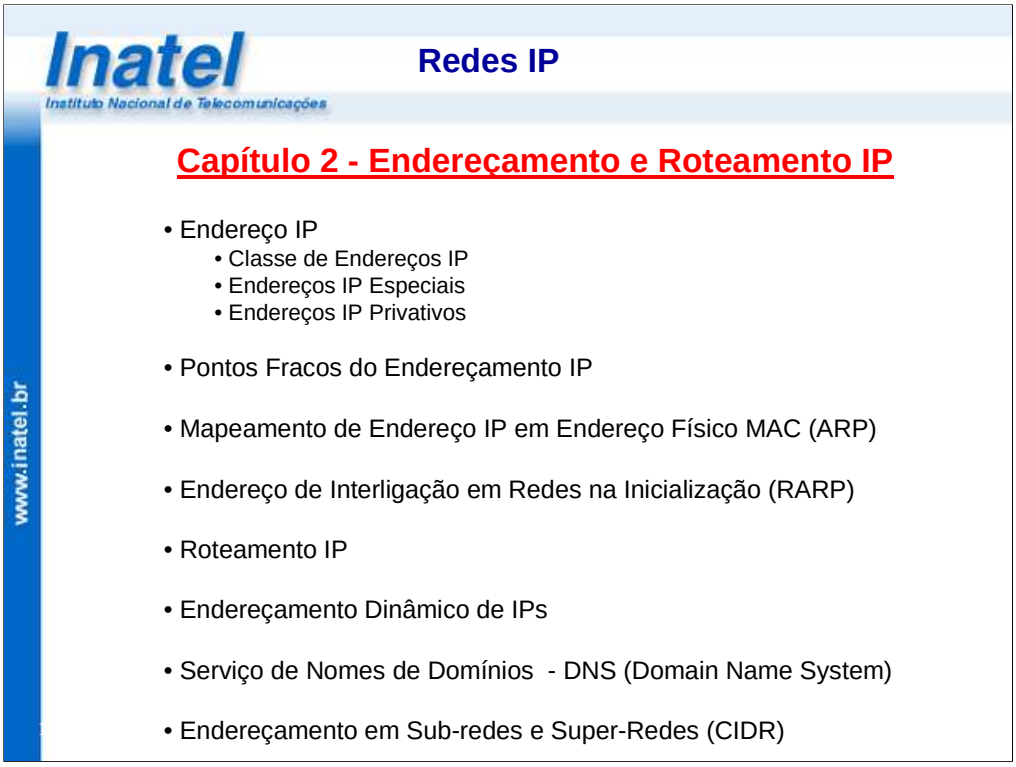
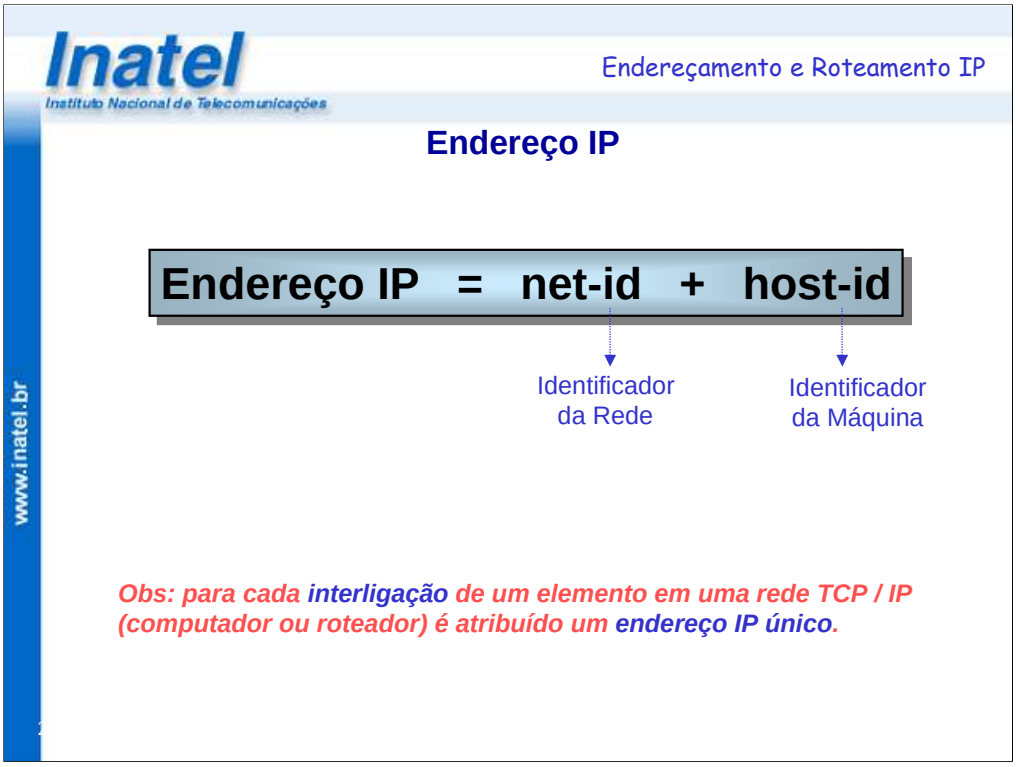


Neste capítulo, será mostrado como uma rede TCP / IP identifica seus elementos: através do endereço IP. Este endereço IP é separado em classes e existem endereços para uso em casos especiais e até mesmo para uso privado.

Apesar de ser bastante eficiente, o endereçamento IP possui alguns pontos fracos, que serão mostrados neste capítulo.

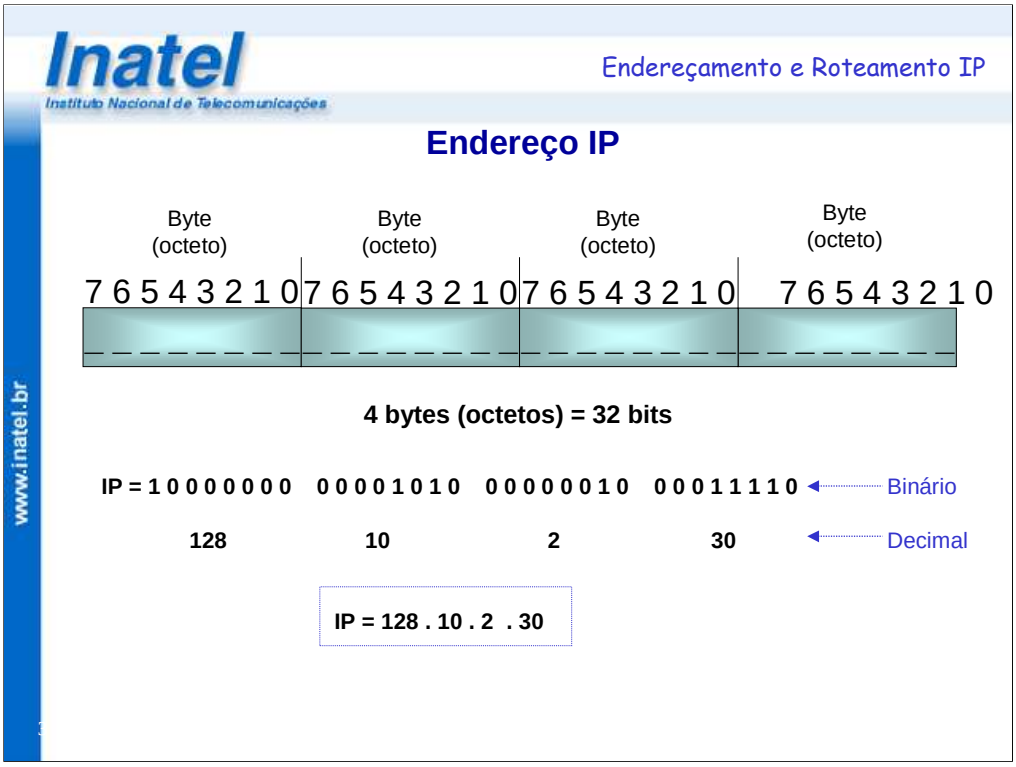
Outras questões serão abordadas, como: mapeamento de endereço IP em endereço MAC utilizando os protocolos ARP e RARP, como é feito o esquema de roteamento na rede TCP / IP, fragmentação dos pacotes através da rede, serviços de nomes, sub-redes e super-redes que são as novas maneiras de se usar o endereço IP na Internet.



A interligação em redes TCP / IP é uma *estrutura virtual, totalmente implantada em software*. Assim, os projetistas estão livres para escolher formatos e tamanhos de pacotes, endereços, técnicas de entrega a assim por diante; nada é orientado pelo hardware. Para os endereços, os projetistas de TCP / IP optaram por um esquema análogo ao endereçamento de rede física, no qual a cada host da interligação em redes é atribuído um endereço com número inteiro de 32 bits, denominado de endereço IP. A parte interessante do endereçamento da interligação em redes é que os números são escolhidos cuidadosamente para tornar o roteamento eficiente. Especificamente, um endereço IP codifica a identificação da rede à qual um host se acopla, assim como a identificação de um único host nessa rede.

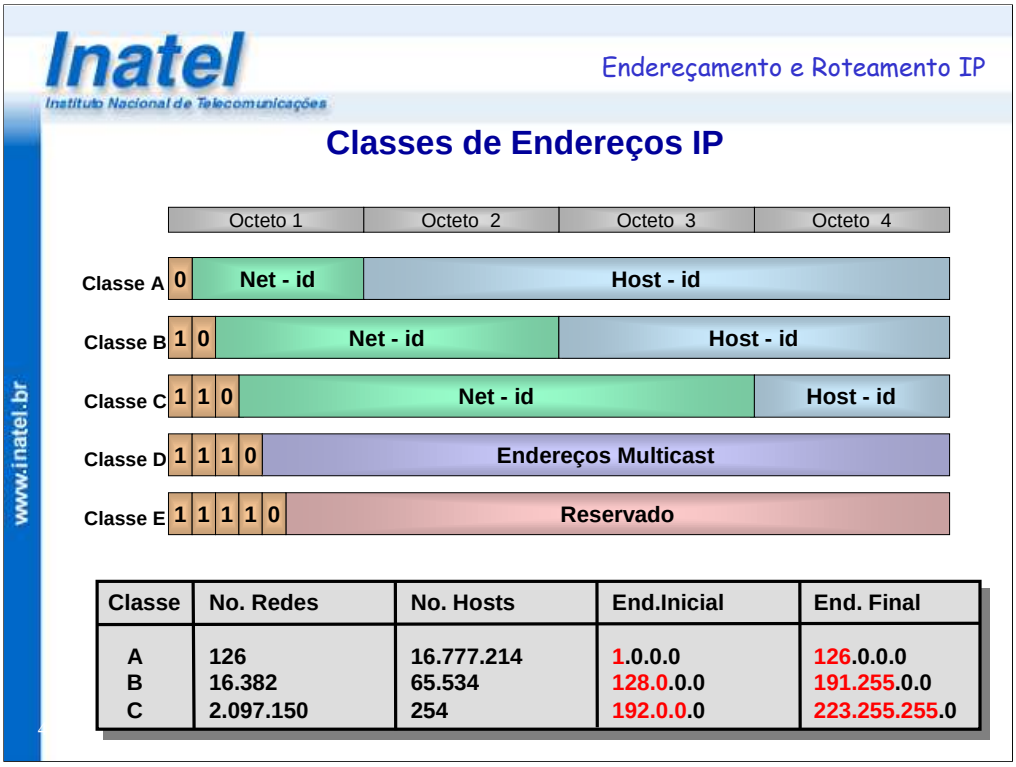
Podemos então resumir que a cada host de uma interligação em redes TCP / IP é atribuído um endereço de interligação em redes único de 32 bits que é usado em todas as comunicações com aquele host.

Conceitualmente, cada endereço é um par (net-id, host-id) em que net-id identifica a rede e host-id identifica um host naquela rede.



O endereço IP possui 32 bits, escritos como quatro números inteiros decimais separados por pontos, no qual cada número inteiro fornece o valor de um octeto de endereço IP.

Usaremos a notação decimal com ponto para expressar os endereços IP, durante todo o restante deste texto.



Classe	No. Redes	No. Hosts	End.Inicial	End. Final
A	126	16.777.214	1.0.0.0	126.0.0.0
B	16.382	65.534	128.0.0.0	191.255.0.0
C	2.097.150	254	192.0.0.0	223.255.255.0

A forma de dividir os 4 octetos em quantos octetos irão representar o identificador da rede e quantos irão representar o identificar do host, é chamada de classes de endereços IP. Portanto, dado um endereço IP, sua classe pode ser determinada a partir dos três bits de alta ordem (mais significativos), sendo dois bits suficientes para distinguir entre as três classes principais.

Um endereço classe A consiste em endereços que tem uma porção de identificação de rede (net-id) de 1 byte (octeto) e uma porção de identificação de máquina (host-id) de 3 bytes. Endereços classe B utiliza 2 bytes para a rede e 2 bytes para a máquina, enquanto que um endereço classe C utiliza 3 bytes para a rede e apenas 1 byte para a identificação das máquinas.

Nesta forma de divisão é possível acomodar um pequeno número de redes muito grandes (classe A), um médio número de redes médias (classe B) e um grande número de redes pequenas (classe C).

A classe D é uma classe especial para identificar endereços de grupo (multicast) e a classe E é uma classe reservada para uso futuro.

Classes de Endereços IP

NETID: Identifica a Rede, HOSTID: Identifica o HOST na rede

Classe A: NNN . HHH . HHH . HHH

1 – 126

0 . 0 . 1 – 255 . 255 . 254

Classe B: NNN . NNN . HHH . HHH

128 . 0 – 191 . 255

0 . 1 – 255 . 254

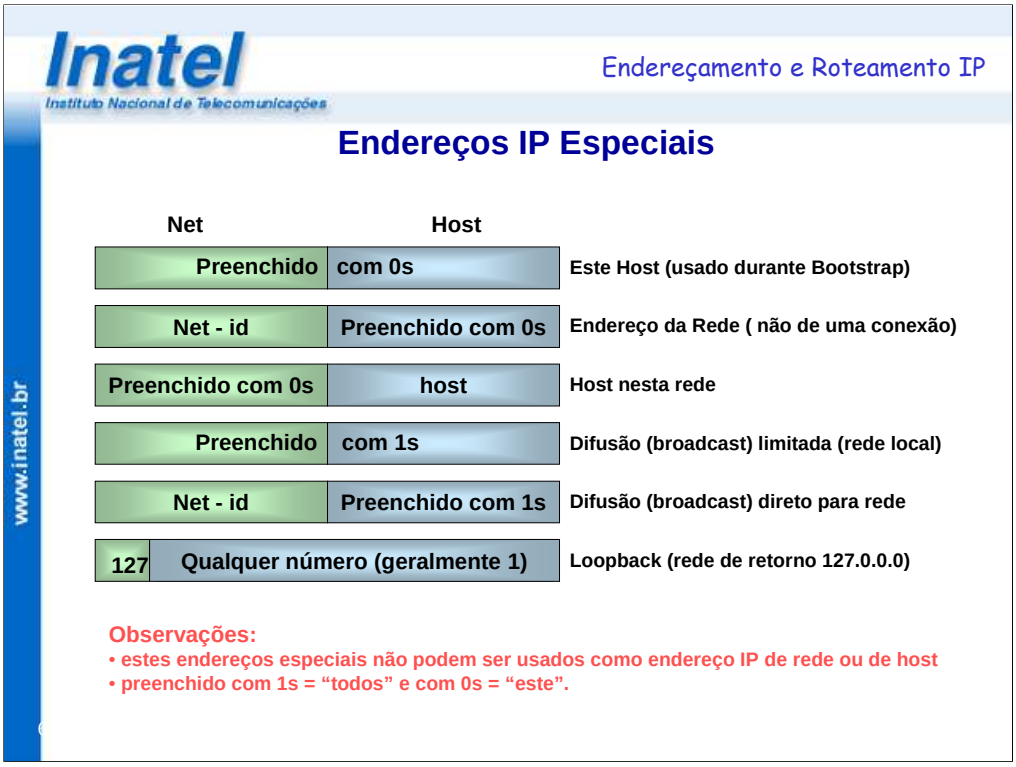
Classe C: NNN . NNN . NNN . HHH

192 . 0 . 0 – 223 . 255 . 255

1 – 254

Classe D: 224 . 0 . 0 . 1 ..... 239 . 255 . 255 . 254

Classe E: 240 . 0 . 0 . 1 ..... 247 . 255 . 255 . 254



Qualquer campo de um endereço IP com valor 0 ou 1 em todos os bits, tem significado especial:

Todos os bits 0: significa que este host (bits host-id=0) ou esta rede (bits net-id=0).

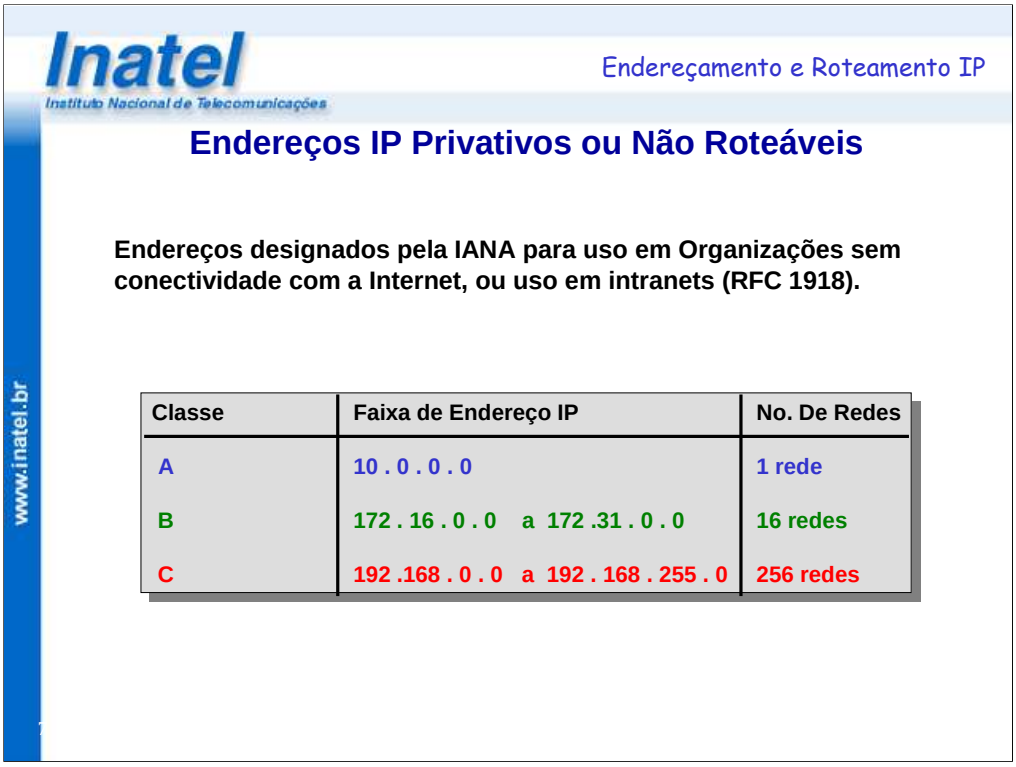
Todos os bits 1: significa todas as redes ou todos os hosts.

Com estes conceitos podemos definir os seguintes endereços IP especiais:

Endereço de Rede: representado por todos os bits de host-id com valor 0, identifica a própria rede e não uma interface de rede específica.

Endereço de Broadcast: representado por todos os bits de host-id com valor 1, identifica todas as máquinas na rede especificada em net-id.

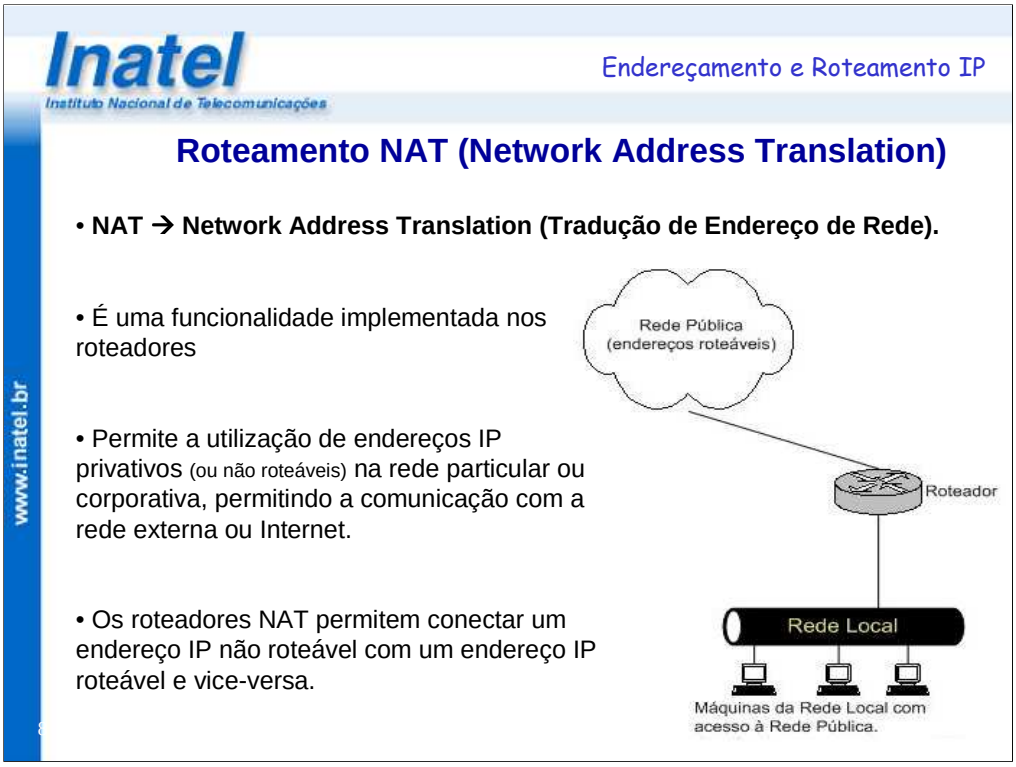
Nestas duas formas anteriores, podemos definir que os primeiros endereços IP e os últimos de cada classe não podem ser usados por nenhuma interface de rede.



