Resumo Ragnarok: Camada de Rede

**Introdução**

**Internet**: Conjunto de redes de escala mundial, ligadas pelo protocolo IP.

**TCP/IP:** Família de protocolos de comunicação, com serviços e acesso independente de tecnologia. Permite a interconexão de redes físicas diferentes. Interconexão realizada por roteadores.

**Protocolo IP:** Não orientado a conexão, roteamento melhor esf orço. Não confiável, sem controle de fluxo e de erros e é simples. Roteamento baseado no endereço da rede de destino.

A camada de rede é responsável por prover conectividade ao computador e selecionar caminhos para que os pacotes de dados possam trafegar. Isso é feito através do roteamento. Um motivo para que a conectividade possa ocorrer tem relação com o IP (Internet Protocol). O endereçamento IP permite que o dispositivo possa se conectar em uma mesma rede ou em redes diferentes.

**Camada de Rede: Objetivos**

* Transferência de pacotes da origem para o destino;
* Vários saltos (hops) intermediários no caminho;
* **Elementos de rede:** Roteadores ou Comutadores;
* **Principal função:** Roteamento (encaminhamento);
* **Outras funções:** Controle de Congestionamento, Negociação de QoS e Interconexão de redes.

**Camada de Rede: Funcionamento:**

* Transporta segmentos do host transmissor para o receptor;
* No lado transmissor, encapsula os segmentos em datagramas;
* No lado receptor, entrega os segmentos à camada de transporte;
* Protocolos da camada de rede em cada host roteador;
* Roteador examina campos de cabeçalho em todos os datagramas IP que passam por ele;

**Roteamento:** Estabelecimento dos melhores caminhos (rotas).

**Encaminhamento:** Processo de despachar cada pacote ao seu destino ou sistema intermediário.

**Serviços oferecidos à camada de Transporte**:

* Devem ser independentes da tecnologia da sub-rede.
* Proteção contra tipo, quantidade e topologia das sub-redes.
* Endereços devem ter plano de numeração único.
* **Tipos de Serviço:** Orientado a conexão e Sem conexão.

Circuitos Virtuais x Datagramas

**Circuitos Virtuais (CVs)**

* Analogia aos circuitos físicos da rede telefônica;
* Rede complexa e segura;
* Geralmente orientada a conexões (conexões na camada de rede são geralmente chamadas de circuitos virtuais);
* Usado para estabelecer, manter e encerrar conexões;
* Usados em ATM, frame-relay e X-25;
* Não é usado na Internet atualmente.

**Datagramas**

* Analogia com o serviço de “telegramas”;
* Rede simples e não confiável;
* Geralmente é não orientada a conexões (mas pode ser);
* Não é estabelecida conexão na camada de rede;
* Roteadores: sem estado sobre conexões fim-a-fim;
* O conceito “conexão” não existe na camada de rede;
* Pacotes são encaminhados para o endereço do host de destino;
* Pacotes para o mesmo destino podem seguir diferentes rotas.

|  |  |  |
| --- | --- | --- |
| **Questão** | **Datagrama** | **Circuito Virtual (VC)** |
| Configuração de Circuito | Não necessária | Requerida |
| Endereçamento | Cada pacote contém endereços da fonte e do destino | Cada pacote contém um número de CV |
| Informação de Estado | Subrede não mantém informação de estado | Cada CV requer espaço na tabela da subrede |
| Roteamento | Cada pacote é roteado independentemente | Rota escolhida quando CV é configurado; todos os pacotes seguem esta rota |
| Efeito de falhas do roteador | Nenhuma, exceto a perda de pacotes durante a falha | Todos os CVs que passam pelo roteador com falha são Terminados |
| Controle de congestionamento | Difícil | Fácil se memória suficiente tiver sido alocada a priori para cada CV |

Roteamento

**Algoritmo de roteamento:**

* Parte da camada de rede responsável pela decisão sobre a linha de saída para a transmissão de um pacote;
* Pode ser implementado por um protocolo de roteamento, ou executado de maneira estática.

