Camada de Rede

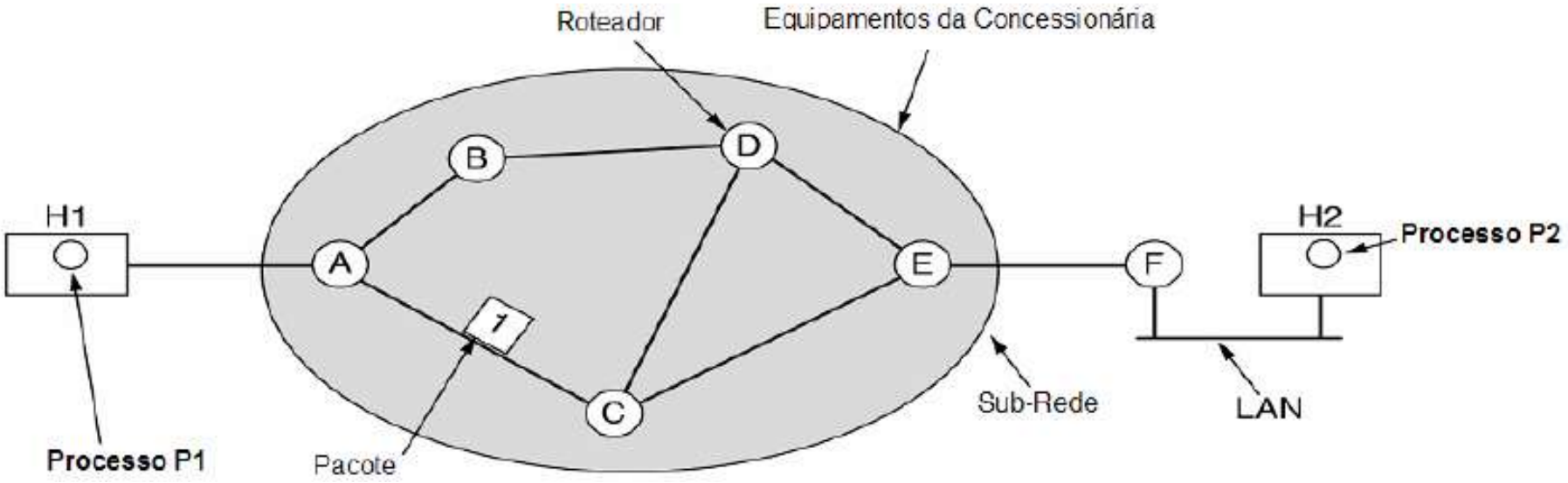
- A tarefa desta camada basicamente é transferência de pacotes da origem para o destino
- Funcionalidades:
  - Encontrar a **rota que os pacotes devem seguir, pois em** muitos casos, a transmissão fim a fim necessita vários saltos (hops).
  - Evitar a sobrecarga (congestionamento) de certas rotas em relação a outras
  - Compatibilizar os pacotes caso origem e o destino estejam em redes diferentes
- É a camada mais baixa que lida com a transmissão fim a fim

Camada de Rede - Conceitos

- **Roteador**
  - Computador de finalidade especial dedicado a interconexão de redes
  - Elemento de comutação que, ao receber um pacote deve escolher uma linha de saída para encaminhá-lo.
  - Hardware utilizado para compatibilizar duas redes distintas, que podem usar diferentes tecnologias como meio físico, endereçamento, formato de pacotes, etc.
  - Em linhas gerais cada roteador possui internamente tabelas que contém informações utilizadas para o roteamento dos pacotes
- **Sub-Rede**
  - O conjunto de linhas de comunicação e roteadores, excetuando-se os hosts

Camada de Rede – Comunicação

- **Comunicação Store-and-Forward**
  - Host(H1) com um pacote a enviar o transmite para o roteador mais próximo (A)
  - O pacote é armazenado até chegar totalmente
  - Após ser verificado (checksum) pacote é encaminhado para o próximo roteador (B) ao longo do caminho
  - Processo se repete até que o pacote chegue até o destino (H2)



Camada de Rede - Serviços

- Serviços oferecidos pela camada de rede à camada de transporte – Premissas:
  - Os serviços devem ser independentes da tecnologia de roteadores.
  - A camada de transporte deve ser isolada do número,do tipo e da topologia dos roteadores presentes.
  - Os endereços de rede que se tornaram disponíveis para a camada de transporte devem usar um plano de numeração uniforme, mesmo nas LANs e WANs.
  - Caso a mensagem seja mais longa que o tamanho máximo de pacote, a camada de rede irá dividi-la e em seguida enviar cada um dos pacotes ao roteador.
- A camada de rede pode oferecer dois tipos de serviços à camada de transporte:
  - Serviços sem Conexão (Datagrama)
  - Serviços Orientados à Conexão (Circuito Virtual)

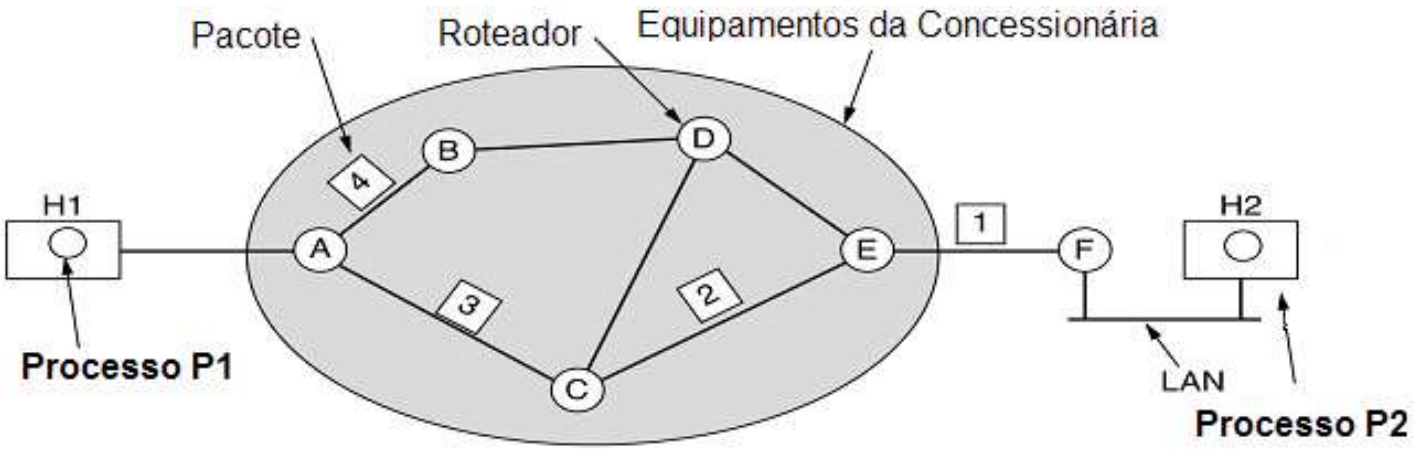
Camada de Rede - Serviços

- Serviços sem Conexão (Datagrama)
  - Também conhecidos como Datagrama.
  - Pacotes serão colocados individualmente na sub-rede e roteados de modo independente uns dos outros.
  - Não é necessário nenhuma configuração antecipada.
- Serviços Orientados à Conexão (Circuito Virtual)
  - Serviço possui 3 fases distintas: Estabelecimento de Conexão; Transferência de Dados e Encerramento
  - Inicialmente deve ser estabelecida uma conexão. Desta forma, escolhe-se uma rota desde a máquina de origem até a máquina de destino, como parte da configuração desta conexão.
  - Esta rota é armazenada em tabelas internas dos roteadores.
  - A rota é usada por todo o tráfego que flui pela conexão, semelhante ao que ocorre no sistema telefônico. Quando a conexão é liberada, o circuito virtual também é encerrado

Serviços – Exemplo

- Independente do tipo de serviço. A camada de rede pode ser vista como dois processos pares, em máquinas diferentes que deseja enviar pacotes de dados um para o outro
- Caso um processo (P1) deseje enviar uma longa mensagem para o processo P2.
- P1 entrega a mensagem à camada de transporte, com instruções para que ela seja entregue a P2 do host H2.
- O código da camada de transporte funciona em H1, em geral dentro do sistema operacional.
- Ele acrescenta um cabeçalho de transporte ao início da mensagem e entrega o resultado à camada de rede, que basicamente é outro procedimento do sistema operacional.
- Caso a mensagem seja mais longa que o tamanho máximo de pacote, a camada de rede irá dividi-la e em seguida enviar cada um dos pacotes ao roteador.
- Nesse ponto, a concessionária de comunicações assume o controle.
- Cada roteador possui internamente tabelas que contém informações utilizadas para o roteamento dos pacotes

Serviços Sem Conexão - Exemplo



Tabelas de A

inicialmente	
A	–
B	B
C	C
D	B
E	C
F	C

Dest. Linha

obstrução	
A	–
B	B
C	C
D	B
E	B
F	B

Tabela de C

A	A
B	A
C	–
D	D
E	E
F	E

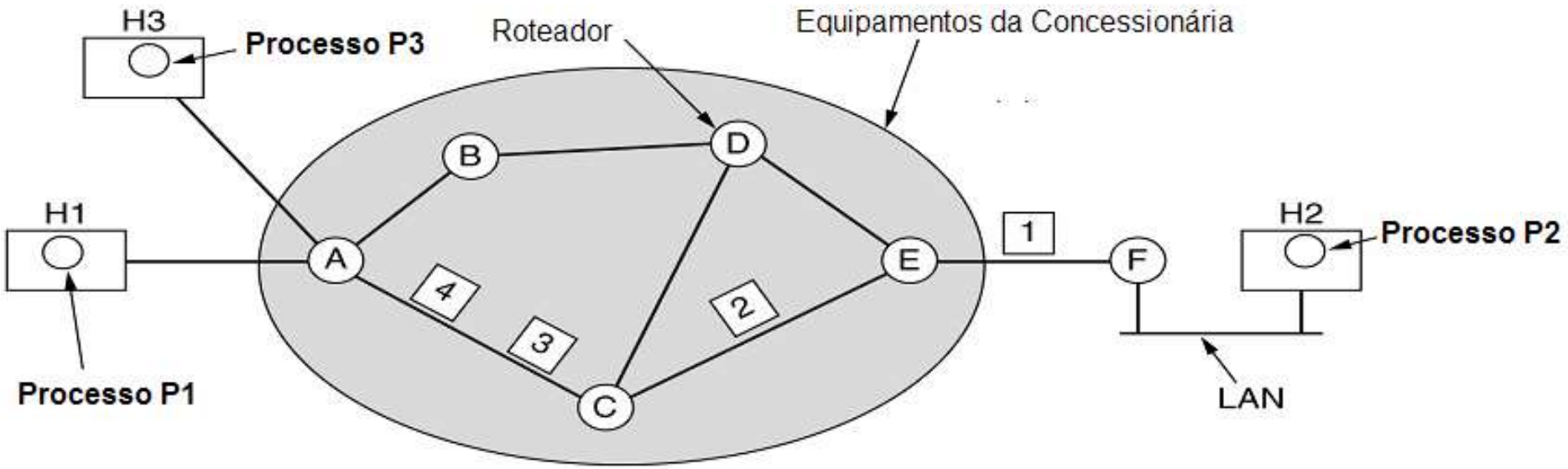
Tabela de E

A	C
B	D
C	C
D	D
E	–
F	F

Serviços Sem Conexão - Exemplo

- Roteador contém uma tabela interna que consiste de um par que contém um destino e a linha de saída a ser utilizada para esse destino.
- Inicialmente a camada de rede divide os pacotes e o entrega ao roteador (A) usando algum protocolo ponto a ponto como, por exemplo, o PPP Pacotes(1,2,3) são recebidos pelo roteador (A) e são armazenados por algum tempo (checksum)
- Em seguida, cada um deles é encaminhado para C, de acordo com a tabela de A.
- Pacote 1 é encaminhado para E e depois para F.
- Chegando a F, ele foi encapsulado em um quadro da camada de enlace de dados e transmitido para H2 pela LAN. Os pacotes 2 e 3 seguiram a mesma rota.
- Entretanto, devido a uma obstrução de tráfego no caminho ACE, por exemplo, a tabela de A é modificada e o pacote 4 é enviado por uma rota diferente.

Serviços Com Conexão - Exemplo



Tabelas de A

H1	1
H3	1

entrada

C	1
C	2

saída

Tabelas de C

A	1
A	2

E	1
E	2

Tabelas de E

C	1
C	2

F	1
F	2

Serviços Com Conexão

- O host H1 estabeleceu a conexão 1 com o host H2. Ela é memorizada como a primeira entrada (1) de cada uma das tabelas de roteamento.
- A primeira linha da tabela de A informa que, se um pacote contendo o identificador de conexão 1 chegar de H1, ele será enviado ao roteador C e receberá o identificador de conexão 1.
- De modo semelhante, a primeira entrada em C faz o roteamento do pacote para E, também com o identificador de conexão 1.
- Caso H3 deseje estabelecer uma conexão com H2. Ele escolhe o identificador de conexão 1, visto que é a primeira conexão e informa à sub-rede a necessidade de se estabelecer o circuito virtual.
- Isso conduz à segunda linha nas tabelas. Observe que nesse caso temos um conflito porque, embora A possa distinguir facilmente os pacotes da conexão 1 provenientes de H1 dos pacotes da conexão 1 que vêm de H3, C não tem como fazer o mesmo.
- Para evitar conflitos, o roteador A atribui um identificador de conexão diferente ao tráfego de saída correspondente a segunda conexão.
- Todos os pacotes irão trafegar seguindo a rota estabelecida na conexão

Comparação entre os Serviços

CRITÉRIO	Sub-rede de datagramas	Sub-rede de circuitos virtuais
Configuração de circuitos	Desnecessária	Obrigatória
Endereçamento	Cada pacote contém os endereços de origem e de destino completos	Cada pacote contém um número de circuito virtual curto
Informações sobre o estado	Os roteadores não armazenam informações sobre o estado das conexões	Cada circuito virtual requer espaço em tabelas de roteadores por conexão
Roteamento	Cada pacote é roteado independentemente	A rota é escolhida quando o circuito virtual é estabelecido; todos os pacotes seguem essa rota
Efeito de falhas no roteador	Nenhum, com exceção dos pacotes perdidos durante a falha	Todos os circuitos virtuais que tiverem passado pelo roteador que apresentou o defeito serão encerrados
Qualidade de serviço	Difícil	Fácil, se for possível alocar recursos suficientes com antecedência para cada circuito virtual
Controle de congestionamento	Difícil	Fácil, se for possível alocar recursos suficientes com antecedência para cada circuito virtual

Algoritmos de Roteamento

- A principal função da camada de rede é rotear pacotes da máquina de origem para a máquina de destino.
- Na maioria das sub-redes, os pacotes necessitarão de vários saltos (hops) para cumprir o trajeto.
- Os algoritmos que escolhem as rotas e as estruturas de dados que eles utilizam constituem um dos elementos mais importantes do projeto da camada de rede.
- O algoritmo de roteamento é responsável pela decisão sobre a linha de saída a ser usada na transmissão do pacote de entrada.
- O algoritmo sempre será utilizado independente se a subrede utiliza datagramas ou circuitos virtuais.