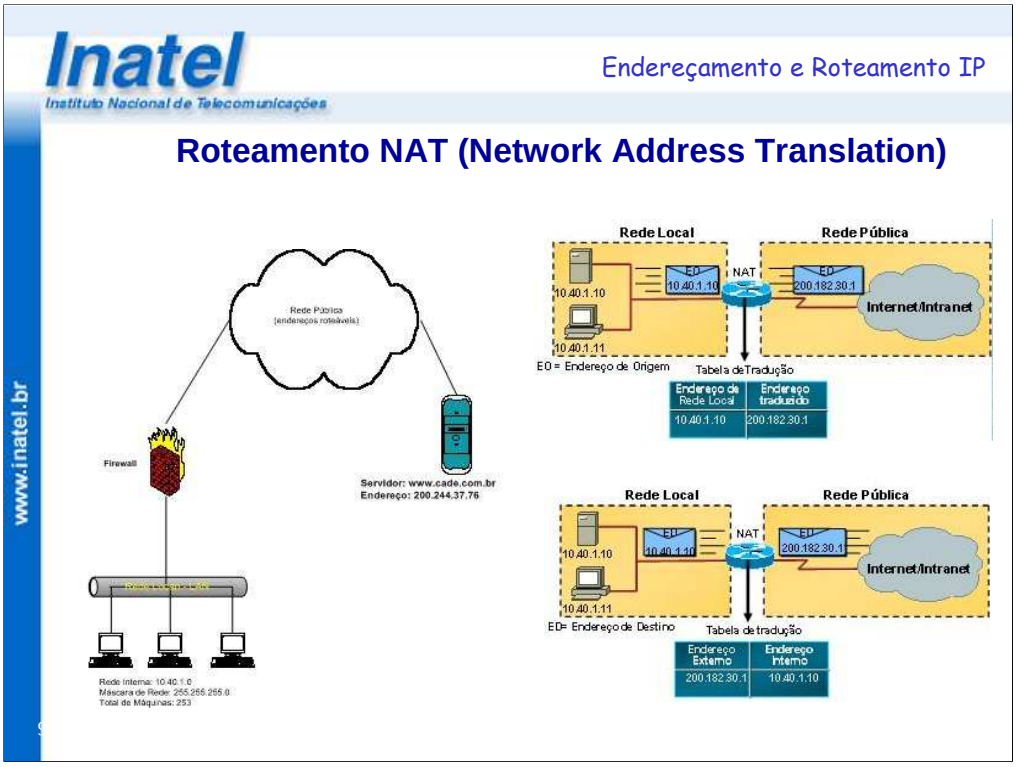
A regra de endereçamento IP na Internet diz que os endereços IP são únicos globalmente, reservando parte do espaço de endereços para redes que são usadas exclusivamente dentro de uma única organização e que não requerem conectividade com a Internet. Três variações de endereços foram reservadas pele IANA para este propósito:

1 rede classe A:	10
16 redes classe B:	172.16 a 172.31
256 redes classe C:	192.168.0 a 192.168.255

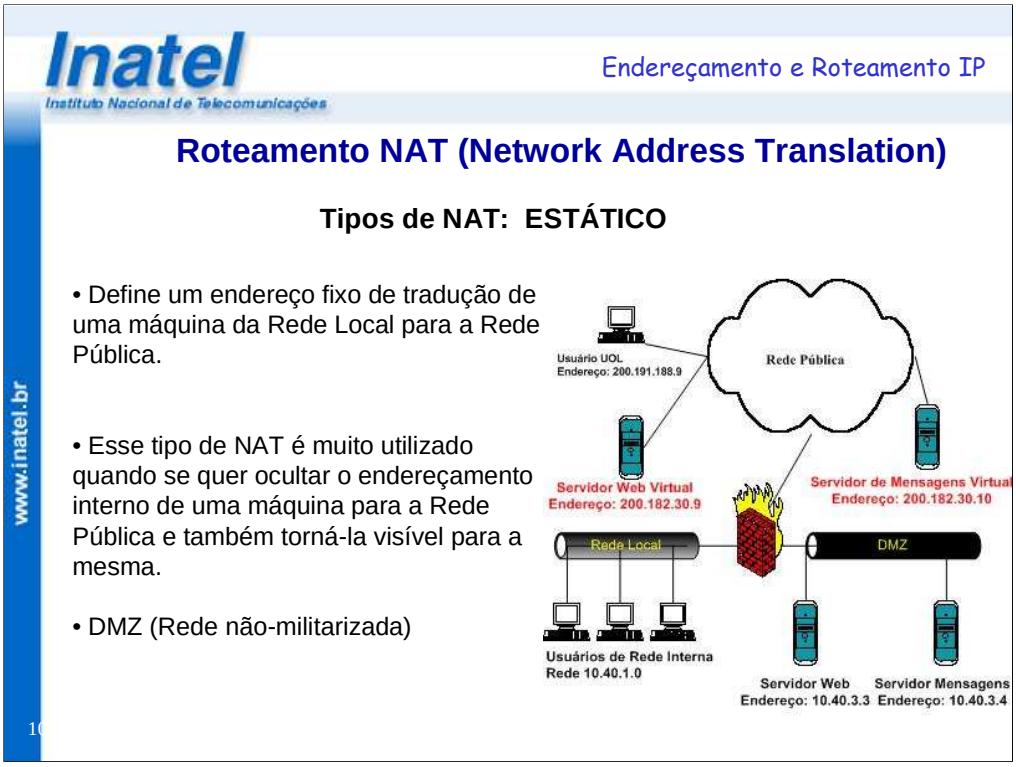
Qualquer organização pode usar quaisquer endereços deste intervalo sem referência a qualquer organização. Entretanto, como estes endereços não são únicos globalmente, eles não podem servir de referência a hosts de outra organização.



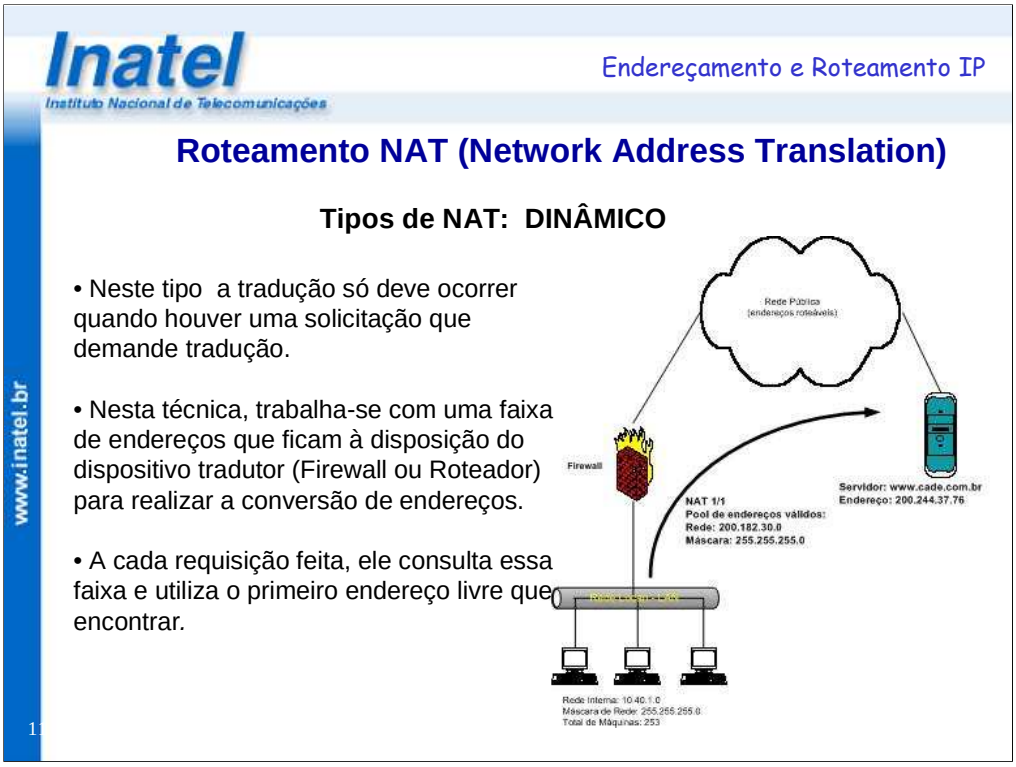
NAT é a abreviação de Network Address Translation (Tradução de Endereço de Rede). O NAT é constituído de roteadores especiais, que permitem a utilização de endereços IP privativos ou não roteáveis na sua rede particular ou corporativa, permitindo a comunicação com a rede externa ou Internet. Ou seja, os roteadores NAT permitem conectar um endereço IP não roteável com um endereço IP roteável e vice-versa.



Nessa rede, os usuários da rede local 10.40.1.0/24 pretendem acessar o servidor do site de busca ***Cadê*** , que possui um endereço roteável (válido) 200.244.37.76. O administrador dessa rede seguiu a RFC 1918mas agora encontra um problema: como sua rede 10.40.1.0 /24 - não roteável - vai acessar o servidor do ***Cadê*** ? A resposta é óbvia: fazendo um NAT, no caso do exemplo, no Firewall. Conforme foi mencionado anteriormente, este poderia ser feito no roteador sem problemas.Com o NAT habilitado, o usuário ao chamar a página Web em questão no seu *browser*, fará com que asua máquina envie um pacote endereçado a 200.244.37.76. O endereço IP da origem (por exemplo10.40.1.10) e a porta de origem (por exemplo a 1500) estão no pacote, assim como o endereço de destino(200.244.37.76) e a porta de destino(80).Quando o pacote chega ao Firewall, ele o deencapsulará e o reescreverá.O pacote que ele enviará para a Rede Pública conterá o endereço da interface do Firewall que está a ela conectada - ou um outro endereço previamente acertado que seja roteável - como endereço de origem, a porta de origem alocada de uma lista de portas livres no Firewall e o resto do pacote será uma cópia do pacote original.



A figura mostra uma arquitetura básica utilizada por administradores de rede para publicarem para a rede pública máquinas que deverão ser acessadas por outras redes sem, no entanto,divulgar seu endereçamento interno.Verifica-se que existem 2 servidores alocados na rede chamada DMZ (Rede não-militarizada).Essa rede é normalmente configurada para abrigar servidores que apresentam maiores riscos de serem atacados por usuários externos. Os servidores precisam ter endereços declarados na Rede Pública e, por isso, podem ser vistos por qualquer usuário - por isso a preocupação com a questão de segurança sobre essas máquinas é redobrada. Ao se colocar as mesmas em uma rede separada da rede principal (Rede Local), caso um usuário mal-intencionado consiga "atacar"a máquina e ganhar acesso à rede, este terá acesso a uma rede sem máquinas que contenham informações de maior importância. Nessa situação, é feito um NAT Estático da máquina Web de 10.40.3.3 para 200.182.30.9.



Existem 2 modelos de NAT dinâmico: **Conversão 1x1:** este modelo é pouco utilizado pois não auxilia no controle da utilização de endereços públicos. Ele diz que cada máquina solicitante da rede interna terá um endereço de tradução na rede pública. Apenas apresenta a vantagem de "esconder" o endereçamento interno.

Conversão N x M (N > M)**:** este modelo é utilizado quando a quantidade de endereços na rede interna é maior que o número de endereços presentes na faixa . É um misto entre a conversão 1x1 e o PAT (a ser definido em seguida). O tradutor, ao receber as requisições, vai utilizar os endereços da faixa como se estivesse fazendo uma conversão 1x1. Ao esgotarem-se os endereços , ele começa a fazer Port Address Translation - PAT.

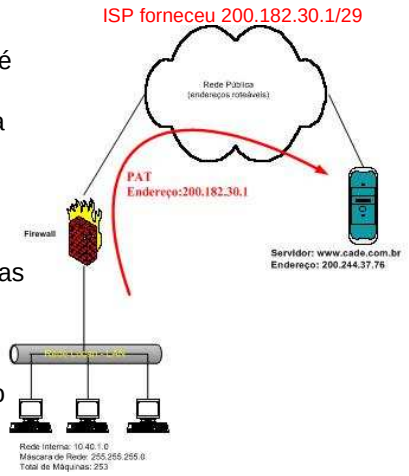
Roteamento NAT (Network Address Translation)

Tipos de NAT: PAT (Port Address Translation)

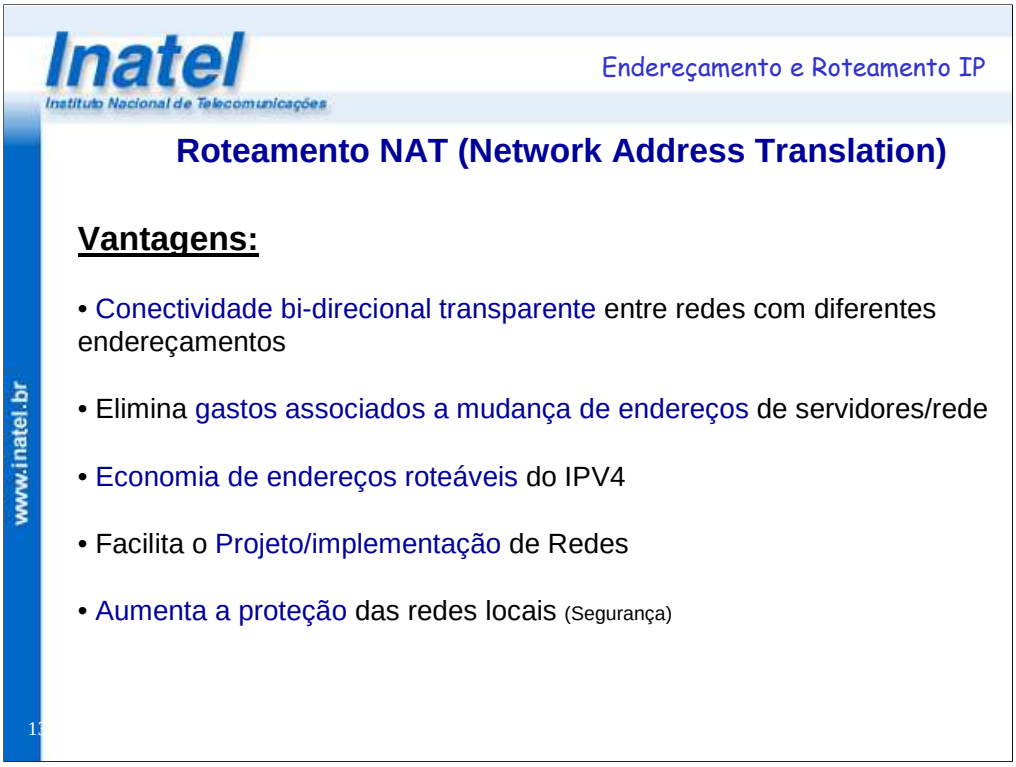
• É o tipo de NAT que mais economiza endereços válidos(roteáveis) pois a tradução é feita no modelo N para 1, ou seja, todos os endereços da Rede Local são traduzidos para um único endereço válido.

• Esse tipo de NAT é, na verdade, um caso especial do NAT dinâmico pois neste caso, assim como no anterior, as traduções são feitas sob demanda, ou seja,só existe a tradução quando houver uma requisição realizada.

• Este modelo apresenta uma limitação para o número máximo de conexões simultâneas (número máximo de portas = 65535).



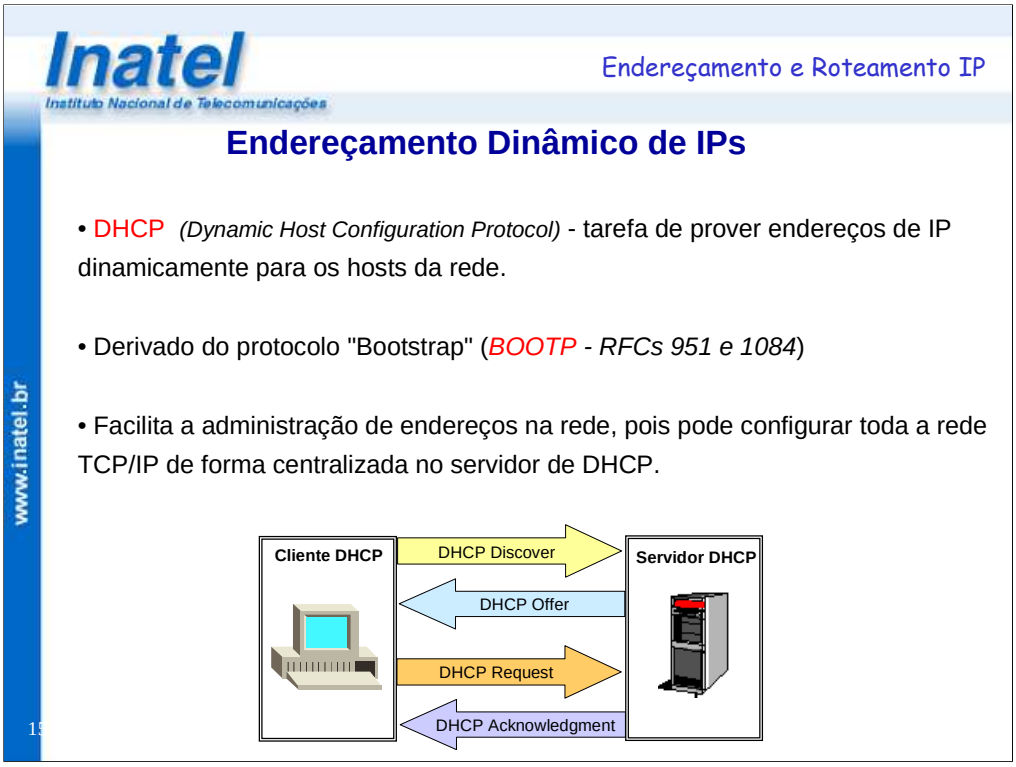
No exemplo, temos usuários alocados na rede 10.40.1.0/24 - Rede Interna - querendo acessar o site http://www.cade.com.br/ . O administrador de rede tem aproximadamente 250 usuários em sua rede local querendo acessar a Internet porém o ISP (Internet Service Provider) só lhe forneceu um range de endereços 200.182.30.0 /29 - o que dá apenas 2 endereços válidos. Com essa escassez de endereços, a única solução para garantir acesso simultâneo a todos é o PAT. Nesse exemplo, todos saem com o endereço 200.182.30.1.



Conectividade bi-direcional transparente entre redes com diferentes endereçamentos: Faz com que seja transparente para os elementos de rede que não estejam diretamente envolvidos com a tradução a utilização do NAT. Para eles, o pacote IP que sofreu NAT é um pacote como outro qualquer. **Eliminam se gastos associados a mudança de endereços de servidores/rede:** qualquer máquina que quisesse ser visualizada na Rede Pública deveria possuir endereços válidos. Nessa situação, qualquer alteração no endereçamento de uma simples máquina implicaria em replicar a mudança em TODAS as máquinas roteáveis. **Economia de endereços roteáveis do IPV4:** Há uma melhor gerência do endereçamento IP com o NAT. Costuma-se dizer que passa a haver uma utilização racional de endereços. **Facilita o desenho/implementação de Redes:** Devido ao uso racional de IPs, há uma menor preocupação com a criação dos mapas de endereçamento de rede, facilitando a implementação/interligação das mesmas. **Aumenta a proteção das redes locais:** O NAT evita que se precise publicar o endereçamento interno das redes locais nas redes públicas. Assim,fica mais difícil para um usuário mal-intencionado montar qualquer tipo de ataque direto à Rede Interna.

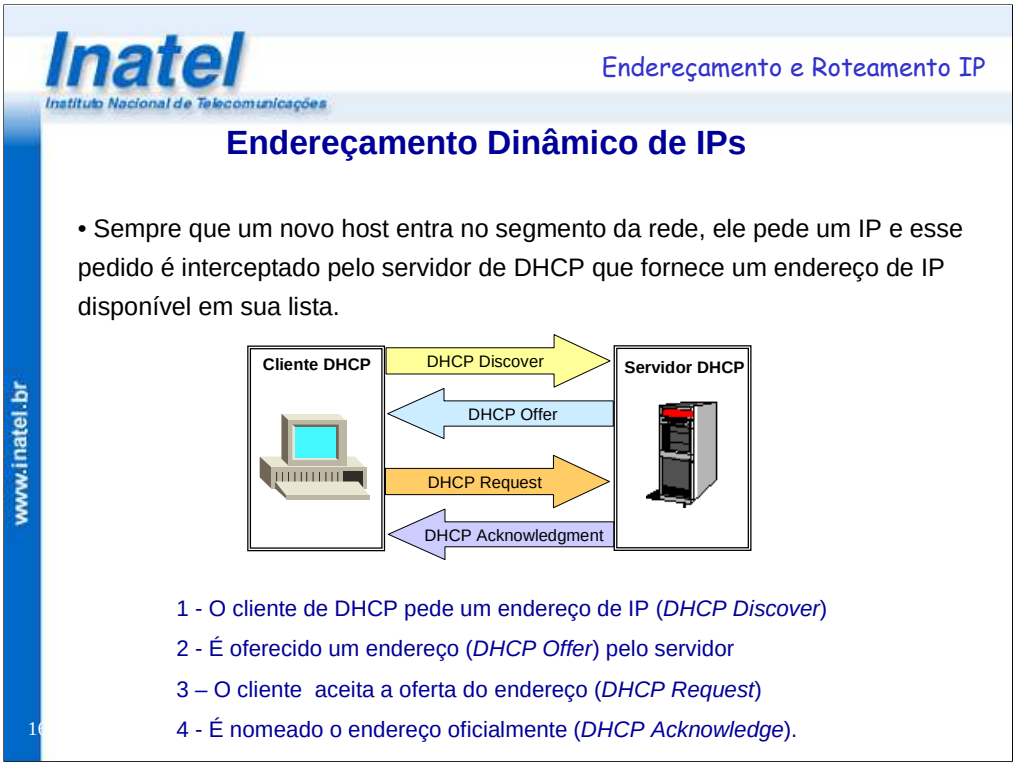
Impossibilidade de se rastrear o caminho do pacote: Com a utilização da tradução, fica impossível se utilizar o comando *traceroute <endereço de destino>*para se identificar o caminho que o pacote segue até encontrar o seu destino, pois o elemento tradutor não permite a tradução reversa (resposta da rede externa para a local) com resposta indicando "esgotado tempo limite" - TTL (*Time to Live*).

Aumento do processamento no dispositivo tradutor: O NAT requer que a máquina que fará a tradução altere o pacote IP. Essa manobra exige que a máquina deixe dedicado para essa tarefa parte do seu potencial de processamento. Por essa razão, há que se ter cuidado na escolha do tradutor. Deve-se atentar para a demanda extra de processamento.



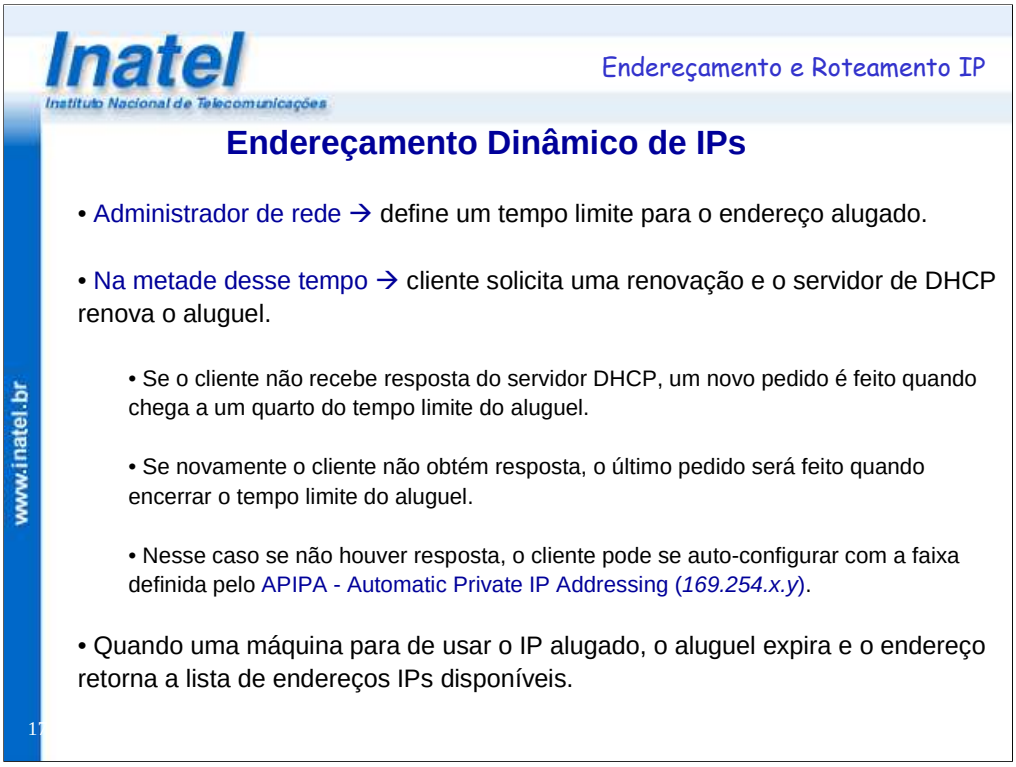
O Dynamic Host Configuration Protocol é derivado do Protocolo "Standard Bootstrap" (*BOOTP - RFCs 951 e 1084*) que o concedeu a tarefa de prover endereços de IP Dinâmicos. Além de fornecer endereços de IP dinamicamente, o DHCP provê todos os dados de configuração requeridos pelo TCP/IP além de dados adicionais requeridos para servidores específicos.

Para que os endereços não se percam (*caso um cliente se conecte só uma vez*), os administradores de rede definem um tempo limite para o endereço alugado. Quando chega a metade desse tempo, o cliente solicita uma renovação e o servidor de DHCP estende o aluguel. Caso o cliente não receba resposta do servidor DHCP, um novo pedido é feito quando chega a um quarto do tempo limite do aluguel. Se novamente o cliente não obtém resposta o último pedido será feio quando encerrar o tempo limite do aluguel. Nesse caso se não houver resposta o cliente pode se auto-configurar com a faixa definida pelo APIPA (*169.254.x.y*) onde x e y são números aleatórios (*x = 0 - e 255, y = 1 - 254*). Caso o número escolhido já esteja em uso na rede outro número será escolhido até que se encontre um disponível. Quando uma máquina para de usar o IP alugado (*por exemplo, caso ela tenha sido movida para outro segmento*), o aluguel expira e o endereço retorna a lista de realocação.