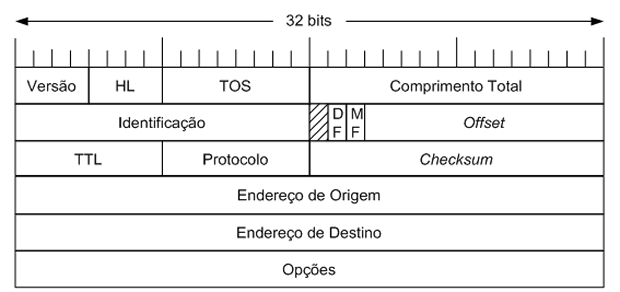
**Protocolo de roteamento:**

* Software utilizado pelos roteadores para que eles estabeleçam tabelas de roteamento consistentes.
* Qualquer protocolo de roteamento deve comunicar informação da topologia da rede para todos os outros roteadores, para tomar decisões de roteamento.

**Datagramas:** Decisão de roteamento deve ser tomada a cada pacote. **Circuitos Virtuais:** Decisão de roteamento é tomada somente no estabelecimento da conexão e também chamado de roteamento por sessão.

Existem duas classes de tipos de roteamento: **estático e dinâmico.** Não adaptativo (**estático**): Decisão do roteamento não é baseada em estimativas de tráfego atual e da topologia. Adaptativo (**dinâmico**): Muda decisões de roteamento para refletir mudanças na topologia, bem como, no tráfego. Leva em consideração: custo, caminho, carga e tamanho da fila.

Cabeçalho IP



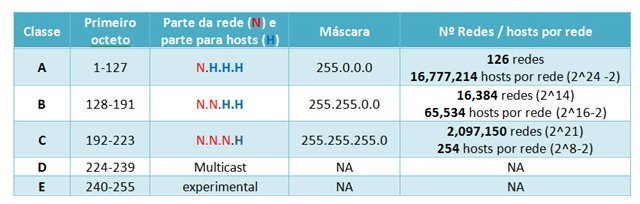
**Fragmentação e remontagem**

* MTU - Maximum Transmit Unit é o tamanho máximo que um pacote pode ter;
* MTU é padronizado de acordo com a interface física;
* Cada roteador deve fragmentar o pacote antes de encaminhá-lo para uma interface, no caso do tamanho original ser maior que o MTU;
* Cada fragmento pertence a um mesmo pacote original (mesmo valor no campo Fragmentation Identifier)

**Processo:**

1. O primeiro bit do campo "Flags" deve ser 0 (reservado);
2. Fragmentos não podem novamente ser fragmentados, logo "May Fragment" deve ser 0;
3. O flag "More Fragments" indica se existe mais fragmentos adiante;
4. 1o. fragmento: "Fragment Offset" = 0, tam=1500 (20+1480) bytes;
5. 2o. fragmento: "Fragment Offset" = 185 (1480/8), tam=1500;
6. 3o. fragmento: "Fragment Offset" = 370 ((1480×2)/8), tam=1500;
7. 4o. fragmento: "Fragment Offset" = 555 ((1480×3)/8), tam=24 (20+4).

Endereçamento IP



**Sub-redes privadas – Não podem ser publicados na Internet:**

* 10.x.x.x *(classe A)*
* 172.16.x.x - 172.31.x.x *(classe B)*
* 192.168.x.x *(classe C)*

**Endereços especiais**

* 0.0.0.0 este host
* 0.0.0.124 host 124 nesta rede
* 255.255.255.255 todos os hosts desta rede
* N.N.N.255 todos os hosts da rede N.N.N
* 127.X.X.X Loopback

**Alguns endereços inválidos para hosts**

* 10.1.0.0 IP do host não pode ser 0
* 10.1.0.255 IP do host não pode ser 255
* 10.123.255.4 Subrede não pode ter valor 255
* 0.12.16.89 Parte do endereço não pode ter valor 0
* 255.9.56.45 Parte do endereço não pode ter valor 255
* 10.34.255.1 Parte do endereço não pode ter valor 255

Máscara de sub-rede

* A máscara serve para indicar qual parte do endereço IP identifica o endereço de rede e qual parte identifica o endereço do host
* Os bits 1 indicam a parte do endereço da rede
* Os bits 0 indicam a parte do endereço do host