Algoritmos de Roteamento

- O algoritmo de roteamento deve possuir algumas propriedades:
  - **Correção -** Calcula rotas corretas para todos os destinos, não podendo indicar uma rota inexistente e a também falhar no envio
  - **Simplicidade -** Algoritmo deve ser eficiente sem sobrecarregar a máquina, permitindo ao administrador da rede o entendimento do mesmo.
  - **Robustez –** Capacidade de aceitar as alterações na topologia e no tráfego; assimilar falhas nos hosts; etc. sem necessidade de interromper todos os hosts e sem reinicialização da rede
  - **Estabilidade –** Algoritmo rapidamente converge para um estado correto. Por exemplo, ao alterar a topologia da rede tabelas ficam momentaneamente incorretas até convergir a um estado correto.
  - **Eqüidade –** Possibilidades iguais para todos os hosts enviarem pacotes
  - **Otimização –** Busca a minimização do retardo médio de pacote; maximização da vazão (throughput) total da rede; menor caminho; etc.

Algoritmos de Roteamento

- Os algoritmos de roteamento podem ser divididos em dois grupos:
  - Roteamento não adaptativo (estático)
    - Não baseiam suas decisões de roteamento em medidas ou estimativas do tráfego ou mesmo na topologia da rede.
    - A escolha da rota a ser utilizada para ir de I até J (para todo I e todo J) é previamente calculada antes do uso e transferida para os roteadores quando a rede é inicializada.
  - Roteamento adaptativo (dinâmico)
    - Mudam suas decisões de roteamento para refletir mudanças na topologia e, normalmente, também no tráfego.
    - Pode alterar sua rota quando a cada ΔT segundos, quando a carga se altera ou quando a topologia muda
    - Algoritmo pode obter informações no roteador local; em roteadores adjacentes ou de todos os roteadores
    - Utilizam várias unidades métricas para a otimização (por exemplo, distância, número de hops ou tempo de trânsito estimado; custo comunicação;)

Roteamento não Adaptativo(Estático)

- Pode ser utilizado em uma rede com um número limitado de roteadores
- Uma tabela de roteamento estático é construída manualmente pelo administrador do sistema, e pode ou não ser divulgada para outros dispositivos de roteamento na rede.
- Tabelas estáticas não se ajustam automaticamente a alterações na rede, portanto devem ser utilizadas somente onde as rotas não sofrem alterações.
- Algumas vantagens do roteamento estático são a segurança obtida pela não divulgação de rotas que devem permanecer escondidas; e a redução do overhead introduzido pela troca de mensagens de roteamento na rede

Roteamento Adaptativo (Dinâmico)

- Redes com mais de uma rota possível para o mesmo ponto devem utilizar roteamento dinâmico.
- Uma tabela de roteamento dinâmico é construída a partir de informações trocadas entre protocolos de roteamento.
- Os protocolos são desenvolvidos para distribuir informações que ajustam rotas dinamicamente para refletir alterações nas condições da rede.
- Protocolos de roteamento podem resolver situações complexas de roteamento mais rápida e eficientemente que o administrador do sistema.
- Protocolos de roteamento são desenvolvidos para trocar para uma rota alternativa quando a rota primária se torna inoperável e para decidir qual é a rota preferida para um destino.
- Em redes onde existem várias alternativas de rotas para um destino devem ser utilizados protocolos de roteamento.

Algoritmos de Roteamento

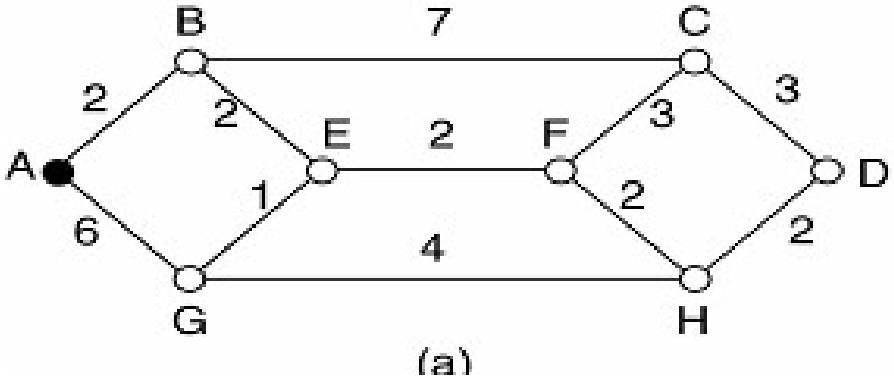
- Os algoritmos de roteamento podem ainda ser classificados como:
  - Algoritmos Intra-domínio
    - Estes são algoritmos que são executados por roteadores de dentro de um determinado Sistema Autônomo (AS-Autonomous System).
    - Sistema autônomo é uma coleção de roteadores sob uma mesma administração
    - Permitem que sejam definidas as rotas para dentro da rede de uma determinada organização.
  - Algoritmos Inter-domínios
    - Estes são algoritmos que são executados por roteadores que estão nos limites dos domínios.
    - Permitem a definição das rotas que são utilizadas para a comunicação com equipamentos de fora de um determinado Sistema Autônomo.

Exemplos de Algoritmos

- Não Adaptativos (Estáticos)
  - Roteamento pelo caminho mais curto (Shortest Path Routing)
  - Inundação (flooding)
- Adaptativos (Dinâmicos)
  - Roteamento com vetor de distância (Distance Vector Routing)
  - Roteamento por estado de enlace (Link State Routing)
  - Normalmente utilizado inter-domínios

Roteamento Pelo Caminho Mais Curto

- O algoritmo de roteamento pelo caminho mais curto (Shortest Path Routing) funciona da seguinte forma:
  - A rede é representada por um grafo onde cada vértice é um roteador;
  - O peso das arestas é uma função qualquer que leva em consideração alguma variável da rede (distância, velocidade do link, tráfego, hops, etc);
  - O algoritmo consiste em encontrar o menor caminho no grafo (Dijkstra, 1959)



(a)

Roteamento por Inundação

- O algoritmo de roteamento inundação (flooding) funciona da seguinte forma:
  - Todos os pacotes são enviados para todos os roteadores vizinhos exceto de onde veio;
  - Produz um grande tráfego enorme na rede;
  - A fim de evitar que um pacote seja transmitido indefinidamente, pode ser utilizado um contador de saltos(hops) no cabeçalho de cada pacote. O contador é decrementado em cada salto, sendo o mesmo descartado quando o contador atingir zero.
  - Este algoritmos não é muito utilizado na prática mas em algumas aplicações é interessante:
    - Aplicações militares devido a sua grande robustez;
    - Aplicações de banco de dados distribuídos;
    - Flooding sempre encontra o menor caminho, visto que os pacotes seguem por todas as rotas possíveis.

Roteamento com vetor de distância

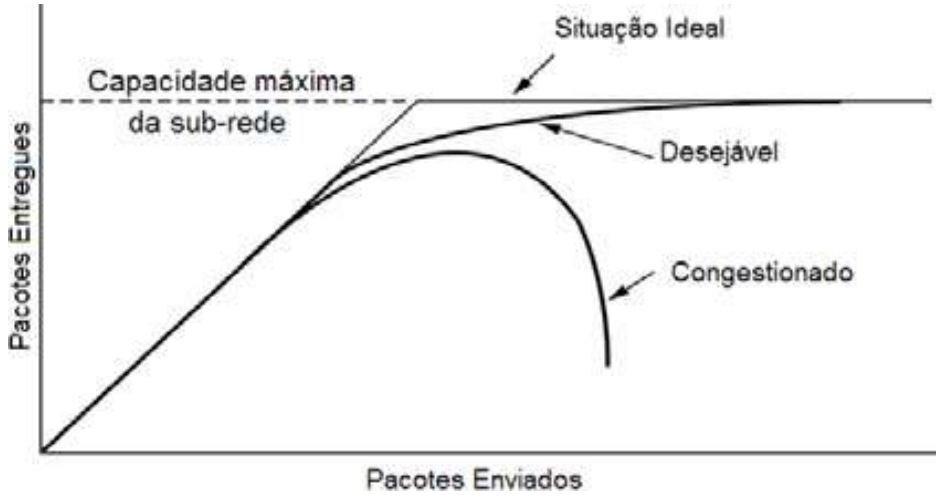
- O roteamento com vetor de distância (Distance Vector Routing) funciona da seguinte forma:
  - Cada roteador mantém uma tabela (vetor de rotas) contendo as distâncias até todos os outros roteadores da mesma sub-rede;
  - A distância pode ser medida em função de variáveis diferentes, mas em geral é usado o tempo de ida e volta de um pacote ECHO (ping);
  - De tempos em tempos cada roteador envia a sua tabela para todos os seus roteadores vizinhos;
  - Quando o roteador recebe as tabelas dos seus vizinhos, soma o tempo de propagação até eles, e encontra a menor distância até cada roteador
  - Converge para uma situação ideal muito lentamente, principalmente quando um roteador sai do ar.
  - Foi utilizado no início da ARPANET até 1979;

Roteamento por estado de enlace

- O roteamento por estado de enlace (Link State Routing) funciona da seguinte forma:
  - Descobre quais são seus vizinhos e obtém os seus endereços;
    - O roteador envia um pacote HELLO;
  - Mede o atraso ou custo para cada vizinho;
  - Constrói um pacote de resposta a respeito das informações coletadas;
    - O pacote contém: quem construiu, a idade do pacote e uma lista de vizinhos e tempo de acesso p/ cada um;
    - Um problema é determinar quando esses pacotes serão construídos (periodicamente ou quando ocorrer um determinado evento);
  - Envia esse pacote para todos os outros roteadores;
    - A distribuição é feita utilizando a inundação(flooding). Se a distribuição não for bem feita, a topologia pode ficar inconsistente contendo loop ou máquinas inatingíveis;
    - Cada roteador mantém uma lista dos pacotes que já recebeu para eliminar os repetidos;
    - Se um pacote é mais antigo que um já recebido, o mesmo e descartado
  - Calcula o menor caminho (shortest path) para todos os outros roteadores;

Controle de congestionamento

- O congestionamento ocorre quando há pacotes demais presentes em uma parte da sub-rede.
- O congestionamento acarreta uma queda no desempenho da rede, visto que o trafego aumenta muito e os roteadores não são capazes de suporta-lo.
- Os pacotes são perdidos e no caso de um tráfego muito intenso nenhum pacote será entregue.



Controle de congestionamento

- As causas dos congestionamentos podem ser diversas:
  - Pacotes chegando de diversas linhas com o mesmo destino, neste caso todos irão seguir a mesma saída e haverá uma fila
  - Falta de memória nos roteadores para manipular todos os pacotes CPUs dos roteadores lentas em tarefas operacionais (enfileiramento em buffers, atualização de tabelas etc.), poderão surgir filas mesmo que haja capacidade de linha suficiente
  - Linhas de baixa velocidade
- O controle de congestionamento é diferente do controle de fluxo.
  - Controle de Fluxo é um conceito ponto-a-ponto. Evitar que transmissor rápido sobrecarregue um receptor lento
  - Controle de congestionamento envolve o desempenho global no transporte de pacotes da sub-rede.
  - Host pode receber uma mensagem "reduzir velocidade", porque o receptor está sobrecarregado, ou porque a rede não é capaz de tratá-la

Controle de congestionamento

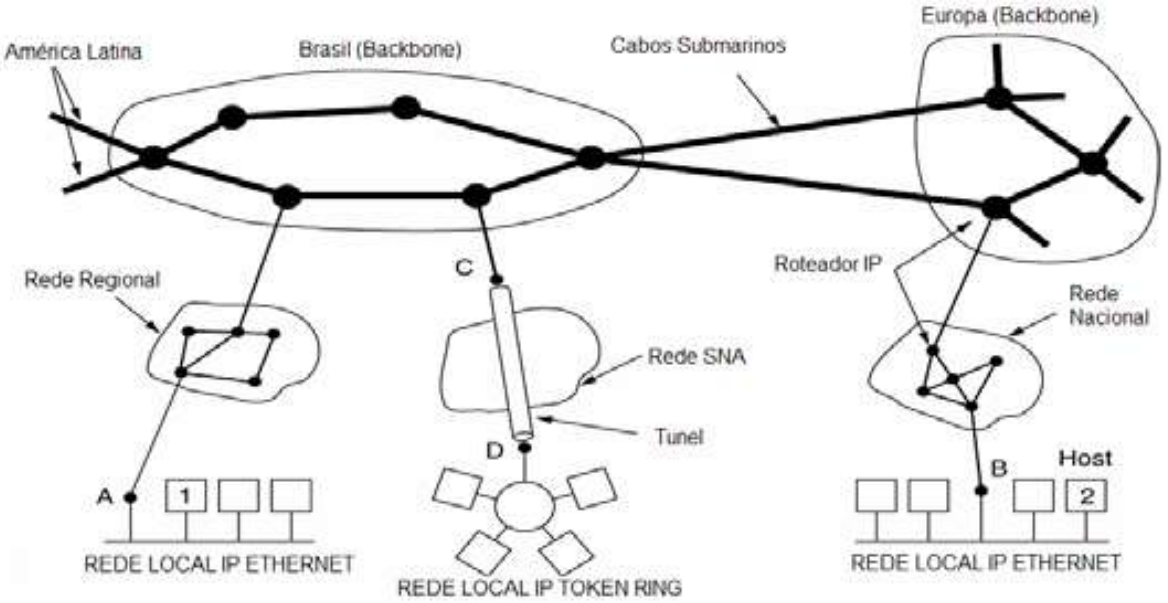
- Existem dois algorimtos para se lidar com o congestionamento
  - Loops Abertos (Open-loop)
    - Criados para evitar o congestionamento e durante o funcionamento pequenos ajustes podem ser feitos;
    - Baseia-se na prevenção do congestionamento
  - Loops Fechado (Closed-loop)
    - São criados para responderem a feedbacks do sistema adaptando-se à situação atual;
    - Baseia-se na correção do congestionamento