Como você pode notar isso facilita a vida do administrador de rede pois ele pode configurar toda sua rede TCP/IP de forma centralizada no servidor de DHCP. Sempre que um novo host entra no segmento da rede, ele é servido pôr este servidor. A máquina pede um IP e esse pedido é interceptado pelo servidor de DHCP que fornece um endereço de IP disponível em sua lista.

Para que os endereços não se percam (*caso um cliente se conecte só uma vez*), os administradores de rede definem um tempo limite para o endereço alugado. Quando chega a metade desse tempo, o cliente solicita uma renovação e o servidor de DHCP estende o aluguel.



Caso o cliente não receba resposta do servidor DHCP, um novo pedido é feito quando chega a um quarto do tempo limite do aluguel. Se novamente o cliente não obtém resposta o último pedido será feio quando encerrar o tempo limite do aluguel. Nesse caso se não houver resposta o cliente pode se auto-configurar com a faixa definida pelo APIPA (*169.254.x.y*) onde x e y são números aleatórios (*x = 0 - e 255, y = 1 - 254*). Caso o número escolhido já esteja em uso na rede outro número será escolhido até que se encontre um disponível. Quando uma máquina para de usar o IP alugado (*por exemplo, caso ela tenha sido movida para outro segmento*), o aluguel expira e o endereço retorna a lista de realocação.

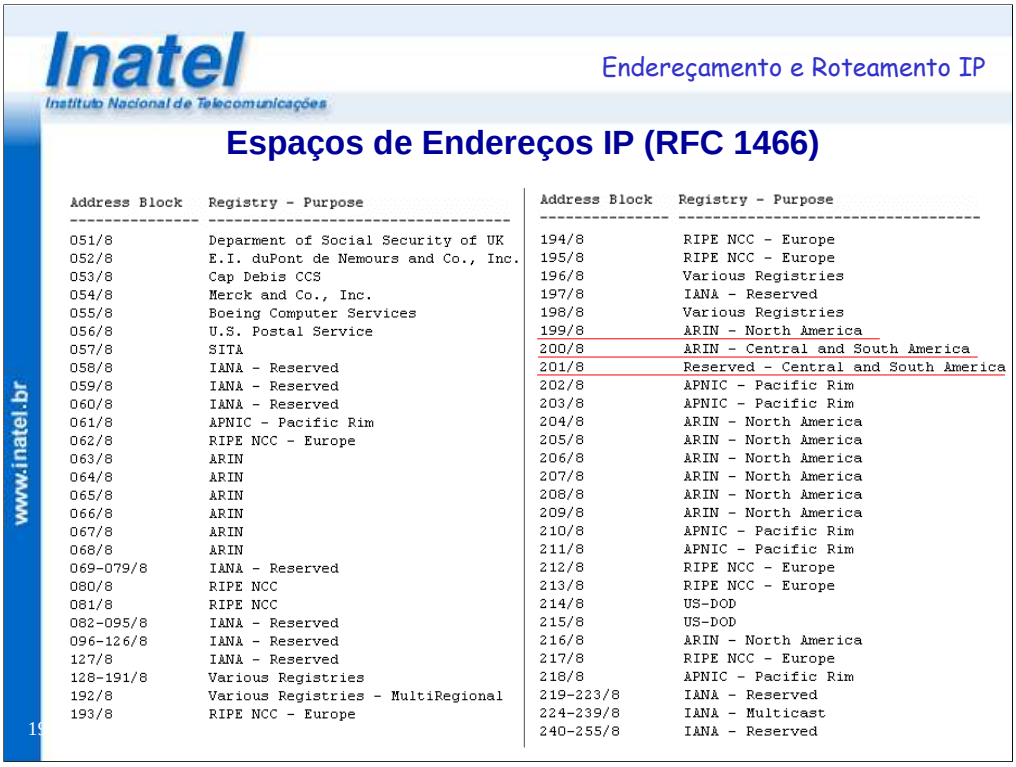
Address Block	Registry - Purpose
000/8	IANA - Reserved
001/8	IANA - Reserved
002/8	IANA - Reserved
003/8	General Electric Company
004/8	Bolt Beranek and Newman Inc.
005/8	IANA - Reserved
006/8	Army Information Systems Center
007/8	IANA - Reserved
008/8	Bolt Beranek and Newman Inc.
009/8	IBM
010/8	IANA - Private Use
011/8	DoD Intel Information Systems
012/8	AT&T Bell Laboratories
013/8	Xerox Corporation
014/8	IANA - Public Data Network
015/8	Hewlett-Packard Company
016/8	Digital Equipment Corporation
017/8	Apple Computer Inc.
018/8	MIT
019/8	Ford Motor Company
020/8	Computer Sciences Corporation
021/8	DDN-RVN
022/8	Defense Information Systems Agency
023/8	IANA - Reserved
024/8	ARIN - Cable Block

Address Block	Registry - Purpose
025/8	Royal Signals and Radar Establishment
026/8	Defense Information Systems Agency
027/8	IANA - Reserved
028/8	DSI-North
029/8	Defense Information Systems Agency
030/8	Defense Information Systems Agency
031/8	IANA - Reserved
032/8	Norsk Informasjonsteknologi
033/8	DLA Systems Automation Center
034/8	Halliburton Company
035/8	MERIT Computer Network
036/8	IANA - Reserved (Formerly Stanford University - Apr 93)
037/8	IANA - Reserved
038/8	Performance Systems International
039/8	IANA - Reserved
040/8	Eli Lily and Company
041/8	IANA - Reserved
042/8	IANA - Reserved
043/8	Japan Inet
044/8	Amateur Radio Digital Communications
045/8	Interop Show Network
046/8	Bolt Beranek and Newman Inc.
047/8	Bell-Northern Research
048/8	Prudential Securities Inc.
049/8	Joint Technical Command Returned to IANA

A tabela ilustra a alocação dos blocos de endereço IP com os respectivos registros no IANA para tal uso.

Note que na tabela estão marcados alguns endereços registrados a título apenas de demonstração.

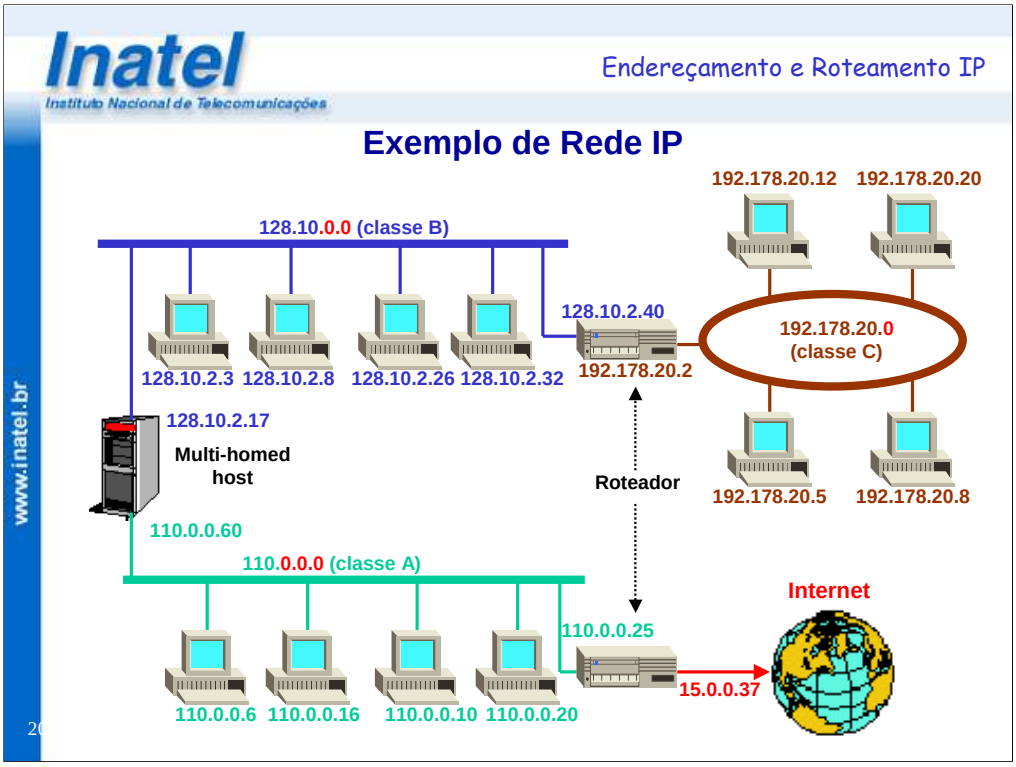
Esta tabela está presente na RFC de número 1466.



Address Block	Registry - Purpose
051/8	Deparment of Social Security of UK
052/8	E.I. duPont de Nemours and Co., Inc.
053/8	Cap Debis CCS
054/8	Merck and Co., Inc.
055/8	Boeing Computer Services
056/8	U.S. Postal Service
057/8	SITA
058/8	IANA - Reserved
059/8	IANA - Reserved
060/8	IANA - Reserved
061/8	APNIC - Pacific Rim
062/8	RIPE NCC - Europe
063/8	ARIN
064/8	ARIN
065/8	ARIN
066/8	ARIN
067/8	ARIN
068/8	ARIN
069-079/8	IANA - Reserved
080/8	RIPE NCC
081/8	RIPE NCC
082-095/8	IANA - Reserved
096-126/8	IANA - Reserved
127/8	IANA - Reserved
128-191/8	Various Registries
192/8	Various Registries - MultiRegional
193/8	RIPE NCC - Europe

Address Block	Registry - Purpose
194/8	RIPE NCC - Europe
195/8	RIPE NCC - Europe
196/8	Various Registries
197/8	IANA - Reserved
198/8	Various Registries
199/8	ARIN - North America
200/8	ARIN - Central and South America
201/8	Reserved - Central and South America
202/8	APNIC - Pacific Rim
203/8	APNIC - Pacific Rim
204/8	ARIN - North America
205/8	ARIN - North America
206/8	ARIN - North America
207/8	ARIN - North America
208/8	ARIN - North America
209/8	ARIN - North America
210/8	APNIC - Pacific Rim
211/8	APNIC - Pacific Rim
212/8	RIPE NCC - Europe
213/8	RIPE NCC - Europe
214/8	US-DOD
215/8	US-DOD
216/8	ARIN - North America
217/8	RIPE NCC - Europe
218/8	APNIC - Pacific Rim
219-223/8	IANA - Reserved
224-239/8	IANA - Multicast
240-255/8	IANA - Reserved

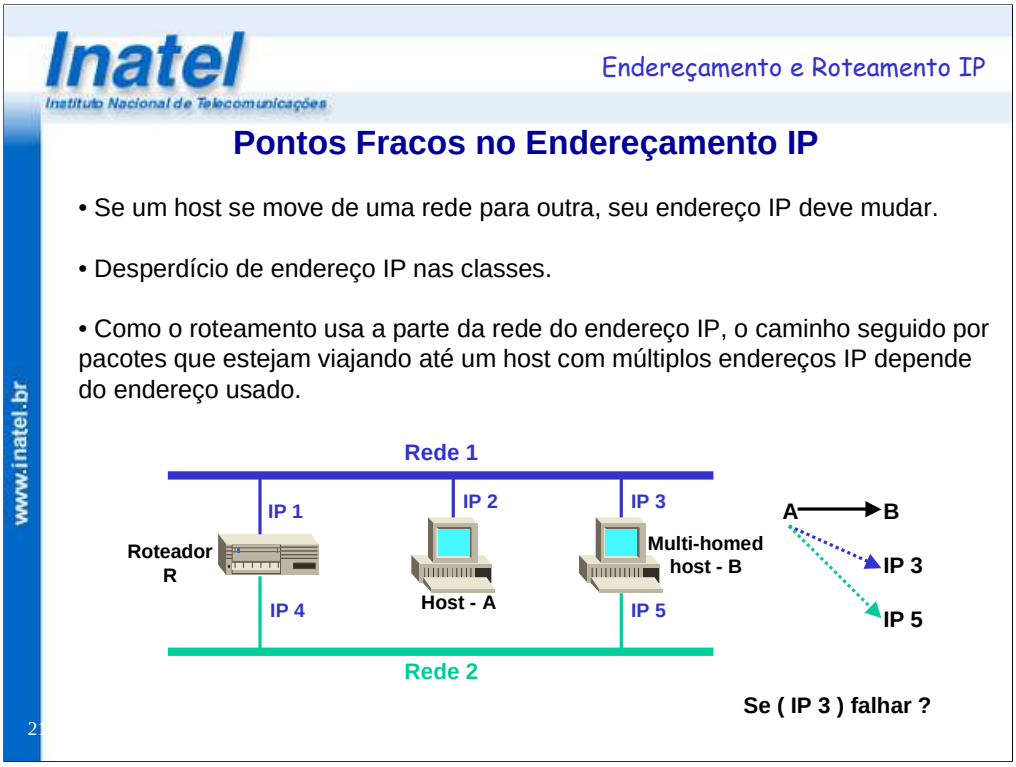
Continuação da tabela anterior.



O exemplo mostra três redes e os números de rede que lhes foram designados: uma rede Ethernet classe B 128.10.0.0, outra rede Ethernet classe A 110.0.0.0 e uma rede Token-Ring classe C 192.178.20.0.

Os roteadores servem para interligar as redes e para cada conexão com cada rede ele possui um endereço IP (lembre-se o endereço IP representa uma conexão de rede e não uma máquina!).

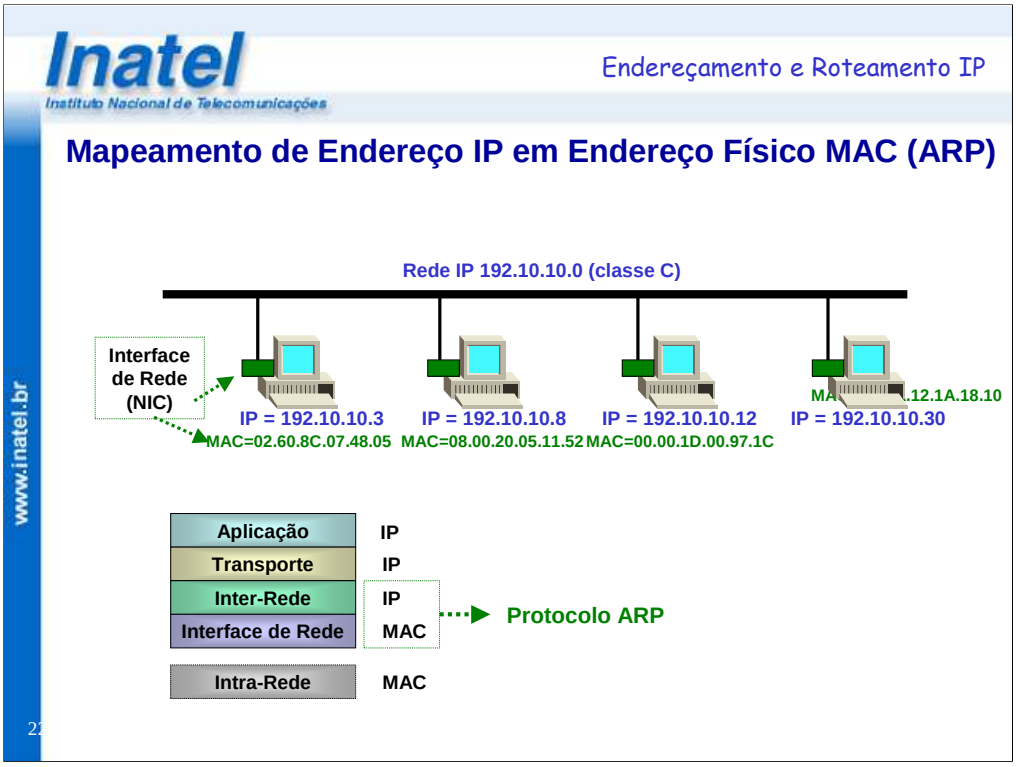
Aparece também neste exemplo um host Multi-Homed tem conexões com as duas redes Ethernet, podendo assim alcançar destinos diretamente em qualquer rede. Apesar de um host multi-homed poder ser configurado para rotear pacotes entre as duas redes, a maioria dos sites usa computadores dedicados como roteadores, a fim de evitar sobrecarregar sistemas de computação convencionais com o processamento requerido pelo roteamento.



Codificar informações de redes em um endereço de interligação em redes pode Ter suas desvantagens. A mais óbvia delas é que os endereços referem-se às conexões de redes, e não às máquinas (hosts). Surge aí a primeira falha no endereçamento IP: se um host se mover de uma rede para outra, seu endereço IP deverá mudar também.

Outra falha do esquema do endereçamento da interligação em redes é que, quando qualquer rede classe C cresce além de 254 hosts, deve ter seu endereço mudado para um endereço classe B.

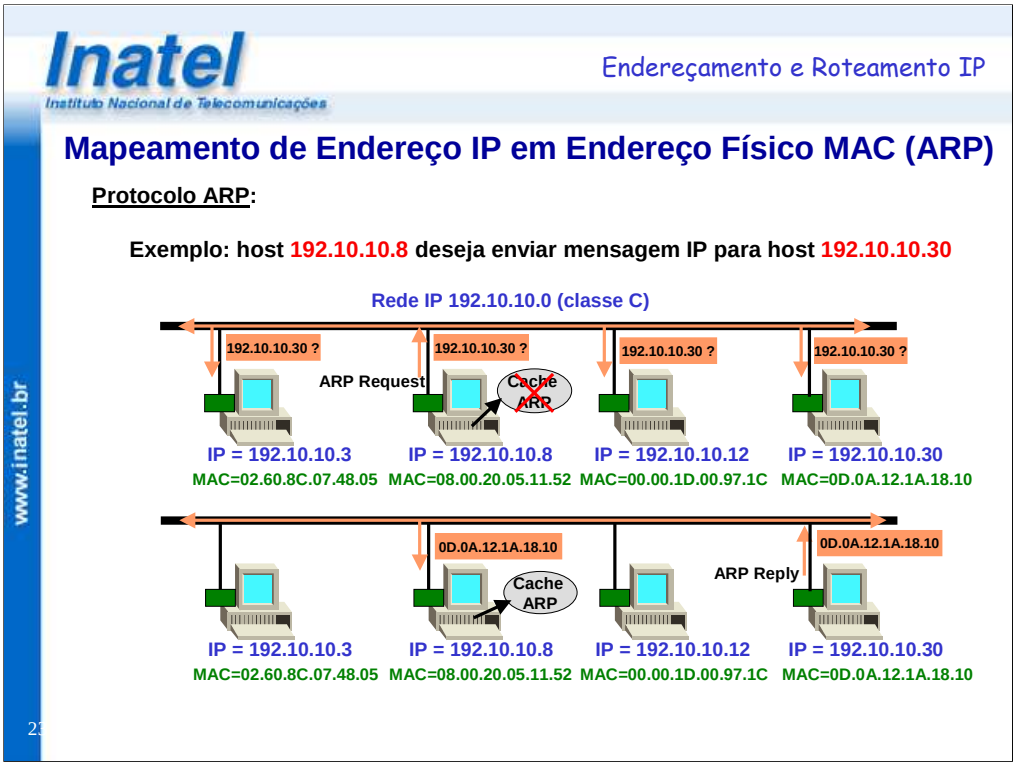
A principal falha do esquema de endereçamento da interligação em redes é com relação ao roteamento. O roteamento baseia-se em endereços de interligação em redes, com a parte net-id de um endereço usado para tomar decisões de roteamento. Se considerarmos que um host com duas conexões para a interligação em redes ele deverá ter dois endereços IP. Como o roteamento usa a parte da rede (net-id) do endereço IP, o caminho seguido por pacotes que estejam viajando até um host com múltiplos endereços IP depende do endereço usado.



Os protocolos de rede compartilhada como Ethernet, Token-Ring e FDDI possuem um endereço próprio para identificar as diversas máquinas situadas na rede. Nessas redes, o endereçamento utilizado é chamado de endereço físico ou endereço MAC (Medium Access Control), formado por 6 bytes, que identificam o fabricante da interface de rede e um número da interface de rede para aquele fabricante.

Este tipo de endereçamento só é útil para identificar diversas máquinas, não possuindo nenhuma informação capaz de distinguir redes distintas. Para que uma máquina com protocolo TCP / IP envie um pacote para outra máquina situada na mesma rede, ela deve se basear no protocolo físico da rede, já que é necessário saber o endereço físico da interface de rede (endereço MAC). Como o protocolo TCP / IP só identifica uma máquina pelo endereço IP, deve haver um mapeamento entre o endereço IP e o endereço MAC. Este mapeamento é realizado pelo protocolo ARP (Address Resolution Protocol).

O protocolo ARP permite que um host encontre o endereço físico (MAC) de um host de destino na mesma rede física, apresentando somente o endereço IP de destino.



O funcionamento do protocolo ARP é descrito a seguir:

- O host origem verifica que o host destino está na mesma rede local, determinado através dos endereços IP de rede origem e destino e suas respectivas classes.

- O protocolo IP do host origem verifica que ainda não possui um mapeamento do endereço MAC para o endereço IP do host de destino.

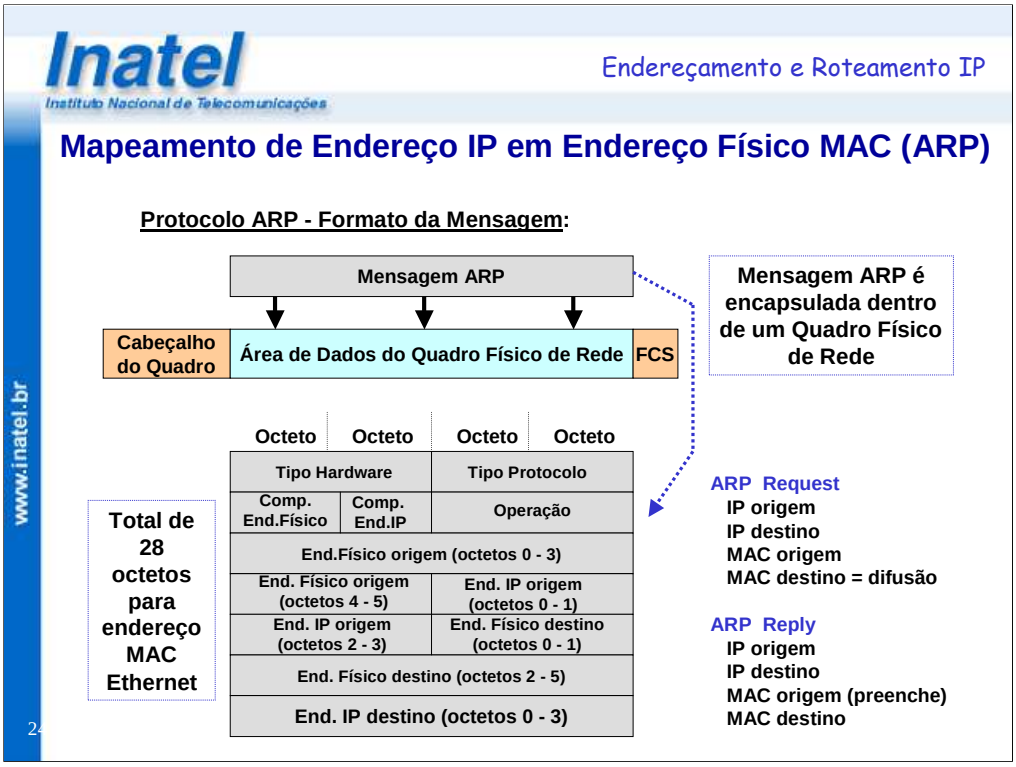
- O protocolo IP solicita ao protocolo ARP que o endereço MAC é necessário.

- O protocolo ARP envia um pacote ARP Request com o endereço MAC destino de broadcast (difusão para todos os hosts da rede).