**Máscaras default:**

Classe A - 255.0.0.0 11111111.00000000.00000000.00000000

Classe B - 255.255.0.0 11111111.11111111.00000000.00000000

Classe C - 255.255.255.0 11111111.11111111.11111111.00000000

**Processo:**

1. Executa-se um AND lógico entre os bits da máscara e endereço IP e obtém o Network Address
2. No endereço da rede todos os bits do host são 0
3. No endereço de broadcast, todos os bits do host são 1

**Exemplo:**

10001100.10110011.11110000.11001000 140.179.240.200 Endereço IP 11111111.11111111.00000000.00000000 255.255.000.000 Máscara classe B ------------------------------------------------------------------------------------------------

10001100.10110011.00000000.00000000 140.179.000.000 Network Address 10001100.10110011.11111111.11111111 140.179.255.255 Network Broadcast

CIDR: Classless InterDomain Routing

* A porção de endereço de rede tem tamanho arbitrário
* Formato do endereço: a.B.C.D/x, em que x é o número de bits na parte de rede do endereço
* Inicialmente somente grupos de endereços Classe C foram utilizados
* Como o mesmo procedimento já foi também aplicado às antigas classes A e B, pode-se dizer que de fato o endereçamento em classes está descaracterizado e completamente substituído pelo CIDR (obs.: vários endereços de antigas classes A e B foram progressivamente realocados)

Como um sistema final pode obter um endereço IP?

* Configuração estática
* DHCP: Dynamic Host Configuration Protocol: obtém dinamicamente endereços IP de um servidor
  + “plug-and-play”

**Como uma rede obtém a parte de sub-rede do endereço IP ?**

* Obtém a porção alocada no espaço de endereço do seu provedor ISP (ICANN: internet corporation for assigned names and numbers)
  + Aloca endereços, gerência DNS, atribui nomes de domínios...

DHCP

É um protocolo utilizado em redes de computadores que permite às máquinas obterem um endereço IP automaticamente.

Por meio dele um servidor é capaz de distribuir automaticamente endereços de IP diferentes a todos os computadores à medida que eles fazem a solicitação de conexão com a rede. Essa distribuição dos IPs é feita em um intervalo pré-definido configurado no servidor. Sempre que uma das máquinas for desconectada o IP ficará livre para o uso em outra.

**Funcionamento:**

* Quando um cliente conecta-se a uma rede ele envia um pacote com um pedido de configurações DHCP.
* O servidor DHCP gerencia uma faixa fixa de IPs disponíveis juntamente com as informações e parâmetros necessários (gateway padrão, nome de domínio, DNS, etc).
* Quando este servidor recebe um pedido, ele entrega um destes endereços e configurações para o cliente.

**Segurança:**

* Mensagens não são autenticadas
  + Alguém pode forjar um DHCP server ou um cliente
* Cliente não pode confiar no servidor e vice-versa

**Configuração**:

* Para redes com mais de um servidor
  + Servidores na rede não podem trocar informações
  + Não existe um protocolo server-server
  + Devem ter espaços de endereçamento disjuntos para evitar distribuição de IP duplicados
* Servidores são configurados manualmente

**Modos de Operação:**

* **Automática**, no qual uma quantidade de endereços de IP (dentro de uma faixa) é definida para ser utilizada na rede. Neste caso, sempre que um dos computadores de uma rede solicitar a conexão com ela, um destes IPs será designado para a máquina em questão.
* Na **dinâmica** o procedimento é bem parecido com o efetuado pela automática, porém a conexão do computador com determinado IP é limitada por um período de tempo pré-configurado que pode variar conforme desejado pelo administrador da rede.
* No modo **manual** o DHCP aloca um endereço de IP conforme o valor de MAC (Medium Access Control) de cada placa de rede de forma que cada computador utilizará apenas este endereço de IP. Utiliza-se este recurso quando é necessário que uma máquina possua um endereço de IP fixo.

NAT

**NAT (Network Address Translation)** é uma técnica que consiste em reescrever os endereços IP de origem de um pacote que passam por um routes ou firewall de maneira que um computador de uma redes interna tenha acesso ao exterior (internet).

