede ethernet. Este padrão define um sistema de marcação ou “tagueamento” (tagging) de VLAN nos quadros ethernet, além dos procedimentos a serem usados pelas *bridges* e *switches* no manuseio de tais quadros “tagueados”. Um quadro onde esteja incluído o campo VLAN é considerado “tagged” (marcado ou “tagueado”).

O campo PCP do IEEE802.1Q, de três bits, contém provisão do nível de prioridade para o quadro, que admite 8 níveis de prioridade para as diferentes classes de tráfego. Este campo especifica valores entre 0 e 7, mais conhecidos como CS0 a CS7, que são usados para diferenciar o tráfego.

Pela ordem de prioridade, estes 7 níveis são:

1 – básico (*background*);
0 – melhor empenho (*best effort*), que é o padrão (*default*);
2 – excelente empenho (*excellent effort*);
3 – aplicação crítica (*critical application*);
4 – vídeo;
5 – voz (*voice*);
6 – controle de redes interconectadas (*internetwork control*);
7 – controle da rede (*network control*).

- **Tipo do Serviço** (*Type of Service* – ToS): atualmente, ToS é chamado de Serviço Diferenciado (*Differentiated Service* – DS), e está definido no campo DSCP de 6 bits do cabeçalho ipv4. Em ipv6, está no campo Classe de Tráfego do cabeçalho IP, de 8 bits, mas usa apenas os 6 primeiros para DS. DiffServ permite o reaproveitamento do campo TOS, que no novo *layout* passou a se chamar campo DS a partir de 1998.

- **Serviços Diferenciados** (*Differentiated Services* – DiffServ): DiffServ ou DS é um protocolo para especificar e controlar tráfego na rede. Usa classes para dar prioridades a determinados tráfegos. É o método mais avançado para gerenciamento de tráfego CoS. Mas no modo "clássico" ou "antigo", usa apenas 3 bits do cabeçalho IP, e não 6.

Geralmente dentro da rede é criado um pequeno conjunto de diferentes classes de tráfego. Para cada classe, precisa ser especificado uma prioridade. Nos pacotes (camada 3) isto é geralmente marcado como um tipo pelos bits do DiffServ (Serviços Diferenciados), e nos quadros é marcado nos bits PCP. Classe de serviço é um parâmetro usado para diferenciar os tipos de cargas contidos nos pacotes que estão sendo transmitidos, e o objetivo é associar prioridades ou níveis de acesso a determinados fluxos de dados.

Na forma "clássica" do CoS, no conceito ToS, usa apenas 3 bits do atual campo DS (Serviços Diferenciados) do cabeçalho IP. Isso é tratado como **Precedência IP** (IP Precedence), onde o administrador da rede pode atribuir valores de 0 a 7 para classificar e priorizar os tipos de tráfegos. Muitas aplicações e roteadores suportam Precedência IP, onde o tráfego é diretamente rotulado e portanto contém a marcação QoS.

Porém, o esquema de Precedência IP somente permite a especificação de prioridade relativa de um pacote. Não contém provisão para especificar uma precedência de descarte diferente para pacotes com certas prioridades. Então, tudo vai bem enquanto o tráfego priorizado por CoS tiver largura de banda suficiente, mas uma vez esgotado este recurso não existe mecanismo para conter o congestionamento de pacotes.

Uma capacidade adicional para CoS é o Controle de Fluxo de Admissão (*Admission Flow Control*), que limita o número de fluxos a um número máximo permitido pela ocupação da largura de banda. Por exemplo, se limitar a 10 o número máximo de *stream* áudio, onde cada um tem cerca de 64 Kbps, o consumo total de largura de banda ficará em torno de 640 Kbps.

No modo de confiança (*trust mode*), os dispositivos de rede (roteadores, switches, etc)

podem ser configurados para usarem os valores de CoS dos pacotes que estão chegando, vindo de outros dispositivos. Mas também podem reescrever, nos pacotes, o valor CoS para algo completamente diferente. Muitos Provedores de Serviço de Internet (*Internet Service Providers* - ISP) não confiam nas marcações QoS dos pacotes que vêm dos seu clientes, então geralmente o uso de CoS é limitado à rede interna da organização.

3.4 – DiffServ

No caso de CoS acima, apenas três bits não permite muita sofisticação no gerenciamento do tráfego, então foi desenvolvido o protocolo DS (DiffServ) para gerenciamento de pacotes com uma abordagem diferente de simples rotulação de prioridade (Precedência IP).

DS usa uma indicação de como determinado pacote tem que ser encaminhado, que é conhecido como Comportamento Por Salto (*Per Hop Behavior* – PHB). O PHB descreve o nível de um serviço em particular em termos de largura de banda, teoria de filas (*queueing theory*) e decisões de descarte de pacotes.

O DiffServ usa uma aproximação simples com menos sobrecarga de sinalizações, onde no caminho inteiro os nós de rede intermediários não estão cientes do QoS. Os pacotes são classificados e marcados para receber um particular encaminhamento PHB.