- .A mensagem ARP enviada é encapsulada em um pacote físico (frame ou quadro) da rede utilizada.

- Todos os hosts recebem o pacote ARP, mas somente aquela que possui o endereço IP especificado responde. O host destino já instala em sua tabela ARP (cache ARP) o mapeamento do endereço IP de origem para o endereço MAC de origem.

- A resposta é enviada no pacote físico (frame ou quadro) da rede, encapsulado através de uma mensagem ARP Reply endereçado diretamente para o host origem.



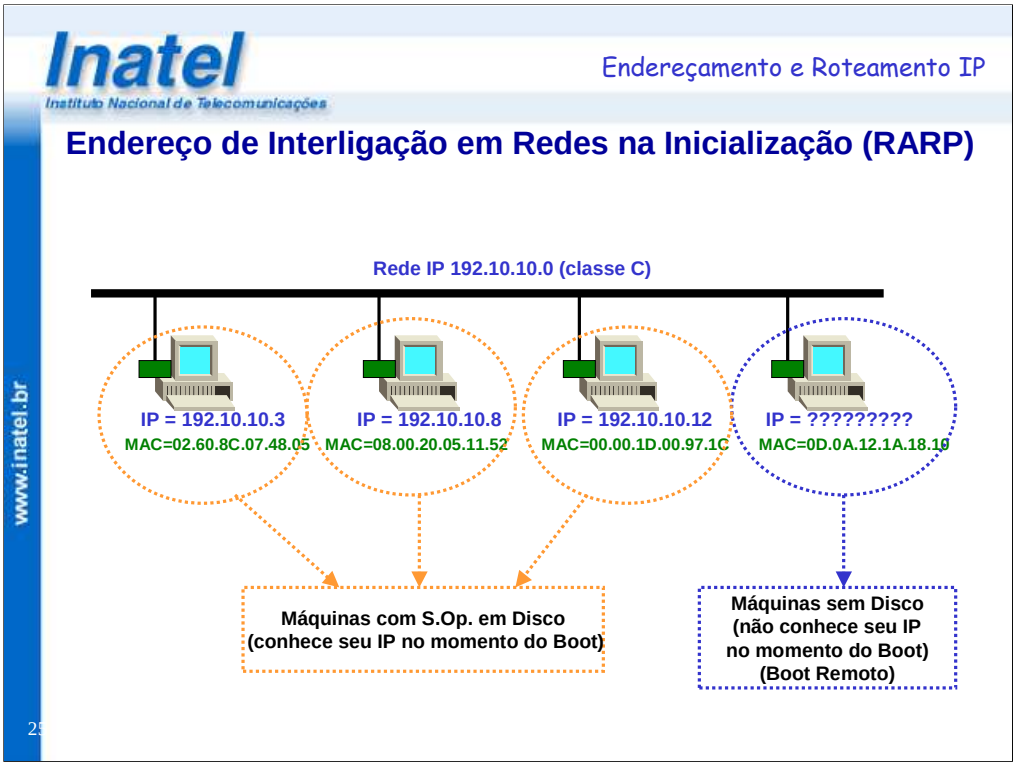
Quando o ARP faz o trajeto de uma máquina (host) para outra, devem ser carregados em quadros físicos (frames). Para identificar um quadro carregando uma mensagem ARP, o transmissor designa um valor especial para o tipo de campo no cabeçalho do quadro e coloca a mensagem ARP no campo de dados do quadro. Quando um quadro chega ao computador, o software de rede usa o tipo de quadro para determinar seu conteúdo.

O formato da mensagem ARP é suficientemente geral para permitir que seja usado com endereços físicos diversos e endereços de protocolo diversos. A mensagem ARP possui 28 octetos para um hardware de rede Ethernet (onde os endereços físicos são de 6 octetos) ao converter endereços de protocolo IP (4 octetos).

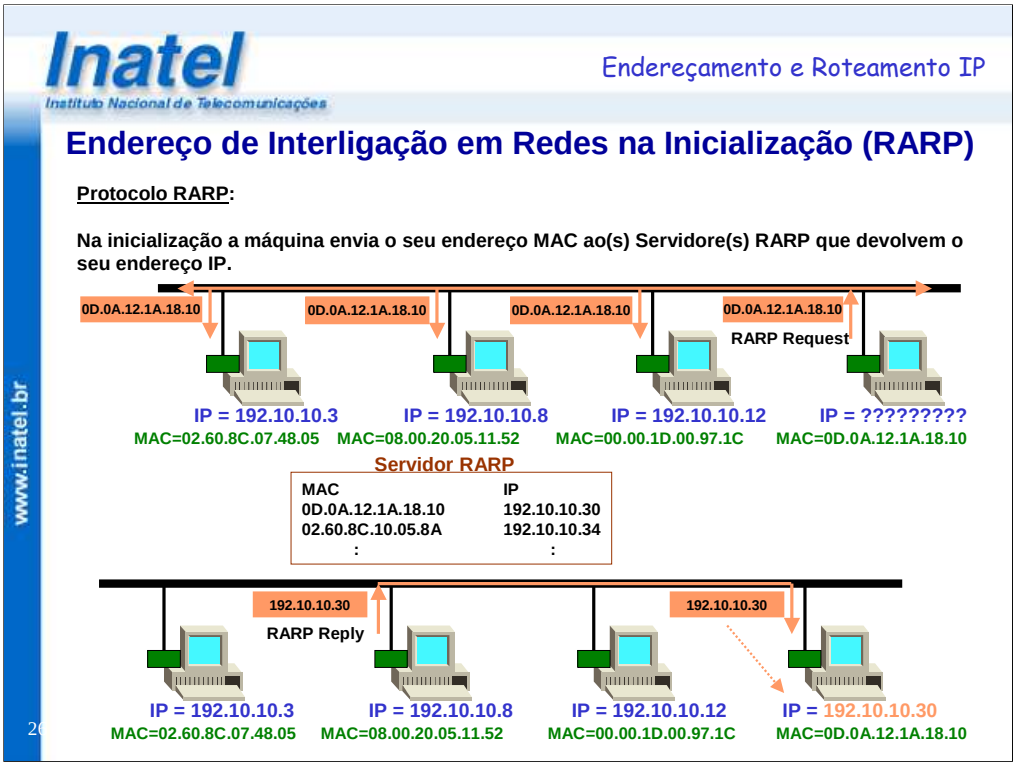
A seguir são descritos os campos do formato da mensagem ARP:

Tipo de Hardware: especifica um tipo de interface de hardware para o qual o transmissor pede uma resposta.

Tipo de protocolo: especifica o tipo de endereço de protocolo de alto nível que o remetente forneceu.

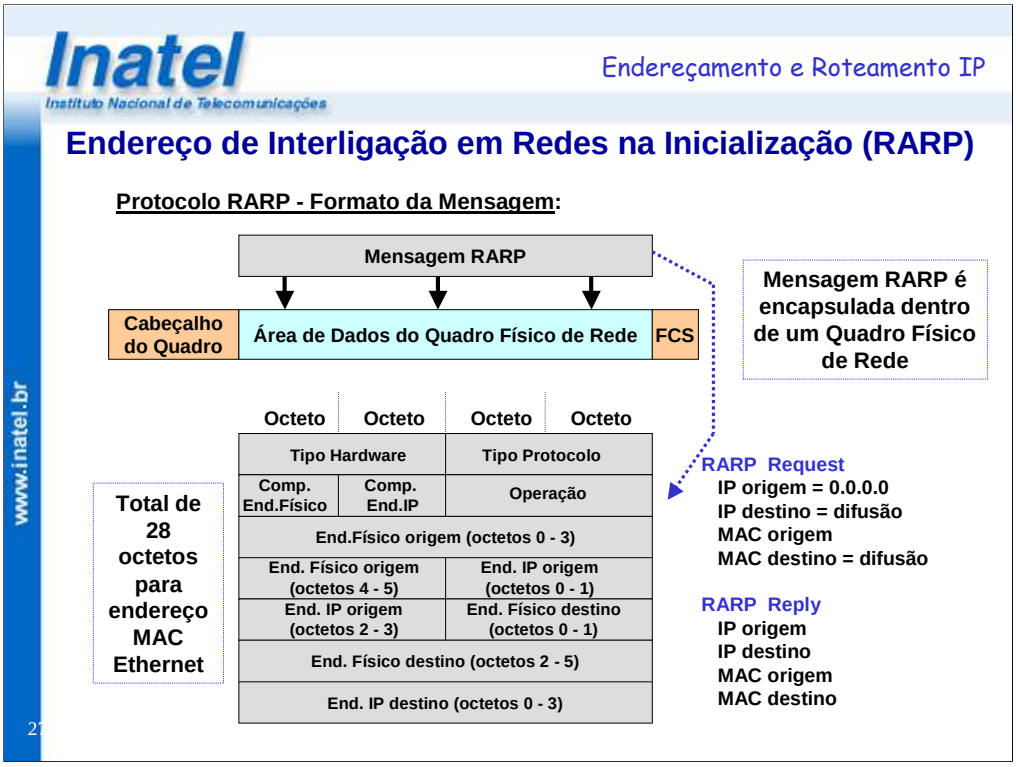


Normalmente, o endereço IP de uma máquina é mantido em sua memória secundária, onde o Sistema Operacional o encontra durante a inicialização. A questão é: como uma máquina sem um disco permanentemente conectado determina seu endereço IP ? O problema é crucial para estações de trabalho que armazenam arquivos de inicialização em um servidor remoto, porque essas máquinas precisam de um endereço IP antes que possam utilizar os protocolos padrão de transferência arquivos TCP / IP para obter sua imagem de inicialização.

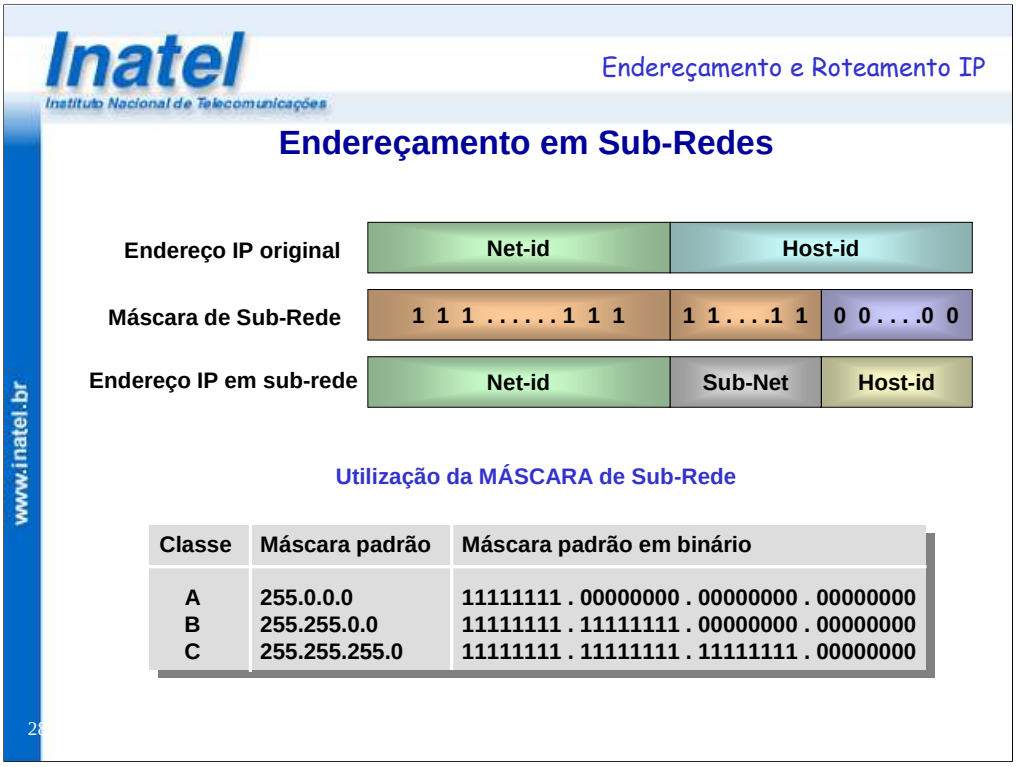


Assim como na mensagem ARP, a mensagem RARP é enviada de uma máquina a outra encapsulada na porção de dados de um quadro de rede.

No protocolo RARP o transmissor difunde uma solicitação RARP que se especifica tanto como máquina transmissora quanto como máquina, e fornece seu endereço de rede física (MAC) no campo de endereço de hardware de destino. Todas as máquinas recebem a solicitação, mas apenas aquelas autorizadas a fornecer o serviço RARP processam a informação e enviam uma resposta; tais máquinas são conhecidas como Servidores RARP. Os servidores respondem às solicitações preenchendo o campo de endereço de protocolo de destino, mudando o tipo de mensagem para RARP Reply e enviando a resposta de volta diretamente à máquina que fez a solicitação. A máquina que originou o RARP recebe respostas de todos os servidores RARP, mesmo se apenas a primeira for necessária.



O formato da mensagem RARP é idêntico ao da mensagem ARP. E seus campos possuem a mesma descrição já mostrada no formato da mensagem ARP.



Devido ao crescimento explosivo da Internet, o princípio de endereços IP atribuídos tornou-se inflexível demais para permitir mudanças fáceis para as configurações de rede local. Estas mudanças podem ocorrer quando:

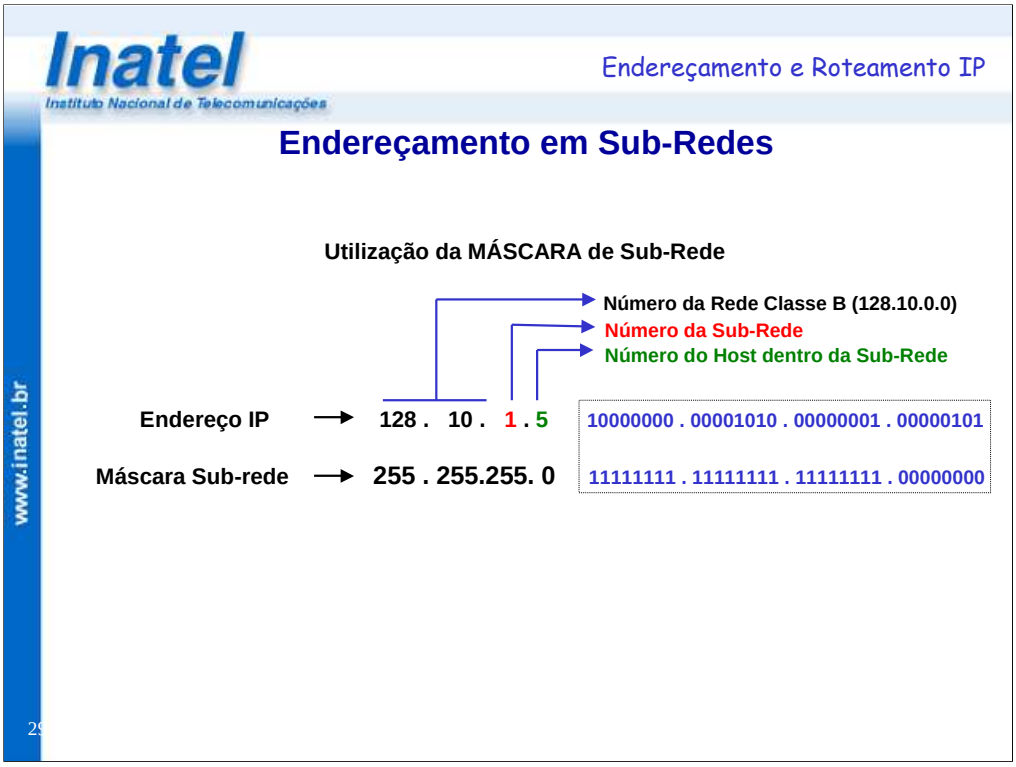
• Um novo tipo de rede física é instalado em um local;

• O crescimento do número de hosts exige a divisão da rede local em duas ou mais redes separadas;

• As distâncias crescentes exigem a divisão de uma rede em redes menores, com gateways (roteadores) entre elas.

Para evitar a solicitação de endereços adicionais de rede IP nestes casos, o conceito de sub-redes foi introduzido. A designação de sub-redes pode ser feita localmente, já que toda a rede ainda aparece como uma rede IP para o mundo externo.

A parte do número do host do endereço IP (host-id) é subdividida em um número de sub-rede e um número de host. A rede principal consiste agora de um número de rede e um número de sub-rede e o endereço IP é interpretado como:

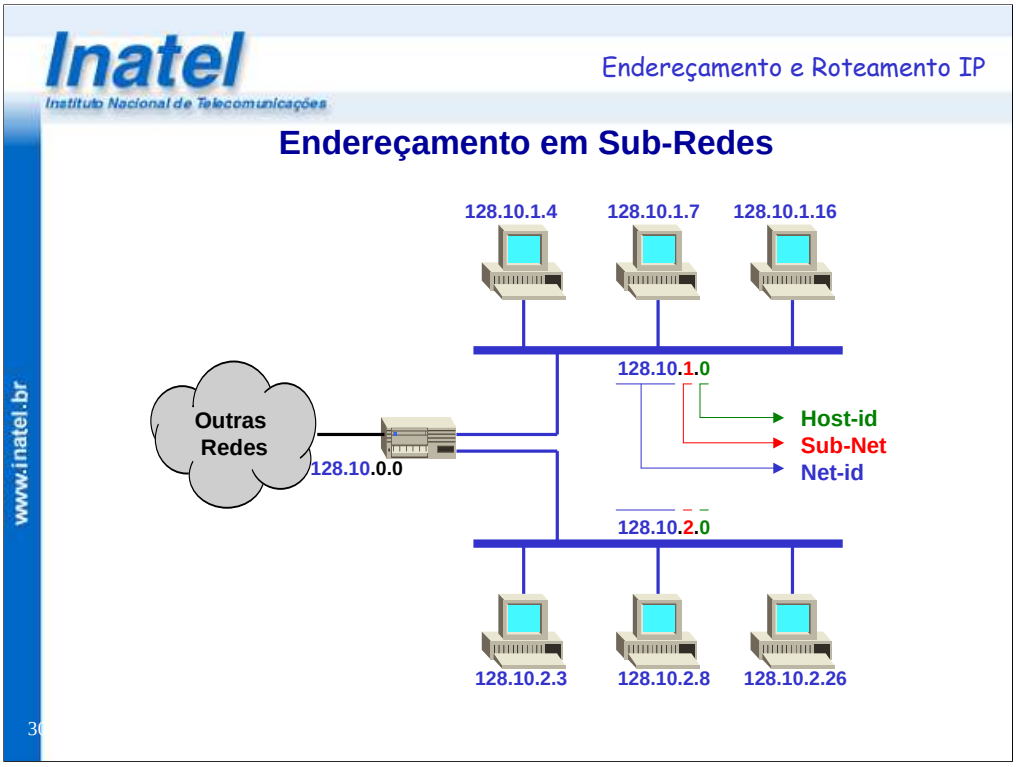


O tratamento especial de todos os bits 0 e todos os bits 1 aplicam-se a cada uma das partes de um endereço IP subdividido em redes. Por exemplo uma rede classe B subdividida em redes, que tem uma parte local de 16 bits, poderia usar um dos seguintes esquemas para sub-redes:

• O primeiro byte do endereço local é o número da sub-rede; o segundo byte é o número do host. Isto nos dá $2^8 – 2$ (254) sub-redes possíveis, cada uma tendo até $2^8 – 2$ (254) hosts. A máscara de sub-rede é 255.255.255.0

• Os 12 primeiros bits do endereço local são usados para o número de sub-rede e os últimos 4 para o número do host. Isto nos dá $2^{12} – 2$ (4094) sub-redes possíveis mas apenas $2^4 – 2$ (14) hosts em cada sub-rede. A máscara de sub-rede é 255.255.255.240

Há muitas outras possibilidades possíveis. O administrador tem a escolha de definir tanto um número grande de sub-redes com um número pequeno de hosts em cada uma, quanto um número pequeno de sub-redes com muitos hosts em cada.



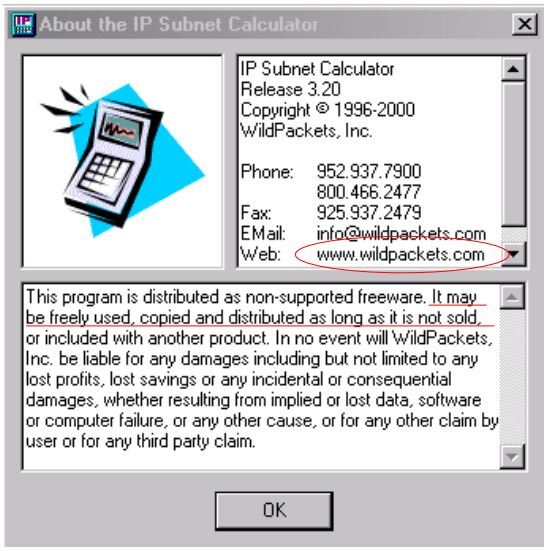
No exemplo mostrado, o site está usando o único endereço de rede classe B 128.10.0.0 para duas redes. Exceto o roteador deste site, todos os roteadores das outras redes roteiam como se ela fosse uma única rede física. Uma vez que um pacote alcance o roteador, ele deve ser enviado pela rede física correta até seu destino.

Para tornar eficaz a opção da rede física, o site optou por utilizar o terceiro octeto do endereço para estabelecer um distinção entre as duas redes (sub-redes). O administrador atribui às máquinas de uma rede física endereços da forma 128.10.1.X, e às máquinas da outra rede física endereços da forma 128.10.2.X, onde X representa o número do host dentro da sub-rede.

Para escolher uma das redes físicas (sub-redes) o roteador do site examina o terceiro octeto do endereço de destino e roteia os datagramas para a sub-rede selecionada.

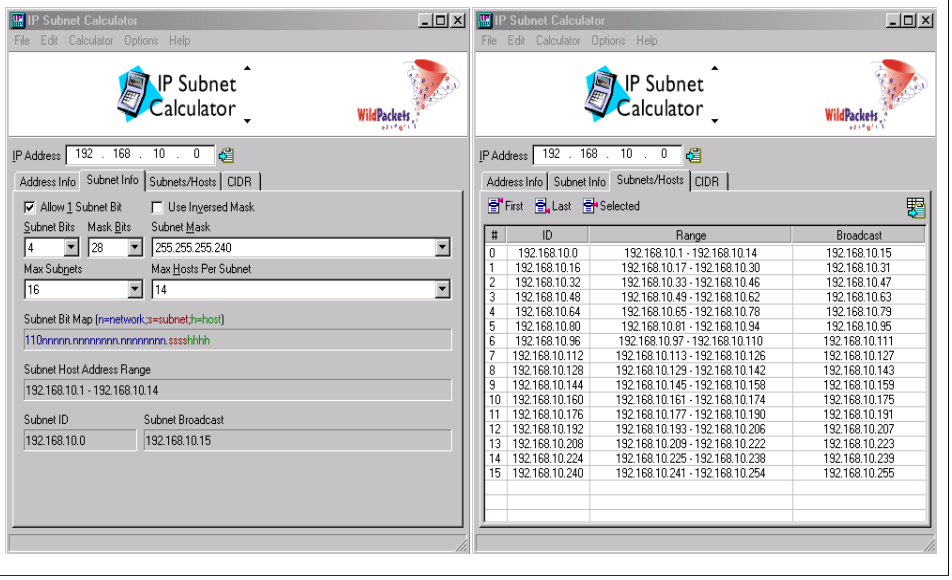
Conceitualmente, adicionar as sub-redes altera apenas ligeiramente a interpretação dos endereços IP.