A motivação para a criação do NAT foi a necessidade das redes internas/privadas se comunicarem com a internet/redes públicas. Como os IPs internos, como 10.0.0.0/8, não podem ser passados para a Internet, pois não são roteador nela, a solução foi fazer um mapeamento baseado no IP interno e na porta local do computador. Com esses dois dados, o NAT gera um número de 16 bits usando a tabela hash, este número é então escrito no campo da porta de origem.

**Motivação: redes locais podem utilizar apenas um endereço IP:**

* Não é preciso alocar uma gama de endereços do ISP, apenas um endereço IP é usado para todos os dispositivos;
* Podem-se alterar os endereços dos dispositivos na rede local sem precisar notificar o mundo exterior;
* Pode-se mudar de ISP sem alterar os endereços dos dispositivos na rede local;
* Dispositivos da rede local não são explicitamente endereçáveis ou visíveis pelo mundo exterior (um adicional de segurança).

**Implementação: o roteador NAT deve:**

* Tratar datagramas que saem: substituir (endereço IP de origem, porta #) de cada datagrama para (endereço IP do NAT, nova porta #)
  + . . . clientes/servidores remotos responderão usando (endereço IP do NAT, nova porta #) como endereço de destino
* Lembrar (na tabela de tradução do NAT) cada (endereço IP de origem, porta #) para o par de tradução (endereço IP do NAT, nova porta #).
* Tratar datagramas que chegam: substituir (endereço IP do NAT, nova porta #) nos campos de destino de cada datagrama pelos correspondentes (endereço IP de origem, porta #) armazenados da tabela NAT

**Campo número de porta com 16 bits:**

* Possibilidade de 65536 conexões simultâneas com um único endereço de LAN!!!

**NAT é controverso:**

* Roteadores deveriam processar somente até a camada de rede (Layer 3): Violação do argumento fim-a-fim
* A possibilidade de NAT deve ser levada em conta pelos desenvolvedores de aplicações: ex.: aplicações P2P ß A escassez de endereços deveria ser resolvida pelo IPv6

IPV6

**Motivação inicial:**

* O espaço de endereços de 32 bits está próximo de ser completamente alocado.

**Motivação adicional:**

* Melhorar o formato do cabeçalho para permitir maior velocidade de processamento e de transmissão;
* Mudanças no cabeçalho para incorporar mecanismos de controle de serviço (i.e., Quality-of-Service, QoS).

**Formato do datagrama IPv6:**

* Cabeçalho fixo de 40 Bytes
* Não é permitida fragmentação (a fragmentação e remontagem tomam muito tempo, retirando essa funcionalidade dos roteadores acelera o repasse de datagramas IP)

**Classe de tráfego:** permitir definir prioridades diferenciadas para vários fluxos de informação. **Rótulo de Fluxo:** identifica datagramas do mesmo “fluxo” (conceito de “fluxo” não é bem definido). **Próximo Cabeçalho:** identifica o protocolo da camada superior ou um header

**Diferenças com o IPV4:**

1. Formato do datagrama IPv6: Cabeçalho fixo de 40 Bytes
2. Não é permitida fragmentação (a fragmentação e remontagem tomam muito tempo, retirando essa funcionalidade dos roteadores acelera o repasse de datagramas IP)
3. Classe de tráfego: permitir definir prioridades diferenciadas para vários fluxos de informação Rótulo de Fluxo: identifica datagramas do mesmo “fluxo” (conceito de “fluxo” não é bem definido).
4. Próximo Cabeçalho: identifica o protocolo da camada superior ou um header
5. Endereços de 128 bits (i.e., 16 Bytes)
6. Checksum: removido inteiramente para reduzir o tempo de processamento em cada salto
7. Options: são permitidas, mas são alocadas em cabeçalhos suplementares, indicados pelo campo “Next header”
8. ICMPv6: nova versão de ICMP
9. Tipos de mensagens adicionais , ex.: “Packet Too Big”
10. Funções de gerenciamento de grupos multicast

Encaminhamento

Encaminhamento é o processo de despacho de um pacote ao seu destino ou sistema intermediário. Envolve a transferência de um pacote de um enlace de entrada para um enlace de saída dentro de um roteador.