Controle de congestionamento

- O controle de congestionamento baseia-se nos seguintes princípios:
  - Monitorar o sistema para detectar quando e onde ocorre congestionamento. Pode ser monitorado:
    - A percentagem de todos os pacotes descartados por falta de espaço em buffer;
    - A média dos comprimentos de fila;
    - O número de pacotes interrompidos por alcançarem o tempo limite e que são retransmitidos;
    - O retardo médio de pacotes
  - Enviar essas informações para lugares onde alguma providência possa ser tomada.
    - Enviar um pacote à origem ou às origens de tráfego, anunciando o problema. Esses pacotes extras aumentam a carga exatamente no momento em que a rede já está congestionada.
    - Reservar um bit ou um campo em todos os pacotes. O bit será igual a 1 caso esteja ocorrendo o congestionamento, isto irá alertar os vizinhos que receberem os pacotes
    - Outra abordagem é fazer com que os hosts ou roteadores enviem pacotes de sondagem periodicamente para perguntar de forma explícita sobre o congestionamento
  - Ajustar a operação do sistema para corrigir o problema

Camada de Rede na Internet

- O Internet Protocol (IP) foi projetado tendo como objetivo a interligação de redes
- IP deve fornecer a melhor forma possível de transportar datagramas da origem para o destino, independente dessas máquinas estarem na mesma rede ou de haver outras redes entre elas.
- Protocolo não fornece garantia de entrega.



Comunicação na internet

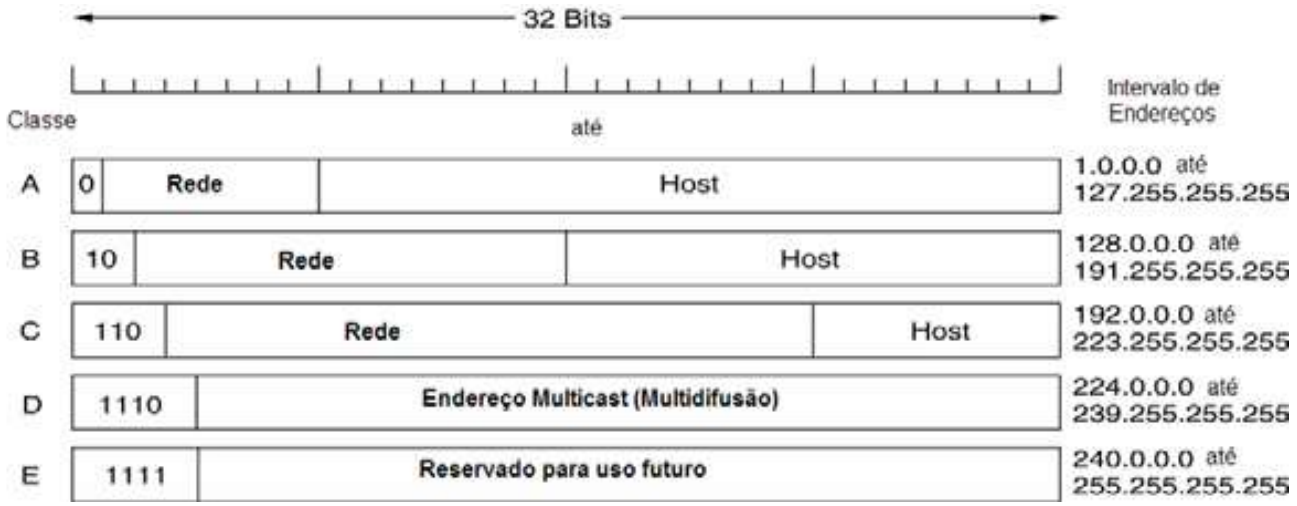
- A camada de transporte recebe os fluxos de dados e os divide em datagramas.
- Cada datagrama pode ter até 64 Kbytes, porém geralmente eles têm no máximo 1500 bytes que equivale a um quadro Ethernet
- Datagrama é transmitido pela Internet. Normalmente é fragmentado em unidades menores durante o percurso até o destino.
- Quando todos os fragmentos finalmente chegam à máquina de destino, eles são remontados pela camada de rede no datagrama original.
- Em seguida, esse datagrama é entregue à camada de transporte, que o insere no fluxo de entrada do processo de recepção

Endereços IP

- Na Internet, cada host e cada roteador tem um endereço IP que codifica seu número de rede e seu número de host.
- A combinação é exclusiva: em princípio, duas máquinas na Internet nunca têm o mesmo endereço IP.
- Endereços IP têm 32 bits e são usados nos campos Source address e Destination address dos pacotes IP.
- O endereço IP normalmente é escrito na forma decimal 192.160.0.1
- Um endereço IP não se refere realmente a um host, na verdade, ele se refere a uma interface de rede; assim, se um host estiver em duas redes, ele precisará ter dois endereços IP.

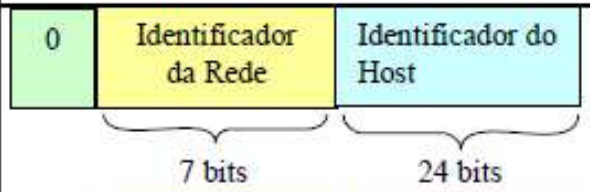
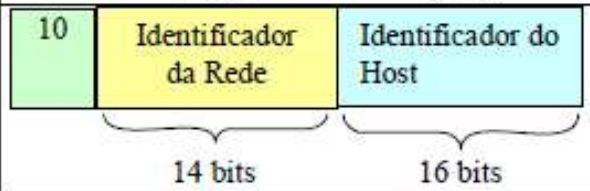
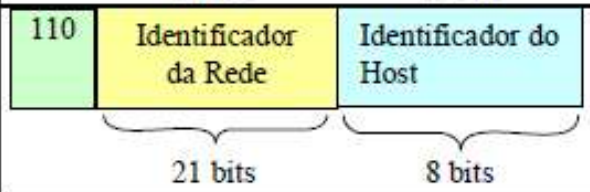
Endereços IP

- O endereço é divido em duas partes: Uma identifica a rede e a outra
- identifica o host.
- O endereço IP mais baixo é 0.0.0.0 e o mais alto é 255.255.255.255
- Na classe A o endereço varia da seguinte forma: w.0.0.1 até w.255.255.254
- Na classe B endereço varia da seguinte forma: w.x.0.1 w.x.255.254
- Na classe C endereço varia da seguinte forma: w.x.y.1 w.x.y.254

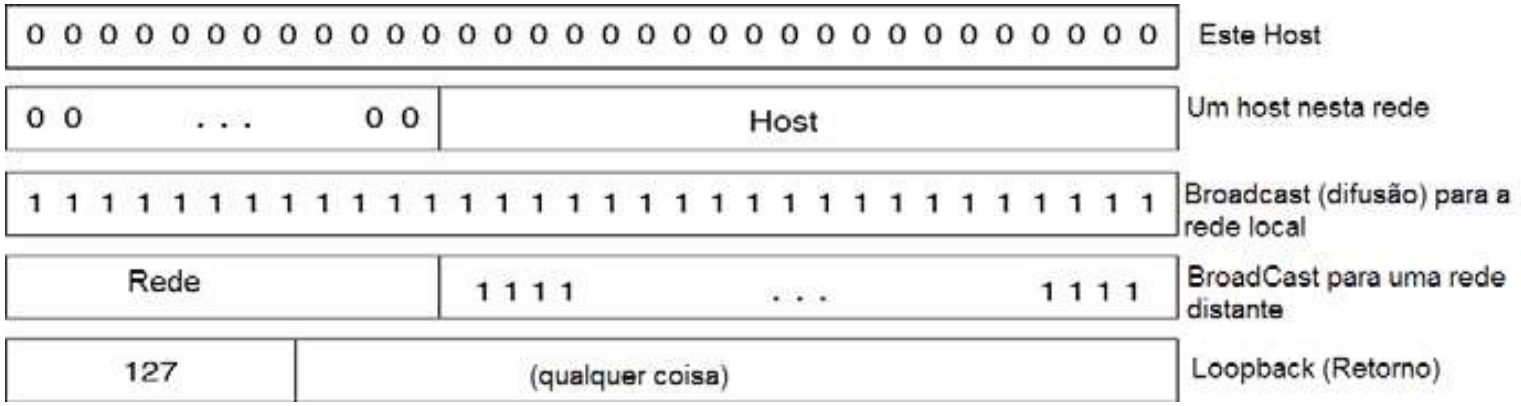


Endereços IP

- Todos os hosts de uma rede devem possuir o mesmo identificador de rede

Classe	Formato do Endereço	Organização da Rede	Intervalo dos endereços da classe
A		127 redes com até 16777216 hosts.	de 1.0.0.0 até 127.255.255.255.
B		16384 redes com até 65535 hosts.	de 128.0.0.0 até 191.255.2555.255.
C		2097152 redes com até 255 hosts.	de 192.0.0.0 até 233.255.2555.255.

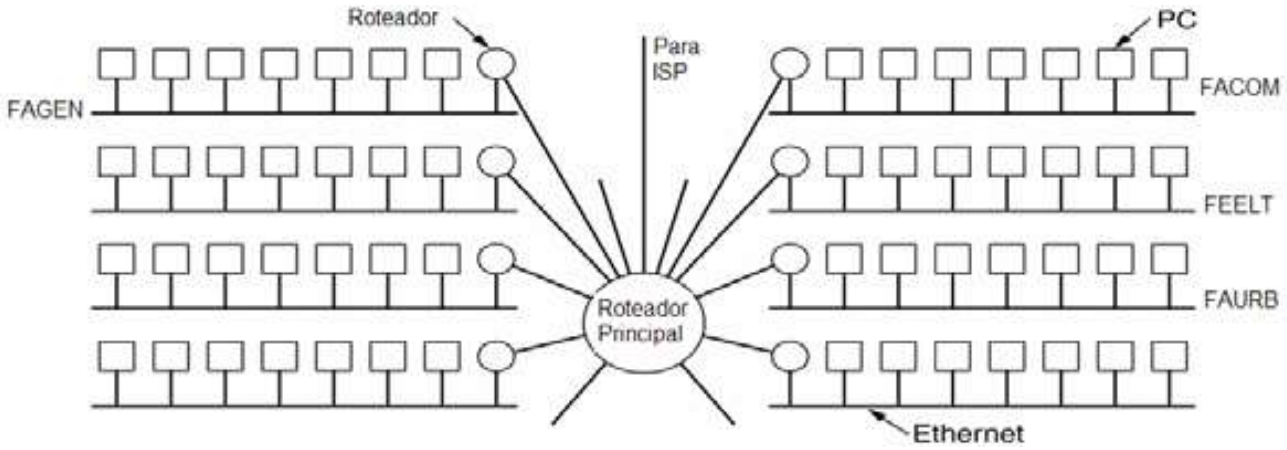
Endereços IP Especiais



- O endereço no formato 127.x.y.z são reservados para teste de loopback (retorno).
- Neste caso os pacotes enviados para esse endereço não são transmitidos; eles são processados localmente e tratados como pacotes de entrada
- Broadcast (Difusão) – Envio da mesma informação para todos
- Multicast (Multidifusão) – Envio da mesma informação para um grupo.

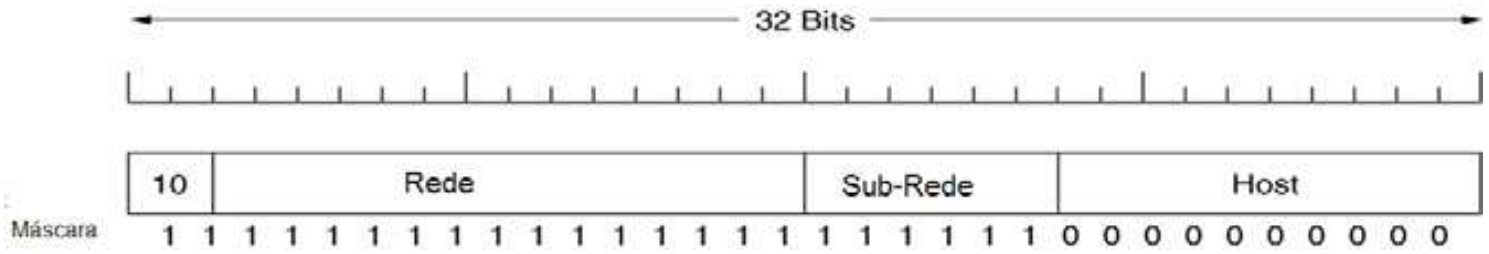
Sub-Rede

- Todos os hosts de uma rede devem ter o mesmo identificador de rede.
- Essa propriedade do endereçamento IP poderá causar problemas à medida que as redes crescem.
- Para resolver este problema pode uma rede pode ser divida em diversas partes para uso interno.
- A solução para esses problemas é permitir que uma rede seja dividida em diversas partes para uso interno, mas externamente continue a funcionar como uma única rede.