Uso de uma ferramenta de software para calcular Sub-nets

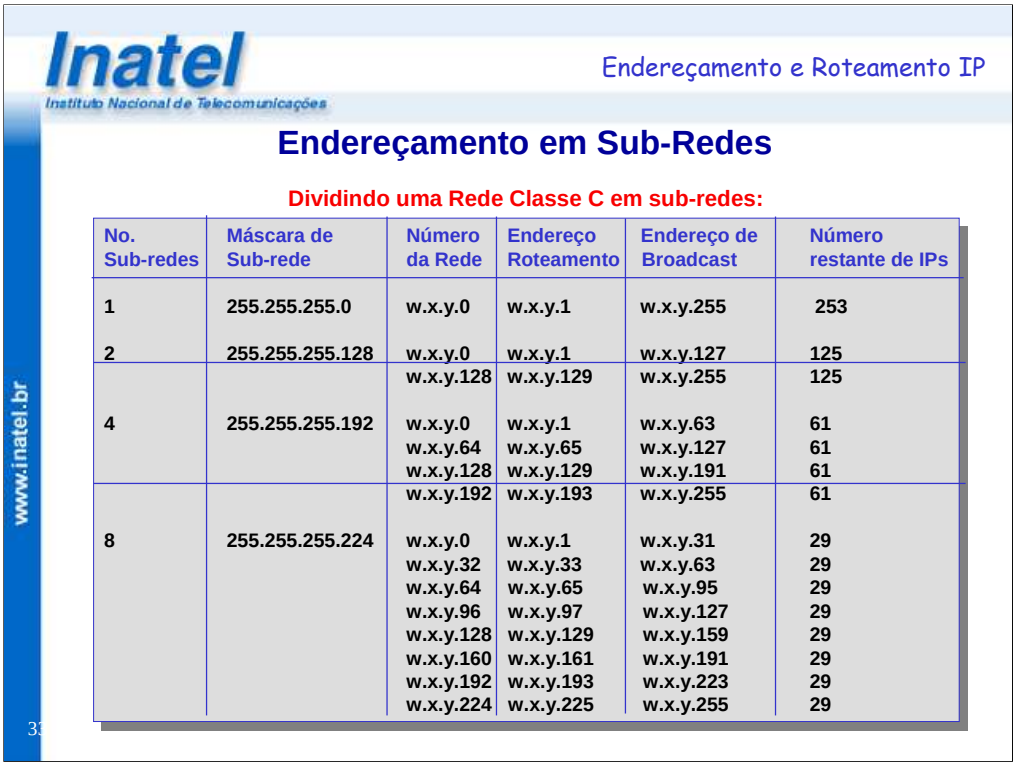


Para o cálculo de sub-redes e super-redes (CIDR) podemos utilizar uma ferramenta de software com distribuição livre. Trata-se do “IP Subnet Calculator”.

Uso de uma ferramenta de software para calcular Sub-nets

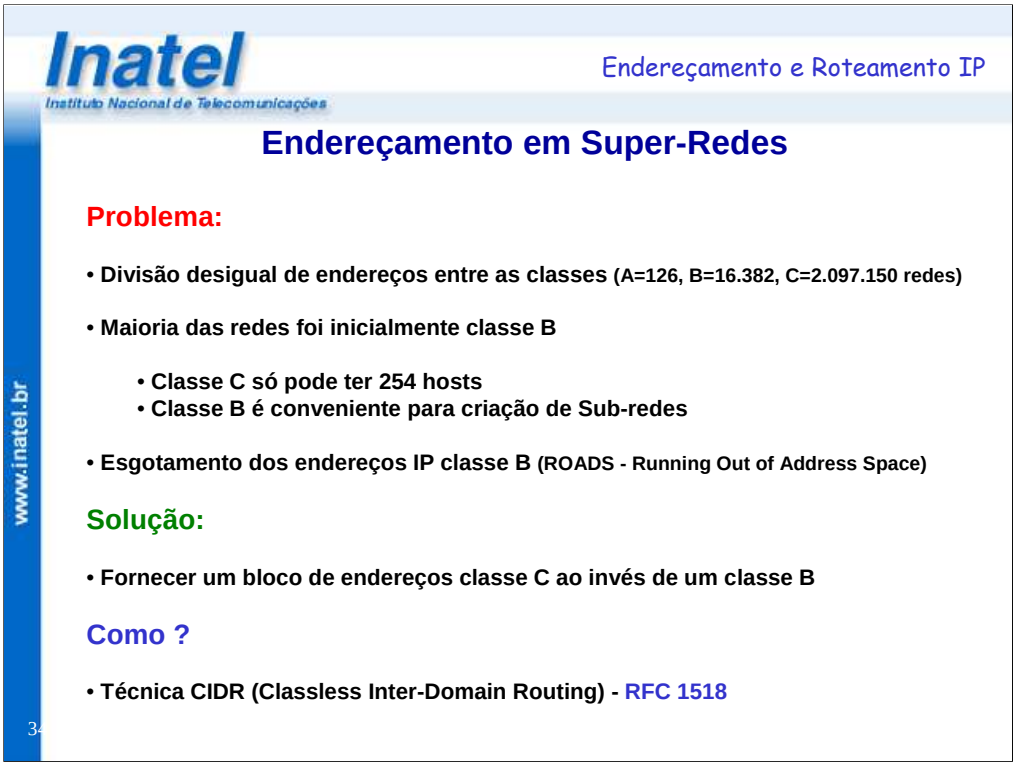


Um exemplo de calculo de sub-redes utilizando a calculadora IP Subnet Calculator.



No. Sub-redes	Máscara de Sub-rede	Número da Rede	Endereço Roteamento	Endereço de Broadcast	Número restante de IPs
1	255.255.255.0	w.x.y.0	w.x.y.1	w.x.y.255	253
2	255.255.255.128	w.x.y.0	w.x.y.1	w.x.y.127	125
		w.x.y.128	w.x.y.129	w.x.y.255	125
4	255.255.255.192	w.x.y.0	w.x.y.1	w.x.y.63	61
		w.x.y.64	w.x.y.65	w.x.y.127	61
		w.x.y.128	w.x.y.129	w.x.y.191	61
		w.x.y.192	w.x.y.193	w.x.y.255	61
8	255.255.255.224	w.x.y.0	w.x.y.1	w.x.y.31	29
		w.x.y.32	w.x.y.33	w.x.y.63	29
		w.x.y.64	w.x.y.65	w.x.y.95	29
		w.x.y.96	w.x.y.97	w.x.y.127	29
		w.x.y.128	w.x.y.129	w.x.y.159	29
		w.x.y.160	w.x.y.161	w.x.y.191	29
		w.x.y.192	w.x.y.193	w.x.y.223	29
		w.x.y.224	w.x.y.225	w.x.y.255	29

A divisão de uma rede classe C em sub-redes, é bastante simples. Para facilitar a visualização desta divisão, a tabela ilustra a divisão em uma, duas, quatro e oito sub-redes, com correspondentes números para máscaras de sub-rede, número da rede, endereço de broadcast, e endereços de roteamento.



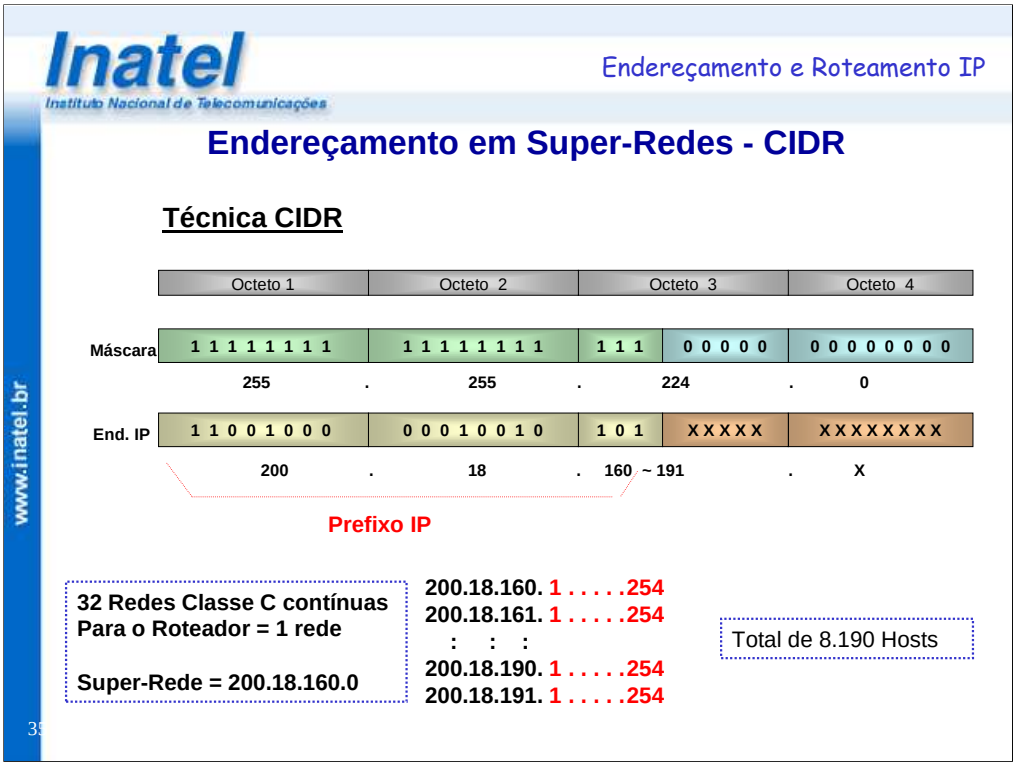
O esquema de endereçamento em super-rede tem uma abordagem aposta à do endereçamento em sub-rede. Em vez de utilizar um único endereço IP de rede para múltiplas redes físicas em determinada organização, a divisão em super-rede permite o uso de muitos endereços IP de rede para uma única organização. Para entender por que a divisão em super-rede foi adotada, é preciso conhecer três fatos:

• O endereçamento IP não divide endereços de rede igualmente entre as classes, ou seja, na classe A podemos ter 126 redes diferentes, enquanto que na classe B podemos ter 16.382 redes e na classe C 2.097.150 redes.

• Os números de classe C foram solicitados aos poucos. A maioria dos endereços IP fornecidos foram de classe B.

• Nesta proporção, todos os números classe B seriam esgotados em alguns anos. O problema se tornou conhecido como Esgotamento do Espaço de Endereço (ROADS – Running Out of Address Space). Isto porque há duas razões para as organizações preferir números de rede da classe B: um endereço classe C não suporta mais de 254 hosts e um endereço classe B pode-se facilmente implementar a divisão em sub-redes.

Para conservar os números classe B, o esquema de super-rede atribui a uma organização um bloco de endereços classe C, em vez de um único número classe B.



O CIDR não roteia de acordo com a classe do número da rede (por isso o termo sem classe) mas apenas de acordo com os bits de ordem mais alta do endereço IP, chamados de prefixo IP. Cada entrada na tabela de roteamento CIDR contém um endereço IP de 32 bits e uma máscara de rede de 32 bits, que juntos fornecem o comprimento e o valor do prefixo IP.

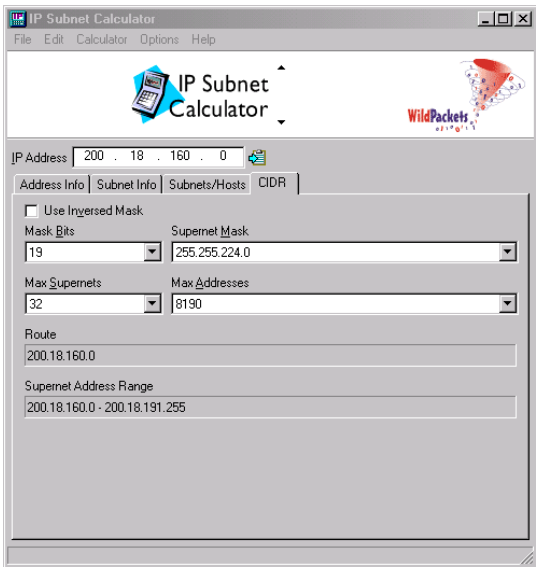
Por exemplo, para endereçar um bloco de 32 endereços classe C com uma única entrada na tabela de roteamento, a representação seguinte seria suficiente:

End. IP 200.18.160.0

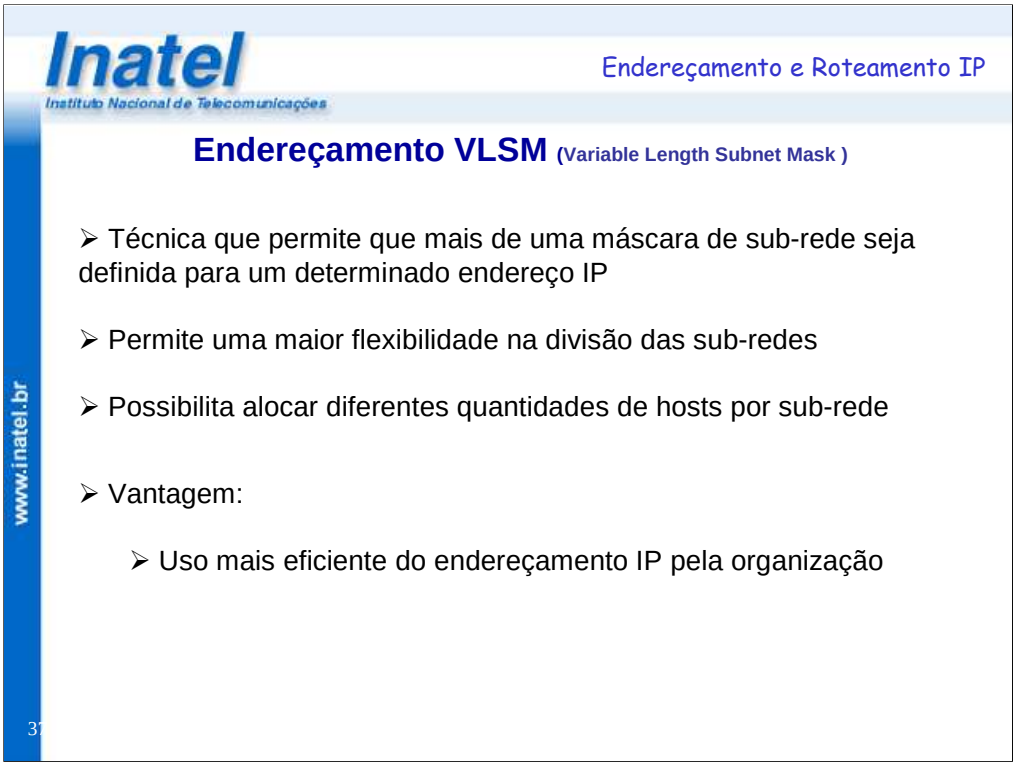
Máscara 255.255.224.0

Isto iria do ponto de vista do backbone, referir-se ao intervalo de rede classe C de 200.18.160.0 a 200.18.191.0 como uma única rede devido ao prefixo IP idêntico.

Uso de uma ferramenta de software para calcular CIDR

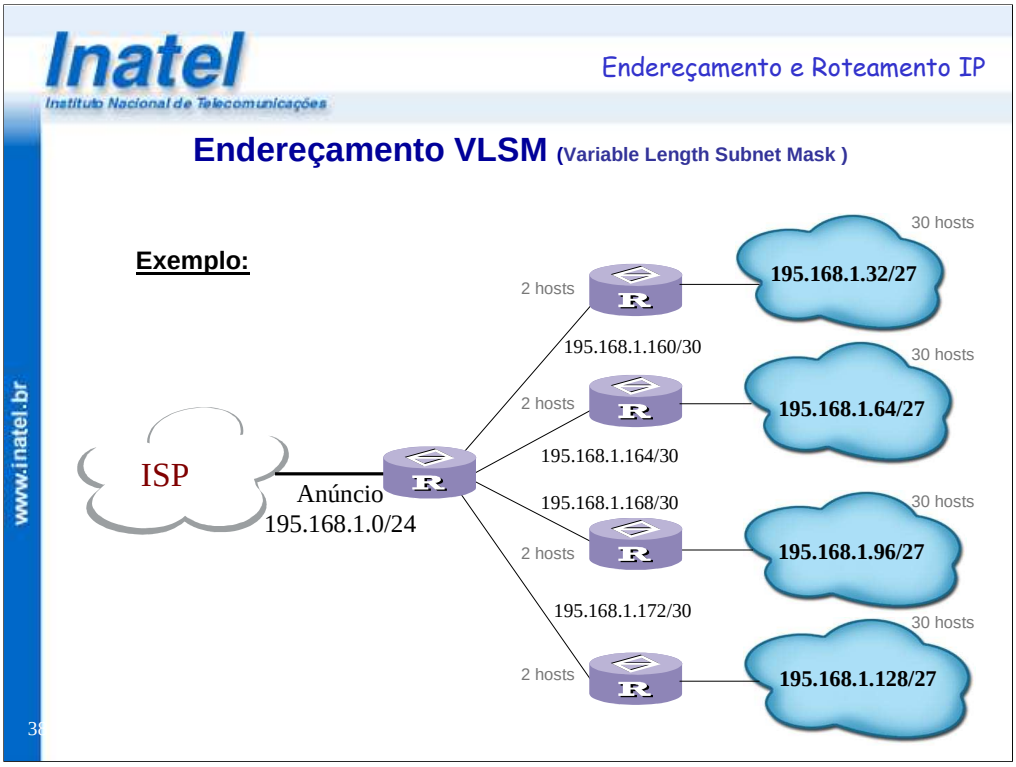


Um exemplo de calculo de super-rede (CIDR) utilizando a calculadora IP Subnet Calculator

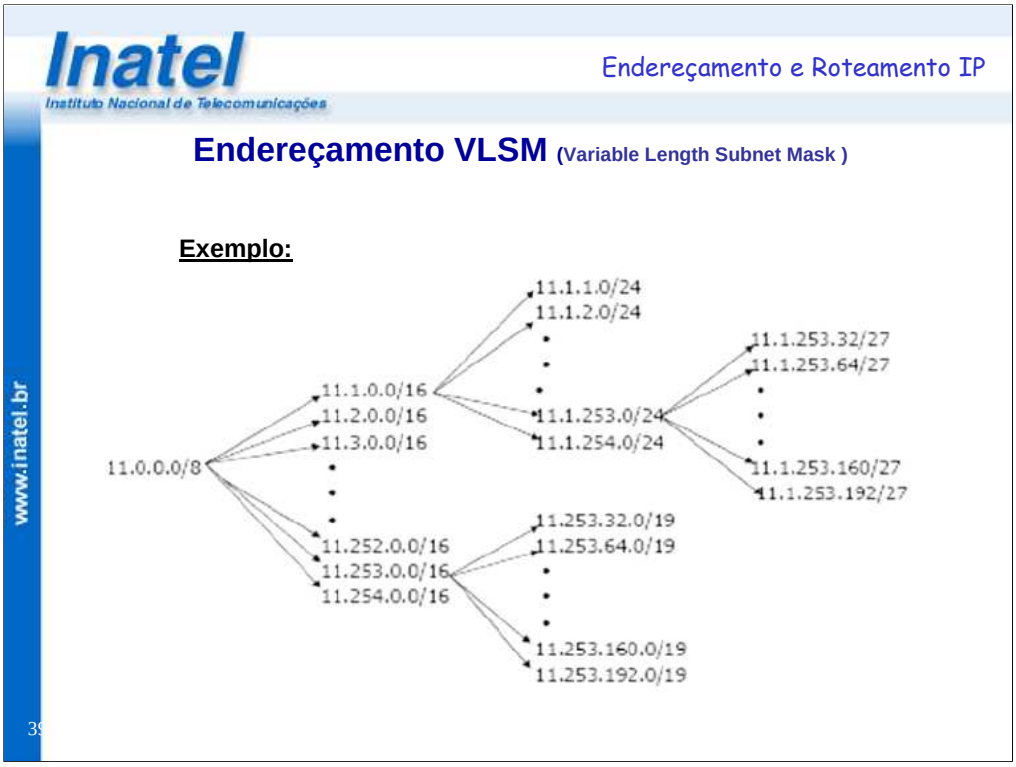


VLSM nada mais é do que a segmentação lógica de subredes. Ou seja, para criar subredes, você segmentou uma determinada rede. VLSM consiste em segmentar as subredes criadas, em blocos não necessariamente do mesmo tamanho. Daí o nome "subredes de tamanho variável".

Protocolos de roteamento modernos, tais como OSPF e I-I-IS, permitem a implementação do VLSM por prover o tamanho do prefixo de rede estendido ou o valor da máscara junto com cada informação de rota. Isto permite que cada sub-rede seja alertada com seu correspondente tamanho do prefixo ou máscara. Se os protocolos de roteamento não disporem das informações de prefixo, um roteador deveria assumir o prefixo que já está aplicado localmente, ou fazer uma busca em uma tabela de prefixos estática que contenha toda a informação de máscara necessária. A primeira alternativa não poderia garantir que o prefixo correto seria aplicado, enquanto que as tabelas estáticas estariam sujeitas a erros humanos e seriam difíceis de ser manter. Um protocolo que também ser utilizado seria o RIP-2, apresentando melhorias sobre o RIP-1 já que suporta informações de prefixo de rede estendidos.



Grandes redes podem obter centenas de milhares de subredes. Por esta razão, não é desejável mantê-las em uma tabela de roteamento. A sumarização de rotas, como descrito no **RFC 1518** permite o resumo de múltiplas pequenas subredes dentro de uma única grande rota.



O VLSM também permite uma divisão recursiva de um endereço de uma organização para que haja uma redução da quantidade de informação no nível mais alto. Conceitualmente, uma rede é divida em sub-redes, algumas dessas sub-redes são divididas em outras sub-redes, e algumas dessas sub-sub-redes são dividas em sub2-sub-redes. Isto permite que as informações detalhadas de roteamento da estrutura de uma sub-rede fiquem invisíveis para os roteadores em outros grupos de sub-redes.

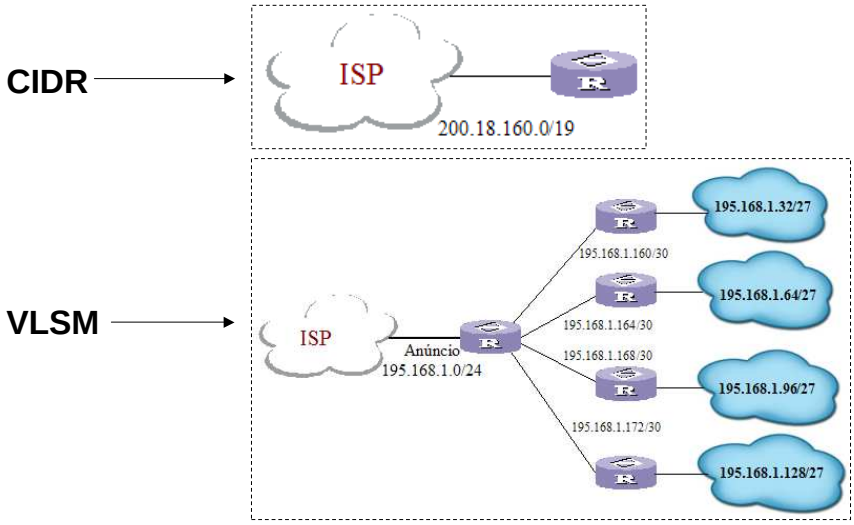
Para a implementação correta do VLSM, três pré-requisitos devem ser alcançados:

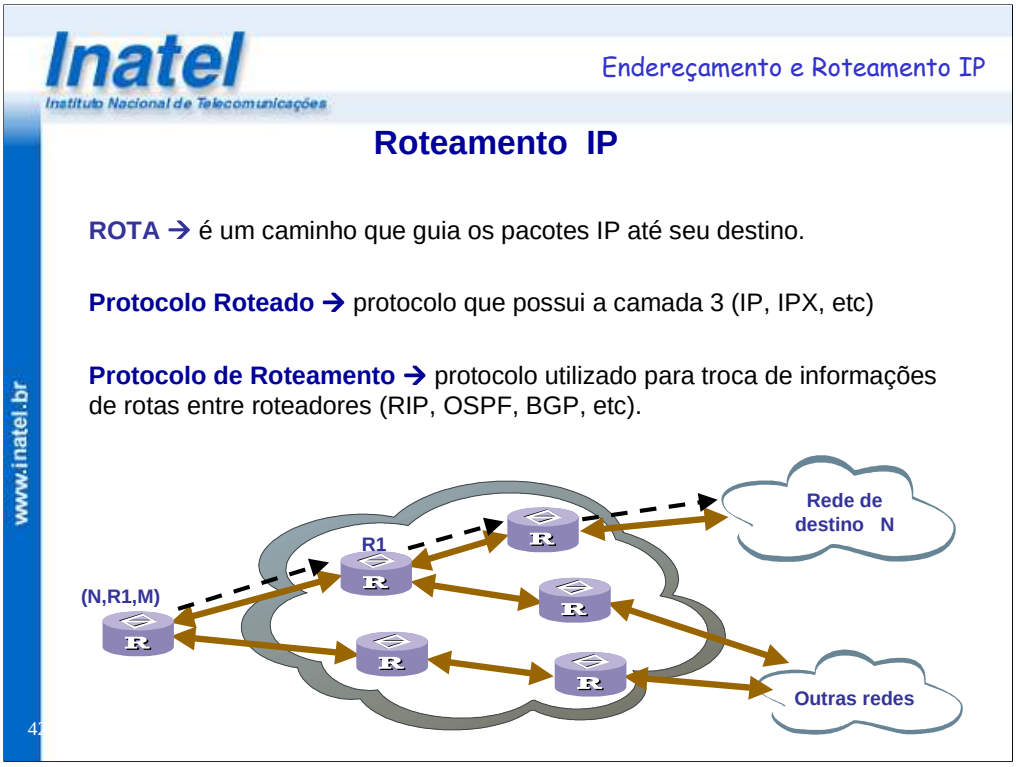
1- Os protocolos de roteamento devem suportar informações de prefixos de rede estendidos com uma fácil identificação.