A arquitetura de Serviços Diferenciados (DS ou DiffServ) foi desenvolvida para ser relativamente simples, com métodos grosseiros para prover diferentes níveis de serviços para o tráfego na internet. A vantagem do DS é que muitos *streams* de tráfego podem ser agregados num único de menor número de agregados de comportamento e encaminhado usando o mesmo PHB no roteador, desse modo simplificando o processamento e o armazenamento associado. PHB é uma descrição externa do comportamento observável de roteamento de um nó DS aplicado a um particular agregado de comportamento DS. Não existe sinalização além do que é carregado no DSCP (DS *Code Point*) de cada pacote. Não é requerido processamento adicional na parte central da rede DS, pois o QoS é invocado numa base de pacote a pacote.

O DSCP usa apenas os 6 primeiros bits, e os dois últimos, de ECN (*Explicit Congestion Notification*), são ignorados.

As redes que implementam DiffServ são chamadas de Domínios DS. Um domínio DS consiste de um conjunto coerente de nós de redes, os quais suportam mecanismos DiffServ (nós com DS habilitado) e que pertencem ao mesmo domínio administrativo (por exemplo, um Sistema Autônomo[2]).

Mas a internet é a união de diferentes Sistema Autônomos que estão sob a administração de muitos ISP[3], e que competem entre si. Entretanto, foi tomado muito cuidado no desenvolvimento da arquitetura DS para considerar estes domínios administrativos na internet, e definir mecanismos

2 Sistema Autônomo é uma rede ou um conjunto de redes sob uma gestão comum, e que possuem características e políticas de roteamento comuns.

3 ISP: *Internet Service Provider* (Provedor de Acesso à Internet).

para que ocorra uma transição controlada de fluxos de dados com qualidade suportada entre os diferentes domínios administrativos.

3.5 – IntServ e RSVP

IntServ (Serviços Integrados) é um modelo com objetivo de garantir QoS. Ao contrário de Serviços Diferenciados (DiffServ), que tem pouca granularidade, IntServ é um sistema de muita granularidade. Em IntServ, o fluxo de dados é tratado como uma reserva indivual.

Basicamente, a idéia por trás do IntServ é que qualquer roteador no sistema suporte IntServ, e que qualquer aplicação que exija um determinado nível de garantia QoS seja responsável por fazer reservas individuais. A Especificação do Fluxo (*Flow Specs*) descreve em que consistem estas reservas, e RSVP (*Resource Reservation Protocol*) é o mecanismo responsável por fazer as reservas.

O Protocolo de Reserva de Recursos (*Resource Reservation Protocol* - RSVP) é usado para requisitar uma qualidade de serviço na rede, e pode ser usado tanto em roteadores quanto nas máquinas dos usuários finais. Todas as máquinas capacitadas a enviar dados QoS na rede enviam uma mensagem PATH (caminho) a cada 30 segundos, a qual se espalha pela rede. Aqueles dispositivos que desejarem escutar a mensagem precisam responder com um RESV (diminutivo de *Reserve* – reserva), que então traça o caminho (*path*) de volta para a origem. A mensagem retornada, RESV, contém as Especificações de Fluxo (*Flow Specs*).

Os roteadores entre quem solicita (*sender*) uma reserva e quem escuta (*listener*) precisam decidir se podem suportar a reserva requisitada, e caso não seja possível devem enviar uma mensagem de rejeição para que o solicitante saiba disso. Mas uma vez aceita a reserva, eles devem transportar o tráfego. Aceito o tráfego, os roteadores então estocam e policiam a natureza do fluxo de dados, e se nada for escutado durante um certo intervalo de tempo, então o roteador vai declarar tempo esgotado (*time out*) e cancelar a reserva. Individualmente, os roteadores precisam policiar o tráfego para verificar se correspondem às Especificações de Fluxo.

IntServ e RSVP proveem uma estrutura para controle detalhado do fluxo individual QoS à medida que ele passa através dos roteadores. Porém, a escalabilidade da solução IntServ deixa dúvidas devido a sua complexa classificação e agendamento por fluxo. IntServ requer que os roteadores estoquem e processem cada fluxo individual que passe por ele, o que se torna muito pesado na internet. Pior, todos os estados (que são muitos) precisam ser estocados em cada roteador e, como resultado, IntServ funciona bem em pequena escala, mas à medida que o sistema cresce e atinge o tamanho da internet, fica muito difícil manter a trilha de todas estas reservas.

3.6 – Diferenças entre DiffServ e IntServ

DiffServ é diferente de IntServ em muitos aspectos, e a principal diferença é que DiffServ

distingue um número pequeno de classes de roteamento ao invés de fluxos individuais. IntServ usa o protocolo de sinalização RSVP para reservar recursos para fluxos individuais, enquanto DiffServ aloca recursos numa base por classe definida no cabeçalho IP. Cada roteador ao longo do caminho examina o cabeçalho IP e toma uma decisão QoS baseado na política configurada no roteador. Como resultado, toda a informação que o roteador necessita para manusear o pacote está contida no cabeçalho do pacote, e então os roteadores não precisam aprender e estocar informações sobre fluxos individuais. Para DiffServ, não existe problema de escalabilidade associado à necessidade do roteador manusear o pacote por fluxo. Cada roteador é tratado independente do outro, entretanto, para prover QoS consistente, os roteadores precisam ser configurados com políticas similares.