Sub-Rede

- A Máscara de Sub-rede consiste em 32 bits em notação decimal pontuada.
  - bits 1 indicam o endereço da sub-rede
  - bits 0 o endereço do host.
- Máscaras Default
  - classe A: 255.0.0.0 ou 11111111.00000000. 00000000. 00000000
  - classe B: 255.255.0.0 ou 11111111. 11111111. 00000000. 00000000
  - classe C: 255.255.255.0 ou 11111111. 11111111. 11111111. 00000000
- Ao receber um pacote o roteador realiza uma operação AND lógica bit a bit entre o endereço IP e a máscara de sub-rede a fim de descobrir qual o endereço de destino



Exemplo - Sub-rede

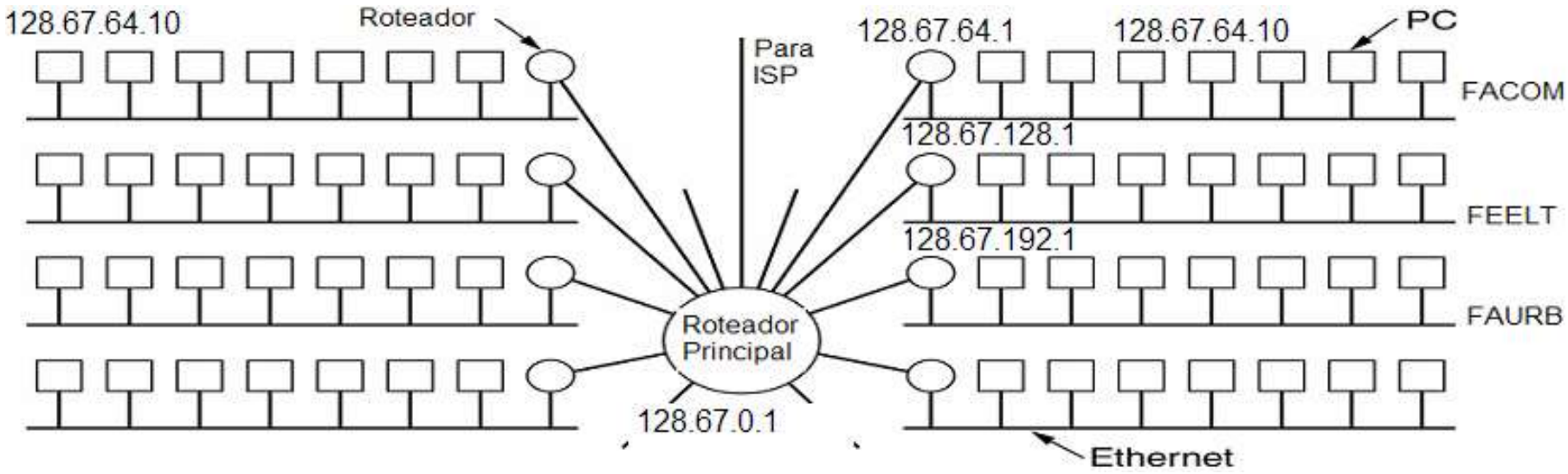
- Para dividir a rede em 4 sub-redes, em um endereço da classe, pode-se utilizar os dois primeiros bits do terceiro octeto a fim de indicar a sub rede ou seja: 11111111. 11111111. **xx000000. 00000000**
- Desta forma as seguintes máscaras de sub-rede devem ser utilizadas
  - 11111111. 11111111. **00000000. 00000000 (255.255.0.0)**
  - 11111111. 11111111. **01000000. 00000000 (255.255.64.0)**
  - 11111111. 11111111. **10000000. 00000000 (255.255.128.0)**
  - 11111111. 11111111. **11000000. 00000000(255.255.192.0)**
- Exemplo: a rede 128.67.x.x seria dividida em 4 sub-redes:

00	1: 128.67.0.0 a 128.67.63.255	16.384 hosts (2^{14})
01	2: 128.67.64.0 a 128.67.128.255	16.384 hosts (2^{14})
10	3: 128.67.128.0 a 128.67.191.255	16.384 hosts (2^{14})
11	4: 128.67.192.0 a 128.67.255.255	16.384 hosts (2^{14})

- Uma notação alternativa para representar a máscara de sub-rede é utilizar /nn, sendo n o número de bits na máscara
  - No exemplo acima /18

Exemplo - Sub-rede

- A rede do campus poderia ser dividida conforme mostrado abaixo.
- O endereço de um computador da sub-rede da FACOM teria o seguinte endereço: 128.67.64.10 e sua máscara de subrede seria: 255.255.64.0



CIDR
Classless InterDomain Routing

- Tentativa de solucionar o problema da escassez de endereços IPv4
- Não leva em conta as classes
- Considera a alocação de endereços em blocos continuos conforme a necessidade
- Necessita alterar a maneira como o ocorre o encaminhamento de pacotes
- Permite a agregação de várias rotas em uma única

University	First address	Last address	How many	Written as
Cambridge	194.24.0.0	194.24.7.255	2048	194.24.0.0/21
Edinburgh	194.24.8.0	194.24.11.255	1024	194.24.8.0/22
(Available)	194.24.12.0	194.24.15.255	1024	194.24.12/22
Oxford	194.24.16.0	194.24.31.255	4096	194.24.16.0/20

- Realiza um AND com a máscara aa fim de verificar se há um casamento com o endereço definido
- C: 11000010 00011000 00000000 00000000 11111111 11111111 11111000 00000000
- E: 11000010 00011000 00001000 00000000 11111111 11111111 11111100 00000000
- O: 11000010 00011000 00010000 00000000 11111111 11111111 11110000 00000000
- IP: 11000010 00011000 00010001 00000100 (194.24.17.4)
- Endereço da Rota Agregado - 194.24.0.0/19

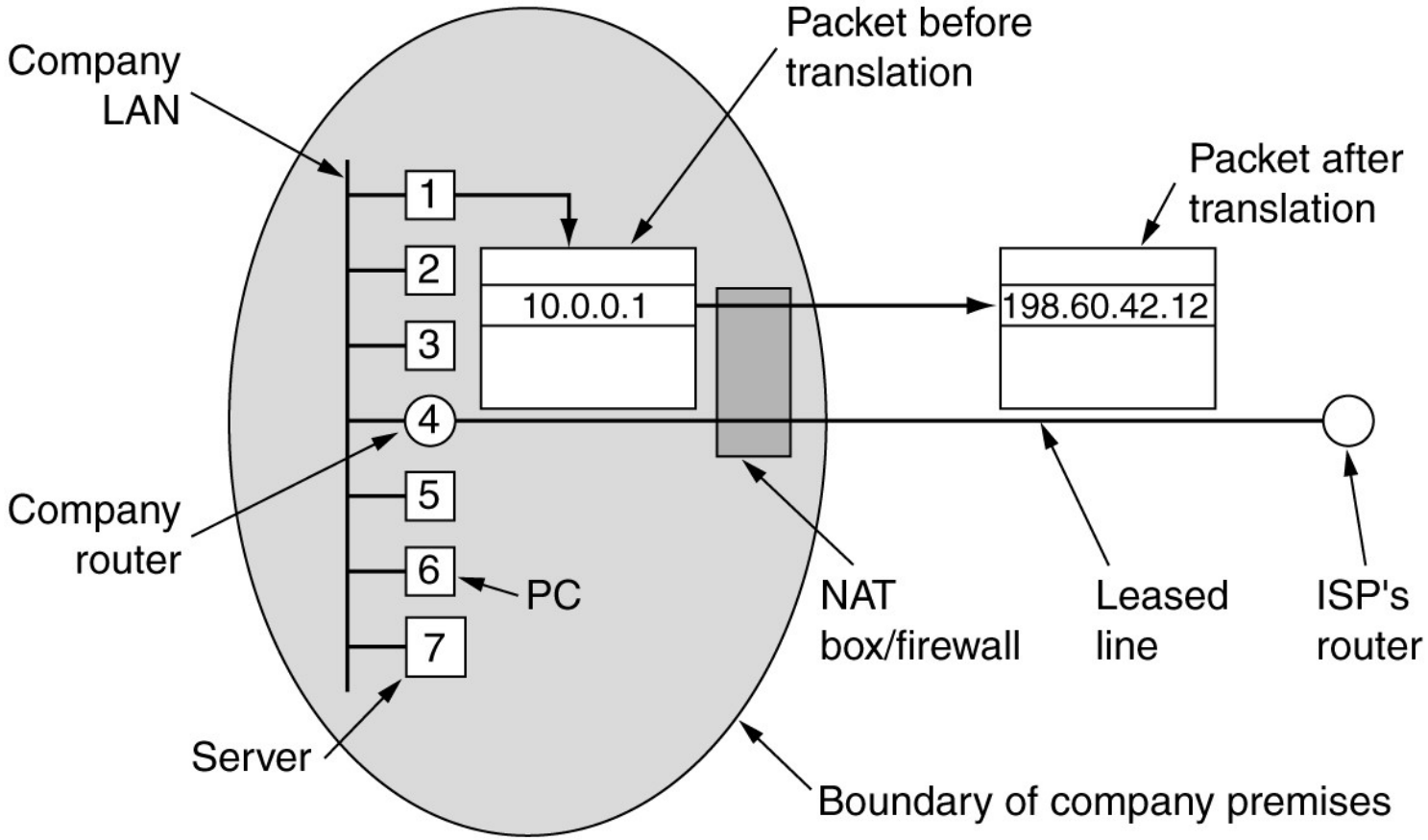
CIDR
Exemplo

- C: 11000010 00011000 00000000 00000000 11111111 11111111 11111000 00000000
- 11000010 00011000 00010001 00000100
- 11000010 00011000 00010000 00000000
- E: 11000010 00011000 00001000 00000000 11111111 11111111 11111100 00000000
- 11000010 00011000 00010001 00000100
- 11000010 00011000 00010000 00000000
- O: **11000010 00011000 00010000** 00000000 11111111 11111111 11110000 00000000
- 11000010 00011000 00010001 00000100
- **11000010 00011000 00010000 00000000**

NAT – Network Address Translation

- Endereços IPv4 escassos
- Internet em residências e pequenos negócios via ADSL ou Cabo
- IPs internos não são válidos na Internet
- Tradução de endereços por um dispositivo
  - Envio: Endereço da estação é trocado pelo endereço válido
  - Recebimento: Endereço válido é substituído pelo Endereço interno
- Entre outros problemas associados
  - Viola o principio de que cada host possui um IP na Internet
  - Necessita manter o estados de todas as conexões que possam pelo dispositivo
  - RFC 2993 – Architectural Implications of NAT

NAT – Network Address Translation



Protocolo IP

- Campos do protocolo

IPv4 Header Format

Offsets	Octet	0								1								2								3							
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version				IHL				DSCP						ECN		Total Length															
4	32	Identification																Flags			Fragment Offset											
8	64	Time To Live								Protocol								Header Checksum															
12	96	Source IP Address																															
16	128	Destination IP Address																															
20	160	Options (if IHL > 5)																															

- Version – IPv4 possui o valor 4
- IHL (Internet Header Length) – Indica o número de palavras de 32 bits existentes no header considerado que o mesmo pode ter tamanho variável (campo Options)
- DSCP (Differentiated Services Code) – Anteriormente conhecimento como Type of Service (ToS) este campo indica diferentes classes de serviço que *poderiam* influenciar a maneira como são encaminhados em função do tipo de aplicação (video, VoIP) e desta forma o pacote seria direcionado para filas apropriadas quanto a requisitos de latência e vazão, por exemplo.
- Explicit Congestion Notification (ECN) – Permite uma notificação fim-a-fim entre os os participantes da comunicação (endpoints). Opcional. Deve ser suportado pela rede
- Total Length – Indica o tamano total do Pacote (Header + Data) em bytes

Protocolo IP

- Campos do protocolo

IPv4 Header Format

Offsets	Octet	0								1								2								3							
Octet	Bit	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
0	0	Version				IHL				DSCP						ECN		Total Length															
4	32	Identification																Flags			Fragment Offset											
8	64	Time To Live								Protocol								Header Checksum															
12	96	Source IP Address																															
16	128	Destination IP Address																															
20	160	Options (if IHL > 5)																															

  - Identification – Campo utilizado para identificar fragmentos. Utilizado no processo de montagem do pacote após a fragmentação. Incrementado a cada pacote enviado
  - Flags – Três bits com a seguinte semântica (R-DF-MF). Bit 0 (Reservado); Bit 1 (Fragmenta se necessário); Bit 2 (Indica que o Datagrama contém mais fragmentos)
  - Fragment Offset – Caso haja fragmentação, contém o offset a partir do inicio do pacote
  - Time to Live (TTL) – Previne datagramas fiquem sendo encaminhados em ciclos. Limita o tempo de vida do pacote. Expresso em segundos. Na prática conta o número de saltos do pacote, sendo que cada roteador decrementa este campo em um. Ao chegar em zero o mesmo é descartado
  - Protocol – Define o protocolo que está contido no campo de dados (payload) do protocol
  - Header Checksum – campo de verificação com 16 bits, visando detecção de erros no cabeçalho. Verificado em cada roteador do trajeto
  - Source IP / Destination IP – Endereços de origem e destino
  - Options – Lista de opções para um datagrama. Opcional e pode não estar presente