2- Todos os roteadores devem implementar um algoritmo de direcionamento baseado na maior coincidência (longest match).

3- Para que ocorra um agrupamento de rotas, os endereços devem ser associados de forma topológica e lógica.

Diferença entre CIDR e VLSM



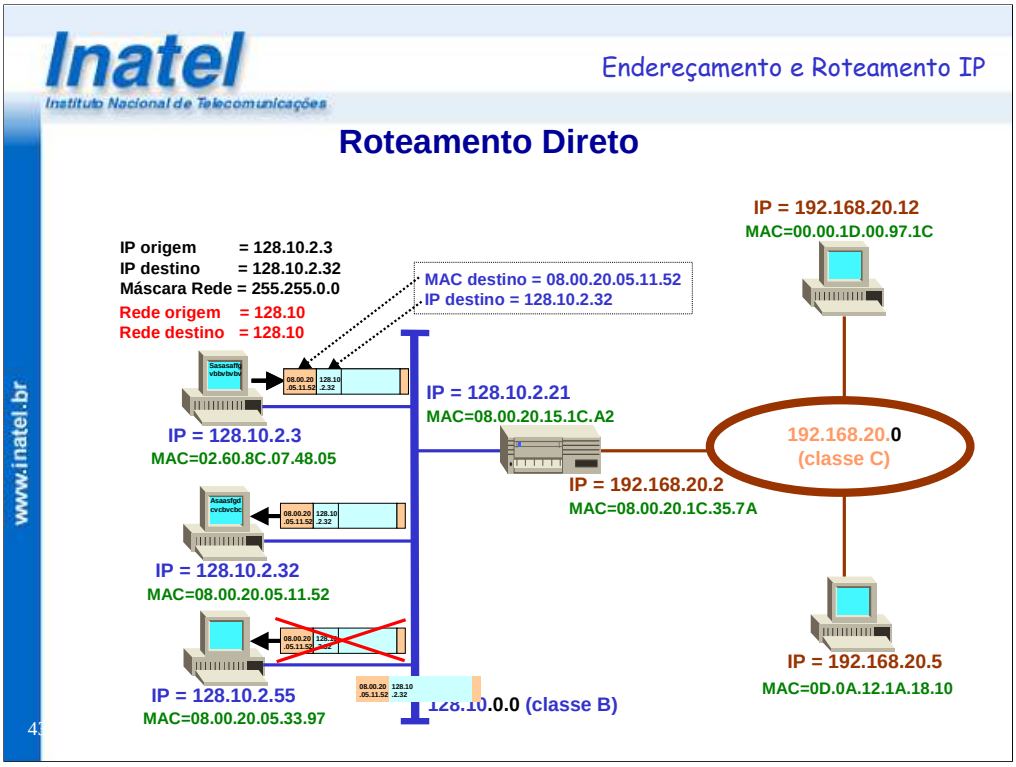


Em um sistema de comutação por pacotes, roteamento refere-se ao processo de selecionar um caminho pelo qual são enviados os pacotes; roteador se refere a um equipamento ou computador que executa tal seleção.

Por causa da similaridade dos dois termos, existe uma freqüente confusão entre protocolo *roteado* e protocolo de *roteamento*.

Protocolo roteado: qualquer protocolo de rede que fornece informaçõees suficientes no seu endereço de camada de rede para permitir que um pacote seja encaminhado de um host para outro, baseado no esquema de endereçamento.

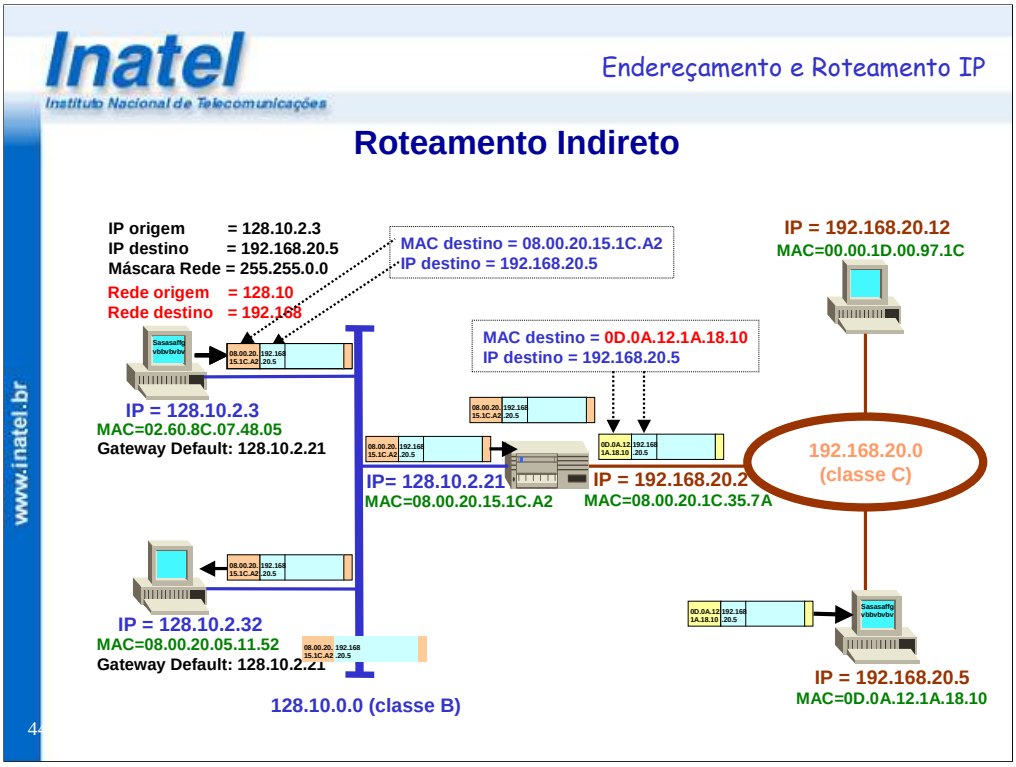
Protocolos de roteamento: suportam um protocolo roteado fornecendo mecanismos para compartilhar as informações de roteamento. As mensagens do protocolo de roteamento se movem entre os roteadores. Um protocolo de roteamento permite que os roteadores se comuniquem com os outros roteadores para atualizar e manter tabelas.



Podemos classificar o roteamento em dois tipos: direto e indireto. O encaminhamento direto é a transmissão de um datagrama, através de uma única rede física para outra máquina. Duas máquinas só podem executar o roteamento direto se ambas se conectarem diretamente a uma mesma rede física.

No roteamento direto, uma máquina de uma determinada rede física pode enviar um quadro físico diretamente a outra máquina conectada na mesma rede. Para transferir um datagrama IP, o transmissor encapsula o datagrama em um quadro físico, mapeia o endereço IP de destino em um endereço físico (MAC) e, para entrega-lo, usa o hardware de rede.

A transmissão de um datagrama IP entre duas máquinas de uma única rede física não envolve roteadores.



O roteamento indireto é mais complexo do que o roteamento direto, porque o transmissor deve identificar um roteador para o qual o datagrama possa ser enviado. O roteador deve então repassar o datagrama para a sua rede de destino. Quando uma máquina deseja enviar algo para a outra, ela encapsula o datagrama em um quadro físico e o envia ao roteador mais próximo. Quando um quadro acessa o roteador, o software extrai o datagrama encapsulado e o software IP seleciona o próximo roteador ao longo do caminho, em direção ao destino. O datagrama é novamente colocado em um quadro e enviado através da próxima rede física para o roteador seguinte, e assim por diante até que possa ser entregue diretamente.

Quando uma máquina enviar uma mensagem IP para outra rede, ela deve seguir os seguintes passos:

- Determinar se a máquina destino está em outra rede e por isto deve-se enviar a mensagem para um roteador

- Determinar, através da tabela de rotas da máquina origem, qual roteador é o correto para se enviar a mensagem

- Descobrir, através do protocolo ARP, qual o endereço MAC do roteador

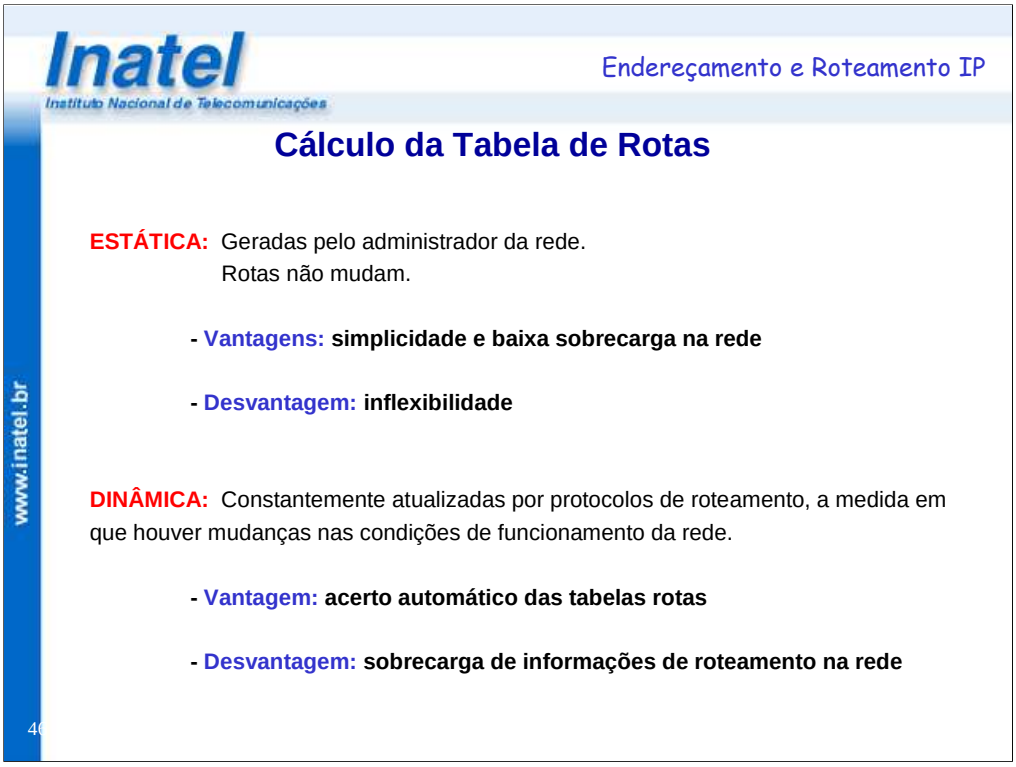
- Enviar a mensagem IP com o endereço de nível de rede apontado para o roteador e o endereço IP (na mensagem IP) endereçado para a máquina destino.

Existem 3 tipos de rotas:

As Rotas Diretas, que são encontradas pelo próprio equipamento, através do protocolo de enlace. Para este tipo de rota não é necessário nenhuma configuração, pois o próprio equipamento de camada 3 "descobre" estas rotas e acrescenta em sua tabela de rotas.

As Rotas Estáticas são configuradas manualmente no equipamento. É uma configuração simples, mas tem que ser alterada toda vez que acontecer alguma mudança com a rota configurada.

As Rotas Dinâmicas são configuradas automaticamente por protocolos chamados de protocolos de roteamento. Tem a vantagem de alterar automaticamente a tabela de rotas quando acontece alguma mudança, mas tem a desvantagem de usar o enlace da rede (banda) para transmissão de suas mensagens.



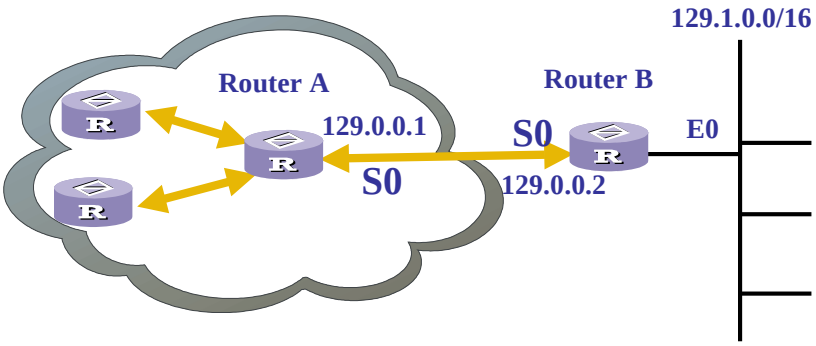
A alimentação das informações na tabela de rotas pode ser de modo estático ou dinâmico ou ambos simultaneamente.

Na alimentação estática, as rotas são preenchidas manualmente, geralmente pela configuração inicial da máquina.

Na alimentação dinâmica, protocolos como RIP, RIP2, OSPF ou BGP4 são responsáveis pela aquisição de informações sobre a topologia da rede e a publicação de rotas na tabela de rotas dos roteadores envolvidos.

Cada tipo de roteamento apresenta vantagens e desvantagens. No roteamento estático a simplicidade e a baixa sobrecarga na rede são as principais vantagens, entretanto como desvantagens podemos citar a inflexibilidade na mudança das rotas pré-estabelecidas. Já no roteamento dinâmico a principal vantagem é a resposta automática da tabela de rotas com relação às mudanças sofridas pela rede, mas o uso deste tipo de roteamento, em parte, sobrecarrega a rede com as informações de roteamento.

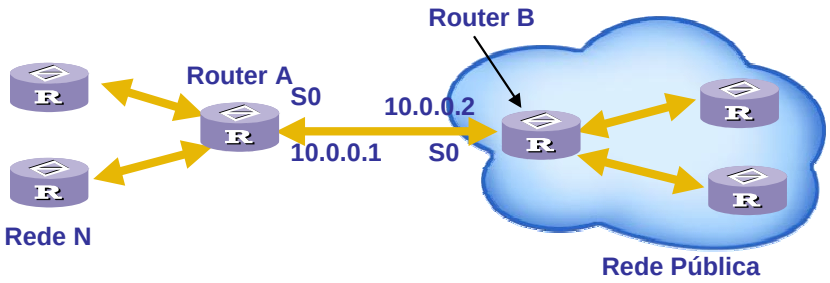
Rotas Estáticas



No Router A,

ip route-static 129.1.0.0 255.255.0.0 129.0.0.2
ou
ip route-static 129.1.0.0 16 129.0.0.2
ou
ip route-static 129.1.0.0 16 s0

Rotas Estáticas – Rota Default



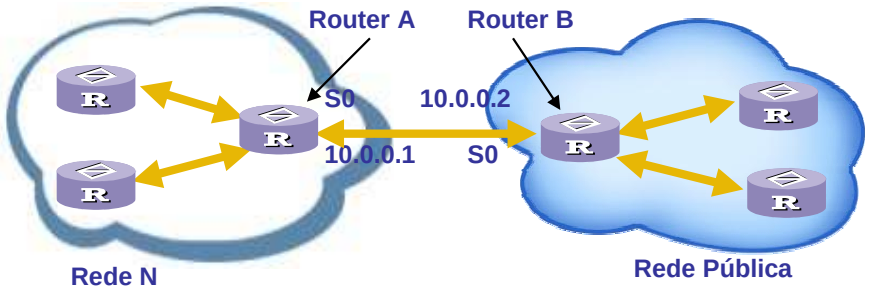
No roteador A,

ip route-static 0.0.0.0 0.0.0.0 10.0.0.2

Na internet, 99.99% dos roteadores configuram rotas default.

Não se trata apenas de configurar a rota estática manualmente. Algumas vezes a rota default pode ser gerada pelo protocolo de roteamento dinâmico.

Rotas Estáticas – Rota Default com "Self-Loop"



Configuração do roteador A:
ip route 20.0.0.0 8 10.0.0.2

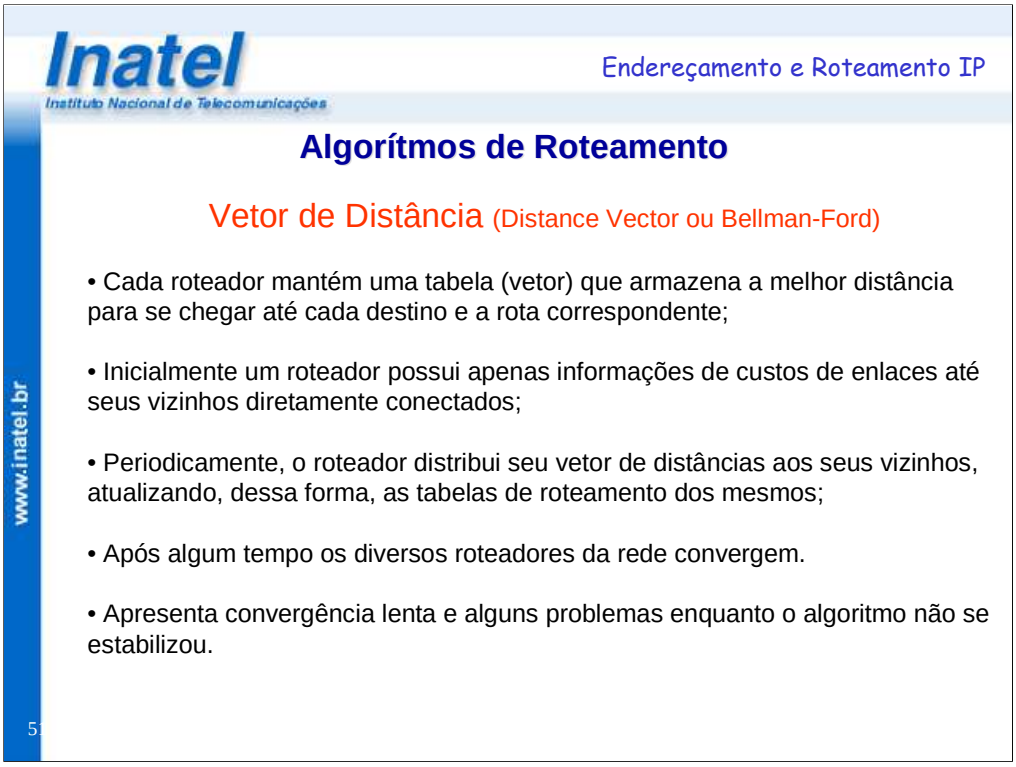
Configuração do roteador B:
ip route 20.0.0.0 8 10.0.0.1

A rota em "self-loop" é um problema para a rede e deve ser evitada.

Protocolos de Roteamento

- **Qual é o propósito dos protocolos de roteamento ?**
  - ➔ Cálculo automático de rotas.

- **Como ele faz isso?**
  - ➔Todos os roteadores enviam suas informações sobre rotas ao roteador vizinho, desta forma cada roteador da rede vai receber informações de roteamento.
  - ➔Baseado em seu algorítmo, o roteador calcula a rota final (valores next hop e métrica).



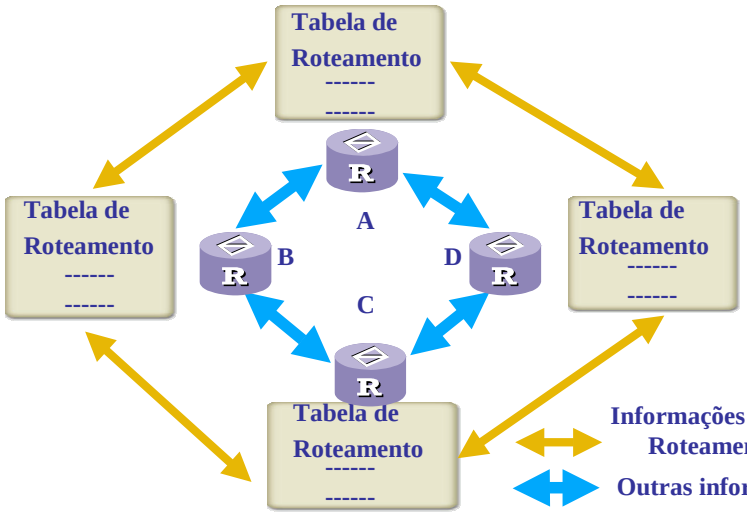
No roteamento com vetor de distância, cada roteador possui uma tabela com a melhor distância conhecida aos diversos destinos alcançáveis, bem como a saída a ser usada para chegar lá. Essa tabela é atualizada a partir de troca de informações com os vizinhos, que informam sobre seus vizinhos, e assim por diante.

Esse algoritmo também é conhecido pelo nome de seus desenvolvedores (**Bellman-Ford** (1957) e **Ford-Fulkerson** (1962)), e foi usado originalmente na Internet com o nome de RIP.

A métrica permitida pelo algoritmo pode ser número de hops (usado no RIP), retardo de tempo em mili-segundos, número de pacotes na fila ou algo semelhante. O slide ilustra um exemplo de funcionamento do algoritmo de roteamento vetor de distância.

Algorítmos de Roteamento

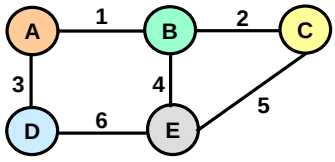
Vetor de Distância (Distance Vector ou Bellman-Ford)



Algorítmos de Roteamento

Vetor de Distância (Distance Vector ou Bellman-Ford)

Exemplo:



Letras → Roteadores
Números → Enlaces
Métrica → Distância (salto)

Os roteadores da rede tem suas tabelas de rotas inicial, a seguir:

A

Destino	Rota	Métrica
B	Direta	0
D	Direta	0

B

Destino	Rota	Métrica
A	Direta	0
C	Direta	0
E	Direta	0

C

Destino	Rota	Métrica
B	Direta	0
E	Direta	0

D

Destino	Rota	Métrica
A	Direta	0
E	Direta	0

E

Destino	Rota	Métrica
B	Direta	0
C	Direta	0
D	Direta	0

Inicialmente, cada *gateway (ou roteador)* possui uma tabela contendo uma entrada para cada rede à qual está conectado. A cada rede especificada na tabela está associada a distância entre a mesma e o *gateway* que mantém a tabela. Esta distância pode ser medida em *hops* (número de *gateways* a atravessar para atingir uma rede). Inicialmente, os campos de distância devem valer zero, pois somente as redes às quais o *gateway* está diretamente conectado são especificadas na tabela.

Periodicamente, cada *gateway* envia uma cópia de sua tabela para todo o *gateway* que possa atingir diretamente. O *gateway* que recebe a tabela, a compara com a sua própria e modifica esta última nos seguintes casos:

• se o *gateway* emissor conhecer um caminho maios curto para determinada rede, ou seja se a distância apresentada na tabela do emissor for menor do que a da tabela do receptor;

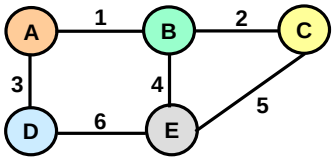
• se o *gateway* emissor apresentar uma rede que o receptor não conhece, ou seja, se na tabela do emissor existir uma entrada que não está presente na tabela do receptor; esta entrada é inserida na tabela do receptor;

• se uma rota que passa pelo emissor tiver sido modificada, ou seja, se a distância associada a uma rede que passa pelo emissor tiver mudado.