Roteamento

Envolve todos os roteadores de uma rede, cujas interações por meio de protocolos de roteamento determinam as rotas que os pacotes devem seguir

**O roteamento IP consiste em duas fases:**

1. Como repassar um pacote de uma interface de entrada para uma interface de saída de um roteador (encaminhamento de pacotes).

2. Como localizar e configurar uma rota?

Existem duas classes de tipos de roteamento: **estático e dinâmico.**

* Não adaptativo (**estático**): Decisão do roteamento não é baseada em estimativas de tráfego atual e da topologia;
* Adaptativo (**dinâmico**): Muda decisões de roteamento para refletir mudanças na topologia, bem como, no tráfego. Leva em consideração: custo, caminho, carga e tamanho da fila.

**Sistemas Autônomos**

Um sistema autônomo (AS) é uma região da internet que é administrada por uma entidade única(tipo a RNP).

Roteamento é feito de maneira diferente se estamos dentro de um sistema autônomo (**roteamento intradomínio**) ou se devemos permitir que sistemas autônomos se comuniquem (**roteamento interdomínio**).

O roteamento que envolve mais de um sistemas autônomos é dito interdomínio, esse roteamento assume que a internet consiste de uma coleção de AS’s interconectados, geralmente existe um roteador dedicado em cada AS que trata o tráfego interdomínio. Os protocolos mais famosos são EGP e BGP. Exemplos: Rede da RNP. Rede de Backbone da Embratel.

**Roteamento Intradomínio e Interdomínio**

**Roteamento Intradomínio**

* Roteamento dentro de um AS
* Ignora o que está fora do AS
* Protocolos para roteamento intradomínio também são chamados de Interior Gateway Protocols ou IGP’s.
* Protocolos mais populares são:
  + RIP (simples, antigo)
  + OSPF (melhor)

**Roteamento Interdomínio**

* Roteamento entre AS’s
* Assume que a Internet consiste de uma coleção de AS’s interconectados
* Geralmente, há um roteador dedicado em cada AS que trata o tráfego interdomínio.
* Protocolos para roteamento interdomínio também são chamados Exterior Gateway Protocols ou EGP’s.
* Protocolos de Roteamento:
  + EGP
  + BGP (mais recente)

**Componentes de um Algoritmo de Roteamento**

* Um procedimento para enviar e receber informações de alcançabilidade sobre a rede para outros roteadores.
* Um procedimento para computação de rotas ótimas.
  + Rotas são calculadas utilizando algoritmo de menor caminho (shortest path):
    - Objetivo: Dada uma rede, a cada enlace é atribuído um custo. Procurar o caminho de menor custo entre duas redes com custo mínimo.
* Um procedimento para reagir e anunciar mudanças na topologia.

**Abordagens para o Roteamento de Menor Caminho**

Existem duas estratégias de roteamento na Internet:

**Vetor de Distância:**

Aqui, as rotas são anunciadas como vetores de distância e direção. A distância é definida em termos de uma métrica como contagem de saltos e a direção é dada simplesmente pelo roteador do próximo salto ou pela interface de saída. Os protocolos do vetor de distância normalmente usam o algoritmo Bellman-Ford para determinar a melhor rota. Esse algoritmo não permite que um roteador aprenda a topologia exata de redes interconectadas. O roteador só conhece as informações de roteamento recebidas de seus vizinhos. Os protocolos de roteamento do vetor de distância não têm um mapa real da topologia da rede.

* Cada nó conhece a distância (=custo) para seus vizinhos diretamente conectados
* Um nó envia periodicamente uma lista de atualizações de roteamento para seus vizinhos
* Se todos os nós atualizam seus custos, as tabelas de roteamento eventualmente convergem
* Novos nós anunciam sua chegada a seus vizinhos

**Funcionam melhor em situações nas quais:**

* A rede é simples e fixa e não requer um design hierárquico especial.
* Os administradores não têm conhecimentos suficientes para configurar e solucionar os problemas dos protocolos link-state.
* Redes de tipos específicos, como redes hub-and-spoke, estão sendo implementadas.
* Os tempos de convergência inesperada em uma rede não são uma preocupação.