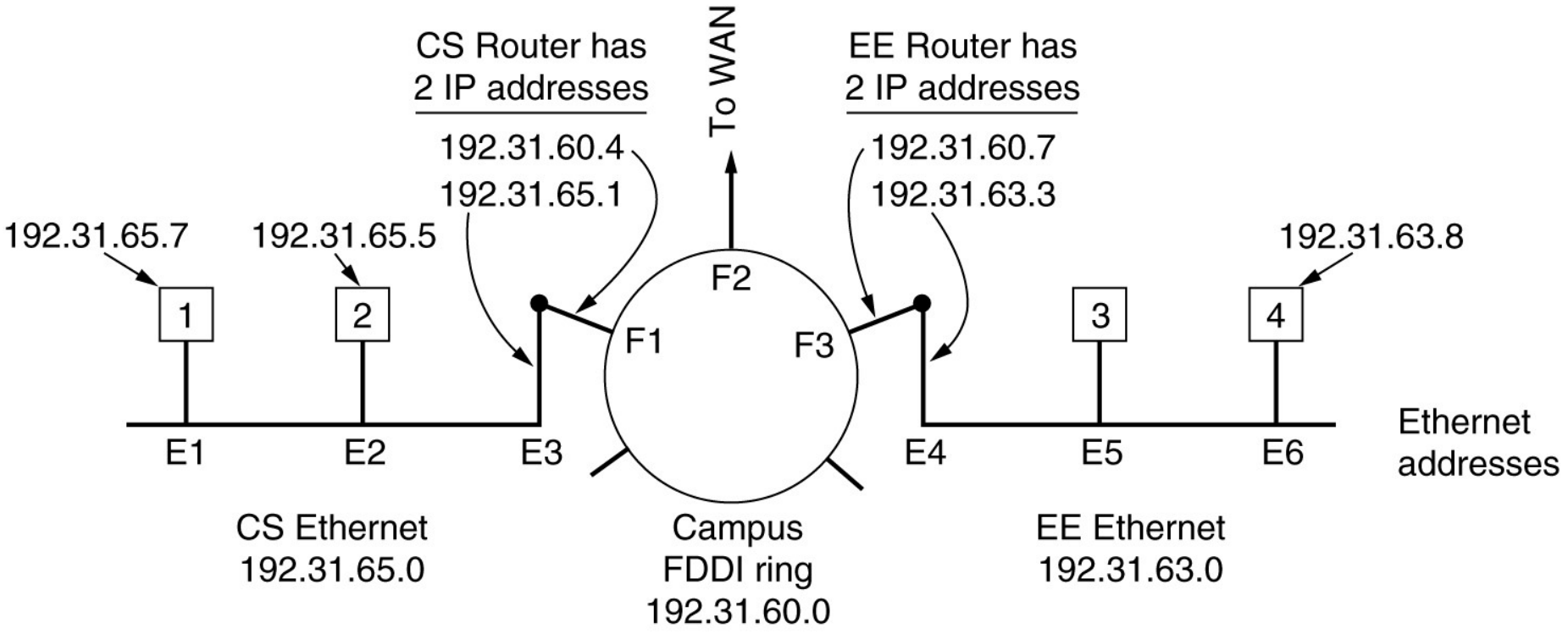
ICMP
Internet Control Message Protocol

- Utilizado pelos roteadores para reportar erros
- Não está relacionado ao transporte de dados de aplicações mas somente a informações entre roteadores

Message type	Description
Destination unreachable	Packet could not be delivered
Time exceeded	Time to live field hit 0
Parameter problem	Invalid header field
Source quench	Choke packet
Redirect	Teach a router about geography
Echo request	Ask a machine if it is alive
Echo reply	Yes, I am alive
Timestamp request	Same as Echo request, but with timestamp
Timestamp reply	Same as Echo reply, but with timestamp

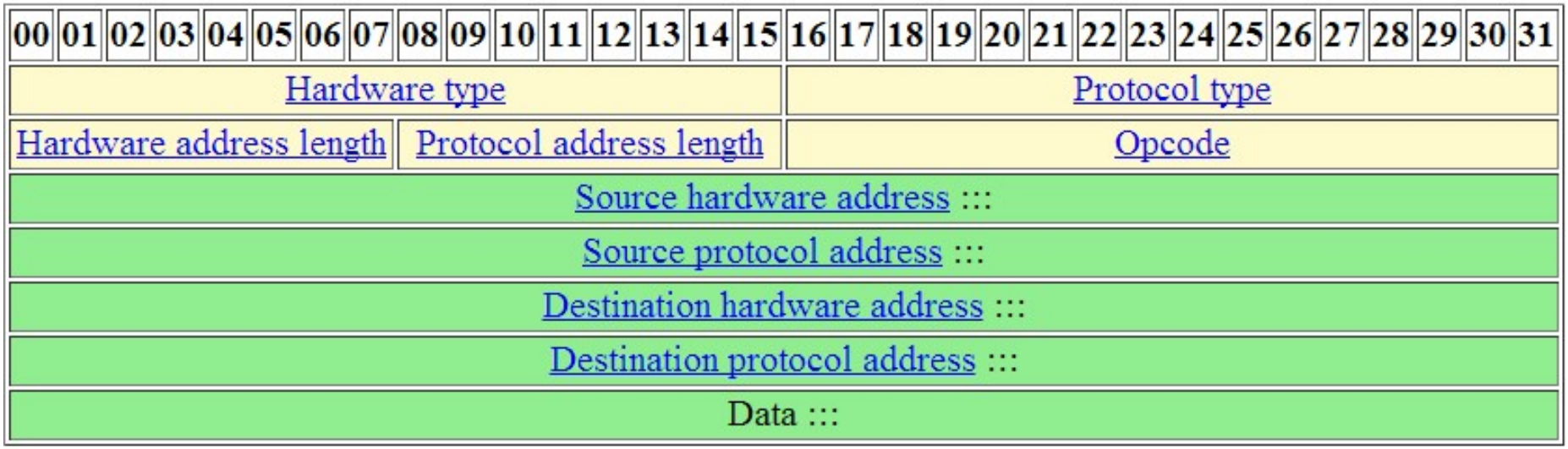
ARP
Address Resolution Protocol

- Permite o Mapeamento de endereços IP
- Um pacote é enviado na rede Ethernet (em broadcast) a fim de descobrir o MAC Address associadao a um dado IP



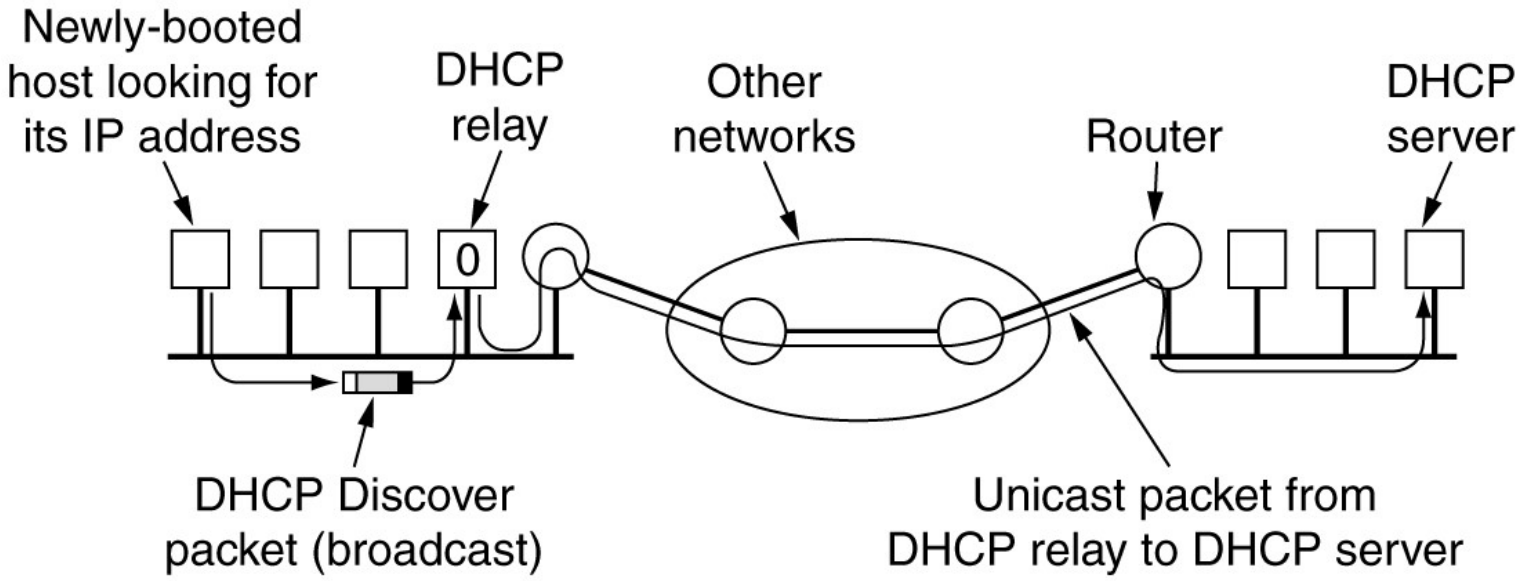
ARP
Address Resolution Protocol

- Cabeçalho do Protocolo



DHCP
Dynamic Host Configuration Protocol

- Baseado em um Servidor (DHCP Server) que atribui endereços IPs aos hosts solicitantes
- Servidor não precisa estar na mesma LAN. Neste caso é necessário um agente (DHCP Relay)
- Ao iniciar uma máquina envia um DHCP DISCOVER. Um IP dinâmico é cedido por um período de tempo pelo Server
- O DHCP é um protocolo da camada de aplicação



Internet
Roteamento

- Internet é composta de um conjunto de Sistemas Autônomos (*AS*)
- Cada AS é operado por uma empresa diferente e pode utilizar um algoritmo próprio de roteamento
- Roteamento interno a um AS (Interior Gateway Protocol)
  - RIP (Routing Information Protocol)
  - OSPF
- Roteamento entre ASes (Exterior Gateway Protocol)
  - BGP

OSPF
Open Shortest Path First

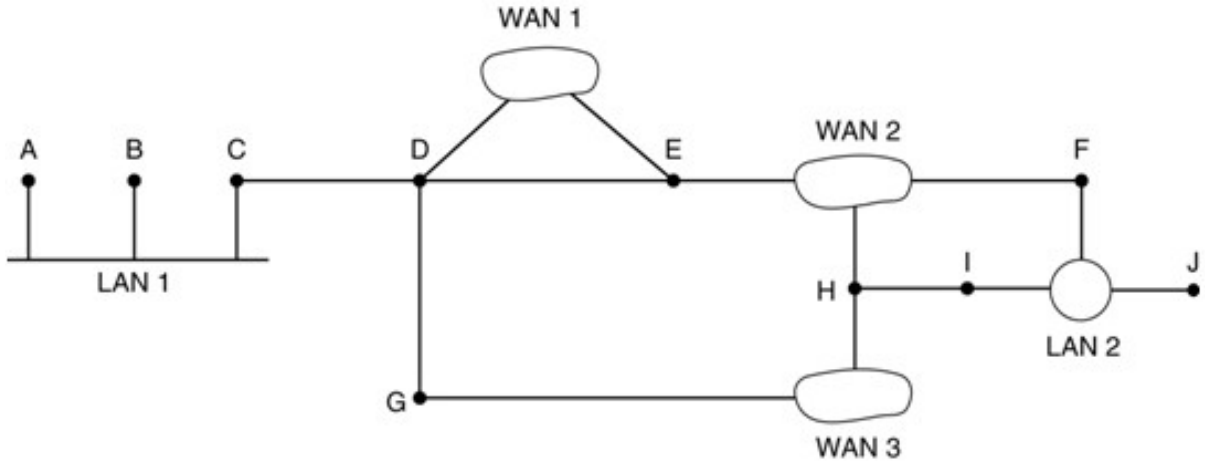
- Motivação
  - Inicialmente foi utilizado na Internet o RIP (Routing Information Protocol)
  - RIP adequado para redes pequenas
  - Convergência lenta
- OSPF
  - OSPFv1 – RFC 1131 (1989)
  - OSPFv2 – RFC 1247 (1991) – RFC 1583 (1994) – RFC 2178 (1997) – RFC 2328 (1998)
- Principal protocolo atualmente utilizado (IGP)
- Baseado no estado do enlace
- Cálculo do menor caminho a partir deste estado

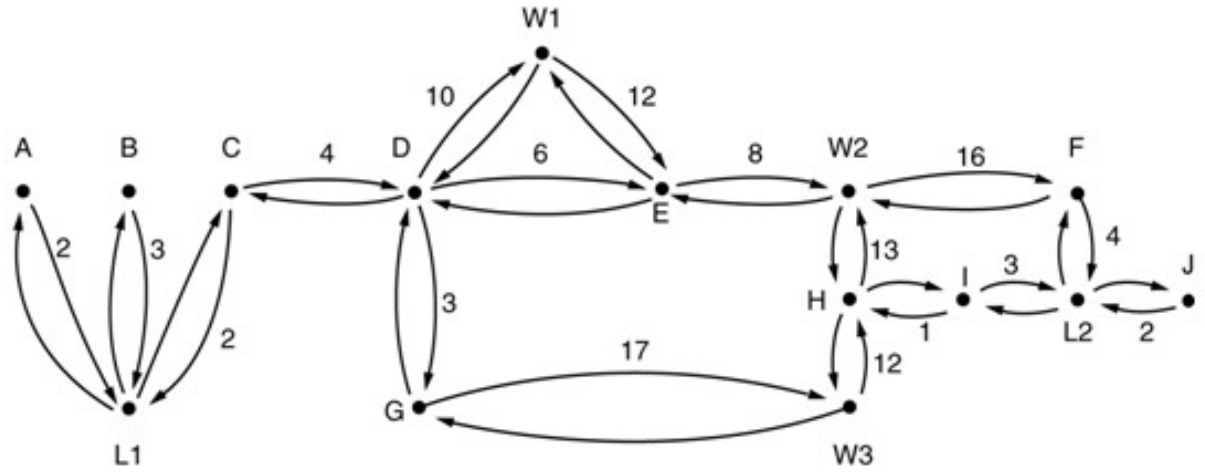
OSPF - Open Shortest Path First Características

- Admite uma variedade de unidades de medida deste estado (distância, retardo, etc.)
- Adaptativo conforme alterações na topologia da rede
- Realizar o balanceamento de carga, dividindo-a por várias linhas
- Suporte a sistemas hierárquicos, desta forma não seria necessário conhecer toda a topologia da rede
- Oferece um certo nível de segurança
- Tipos de Conexões e redes suportados
  - Linhas ponto-a-ponto entre dois roteadores
  - Redes Multiacesso com Difusão (LAN)
  - Redes Multiacesso sem Difusão (WAN baseadas em pacotes)
    - Em uma rede multiacesso pode haver vários roteadores que podem comunicar-se entre si

OSPF - Open Shortest Path First
Tipos de Rede

- Exemplo de Sistema Autônomo (AS)





OSPF
Dados Associados

- Em cada roteador existem os seguintes dados associados
  - Tabela de Roteamento
    - Única para cada roteador
    - Gerada a partir dos dados de cada enlace
  - Banco de dados de Adjacências
    - Lista de todos os vizinhos que um roteador estabeleceu uma comunicação bidirecional
  - Banco de Dados Topológico
    - Lista de todos os roteadores de uma rede
    - Contém informações de link state de todos os roteadores