Algorítmos de Roteamento

Vetor de Distância (Distance Vector ou Bellman-Ford)

Exemplo - continuação:



Letras → Roteadores
Números → Enlaces
Métrica → Distância (salto)

Supondo que **A** envie primeiro sua tabela de rotas, os roteadores **B** e **D** atualizarão suas tabelas conforme o seguinte:

B

Destino	Rota	Métrica
A	Direta	0
C	Direta	0
E	Direta	0
D	*1*	*1*

D

Destino	Rota	Métrica
A	Direta	0
E	Direta	0
B	*3*	*1*

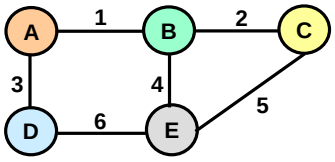
Agora **B** transmite sua tabela a seus vizinhos (**A**, **C** e **E**).
D faz o mesmo para **A** e **E**.
A ao receber a mensagem de **B** e **D**, atualiza sua tabela conforme abaixo:

Destino	Rota	Métrica
B	Direta	0
D	Direta	0
C	*1*	*1*
E	*1*	*1*

Algorítmos de Roteamento

Vetor de Distância (Distance Vector ou Bellman-Ford)

Exemplo - continuação:



Letras → Roteadores
Números → Enlaces
Métrica → Distância (salto)

A ao receber a mensagem de **B,** atualiza sua tabela conforme abaixo:

Destino	Rota	Métrica
B	Direta	0
D	Direta	0
C	*1*	*1*
E	*1*	*1*

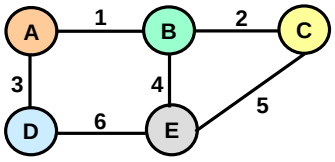
Quando um roteador recebe uma tabela de atualização de outro roteador, ele **verifica cada rota e mantém em sua tabela as rotas de menor métrica** com mesmo destino.

Assim, os roteadores vão trocando mensagens e se atualizando até as tabelas convergirem.

Algorítmos de Roteamento

Vetor de Distância (Distance Vector ou Bellman-Ford)

Exemplo - continuação:



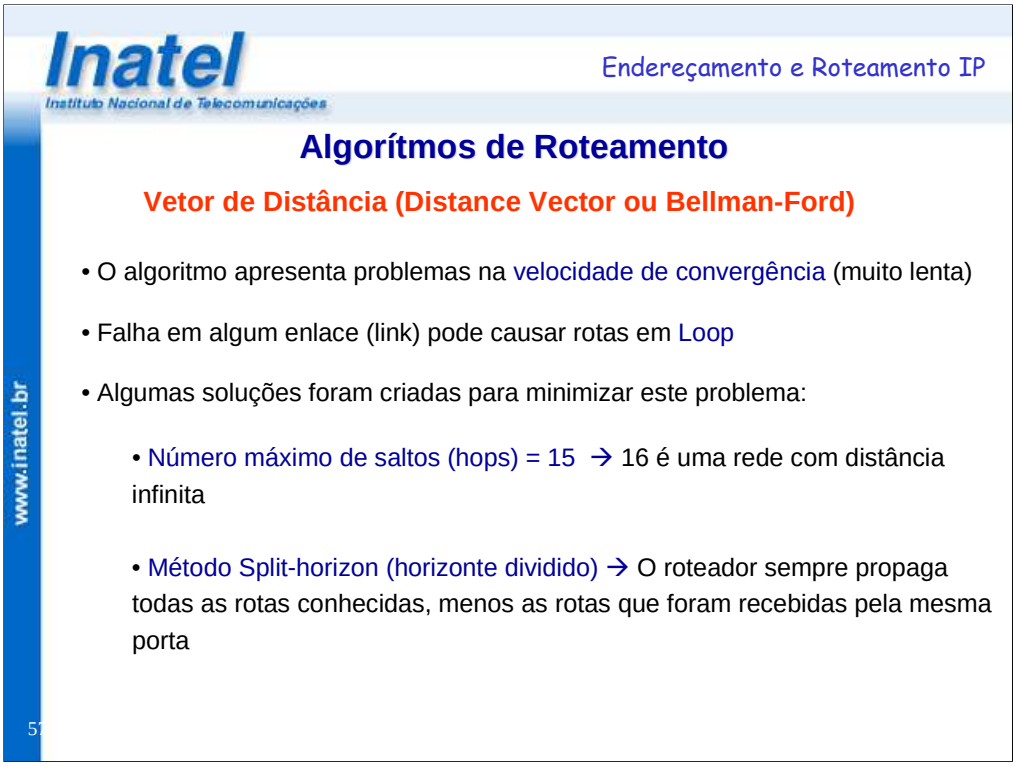
Letras → Roteadores
Números → Enlaces
Métrica → Distância (salto)

A tabela de rotas do roteador **A** depois da convergência será:

Destino	Rota	Métrica
B	Direta	0
D	Direta	0
C	*1*	*1*
E	*1*	*1*

A tabela de rotas do roteador **C** depois da convergência será:

Destino	Rota	Métrica
B	Direta	0
E	Direta	0
A	*2*	*1*
D	*5*	*1*



O algoritmo de vetor de distância tem um problema na velocidade de convergência quando ocorrem problemas na rede. Falhas na rede demoram a ser eliminadas, podendo ocasionar loops, que são pacotes que ficam sendo repassados de um roteador para outro durante certo tempo.

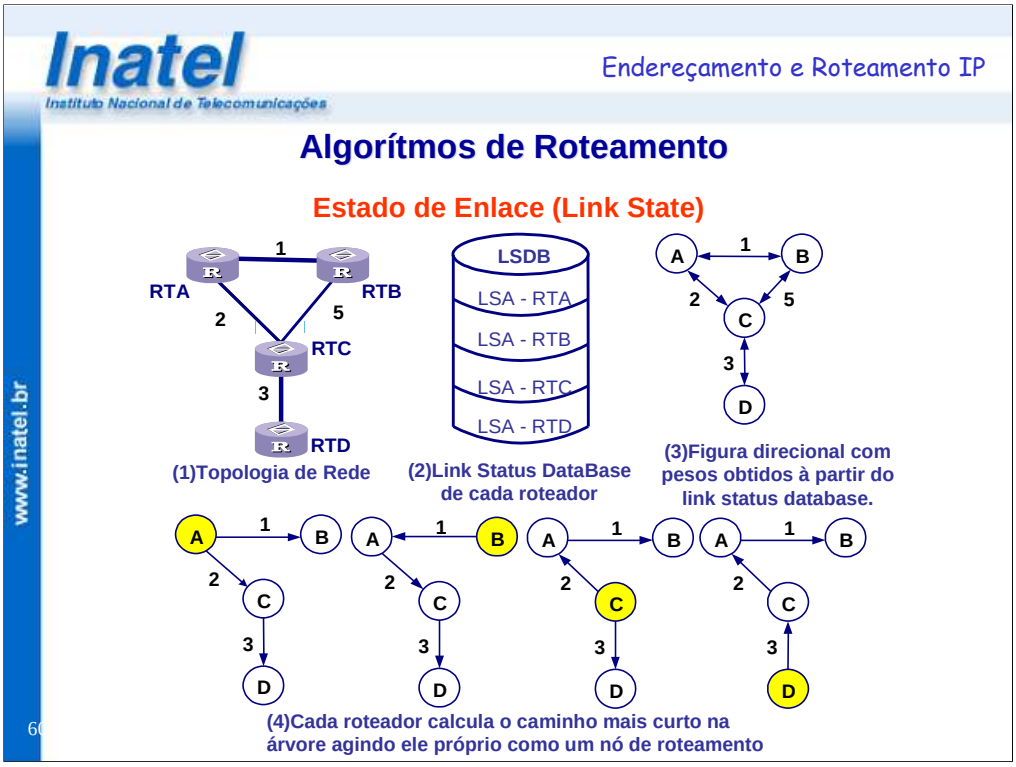
Soluções:

• Split-Horizon (dividir o horizonte): O roteador sempre propaga todas as rotas conhecidas, menos as rotas que foram recebidas pela mesma porta.

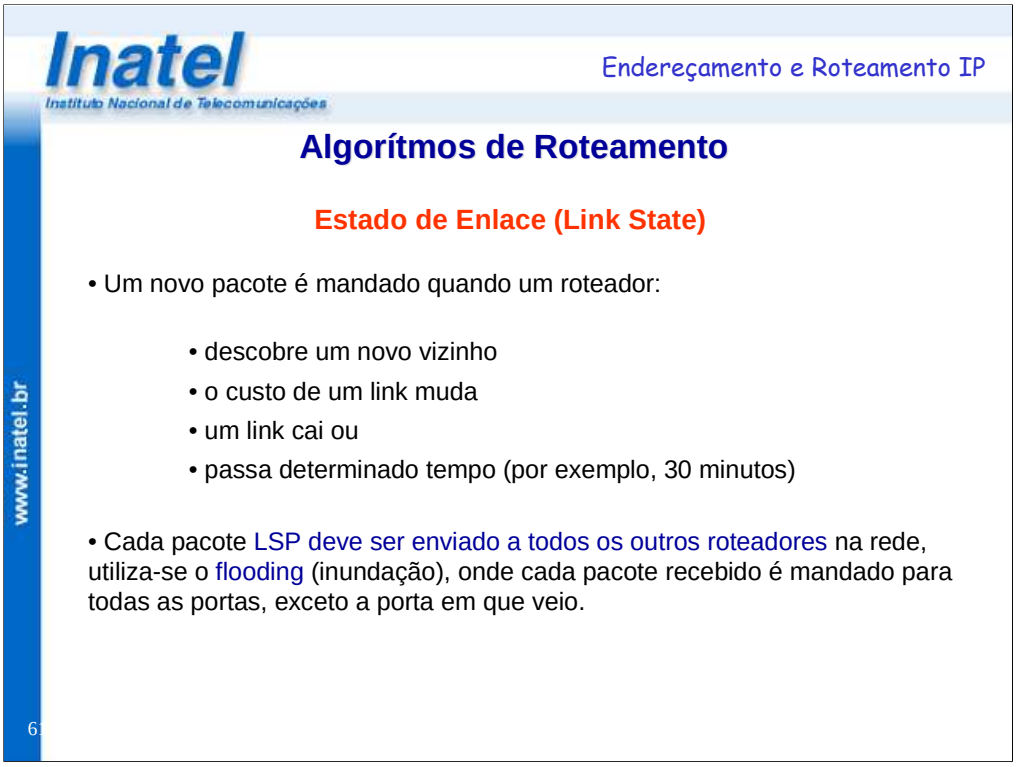
• *Hold Down Time* – Se um link "falhar", o roteador ignora todas atualizações para aquela rede por um tempo (por exemplo, 180s). Assim, todos roteadores esquecem o caminho da rede com problemas, e um link quebrado não será propagado. Este tempo pode ser um problema para os pacotes sendo transmitidos para aquela rede caso exista outra rota alternativa.

• *Poison Reverse and Triggered Updates* – Se um link "falhar", modifica a informação da rota para uma distância infinita (16) e propaga imediatamente essa informação adiante (não dá tempo de receber a atualização de outro nó). Assim, se houver alguma outra rota para a rede, essa rota será utilizada, pois será melhor que uma distância infinita.

O algoritmo de estado de enlace (link State) funciona de modo diferente do vetor-distância, ao invés de ter na tabela as melhores rotas, todos os nós possuem todos os links da rede. Cada rota contém o identificador de interface, o número do enlace e a distância ou métrica. Com essas informações os nós (roteadores) descobrem sozinhos a melhor rota.



O objetivo é que cada roteador envolvido tenha um banco de dados completo de toda a topologia da rede, para conseguir traçar o caminho mais curto através de um algoritmo, como o de Dijkstra, por exemplo. Este banco de dados deve ser o mesmo em todos os roteadores, a fim de que todos tomem as mesmas decisões.



Descobrir seus vizinhos e aprender seus endereços de rede: Para o roteador saber quem são seus vizinhos, **pacotes Hello** (multicast para 224.0.0.5 =

"todos roteadores") são enviados para as portas de tempos em tempos. Se um roteador recebe um pacote Hello ele responde com outro pacote contendo seu nome. Os nomes dos roteadores não podem ser duplicados. Os pacotes Hello também são utilizados para saber se um link está operacional

Medir o retardo para cada um dos vizinhos: A forma mais simples de medir o retardo é enviar pacotes de ECHO para o vizinho e esperar resposta. A média de vários tempos de resposta dividida por dois é uma estimativa do retardo. O tamanho da fila e a carga na rede também podem ser levados em consideração.

No algoritmo de estado de enlace, outras métricas podem ser medidas e levadas em consideração, como a velocidade das linhas, sua segurança, e assim por diante.

O pacote LSP (*Link State Packet*): Os roteadores utilizam o pacote LSP (*Link State Packet*) para trocar as informações de rotas com os outros roteadores.

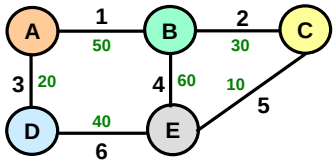
Enviar esse pacote a todos outros roteadores (*Flooding*): Um pacote é mandado quando um roteador **descobre um novo vizinho**, **o custo de um**

link muda, **um link cai** ou **passa determinado tempo** (30minutos por exemplo). Como cada LSP **deve ser enviado a todos os outros roteadores na rede**, utiliza-se flooding (inundação), onde cada pacote recebido é mandado para todas as portas, exceto a porta em que veio.

Algorítmos de Roteamento

Estado de Enlace (Link State)

Exemplo:



Letras → Roteadores
Números → Enlaces
Métrica → velocidade do enlace
(> bps .... < métrica)

Após a convergência do algoritmo, o banco de dados montado pelo roteador **A** será:

De	Para	Rota	Métrica
A	A	Direta	0
A	B	1	50
A	D	3	20
B	A	1	50
B	C	2	30
B	E	4	60
C	B	2	30
C	E	5	10
D	A	3	20
D	E	6	40
E	B	4	60
E	C	5	10
E	D	6	40

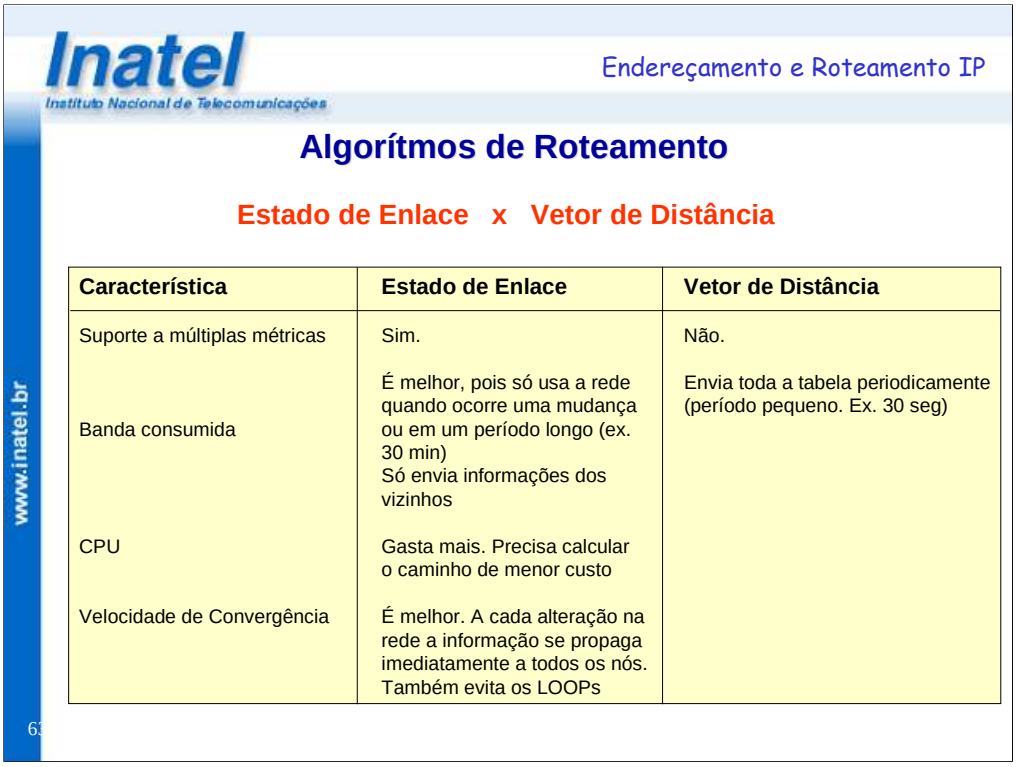
E a tabela de rotas do roteador **A** ficará assim:

Destino	Rota	Métrica
A	Direta	0
B	1	50
D	3	20
C	3	70
E	3	60

www.inatel.br

Calcular o caminho mais curto para cada um dos roteadores: Após as informações serem distribuídas por flooding, o algoritmo de Dijkstra pode ser

usado para encontrar o caminho mais curto para cada um dos outros roteadores. Cada roteador, tendo posse dos LSP´s de todos os outros roteadores de sua área, deve construir uma árvore lógica para chegar a qualquer outro roteador pelo caminho com menor custo.



Característica	Estado de Enlace	Vetor de Distância
Suporte a múltiplas métricas	Sim.	Não.
Banda consumida	É melhor, pois só usa a rede quando ocorre uma mudança ou em um período longo (ex. 30 min) Só envia informações dos vizinhos	Envia toda a tabela periodicamente (período pequeno. Ex. 30 seg)
CPU	Gasta mais. Precisa calcular o caminho de menor custo	
Velocidade de Convergência	É melhor. A cada alteração na rede a informação se propaga imediatamente a todos os nós. Também evita os LOOPs	

Estado de Enlace x Vetor de Distância

Suporte a múltiplas métricas – Como cada roteador tem consciência de toda a rede, basta a medição de outras métricas para que o algoritmo consiga decidir pelo menor custo dependendo da necessidade da aplicação.

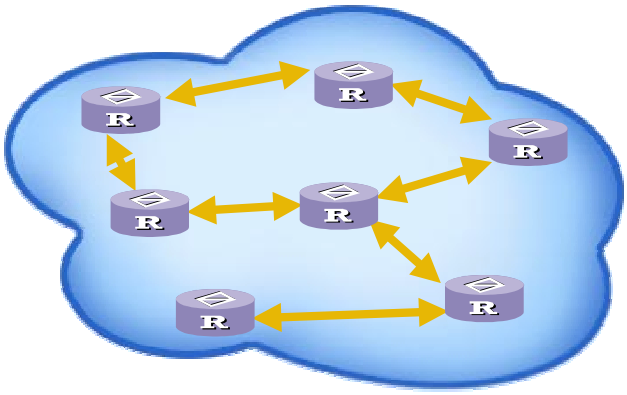
Banda Consumida – O algoritmo de Estado de Enlace é melhor, pois as atualizações ocorrem somente quando algo muda na rede, e não a cada 30 segundos, por exemplo. Além disso, o Estado de Enlace necessita enviar somente as informações dos vizinhos, e não toda sua tabela;

CPU – O Estado de Enlace gasta mais, pois deve calcular a árvore de menor custo;

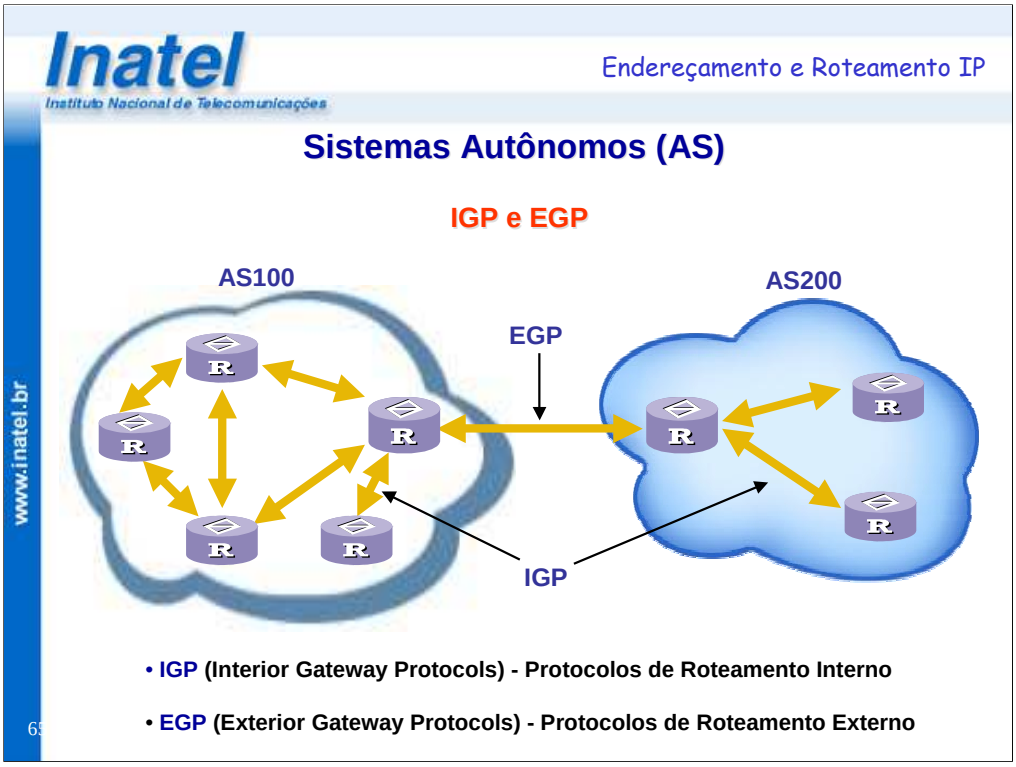
Velocidade de Convergência – O Estado de Enlace é melhor, pois a cada alteração na rede, essa informação se propaga imediatamente a todos os nós. Isso dá uma vantagem adicional, ou seja, **evita loops na rede**.

Sistemas Autônomos (AS)

Conjunto de roteadores que obedecem as mesmas estratégias de roteamento e é gerenciado por organizações unificadas.



Um AS é definido como uma porção lógica das redes IP maiores que são administradas por uma única autoridade. Normalmente, o AS abrangeria a inter-rede dentro de uma organização.



Os protocolos de roteamento dinâmicos podem ser divididos em dois grupos:

IGPs: Interior Gateway Protocols

EGPs: Exterior Gateways Protocols

Os protocolos IGPs permitem que os roteadores troquem informações de roteamento dentro de um Sistema Autônomo (AS). Já, os protocolos EGPs permitem a troca de informação de alcance entre AS administrados separadamente.