Exemplo: **RIP**

* O protocolo de informações de roteamento (RIP, Routing Information Protocol) foi especificado originalmente em RFC 1058. Suas principais características são:
* A métrica usada para a seleção de caminhos é a contagem de saltos.
* Se a contagem de saltos de uma rede for maior do que 15, o RIP não poderá fornecer uma rota a essa rede.
* Por padrão, as atualizações de roteamento são broadcast ou multicast a cada 30 segundos.

**Estado de Enlace:**

Este algoritmo trabalha baseado na ideia de que cada roteador possui informações sobre as redes que estão conectadas a ele e, periodicamente, testa para determinar se cada enlace está ativo. Com estas informações cada roteador divulga uma lista sobre o status de cada conexão, dizendo se estas estão ativas ou inativas. Baseado nessas informações, quando um roteador recebe um conjunto de mensagens sobre o estado dos enlaces das redes próximas a ele, é aplicado o algoritmo Dijskstra. Este algoritmo é aplicado baseado nas informações de cada roteador e é feito localmente a cada um destes, para o cálculo das melhores rotas para todos os destinos a partir de uma mesma origem.

São basicamente cinco passos:

1. Descobrir seus vizinhos e aprender seus endereços de rede;
2. Medir o retardo ou o custo até cada um deles;
3. Criar um pacote que informe tudo o que ele aprendeu;
4. Enviar esse pacote a todos os roteadores;
5. Calcular o caminho mais curto até cada um dos roteadores.

Em termos de expansão, este algoritmo tem vantagem sobre o vetor distância, pois o cálculo do melhor caminho é feito localmente e não depende do cálculo de roteadores intermediários.

* Cada nó sabe a distância para seus vizinhos.
* A informação de distância (=estado de enlace) é difundida para todos os nós da rede.
* Cada nó calcula sua tabela de roteamento de forma independente.

**Contagem ao Infinito**

A razão para o problema da contagem ao infinito é que cada roteador tem apenas “next-hop-view” não através de quem está indo. Por exemplo, no primeiro passo, **A** não percebe que sua rota (com custo 2) para **C** vêm através do nó **B**. Há duas possíveis soluções.

1. Sempre passar o caminho inteiro na mensagem de atualização **(Path vectors ou Vetor de Distância por Caminho)**. Se as tabelas são grandes, as mensagens de roteamento vão requerer largura de banda substancial – BGP usa esta solução.
2. Nunca anunciar o custo para um vizinho se este é o próximo salto no caminho atual (**Horizonte Dividido**). Exemplo: **A** não enviaria a primeira atualização de rota para **B**, visto que **B** é o próximo salto na rota atual de **A** para **C** (**A** não manda antes de **B**). Também falha, dependendo da situação.

*O problema da contagem para o infinito pode ocorrer também no roteamento por estado do enlace? Por quê?*

Não, pois no roteamento por estado de enlace cada roteador guarda um grafo com a topologia total da rede atualizado por mensagens de LSA’s trocadas entre eles, usando a técnica de inundação. Caso algum link quebre, os roteadores se comunicam e atualizam seus grafos da rede através das LSA’s considerando o link quebrando. Assim não é possível ocorrer o problema de contagem para o infinito.

**Link State Database**

* O conjunto de todos os LSAs recebidos pelos nós é chamado de link-state database
* Cada roteador possui uma base de dados idêntica
  + Cada roteador necessita da descrição completa da rede para calcular as rotas
* Se roteadores vizinhos se comunicam pela primeira vez eles irão trocar as informações de suas bases de dados
* As bases de dados são sincronizadas através do mecanismo de disseminação dos dados do algoritmo *(flooding confiável)*

**Descoberta de Vizinhos**

* Os roteadores transmitem pacotes de hello para todas as interfaces controladas pelo OSPF.
* Se dois roteadores compartilham um link eles se tornam vizinhos e estabelecem a adjacência.
* Depois de se tornarem vizinhos os roteadores trocam suas bases de informações.

**Distribuição de Informações de Roteamento**

* LSA-Updates são transmitidos através de um Flooding Confiável (Reliable Flooding)
* Exemplo: Flooding do LSA do roteador 10.10.10.1