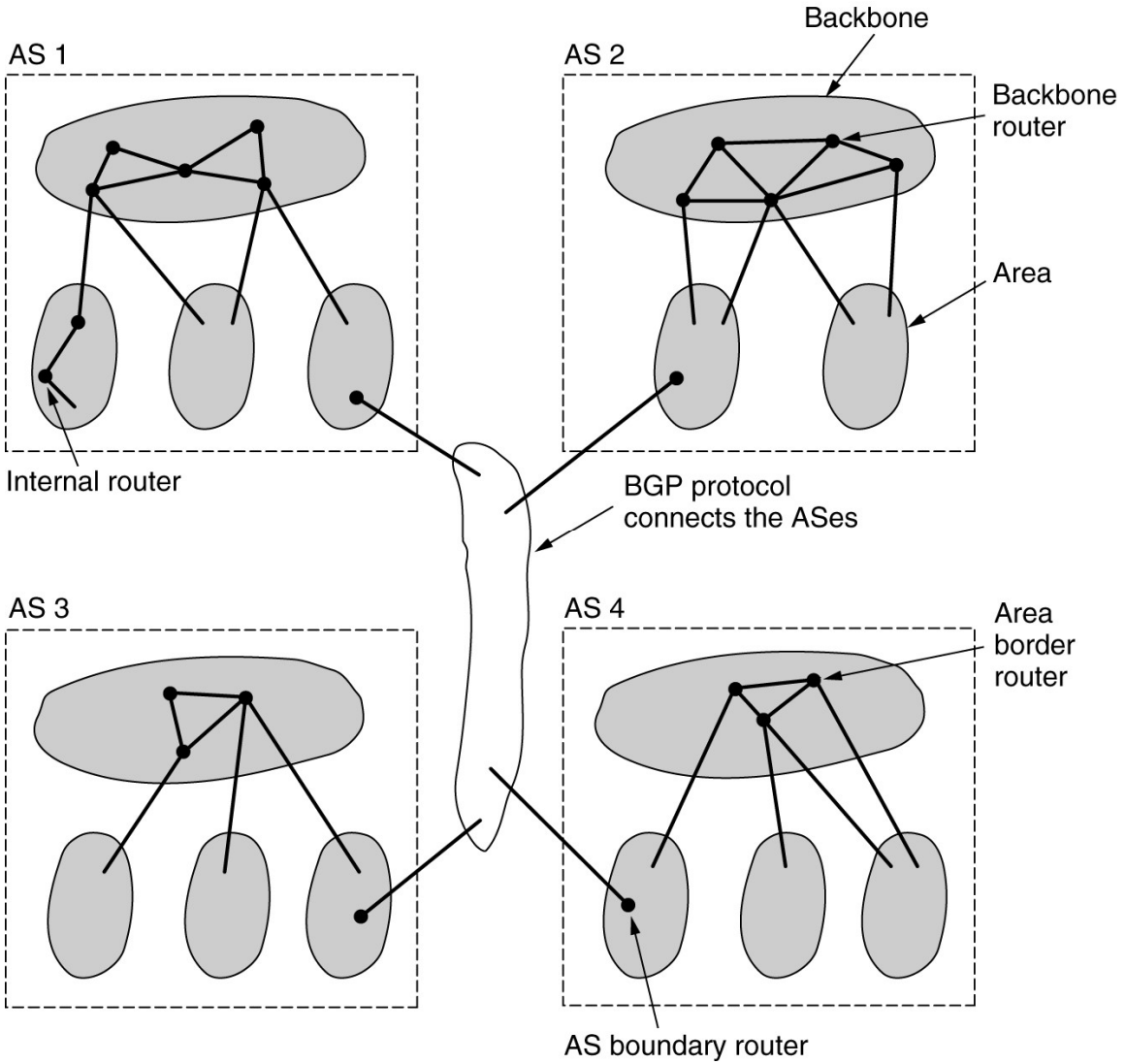
OSPF

- Como um AS pode ser uma rede muito ampla o OSPF prevê a divisão da mesma em diferentes áreas
- Divisão em áreas permite uma visão hierárquica da rede do AS
- Uma área é uma rede ou conjunto de redes contiguas que não se sobrepõem
- Fora de uma área a topologia e os detalhes da sub-rede não são visíveis
- Roteadores de uma mesma área compartilham o mesmo banco de dados e o mesmo algoritmo de caminho mais curto
- Áreas existentes
  - Área 0 – Backbone – Responsável por interligar todas as outras áreas
  - Área i – Conectada ao backbone
  - Cada roteador conectado a duas ou mais áreas faz parte do backbone (núcleo)

OSPF
Classes de Roteadores

- Roteadores Internos
  - Ficam inteiramente em uma área
- Roteadores de borda (border) de área (*Area Border Router*)
  - Conectam duas ou mais áreas
- Roteadores de backbone
  - Ficam no backbone
- Roteadores de fronteira (boundary) do AS (*ASBorder Router*)
  - Interagem com outros roteadores de outros AS

OSPF
Exemplo de Áreas



OSPF
Mensagens do Protocolo

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

- Ao ser inicializado o roteador enviar um HELLO para todas as suas linhas
  - A partir desta resposta roteador descobre seus vizinhos
  - No caso de uma LAN, como todos os roteadores são vizinhos
  - Em uma LAN um roteador é designado para representar todos os outros (DR – Designated Router)
  - As informações são trocadas entre roteadores vizinhos, adjacentes entre si
- Link State Update
  - Informa o estado e fornece os custos (atualiza o banco de dados da topologia)
  - Enviadas periodicamente por inundação a todos os vizinhos adjacentes
  - Possuem um número de sequencia e somente a mais recente é considerada
  - Enviadas quando uma linha é ativada ou quando seu custo se altera

OSPF
Mensagens do Protocolo

Message type	Description
Hello	Used to discover who the neighbors are
Link state update	Provides the sender's costs to its neighbors
Link state ack	Acknowledges link state update
Database description	Announces which updates the sender has
Link state request	Requests information from the partner

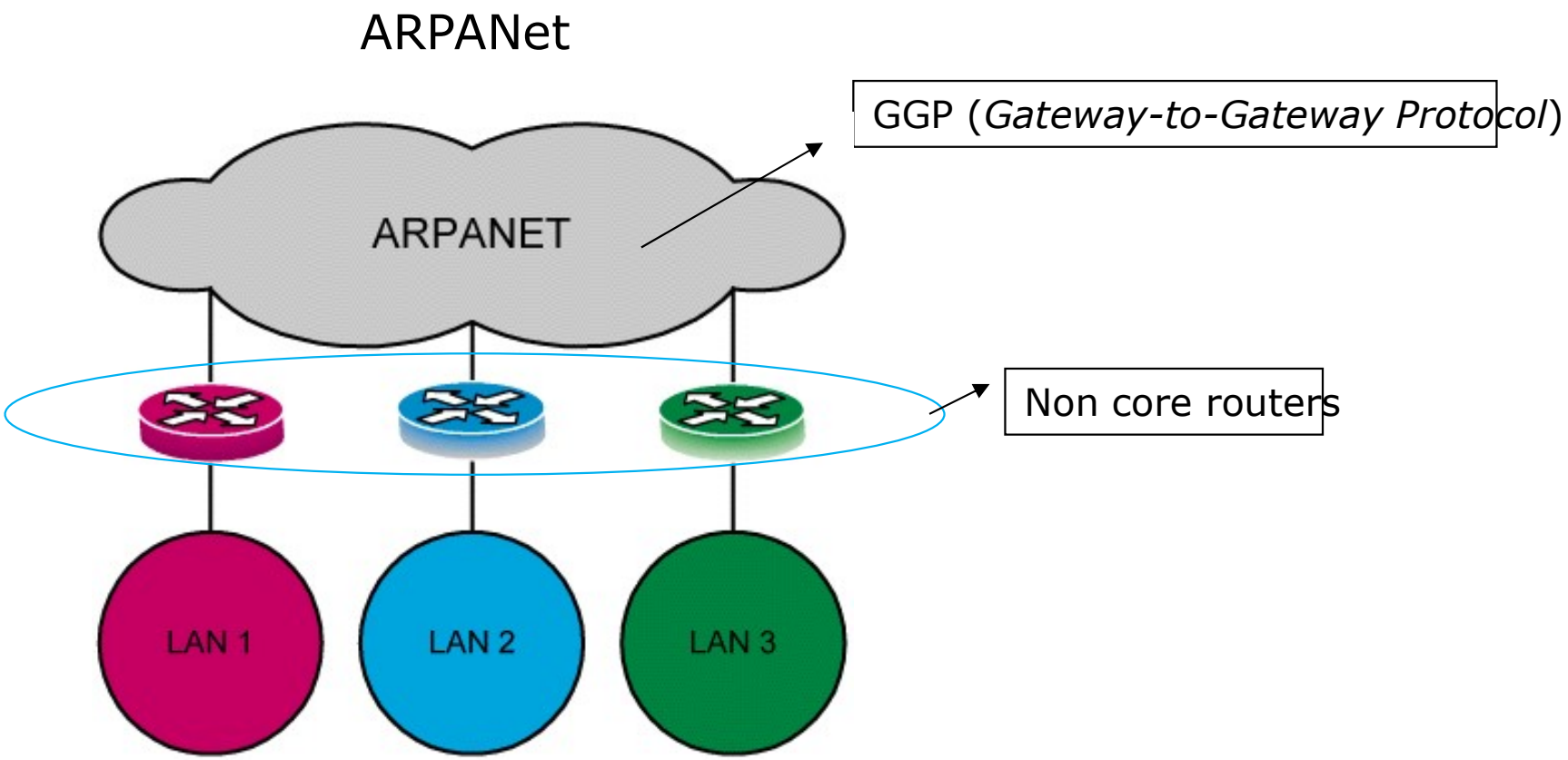
- Database Description
  - Fornece os números de sequencia de todas as entradas de estado de enlace mantidas pelo transmissor (Banco de Dados de Adjacencias)
  - Receptor desta mensagem pode utilizar sequencias mais novas para atualizar sua própria base
- LINK STATE REQUEST
  - Um roteador pode solicitar a outro o seu estado de enlace
  - Permite que roteadores adjacentes se atualizem
  - Após atualização informações são enviadas para toda a área
- Resumo
  - A apartir destas mensagens, roteadores de cada área constroem seu grafo;
  - Roteadores de backbone aceitam informações de roteadores de borda de área e vice-versa

Histórico
Roteamento na Internet

- Arquitetura proposta na época
  - Um conjunto reduzido e centralizado de roteadores no núcleo da rede (*Core routers*)
    - Mantinham rotas para todos os possíveis destinos na internet
    - Administrados pelo INOC (*Internert network Operation Center*)
    - Desenvolvimento do protocolo GGP (*Gateway-to-Gateway Protocol*) para atualização automática das tabelas de rotas
      - Baseado no algoritmo vetor-distância (*Bellman Ford*)
  - Um conjunto maior de roteadores (*Non cores routers*) com rotas parciais
    - Administrados pelas instituições de pesquisa

Histórico
Roteamento na Internet

- Arquitetura proposta na época



Histórico
Roteamento na Internet

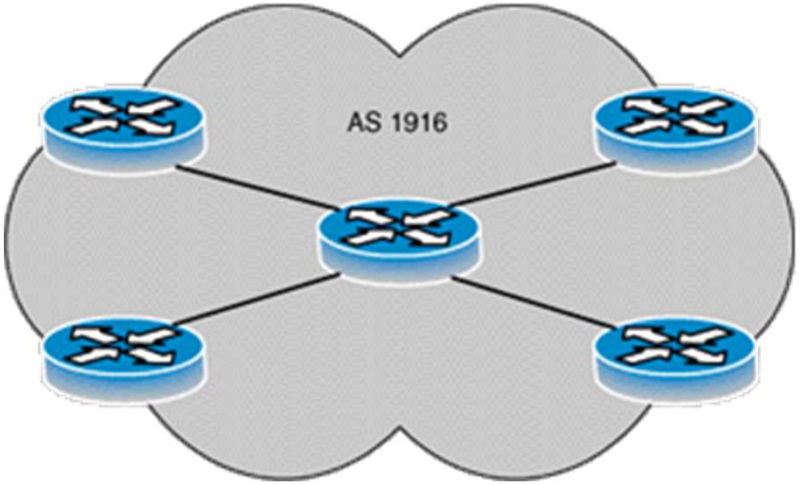
- Limitações da arquitetura proposta na época
  - *Backbone* complexo de cada site (instituição)
    - Impossibilidade de conectar diretamente as redes
  - Levava em conta apenas interligação com apenas uma internet
  - Não contemplava questões administrativas
- Solução
  - Desenvolver um mecanismo para possibilitar a comunicação com o "mundo exterior"
  - Nasce e firma-se o conceito de AS
  - Desenvolvimento do EGP (Que também mostrou-se limitado)
  - Desenvolvimento do BGP em substituição ao EGP

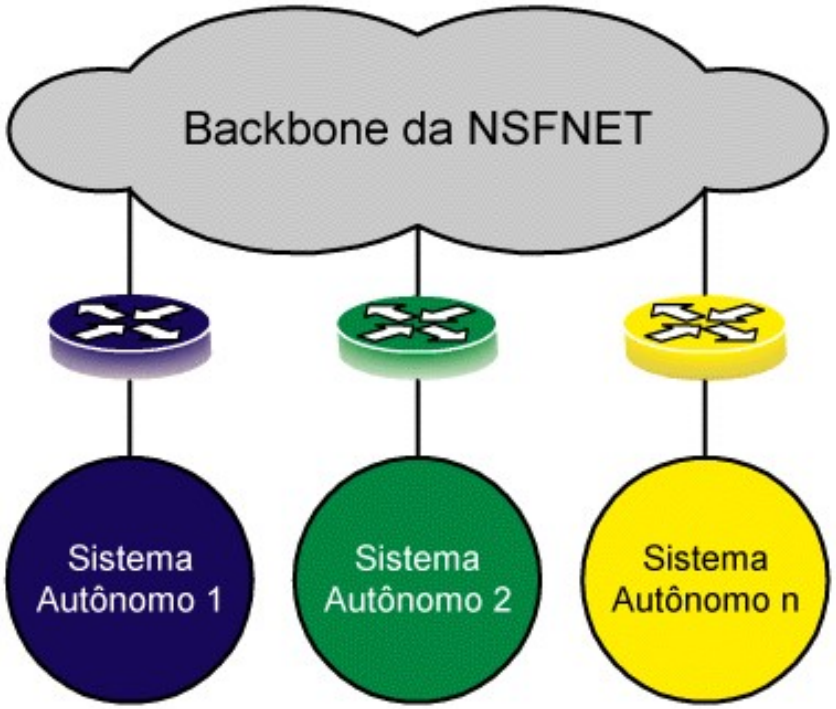
Histórico
Roteamento na Internet

- Sistema Autônomo (*Autonomous System*)
  - Um conjunto de redes e roteadores controlados por uma única autoridade administrativa
- Segundo a RFC 1930 (Definição formal)
  - Um conjunto de roteadores controlados por uma **única administração técnica**, usando um **protocolo interior e métricas comuns** para rotear pacotes dentro do **AS**, e usando um **protocolo exterior** para rotear pacotes para **outros ASs**.
  - Requisito básico: uma política de roteamento única
  - A política de roteamento define como são tomadas as decisões de roteamento na internet.