Protocolos de Roteamento

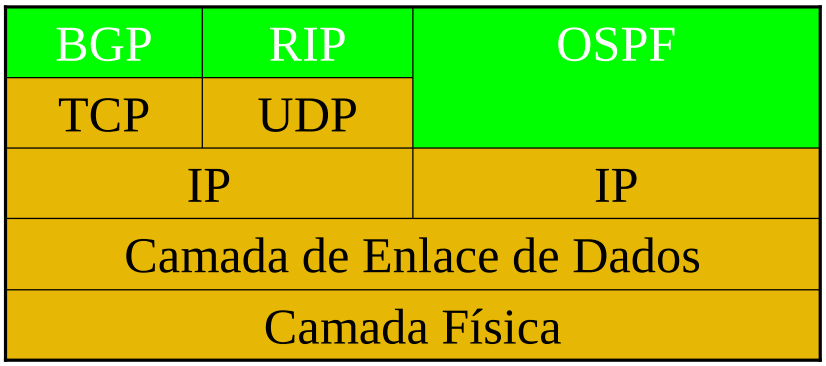
Protocolos IGPs:

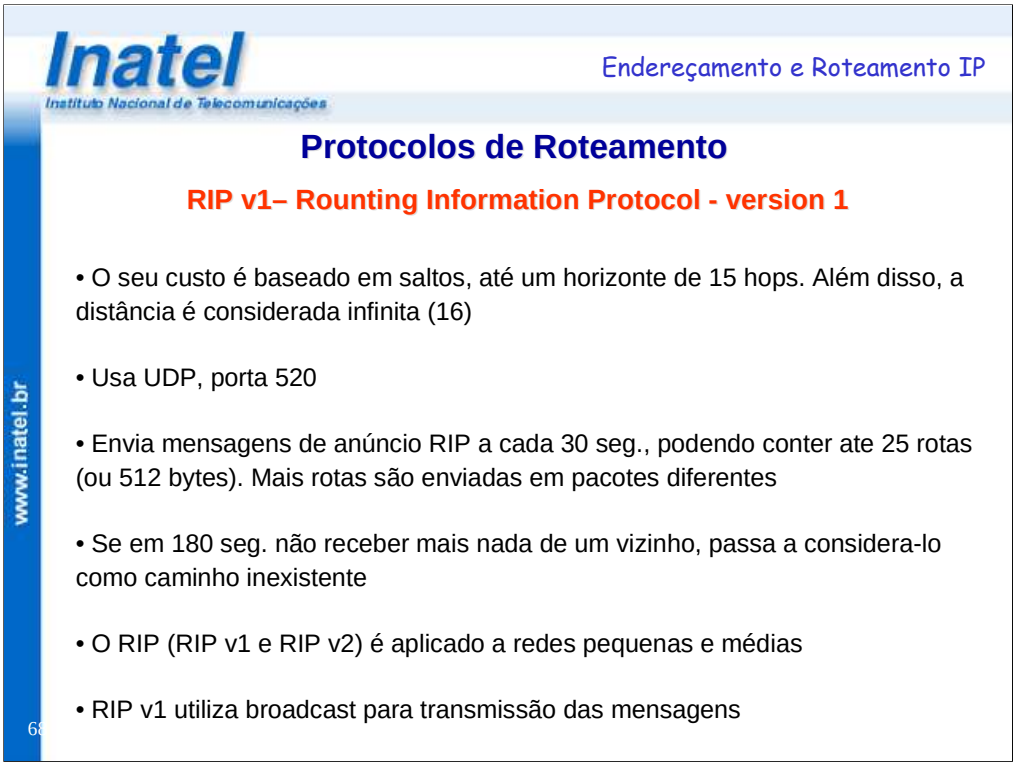
- RIP (Routing Information Protocol) - RFC 1058 — Vetor de Distância
- RIP-2 (RIP version 2) - RFC 1723
- OSPF (Open Shortest Path First) - RFC 2178 — Estado de Enlace

Protocolos EGPs:

- EGP (Exterior Gateway Protocol) - RFC 904 — Estado de Enlace
- BGP (Border Gateway Protocol) - RFC 1771 — Vetor de Distância

Protocolos de Roteamento





O RIP é um protocolo do tipo IGP destinado a redes de tamanho limitado. Foi criado inicialmente pela Xerox e baseado no algoritmo de vetor de distância visto anteriormente. Está definido na RFC 1058.

Suas características são:

Utiliza o número de hops como métrica;

Comunica-se com seus vizinhos a cada 30s. Se uma rota não é re-anunciada em 180s, ela é considerada como infinita e removida da tabela;

Transmite toda tabela de rotas para vizinhos sempre (e não só alterações da tabela), o que gera bastante tráfego;

Distância máxima de 15 hops. Acima disso o destino é considerado inalcançável;

Utiliza protocolo UDP (porta 520) com pacotes de 576 bytes, o que demanda múltiplos pacotes para enviar a tabela inteira. Por exemplo, uma tabela com 300 entradas necessita de 12 pacotes;

Utiliza sempre o caminho mais curto;

Possui uma convergência lenta, pois uma nova rota vai sendo enviada a cada 30 segundos de vizinho em vizinho, demorando a atingir um destino distante;

O RIP versão 2 está definido na RFC 1723, e possui compatibilidade com o RIP v1, porém, adiciona uma série de melhorias, como as descritas a seguir:

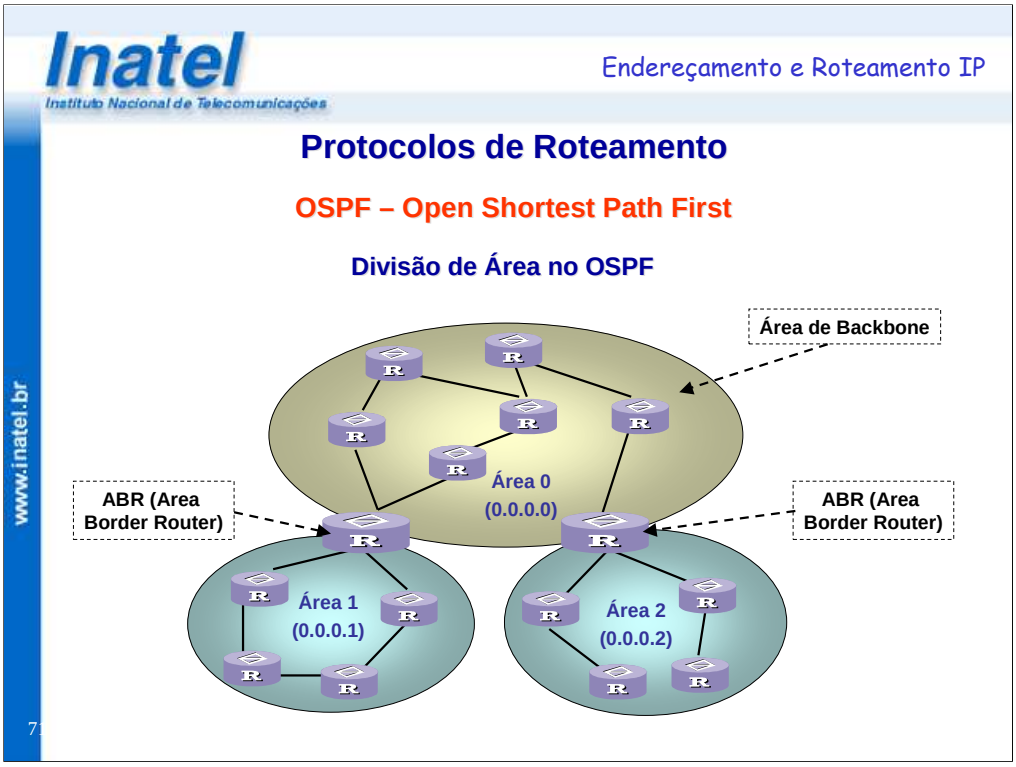
- Autenticação: só troca mensagens de roteamento com roteadores autorizados
- Máscara de subrede: útil para *classless*;
- Multicast: RIPv2 usa o IP multicast 224.0.0.9. O RIP v1 utiliza broadcast.

Suporta VLSM

Aprende rotas externas, vindas de outros sistemas autônomos

O OSPF utiliza como base o algoritmo de estado de enlace, está definido na RFC 2178, e possui as seguintes características:

- Usa outras métricas além da contagem de hops;
- Permite máscara de subrede;
- Autentica troca de rotas;
- Aprende de rotas externas (vindas de outros sistemas autônomos);
- Possui rápida convergência (que é um dos problemas do RIP);
- A informação é enviada novamente (*flooding*) toda vez que um roteador descobre um novo vizinho, o custo de um link muda, um link cai ou passa determinado tempo (30 minutos no caso do OSPF).



Muitos Sistemas Autônomos na Internet são grandes e difíceis de gerenciar. O OSPF permite que eles sejam divididos em áreas, a fim de diminuir o tráfego de roteamento.

Grandes áreas são divididas em áreas menores, e cada área tem uma identificação de 32 bits. A área de backbone é a área 0 (0.0.0.0), e as outras podem ser “0.0.0.1”, “0.0.0.2”, e assim por diante. As áreas se comunicam entre si via backbone, ou seja, todas as áreas devem estar ligadas ao backbone central, conforme mostra a figura.

Os roteadores de borda de área (ABR – *Area Border Routers*) sumarizam a topologia e a transmitem para a área de backbone que, por sua vez, retransmitem essa informação para as outras áreas. Durante a operação normal, podem ser necessários três tipos de rotas: na mesma área (usam algoritmo completo entre eles), entre áreas (usam o ABR para ir e voltar ao backbone) e entre sistemas autônomos (necessitam de protocolo de roteamento externo). Para suportar isso, o OSPF define quatro classes de roteadores:

Roteadores internos: ficam inteiramente dentro de uma área;

Roteadores de borda de área (ABRs): conectam uma área com o backbone;

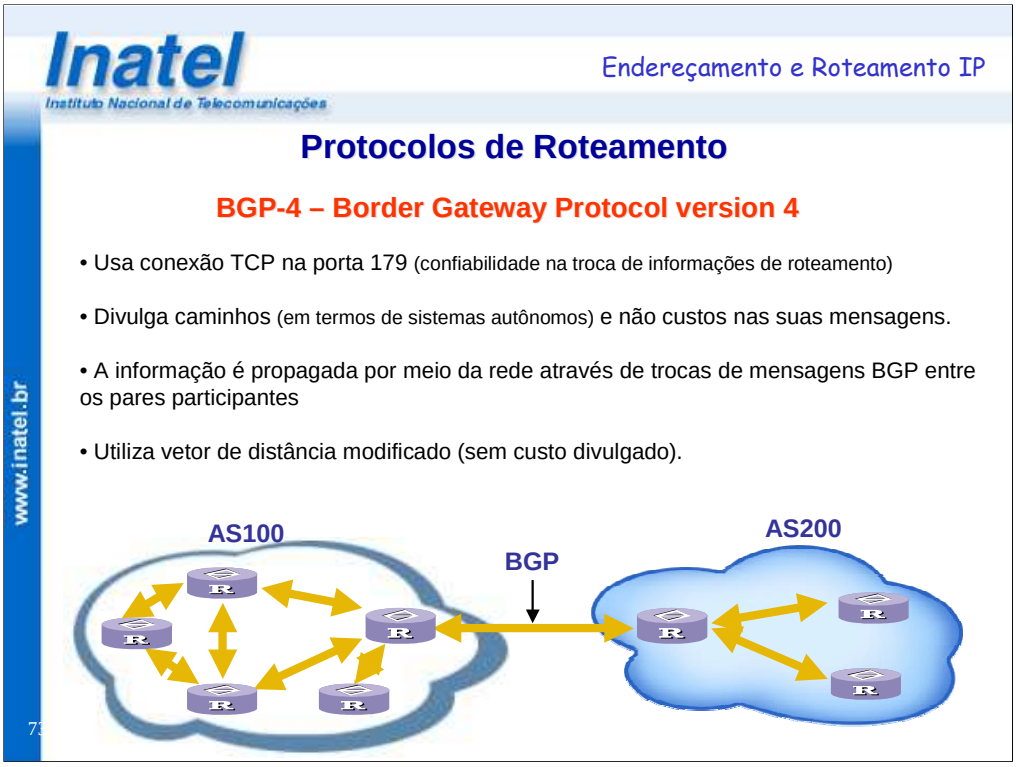
Roteadores de backbone: redirecionam pacotes dentro do backbone;

Roteadores de fronteira do AS: interagem com roteadores de outros AS.

Protocolos de Roteamento

RIP v1 x RIP v2 x OSPF

	RIP-1	RIP-2	OSPF
RFC	RFC 1058	RFC 1723	RFC 2178
Algoritmo	Vetor de distância	Vetor de distância	Estado de enlace
Suporte a múltiplas métricas	Não (hops)	Não (hops)	Sim
Suporte a CIDR	Não	Sim	Sim
Suporte a múltiplos caminhos	Não (o mais curto)	Não (o mais curto)	Sim
Máximo diâmetro da rede	15 Hops	15 Hops	65.535
Intervalo de Atualização	30 segundos	30 segundos	A cada mudança ou 30min
Atualizações	Toda a tabela de rotas	Toda a tabela de rotas	Somente vizinhos (LSP)
Intervalo para eliminação	180s	180s	
Autenticação	Não	Sim (senha aberta)	Sim (senha MD5)
Velocidade de Convergência	Devagar	Devagar	Rápido
Suporta AS#	Não	Sim	Sim



O BGP é um protocolo entre sistemas autônomos, definido na RFC 1761 e usado desde 1989, entretanto, seu uso cresceu mesmo nos últimos anos. Sua função principal é trocar informações de acessibilidade com outros sistemas BGP com o objetivo de traçar um grafo da conectividade do AS, visando eliminar problemas de laços e reforçar decisões de políticas do sistema (como, por exemplo, a política de anunciar para os AS vizinhos somente as rotas que ele usa, refletindo o paradigma "hop-by-hop" usado na Internet).

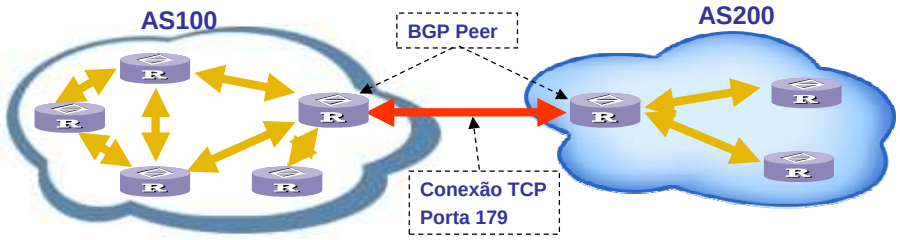
A versão 4 do BGP (RFC 1771, 1772 e 1773) possui um novo conjunto de mecanismos para suportar CIDR (*Classless Interdomain Routing*), como suporte para anunciar um prefixo IP (*extended network prefix* – ex. /24) e eliminar o conceito de "classes". Além disso, inclui mecanismos para agregação de rotas.

O BGP é fundamentalmente um protocolo de vetor de distância, porém, ao contrário do RIP, cada roteador mantém o controle total de cada caminho (da origem ao destino) utilizado por cada roteador envolvido na rede.

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

- **BGP *Peers***
  - Inicialmente um roteador BGP deve reconhecer e autenticar o seu *peer*
    - Os dois *peers* estabelecem uma conexão TCP
  - Cada *peer* envia uma informações de alcançabilidade positiva ou negativa
    - Divulgação das rotas ativas e inativas de cada um
  - Troca de mensagens contínua para confirmação das rotas e também da conexão entre os roteadores



Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

Mensagens:

- **Open** → utilizadas para o estabelecimento de uma conexão BGP;
- **Update** → utilizadas para os anúncios propriamente ditos, incluindo rotas que devem ser incluídas na tabela e também rotas que devem ser removidos da tabela BGP.
- **Notification** → reportam erros e serve para representar possíveis problemas nas conexões BGP.
- **Keepalive** → são utilizadas para manter a conexão entre roteadores BGP caso não existam atualizações através de mensagens UPDATE.

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

BGP table version is 1660291, local router ID is 200.10.20.30
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	AS-Path	Origin
*>i12.0.48.0/20	198.32.252.254		100	0	11537 10578 1742	**i**
*>i12.6.208.0/20	198.32.252.254		100	0	11537 10578 1742	**i**
*>i12.6.252.0/24	198.32.252.254		100	0	11537 10578 14325	**?**
*>i12.16.126.192/26	198.32.252.254		100	0	11537 10578 14325	**?**
*>i12.144.59.0/24	198.32.252.254		100	0	11537 10466 13778	**i**

- ***** → mostra que estes estão definidos como melhores caminhos para as redes em questão.
- **Network** → a rede anunciada na forma de bloco CIDR
- ***Next Hop*** → como próximo roteador que os pacotes para esta rede deverão ser enviados.
- ***Local Preference*** → com valor 100
- **AS_Path** → mostra a seqüência de sistemas autônomos até a chegada a rede destino
- **Origin** → define a procedência do anúncio pelo AS, i (IGP), e (EGP) ? (indefinido ou incompleto).

1) Se o *next hop* não for alcançável, a rota é ignorada

2) Será preferida a rota que tiver maior valor de *Weight*, que se trata de um parâmetro proprietário da Cisco, utilizado localmente em um roteador. Caso o equipamento não seja Cisco, este passo do algoritmo não será efetuado

3) Caso o parâmetro anterior seja o mesmo, será preferida a rota que tiver o maior valor de *Local Preference* (LOCAL_PREF)

4) Caso o valor de *Local Preference* seja o mesmo, será preferida a rota com menor AS_PATH

5) Caso o AS_PATH tenha o mesmo tamanho, será preferida a rota com menor tipo ORIGIN, ou seja, serão priorizados os anúncios tipo IGP (i), seguido pelos EGP (e) e INCOMPLETE (?)

6) Caso o tipo ORIGIN seja o mesmo, será preferida a rota o atributo MED mais baixo caso as rotas tenham sido aprendidas a partir do mesmo AS.

7) Caso as rotas tenham o mesmo valor de MED, será preferida a rota por eBGP a iBGP.

8) Se o valor de MED for o mesmo, será preferido o anúncio vindo do roteador conectado via IGP mais próximo deste.

9) Se o caminho interno for o mesmo, o atributo BGP ROUTER_ID será o responsável pela decisão (*tiebreaker*). Neste caso, será preferido o caminho cujo roteador possuir o menor ROUTER_ID, que nas implementações Cisco é definido como IP da interface *loopback* se esta estiver configurada. No caso do roteador não possuir interface loopback configurada, será escolhido o IP mais alto do roteador. Vale lembrar que para cada fabricante o ROUTER_ID pode ser baseado em outras informações.

1) Se o *next hop* não for alcançável, a rota é ignorada

2) Será preferida a rota que tiver maior valor de *Weight*, que se trata de um parâmetro proprietário da Cisco, utilizado localmente em um roteador. Caso o equipamento não seja Cisco, este passo do algoritmo não será efetuado

3) Caso o parâmetro anterior seja o mesmo, será preferida a rota que tiver o maior valor de *Local Preference* (LOCAL_PREF)

4) Caso o valor de *Local Preference* seja o mesmo, será preferida a rota com menor AS_PATH

5) Caso o AS_PATH tenha o mesmo tamanho, será preferida a rota com menor tipo ORIGIN, ou seja, serão priorizados os anúncios tipo IGP (i), seguido pelos EGP (e) e INCOMPLETE (?)

6) Caso o tipo ORIGIN seja o mesmo, será preferida a rota o atributo MED mais baixo caso as rotas tenham sido aprendidas a partir do mesmo AS.

7) Caso as rotas tenham o mesmo valor de MED, será preferida a rota por eBGP a iBGP.

8) Se o valor de MED for o mesmo, será preferido o anúncio vindo do roteador conectado via IGP mais próximo deste.

9) Se o caminho interno for o mesmo, o atributo BGP ROUTER_ID será o responsável pela decisão (*tiebreaker*). Neste caso, será preferido o caminho cujo roteador possuir o menor ROUTER_ID, que nas implementações Cisco é definido como IP da interface *loopback* se esta estiver configurada. No caso do roteador não possuir interface loopback configurada, será escolhido o IP mais alto do roteador. Vale lembrar que para cada fabricante o ROUTER_ID pode ser baseado em outras informações.

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

Organização da Internet em PTT

- Organização da Internet
  - Diversos *backbones* pertencentes a empresas que inicialmente forneciam serviços de telefonia.
  - a fim de garantir a ligação entre todos os pontos, independente de qual *backbone* estivessem conectados, foi necessário criar pontos de conexão entre esses *backbones* para não isolá-los dos demais

- *Network Access Point* (NAP) ou Ponto de Troca de Tráfego (PTT)
  - melhorar principalmente o tempo de resposta e diminuir os gastos
  - diferentes *backbones* estabelecem conexões locais com o intuito de trocar tráfego e conseguir fornecer um serviço de maior qualidade, mais econômico e com tempo de acesso mais baixo

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

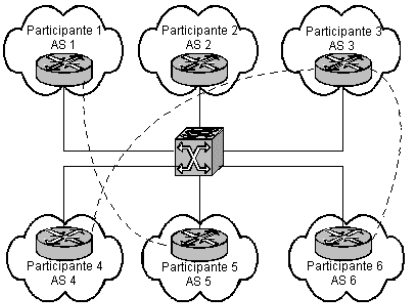
PTT

- Formado por um conjunto de roteadores localizados em um único ponto neutro
  - Formam uma rede local de alta velocidade, geralmente com tecnologias de nível 2 como *Fast Ethernet* ou *Gigabit Ethernet*
  - Cada roteador representa um sistema autônomo que deseja trocar tráfego com pelo menos um dos participantes.
  - Para a conexão entre todos os roteadores, existe um comutador ou *switch* com alto poder de processamento

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

Arquiteturas de PTT



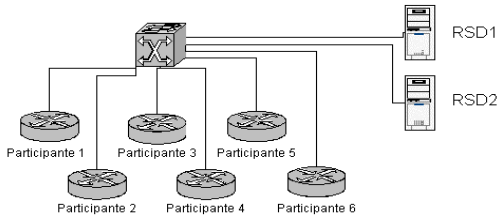
Neste tipo de PTT todos os participantes devem simplesmente estar na rede do PTT.

Os acordos de troca de tráfego neste tipo de PTT são chamados bilaterais, ou seja, se um determinado participante desejar trocar tráfego com um conjunto de ASs, este deverá estabelecer uma conexão BGP com cada um destes participantes diretamente. Neste caso, na entrada de um novo participante no PTT, se este desejar trocar tráfego com N participantes, N novas sessões BGP serão estabelecidas adicionalmente no PTT.

Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

Arquiteturas de PTT



Neste caso, além dos roteadores de cada participante, são utilizadas estações de trabalho com *software* que implementam o protocolo BGP e suas operações.

Sendo assim, os dois computadores trabalham como centralizadores das rotas anunciadas pelos participantes do PTT. Todos os participantes não mais estabelecem sessões entre eles, mas estabelecem sessões BGP com os dois *Route Servers*. Neste caso, são estabelecidos acordos de tráfego multilaterais.

Protocolos de Roteamento

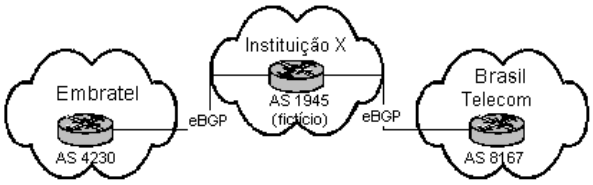
BGP-4 – Border Gateway Protocol version 4

Utilização de Políticas

As políticas aplicadas são relacionadas ao ato de troca de tráfego que tem relação direta com seus anúncios.

• **Trânsito** → AS anuncia-se como caminho não somente para suas redes, mas para todas as demais que ele conhece. Outros que não desejam fornecer trânsito apenas anunciam suas próprias redes.

• **Peering** → AS anuncia suas rotas recebidas a apenas um conjunto restrito de ASs. É feita geralmente mediante acordos entre ASs, como é o caso dos NAPs ou PTTs ou em conexões particulares entre ASs.



Protocolos de Roteamento

BGP-4 – Border Gateway Protocol version 4

Utilização de Políticas

• **Route Dampening** → uma espécie de "punição" que determinado AS pode levar caso seus anúncios sofram instabilidades na tabela de roteamento constantes. Isso faz com que determinado AS não seja "ouvido" pelos demais ASs por um tempo determinado, mantendo a estabilidade até que aquele anúncio estabilize. Isso evita que isso seja propagado por toda a Internet, consumindo banda e CPU de milhares de roteadores na inclusão/exclusão em suas tabelas de rotas BGP.

• Alguns backbones implementam tal funcionalidade, estabelecendo seus tempos de punição.

• Outros recursos de políticas podem ser aplicados de acordo com a necessidade de administrador do AS, podendo filtrar tipos determinados de anúncios, baseado em algum parâmetro do protocolo BGP, aceitando ou filtrando tais anúncios.

• A utilização de políticas em um sistema autônomo é uma das tarefas mais importantes de um administrador, visto que sua configuração pode refletir em uma melhora no acesso a outras redes

Indice de Desempenho dos Protocolos de Roteamento

- **Precisão**
  A melhor rota deve ser encontrada, e não deve existir “self-loop”.
- **Convergência Rápida**
  Quando há modificações na estrutura da topologia de rede, a rota será modificada de acordo no Sistema Autônomo.
- **Baixo overhead**
  O overhead do protocolo (custo, cpu, bandwidth da rede) é o mínimo.
- **Segurança**
  Com os mecânismos de segurança, o protocolo torna-se menos vulnerável a ataques.
- **Aplicação Universal**
  Aplicação geral em redes de diferentes topologias e tamanhos.

Indice de Desempenho dos Protocolos de Roteamento