Histórico
Roteamento na Internet

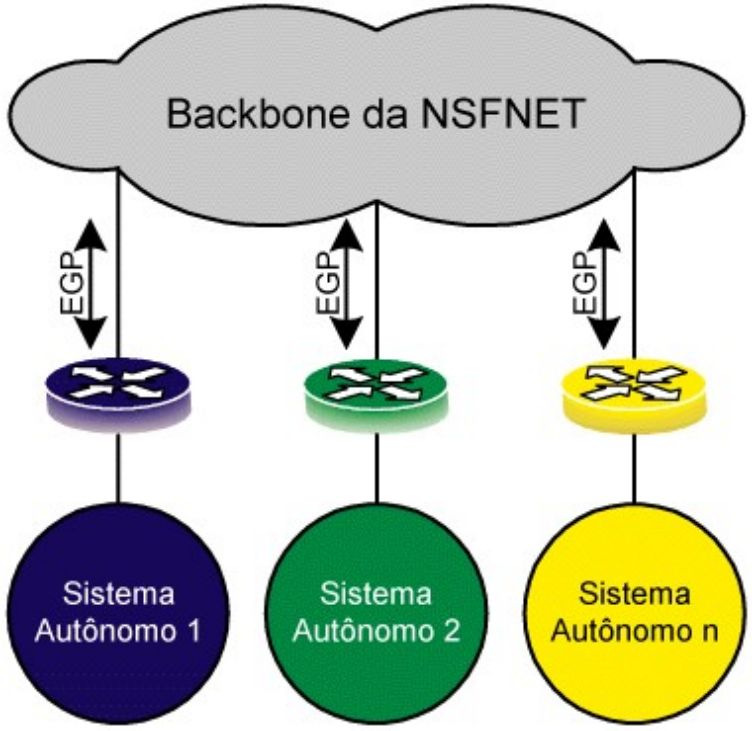
- Conjunto de redes compartilhando a mesma política
- Utilizam um único protocolo de roteamento
- Estão sob a mesma administração técnica
  - Exemplos de ASs





Histórico
Roteamento na Internet

- O EGP apresentou deficiências insustentáveis, como restrições em topologia, incapacidade de evitar "loops" de roteamento e pouca flexibilidade para a configuração de políticas de roteamento
- Um grande desafio para os projetistas era a solução de como transformar uma arquitetura internet para não depender de um sistema centralizado (*core routers*) - deixando uma topologia organizada hierarquicamente e iniciando outra, com diferente estrutura



Histórico
Roteamento na Internet

- Solução:
  - Roteadores utilizados para trocar informações dentro de um **AS**
    - Roteadores interiores (*Internal Routers*)
      - Utilizam algum protocolo IGP (*Interior Gateway Protocol*)
        - RIP, OSPF, IS-IS, IGRP, EIGRP
  - Roteadores utilizados para trocar informações entre **ASs**
    - Roteadores exteriores (*External Routers*)
      - Utilizam algum protocolo EGP (*External Gateway Protocol*)
        - BGP-4 (RFC 4271)
      - Consideram blocos CIDR (Super Redes)

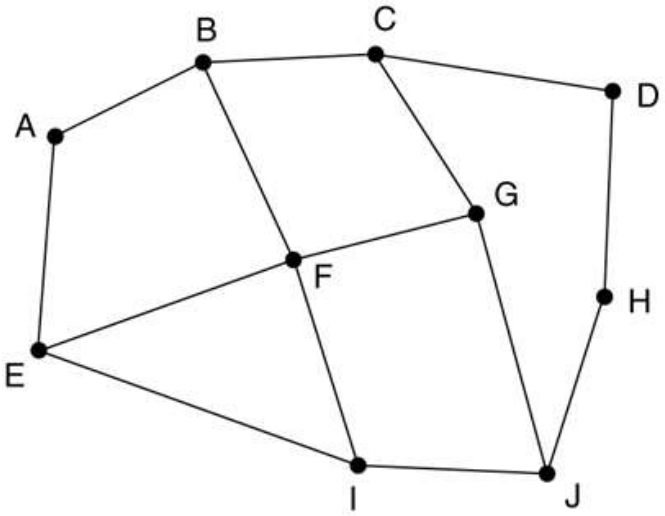
Border Gateway Protocol (BGP)

- Responsável pelo roteamento entre ASes (Exterior Gateway Protocol)
- Protocolo entre ASes deve preocupar-se com a política
  - Um SA pode não estar interessando em transportar tráfego de controle entre dois outros SAs vizinhos a ele
  - Por outro lado ele pode interessar-se por este tráfego caso um deles tenha pago por isto
  - Um operador de telecomunicações interesse pelo transporte de dados de seus clientes, mas não de outros
- Protocolo projetado com o objetivo expressar vários tipos de políticas de roteamento no trafego entre SAs
- Exemplos de Políticas
  - Nenhum trafego de trânsito deve passar por certos SAs
  - Nunca coloque o lêmen em uma rota que comece no Pentágono
  - Não usar os Portugal para ir da França até a Alemanha
  - O tráfego que começar ou terminar na Apple não deve transitar pela Microsoft
- Políticas não fazem parte diretamente do protocolo
  - São configuradas manualmente em cada roteador BGP (ou incluídas utilizando algum tipo de script)

Border Gateway Protocol (BGP)

- Para um roteador BGP o que importa são os ASes e linhas que os conectam
- Linhas possuem três categorias
  - Stub – possuem apenas uma conexão com o grafo BGP e não podem ser utilizadas para tráfego entre ASes. Suporta apenas tráfego local
  - *Multihome* – Estão conectadas a mais de um AS e podem ser utilizadas para tráfego entre ASes a menos que recusem
  - Trânsito – Desejam trafegar pacotes de terceiros, podendo possuir alguma restrição e normalmente receber por isto

Border Gateway Protocol (BGP)



Informação que F recebe de seus vizinhos sobre D

De B: utilizo "BCD"
De G: utilizo "GCD"
De I: utilizo "IFGCD"
De E: utilizo "EFGCD"

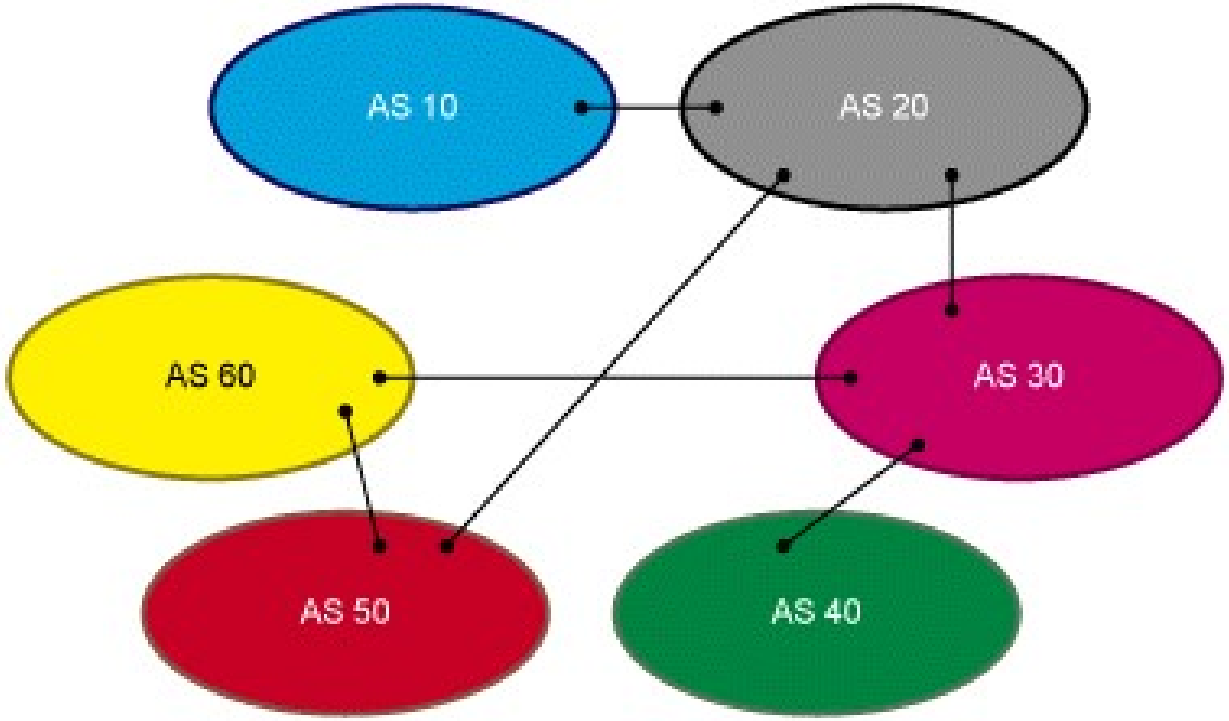
- BGP e fundamentalmente um protocolo de vetor de distância
- Cada roteador mantém qual caminho está sendo utilizado
- E informa a seus vizinhos o caminho exato que utiliza para chegar a outros pontos
- Exemplo
  - Suponha que F que chegar a D
  - Após receber o caminho de seus vizinhos descarta I e E (pois passam por F)
  - Nas rotas que sobram o roteador calcula a “distância” levando em conta as políticas e utiliza a menor distância
    - A função de contagem não é definida pelo protocolo
  - Suponha que escolha G (FGCD)
  - Caso G falhe ao receber informações dos três vizinhos restantes, escolheria então “FBCD”

Border Gateway Protocol – BGP4

- RFC 4271 (http://www6.ietf.org/rfc/rfc4271)
- Roteamento entre ASs
- Suporte a Super-redes (CIDR – *Classless Interdomain Router*)
- Interage com IGPs: RIP, OSPF, etc.
- Usa TCP porta 179
- Estabelece sessões BGP
  - Estabelecimento de conexão TCP entre os roteadores
  - Envio de tabela de rotas completas apenas uma vez
  - Atualização parcial da tabela (Incremental)
  - Mensagens de **KEEPALIVE** para manter a sessão

Border Gateway Protocol – BGP-4

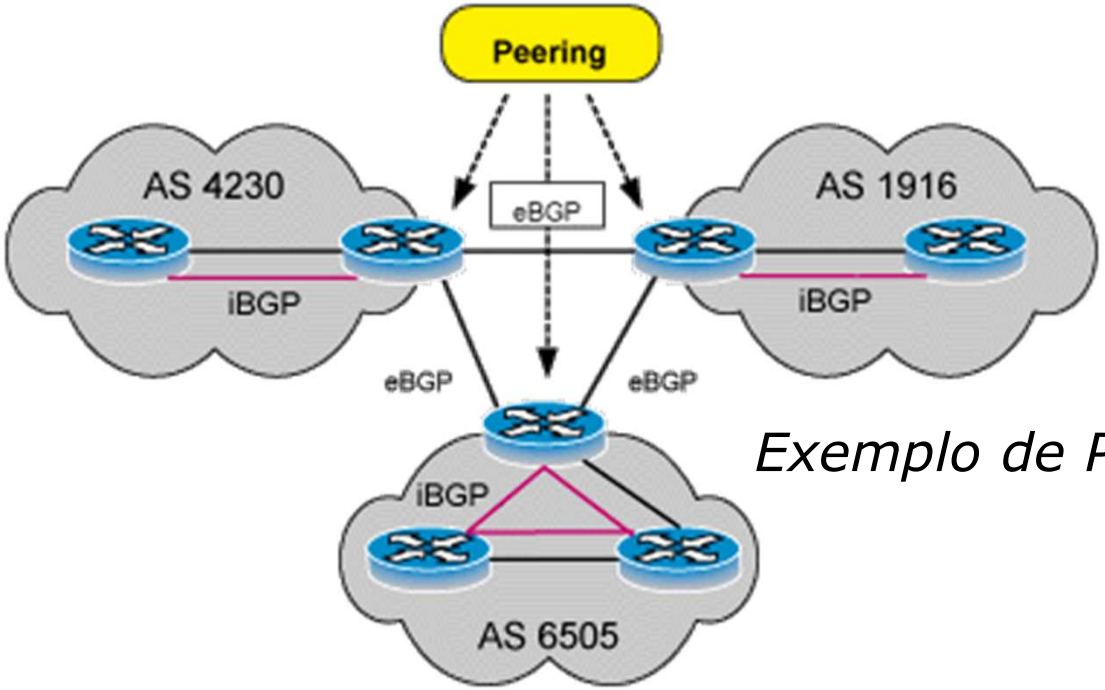
- Mensagens de aviso são enviadas quando ocorrem erros ou outras situações especiais;
- Caso uma conexão verifique um erro, uma mensagem é enviada e a conexão fechada, encerrando a sessão



A atual arquitetura da Internet, onde ASes comunicam-se via BGP-4

Border Gateway Protocol – BGP-4

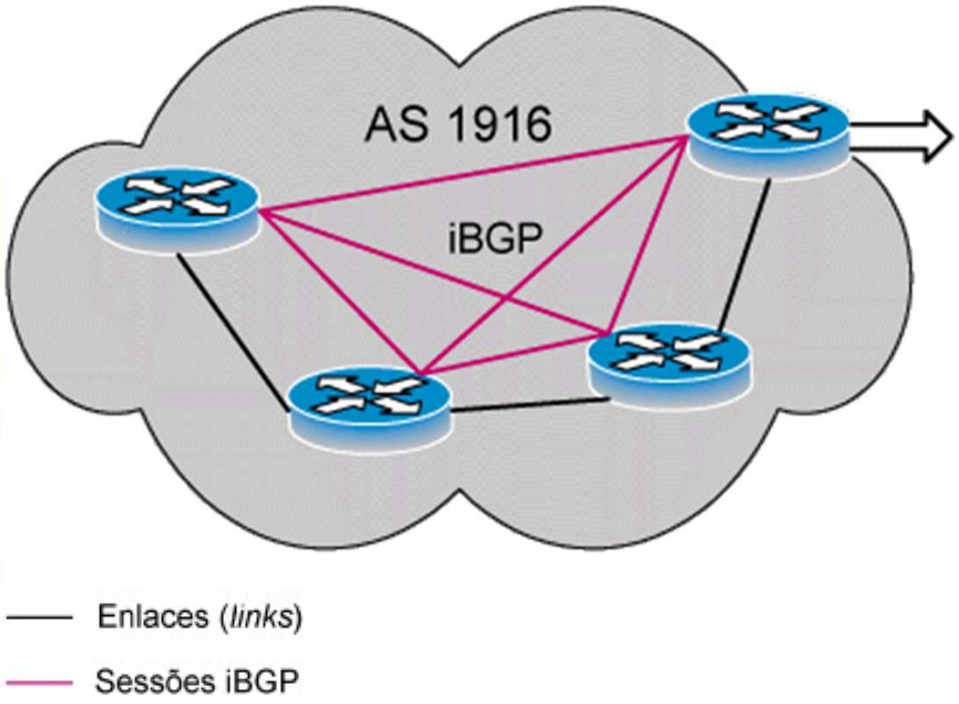
- Neighbors, Peers, eBGP e iBGP
  - Sistemas (roteadores) que são "vizinhos BGP" (BGP *neighbors*) comunicam-se através de "sessões" estabelecidas entre eles
  - Os roteadores de "borda" (*border routers*) de ASes vizinhos são considerados *peers*
    - *Esses peers são as "fronteiras políticas" dos ASes, que trocam tráfego de acordo com as regras definidas pelos ASes participantes*



Exemplo de Peers, Neighbors, eBGP e iBGP

Border Gateway Protocol – BGP-4

- O algoritmo do eBGP trabalha, basicamente, anunciando todas rotas que conhece, enquanto o do iBGP faz o possível para não anunciar rotas

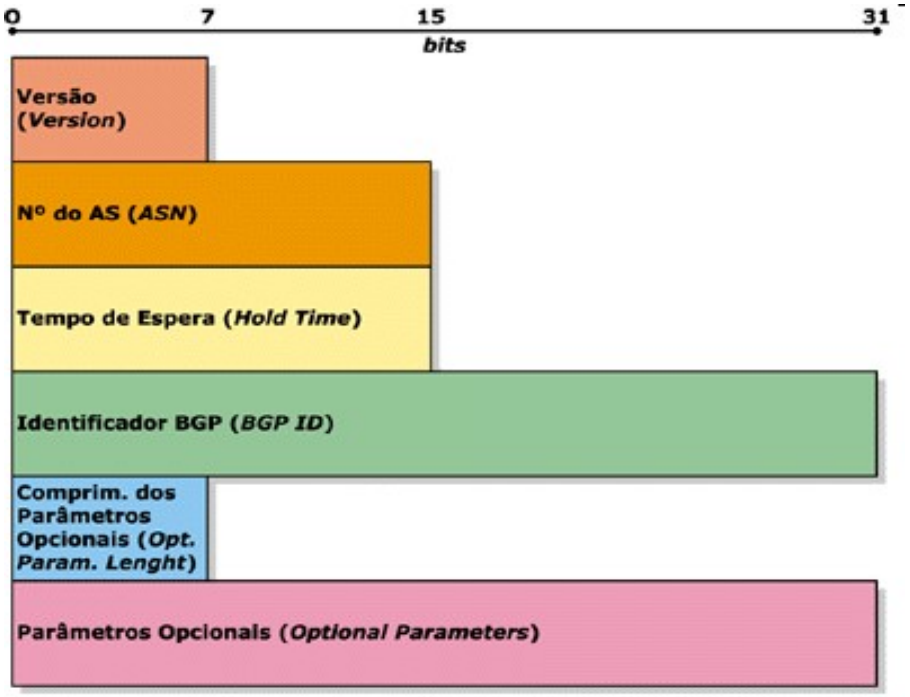


para fazer o iBGP funcionar adequadamente dentro de um AS é necessário estabelecer sessões BGP entre todos os roteadores que "falam" iBGP, formando uma "malha completa" (*full mesh*) de sessões iBGP dentro do AS

BGP4
MENSAGENS

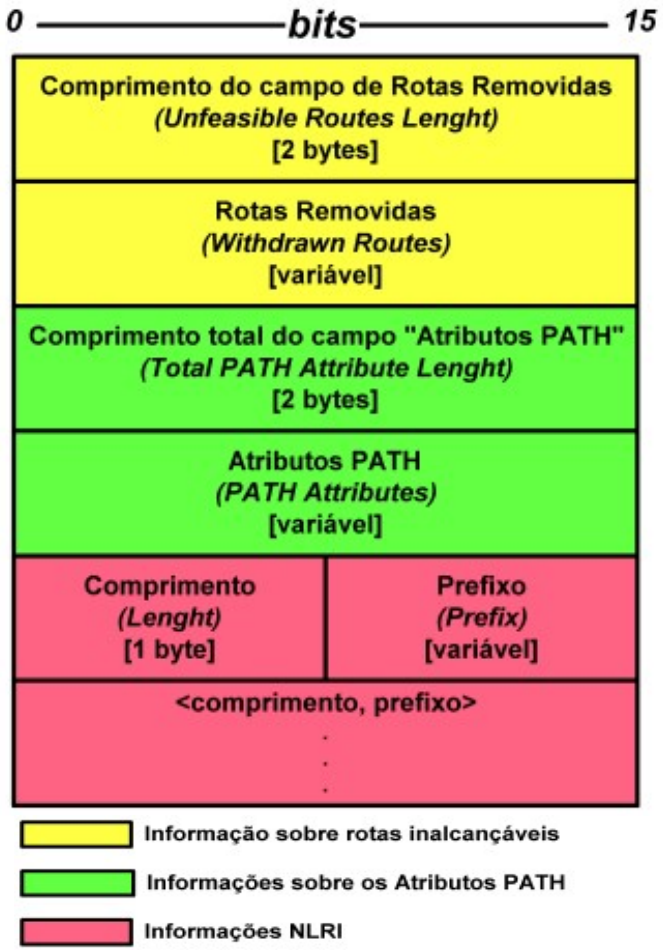
- Mensagens trocadas em sessões BGP
  - Todas as mensagens são compostas de, no mínimo, um cabeçalho e, opcionalmente, uma parte de dados.
  - Campo Marcador (*Marker*)
    - Serve para verificar a autenticidade da mensagem recebida e se houve perda de sincronização entre os roteadores vizinhos BGP
  - Campo Comprimento (*Length*)
  - Campo Tipo (*Type*)
    - Deve conter um número que representa o código de um tipo de primitiva
      - **OPEN, UPDATE, NOTIFICATION, KEEPALIVE**

BGP4 OPEN



- Enviada para se iniciar a abertura de uma sessão BGP entre *neighbors* ou *peers* BGP
  - Versão (Version) – 3 ou 4
  - Número do AS (AS Number)
    - Deve conter o número do AS a qual o roteador pertence
  - Tempo de espera (Hold Time)
    - Valor, em segundos, do maior tempo de espera (hold time) permitido entre mensagens do tipo UPDATE ou KEEPALIVE
  - Identificador BGP
    - Normalmente o IP do roteador
  - Comprimento dos Parâmetros Opcionais (Optional Parameters Length)
  - Parâmetros Opcionais

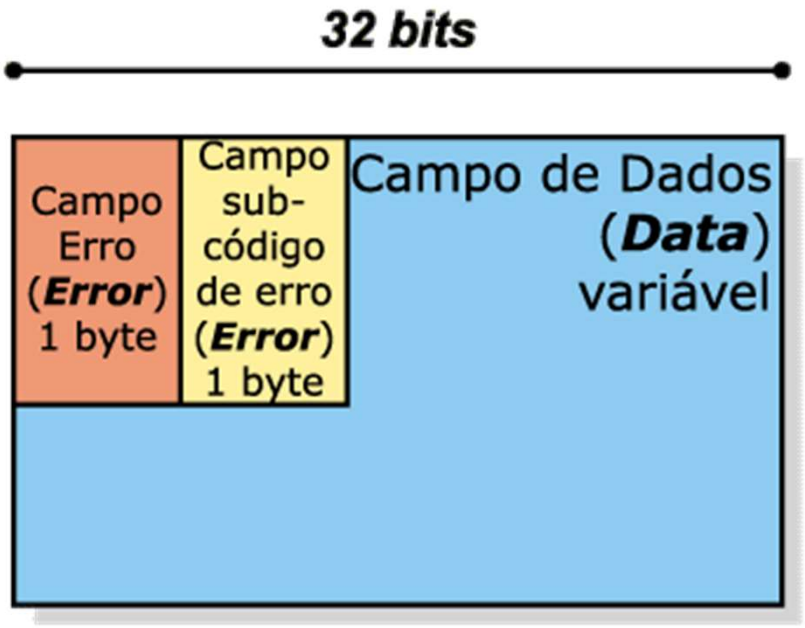
BPG4 UPDATE



- Comprimento das Rotas Removidas ou Inalcançáveis (Unfeasible Routes Length)
  - Neste campo, é indicado o comprimento total, em bytes, do total de rotas removidas
- Rotas Removidas (Withdrawn Routes)
  - Este campo inclui uma lista de prefixos de endereços para rotas que devem ser removidas da tabela de rotas BGP (CIDR)
- Comprimento Total do Atributo PATH (Total Path Attribute Length)
  - Deve indicar o comprimento total, em bits, do campo Atributos PATH. O valor contido neste campo deve permitir a determinação do comprimento do campo *Network Layer Reachability Information* (NLRI)
- Atributos PATH (PATH Attributes)
  - Composto de um
- Comprimento (Length)
  - Deve indicar o comprimento total, em bits, do total de rotas removidas. Um comprimento igual a 0 (zero), indica que, nesta mensagem UPDATE, não há rotas a serem removidas.
- Prefixo - (Prefix)
  - Contém prefixos de endereços IP seguidos de bits suficientes para fazer o final deste campo terminar "arredondado" em bytes completos. O valor dos bits complementares não têm importância
- <Comprimento, Prefixo>
  - Informações que representam as redes alcançáveis
    - Comprimento – Máscara CIDR
    - Prefixo – Endereço da subrede
  - Exemplo: /25, 204.149.16.128; /23, 206.134.32; /8, 10

BPG4 NOTIFICATION

- Este tipo de mensagem é enviada no caso de detecção de erros durante ou após o estabelecimento de uma sessão BGP

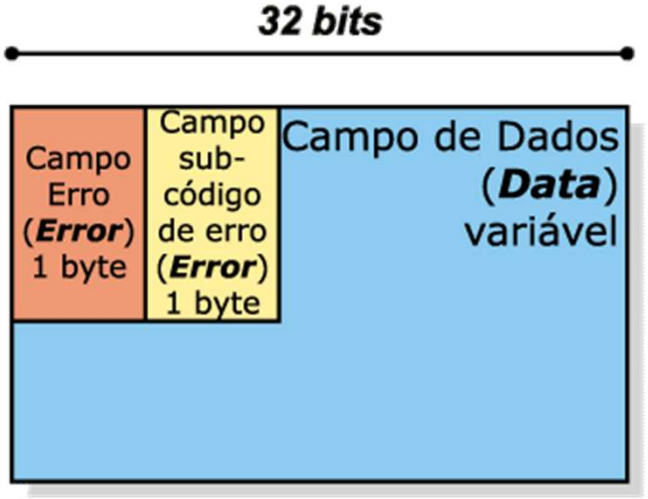


- **Campo Erro (*Error*)**
  - Deve conter o tipo da notificação
- **Campo Sub Código de Erro (*Error subcode*)**
  - Deve conter um valor que fornece maiores informações sobre o erro
- **Campo de Dados (*Data*)**
  - Pode conter dados referentes ao erro, como por exemplo, um cabeçalho mal formado (inválido), um número de AS inválido.

BPG4 NOTIFICATION

- Tabela de códigos e subcódigos

Códigos de Erro	**Sub códigos de Erro**
1 - Erro no cabeçalho da mensagem	1 - Conexão não sincronizada 2 - Comprimento da mensagem inválido 3 - Tipo de mensagem inválido
2 - Erro na mensagem OPEN	1 - Número de versão não suportado 2 - Número de AS vizinho inválido 3 - Identificador BGP inválido 4 - Parâmetro opcional não suportado 5 - Falha na autenticação 6 - Tempo de espera inaceitável
3 - Erro na mensagem UPDATE	1 - Lista de atributos mal formada 2 - Atributo *Well-Known* desconhecido 3 - Atributo *Well-Known* faltando 4 - Erro nas *flags* de atributos 5 - Erro no comprimento do atributo 6 - Atributo origem inválido 7 - *Loop* de roteamento em AS 8 - Atributo NEXT_HOP inválido 9 - Erro no atributo Opcional 10 - Campo de rede inválido 11 - *AS_path* mal formado



BPG4 KEEPALIVE

- São mensagens trocadas periodicamente com o propósito de verificar se a comunicação entre os vizinhos está ativa
- A mensagem do tipo KEEPALIVE é composta apenas do cabeçalho padrão das mensagens BGP, sem dados transmitidos após o cabeçalho.
- O tempo máximo permitido para o recebimento de mensagens KEEPALIVE ou UPDATE é definido pelo *hold time*, das mensagens OPEN
- Para manter aberta a sessão, a mensagem de KEEPALIVE deve ser enviada antes que o prazo definido no *hold time* expire